

Αφιέρωση

Αφιερώνω την εργασία αυτή στην γυναίκα μου Στέλλα, που κάνει κάθε όνειρο στην ζωή μου πραγματικότητα. Συγκεκριμένα χάρις στην πίστη της σε εμένα, στην συνεχή στήριξη της σε όλη την διάρκεια εκπόνησης της εργασίας, αλλά χάρις και σε δικές της θυσίες μπόρεσε να γίνει η εργασία, πραγματικότητα.

Επίσης αφιερώνω αυτήν την εργασία στους γονείς μου Ιωάννη και Κυριακή, οι οποίοι μου έδωσαν παιδεία και δεν έκαναν τίποτα για τον εαυτό τους ώστε να κάνουν τα πάντα, προκειμένου να αποκτήσουν και τα τρία τους παιδιά ανώτατη εκπαίδευση.

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου κ. Κώστα Γουλιάνα για τις γνώσεις πάνω στα Τεχνητά Νευρωνικά Δίκτυα, που μετέφερε σε εμένα μέσα από την εκπαιδευτική διαδικασία. Επίσης για την υποστήριξη του, την εμπιστοσύνη που επέδειξε ότι μπορώ να ολοκληρώσω την εργασία και τις ώρες που αφιέρωσε.

Πρόλογος

Η ολοκλήρωση των σπουδών ενός Μηχανικού Πληροφορικής στο Α.Τ.Ε.Ι. Θεσσαλονίκης, δίνει μια σημαντική προοπτική για την ενασχόληση με ένα άγνωστο αντικείμενο. Μια τέτοια εργασία μπορεί είτε να χρησιμοποιηθεί ως εφόδιο για την επαγγελματική αποκατάσταση στην μετέπειτα επαγγελματική ζωή, ή απλά για το άνοιγμα ενός νέου παράθυρου στην γνώση. Σκεφτόμενος το δεύτερο, επέλεξα να ασχοληθώ με ένα θέμα σχετικό με τα Τεχνητά Νευρωνικά Δίκτυα.

Το 2001 σε ένα περίπτερο της παγκόσμιας έκθεσης πληροφορικής Cebit, είχα δοκιμάσει την λειτουργία ενός συστήματος αναγνώρισης προσώπου (facial recognition) το οποίο πήρε δύο εικόνες, προφίλ και ανφάς από το πρόσωπο μου. Ενώ μέχρι τότε δεν με «γνώριζε», μπορούσε από εκείνη την στιγμή να με αναγνωρίσει σε οποιαδήποτε γωνία και αν είχα στραμμένο το πρόσωπο μου προς την κάμερα. Έκτοτε μου δημιουργούσε έντονο ενδιαφέρον το ενδεχόμενο της δημιουργίας ενός υπολογιστικού συστήματος που «αντιλαμβάνεται» τον κόσμο οπτικά. Σήμερα έχει περάσει ακριβώς μία δεκαετία από εκείνη την στιγμή με επιστήμη των υπολογιστών να συνεχίζει την αλματώδη πρόοδο της. Το προσωπικό μου ενδιαφέρον για το θέμα παρέμεινε αμείωτο όλα αυτά τα χρόνια, και ως θέμα η τεχνητή όραση παραμένει στην πρώτη γραμμή της έρευνας. Γι αυτό αποφάσισα και ξεκίνησα την υλοποίηση της πτυχιακής εργασίας με θέμα **«Αναγνώριση πραγματικών αντικειμένων από τις φωτογραφίες τους, με τη χρήση αντικειμενοστραφούς νευρωνικού δικτύου»**

Στο εξάμηνο αυτό βρέθηκα, όπως όλοι στην χώρα μας, αντιμέτωπος με κοινωνικά γεγονότα που είχαν άμεσο αντίκτυπο στην προσωπική μου κατάσταση και το λιγότερο αποσπούσαν την προσοχή από την υλοποίηση του στόχου μου. Έτσι μέσα σε αυτήν την πραγματικότητα, η όλη προσπάθεια φάνταζε σαν μια περιπέτεια του Don Quijote στην οποία πολεμούσε με ανεμόμυλους. Αν όμως αυτό σημαίνει ότι, η δύναμη της φαντασίας πηγαίνοντας κόντρα στην κοινή λογική καταφέρνει να μην περιοριστεί μέσα στα στενά πλαίσια της πεζής πραγματικότητας, τότε ας είναι ο υπολογιστής ο Rocinante. Προσθέτοντας στην φαντασία, σκληρή δουλειά, επιμονή και υπομονή έγινε εφικτό το αποτέλεσμα αυτής της εργασίας, το οποίο περιγράφεται μέσα σε αυτό το σύγγραμμα.



Ένας σύγχρονος ... ανεμόμυλος του Δον Κιχώτη, προϊόν της φαντασίας κάποιων ανθρώπων, το Bahrain World Trade Center. Ολοκληρώθηκε το 2008 και είναι το πρώτο κτήριο ουρανοξύστη που ενσωματώνει στην κατασκευή του ανεμογεννήτριες.

Πιθανές χρήσεις του συγγράμματος

Εκτός από τον βασικό του σκοπό, ο οποίος είναι να αποτυπώσει την πτυχιακή εργασία που υλοποιείται με το Horasis, το παρόν σύγγραμμα δημιουργήθηκε με γνώμονα τον πιθανό αναγνώστη του. Έτσι προσπαθεί να παραθέσει τις πληροφορίες ως εργασία της επιστήμης των υπολογιστών, χωρίς όμως να θεωρεί ότι ο αναγνώστης θα έχει τις ίδιες γνώσεις με τον συγγραφέα. Επίσης επεκτείνεται, έστω και ως υποσημείωση, σε γενικές γνώσεις που ίσως δεν σχετίζονται με τον σκοπό του, αλλά μπορεί να φανούν ενδιαφέρουσες σε οποιονδήποτε τις διαβάσει. Πιθανά η εργασία αυτή θα μπορούσε να χρησιμοποιηθεί από:

- Φοιτητές επιστήμης των υπολογιστών, ως μέσο για την γνωριμία με την τεχνητή νοημοσύνη, τα τεχνητά νευρωνικά δίκτυα και τον σύγχρονο αντικειμενοστραφή προγραμματισμό. Επίσης ως εργαλείο απόκτησης γενικών γνώσεων πληροφορικής και ως εγχειρίδιο αναφοράς για παρόμοιες εργασίες.
- Καθηγητές και επιστήμονες πληροφορικής, για την αξιολόγηση των νέων προτάσεων και καινοτομιών που εισάγει η εργασία. Επίσης σαν οδηγό προς μια πλούσια βιβλιογραφία που παρατίθεται για όλα τα θέματα με τα οποία ασχολείται.
- Φοιτητές άλλων επιστημών που αγγίζει η εργασία, για μια γρήγορη και ευκολονόητη αναφορά σε δικά τους θέματα, πριν τα διδαχθούν πληρέστερα, στον δικό τους κύκλο σπουδών.
- Καθηγητές και επιστήμονες άλλων επιστημών που αγγίζει η εργασία, για επίδειξη της προοπτικής που μπορεί να έχει ένα διεπιστημονικό ερευνητικό έργο τεχνητής νοημοσύνης.

Επίσης εκτός από τους ακαδημαϊκούς κύκλους ίσως η εργασία μπορεί να φανεί ενδιαφέρουσα σε:

- Αναγνώστες εκτός του χώρου της Πληροφορικής, για πληροφόρηση τους με έννοιες που δεν είναι γνωστές εκτός αυτού.
- Ειδήμονες από οποιονδήποτε τομέα που μπορούν να αξιολογήσουν τις Ελληνικές μεταφράσεις για τους Αγγλικούς όρους της ξένης βιβλιογραφίας.

Περίληψη

Με την πτυχιακή εργασία επιχειρείται μία νέα προσέγγιση στο πεδίο της τεχνητής όρασης, πάνω στο πρόβλημα της αναγνώρισης πραγματικών αντικειμένων από σταθερές εικόνες τους. Τα οπτικά πρότυπα που χρησιμοποιούνται, είναι έγχρωμα καρέ μιας ροής βίντεο, που καταγράφονται μέσω του συστήματος από απλές web κάμερες. Κατόπιν επισταμένης έρευνας σε πηγές της νευροεπιστήμης, η εργασία μοντελοποιεί όλα τα κύτταρα που υπάρχουν μέσα στον αμφιβληστροειδή χιτώνα του ανθρώπινου ματιού σε έναν τεχνητό αμφιβληστροειδή, που συνίσταται από αντικειμενοστραφή τεχνητά νευρωνικά δίκτυα (ΤΝΔ). Επίσης η λειτουργία για συγκεκριμένους νευρώνες του πρωτοταγούς οπτικού φλοιού προσομοιώνεται με αντίστοιχα ΤΝΔ, με την γενικότερη εφαρμογή της βιομιμητικής και του αντικειμενοστραφούς προγραμματισμού σε αυτά. Όλα αυτά τα ΤΝΔ εκπαιδεύονται αυτόνομα με τον αλγόριθμο back propagation και συνδέονται μεταξύ τους δημιουργώντας ένα μεγαλύτερο. Αυτό το ΤΝΔ μετατρέπει μια αρχική εικόνα τύπου bitmap που δίδεται ως είσοδος, σε μια συμπιεσμένη οπτική πληροφορία στην έξοδο του. Με αυτού του τύπου πληροφορίες ως πρότυπα εκπαιδεύεται ένα ΤΝΔ για να κατηγοριοποιεί τα πραγματικά αντικείμενα. Στην υλοποίηση του λογισμικού αξιοποιούνται οι σύγχρονες κατευθύνσεις της μηχανικής λογισμικού, όπως είναι ο προγραμματισμός με RAD, η αντικειμενοστραφή σχεδίαση και η τριών επιπέδων αρχιτεκτονική. Επίσης επιδεικνύονται οι δυνατότητες του παράλληλου προγραμματισμού σε υλικό πολυπύρηνου επεξεργαστή, υλοποιώντας την δυνατότητα να τρέχει μια αυτόνομη εκπαίδευση back propagation σε κάθε πυρήνα. Τέλος παρέχεται μια προχωρημένη γραφική διασύνδεση στον χρήστη, καθιστώντας δυνατή την οπτική επισκόπηση της κατάστασης του κάθε δικτύου κατά την εκπαίδευση, αλλά και την οπτικοποίηση των τιμών που προκύπτουν κατά την ανάκληση .

Abstract

With this thesis a new approach in the field of computer vision is attempted, on the problem of real life objects recognition using steady images. The visual patterns being used, are colored still frames of a video feed, captured by the system from simple web cameras. Following a detailed research on neurosciences sources, this project is modeling all cells that exist inside the human eye retina into an artificial retina, that is composed by object oriented artificial neural networks (ANNs). Also the functionality of certain neurons in the primary visual cortex is simulated by corresponding ANNs, with the general application of biomimicry and object oriented programming on them. All these ANNs are trained autonomously with the back propagation algorithm and are interconnected creating one larger. This ANN converts an original bitmap picture, given as input, into a new compressed visual information on its exit. With that kind of information as patterns, a ANN is trained to categorize the real objects. For the implementation of the software, modern guidelines of software engineering are utilized, such as RAD programming, object oriented design and three tiered architecture. Also the capabilities of parallel programming are displayed on multicore processor hardware, by implementing the capability of running an autonomous back propagation training task in each processor core. Finally an advanced graphical user interface is provided to the user, making possible the visual inspection of each network state during training, and also the visualization of the values that arise during recall.

Περιεχόμενα

Αφιέρωση	1
Ευχαριστίες.....	2
Πρόλογος	3
Πιθανές χρήσεις του συγγράμματος.....	4
Περίληψη.....	5
Abstract	5
Περιεχόμενα.....	7
1. Επισκόπηση	13
Εισαγωγή	13
<i>Η γενική περιγραφή και οι στόχοι της εργασίας.....</i>	<i>14</i>
<i>Η ονοματοδοσία</i>	<i>15</i>
Εργαλεία και υλικό που χρησιμοποιήθηκαν.....	17
<i>Η επιλογή του κατάλληλου εργαλείου RAD για την υλοποίηση του αντικειμενοστραφούς δικτύου.....</i>	<i>17</i>
<i>Το περιβάλλον προγραμματισμού Delphi</i>	<i>18</i>
<i>Η βάση δεδομένων του συστήματος</i>	<i>19</i>
<i>Το hardware που χρησιμοποιήθηκε και οι ελάχιστες απαιτήσεις</i>	<i>19</i>
<i>Η σκηνή φωτογράφισης του Horasis</i>	<i>20</i>
2. Επιστημονικό Πλαίσιο	21
Γνωστική Επιστήμη, Νευρωνικά Δίκτυα και Συνδεδασμός	21
<i>Τα τεχνητά νευρωνικά δίκτυα στην πράξη</i>	<i>21</i>
<i>Συνδεδασμός, Νευροεπιστήμη και Βιομημητική</i>	<i>22</i>
Η επιστημονική τοποθέτηση της εργασίας Horasis.....	24
<i>Η τεχνητή όραση και η εργασία Horasis</i>	<i>24</i>
Σύγχρονος Αντικειμενοστραφής Προγραμματισμός	26
<i>Η φιλοσοφία RAD.....</i>	<i>27</i>
<i>Παράλληλος Προγραμματισμός</i>	<i>28</i>
<i>Η βέλτιστη χρήση των αντικειμένων</i>	<i>28</i>
<i>Οι ομοιότητες της σωστής οργάνωσης των αντικείμενων με την οργάνωση των βιολογικών μονάδων.....</i>	<i>31</i>
Προοπτικές για την πρόοδο των τεχνητών νευρωνικών δικτύων	32
<i>Η μελλοντική πρόοδος του λογισμικού και τι μπορεί να συνεισφέρει</i>	<i>32</i>

<i>Η σημασία της αύξησης της υπολογιστής ισχύος και προοπτικές για μελλοντικό υλικό.....</i>	<i>32</i>
<i>Η σημασία της μείωσης του κόστους του υλικού για την αλματώδη πρόοδο των νευρωνικών δικτύων</i>	<i>33</i>
<i>Νευροεπιστήμη, το μέλλον.....</i>	<i>35</i>
3. Τα Τεχνητά Νευρωνικά Δίκτυα	37
Ο Βιολογικός Νευρώνας.....	37
<i>Περιγραφή του βιολογικού νευρώνα και της σύναψης.....</i>	<i>37</i>
<i>Η εσωτερική λειτουργία του νευρώνα</i>	<i>39</i>
<i>Διαφορετικοί τύποι συνάψεων</i>	<i>40</i>
Ο Τεχνητός Νευρώνας.....	41
<i>Οι συναρτήσεις ενεργοποίησης σιγμοειδούς μορφής</i>	<i>42</i>
Το Τεχνητό Νευρωνικό Δίκτυο και η εκπαίδευσή του.....	45
<i>Ανάκληση στο δίκτυο MLP του Horasis, με σιγμοειδή συνάρτηση ενεργοποίησης.....</i>	<i>47</i>
<i>Η εκπαίδευση του δικτύου με επίβλεψη</i>	<i>47</i>
<i>Η μαθηματική βάση του αλγόριθμου εκπαίδευσης Back-Propagation</i>	<i>49</i>
<i>Ο τερματισμός της εκπαίδευσης.....</i>	<i>51</i>
<i>Η έννοια της ορμής στο back propagation και η σημασία του βήματος εκπαίδευσης</i>	<i>52</i>
<i>Τα βήματα του αλγόριθμου back-propagation.....</i>	<i>53</i>
4. Τα Αντικείμενα Τεχνητών Νευρώνων	57
Ο Αντικειμενοστραφής Τεχνητός Νευρώνας	57
<i>Η κλάση του νευρώνα και τα συστατικά της.....</i>	<i>58</i>
<i>Η σύναψη ως αντικείμενο</i>	<i>58</i>
<i>Το συναπτικό βάρος ως ενεργό αντικείμενο.</i>	<i>59</i>
<i>Συναπτικές ρυθμίσεις για την λειτουργία του αντικειμένου-βάρους</i>	<i>60</i>
Το Αντικειμενοστραφές Νευρωνικό Δίκτυο.....	61
<i>Το δίκτυο και τα στρώματα των νευρώνων.....</i>	<i>61</i>
<i>Οι διασυνδέσεις των στρωμάτων</i>	<i>61</i>
Τα Πρότυπα Εκπαίδευσης	63
<i>Το αντικείμενο ως ερέθισμα εισόδου και η συλλογή από πρότυπα εκπαίδευσης.....</i>	<i>63</i>
<i>Ανάκληση του αντικειμενοστραφούς δικτύου.....</i>	<i>63</i>
<i>Η τριών επιπέδων σχεδίαση και η διαχείριση των προτύπων ως δεδομένα</i>	<i>64</i>
Η Μορφή του Δικτύου.....	66
<i>Ο γονότυπος του αντικειμενοστραφούς νευρωνικού δικτύου</i>	<i>66</i>
<i>Δυνατότητες αποθήκευσης και επαναφοράς μια μορφής δικτύου</i>	<i>66</i>
5. Η Εκπαίδευση του Δικτύου	69

Η Οπτικοποίηση της μορφής του δικτύου	69
Διαδικασία της Παράλληλης Εκπαίδευσης.....	71
<i>Μέγιστη ταχύτητα με πραγματικά παράλληλη λειτουργία σε πυρήνα επεξεργαστή.</i>	<i>71</i>
<i>Το αντικείμενο κατάστασης του αντικειμενοστραφούς ΤΝΔ</i>	<i>72</i>
<i>Κάτω από το καπώ και μέσα στο νήμα εκπαίδευσης back propagation</i>	<i>73</i>
<i>Το νέο κριτήριο τερματισμού του back propagation</i>	<i>78</i>
Οπτικοποίηση της Εκπαίδευσης	79
<i>Το επίπεδο εμφάνισης (presentation layer) της εκπαίδευσης</i>	<i>79</i>
<i>Το παράθυρο παρακολούθησης εποχών και μέσου τετραγωνικού σφάλματος.....</i>	<i>80</i>
<i>Η σελίδα γραφικής παράστασης του μέσου τετραγωνικού σφάλματος.....</i>	<i>81</i>
<i>Η χρωματική αναπαράσταση του σφάλματος των νευρώνων.....</i>	<i>83</i>
<i>Φόρμα επισκόπησης νευρώνα</i>	<i>84</i>
6. Τα Οπτικά Πρότυπα	85
Η εικόνα στον υπολογιστή	85
<i>Η τεχνολογία της ψηφιακής κάμερας.....</i>	<i>85</i>
<i>Η διασύνδεση με Microsoft DirectShow</i>	<i>88</i>
<i>Η μορφή μνήμης της εικόνας BMP.....</i>	<i>89</i>
Χρωματικά Μοντέλα	90
<i>Η σύγχρονη χρωματική θεωρία και το προσθετικό μοντέλο RGB.....</i>	<i>90</i>
<i>Το μοντέλο HSL και ο τροχός των χρωμάτων</i>	<i>90</i>
<i>Μετατροπές μεταξύ μοντέλων RGB και HSL</i>	<i>93</i>
<i>Τα αμιγή χρώματα και οι τιμές της απόχρωσης τους.....</i>	<i>95</i>
<i>Μετατροπή από χρώμα RGB σε μετρήσιμη ποσότητα φωτός.....</i>	<i>96</i>
Δειγματοληψία των εικόνων στο Horasis	97
<i>Η ρύθμιση των δύο καμερών στην σκηής φωτογράφισης.</i>	<i>97</i>
<i>Η φόρμα της λήψης φωτογραφιών</i>	<i>98</i>
<i>Η σχεσιακή βάση δεδομένων των προτύπων</i>	<i>99</i>
<i>Σχηματική αναπαράσταση του συστήματος λήψης οπτικών προτύπων του Horasis.....</i>	<i>101</i>
7. Η Ανθρώπινη Όραση	103
Το Ανθρώπινο Μάτι.....	103
<i>Γενική περιγραφή του οφθαλμού</i>	<i>103</i>
<i>Αντιστοιχία της ψηφιακής κάμερας με τον οφθαλμό.</i>	<i>105</i>
Αμφιβληστροειδής Χιτώννας	106
<i>Τα συστατικά και η γενική συνδεσμολογία του αμφιβληστροειδούς.....</i>	<i>106</i>
<i>Φωτοϋποδοχείς.....</i>	<i>108</i>
<i>Οριζόντια κύτταρα</i>	<i>110</i>

<i>Δίπολα κύτταρα και η πρώτη παρουσία του χρωματικού ανταγωνισμού</i>	111
<i>Κυκλικά υποδεκτικά πεδία</i>	112
<i>Γαγγλιακά κύτταρα</i>	113
<i>Βραχύινα κύτταρα</i>	116
Ο Οπτικός Φλοιός V1.....	117
<i>Η έξοδος του αμφιβληστροειδούς ως είσοδος στον εγκέφαλο</i>	117
<i>Απλοί V1 νευρώνες και υπερστήλες νευρώνων.</i>	119
<i>Σύνθετοι V1 νευρώνες</i>	122
8. Η Αρχιτεκτονική του Horasis	123
Η Διεργασία Ανάπτυξης Λογισμικού του Horasis	123
<i>Η αυξητική και επαναληπτική ανάπτυξη του Horasis</i>	123
<i>Η αρχή UP της μοντελοποίησης το όραμα και η τεκμηρίωση της αρχιτεκτονικής</i>	125
Η Αρχιτεκτονική του Horasis	126
<i>Το γενικό αρχιτεκτονικό πλάνο του συστήματος Horasis</i>	126
<i>Τα τμήματα της αρχιτεκτονικής του Horasis</i>	128
<i>Τα υποσυστήματα του τεχνητού αμφιβληστροειδή</i>	130
9. Η Μοντελοποίηση της Όρασης	133
Τεχνητοί Φωτοϋποδοχείς.....	133
<i>Περικοπή της εικόνας σε τετράγωνο</i>	133
<i>Πρόσβαση στις χρωματικές συνιστώσες ενός εικονοστοιχείου μέσω scanlines</i>	134
<i>Το αντικείμενο φωτοϋποδοχέων σημείου</i>	135
Τεχνητός Αμφιβληστροειδής.....	137
<i>Τεχνητά οριζόντια κύτταρα και ο χάρτης υποδεκτικού πεδίου</i>	137
<i>Η έναρξη των διαύλων χρωματικού ανταγωνισμού και τα τεχνητά δίπολα κύτταρα</i>	139
<i>Η εκπαίδευση του ΤΝΔ που περιέχει τεχνητούς δίπολους νευρώνες των τριών κωνίων ανάλογα με την απόχρωση</i>	140
<i>Η εκπαίδευση του ΤΝΔ που περιέχει τεχνητούς δίπολους νευρώνες ενός ραβδίου ανάλογα με το φώς</i>	143
<i>Γαγγλιακά Κύτταρα</i>	145
<i>Τεχνητός ρετινοτοπικός χάρτης και η προκαθορισμένη τοποθέτηση του κέντρου του</i>	147
<i>Χρήση του ρετινοτοπικού χάρτη</i>	149
<i>Η μαθηματική έκφραση κυκλικού υποδεκτικού πεδίου με ανταγωνισμό κέντρου και περιφέρειας</i>	150
<i>Η εκπαίδευση του ΤΝΔ που αντιπροσωπεύει ένα τεχνητό γαγγλιακό κύτταρο</i>	152
<i>Τεχνητά βραχύινα κύτταρα</i>	156
Τεχνητός Οπτικός Φλοιός.....	157
<i>Μικρά Απλά Κύτταρα V1</i>	157

<i>Μεγάλα Απλά Κύτταρα V1</i>	<i>160</i>
<i>Η εκπαίδευση του ΤΝΔ που αντιπροσωπεύει ένα τεχνητό απλό κύτταρο V1.</i>	<i>161</i>
<i>Εσωτερική αναπαράσταση της εικόνας του αντικειμένου.....</i>	<i>162</i>
<i>Συνδεσμολογία ΤΝΔ στο Horasis</i>	<i>164</i>
10. Οργάνωση και Υλοποίηση	165
<i>Η αυξητική και επαναληπτική υλοποίηση, συνοπτικά.....</i>	<i>165</i>
Τα Τεχνητά Νευρωνικά Δίκτυα του Horasis και η εκπαίδευση τους.....	169
<i>Το αντικείμενο διαχείρισης των διαφορετικών τύπων τεχνητών νευρωνικών δικτύων.</i>	<i>169</i>
<i>Οργάνωση των προτύπων σε συλλογές.</i>	<i>171</i>
<i>Φόρμα προτύπων εκπαίδευσης τεχνητών δίπολων κυττάρων.....</i>	<i>172</i>
<i>Φόρμα προτύπων εκπαίδευσης τεχνητών γαγγλιακών κυττάρων</i>	<i>174</i>
<i>Φόρμα προτύπων εκπαίδευσης τεχνητών απλών κυττάρων V1.....</i>	<i>176</i>
<i>Φόρμα προτύπων εκπαίδευσης ενοποιημένης απόκρισης V1</i>	<i>180</i>
Οπτική Αναγνώριση των Πραγματικών Αντικειμένων.....	182
<i>Ανάπτυξη του τεχνητού αμφιβληστροειδούς</i>	<i>182</i>
<i>Ανάπτυξη του τεχνητού οπτικού φλοιού</i>	<i>184</i>
<i>Φόρμα οπτικοποίησης εσωτερικής λειτουργίας του Horasis.....</i>	<i>185</i>
<i>Φόρμα μετατροπών οπτικών προτύπων του πραγματικού αντικειμένου</i>	<i>191</i>
<i>Φόρμα προτύπων εκπαίδευσης τεχνητού νευρωνικού δικτύου κατηγοριοποίησης</i>	<i>193</i>
Παραμετροποίηση της Λειτουργίας.....	195
<i>Οπτικές παράμετροι.....</i>	<i>195</i>
<i>Χρωματικές Παράμετροι</i>	<i>196</i>
<i>Γεωμετρικές Παράμετροι.....</i>	<i>197</i>
<i>Αλγεβρικές Παράμετροι</i>	<i>199</i>
<i>Νευρωνικές Παράμετροι</i>	<i>203</i>
11. Συμπεράσματα και Προτάσεις.....	205
Επιτεύγματα Του Horasis	205
<i>Χρήση αντικειμενοστραφούς σχεδιασμού και προγραμματισμού στα τεχνητά νευρωνικά δίκτυα.....</i>	<i>205</i>
<i>Μεταφορά του βιολογικού μοντέλου της όρασης σε λογισμικό τεχνητών νευρωνικών δικτύων.</i>	<i>206</i>
<i>Ολοκληρωμένος σχεδιασμός τεχνητής όρασης και μεταφορά του σε ένα πλήρως λειτουργικό λογισμικό.....</i>	<i>207</i>
<i>Λειτουργία αναγνώρισης οπτικών αντικειμένων που στηρίζεται στο χρώμα.....</i>	<i>207</i>
<i>Μη αλγοριθμικός τρόπος εξαγωγής πληροφοριών σε σχέση με σχήμα.</i>	<i>208</i>
<i>Δημιουργία οπτικών προτύπων από καρέ ροής βίντεο σε μη ιδανικές συνθήκες.</i>	<i>209</i>
<i>Συμπίεση οπτικής πληροφορίας χωρίς να χάνει τα βασικά χαρακτηριστικά της.....</i>	<i>209</i>

<i>Χρήση παράλληλης επεξεργασίας.....</i>	210
<i>Εισαγωγή νέων προτάσεων και παραγωγή νέων όρων στην τεχνητή όραση</i>	210
Προτάσεις, Ονοματολογικές Συμβάσεις, Νέες Έννοιες,	211
<i>Προτάσεις και καινοτομίες πάνω στα τεχνητά νευρωνικά δίκτυα.....</i>	211
<i>Μεταφράσεις ξένων όρων</i>	212
<i>Ονομασίες κατά σύμβαση που αφορούν τεχνητά νευρωνικά δίκτυα</i>	213
<i>Νέες έννοιες και καινοτομίες πάνω στην τεχνητή όραση</i>	213

Βιβλιογραφία και Παραπομπές URL	215
Πίνακας Εικόνων	227

1. Επισκόπηση

Εισαγωγή

Η τεχνητή όραση αποτελεί σήμερα, ένα από τα θέματα αιχμής της επιστήμης των υπολογιστών του 21^{ου} αιώνα, με το οποίο ασχολείται ο κλάδος της Τεχνητής Νοημοσύνης ή ΤΝ για συντομία. Ήδη από τις δύο τελευταίες δεκαετίες του προηγούμενου αιώνα έχουν δημιουργηθεί οι κατάλληλες συνθήκες που μπορεί να οδηγήσουν σε σημαντική πρόοδο. Η ανάπτυξη του υλικού των υπολογιστών, της προγραμματιστικής φιλοσοφίας και εργαλείων αλλά κυρίως η δημιουργία της Γνωστικής Επιστήμης ως ένα διεπιστημονικό κλάδο, δίνουν την απαιτούμενη ώθηση σε οποιαδήποτε προσπάθεια του ανθρώπου να κάνει την μηχανή να «σκέφτεται» και να «βλέπει» τον κόσμο.

Η πρόδος της υπολογιστικής ισχύος, που ακολούθησε τον νόμο του Moore¹, αλλά και η δυνατότητα απόκτησης της με μικρό κόστος παρέχει πλέον την δυνατότητα να δοκιμαστεί εύκολα μια πληθώρα επιστημονικών προσεγγίσεων. Πλέον η δυνατότητα πραγματικής παράλληλης επεξεργασίας βρίσκεται ήδη πάνω στα γραφεία από τον απλούστερο προσωπικό υπολογιστή που έχει πλέον πολλούς πυρήνες επεξεργασίας. Στο μέλλον η αύξηση του πλήθους τους, θα δημιουργήσει ακόμα πιο γρήγορους υπολογιστές που απαιτούνται για την έρευνα ΤΝ.

Από την άλλη η πρόδος της μηχανικής λογισμικού και της αντικειμενοστραφούς προγραμματιστικής τεχνολογίας, μας έδωσε ισχυρά εργαλεία για την ταχύτερη ανάπτυξη εφαρμογών που είναι γνωστά με τα αρχικά RAD². Με αυτά ελαχιστοποιείται ο χρόνος μεταφοράς ενός θεωρητικού μοντέλου σε ένα πλήρες λειτουργικό πρόγραμμα που το υλοποιεί. Προχωρημένες προγραμματιστικές δυνατότητες, όπως η αξιοποίηση των πολλαπλών πυρήνων για πραγματική παράλληλη επεξεργασία, πρόσβαση σε πολυμέσα, σχεδιασμός γραφικών, διαχείριση δεδομένων μπορούν πολύ εύκολα να ενοποιηθούν στο ίδιο σύστημα που λειτουργεί το μοντέλο. Με την χρήση κλάσεων και αντικειμένων για την μοντελοποίηση, δημιουργείται ένας νέος τρόπος σκέψης στον προγραμματιστή, που μπορεί να παράξει ριζοσπαστικές υλοποιήσεις ικανές να βοηθήσουν προσθετικά στην μαθηματική βάση του μοντέλου.

Ίσως όμως το πιο σημαντικό βήμα έγινε μέσα από την οργάνωση του νέου κλάδου της γνωστικής επιστήμης, που πλέον εναγκαλίζει ως πεδίο την ΤΝ. Ένα ρεύμα μέσα σε αυτήν που ονομάζεται συνδεσιασμός (connectionism) προσπαθεί να φθάσει σε επιτεύγματα χρησιμοποιώντας τα τεχνητά νευρωνικά δίκτυα, τα οποία προσομοιώνουν τα πραγματικά δίκτυα που δημιουργούν οι βιολογικοί νευρώνες. Εξετάζοντας τις διανοητικές λειτουργίες του ανθρώπου, φτιάχνονται μοντέλα που στηρίζονται στην μελέτη της βιολογικής πραγματικότητας από την νευροεπιστήμη, ένα άλλο σημαντικό πεδίο της γνωστικής επιστήμης. Σήμερα υπάρχει τεράστιο ερευνητικό ενδιαφέρον παγκοσμίως για την νευροεπιστήμη, κυρίως για την ίαση σημαντικών εκφυλιστικών ασθενειών του νευρικού συστήματος. Από αυτό μπορεί να προκύψουν στο μέλλον, νέα σημαντικά δεδομένα που μπορούν να αξιοποιηθούν και από την τεχνητή νοημοσύνη, για την δημιουργία «έξυπνων» μηχανών.

¹ Ο νόμος του Moore λέγει ότι ο αριθμός των τρανζίστορς σε ένα μικροεπεξεργαστή, θα διπλασιάζεται κάθε περίπου δύο χρόνια.

² RAD = Rapid Application Development

Η γενική περιγραφή και οι στόχοι της εργασίας

Με την παρούσα εργασία επιχειρείται μια αξιοποίηση των νέων δυνατοτήτων της υπολογιστικής τεχνολογίας και της μηχανικής λογισμικού. Επίσης γίνεται μια περιήγηση μέσα από τους διαδρόμους της γνωστικής επιστήμης, της τεχνητής νοημοσύνης από την σκοπιά του συνδεδεασμού και της νευροεπιστήμης, έχοντας ως άρμα τα τεχνητά νευρωνικά δίκτυα. Ο προορισμός είναι η προσομοίωση της λειτουργίας των νευρώνων του ανθρώπινου οφθαλμού αλλά και κάποιων από τους νευρώνες του εγκεφάλου που ανήκουν στον οπτικό φλοιό. Ο τελικός στόχος είναι η να προστεθεί ένα ακόμα λιθαράκι στο οικοδόμημα της τεχνητής όρασης, αφενός με την δημιουργία μια νέα διαμόρφωσης της οπτικής πληροφορίας από έγχρωμες φωτογραφίες πραγματικών αντικειμένων και αφετέρου με την αναγνώριση του από τον υπολογιστή με την αξιοποίηση της.

Η προσπάθεια ακολουθεί την ιδέα του βιολογικού ρεαλισμού, επιχειρώντας να υλοποιήσει τις παρατηρήσεις της νευροεπιστήμης ως λογισμικό τεχνητών νευρωνικών δικτύων. Το σύστημα είναι στην ουσία ένα μεγάλο νευρωνικό δίκτυο που αποτελείται από πολλά επιμέρους υποδίκτυα, έχοντας την έξοδο του κάθε υποδικτύου να τροφοδοτεί την είσοδο του επομένου. Η αρχική είσοδος είναι μια τριχρωματική εικόνα υψηλής ανάλυσης σε μορφή bitmap³, στην οποία για το κάθε εικονοστοιχείο (pixel⁴) υπάρχουν τρία δεδομένα, οι χρωματικές συνιστώσες του κόκκινου, πράσινου και μπλε που έχει καταγράψει ο φωτογραφικός αισθητήρας. Όλο το δίκτυο έχει δομηθεί από αντικείμενα - τεχνητά νευρωνικά κύτταρα που το κάθε ένα έχει συστατικά αντίστοιχα με τα συστατικά των βιολογικών νευρωνικών κυττάρων. Οι αντικειμενοστραφείς νευρώνες σχηματίζουν συνάψεις δηλαδή συνδέσεις μεταξύ τους, που και αυτές με την σειρά τους είναι αντικείμενα.

Οι στόχοι της πτυχιακής εργασίας πέραν από τον βασικό της σκοπό που υποδηλώνει ο επιθετικός της προσδιορισμός είναι πολλοί. Πρώτα από όλα είναι να προτείνει την εφαρμογή της βιομημητικής στην ανάπτυξη λογισμικού τεχνητής νοημοσύνης, με την μεταφορά της συμπεριφοράς των νευρώνων της όρασης, από την νευροεπιστήμη, σε ένα υλοποιημένο μοντέλο ΤΝΔ. Επίσης προτείνει ότι η υλοποίηση πρέπει να έχει διακριτά αντικείμενα λογισμικού για κάθε διαφορετικό βιολογικό τμήμα του νευρωνικού δικτύου. Έτσι πηγαίνουμε στον δεύτερο στόχο της που είναι να δείξει τα οφέλη από την δημιουργία αντικειμενοστραφών ΤΝΔ με την χρήση σύγχρονων προγραμματιστικών εργαλείων και κατευθύνσεων, για την υλοποίησή τους. Ο τρίτος στόχος είναι να δείξει πως μπορεί να αξιοποιηθεί το σύγχρονο υλικό που δίνει δυνατότητες παράλληλης επεξεργασίας σε λύσεις μεγάλης πολυπλοκότητας που δεν ήταν δυνατόν να υλοποιηθούν έως τώρα. Ο τέταρτος και πιο ουσιαστικός στόχος είναι τηρώντας τα δύο παραπάνω να παράγει ένα χρήσιμο αποτέλεσμα στην οπτική αναγνώριση πραγμάτων του χώρου. Όλοι οι στόχοι βρίσκονται κάτω από την σκέπη της επιστήμης των υπολογιστών, περνώντας ταυτόχρονα τα σύνορα της γνωστικής επιστήμης, της νευροεπιστήμης και της βιολογίας για την επίτευξη τους.

Η εργασία αποτελεί τέλος ένα πρόσφορο έδαφος για να παρατεθούν σκέψεις, νέες ιδέες, ίσως και πρωτοεμφανιζόμενοι όροι σχετικοί με τους στόχους της και τους τομείς πάνω στους οποίους δραστηριοποιείται.

³ Bit-map είναι μια μορφή στην οποία η εικόνα έχοντας καταταμηθεί σε ένα πλήθος ομοίου μεγέθους τετραγωνισμένων κομματιών (pixel), αναπαριστάται πίνακας από bit σε αντίστοιχη θέση με τις συντεταγμένες του κάθε pixel.

⁴ Το Pixel παράγεται από τα τις λέξεις Picture Element.

Η ονοματοδοσία

Το όνομα που δόθηκε στο σύστημα λογισμικού που υλοποιεί η εργασία είναι **«Όρασις»** από την αντίστοιχη αρχαία ελληνική λέξη που περιγράφει την αίσθηση. Με λατινικούς χαρακτήρες θα αναγράφεται αντίστοιχα **Horasis**⁵, και θα χρησιμοποιείται πλέον αυτός ο τίτλος μέσα στην εργασία ώστε να μην συγχέεται με την λέξη όραση και να δείχνει κατευθείαν ότι αναφερόμαστε στο σύστημα λογισμικού.



1-1: Έμβλημα εφαρμογής horasis.

Ο τίτλος αυτός επιλέχθηκε επειδή είναι απλός, ευκολομνημόνευτος και περιγραφικός για την λειτουργία του συστήματος. Με αυτήν την σκέψη επιλέχθηκε και ένα έμβλημα (logo) για την εφαρμογή. Είναι και αυτό με την σειρά του εμπνευσμένο από την αρχαιότητα και είναι το μάτι του Ώρου⁶ (Horus), ένα αρχαίο Αιγυπτιακό σύμβολο προστασίας αλλά και ιερογλυφικό. Αξίζει να σημειωθεί ότι το κεφάλι του Ώρου, ως προφίλ με ματιά προς τα δεξιά, είναι σήμερα το εμπορικό σήμα της Αιγυπτιακής αεροπορικής εταιρίας EgyptAir. Έτσι ένα τέτοιο σύμβολο εκτός του ότι είναι άμεσα αναγνωρίσιμο και έχει μεγάλη ιστορική αξία, είναι το πλέον περιγραφικό για την λειτουργία του συστήματος Horasis.

Με τον δανεισμό του ονόματος και του συμβόλου από την αρχαιότητα, δίδεται φόρος τιμής στον Ελληνικό πολιτισμό που δημιούργησε τις βάσεις της επιστημονικής σκέψης και στον Αιγυπτιακό πολιτισμό που αξιοποιώντας το φυσικό περιβάλλον⁷, κληροδότησε μεγάλα τεχνουργήματα.



1-2: Πρότυπα λογότυπα για το σύστημα στα Αγγλικά και στα Ελληνικά
Σχεδίαση: Παντελής Ι. Καπλάνογλου

⁵ Επειδή στο πολυτονικό η λέξη «όρασις» έχει δασεία, στην λατινική της ονομασία τοποθετείται το λατινικό γράμμα H στην αρχή της ονομασίας.

⁶ Ο Ώρος είναι ένας από τους σημαντικότερους αρχαίους Αιγυπτιακούς θεούς που απεικονίζονταν ως γεράκι ή άνθρωπος με κεφαλή γερακιού. Ήταν ο θεός του ουρανού και θεωρείτο ως προστάτης των ανθρώπων.

⁷ Η αξιοποίηση του νερού του Νείλου ήταν το βασικό στοιχείο του πολιτισμού τους. Επίσης μια πρόσφατη ριζοσπαστική θεωρία υποστηρίζει ότι η ανέγερση των ογκόλιθων των Πυραμίδων και των οβελίσκων, έγινε με την βοήθεια του ανέμου με τεράστιους «χαρταετούς» που έσερναν τα κομμάτια.

Εργαλεία και υλικό που χρησιμοποιήθηκαν

Ως πληροφοριακό σύστημα το Horasis αποτελείται από τρία διακριτά μέρη. Έναν πολυπύρρηνο ⁸ προσωπικό υπολογιστή με λειτουργικό σύστημα Windows, ένα λογισμικό φτιαγμένο με προγραμματιστικά εργαλεία RAD που αποτελεί την υλοποίηση του συστήματος, και μια κατασκευή λήψης φωτογραφιών αντικειμένων που έχει στερεωμένες δύο κάμερες web και λειτουργεί ως σκηνικό λήψης των φωτογραφιών.

Από την αρχή ελήφθησαν μερικές αποφάσεις που καθόρισαν την μορφή και ύπαρξη αυτών των τριών μερών. Καταρχήν η απόφαση να ληφθούν φωτογραφίες απευθείας από απτά και συνηθισμένα αντικείμενα, δημιούργησε την ανάγκη χρήσης της κάμερας. Για την απλοποίηση των φωτογραφιών που θα χρησιμοποιούσε το σύστημα, αποφεύγοντας τις πληροφορίες του φόντου και περιορίζοντας τις σκιάς, κατασκευάστηκε μια σκηνή με μαύρο φωτογραφικό χαρτί. Δεύτερον, επειδή το σύστημα επρόκειτο να χρησιμοποιήσει παράλληλη επεξεργασία για την επιτάχυνση της παραγωγής των αποτελεσμάτων, ο προσωπικός υπολογιστής έπρεπε να έχει παραπάνω από έναν επεξεργαστή (CPU). Τρίτον για να υλοποιηθεί η ιδέα του Horasis χρειαζόταν η συγγραφή ενός μεγάλου σε μέγεθος κώδικα και η σκέψη ήταν να χρησιμοποιηθεί ένα προγραμματιστικό εργαλείο RAD για την ευκολία και ταχύτητα που προσφέρει στην ανάπτυξη. Καθώς η ιδέα θα απαιτούσε αρκετά πολύπλοκη υλοποίηση, που μόνο η χρήση του αντικειμενοστραφή προγραμματισμού θα μπορούσε να ανταπεξέλθει με ευρωστία, η χρήση ενός εργαλείου RAD άνοιγε τον δρόμο στο να γίνει η υλοποιηθεί η ιδέα.

Η επιλογή του κατάλληλου εργαλείου RAD για την υλοποίηση του αντικειμενοστραφούς δικτύου.

Για την επιλογή του εργαλείου RAD, έγινε αρχικά ο συμβιβασμός ότι θα τρέχει σε προσωπικούς υπολογιστές PC και σε λειτουργικό σύστημα Windows, όπου υπάρχουν πολλές επιλογές για εργαλεία RAD. Η δυνατότητα μεταφερσιμότητας σε άλλο τύπο υπολογιστών δεν ήταν στους στόχους της εργασίας και γι αυτό δεν υλοποιήθηκε σε γλώσσα C++.

Η λύση χρειαζόταν επιπρόσθετα προχωρημένες λειτουργίες που απαιτούσαν την διασύνδεση με τα Windows, όπως η διαχείριση της κάμερας, η διαχείριση και επεξεργασία εικόνων bitmap αλλά και η σχεδίαση γραφικών. Επιπρόσθετα, η γλώσσα προγραμματισμού έπρεπε να είναι ευρέως διαδεδομένη ώστε να υπάρχει πληθώρα λύσεων και βιβλιοθηκών στο Internet. Έτσι η επιλογή στένεψε αρχικά σε τρεις γλώσσες προγραμματισμού.

- Java
- C#
- Object Pascal / Delphi

Όταν όμως το θέμα ταχύτητα ήρθε να προστεθεί στα κριτήρια επιλογής, τότε η μόνη γλώσσα από τις τρεις που παρήγαγε εκτελέσιμο απευθείας στην γλώσσα μηχανής του επεξεργαστή (native), χωρίς την ανάγκη μεταγλώττισης ενός ενδιάμεσου κώδικα ⁹ [1] [2] ήταν η Delphi. Επίσης η Object Pascal είναι μια αρκετά ευκολονόητη γλώσσα με τον κώδικα να μπορεί εύκολα να διαβαστεί, γι αυτό το συντακτικό της χρησιμοποιείται συχνά ως ψευδοκώδικας. Αν και 32bit,

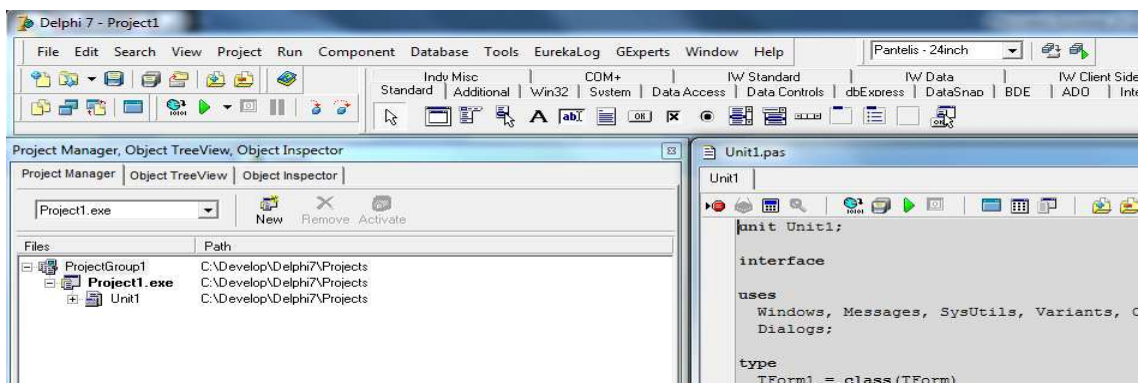
⁸ Πολυπύρρηνος (multicore) υπολογιστής που αποτελείται από επεξεργαστή που έχει πολλούς πυρήνες επεξεργασίας, έχοντας στην πράξη παραπάνω του ενός επεξεργαστές.

⁹ Η Java παράγει τον java bytecode που κατόπιν μεταγλωττίζεται από τον JIT Compiler σε native κώδικα του επεξεργαστή ενώ η C# παράγει την CIL (MSIL) που μεταγλωττίζεται από τον .NET Runtime JIT Compiler.

το εκτελέσιμο που παράγει ο μεταγλωττιστής της Delphi είναι βελτιστοποιημένο για την ταχύτερη εκτέλεση. Πολλές από τις κεντρικές ρουτίνες της βιβλιοθήκης της delphi είναι γραμμένες απευθείας σε γλώσσα μηχανής assembly, που καθιστά την γλώσσα ταχύτερη για τις ανάγκες των υπολογισμών που εκτελούν τα τεχνητά νευρωνικά δίκτυα. Από τις διαφορετικές εκδόσεις της Delphi χρησιμοποιήθηκε η έκδοση 7 που υπάρχει εδώ και 10 χρόνια, ώστε αφενός να αποφευχθούν προβλήματα με τυχόν άλυτα προβλήματα των νεότερων εκδόσεων, αφετέρου να αξιοποιηθούν οι συσσωρευμένες γνώσεις που υπάρχουν στο διαδίκτυο σε όσες περιπτώσεις θα χρειαζόταν.

Το περιβάλλον προγραμματισμού Delphi

Η Delphi [3][4] έχει χρησιμοποιηθεί σε πολλές μεγάλες επιχειρησιακές λύσεις και είναι επιλογή πολλών μεγάλων εταιριών ακόμα και στην Ελλάδα¹⁰. Αποτελείται από ένα Ολοκληρωμένο Περιβάλλον Ανάπτυξης (IDE¹¹) που ενσωματώνει τον «έξυπνο» μεταγλωττιστή Object Pascal, ένα Debugger, Linker κ.α. Επίσης εμπεριέχει μια βιβλιοθήκη (framework) με έτοιμες κλάσεις και ρουτίνες που συνίσταται από την Run Time Library (RTL) και την Visual Component Library (VCL), αποτελώντας ουσιαστικά τον πρόδρομο του Microsoft .NET Framework. Αυτό φαίνεται αν θυμηθούμε ότι αρχικά η Delphi δημιουργήθηκε με την καθοδήγηση του Anders Hejlsberg [5] ο οποίος μετακινήθηκε από την Borland, την τότε εταιρία παραγωγής της Delphi, στην Microsoft το 1996, και είναι έκτοτε ο γενικός αρχιτέκτων του .NET framework.



1-3: Βασικά στοιχεία του IDE της Delphi 7, στο επάνω μέρος της οθόνης

Η βιβλιοθήκη παρέχει αρκετές λειτουργίες προς τον προγραμματιστή και πρακτικά μπορούμε να πούμε ότι με αυτήν μπορεί κάποιος να γράψει οποιοδήποτε τύπου εφαρμογή χωρίς την ανάγκη εξωτερικών βιβλιοθηκών. Επίσης υπάρχει η δυνατότητα για εύκολο χειρισμό των νημάτων (thread), με απόκρυψη των αντικειμένων των Windows που απαιτούνται για αυτά και επίσης χρήση συναρτήσεων από DLL με δήλωση μόνο μιας γραμμής στον πηγαίο κώδικα. Ως μια γλώσσα Object Pascal, η Delphi υλοποιεί όλες τις απαιτούμενες λειτουργίες για τα αντικείμενα. Κληρονομικότητα, πολυμορφισμός, interfaces, δυναμικού πίνακες (dynamic arrays) αλλά και επιπρόσθετα τον μηχανισμό πολυμορφικής δημιουργίας αντικειμένων virtual construction¹² και την δυνατότητα μεταβλητής που περιέχει κλάση (class reference). Επίσης υπάρχουν ονοματολογικές συμβάσεις (code convention) της Delphi, όπως η ονομασία όλων των κλάσεων

¹⁰ Η εταιρία Altec Software που παράγει τα δημοφιλή προγράμματα «Κεφάλαιο» και «Ατλαντίς» χρησιμοποιεί την Delphi για την ανάπτυξη των εφαρμογών της. Επίσης χρησιμοποιείται από την εταιρία Soft One.

¹¹ IDE = Integrated Development Environment

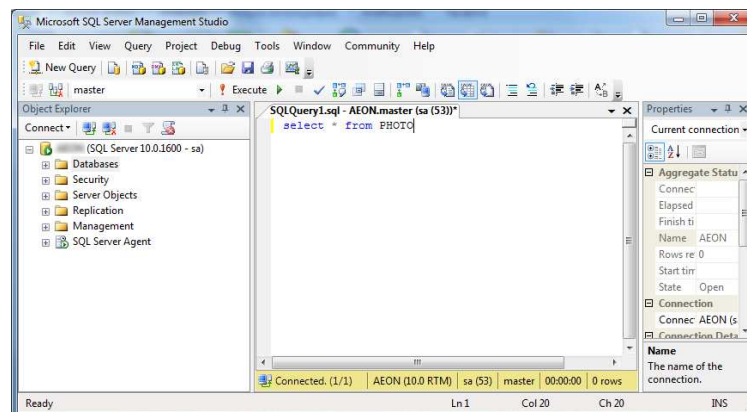
¹² Μια κλάση Delphi μπορεί να έχει έναν constructor που είναι virtual, ώστε να υπάρχει διαφοροποίηση της δημιουργίας του αντικειμένου σε κάθε απόγονο.

με το γράμμα «Τ» ως πρόθεμα και όλων των private πεδίων μιας κλάσης με το γράμμα «F» ως πρόθεμα που προάγουν την χρήση τυποποίησης στον προγραμματισμό.

Σήμερα η Delphi συνεχίζει να αναπτύσσεται κάτω από την σκέπη της Embarcadero Technologies, και με την πολυαναμενόμενη έκδοση XE2, θα έχει δυνατότητες παραγωγής native 64bit εφαρμογών αλλά και μεταφερισιμότητας σε Mac και IOS [6]. Η υλοποίηση του Horasis μπορεί εύκολα να μεταφερθεί σε μια νεότερη έκδοση της Delphi για να εκμεταλλευθεί τις καινούργιες δυνατότητες που προσφέρει, καθώς ούτε η σύνταξη ούτε τα βασικά εργαλεία της βιβλιοθήκης RTL/VCL αλλάζουν από έκδοση σε έκδοση.

Η βάση δεδομένων του συστήματος

Το σύστημα διαχειρίζεται πολλούς διαφορετικούς τύπους δεδομένων, ξεκινώντας από τις φωτογραφίες σε μορφή bitmap ως τις διαμορφώσεις στις οποίες αποθηκεύονται οι καταστάσεις των εκπαιδευμένων δικτύων. Για την αποθήκευση και καλύτερη διαχείριση των δεδομένων, πάρθηκε η απόφαση να χρησιμοποιηθεί μια σχεσιακή βάση δεδομένων (RDBMS¹³) [7] [8]. Η δυνατότητα εύκολης διαχείρισης, μεγάλης απόδοσης σε ταχύτητα αλλά και η πληθώρα τεκμηρίωσης που υπάρχει, οδήγησαν στην επιλογή του Microsoft SQL Server 2008¹⁴ ως την βάση δεδομένων του συστήματος [9]. Ακόμα ένας λόγος ήταν η προγραμματιστική ευκολία που υπάρχει, με την χρήση του ADO/OleDb [10], για πρόσβαση στα δεδομένα μέσω της Delphi. Η έκδοση που χρησιμοποιείται είναι η Express Edition που είναι ελεύθερη για χρήση, έχοντας τις ίδιες δυνατότητες και λειτουργίες με την επί πληρωμή, με μοναδικό περιορισμό την ύπαρξη ενός ορίου 4GB στον όγκο των δεδομένων που μπορεί να αποθηκεύσει.



1-4: SQL Server Management Studio: Ένα ολοκληρωμένο περιβάλλον διαχείρισης και εκτέλεσης ερωτημάτων SQL

Το hardware που χρησιμοποιήθηκε και οι ελάχιστες απαιτήσεις

Ο υπολογιστής που χρησιμοποιήθηκε για την ανάπτυξη του συστήματος και τις εκπαιδεύσεις των δικτύων στηρίζεται στην κεντρική μονάδα επεξεργασίας (CPU) AMD Phenom II x4 965, ο οποίος διαθέτει 4 πυρήνες και συχνότητα λειτουργίας στα 3.4GHz [11]. Η κεντρική μνήμη του υπολογιστή ήταν 4 GB RAM, DDR2. Το σύστημα αξιοποίησε τους επιπλέον πυρήνες τρέχοντας παράλληλες εκπαιδεύσεις στους 3 από τους 4 πυρήνες, έχοντας τον πρώτο ελεύθερο για τον χειρισμό της διασύνδεσης χρήση (UI¹⁵). Για την λήψη των πρότυπων εικόνων χρησιμοποιήθηκε

¹³ RDBS = Relational DataBase Management System

¹⁴ Η έκδοση 2008 δεν έχει μεγάλες διαφορές σε σχέση με την αμέσως προηγούμενη της 2005, αλλά έχει βελτιωμένο ως προς την ευκολία της πρόγραμμα εγκατάστασης (installer)

¹⁵ UI = User Interface

web camera με δυνατότητα λήψης καρέ HD¹⁶ σε ανάλυση 720 γραμμών, ώστε να πάρουμε τελικά μια τετραγωνισμένη η εικόνα στα 500 x 500 pixels. Με αυτό το μέγεθος εικόνας, διαπιστώθηκε στην διάρκεια της ανάπτυξης της εφαρμογής ότι με 4GB μνήμης το σύστημα λειτουργούσε οριακά λόγω του μεγάλου πλήθους των αρχικών δεδομένων. Για καλύτερη εμφάνιση των φορμών της εφαρμογής χρησιμοποιήθηκε ανάλυση 1920 x 1080 σε οθόνη 24" με αναλογία 16:9. Έτσι οι ελάχιστες απαιτήσεις για την λειτουργία του συστήματος, διαμορφώνονται σε PC με 2 πυρήνες, 4 GB RAM άρα και 64bit Windows, HD web camera και οθόνη widescreen με ανάλυση τουλάχιστον 1600x1040.

Η σκηνή φωτογράφισης του Horasis

Για την δυνατότητα εύκολης λήψης φωτογραφιών των αντικειμένων, ώστε να δημιουργείται μια οποιαδήποτε συλλογή από πρότυπα εκπαίδευσης, κατασκευάστηκε μια φωτογραφική σκηνή με απλά υλικά και εργαλεία DIY¹⁷. Αυτή αποτελείται από μια ξύλινη βάση για τις κάμερες, που



1-6: Βάση στήριξης δύο καμερών web με κανάλι-οδηγό
©2011 Horasis

περιλαμβάνει δύο μεταλλικές βέργες πάνω στις οποίες είναι στερεωμένες οι κάμερες. Οι βέργες είναι τοποθετημένες κάθετα ως προς την βάση και βιδώνονται μέσα σε ένα οριζόντιο κανάλι που είναι κομμένο πάνω σε μια επιφάνεια από hardboard¹⁸. Το κανάλι έχει την μορφή ενός τόξου κύκλου και έτσι δίνεται η δυνατότητα για ρύθμιση της σύγκλιση των δύο καμερών, υλοποιώντας μια απλή αντιστοιχία με την θέση των ματιών στο πρόσωπο. Ακριβώς απέναντι και σε μια μικρή απόσταση που εξυπηρετεί την εστίαση των δύο καμερών

στην προεπιλεγμένη σύγκλιση τους, βρίσκεται ένα κυλινδρικό βάθρο, στο οποίο τοποθετούνται τα αντικείμενα που πρόκειται να φωτογραφηθούν. Το βάθρο είναι σταθερά συνδεδεμένο την βάση πάνω σε μια κάθετη ξύλινη δοκό, ώστε τα αντικείμενα να βρίσκονται πάντα στην ίδια θέση ως προς τις κάμερες. Γύρω από το βάθρο υπάρχει ένα παραβάν φτιαγμένο από τρία πλαίσια το οποίο κλείνει την σκηνή, εμποδίζοντας την λήψη πληροφοριών φόντου. Το βάθρο και το παραβάν είναι επενδεδυμένα με μαύρο φωτογραφικό χαρτόνι (παλλάδιο), που έχει την ιδιότητα να απορροφά τις ακτίνες του φωτός εξαλείφοντας τις αντανακλάσεις. Με αυτόν τον τρόπο το περιθώριο γύρω από το αντικείμενο είναι αρκετά απλοποιημένο για να εξυπηρετήσει τους στόχους της εργασίας. Ο φωτισμός μπορεί να είναι διάχυτος ατμοσφαιρικός (ambient), ή εναλλακτικά να τοποθετηθεί μια μικρή βοηθητική πηγή φωτός πίσω από την σκηνή.



1-5: Η σκηνή φωτογράφισης του Horasis
©2011 Horasis

¹⁶ HD = High Definition

¹⁷ DIY = Do It Yourself . στα ελληνικά φτιαχτό μόνος σου, είναι εργαλεία για ατομική και οικιακή χρήση.

¹⁸ Το hardboard είναι ένα συνθετικό υλικό από συμπιεσμένες ίνες ξύλου.

2. Επιστημονικό Πλαίσιο

Γνωστική Επιστήμη, Νευρωνικά Δίκτυα και Συνδεδασμός

Η γνωστική ή γνωσιακή επιστήμη (cognitive science) είναι η διεπιστημονική μελέτη του ανθρώπινου νου και νοημοσύνης που εναγκαλίζει την φιλοσοφία, ψυχολογία, νευροεπιστήμη, γλωσσολογία, ανθρωπολογία αλλά και την τεχνητή νοημοσύνη ή απλά ΤΝ (Artificial Intelligence ή AI)[12]. Η τεχνητή νοημοσύνη, όπως ορίστηκε το 1956 από τον John McCarthy, είναι η επιστήμη και η μηχανική της δημιουργίας έξυπνων μηχανών και την γνωρίζουμε συνήθως ως ένα κομμάτι της επιστήμης των υπολογιστών (computer science). Σήμερα πλέον αποτελεί μια γέφυρα που ενώνει τα δύο παραπάνω επιστημονικά πεδία και δίνει την δυνατότητα της αξιοποίησης της έρευνας του ενός από το άλλο.

Στην εποχή μας βλέπουμε νέους τομείς να εμφανίζονται όπως η υπολογιστική νευροεπιστήμη (computational neuroscience) που χρησιμοποιεί εκτός από τα μαθηματικά και την φυσική, την επιστήμη των υπολογιστών και την ηλεκτρονική για να περιγράψει και να μοντελοποιήσει την λειτουργία των νευρωνικών κυττάρων, όπως την ανακαλύπτει η έρευνα που γίνεται σε βιολογικούς οργανισμούς. Τα αποτελέσματα αυτής της συμπίεσης αναμένονται εντυπωσιακά, καθώς βρισκόμαστε μόλις στην τρίτη δεκαετία της ύπαρξης της Γνωστικής Επιστήμης βιώνοντας παράλληλα την ωρίμανση της Επιστήμης των Υπολογιστών.

Τα τεχνητά νευρωνικά δίκτυα στην πράξη

Ένα βασικό εργαλείο που χρησιμοποιείται από όλους τους προαναφερθέντες τομείς της επιστημονικής έρευνας και αποτελεί το θέμα της παρούσης εργασίας, είναι το τεχνητό νευρωνικό δίκτυο (artificial neural network) ή ΤΝΔ για συντομία. Τα δίκτυα αυτά είναι υπολογιστικά μοντέλα που προσομοιώνουν την λειτουργία των βιολογικών δικτύων που σχηματίζονται στην πραγματικότητα μεταξύ των κυττάρων του νευρικού συστήματος ή αλλιώς των νευρώνων. Στην ουσία είναι ένα πρόγραμμα υπολογιστή ή ένα ηλεκτρονικό κύκλωμα που λειτουργεί παρόμοια με τους βιολογικούς νευρώνες, μεταφέροντας πληροφορίες από τον ένα τεχνητό νευρώνα στον άλλο. Ανάμεσα στις λειτουργίες που εκτελούν, είναι η προσέγγιση μιας μαθηματικής συνάρτησης, η προσομοίωση κάποιας νοητικής λειτουργίας και η εμφάνιση κάποιας επιθυμητής συμπεριφοράς. Στο Τμήμα Πληροφορικής του Α.Τ.Ε.Ι. Θεσσαλονίκης, διδάσκονται ήδη από την δεκαετία του 1990, δίνοντας τις απαραίτητες βάσεις που χρειάζεται ο οποιοσδήποτε απόφοιτος, ώστε να μπορεί να αναπτύξει τις γνώσεις του περαιτέρω για ερευνητικούς ή για επαγγελματικούς λόγους.

Τα τεχνητά νευρωνικά δίκτυα υπάρχουν εδώ και πολύ καιρό στην επιστήμη, αφού από την δεκαετία του 1940 προτάθηκαν ως θεωρητικά μοντέλα από τον Warren McCulloch και τον Walter Pitts [13] για την μελέτη του ανθρώπινου νου. Μέσα σε 7 δεκαετίες έχουν φτάσει από την θεωρητική σύλληψη και την πρώτη υλοποίηση τους στην πρακτική εφαρμογή τους στην καθημερινή ζωή, σε λογισμικό όπως τα OCR¹⁹ δηλαδή τα συστήματα Οπτικής Αναγνώρισης Χαρακτήρων και στα Συστήματα Αναγνώρισης Προσώπων (Facial Recognition).

¹⁹ OCR = Optical Character Recognition

Εφαρμογές όπως τα OCR έχουν φθάσει μέχρι το επίπεδο του καταναλωτή, έτσι σήμερα μπορεί ο οποιοσδήποτε να προμηθευτεί ένα τέτοιο λογισμικό, που στηρίζει την λειτουργία του σε ένα ΤΝΔ [14], και είναι ικανό να σαρώσει μια τυπογραφική σελίδα και να την μετατρέψει σε κείμενο. Τα αλφάβητα που μπορεί να αναγνωρίζει μπορεί να ανήκουν σε οποιαδήποτε γλώσσα του κόσμου [15]. Επίσης βλέπουμε την ύπαρξη σε πολλά δημόσια έγγραφα και φόρμες, μικρών τετραγωνιδίων μέσα στα οποία πρέπει να γράψουμε τα γράμματα όταν συμπληρώνουμε για παράδειγμα το όνομα μας. Αυτά υποδεικνύουν ότι έχουν σχεδιαστεί, ώστε το έγγραφο να καταχωρηθεί αυτοματοποιημένα, με σάρωση από ένα σύστημα OCR.

The image shows a form titled 'ΕΘΝΙΚΟ ΚΤΗΜΑΤΟΛΟΓΙΟ' (National Land Registry) with the subtitle 'ΔΗΛΩΣΗ ΝΟΜΟΥ 2308/1995 (Α.1)'. It contains several sections for data entry:

- ΔΗΜΟΣ:** _____
- ΤΟΠΙΚΟ ΔΙΑΜΕΡΙΣΜΑ:** _____
- ΚΟΙΝΟΤΗΤΑ:** _____
- ΝΟΜΟΣ:** _____
- ΑΡ. ΠΡΩΤΟΚΟΛΛΟΥ:** _____
- ΣΤΟΙΧΕΙΑ ΦΥΣΙΚΟΥ ΠΡΟΣΩΠΟΥ:**

Επίσημο (I)	Όνομα (II)	Α.Φ.Μ. (II)
Τόπος Γέννησης (II)	Ημερομηνία Γέννησης (II)	Ήλικος / Μέθος / Έτος (II)
Αριθμός Ταυτότητας / Διαβατηρίου (II)	Επίδομα Αρσής (I)	Ημερομηνία Έγγραφης (II)
Επίσημο Πατρός (II)		Ήλικος / Μέθος / Έτος (II)
		Όνομα Πατρός (II)

2-1: Έντυπο του Εθνικού Κτηματολογίου στο οποίο τα στοιχεία του δικαιούχου συμπληρώνονται μέσα σε τετραγωνίδια



2-2: Το σύμβολο του βιομετρικού διαβατηρίου που εμφανίζεται στο εξώφυλλο των νέων διαβατηρίων

Από την άλλη εφαρμογές όπως η αναγνώριση προσώπων και η αναγνώριση ιριδογραφήματος [16], αφορούν αποκλειστικά μεγάλους οργανισμούς ή κρατικούς φορείς με χαρακτηριστικό

παράδειγμα τα διαβατήρια της Ευρωπαϊκής Ένωσης [17]. Για την φωτογραφία του κατόχου διαβατηρίου πλέον υπάρχουν συγκεκριμένες απαιτήσεις και αυστηρή τυποποίηση, όπως το να έχουν οι γυναίκες τα μαλλιά τους μαζεμένα πίσω και να μην φοριούνται γυαλιά. Από τις φωτογραφίες εξάγονται βιομετρικά χαρακτηριστικά, δηλαδή μοναδικά μετρήσιμα βιολογικά γνωρίσματα που ταυτοποιούν έναν άνθρωπο, όπως η απόσταση μεταξύ του κέντρου των οφθαλμών ή το μήκος του στόματος. Όλα αυτά τροφοδοτούνται σε σύστημα, που ίσως είναι ένα τεχνητό νευρωνικό δίκτυο, και το οποίο έχει την δυνατότητα να αναγνωρίζει το πρόσωπο του κατόχου του διαβατηρίου. Τα βιομετρικά χαρακτηριστικά είναι ήδη αποθηκευμένα με κρυπτογράφηση, σε ένα μικροσίπ μνήμης πάνω στο διαβατήριο [17], οπότε μπορεί να γίνει διασταύρωση ότι το άτομο που φέρει το διαβατήριο, έχει τα ίδια βιομετρικά χαρακτηριστικά με αυτό που υπάρχει στην φωτογραφία του διαβατηρίου.

Συνδεδασμός, Νευροεπιστήμη και Βιομιμητική

Μέσα στους κύκλους της γνωστικής επιστήμης έχει γεννηθεί ένα επιστημονικό ρεύμα που ονομάζεται στα αγγλικά connectionism ένας όρος που στα ελληνικά χρησιμοποιείται ως συνδεδασμός. Ο συνδεδασμός είναι στην ουσία ένα κίνημα της γνωστικής επιστήμης το οποίο επιχειρεί να εξηγήσει τις ανθρώπινες νοητικές ικανότητες χρησιμοποιώντας τα τεχνητά νευρωνικά δίκτυα [18]. Η βασική του θεώρηση είναι ότι όλα τα διανοητικά φαινόμενα μπορούν να περιγραφούν από διασυνδέσεις πολλών απλών μονάδων που δημιουργούν σύνθετα δίκτυα. Αυτός είναι και ο λόγος που ο συνδεδασμός έχει ταυτιστεί με την χρήση της τεχνολογίας των τεχνητών νευρωνικών δικτύων, δημιουργώντας ένα σημείο τομής της γνωστικής επιστήμης με την επιστήμη των υπολογιστών.

Εκτός όμως από την επιστήμη των υπολογιστών η δεύτερη σημαντική επιστήμη με την οποία εφάπτονται τα ΤΝΔ είναι η νευροεπιστήμη. Αυτό γίνεται με την προσπάθεια μοντελοποίησης των πειραματικών παρατηρήσεων, που παράγει η νευροεπιστήμη ανακαλύπτοντας τον βιολογικό κόσμο. Για μια βιολογική οντότητα δημιουργείται μια τεχνητή μονάδα και η παρατηρηθείσα συμπεριφορά στην φύση κωδικοποιείται ως ένα απλό μοντέλο στον υπολογιστή. Για την μελέτη της λειτουργίας ενός συνόλου από νευρώνες που κατέγραψαν βιολογικά

πειράματα της νευροεπιστήμης, ο συνδεσιασμός θα φτιάξει ένα ΤΝΔ που θα λειτουργεί όπως το βιολογικό νευρωνικό δίκτυο. Η κεντρική φιλοσοφία της επιστημονικής μεθόδου που ακολουθείται είναι απλή και στηρίζεται σε μια ευθεία σκέψη που να περιγραφεί απλοϊκά με μια φράση πρώτου προσώπου: **«Παρατηρώ κάτι πως γίνεται στην φύση, ώστε να φτιάξω κάτι τεχνητό που λειτουργεί με τον ίδιο τρόπο, προσπαθώντας να έχω το ίδιο αποτέλεσμα»**



2-4: Φυσικές σπείρες (arbyreed/Flickr)

Η ίδια ακριβώς σκέψη είναι και αυτή που οδηγεί την Βιομιμητική (Biomimicry) [19]. Η Βιομιμητική παρατηρεί την φύση ως προς τα μοντέλα, συστήματα, διαδικασίες και στοιχεία για να εμπνευστεί από αυτά παράγοντας λύσεις σε ανθρώπινα προβλήματα. Και αυτό γιατί η φύση μέσα από 3.8 δισεκατομμύρια χρόνια εξέλιξης, έχει ήδη δημιουργήσει αποδοτικές λύσεις σε προβλήματα που απασχολούν τον άνθρωπο. Ιστορικά μπορεί να αναχθεί προς τα πίσω από τον Leonardo Davinci που παρατηρούσε

τα πουλιά [20] για να δημιουργήσει την δική του ιπτάμενη μηχανή, τον κοχλία του Αρχιμήδη [21][22] που υλοποιούσε τις σπειροειδείς μορφές που παρατηρούνταν στην φύση, μέχρι τον αρχέγονο άνθρωπο που έφτιαξε τα πρώτα πριόνια όπως τις οδοντοστοιχίες των σαρκοφάγων αρπακτικών ζώων. Αυτό ο τρόπος σκέψης έχει βοηθήσει ή έχει παράγει απευθείας τρανταχτά τεχνολογικά επιτεύγματα. Οι παρατηρήσεις των αδερφών Wright πάνω στην πτήση των περιστερών, βοήθησαν στην δημιουργία του πρώτου επιτυχημένου σχεδίου αεροπλάνου που πέταξε το 1904. Αργότερα το 1940 έγινε μια ανακάλυψη καθαρά στηριγμένη στην βιομιμητική, το Velcro²⁰ από τον Ελβετό επιστήμονα George De Mestral. Παρατήρησε το πως οι κολλισίδες²¹, κολλούσαν στα ρούχα και στο τρίχωμα των σκύλων του, λόγω των μικροσκοπικών αγκίστρων που είχαν, κατασκευάζοντας ακριβώς την ίδια υφή στο Velcro [23]. Ένα σύγχρονο και τρανταχτό παράδειγμα είναι η νέα στολή κολύμβησης Speedo LZR Racer που παρουσιάστηκε το 2008 [24] και υλοποιεί τον τρόπο λειτουργίας των λεπιών του καρχαρία. Με αυτόν τον τρόπο όπως και στους καρχαρίες, μειώνεται η αντίσταση του κολυμβητή. Το αποτέλεσμα της τεχνολογίας αυτή φάνηκε στην Ολυμπιάδα του Πεκίνου, με το 88% των μεταλλίων να πηγαίνει σε αθλητές που την χρησιμοποιούσαν και με την οποία έσπασαν 23 παγκόσμια ρεκόρ στην κολύμβηση [25].



2-3: Φυσικοί γάντζοι (Zephyris/wikimedia commons)

Όλα αυτά δίνουν ελπίδες για πρόοδο και στον χώρο της τεχνητής νοημοσύνης, μέσω του συνδεσιασμού που ασκεί στην ουσία βιομιμητική. Η διαφορά μεταξύ των δύο έγκειται στο ότι με τον συνδεσιασμό ενδιαφερόμαστε κυρίως για την μελέτη του τελικού αποτελέσματος της διάνοησης και όχι αποκλειστικά με την μηχανική που αυτή την διέπει. Κυρίως γιατί η μηχανική αυτή είναι υπερβολικά σύνθετη και είναι αδύνατον με τα σημερινά μέσα να κατανοηθεί και να προσομοιωθεί πλήρως από την γνωστική επιστήμη. Όμως με τα διαθέσιμα ερευνητικά μέσα της νευροεπιστήμης να βελτιώνονται συνεχώς τα τελευταία χρόνια, δίνεται η δυνατότητα για πειράματα που πριν θα ήταν αδύνατο να διεξαχθούν [26]. Η μελλοντική πρόκληση είναι να μεταφερθούν τα αποτελέσματα που θα παραχθούν στον υπολογιστή, δημιουργώντας μια νέα γενιά τεχνητών νευρωνικών δικτύων.

²⁰ Το Velcro είναι ένα είδος συνδετικού που συναντούμε συνήθως σε ρούχα και παπούτσια.

²¹ Κολλισίδες είναι οι καρποί ορισμένων φυτών που έχουν γάντζους, όπως το φυτό Εχίνωψ.

Η επιστημονική τοποθέτηση της εργασίας Horasis

Η εργασία Horasis ακολουθεί την επιστημονική θεώρηση του συνδεδιασμού και προσπαθεί να μεταφέρει με ένα απλοποιημένο τρόπο τις σημαντικές ανακαλύψεις για την λειτουργία της ανθρώπινης όρασης, που έχουν γίνει μέσω της νευροεπιστήμης, ως ένα πλήρως λειτουργικό λογισμικό υπολογιστών. Οι παρατηρήσεις αυτές αφορούν το ανθρώπινο μάτι με τους νευρώνες που υπάρχουν σε αυτό καθώς και μερικούς από τους πρώτους νευρώνες που λειτουργούν στον εγκέφαλο για την αίσθηση της όρασης. Η μεταφορά τους επιχειρείται με την υλοποίηση αντίστοιχων ΤΝΔ ανά περίπτωση, που ενώνονται όλα σε ένα μεγάλο τεχνητό ΤΝΔ, όπως ακριβώς συμβαίνει και με τους βιολογικούς νευρώνες της όρασης. Το «Παρατηρώ κάτι πως γίνεται στην φύση, ώστε να φτιάξω κάτι τεχνητό που λειτουργεί με τον ίδιο τρόπο, προσπαθώντας να έχω το ίδιο αποτέλεσμα» είναι η ουσία της ιδέας που καθοδηγεί αυτήν την προσπάθεια.

Υπάρχει μια αρκετά καλή περιγραφή σε βιολογικούς όρους, της λειτουργίας του αμφιβληστροειδούς χιτώνα, του τμήματος του ματιού στο οποίο καταλήγει το φως, και στο οποίο γίνεται η επεξεργασία του. Έτσι γνωρίζουμε για κύτταρα φωτούποδοχείς του ματιού που διακρίνονται σε ευαίσθητους στο χρώμα και ευαίσθητους στην φωτεινότητα αλλά και για τους νευρώνες που υπάρχουν πιο κάτω από αυτούς. Στο τελευταίο επίπεδο έχουμε νευρώνες που αποτελούν και την «έξοδο» του ματιού οι οποίοι στέλνουν τα δεδομένα στον εγκέφαλο, μέσω του οπτικού νεύρου, σε ένα τμήμα του που ονομάζεται οπτικός φλοιός. Στο πρώτο τμήμα του οπτικού φλοιού γίνεται μια επιπλέον επεξεργασία για την εξαγωγή σχημάτων από τα χρωματικά δεδομένα, που επεξεργάστηκε ο αμφιβληστροειδής. Η «έξοδος» του πρώτου τμήματος του οπτικού φλοιού χρησιμοποιείται ως η πρώτη ύλη για την αναγνώριση των οπτικών προτύπων από τον εγκέφαλο.

Έχοντας το παράδειγμα της βιομημητικής, θεωρούμε ότι τα pixels της εικόνας μας, είναι οι φωτούποδοχείς, και στήνουμε από εκεί και στο εξής τεχνητούς νευρώνες σε υποδίκτυα για το κάθε είδος βιολογικού νευρώνα που βρίσκεται στον αμφιβληστροειδή, όπως και για τους πρώτους που υπάρχουν στον εγκέφαλο, δηλαδή στον πρωτοταγή οπτικό φλοιό. Τα δεδομένα που θα παράγει ως έξοδο το ΤΝΔ του Horasis, θα χρησιμοποιηθούν κατά αντιστοιχία ως οπτικά πρότυπα για την εκπαίδευση ενός δικτύου ταξινόμησης των σε κατηγορίες.

Η τεχνητή όραση και η εργασία Horasis

Μπορούμε να κατατάξουμε την εργασία στις προσπάθειες που γίνονται πάνω στο θέμα της τεχνητής όρασης (computer vision) με την οποία γίνεται η προσπάθεια εξαγωγής πληροφοριών από ένα οπτικό ερέθισμα, δηλαδή μια φωτογραφική σκηνή. Η γενικότερη ιδέα να κάνουμε το υπολογιστικό σύστημα να «κατανοήσει» κάποια πράγματα μέσα σε αυτήν. Απλοϊκά θα λέγαμε ότι η τεχνητή όραση προσπαθεί να κάνει τον υπολογιστή ή το ρομπότ να «δει».

Η όραση είναι μια εξαιρετικά σύνθετη διαδικασία με πολύ περισσότερες παραμέτρους και διαδικασίες, από αυτές που αντιμετωπίζει το Horasis στα πλαίσια μιας πτυχιακής εργασίας. Υπάρχουν διαδικασίες όπως οι κινήσεις σύγκλισης και απόκλισης των οφθαλμών που γίνονται ανάλογα με την απόσταση ενός αντικειμένου από τα μάτια. Επίσης σημαντικές είναι οι σακκαδικές κινήσεις των οφθαλμών (saccades) με τις οποίες «σαρώνουμε» τις λεπτομέρειες πάνω σε μια οπτική σκηνή [27]. Είναι οι γρηγορότερες κινήσεις από όλες τις κινήσεις του ανθρώπινου σώματος με ταχύτητα περιστροφής 700°/sec και διάρκεια 50 msec, με τις οποίες τοποθετείται το κέντρο του ματιού σε διαφορετικά σημεία τις εικόνας επικεντρώνοντας έτσι στις

λεπτομέρειες που υπάρχουν σε αυτά. Όπως ήδη είδαμε όλες αυτές τις διαδικασίες έχουν απλοποιηθεί με την σκηνή φωτογράφισης, καθώς υπάρχει προκαθορισμένη σύγκλιση των καμερών. Επίσης όπως θα δούμε θεωρούμε το κέντρο της όρασης σε ένα σταθερό σημείο. Μελλοντικά το Horasis μπορεί να ενσωματωθεί ως τμήμα μια πολυπλοκότερης υλοποίησης με διεπιστημονικό ενδιαφέρον, προσομοιώνοντας λεπτομερέστερα τα βιολογικά χαρακτηριστικά της όρασης, και υλοποιώντας αυτοματισμούς κίνησης των καμερών. Τότε θα πρέπει να υπάρχει ένα σύστημα αυτομάτου ελέγχου, η για συντομία ΣΑΕ, στο οποίο μια εικόνα θα παράγει ανάδραση κίνησης για λήψη νέας εικόνας. Η ανάδραση θα δημιουργείται από την νευρωνική επεξεργασία που γίνεται στο Horasis.

Επειδή στην πραγματικότητα η ανθρώπινη όραση, εμπλέκει ανώτερες νοητικές λειτουργίες, για να προσομοιωθεί τεχνητά, θα χρειαστούν και άλλα ΤΝΔ που δεν αφορούν τα πρώτα στάδια της. Ένα τέτοιο υπερσύστημα που θα έχει πραγματική δυνατότητα τεχνητής όρασης, θα πρέπει να διαχειρίζεται έννοιες όπως για παράδειγμα τα γεωμετρικά σχήματα, οι γεωμετρικοί όγκοι και η υφή αναγνωρίζοντας, ποσοτικοποιώντας, ταξινομώντας, συγκρίνοντας και γενικεύοντας πολύ πριν να λάβει οπτικά ερεθίσματα που τις περιέχουν.

Η πιο σημαντική όμως λειτουργία για την όραση, που ίσως θεωρείται αδιανόητη να παραχθεί τεχνητά είναι η φαντασία. Βλέποντας μια στατική φωτογραφία όπως την παρακάτω, με την φαντασία μας νιώθουμε το δελφίνι να κινείται. Έτσι σχηματίζουμε μια νέα εικόνα στο μυαλό μας χωρίς να την έχουμε λάβει από τα μάτια μας, με το κεφάλι του δελφινιού μέσα στο νερό και την ουρά του έξω, ενώ για την θάλασσα δεν περιμένουμε να δούμε οποιονδήποτε άλλο χρωματισμό από μπλε, κυανό ή άσπρο.



2-5: Μια απλή εικόνα για την σύνθετη αίσθηση της όρασης

Το πόσο δύσκολο είναι να υλοποιηθεί ένα τέτοιο σύστημα τεχνητής νοημοσύνης στον υπολογιστή, που θα έχει πραγματική τεχνητή όραση, μπορεί να το αναλογιστεί κάποιος αν σκεφτεί το εξής: Ένα παιδί στην πρώτη τάξη του δημοτικού, επειδή καθοδηγείται αποκλειστικά από τα οπτικά χαρακτηριστικά του δελφινιού, ακόμα και αν του διδαχθεί ότι γεννάει τα παιδιά του και ότι αναπνέει με πνεύμονες το κατατάσσει στα ψάρια [28].

Σύγχρονος Αντικειμενοστραφής Προγραμματισμός

Ο προγραμματισμός έχει περάσει προ πολλού από την φάση στην οποία αντιπροσώπευε μια ακολουθία εντολών, και μια συρραφή από διαδικασίες που συνδυαζόταν μεταξύ τους για να υλοποιηθεί η γνωστή ροή είσοδος → επεξεργασία → έξοδος. Όσο η πολυπλοκότητα του υπό ανάπτυξη λογισμικού αυξανόταν, τόσο αυξανόταν και οι ανάγκες για γρηγορότερη υλοποίηση έχοντας παράλληλα την προοπτική της εύκολης συντήρησης. Τα οφέλη από την χρήση του αντικειμενοστραφούς προγραμματισμού[29] και των σχετικών τεχνολογιών που περιστρέφονται γύρω από αυτόν, προέβαλλαν ξεκάθαρα και αποτέλεσαν τον βασικό λόγο για την αλματώδη πρόοδο στο λογισμικό που έγινε τις τελευταίες δύο δεκαετίες.

Με την εμφώλευση (encapsulation) [30] δημιουργούνται αυτόνομα αντικείμενα που δεν χρειαζόταν να βλέπουν ολόκληρο τον πηγαίο κώδικα του λογισμικού για να εκτελέσουν τον σκοπό τους. Έτσι μειώνεται η πολυπλοκότητα αλλά και η σύζευξη (coupling) των διαφορετικών τμημάτων του κώδικα. Επίσης η εμφώλευση μαζί με την κληρονομικότητα, βοήθησαν στην δημιουργία επαναχρησιμοποιούμενων κλάσεων αντικειμένων. Μέσω της κληρονομικότητας όλη οι κοινόχρηστη λογική σε μια ιεραρχία αντικειμένων τοποθετείται στον πρόγονο, και μόνο οι διαφοροποιήσεις από αυτή υλοποιούνται στους απογόνους. Έτσι επιταχύνεται σημαντικά η ανάπτυξη του κώδικα καθώς με σωστή οργάνωση των κλάσεων, μπορεί να αποφευχθεί η επανάληψη στην υλοποίηση, που μαστίζε τον κλασικό προγραμματισμό. Η γενικότερη ευταξία του κώδικα έκανε την δουλειά του προγραμματιστή ως προς την συντήρηση και αναβάθμιση του κώδικα σημαντικά ευκολότερη.

Τα αντικείμενα τα τελευταία χρόνια άρχισαν να μπαίνουν σταδιακά και σε τομείς που εδώ και χρόνια ήταν υποχρεωτική η χρήση του κλασικού διαδικαστικού προγραμματισμού, όπως στα λειτουργικά συστήματα. Οι βιβλιοθήκες δυναμικής διασύνδεσης, οι γνωστές DLL ²² του λειτουργικού συστήματος Windows, είναι στην ουσία συλλογές από συναρτήσεις στις οποίες περνιούνται ως ορίσματα τα δεδομένα. Αυτές οι ρουτίνες εξάγονται προς τα έξω ώστε να χρησιμοποιούνται από τις εφαρμογές. Πλέον με την σταδιακή μετατροπή τους σε .NET assemblies²³, εξάγουν κλάσεις αντικειμένων οι οποίες πλέον μεθόδους στην θέση των απλών συναρτήσεων [31].

Ως τελευταία εξέλιξη, στον τομέα των λειτουργικών που αποτελεί ένα προπύργιο του κλασικού προγραμματισμού, παραθέτουμε ένα λειτουργικό σύστημα το οποίο αποτελείται εξολοκλήρου από αντικείμενα. Το Microsoft Singularity [32] που ξεκίνησε το 2003 και βρίσκεται έως σήμερα σε πειραματικό στάδιο, μπορεί σε λίγα χρόνια να το δούμε να λειτουργεί στο PC μας. Λόγω της αντικειμενοστραφούς αρχιτεκτονικής του και υλοποίησης σε C#, υποστηρίζει Απομονωμένες Διεργασίες Λογισμικού (SIP²⁴). Η κάθε μια από αυτή έχει δική της αυτόματη διαχείριση μνήμης με garbage collector ²⁵προστατεύοντας το σύστημα από διαρροές μνήμης²⁶ (memory leaks) και

²² **DLL = Dynamic Link Library**

²³ Η assembly είναι η βιβλιοθήκη δυναμικής διασύνδεσης του .NET που είναι μεταγλωττισμένες σε ενδιάμεση γλώσσα MSIL σε αντίθεση με τις κλασικές DLL που είναι σε γλώσσα μηχανής x86.

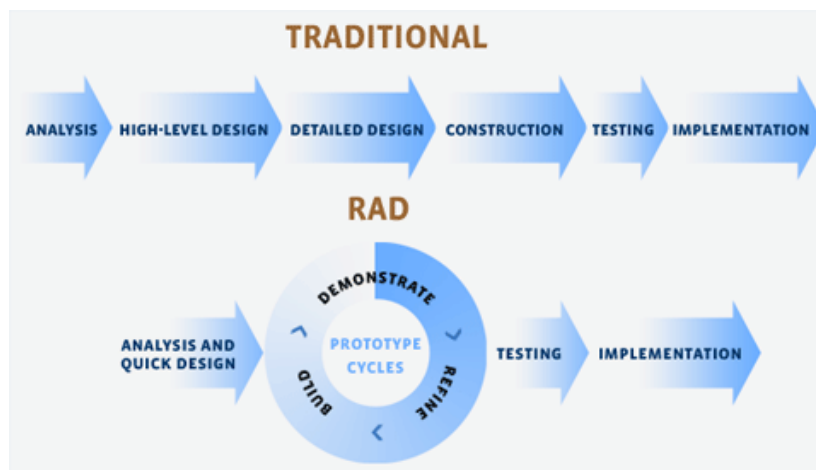
²⁴ **SIP = Software Isolated Process** είναι διεργασίες με αυτόνομα δεδομένα και το κώδικα που επικοινωνούν μεταξύ τους αποκλειστικά μέσω καναλιών επικοινωνίας αποφεύγοντας την κοινή χρήση δεδομένων.

²⁵ Ο Garbage Collector (GC) είναι ένα σύστημα αυτόματης διαχείρισης μνήμης, που στηρίζεται στην μέτρηση αναφορών προς ένα αντικείμενο (reference counting).

την δική της ροή εκτέλεσης [33]. Πρακτικά εφόσον και ο σχεδιασμός hardware το επιτρέψει, θα υπάρχει η δυνατότητα να αντικαθίσταται μια κάρτα γραφικών χωρίς να κλείνει ο υπολογιστής από το ρεύμα ενώ το λειτουργικό σύστημα θα συνεχίζει να τρέχει όλες τις υπόλοιπες λειτουργίες του, καθώς θα απαιτείται μόνο η επανεκκίνηση της SIP που αφορά το υποσύστημα γραφικών [34].

Η φιλοσοφία RAD

Πάνω στον δρόμο της προόδου του αντικειμενοστραφούς προγραμματισμού κινείται η φιλοσοφία της Ταχείας Ανάπτυξης Εφαρμογών (Rapid application development) που είναι γνωστότερη ως RAD. Δημιουργήθηκε ως όρος από τον James Martin το 1990 και είναι ένας κύκλος ανάπτυξης λογισμικού (development lifecycle), ο οποίος στοχεύει στην ταχύτερη ανάπτυξη και ποιοτικότερη παραγωγή εφαρμογών [35]. Βασικό στοιχείο είναι η εναγκαλισμός της μεθοδολογίας του αντικειμενοστραφούς προγραμματισμού και η επαναχρησιμοποίηση συστατικών. Τα προγραμματιστικά περιβάλλοντα RAD όπως η Delphi παρέχουν σημαντικές διευκολύνσεις προς τον προγραμματιστή ενοποιώντας την συγγραφή, μεταγλώττιση αποσφαλμάτωση και χτίσιμο του εκτελέσιμου σε ένα περιβάλλον που λέγεται Ενοποιημένο Περιβάλλον Προγραμματισμού ή IDE από τα αρχικά του Integrated Development Environment. Επίσης τα περιβάλλοντα RAD, συνήθως παρέχουν βιβλιοθήκες που υλοποιούν τα πλέον βασικά και συνηθισμένα εργαλεία που χρειάζεται ένας προγραμματιστής.



2-6: Σύγκριση κλασσικής ανάπτυξης με την RAD

Πλέον η χρήση του RAD για την ανάπτυξη εμπορικών εφαρμογών έχει καθιερωθεί ως de facto²⁷ στην εμπορική ανάπτυξη λογισμικών, κυρίως για την μείωση του κόστους ανάπτυξης και συντήρησης. Όπως μπορούν να προσδώσουν χρηματικά οφέλη για εμπορική χρήση, μπορούν να δημιουργήσουν και ερευνητικά οφέλη, επιτυγχάνοντας την γρήγορη δημιουργία ενός πρωτότυπου συστήματος για κάποια επιστημονική ιδέα. Με αυτήν την λογική αποφασίστηκε να γίνει υλοποίηση του Horasis με εργαλεία RAD λόγω της ανάγκης για ταχεία και εύκολη υλοποίηση. Επιπλέον χρειαζόταν εύκολη διαχείριση της πολυπλοκότητας που προέκυψε με την υλοποίηση της εργασίας. Η δυνατότητα επαναχρησιμοποίησης πολλών από τα συστατικά του Horasis ως βιβλιοθήκη, είναι ένα από τα οφέλη της τεχνολογίας RAD που χρησιμοποιήθηκε.

²⁶ Τα memory leaks είναι μια μάστιγα στην υλοποίηση των εφαρμογών, καθώς ένα συνηθισμένο προγραμματιστικό σφάλμα είναι να μην απελευθερώνεται η μνήμη που χρησιμοποιείται. Σαν γιατρεία προτείνεται η αυτόματη διαχείριση της μνήμης με την χρήση μηχανισμών garbage collection.

²⁷ Λατινικός όρος που σημαίνει: Από τα γεγονότα

Παράλληλος Προγραμματισμός

Ένα ακόμη σημαντικό πλεονεκτήματα του σύγχρονου αντικειμενοστραφή προγραμματισμού, το οποίο αξιοποιεί το Horasis, είναι η ευκολία στην υλοποίηση παράλληλου προγραμματισμού. Η χρήση παράλληλων διεργασιών διευκόλυνε στην διερεύνηση διαφορετικών τύπων ΤΝΔ όπως θα δούμε στα επόμενα κεφάλαια.

Ο παράλληλος προγραμματισμός στα Windows είναι στην ουσία η χρήση των νημάτων (threads) [36]. Στους επεξεργαστές με μόνο ένα πυρήνα²⁸, υπήρχε ψευδοπαράλληλη λειτουργία των νημάτων καθώς για να εκτελεστούν δύο από αυτά, έπρεπε να σταματάει το ένα δίνοντας στην θέση στο άλλο κατά τακτά, καθοριζόμενα από το λειτουργικό σύστημα, χρονικά διαστήματα. Για αλγόριθμους που από την φύση τους χρησιμοποιούσαν όλη την επεξεργαστική ισχύ του πυρήνα όπως μαθηματικές πράξεις, αυτό το γεγονός όχι μόνο δεν ωφελούσε χρονικά, αλλά μπορούσε να παράγει και επιπλέον καθυστερήσεις. Πλέον με τους σύγχρονους πολυπύρηνους επεξεργαστές το κάθε νήμα μπορεί να λειτουργήσει πραγματικά παράλληλα, «φορτωμένο» σε έναν πυρήνα του οποίου θα καταναλώνει όλη την επεξεργαστική ισχύ.

Ο παράλληλος προγραμματισμός μπορεί να έχει τα σημαντικά οφέλη στην ταχύτητα, αλλά η υλοποίηση του ήταν αρκετά δύσκολη στο παρελθόν. Έπρεπε κάποιος να γνωρίζει σε βάθος τον προγραμματισμό σε επίπεδο του λειτουργικού συστήματος και να χρησιμοποιεί δείκτες χειρισμού²⁹ (handles) για την χρήση των νημάτων [36]. Απαιτούνταν επίσης μεγάλη προσοχή στην χρήση των δεδομένων καθώς δύο παράλληλες διαδικασίες δεν πρέπει να τροποποιούν τα ίδια δεδομένα την ίδια χρονική στιγμή. Ο αντικειμενοστραφής προγραμματισμός και η φιλοσοφία RAD απλοποίησε τα πράγματα σημαντικά καθώς πλέον το κάθε νήμα είναι ένα αντικείμενο που έχει τον κώδικα της παράλληλης διαδικασίας ως μέθοδο του, δίπλα στα δικά του δεδομένα. Στην Delphi υπάρχει η κλάση `TThread` που αποκρύπτει τον προγραμματισμό με API³⁰ των windows που απαιτείται για την δημιουργία ενός νήματος. Κληρονομώντας από αυτήν μπορούμε στον απόγονο της να υλοποιήσουμε εύκολα την παράλληλη λογική κάνοντας απλά επικάλυψη (override) της μεθόδου `Execute` με τον κώδικα της παράλληλης διαδικασίας. [4]. Επίσης στην έκδοση 4.0 του .NET Framework η Microsoft εισήγαγε ένα νέο προγραμματιστικό μοντέλο για την συγγραφή παράλληλου και ασύγχρονου κώδικα που απλοποιεί την εργασία του προγραμματιστή, όπως αυτόματο παραλληλισμό για επαναληπτικούς βρόχους `for` και `foreach` απαλλάσσοντας τον προγραμματιστή από την ευθύνη σχεδιασμού [37].

Η βέλτιστη χρήση των αντικειμένων

Τα αντικείμενα έχουν δώσει όπως είδαμε νέα ώθηση για την πρόοδο του προγραμματισμού. Από την χρήση τους, είδαμε τα πρώτα οφέλη του αντικειμενοστραφούς σε σχέση με τον κλασικό προγραμματισμό. Η τελευταία όμως τάση δεν είναι απλώς και μόνο η χρήση τους για την συγγραφή των προγραμμάτων, αλλά καλύτερη δυνατή οργάνωση της χρήσης τους. Πλέον από την μεθοδολογία του αντικειμενοστραφούς προγραμματισμού έχουμε μεταβεί στην προσέγγιση της αντικειμενοστραφούς ανάλυσης και σχεδιασμού (OOAD³¹), με την οποία ασχολείται ο τομέας της μηχανικής λογισμικού [30]. Με αυτήν οποιοδήποτε σύστημα μοντελοποιείται ως μια ομάδα από αντικείμενα που αλληλεπιδρούν και επικοινωνούν και μεταξύ τους. Κάθε αντικείμενο

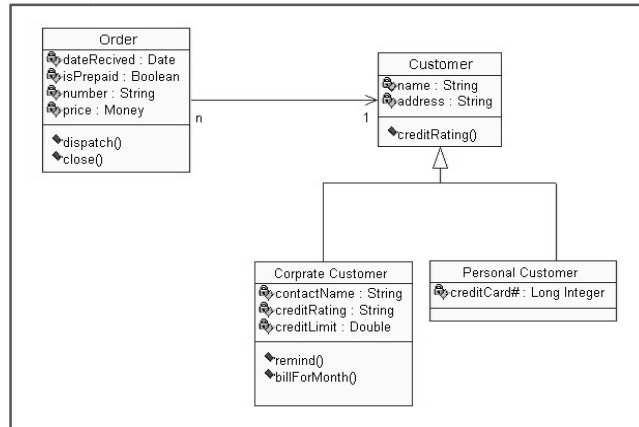
²⁸ Ο πυρήνας επεξεργαστή έχει αυτόνομη μονάδα ελέγχου (CU), αριθμητική και λογική μονάδα (ALU) και δικό του cache. Είναι στην ουσία ένας αυτόνομος επεξεργαστής.

²⁹ Τα handles είναι ένα από τα αντικείμενα του λειτουργικού συστήματος Windows

³⁰ **API** = **A**pplication **P**rogramming **I**nterface

³¹ **OOAD** = **O**bject **O**riented **A**nalysis and **D**esign

αντιπροσωπεύει κάποια οντότητα με συγκεκριμένο σκοπό ύπαρξης στο σύστημα και χαρακτηρίζεται από την κλάση στην οποία ανήκει, την κατάσταση του και την συμπεριφορά του. Τα μοντέλα σχεδιάζονται όχι ως προς την υλοποίηση των αντικειμένων, αλλά ως προς την δομή, την οργάνωση τους, την δυναμική τους συμπεριφορά και τις σχέσεις τους. Για την καταγραφή τους χρησιμοποιούνται πλέον τυποποιημένες σημειογραφίες όπως η Ενοποιημένη Γλώσσα Μοντελοποίησης UML³² που βοηθούν στην κατανόηση του μοντέλου πριν την υλοποίηση [21].



2-7: Διάγραμμα Κλάσεων UML (Fowler & Scott)

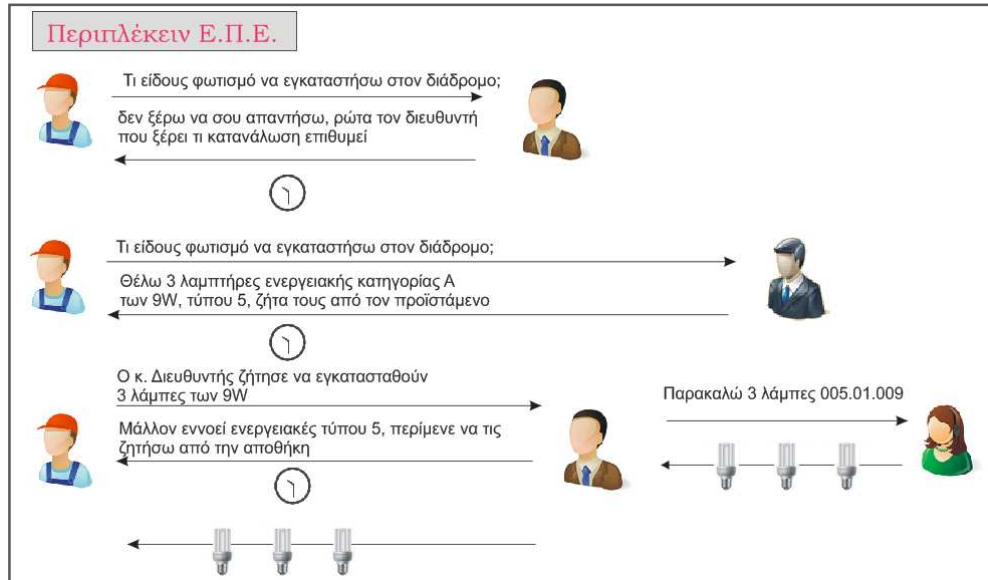
Κατά τον διαδικασία δημιουργίας του μοντέλου λαμβάνονται πλέον σχεδιαστικές αποφάσεις σε αντιστοιχία με τις προγραμματιστικές αποφάσεις που λαμβάνονται όταν αναπτύσσεται ο πηγαίος κώδικας. Οι αποφάσεις αφορούν στην μείωση της πολυπλοκότητας του μοντέλου και την αύξηση της αποδοτικότητας του, ώστε να επιτύχει τον στόχο του. Δύο σημαντικοί τρόποι βελτίωσης που μας δίνει η μεθοδολογία είναι η μείωση της σύζευξης (coupling) και η αύξηση της συνοχής (cohesion)[38], που είναι οι δύο βασικοί δείκτες της σωστής κατάτμησης του συστήματος σε αντικείμενα. Σύμφωνα με ειδικούς και ερευνητές πληροφορικής η επίτευξη μικρής σύζευξης και υψηλής συνοχής οδηγεί σε πιο αξιόπιστα και πιο συντηρήσιμα συστήματα [39].

Για το πρώτο από τα δύο, τα αντικείμενα πρέπει να δημιουργούνται με όσο το δυνατόν λιγότερες εξαρτήσεις μεταξύ τους. Δηλαδή ένα αντικείμενο ή τμήμα του συστήματος πρέπει να είναι όσο το δυνατόν αυτόνομο, και να μην εξαρτάται από πολλά άλλα για να επιτελέσει το σκοπό του. Με την χαλαρή σύζευξη (loose coupling) όταν γίνεται μια αλλαγή σε ένα αντικείμενο, αποφεύγεται ένα κύμα αλλαγών σε πολλά εξαρτώμενα αντικείμενα καθιστώντας ευκολότερη την συντήρηση. Επίσης διευκολύνεται η συναρμολόγηση των τμημάτων του συστήματος καθώς τα αυτόνομα αντικείμενα λειτουργούν ως μικρά «τουβλάκια» που εύκολα συνδέονται μεταξύ τους. Τέλος διευκολύνεται η επαναχρησιμοποίηση, επειδή τα αντικείμενα μπορούν εύκολα να αφαιρεθούν ή να προστεθούν σε οποιοδήποτε σημείο του σχεδιασμού.

Με τον όρο συνοχή περιγράφουμε το πόσο σχετική μεταξύ τους είναι η λειτουργία των αντικειμένων ενός συγκεκριμένου τμήματος λογισμικού. Επίσης δείχνει το πόσο σχετιζόμενες μεταξύ τους είναι οι ευθύνες με τις οποίες είναι επιφορτισμένο ένα αντικείμενο. Αν οι διαφορετικές μεθόδους που υλοποιεί ένα αντικείμενο έχουν πολύ λίγη σχέση μεταξύ τους τότε μιλάμε για χαμηλή συνοχή. Αν μια συγκεκριμένη εκτελεί λειτουργίες πάνω σε άσχετα μεταξύ τους αντικείμενα έχουμε πάλι την ίδια χαμηλή συνοχή. Η αύξηση της συνοχής βοηθάει στην

³² UML = Unified Modeling Language

καλύτερη κατανόηση της λειτουργίας των επιμέρους τμημάτων του συστήματος. Η συντήρηση γίνεται και εδώ ευκολότερη αποφεύγοντας το ίδιο κύμα αλλαγών σε άσχετα μεταξύ τους αντικείμενα. Έχοντας υψηλή συνοχή έχουμε ακόμη πιο καλοσχεδιασμένα αντικείμενα που έχουν συγκεκριμένες και ευκολονόητες ευθύνες λειτουργίας και σαφώς καθορισμένο λόγο ύπαρξης[40]. Με τα παρακάτω παραδείγματα, συγκρίνεται η σύζευξη και η συνοχή, μέσα από δύο διαφορετικές προσεγγίσεις της ίδιας απλής διαδικασίας, σε δύο φανταστικές εταιρίες.



2-8: Υψηλή σύζευξη και χαμηλή συνοχή ©2011 Horasis

Η απλή διαδικασία της αλλαγής λαμπτήρων, η ροή εργασίας που δείχνει η εικόνα 2-8 απαιτεί περισσότερη επικοινωνία και μεταξύ ατόμων που δεν θα έπρεπε να την έχουν (τεχνικός με διευθυντή) εμφανίζοντας υψηλή σύζευξη, και παράλληλα ασάφεια στις αρμοδιότητες του καθενός, δηλαδή χαμηλή συνοχή. Το απτό αρνητικό αποτέλεσμα είναι η αύξηση του χρόνου εκτέλεσης.



2-9: Χαμηλή σύζευξη και υψηλή συνοχή ©2011 Horasis

Στην εικόνα 2-9 υπάρχει χαμηλότερη σύζευξη γιατί οι γραμμές επικοινωνίας είναι λιγότερες, ενώ παράλληλα η επικοινωνία δεν επεκτείνεται σε παραπάνω συνομιλητές από όσους χρειάζονται, άρα υπάρχει και υψηλή συνοχή. Επίσης μειωμένη σύζευξη είναι η αποφυγή δεύτερης

επικοινωνίας για το ίδιο θέμα μεταξύ του τεχνικού και του προϊσταμένου. Εκτός της ταχύτητας κερδίζουμε και στην σαφήνεια με την διαμεσολάβηση μόνο του προϊσταμένου, καθώς δεν μπορεί ο τεχνικός να αλλοιώσει την εντολή όπως στην εικόνα 2-8.

Οι ομοιότητες της σωστής οργάνωσης των αντικείμενων με την οργάνωση των βιολογικών μονάδων

Με αυτές και με πολλές άλλες κατευθύνσεις που υπάρχουν, μπορούμε να σχεδιάσουμε αντικείμενα λογισμικού που να λειτουργούν ως αυτόνομες βιολογικές μονάδες. Ας δούμε τον σχεδιασμό των αντικειμένων σε αντιστοιχία με κάποια όργανα του ανθρώπινου οργανισμού για να καταλάβουμε καλύτερα αυτήν την θεώρηση. Ο εγκέφαλος αν και το σημαντικότερο όργανο στο σύστημα δεν έχει καθόλου σύζευξη με την καρδιά καθώς ο καρδιακός μυς δεν περιμένει εντολές από τον εγκέφαλο για να κινηθεί, αλλά λειτουργεί ακουσίως. Ως προς τον πνεύμονα έχει



2-10: Πνεύμονες (Kurimoto, Fielding ,Musani)

μειωμένη σύζευξη καθώς η αναπνοή είναι ακούσια και εκούσια. Ένας συγκεκριμένος μυς του πνεύμονα έχει την δυνατότητα να ελεγχθεί από τον εγκέφαλο ώστε να κρατήσουμε την αναπνοή μας ή να εκπνεύσουμε τον αέρα από αυτόν. Από την άλλη ο πνεύμονας όπως και όλα τα όργανα έχει αυξημένη συνοχή καθώς έχει αποκλειστική ευθύνη για την λειτουργία του οξυγόνωσης του αίματος και της ταυτόχρονης αποβολής του CO₂ χωρίς να την έχει καταναμεμημένη ή μοιρασμένη με άλλα σημαντικά όργανα. Οι πνευμονικές κυψελίδες που στις οποίες γίνεται τελικά η λειτουργία της ανταλλαγής των αερίων στο αίμα, μπορούν να αντιστοιχηθούν με ένα σύνολο υποαντικείμενων που έχει ένα αντικείμενο για την εκτέλεση της ευθύνης που έχει.

Επίσης σημαντικές είναι και οι σχέσεις μεταξύ των αντικειμένων όπως οι δύο βασικές δηλαδή της σύστασης (composition) και της συνάθροισης (aggregation) [41]. Η πρώτη σχέση είναι το παράδειγμα του πνεύμονα με τις πνευμονικές κυψελίδες και υποδηλώνει ότι ο κύκλος ζωής των συστατικών είναι συνυφασμένος με την ζωή του ίδιου του αντικειμένου. Στην συνάθροιση τα υποαντικείμενα μπορούν να καταστρέφονται ή να αντικαθίστανται και ένα αντίστοιχο παράδειγμα που θα μπορούσαμε να δώσουμε για το ανθρώπινο σώμα είναι το δέρμα. Ενώ ένα μέρος του σώματος όπως το χέρι συνεχίζει να υπάρχει το δέρμα μπορεί να καταστραφεί από τραυματισμό, να ανανεωθεί από μόνο του ή ακόμα να μεταμοσχευθεί από άλλο σημείο του σώματος στην περίπτωση πλαστικής επέμβασης.

Ακολουθώντας τις παραπάνω μεθοδολογίες που δίνει η μηχανική λογισμικού κατά τον σχεδιασμό του Horasis, έγινε η προσπάθεια να δημιουργηθεί ένα αποδοτικός αντικειμενοστραφής σχεδιασμός για όλα τα αντικείμενα. Ακριβώς πάνω σε αυτήν την προγραμματιστική φιλοσοφία στηρίχθηκε η υλοποίηση του αντικειμενοστραφή νευρώνα. Τα υποαντικείμενα που έχει είναι εύκολα και κατανοητά, και αντιστοιχούν στα βιολογικά κομμάτια που αποτελούν τον πραγματικό νευρώνα. Έτσι μεταξύ άλλων η εργασία προσπαθεί να προάγει την χρήση των αρχών της μηχανικής λογισμικού, μεταφέροντας τις στην υλοποίηση των τεχνητών νευρωνικών δικτύων.

Προοπτικές για την πρόοδο των τεχνητών νευρωνικών δικτύων

Η μελλοντική πρόοδος του λογισμικού και τι μπορεί να συνεισφέρει

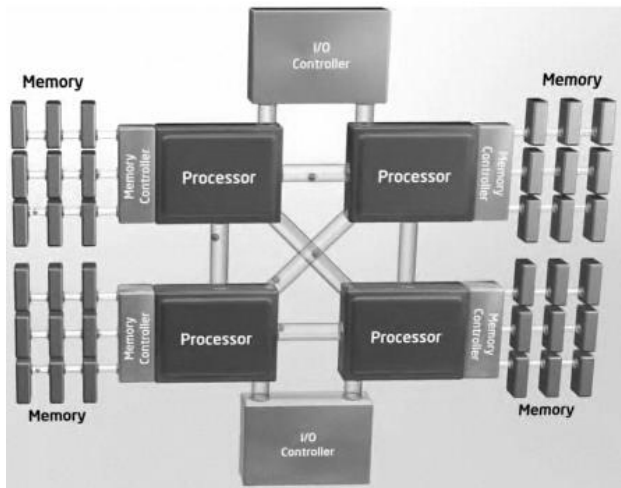
Με την πλήρη αξιοποίηση των σύγχρονων δυνατοτήτων της μηχανικής λογισμικού και του προγραμματισμού, μπορεί να ανοιχθούν νέοι δρόμοι για την υλοποίηση οποιουδήποτε τύπου τεχνητού νευρωνικού δικτύου απαιτηθεί στην τεχνητή νοημοσύνη. Οι παρατηρήσεις της νευροεπιστήμης θα μπορούν επίσης εύκολα να μεταφερθούν δημιουργώντας απόγονες κλάσεις της ίδιας προγονικής κλάσης νευρώνων. Για παράδειγμα αν ανακαλυφθεί ένας νέος τύπος νευρώνα, και χρειαστεί η προσομοίωση της λειτουργίας του, θα χρειαστεί μόνο η τροποποίηση κάποιων μεθόδων σε ένα υποαντικείμενο του νευρώνα κρατώντας τα κοινά χαρακτηριστικά που έχουν όλοι οι νευρώνες.

Εφόσον οι τεχνητοί νευρώνες σχεδιαστούν με την λογική της χαμηλής σύζευξης και υψηλής συνοχής, το συγκεκριμένο σημείο στο οποίο θα υλοποιηθεί η διαφοροποίηση δεν θα απαιτήσει ούτε τον επανασχεδιασμό του νευρώνα και πολύ περισσότερο όλου του νευρωνικού δικτύου. Όπως στην πραγματικότητα κάποιες λειτουργίες και συμπεριφορές εμφανίζονται ως κοινά γνωρίσματα και κάποιες μόνο διαφοροποιούνται. Έτσι θα υπάρξει επαναχρησιμοποίηση έτοιμου λογισμικού που μπορεί να επιταχύνει ραγδαία την έρευνα, με ελάχιστους χρόνους δημιουργίας των νέων μοντέλων. Ένα πλαίσιο λογισμικού (software framework) τεχνητών νευρώνων, όπως αυτό που υλοποιείται για την λειτουργία του Horasis, μπορεί να αποτελέσει στο μέλλον αναπόσπαστο κομμάτι της βιβλιοθήκης των προγραμματιστικών εργαλείων RAD, όπως σήμερα αποτελεί η διασύνδεση με βάσεις δεδομένων που δεν υπήρχε στην αρχή της δεκαετίας του 1990. Αυτό θα δημιουργήσει εκτός από την δυνατότητα ταχείας υλοποίησης των τεχνητών νευρωνικών δικτύων, την δυνατότητα ενσωμάτωση τους σε οποιαδήποτε εφαρμογή άμεσα, εύκολα και αποκρύπτοντας τις λεπτομέρειες υλοποίησης τους από τον προγραμματιστή.

Όπως θα δούμε παρακάτω όταν γνωρίσουμε αναλυτικά τα αντικείμενα των νευρώνων του Horasis, νέοι τρόποι λειτουργίας των νευρωνικών δικτύων μπορούν να υλοποιηθούν, με τα αντικείμενα ΤΝΔ πολύ πριν γίνει δυνατή η διατύπωση του ως μαθηματικά μοντέλα. Επίσης θα δούμε ότι τα υπάρχοντα μαθηματικά μοντέλα εκπαίδευσης των δικτύων όπως το Back-Propagation, μπορούν να εμπλουτιστούν με την χρήση των αντικειμένων.

Η σημασία της αύξησης της υπολογιστής ισχύος και προοπτικές για μελλοντικό υλικό.

Τα πρακτικά επιτεύγματα των τεχνητών νευρωνικών δικτύων, έγιναν ως ένα βαθμό δυνατά και λόγω της ραγδαίας αύξησης της υπολογιστικής ισχύος, που ήταν αποτέλεσμα της προόδου στην Φυσική και την Ηλεκτρονική. Όπως θα δούμε παρακάτω η εκπαίδευση και η λειτουργία των ΤΝΔ εκτελεί αναρίθμητους και χρονοβόρους μαθηματικούς υπολογισμούς. Επίσης όσο ανεβαίνει το πλήθος των νευρώνων ενός ΤΝΔ, το οποίο μπορεί να γίνει αρκετά μεγάλο, η απαίτηση σε υπολογιστική ισχύ και μνήμη αυξάνει εκθετικά. Ένα απλό παράδειγμα που δείχνει το πόσο διαθέσιμη έχει γίνει αυτή η απαιτούμενη ισχύ είναι να δούμε τον πιο γρήγορο επεξεργαστή που υπάρχει σήμερα [42], στο τρίτο τέταρτο του 2011, για απλούς προσωπικούς υπολογιστές PC.



2-11: Σχηματική αναπαράσταση τετραπύρηνου CPU
(www.newtechnology.co.in)

Ο Intel Core i7 990X είναι ένας πολυεπεξεργαστής που έχει 4 πυρήνες παράλληλης επεξεργασίας στο ίδιο PGA³³ chip και τρέχει σε συχνότητα λειτουργίας 3.47GHz. Σε σχέση με τον γρηγορότερο Intel Pentium 4, την τεχνολογία της ίδιας εταιρία που κυκλοφόρησε από το 2000 έως το 2008, και στην ίδια συχνότητα 3,4GHz³⁴ έχει μετρηθεί ως 20 φορές³⁵ πιο γρήγορος σε δοκιμές ταχύτητας. Επίσης η δυνατότητα παράλληλης επεξεργασίας του i7 είναι πραγματική ενώ το παλαιότερο μοντέλο προσομοίωνε την παράλληλη επεξεργασία καθώς είχε μόνο ένα πυρήνα.

Μια σημαντική προοπτική για το μέλλον είναι κάποιοι μαθηματικοί υπολογισμοί ή και λειτουργίες που εκτελούνται στο λογισμικό των τεχνητών νευρωνικών δικτύων να ενσωματωθούν στα σύνολα εντολών (instruction set) των συνηθισμένων επεξεργαστών. Ήδη υπάρχουν προτάσεις για υλοποίηση σε ένα ψηφιακό ηλεκτρονικό κύκλωμα της σιγμοειδούς συνάρτησης, που όπως θα δούμε έχει ιδιαίτερο ενδιαφέρον για τα νευρωνικά δίκτυα [43] [44]. Το υλικό πλέον ακολουθεί τις επιταγές του λογισμικού, όπως φαίνεται και από το παράδειγμα των επεξεργαστών Physics ή PPU³⁶. Με αυτούς υλοποιούνται ως υλικό και έτσι επιταχύνονται οι υπολογισμοί που χρησιμοποιούνται στα μοντέλα των τρισδιάστατων παιχνιδιών για να μετακινούνται, αλληλεπιδρούν και αντιδρούν τα αντικείμενα μέσα στο περιβάλλον του παιχνιδιού [45]. Αντίστοιχα θα μπορούσαν να υλοποιηθούν **Επεξεργαστές Νευρωνικών Δικτύων** (Neural Network Processing Units – NNPU), δηλαδή μονάδες που υλοποιούν ή επιταχύνουν υπολογισμούς, μεταφορά πληροφοριών και διαδικασίες εκπαίδευσεων τεχνητών νευρωνικών δικτύων.

Η σημασία της μείωσης του κόστους του υλικού για την αλματώδη πρόοδο των νευρωνικών δικτύων

Παράλληλα με την αύξηση της ταχύτητας του υλικού, εξίσου σημαντικό ρόλο παίζει και η μείωση του κόστους του υλικού των υπολογιστών, ώστε αυτή η ισχύς να μπορεί να αποκτηθεί εύκολα από κάποιον ερευνητή τεχνητών νευρωνικών δικτύων, που δεν έχει πρόσβαση σε ισχυρές υπολογιστικές μονάδες ή χρηματοδότηση για την έρευνα του. Το κόστος απόκτησης του πιο γρήγορου chip Intel Core i7 990X είναι στα 1000€, αλλά πηγαίνοντας λίγο πιο κάτω σε απόδοση στον Intel Core i7 - 2600 το κόστος γίνεται περίπου 300€ έχοντας 90% της πιο γρήγορης ταχύτητας που υπάρχει [42]. Για την ιστορία, το κόστος του επεξεργαστή με τον οποίο έγινε η

³³ **PGA = Pin Grid Array** είναι ένας τύπος συσκευασίας ολοκληρωμένων κυκλωμάτων που χρησιμοποιείται για εύκολη συναρμολόγηση ενός προσωπικού υπολογιστή.

³⁴ Ο Intel Pentium 4 στα 3.4GHz κυκλοφόρησε το 2008 και είναι ο τελευταίος της σειράς των επεξεργαστών.

³⁵ Με βάση τον δείκτη ταχύτητας επεξεργαστή Passmark, και τα αποτελέσματα των δοκιμών που υπάρχουν στο <http://www.cpubenchmark.net>

³⁶ **PPU = Physics Processing Unit**

ανάπτυξη του Horasis, του AMD Phenom II x4 965, είναι σήμερα 130€, έχει 40% της μέγιστης διαθέσιμης υπολογιστικής ισχύος σε PC, όμως το κόστος του είναι μόνο το 13% του κόστους του γρηγορότερου επεξεργαστή. Μπορούμε να πούμε ότι η απαιτούμενη ισχύς για την περαιτέρω μελέτη και εφαρμογή των τεχνητών νευρωνικών δικτύων υπάρχει, αλλά είναι και προσιτή για όλους. Έτσι αφαιρείται ένα ανυπέρβλητο εμπόδιο στην περαιτέρω ανάπτυξη της τεχνολογίας τους, που είναι το μεγάλο ερευνητικό κόστος.

Εύλογο ερώτημα για κάποιον αναγνώστη, βλέποντας μέσα σε μια εργασία της επιστήμης των υπολογιστών να παρατίθενται οικονομικά στοιχεία όπως τιμές και κόστος είναι: σε τι αποσκοπούν; Πλέον για οποιαδήποτε νέα τεχνολογία υπάρχει ένα ολόκληρο σύστημα οικονομικών παραγόντων που πρέπει να λαμβάνεται υπόψη και που επηρεάζουν την επιτυχία της. Αν μια νέα τεχνολογία μπορεί να αναπτυχθεί με λογικό κόστος έρευνας και ανάπτυξης (R&D) και να παραχθεί με ακόμη μικρότερο κόστος μέσα από μαζική παραγωγή, μπορεί να φτάσει πολύ πιο γρήγορα στον τελικό καταναλωτή. Έτσι αναφέρονται στην εργασία όπου κρίνεται σκόπιμο, πληροφορίες που δείχνουν πως μπορεί να μειωθεί το κόστος μιας πρότασης που κάνει ή που επιδεικνύουν το μικρό κόστος που απαιτείται για την εφαρμογή του Horasis.

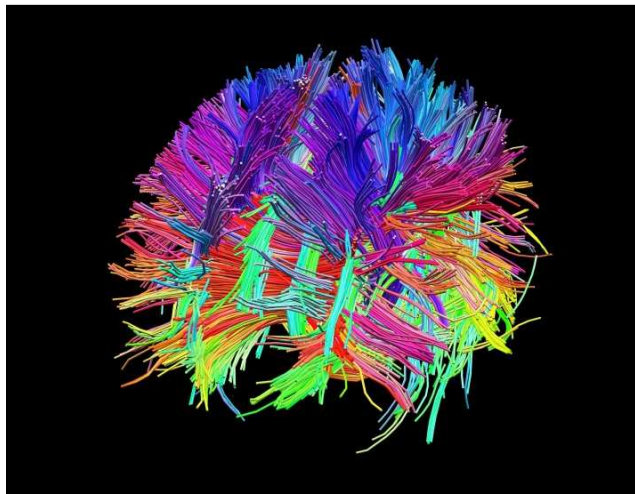
Είναι γεγονός ότι η τεχνολογία, όσο επαναστατική και αν είναι, δεν μπορεί να δημιουργήσει επανάσταση από μόνη της, παρά μόνο αν μπορεί να εφαρμοστεί σε μεγάλη κλίμακα λόγω του μειωμένου κόστους. Ένα παράδειγμα που το αποδεικνύει είναι η ίδια η ιστορία των ηλεκτρονικών υπολογιστών, όπου η εμπορική επιτυχία των πρώτων υπολογιστών σπιτιού (home computers) όπως ο Apple II [46], άνοιξε τον δρόμο για την πρόοδο της Επιστήμης των Υπολογιστών. Η προχωρημένη τεχνολογική σκέψη του Steve Wozniak [47] δεν θα μπορούσε παρά μόνο σε συνδυασμό με την προχωρημένη εμπορική σκέψη του Steve Jobs (1955 – 2011), να βρει την απήχηση και να ξεκινήσει την επανάσταση στον χώρο του υλικού των προσωπικών υπολογιστών. Ένα δεύτερο παράδειγμα είναι το Internet, που βιώνουμε τα τελευταία 15 χρόνια την εντυπωσιακή εξέλιξη της ιστορίας του. Από απόρρητο στρατιωτικό μυστικό της DARPA ³⁷ στα 1970, πέρασε σε μερικούς εκλεκτούς στα 1980, κατόπιν σε πολλούς γνώστες των υπολογιστών στην δεκαετία του 1990, και πλέον μετά τα μέσα της προηγούμενης δεκαετίας σε 30% του παγκοσμίου πληθυσμού σήμερα, έχοντας μια αύξηση 480% μόνο από το 2010 μέχρι το 2011 [48]. Η αλματώδης πρόοδος που συντελέστηκε οφείλεται σε ένα μεγάλο βαθμό και στην εκλαΐκευση και στην εμπορευματοποίηση του.

³⁷ Η **DARPA** - **Defense Advanced Research Projects Agency** είναι η υπηρεσία έρευνας του Αμερικανικού στρατού.

Νευροεπιστήμη, το μέλλον

Η νευροεπιστήμη είναι ίσως ένας από τους πιο δυναμικούς ερευνητικούς τομείς που υπάρχουν σήμερα και είναι η επιστημονική μελέτη του νευρικού συστήματος. Ξεκίνησε αρχικά ως τμήμα της βιολογίας αλλά πλέον συνεργάζεται κάτω από την σκέπη της γνωστικής επιστήμης, με ένα πλήθος άλλων επιστημονικών πεδίων όπως μαθηματικά, χημεία, φυσική, ιατρική, ψυχολογία, γλωσσολογία και επιστήμη των υπολογιστών με την τελευταία να αποτελεί την οπτική γωνία από την οποία την βλέπουμε σε αυτήν την εργασία. Οι ανακαλύψεις που μπορεί να δώσει στο μέλλον, αφού μοντελοποιηθούν με βιολογικούς όρους και με μαθηματικά, είναι αυτές που θα δημιουργήσουν τα τεχνητά νευρωνικά δίκτυα του μέλλοντος.

Ανάμεσα σε πληθώρα ερευνητικών προσπαθειών που θα έπρεπε να αναφερθούν, ίσως το πιο φιλόδοξο ερευνητικό πρόγραμμα που υπάρχει σήμερα είναι το Human Connectome Project δηλαδή η χαρτογράφηση του ανθρώπινου **συνδεσιώματος**³⁸. Συγκρίνεται σε μέγεθος και σημαντικότητα με την χαρτογράφηση του ανθρώπινου γονιδιώματος (Human Genome Project), που άρχισε το 1990 και ολοκληρώθηκε το 2003 και στην οποία ρόλο κλειδί έπαιξε ένας Έλληνας επιστήμονας ο Αριστείδης Πατρινός [49]. Το νέο έργο έχει σαν στόχο χρησιμοποιώντας την αιχμή της τεχνολογίας για την απεικόνιση του εγκεφάλου, να χαρτογραφήσει τα κυκλώματα ενός υγιούς ενήλικου εγκεφάλου. Συγκεντρώνοντας συστηματικά εγκεφαλικά δεδομένα από εκατοντάδες συμμετέχοντες θα προσπαθήσει να φωτίσει πως οι συνδέσεις του εγκεφάλου σχετίζονται με τις εγκεφαλικές λειτουργίες [50].

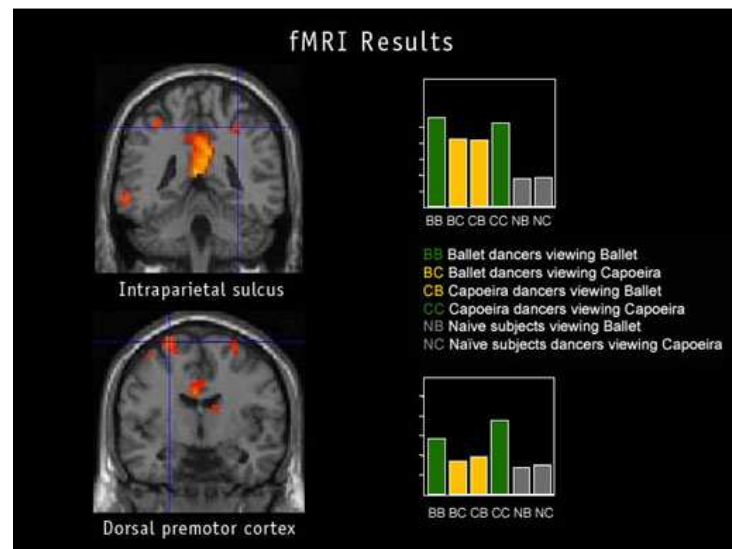


2-12: Χάρτης λευκής ουσίας εγκεφάλου μέσω DSI (<http://www.humanconnectomeproject.org/gallery/>)
 Οπτικοποίηση: Vaughan Green, Λογισμικό TrackVis, Δεδομένα: Randy Buckner Ph.D

Με αυτά τα δεδομένα, θα γίνει δυνατή η κατασκευή στοχευμένων φαρμάκων για την καταπολέμηση νευρολογικών διαταραχών και η εύρεση της νευρολογικής βάσης για τις παθήσεις που εκδηλώνουν χρόνιο πόνο. Γι αυτό παρά την παγκόσμια οικονομική συγκυρία που υπάρχει κατά την έναρξη του, το πρόγραμμα έχει λάβει αρχική χρηματοδότηση 6 εκατομμύρια δολάρια/χρόνο για 5 χρόνια. Από αυτόν τον θησαυρό δεδομένων που θα προκύψει στο μέλλον, μπορούμε εύκολα να φανταστούμε την τεχνητή νοημοσύνη να υλοποιεί τεχνητά νευρωνικά δίκτυα βασισμένα στο ανθρώπινο συνδεσίωμα. Η διεπιστημονική αυτή ανταλλαγή είναι αμφίδρομη και ήδη παίζει σημαντικό ρόλο καθώς η επιστήμη των υπολογιστών χρησιμοποιείται σε όλες τις νέες ερευνητικές τεχνικές και στην διαχείριση των συνόλων δεδομένων (data sets)

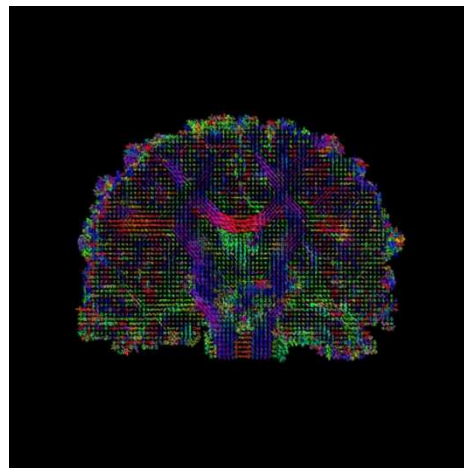
³⁸ Προτείνεται ως ελληνική μετάφραση «συνδεσίωμα» για τον αγγλικό όρο connectome, κρατώντας κοινή ρίζα με το ουσιαστικό **σύνδεσι-ς** και προσθέτοντας την ίδια κατάληξη με το ουσιαστικό γονιδ-**ίωμα**

που προκύπτουν. Ανάμεσα στις νέες μη παρεμβατικές τεχνικές απεικόνισης που χρησιμοποιούνται είναι η Λειτουργική Μαγνητική Τομογραφία fMRI, (Functional MRI) που χρησιμοποιεί τις αλλαγές στην ροή του αίματος και την κατανάλωση οξυγόνου σαν δείκτες για την νευρωνική δραστηριότητα, χρωματίζοντας τις περιοχές του εγκεφάλου που ενεργοποιούνται ανάλογα με την κάθε διαφορετική συμπεριφορά. Στην παρακάτω εικόνα με την μέθοδο fMRI, φαίνεται πως ενεργοποιούνται τα σημεία των εγκεφάλων και τα στατιστικά, σε τρεις ομάδες χορευτών, αθλητών capoeira³⁹ και τρίτων που παρακολουθούν όλες από δύο ερεθίσματα χορογραφίας και επίδειξης capoeira.



2-13: Αποτελέσματα fMRI (<http://www.pbs.org/wgbh/nova/>)

Δύο ακόμη ενδιαφέρουσες τεχνική είναι η HARDI (High Angular Resolution Diffusion Imaging) η οποία ανιχνεύει την διάχυση του νερού πάνω σε ινώδη ιστό, και η DSI (Diffusion Spectrum Imaging) με τις οποίες μπορούν να οπτικοποιηθούν δέσμες από νευροάξονες, τις απολήξεις των νευρώνων [51].



2-14: Στεφανιαία⁴⁰ όψη με HARDI (<http://www.humanconnectomeproject.org/gallery/>)
Οπτικοποίηση David Shattuck, Ph.D

³⁹ Capoeira είναι μια πολεμική τέχνη της Βραζιλίας που τα λακτίσματα της μοιάζουν με χορευτικές κινήσεις

⁴⁰ Στεφανιαία ραφή είναι ο αντίστοιχος ελληνικός όρος για το coronal suture.

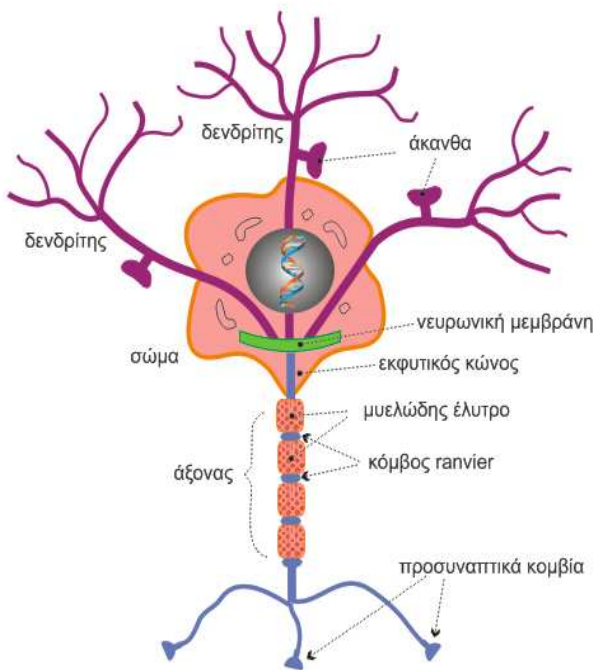
3. Τα Τεχνητά Νευρωνικά Δίκτυα

Ο Βιολογικός Νευρώνας

Όλα τα ερεθίσματα του πραγματικού κόσμου που λαμβάνει ο άνθρωπος μέσω των αισθήσεων του όπως μια εικόνα, ένας ήχος, ή μια μυρωδιά, μεταφέρονται από τα αισθητήρια όργανα στον εγκέφαλο όπου καταλήγουν σε νευρικά κύτταρα ή αλλιώς νευρώνες. Οι νευρώνες συνδέονται μεταξύ τους δημιουργώντας τα βιολογικά νευρωνικά δίκτυα, στο νευρικό σύστημα και τον εγκέφαλο μας. Το πλήθος αυτών των νευρώνων στον εγκέφαλο είναι εκπληκτικά μεγάλο, περίπου 100 δισεκατομμύρια και μπορεί κάποιος εύκολα να θεωρήσει ότι ο αριθμός των συνδέσεων που υπάρχουν μεταξύ των νευρώνων είναι αστρονομικός.

Περιγραφή του βιολογικού νευρώνα και της σύναψης

Ένας συνήθης νευρώνας αποτελείται από τέσσερις μορφολογικά καθορισμένες περιοχές [52]:



- τους δενδρίτες
- το κυτταρικό σώμα ή σώμα
- τον νευροάξονα ή άξονα
- τα προσυναπτικά κομβία
- τις άκανθες

Το κυτταρικό σώμα αποτελεί το κέντρο της λειτουργίας του κυττάρου από το οποίο ξεκινάει ένα πλήθος από ίνες που λέγονται δενδρίτες. Οι δενδρίτες στο τέλος τους διακλαδίζονται σε λεπτότερες αποφυάδες που η κάθε μία έχει δύο βασικά μέρη, το κύριο στέλεχος και την άκανθα. Οι άκανθες είναι εξειδικευμένα σημεία με λεπτό αυχένα και διογκωμένη κεφαλή.

Επίσης από το σώμα εκφύεται μια σωληνοειδής κατασκευή, ο άξονας ο οποίος ξεκινάει από μια

3-1: Σχηματική αναπαράσταση βιολογικού νευρώνα
©2011 Horasis

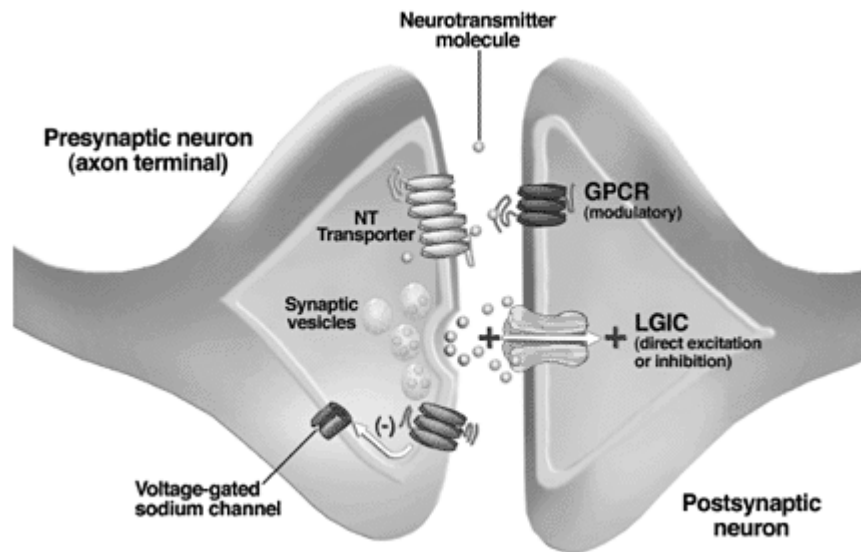
εξειδικευμένη περιοχή του σώματος τον εκφυτικό κώνο. Η λειτουργία του απλοποιημένα είναι η λήψη ενός ηλεκτρικού σήματος μέσω των δενδριτών στο σώμα και η εκπομπή υπό συνθήκες ενός άλλου ηλεκτρικού σήματος από το σώμα προς τον άξονα. Αυτό το σήμα κατόπιν μεταφέρεται προς τους άλλους νευρώνες που είναι συνδεδεμένοι με αυτόν .

Για να διασφαλιστεί η ταχεία αγωγή του ενεργειακού δυναμικού μέσα από έναν μεγάλο άξονα υπάρχει ένα είδος μόνωσης γύρω από την ίνα του άξονα, το μυελώδη έλυτρο. Αυτό αποτελείται από ένα πολύ σημαντικό βιολογικό υλικό του ανθρώπινου σώματος, την μυελίνη⁴¹. Ανά τακτά

⁴¹ Η έλλειψη της μυελίνης είναι η αιτία για τις εκφυλιστικές ασθένειες του νευρικού συστήματος, όπως η σκλήρυνση κατά πλάκας.

διαστήματα το στρώμα μυελίνης διακόπτεται σε σημεία που ονομάζονται κόμβοι Ranvier, στους οποίους αναγεννιάται το δυναμικό ενέργειας. Δηλαδή επανασηματίζεται ο νευρωνικός παλμός που μπορεί να έχει απώλειες κατά την μετάδοση του μέσω του άξονα. Θα λέγαμε ότι ο άξονας μοιάζει με ένα καλώδιο USB ⁴² που εκτός του μονωτικού περιβλήματος που έχει, αν πρέπει να συνδέσει μια συσκευή σε μεγάλη απόσταση πρέπει να τοποθετηθεί ανά 5 μέτρα μια συσκευή επαναληπτή ⁴³(repeater).

Κοντά στην απόληξη του ο άξονας διαιρείται σε λεπτούς κλάδους που συνδέονται με άλλους νευρώνες σε σημεία που ονομάζονται συνάψεις. Τα προσυναπτικά κομβία είναι διογκώσεις που βρίσκονται πάνω στους λεπτούς κλάδους του άξονα και αποτελούν την πλευρά πριν την σύναψη. Στην απέναντι πλευρά της σύναψης, δηλαδή μετασυναπτικά, συνήθως βρίσκονται οι δενδρίτες ενός άλλου νευρώνα. Στην σύναψη δεν υπάρχει ανατομική επαφή αλλά τα δύο κύτταρα χωρίζονται μέσω ενός κενού διαστήματος της συναπτικής σχισμής. Η διαβίβαση του σήματος μέσω αυτού του κενού γίνεται με την έκλυση από το προσυναπτικό κομβίο ουσιών που λέγονται νευροδιαβιβαστές (neurotransmitters), οι οποίοι φθάνουν στην απέναντι «όχθη» της σύναψης και προσδέονται σε ειδικούς υποδοχείς του αποδέκτη. Έτσι βλέπουμε ότι στην ουσία η σύναψη, χρησιμοποιώντας πάλι ένα απλό παράδειγμα από την ηλεκτρονική, είναι μια ασύρματη σύνδεση σαν το υπέρυθρο τηλεχειριστήριο της τηλεόρασης που πρέπει να βρίσκεται κοντά και να έχει κατεύθυνση προς τον δέκτη χωρίς να χρειάζεται να εφάπτεται.



3-2: Σχηματική αναπαράσταση σύναψης (<http://www.niaaa.nih.gov>)
Τα μικρά σφαιρίδια αντιπροσωπεύουν μόρια νευροδιαβιβαστών

⁴² **USB = Universal Serial Bus**

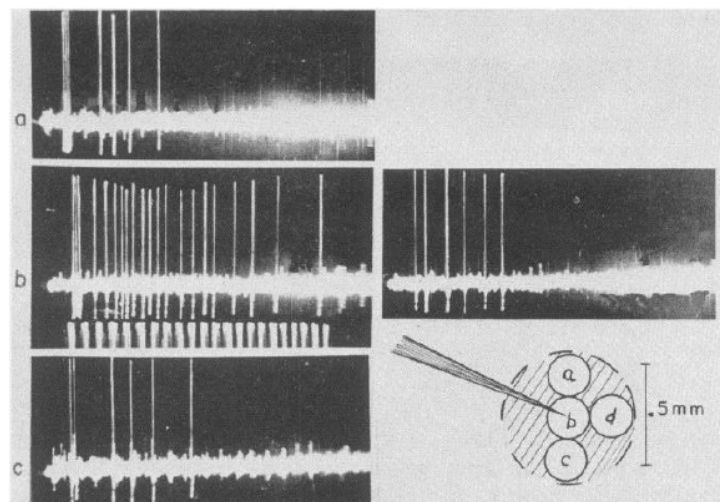
⁴³ Οι επαναληπτές είναι ηλεκτρονικές διατάξεις που αναγεννούν το σήμα όταν αυτό φθάνει με απώλεια σε μεγάλες αποστάσεις.

Η εσωτερική λειτουργία του νευρώνα

Με την διαβίβαση του σήματος από τον άξονα προς όλους τους αποδέκτες με τους οποίους έχει δημιουργήσει συνάψεις, ο νευρώνας διασπείρει την πληροφορία που έχει διαμορφώσει. Η πληροφορία αυτή μπορεί να φθάσει διαφοροποιημένη σε κάθε αποδέκτη. Η κάθε σύναψη μπορεί να τροποποιήσει το αρχικό ηλεκτρικό σήμα που πρόκειται να μεταφερθεί μέσω αυτής, προσδίδοντας του μικρή ή μεγάλη βαρύτητα, η οποία μπορεί να εκφραστεί ως ποσότητα και λέγεται συναπτικό βάρος.

Όταν σε έναν νευρώνα έχουν ληφθεί όλα τα τροποποιημένα από την εφαρμογή των συναπτικών βαρών ηλεκτρικά σήματα, συναθροίζονται στο εσωτερικό του σώματος όπου υπάρχει η νευρωνική μεμβράνη. Η διεργασία αυτή ονομάζεται χωρική άθροιση. Η μεμβράνη βρίσκεται σε κατάσταση ηρεμίας μέχρις ότου εκπληρωθούν κάποιες συνθήκες που σχετίζονται με την λειτουργία του νευρώνα, οπότε και ενεργοποιείται. Αυτή η επεξεργασία και η διαδικασία απόφασης για την ενεργοποίηση ή όχι του νευρώνα ονομάζεται νευρωνική ολοκλήρωση [52]. Έτσι σε μια δεδομένη χρονική στιγμή ο νευρώνας είτε ηρεμεί είτε εκπολώνεται, δηλαδή παράγει ένα νέο ηλεκτρικό σήμα το οποίο μεταφέρεται μέσω του άξονα και του μηχανισμού των συνάψεων, προς στους επόμενους νευρώνες.

Τα σήματα που ανταλλάσσονται μεταξύ των νευρώνων είναι δύο ειδών. Διεγερτικά, δηλαδή που μπορούν να βοηθήσουν στην διέγερση της κυτταρικής μεμβράνης για την παραγωγή του νέου σήματος ή ανασταλτικά, όπως ανακαλύφθηκε από τον John Eccles. Τα ανασταλτικά εμποδίζουν την διέγερση και την παραγωγή του ηλεκτρισμού στην κυτταρική μεμβράνη. Θα μπορούσαμε να παρομοιάσουμε τα δύο είδη των σημάτων όπως τα πεντάλ του γκαζιού και του φρένου για μια μηχανή αυτοκινήτου που είναι η κυτταρική μεμβράνη. Όπως η απόκριση της μηχανής για την κίνηση των τροχών είναι ενιαία, έτσι και η απόκριση του νευρώνα ενοποιεί και τις δύο αντικρουόμενες τάσεις, καθώς τα ηλεκτρικά σήματα που παράγουν οι μεμβράνες δεν είναι σταθερής έντασης, αλλά σύνθετες κυματομορφές [53].

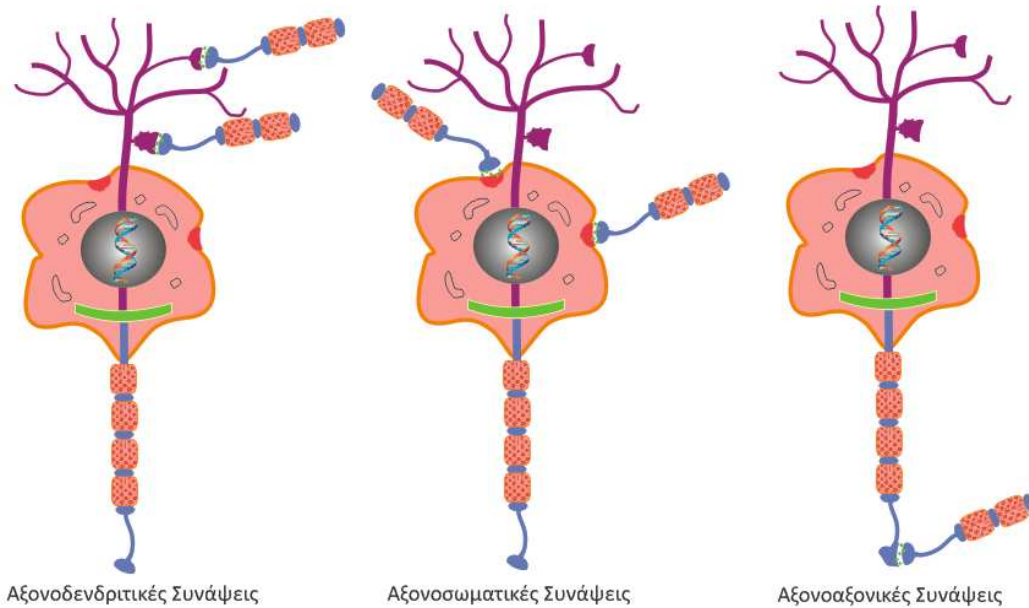


3-3: Ηλεκτρικά σήματα που παράγουν νευρώνες του ματιού (Kuffler) [53]

Διαφορετικοί τύποι συνάψεων

Στους νευρώνες του κεντρικού νευρικού συστήματος έχουν παρατηρηθεί τριών διαφορετικών τύπων συνάψεις, που μπορούν να έχουν διεγερτικό ή ανασταλτικό χαρακτήρα [52].

- Αξονοδενδρικές
- Αξονοσωματικές
- Αξονοαξονικές



3-4: Οι διαφορετικοί τύποι συνάψεων.

©2011 Horasis

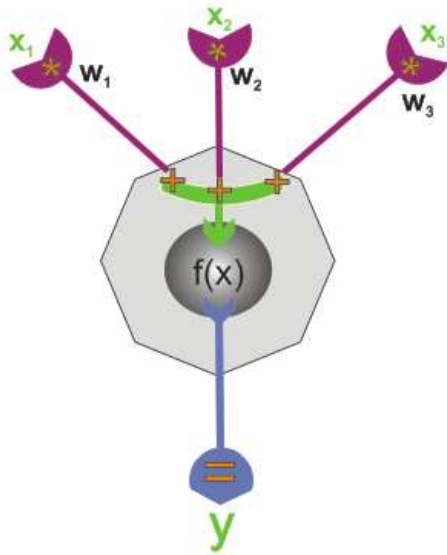
Ο πιο συνηθισμένος και γνωστός τύπος σύναψης είναι η αξονοδενδρική σύναψη όπου ένας άξονας μέσω των κλάδων του συνδέεται και αποστέλλει το ηλεκτρικό σήμα σε έναν δενδρίτη. Η σύναψη μπορεί να γίνει στο κύριο στέλεχος του δενδρίτη ή στην άκανθα του, όπου στην δεύτερη περίπτωση συνήθως έχει διεγερτικό χαρακτήρα.

Όταν ένας άξονας συνδέεται απευθείας με το σώμα ενός άλλου νευρώνα, δηλαδή πάνω στην κυτταρική μεμβράνη του, τότε έχουμε μια αξονοσωματική σύναψη. Αυτή η σύναψη επενεργεί άμεσα πάνω στον τρόπο λειτουργίας της μεμβράνης και τον μηχανισμό εκκίνησης του ηλεκτρικού σήματος καθώς μπορεί να τροποποιήσει εσωτερικά την αγωγιμότητα, μειώνοντας έτσι την εκπόλωση της. Θα λέγαμε με όρους της ηλεκτρονικής, ότι μια τέτοια σύναψη είναι μια μεταβλητή αντίσταση η οποία παρεμβάλλεται μετά την κυτταρική μεμβράνη και πριν τον άξονα.

Οι αξονοαξονικές συνάψεις δεν επιδρούν στον μηχανισμό εκκίνησης του ηλεκτρικού σήματος, αλλά επηρεάζουν την μετάδοση του στους μετασυναπτικούς νευρώνες, ελέγχοντας την ποσότητα του νευροδιαβιβαστή που ελευθερώνεται. Σε μια τέτοια σύναψη ο ένας άξονας παίζει τον ρόλο του ρυθμιστή στον άλλο άξονα. Θα μπορούσαμε να πούμε ότι ο ρυθμιστής λειτουργεί ως ένα κοινό συναπτικό βάρος για όλες τις συνάψεις που δημιουργεί ο ρυθμιζόμενος άξονας.

Ο Τεχνητός Νευρώνας

Στηριζόμενη στη βιολογική πραγματικότητα των νευρώνων, οι McCulloch και Pits, δημιούργησαν την δεκαετία του 1940 ένα απλό μαθηματικό μοντέλο[13], που περιγράφει και υλοποιεί την βασική λειτουργία ενός νευρώνα. Η ενεργοποίηση του νευρώνα αντιπροσωπεύεται από την έξοδο μιας συνάρτησης $y = f(x)$ όπου με $y = 1$ αναπαριστάται η ενεργοποίηση του νευρώνα και με $y = 0$ η κατάσταση ηρεμίας. Στην ουσία η συνάρτηση f προσομοιώνει με



3-5: Σχηματική αναπαράσταση τεχνητού νευρώνα
©2011 Horasis

μαθηματική διατύπωση την νευρωνική ολοκλήρωση του βιολογικού νευρώνα και ονομάζεται συνάρτηση ενεργοποίησης του νευρώνα [54]. Τα δε σήματα που διαβιβάζονται σε ένα πλήθος n συνάψεων περιγράφονται ως μεταβλητές $x_1 \dots x_n$ των οποίων οι τιμές είναι πραγματικοί αριθμοί. Τα διεγερτικά σήματα αναπαριστώνται με θετικούς αριθμούς και τα ανασταλτικά σήματα με αρνητικούς. Τα αντίστοιχα συναπτικά βάρη είναι οι πραγματικοί αριθμοί $w_1 \dots w_n$ οπότε το τροποποιημένο σήμα στην κάθε σύναψη είναι το αποτέλεσμα του γινομένου $w_i x_i$. Έτσι συνάθροιση των τιμών των συνάψεων στην μεμβράνη του νευρώνα μπορεί να παρασταθεί:

$$s = \sum_{i=1}^n w_i x_i \quad (3.1)$$

Η ποσότητα s που αντιπροσωπεύει την συγκέντρωση δυναμικού στην μεμβράνη του νευρώνα, θα ονομάζεται εφεξής **συναπτικό άθροισμα (synaptic summation)** [55] με $s \in \mathbf{R}$. Η πιο βασική ομάδα συναρτήσεων που χρησιμοποιείται για την ενεργοποίηση των νευρώνων είναι οι συναρτήσεις κατωφλίου (threshold functions) οι οποίες έχουν ως σύνολο τιμών το ζεύγος $\{0,1\} \in \mathbf{N}$ [56]. Σε αυτές αν το συναπτικό άθροισμα ξεπεράσει ένα όριο, τότε η συνάρτηση θα επιστρέφει την μονάδα ενεργοποιώντας τον τεχνητό νευρώνα, ενώ κάτω του ορίου θα παραμένει ανενεργός επιστρέφοντας την τιμή μηδέν. Η πιο απλή μορφή μπορεί να γραφεί ως εξής:

$$y = f(s) = \begin{cases} 0, & \text{αν το } s \leq \theta \\ 1, & \text{αν το } s > \theta \end{cases} \quad (3.2)$$

Ο αριθμός θ λέγεται κατώφλι (threshold) και αντιστοιχεί στο κατώφλι που έχει η βιολογική νευρωνική μεμβράνη, που όταν ξεπεραστεί προκαλεί την εκπόλωση του νευρώνα.

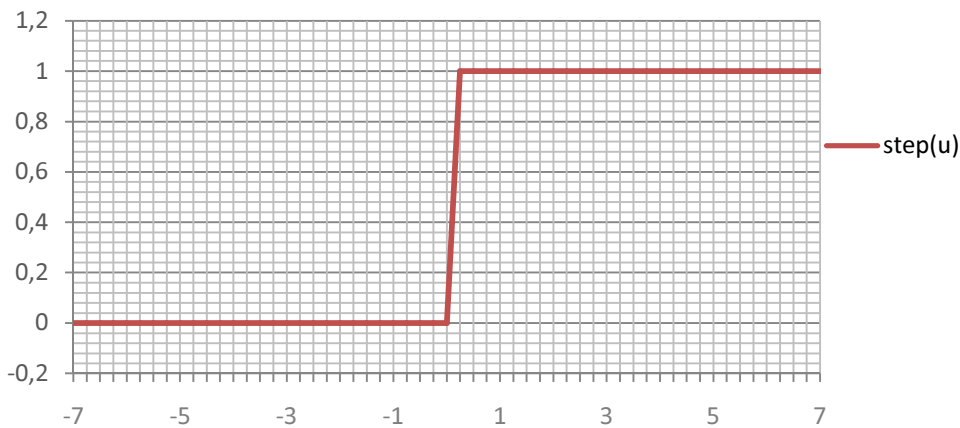
Αν θεωρήσουμε και το κατώφλι ως ακόμα ένα συναπτικό βάρος w_0 , σε μια εικονική σύναψη με σταθερή τιμή εισόδου $x_0 = -1$, παίρνουμε:

$$u = s - \theta = \sum_{i=1}^n w_i x_i - \theta = \sum_{i=0}^n w_i x_i, \text{ με } x_0 = -1 \text{ και } w_0 = \theta$$

Έχοντας απλούστερα την συνάρτηση:

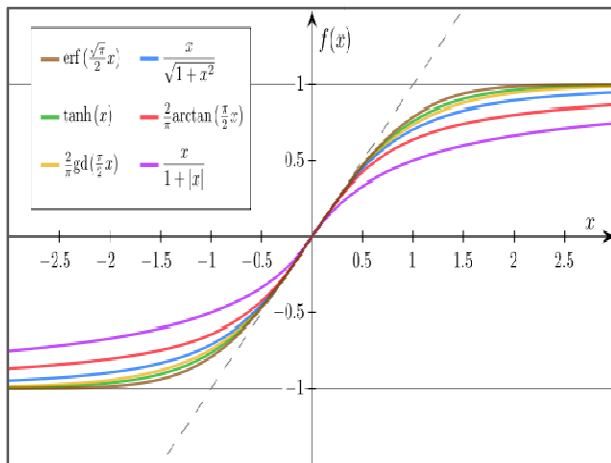
$$y = f(u) = \begin{cases} 0, & \text{αν το } u \leq 0 \\ 1, & \text{αν το } u > 0 \end{cases} \quad (3.3)$$

Η παραπάνω συνάρτηση ενεργοποίησης ονομάζεται και βηματική συνάρτηση 0/1 (step function) [54].



3-6: Γραφική παράσταση της βηματικής συνάρτησης

Οι συναρτήσεις ενεργοποίησης σιγμοειδούς μορφής



3-7: Διάφορες συναρτήσεις σιγμοειδούς μορφής με σύνολο τιμών το [-1,1] (Georg-Johann/Wikimedia Commons)

Η βηματική συνάρτηση από την φύση της δεν έχει ενδιάμεσες καταστάσεις εξόδου εκτός από τα 0 και 1, λειτουργώντας στην ουσία σαν ένας ψηφιακός διακόπτης. Για την ύπαρξη και κάποιας ενδιάμεσης κατάστασης μεταξύ ηρεμίας και ενεργοποίησης, ενσωματώνοντας στην ουσία ασάφεια⁴⁴ (fuzziness) στην απόκριση, χρησιμοποιούνται στα τεχνητά νευρωνικά δίκτυα οι συναρτήσεις σιγμοειδούς μορφής. Το κοινό τους γνώρισμα είναι ότι η καμπύλη τους έχει την μορφή “S”, από όπου κατάγουν και την ονομασία τους, και επίσης ότι περικλείεται μεταξύ δύο οριζοντίων ασύμπτωτων γραμμών.

⁴⁴Στην ασαφή λογική έχουμε τιμές μεταξύ του απολύτους αληθούς (1) και του απολύτους ψευδούς (0)

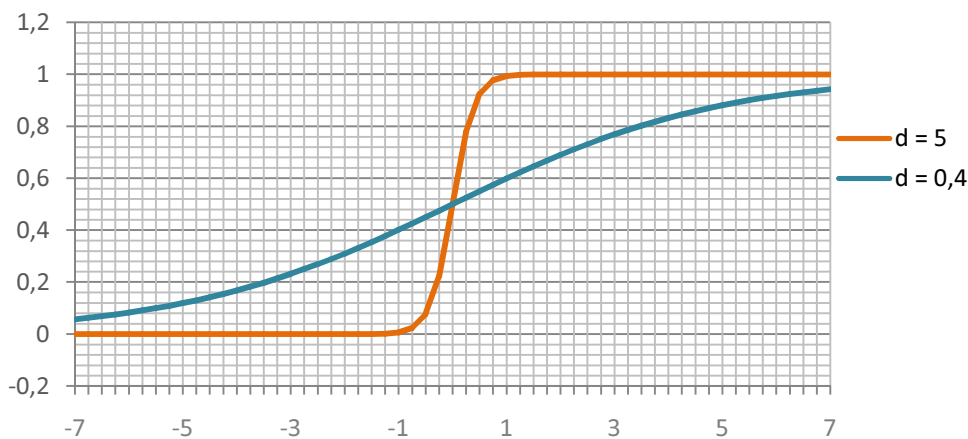
Η λογιστική συνάρτηση (logistic function) επιλέχθηκε από αυτές και χρησιμοποιήθηκε ως βάση για την συνάρτηση ενεργοποίησης, που είναι κοινή για όλους τους τεχνητούς νευρώνες του Horasis. Μια γενική της μορφή είναι:

$$f(u) = a_L + \left(\frac{\alpha_H - \alpha_L}{1 + b e^{-d(u-c)}} \right) \quad (3.4)$$

Το α_L και το α_H είναι οι δύο ασύμπτωτες της καμπύλης, κατώτερη και ανώτερη αντίστοιχα[57]. Για απλούστευση της μελέτης της συνάρτησης δεν συμπεριλαμβάνουμε τους παράγοντες b, c και τους κρατούμε σταθερούς με τιμές $b = 1, c = 0$, οπότε πλέον έχουμε:

$$f(u) = a_L + \left(\frac{\alpha_H - \alpha_L}{1 + e^{-du}} \right) \quad (3.5)$$

Ο παράγοντας d ρυθμίζει την κρημνότητα (steepness) της συνάρτησης και στην ουσία είναι ο συντελεστής ρύθμισης της ταχύτητας μετάβασης μεταξύ των δύο ασύμπτωτων γραμμών α_L και α_H . [58]. Πρακτικά μπορούμε να ονομάσουμε τον d στοχαστικό παράγοντα, καθώς καθορίζει την φύση της στοχαστικής συμπεριφοράς του νευρώνα[57], ενώ στα στοχαστικά νευρωνικά δίκτυα λέγεται και παράμετρος θερμοκρασίας (temperature parameter).[59]



3-8: Γραφική παράσταση της λογιστικής συνάρτησης με $\alpha_L = 0, \alpha_H = 1, b = 1, c = 0$

Μελετώντας την συνάρτηση, παρατηρούμε ότι, όσο πιο μεγάλος γίνεται ο στοχαστικός παράγοντας άνω της μονάδας, η καμπύλη τείνει να γίνει ίδια με την βηματική συνάρτηση, έχοντας πιο «απότομη» ή «γρηγορότερη» μετάβαση από το 0 στο 1. Με μια τέτοια ρύθμιση στην συνάρτηση ενεργοποίησης έχουμε στην ουσία έναν δυαδικό στοχαστικό νευρώνα[57] [60]. Από την άλλη αν η τιμή του d είναι μικρότερη του 1, «επιμηκώνεται» το εύρος τιμών του u για το οποίο η συνάρτηση θα δώσει «ασαφή» απάντηση μεταξύ του 0 και του 1. Αντίστοιχα θα μπορούσαμε να πούμε ότι η μετάβαση από το 0 προς το 1 γίνεται «ομαλότερα» ή «αργότερα».

Θέτουμε τις παραμέτρους $\alpha_L = 0$ και $\alpha_H = 1$ στην σχέση 3.5 ώστε να έχουμε αντίστοιχα τις επιθυμητές τιμές ηρεμίας και ενεργοποίησης του τεχνητού νευρώνα μας. Παράλληλα επιλέγουμε την τιμή του στοχαστικού παράγοντα στο $d = 1$ έχοντας τελικά την απλοποιημένη μορφή της λογιστικής συνάρτησης, που συνήθως αναφερόμαστε σε αυτήν ως σιγμοειδής συνάρτηση και περιγράφεται από τον μαθηματικό τύπο:

$$f(u) = \left(\frac{1}{1+e^{-u}} \right), \quad (3.6)$$

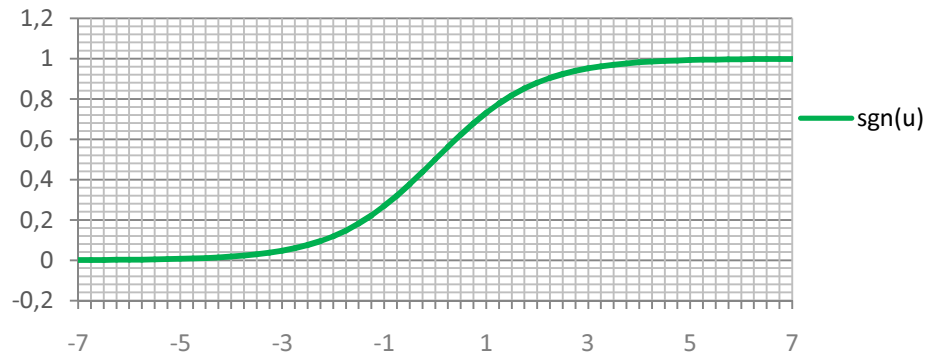
με πεδίο ορισμού $u \in \mathbb{R}$, σύνολο τιμών το $(0,1) \in \mathbb{R}$ και

$f(u) \rightarrow 0$, όταν $u \rightarrow -\infty$

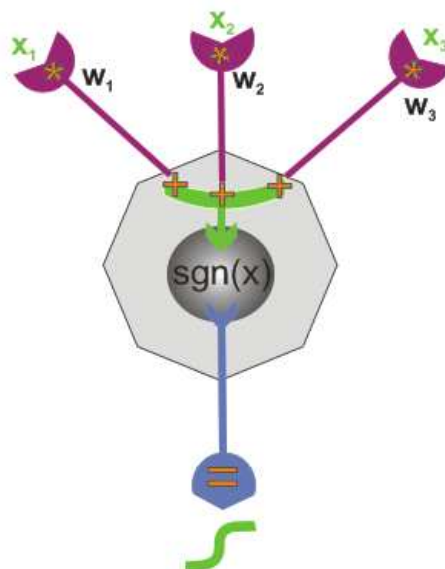
$f(u) = 0.5$, όταν $u = 0$, το οποίο είναι και σημείο συμμετρίας

$f(u) \rightarrow 1$, όταν $u \rightarrow +\infty$

Με αυτόν τον τρόπο, όταν το συναπτικό άθροισμα είναι μεταξύ -4 έως περίπου 4 , ο τεχνητός νευρώνας δεν έχει σαφή απόκριση αληθείας 0 ή 1 , αλλά παράγει ενδιάμεσες δεκαδικές τιμές που περιέχουν ασάφεια. Στην περίπτωση που το άθροισμα ισούται ακριβώς 0 τότε ο νευρώνας ενεργοποιείται κατά το ήμισυ, ισορροπώντας μεταξύ των δύο αποφάσεων.



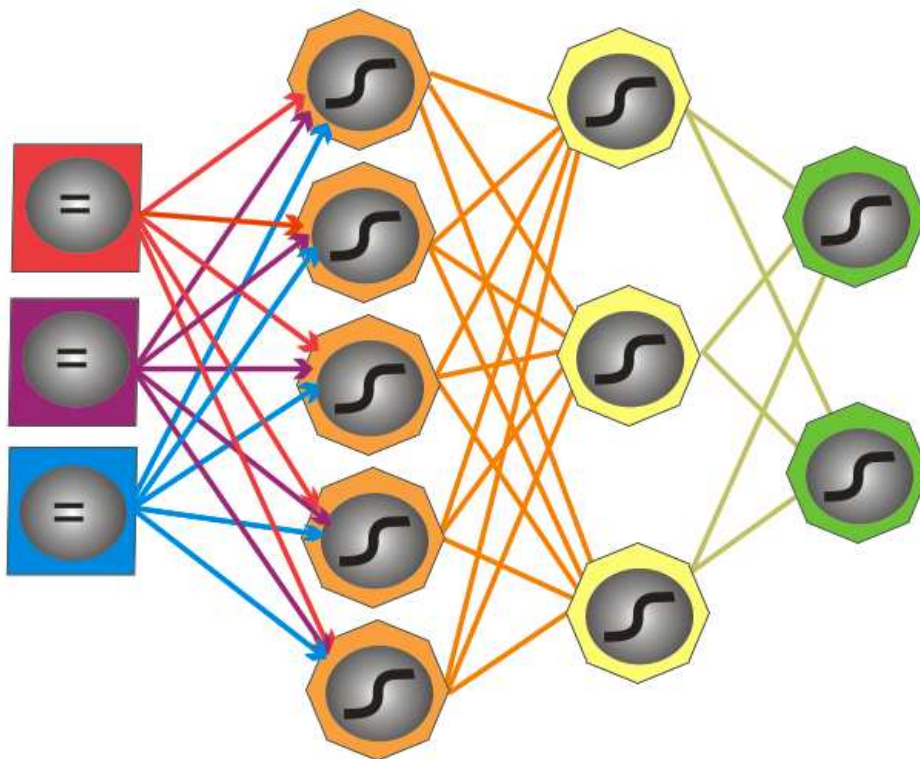
3-9: Γραφική παράσταση της σιγμοειδούς συνάρτησης (λογιστικής με $d = 1$)



3-10: Σχηματική αναπαράσταση τεχνητού νευρώνα με σιγμοειδή ενεργοποίηση

Το Τεχνητό Νευρωνικό Δίκτυο και η εκπαίδευσή του.

Το τεχνητό νευρωνικό δίκτυο στο οποίο επικεντρώνεται η εργασία είναι το Multi Layer Perceptron ή MLP. Το δίκτυο αυτό αποτελείται από ένα πλήθος από στρώματα (layers) νευρώνων που το κάθε ένα από αυτά μπορεί να έχει διαφορετικό πλήθος τεχνητών νευρώνων. Το κύριο χαρακτηριστικό ενός τέτοιου δικτύου είναι ότι ένα στρώμα με δείκτη i τροφοδοτείται από τους νευρώνες του προηγούμενου στρώματος $i - 1$ και μόνον από αυτούς [54]. Έτσι όλες οι εισερχόμενες συνάψεις σε έναν τεχνητό νευρώνα που βρίσκεται σε κάποιο στρώμα, προέρχονται από όλες τις εξόδους νευρώνων που ανήκουν στο αμέσως προηγούμενο στρώμα. Από την σκοπιά της σύναψης θα ορίσουμε ως νευρώνα προορισμού τον νευρώνα στον οποίο καταλήγει η σύναψη και νευρώνα προέλευσης αυτόν που ανήκει στο αμέσως προηγούμενο στρώμα και αποστέλλει την πληροφορία μέσω της σύναψης. Θα τοποθετήσουμε επίσης την σύναψη στην πλευρά του νευρώνα προορισμού, σαν να ανήκει σε αυτόν, για καλύτερη κατανόηση.

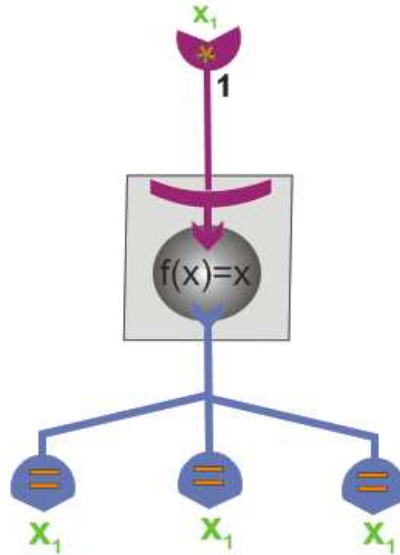


3-11: Ένα δίκτυο MLP του Horasis: Το στρώμα εισόδου με τρία διαφορετικά δεδομένα, χρωματισμένα με κόκκινο, μωβ και γαλάζιο. Το κρυφό στρώμα 1 με πορτοκαλί, το κρυφό στρώμα 2 με κίτρινο και το στρώμα εξόδου με πράσινο. ©2011 Horasis

Το πρώτο στρώμα ονομάζεται στρώμα εισόδου και σε αυτό τροφοδοτούνται τα δεδομένα, ενώ το τελευταίο λέγεται στρώμα εξόδου και σε αυτό παράγεται το αποτέλεσμα της νευρωνικής επεξεργασίας. Τα υπόλοιπα στρώματα που βρίσκονται στο ενδιάμεσο ονομάζονται κρυφά στρώματα. Το MLP δίκτυο που χρησιμοποιείται στο Horasis έχει 4 στρώματα, το στρώμα εισόδου, κρυφό στρώμα 1, κρυφό στρώμα 2 και το στρώμα εξόδου. Οι συνάψεις στο δίκτυο MLP του Horasis ακολουθούν το απλό γενικό κανόνα: Αν σε ένα στρώμα προέλευσης με δείκτη $i - 1$ έχουμε πλήθος νευρώνων n , τότε οποιοσδήποτε νευρώνας που ανήκει στο επόμενο στρώμα προορισμού i θα έχει n συνάψεις μια για τον κάθε νευρώνα που προηγείται. Η κάθε μια από τις αντίστοιχες μεταβλητές $x_0 \dots x_{n-1}$ θα παίρνει την τιμή της από την συνάρτηση ενεργοποίησης

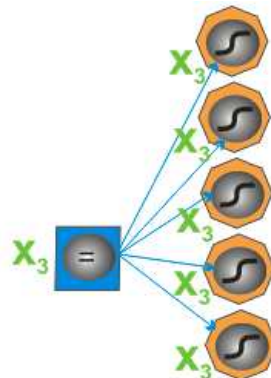
$f(u)$ του αντίστοιχου νευρώνα προέλευσης στις θέσεις 0 έως $n-1$ μέσα στο προηγούμενο στρώμα.

Το στρώμα εισόδου υλοποιεί ένα ιδιότυπο είδος τεχνητού νευρώνα που απλώς μεταφέρει την είσοδο στην έξοδο του, ώστε να διασπαρθεί σε κάθε νευρώνα του πρώτου από τα κρυφά στρώματα. Έχουμε $w_0 = 1$ άρα $u = x_0$ και τελικά την συνάρτηση ενεργοποίησης να είναι $y=f(u)=x_0$



3-12: Σχηματική αναπαράσταση τεχνητού νευρώνα εισόδου σε δίκτυο MLP ©2011 Horasis

Έχοντας διαφορετικά βάρη $w_0 \dots w_{n-1}$ για κάθε νευρώνα του δικτύου, βλέπουμε ότι μια τιμή εξόδου από έναν νευρώνα θα έχει διαφορετική επίδραση στον κάθε νευρώνα προορισμού του αμέσως επομένου στρώματος. Έτσι μια τιμή εισόδου x_i καταλήγει στα συναπτικά αθροίσματα των νευρώνων του κρυφού στρώματος 1 , διαφοροποιημένη αφού ζυγίζεται με διαφορετικό συναπτικό βάρος w_i σε κάθε νευρώνα προορισμού.



3-13: Διασπορά της τιμής εισόδου x_3 σε όλους τους νευρώνες του κρυφού στρώματος 1 ©2011 Horasis

Ανάκληση στο δίκτυο MLP του Horasis, με σιγμοειδή συνάρτηση ενεργοποίησης

Η ανάκληση (recall) είναι η διαδικασία υπολογισμού των τιμών εξόδων όλων των νευρώνων του δικτύου έχοντας δεδομένες τις τιμές των μεταβλητών $x_0 \dots x_{n-1}$ [54]. Έχουμε με όρους βιολογίας την μεταφορά των σημάτων και την νευρωνική ολοκλήρωση κάθε τεχνητού νευρώνα του δικτύου. Γίνονται διαβιβάσεις του σήματος που παράγει στην έξοδο του ένας νευρώνας, μέσω των συνάψεων, στους νευρώνες του επομένου στρώματος. Θα μπορούσαμε να πούμε με βάση το κλασικό μοντέλο είσοδος, επεξεργασία, έξοδος ότι η ανάκληση σε ένα νευρωνικό δίκτυο, είναι η παραγωγή της πληροφορίας στην έξοδο του δικτύου, αφού γίνει μια επεξεργασία υπολογισμού σε κάθε νευρώνα του, με αρχικά δεδομένα τις τιμές στους νευρώνες του στρώματος εισόδου.

Έτσι στα νευρωνικά δίκτυα MLP που χρησιμοποιούνται στο Horasis, η ανάκληση για τον κάθε τεχνητό νευρώνα στο στρώμα εισόδου, έχοντας ως αρχικά δεδομένα δεκαδικές τιμές $x_0 \dots x_{n-1}$ είναι απλά:

$$q_0 = f_{INP}(x_0) = x_0$$

Η ανάκληση για τον κάθε τεχνητό νευρώνα του αμέσως επόμενου κρυφού στρώματος 1, με σιγμοειδή ενεργοποίηση υπολογίζεται από τον μαθηματικό τύπο:

$$z_0 = f_{HID1}(q_0, \dots, q_{n-1}) = \left(\frac{1}{1 + e^{-\sum_0^{n-1} w_i q_i}} \right), \quad (3.7)$$

όπου n είναι το πλήθος των νευρώνων του στρώματος εισόδου

Η ανάκληση για νευρώνα του κρυφού στρώματος 2, είναι:

$$h_0 = f_{HID2}(z_0, \dots, z_{m-1}) = \left(\frac{1}{1 + e^{-\sum_0^{m-1} w_i z_i}} \right), \quad (3.8)$$

όπου m είναι το πλήθος των νευρώνων του κρυφού στρώματος 1

Η ανάκληση των νευρώνων του στρώματος εξόδου, που αποτελεί και την έξοδο του δικτύου υπολογίζεται με:

$$y_0 = f_{OUT}(h_0, \dots, h_{k-1}) = \left(\frac{1}{1 + e^{-\sum_0^{k-1} w_i h_i}} \right), \quad (3.9)$$

όπου k είναι το πλήθος των νευρώνων του κρυφού στρώματος 2

Η εκπαίδευση του δικτύου με επίβλεψη

Η εκπαίδευση ενός τεχνητού νευρωνικού δικτύου είναι μια διαδικασία υπολογισμού των συναπτικών βαρών του, ώστε να ικανοποιείται κάποιο κριτήριο καταλληλότητας [54]. Το κριτήριο καταλληλότητας μπορεί να είναι ένα σύνολο από επιθυμητές τιμές των νευρώνων στην έξοδο του, το οποίο ονομάζεται στόχος εκπαίδευσης. Δηλαδή με κάποιον αλγόριθμο πρέπει να ρυθμιστούν τα συναπτικά βάρη του δικτύου ώστε στην έξοδο του να έρχεται ως αποτέλεσμα ο στόχος εκπαίδευσης. Αυτό το είδος εκπαίδευσης λέγεται εκπαίδευση με επίβλεψη (supervised learning) καθώς κάποιος με την γνώση του υποδεικνύει την επιθυμητή τιμή εξόδου σε σχέση με τα δεδομένα εισόδου και περιμένει από το εκπαιδευμένο δίκτυο να την αναπαράγει.

Η εκπαίδευση ΤΝΔ με επίβλεψη είναι αντίστοιχη με την διαδικασία που ο ενήλικος άνθρωπος μαθαίνει αντικείμενα σε ένα νήπιο δίνοντας του ως στόχους τις ονομασίες τους. Θα πάρουμε ένα απλό παράδειγμα όπου ένας ενήλικας ξέρει πολλά αντικείμενα που είναι μπάλες από την δική του εκπαίδευση και εμπειρία και προσπαθεί να εκπαιδεύσει το παιδί του, κατά τα πρώτα έτη της ζωής του, ώστε να αναγνωρίζει τις μπάλες. Ως δεδομένο εισόδου του δείχνει, και του δίνει να κρατήσει μια μπάλα, που αποτελεί εκείνη την στιγμή πρότυπο για όλες τις μπάλες και παράλληλα δίνει τον στόχο εκπαίδευσης φωνάζοντας «μπάλα». Μετά από πολλές επαναλήψεις και ίσως ένα ικανοποιητικό πλήθος από διαφορετικά πρότυπα αντικείμενα-μπάλες το παιδί μπορεί πλέον να διαχωρίσει την μπάλα προσπαθώντας και αυτό να εκφέρει τις συλλαβές «μπά-λα».

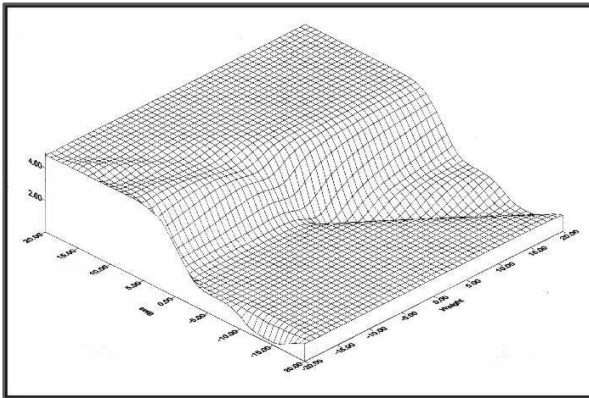
Οι ακριβείς διαδικασίες που συμβαίνουν στο παραπάνω «μαγικό» παράδειγμα από την πραγματική ζωή δεν μπορούν να περιγραφθούν με την υπάρχουσα επιστημονική γνώση. Αυτό που μπορεί να κάνει η επιστήμη σήμερα, είναι να αντιγράψει μια κοινή ανθρώπινη συμπεριφορά και να την μετατρέψει σε μαθηματικά. Η ανθρώπινη συμπεριφορά είναι η συνεχής βελτίωση των σφαλμάτων που κάνει ο άνθρωπος σε σχέση με τους στόχους εκπαίδευσης που του τίθενται. Αυτό μεταφέρεται ως ένας μαθηματικός τρόπος εκπαίδευσης του ΤΝΔ που λειτουργεί με την επαναληπτική διόρθωση των τιμών των συναπτικών βαρών, έχοντας απώτερο σκοπό την μείωση του σφάλματος που έχει η έξοδος του δικτύου σε σχέση με τους επιθυμητούς στόχους.

Έτσι μπορούμε να θεωρήσουμε ως ένα πρότυπο για την εκπαίδευση ενός τεχνητού νευρωνικού δικτύου ένα διάνυσμα $\mathbf{S} = [s_0, \dots, s_{n-1}]$ οι οποίες αντιπροσωπεύουν τα εισερχόμενα σήματα στην είσοδο του δικτύου και ένα διάνυσμα $\mathbf{T} = [t_0, \dots, t_{m-1}]$ και αντιστοιχούν στους επιθυμητούς στόχους που πρέπει παράγουν οι νευρώνες του στρώματος εξόδου. Επίσης μετά από κάποια ανάκληση, προκύπτει και ένα διάνυσμα $\mathbf{Y} = [y_0, \dots, y_{m-1}] = [f_{out_0}, \dots, f_{out_{m-1}}]$ στην έξοδο του δικτύου με τιμές που είναι συνάρτηση των συναπτικών βαρών του δικτύου \mathbf{w}_{ij} . Για την ποσοτικοποίηση του σφάλματος χρησιμοποιείται το μέσο τετραγωνικό σφάλμα (Mean Square Error) ή MSE, που είναι και αυτό σε τελική ανάλυση συνάρτηση των βαρών του δικτύου.

$$J(\mathbf{w}_{ij}) = \left(\frac{1}{m}\right) \sum_{i=1}^m |t_i - y_i|^2 \quad (3.10)$$

Η μαθηματική βάση του αλγόριθμου εκπαίδευσης Back-Propagation

Η ιστορία του αλγόριθμου ξεκινάει το 1969 όπου οι Arthur Bryson και Yu-Chi Ho τον είχαν περιγράψει ως μια μέθοδο βελτιστοποίησης δυναμικών συστημάτων[61][62]. Έτυχε όμως ευρείας αναγνώρισης στις επόμενες δεκαετίες με απαρχή την εργασία του Paul Werbos το 1974[63] για την ανάλυση μοντέλων οικονομικής και πολιτικής πρόβλεψης[54]. Αρκετά αργότερα το 1986 βρήκε τον δρόμο του στον συνδεδασισμό, όταν οι David Rumelhard, Geoffrey Hinton και Ronald Williams[64] [54] πρότειναν τον αλγόριθμο ως μια διαδικασία μάθησης για τεχνητούς νευρώνες δίνοντας του το όνομα back-propagation. Από τότε είναι ένας από τους πιο



3-14: Κατάβαση Δυναμικού

γνωστούς και ευρέως διαδεδομένους αλγόριθμους εκπαίδευσης τεχνητών νευρωνικών δικτύων.

Ο αλγόριθμος προσπαθεί να ελαχιστοποιήσει την συνάρτηση του σφάλματος της εξόδου του δικτύου σε σχέση με χώρο των βαρών $J(\mathbf{w}_{ij})$, χρησιμοποιώντας την μέθοδο της κατάβασης δυναμικού (gradient descent). Ο συνδυασμός των συναπτικών βαρών όλου του δικτύου με τον οποίο ελαχιστοποιείται η συνάρτηση σφάλματος θεωρείται ως η

επίλυση του προβλήματος της εκπαίδευσης του τεχνητού νευρωνικού δικτύου [59]. Σε κάθε βήμα υπολογίζεται η κλίση της συνάρτησης σφάλματος ως προς τα βάρη του δικτύου[54] και για αυτό πρέπει να διασφαλιστεί ότι η συνάρτηση αυτή είναι συνεχής και παραγωγίσιμη[59]. Προφανώς η βηματική συνάρτηση δεν εξυπηρετεί στο back propagation και αν' αυτής έχουν προταθεί διάφορα είδη συναρτήσεων ενεργοποίησης. Οι συναρτήσεις σιγμοειδούς μορφής που ήδη γνωρίσαμε είναι συνεχείς στο \mathbb{R} και διαφορίζονται. Θεωρώντας ότι σε κάθε τεχνητό νευρώνα χρησιμοποιείται μόνο ένα άθροισμα των τιμών των συνάψεων του και καθώς το δίκτυο υπολογίζει μόνο συνθέσεις συναρτήσεων, η συνάρτηση του σφάλματος μπορεί και αυτή να παραγωγιστεί.[59].

Στην συνέχεια θα διερευνήσουμε την λειτουργία του αλγορίθμου εκπαίδευσης back propagation με την απλή σιγμοειδή συνάρτηση $f(\mathbf{u}) = \left(\frac{1}{1+e^{-u}}\right)$. Η παράγωγος της που θα φανεί χρήσιμη παρακάτω είναι:

$$f'(\mathbf{u}) = \frac{e^{-u}}{1+(e^{-u})} = f(\mathbf{u})(1 - f(\mathbf{u})) \quad (3.11)$$

Θεωρούμε ένα βάρος σύναψης σε ένα στρώμα του δικτύου ως \mathbf{w}_{ij} με i τον δείκτη του νευρώνα προορισμού μέσα στο στρώμα και j τον δείκτη του νευρώνα προέλευσης της που ανήκει στο αμέσως προηγούμενο στρώμα. Για τα βάρη διαδοχικών στρωμάτων θα χρησιμοποιηθεί ο συμβολισμός $\mathbf{w}_{THIS_{ij}}$, για του προηγούμενου $\mathbf{w}_{PREV_{ij}}$ και για του επομένου $\mathbf{w}_{NEXT_{ij}}$. Σύμφωνα με την μέθοδο κατάβασης δυναμικού η μεταβολή του \mathbf{w}_{ij} ως προς τον χρόνο t γίνεται χρησιμοποιώντας την παράγωγο της $J(\mathbf{w}_{ij})$ ως προς το \mathbf{w}_{ij} δηλαδή:

$$\frac{\Delta \mathbf{w}_{ij}}{\Delta t} = \frac{\partial J(\mathbf{w}_{ij})}{\partial \mathbf{w}_{ij}} \leftrightarrow \Delta \mathbf{w}_{ij} = \Delta t \frac{\partial J(\mathbf{w}_{ij})}{\partial \mathbf{w}_{ij}} \quad (3.12)$$

Ο χρόνος εδώ θεωρείται ότι αποτελείται από διακριτές χρονικές στιγμές, τις οποίες τις ονομάζουμε εποχές εκπαίδευσης ή απλά εποχές. Με αυτόν τον τρόπο μπορούμε να γράψουμε προσεγγιστικά, για δύο διαδοχικές εποχές k και $k + 1$, την διαφορά σε ένα συγκεκριμένο βάρος του στρώματος εξόδου ως την εξής διακριτή εξίσωση:

$$\Delta w_{OUT_{ij}}(k) = w_{OUT_{ij}}(k + 1) - w_{OUT_{ij}}(k) = -\beta \left(\frac{\partial J(w_{OUT_{ij}}(k))}{\partial w_{OUT_{ij}}(k)} \right) \quad (3.13)$$

Ο συντελεστής β είναι ένας μικρός θετικός αριθμός που λέγεται βήμα εκπαίδευσης (learning constant) και αντιστοιχεί στο ελάχιστο βήμα χρόνου Δt . [54]. Η επιλογή μιας μικρής τιμής για αυτό δίνει καλύτερη προσέγγιση της κατάβασης δυναμικού, ενώ μια μεγαλύτερης γρηγορότερη σύγκλιση στην εύρεση του ελάχιστου σφάλματος. Η επιλογή του βήματος εκπαίδευσης δεν μπορεί να γίνει με κάποιο μαθηματικό υπολογισμό, αλλά αποτελεί αντικείμενο έρευνας δια της δοκιμής και απόρριψης, ανάλογα με το εκπαιδευόμενο δίκτυο. Θα επανέλθουμε στο βήμα εκπαίδευσης παρακάτω, εξετάζοντας την λειτουργία του πάνω στην τελική διατύπωση του μαθηματικού τύπου.

Η εκπαίδευση στοχεύει στην κατάλληλη τροποποίηση ενός συναπτικού βάρους την εποχή k κατά μια ποσότητα, ώστε να το νέο βάρος που προκύπτει στην εποχή $k + 1$, να μειώνει την τιμή της συνάρτησης σφάλματος J .

$$w_{ij}(k + 1) = w_{ij}(k) + \Delta w_{ij}(k) \quad (3.14)$$

Από την μαθηματική επίλυση της κατάβασης δυναμικού [54] προκύπτει ένας κανόνας για την τροποποίηση των βαρών του στρώματος εξόδου σε μια εποχή k , που ονομάζεται κανόνας δέλτα (delta rule) και περιγράφεται από τον τύπο :

$$\Delta w_{OUT_{ij}} = \beta (t_i - y_i) f'_{OUT}(u_i) x_j \quad (3.15)$$

- με β το βήμα εκπαίδευσης,
- t_i τον στόχο για τον νευρώνα i του στρώματος εξόδου
- $y_i = f_{OUT}(u_i)$ η έξοδος της συνάρτησης ενεργοποίησης του νευρώνα i
- u_i το συναπτικό άθροισμα για τον νευρώνα i , ήτοι άθροισμα των γινομένων των σημάτων επί των αντίστοιχων συναπτικών βαρών, για όλες τις συνάψεις
- f'_{OUT} η παράγωγος της συνάρτησης ενεργοποίησης του νευρώνα i
- x_j η εισερχόμενη τιμή σήματος στην σύναψη j του νευρώνα i

Θα θέσουμε ως $\delta_i = (t_i - y_i) f'(u_i)$ το τμήμα της εξίσωσης 3.15 που εμπλέκεται το σφάλμα της εξόδου, οπότε μπορούμε να γράψουμε πιο γενικά την εξίσωση τροποποίησης των βαρών ενός στρώματος ως:

$$\Delta w_{ij} = \beta \delta_i x_j \quad (3.16)$$

Ο παράγοντας του σφάλματος δ είδαμε ότι υπολογίζεται αρχικά μόνο για το στρώμα εξόδου, καθώς εξαρτάται από έναν στόχο που έχει προκαθοριστεί μόνο για τον νευρώνα εξόδου. Για τα προηγούμενα στρώματα η ποσότητα αυτή διαχέεται αναλογικά από στρώμα σε στρώμα και προς

τα πίσω. Έτσι για το αμέσως προηγούμενο κρυφό στρώμα, το σφάλμα του νευρώνα προέλευσης, υπολογίζεται σε εξάρτηση με το σφάλμα του στρώματος εξόδου ως:

$$\delta_j = f'(u_j) \sum_{i=1}^h w_{ij} \delta_i \quad (3.17)$$

όπου w_{ij} τα βάρη όλων των συνάψεων με i τον δείκτη του νευρώνα προορισμού μέσα στο στρώμα και j τον δείκτη του νευρώνα προέλευσης. Ο παραπάνω τύπος ισχύει για κάθε προηγούμενο στρώμα σε σχέση με το επόμενο. Βλέπουμε έτσι το σφάλμα διαδίδεται προς τα πίσω (back propagates) στον νευρώνα προέλευσης από τα επιμέρους σφάλματα των νευρώνων προορισμού, ζυγισμένα επί του συναπτικού βάρους. Κατ' αυτόν τον τρόπο θα τροποποιηθούν τα συναπτικά βάρη όλου του δικτύου με βάση το πόσο συνεισφέρουν στην δημιουργία του τελικού σφάλματος στην έξοδο. Η διαδικασία ενημέρωσης θα διατρέξει όλες τα στρώματα που υπάρχουν προς τα πίσω, μέχρι τις εξερχόμενες συνάψεις που κάνουν οι νευρώνες του στρώματος εισόδου. Όλη η διαδικασία θα επαναληφθεί την επόμενη χρονική στιγμή $k + 1$ τροποποιώντας ξανά τα βάρη και ούτω καθεξής.

Ο τερματισμός της εκπαίδευσης

Οι μαθηματικοί υπολογισμοί που είδαμε εκτελούνται επαναληπτικά με ζητούμενο την ελαχιστοποίηση του σφάλματος της εξόδου. Για τον τερματισμό της επανάληψης μετά από κάποιες εποχές εκπαίδευσης, όταν δηλαδή έχει επέλθει μια κατάσταση που το σφάλμα έχει ελαττωθεί ικανοποιητικά, υπάρχουν τα κριτήρια τερματισμού του back propagation.

Κριτήριο 1: Το μέσο τετραγωνικό σφάλμα να είναι μικρότερο από ένα κατώφλι ϵ που έχουμε ορίσει

$$J(\mathbf{w}_{ij}) = \left(\frac{1}{m}\right) \sum_{i=1}^m |t_i - y_i|^2, \quad J(\mathbf{w}_{ij}) < \epsilon \quad (3.18)$$

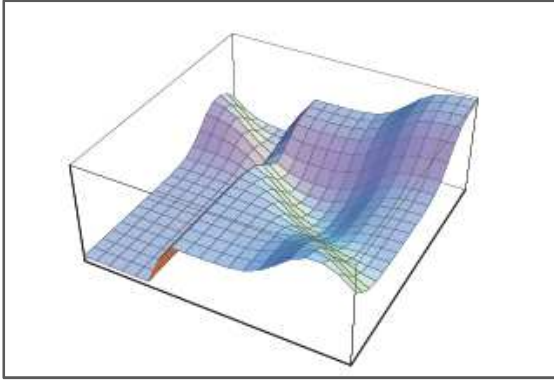
Κριτήριο 2: Το μέσο τετραγωνικό σφάλμα μεταξύ δύο διαδοχικών εποχών να έχει τροποποιηθεί ελάχιστα, βάση μιας μικρής ανοχής α .

$$J(\mathbf{w}_{ij}(k-1)) - J(\mathbf{w}_{ij}(k)) < \alpha \quad (3.19)$$

Κριτήριο 3: Σε μια εποχή τα συναπτικά βάρη του δικτύου έχουν υποστεί αμελητέα τροποποίηση σύμφωνα με ένα κατώφλι μ που έχουμε ορίσει. Αυτό γίνεται αθροίζοντας τις απόλυτες τιμές ή τα τετράγωνα των τροποποιήσεων σε κάθε εποχή.

$$\sum_{i,j} (\Delta_{w_{ij}})^2 < \mu \quad \text{είτε} \quad \sum_{i,j} |\Delta_{w_{ij}}| < \mu \quad (3.20)$$

Στις περιπτώσεις που ικανοποιούνται τα κριτήρια τερματισμού, το δίκτυο έχει φτάσει σε μια κατάσταση όπου θεωρείται εκπαιδευμένο, και ικανό να παράγει με καλή προσέγγιση τους στόχους που δόθηκαν σε κάθε πρότυπο εκπαίδευσης.



3-15: Τοπικά Ελάχιστα

απόκλιση από τους στόχους και για αυτό έχουν προταθεί διάφορες παραλλαγές του back propagation. Μια από αυτές που χρησιμοποιείται στο Hogasis είναι η χρήση της ορμής (momentum) στην τροποποίηση των συναπτικών βαρών.

Πρακτικά ο αλγόριθμος μας είναι μια μπάλα που κατηφορίζει μια πλαγιά της επιφάνειας του σφάλματος, μέχρι να φτάσει το ελάχιστο σημείο που μπορεί να σταθεροποιηθεί. Το πρόβλημα που μπορεί να συμβεί είναι να φτάσουμε σε ένα τοπικό ελάχιστο στην προσπάθεια της ελαχιστοποίησης της συνάρτησης του σφάλματος. Αυτό συμβαίνει σε πολλές περιπτώσεις όταν οι στόχοι για την έξοδο του δικτύου είναι διαφορετικοί από τα 0 ή 1 [59]. Σε ένα τοπικό ελάχιστο μπορεί να υπάρξει μεγάλη

Η έννοια της ορμής στο back propagation και η σημασία του βήματος εκπαίδευσης

Αν σκεφτούμε την προηγούμενη παρομοίωση της εκπαίδευσης back propagation ως μπάλα και κάποια στιγμή που κατηφορίζει την πλαγιά, θεωρούμε ότι έχει κάποια ορμή και έτσι συντηρεί κατά ένα βαθμό την πορεία που είχε στην αμέσως προηγούμενη χρονική στιγμή [54]. Όταν φτάνει σε κάποιο οροπέδιο ενός τοπικού ελαχίστου έχοντας αυτήν την ορμή, η εκπαίδευση δεν χάνει την ταχύτητα της αμέσως, και η ορμή θα την βοηθήσει να φτάσει μέχρι την άκρη του οροπεδίου όπου η «κατηφόρα» θα αναλάβει πάλι να την σπρώξει προς τα κάτω. Με αυτήν την λογική τροποποιείται η κατάβαση δυναμικού, ώστε σε κάθε τροποποίηση ενός συναπτικού βάρους να συνυπολογίζεται κατά ένα ποσοστό η προηγούμενη τροποποίηση που έγινε:

$$\Delta w_{OUT_{ij}}(k) = -\beta \left(\frac{\partial J(w_{OUT_{ij}}(k))}{\partial w_{OUT_{ij}}(k)} \right) + \mu \Delta w_{OUT_{ij}}(k-1) \quad (3.21)$$

Το ποσοστό αυτό, που εμφανίζεται στον παραπάνω τύπο ως η σταθερά μ , λέγεται συντελεστής ορμής και παίρνει τιμές συνήθως μεταξύ 0 και 1. Προφανώς στην περίπτωση που παίρνει την τιμή 0 η ορμή απενεργοποιείται πλήρως και έχουμε τον απλό τύπο του back propagation. Με την ορμή ο τελικός μαθηματικός τύπος τροποποίησης των συναπτικών βαρών σε μια εποχή εκπαίδευσης k για ένα συγκεκριμένο στρώμα νευρώνων διαμορφώνεται ως εξής:

$$\Delta w_{ij}(k) = \beta \delta_i x_j - \mu \Delta w_{ij}(k-1) \quad (3.22)$$

Πλέον έχουμε δύο σταθερές που καθορίζουν τον τρόπο με τον οποίο εκπαιδεύονται τα συναπτικά βάρη και ο σωστός συνδυασμός τους οδηγεί σε μεγάλη ακρίβεια στην λύση και σε γρηγορότερη εύρεση της. Αν αγνοήσουμε την επίδραση της ορμής και επικεντρωθούμε στο βήμα εκπαίδευσης, βλέπουμε ότι μικρές τιμές του δημιουργούν λεπτομερείς τροποποιήσεις των συναπτικών βαρών αλλά παράλληλα απαιτείται μεγαλύτερο πλήθος εποχών για να συγκλίνει σε κάποια λύση. Είναι επειδή το β αντιπροσωπεύει χρόνο και απαιτούνται πολλές μικρές διακριτές χρονικές στιγμές για να παρέλθει ο ίδιος χρόνος t , στον οποίο η εκπαίδευση θα φτάσει σε λύση. Αν από την άλλη δοθεί μεγάλη τιμή η εκπαίδευση φτάνει σε λίγες εποχές στον χρόνο t τροποποιώντας τα βάρη όμως με πολύ πιο χονδροειδή τρόπο με κίνδυνο να αποκλίνουμε από την λύση φτάνοντας σε κάποιο τοπικό ελάχιστο. Το παράδειγμα της μπάλας και εδώ θα μας

βοηθήσει αν σκεφτούμε ότι όσο αυτή κατηφορίζει, αλλάζουν την πορεία της πολλά πόδια με διαδοχικά ελαφρά λακτίσματα ακριβείας, «ντριμπλάροντας» έτσι τα εμπόδια. Αν τα λακτίσματα είναι λιγότερα και με μεγαλύτερη δύναμη, οι «σέντρες» θα καλύψουν γρήγορα την απόσταση προς τον στόχο, αλλά η πορεία που θα πάρει μπορεί να αποκλίνει από τον στόχο η χειρότερα να χτυπήσει σε ένα εμπόδιο και να επιστρέψει πίσω. Αυτό το φαινόμενο εμφανίζεται όταν το MSE αυξομειώνεται, ο αλγόριθμος ταλαντώνεται και τελικά γίνεται ασταθής. Γενικά το βέλτιστο βήμα εκπαίδευσης είναι δύσκολο να προκαθοριστεί γιατί αλλάζει κατά την διάρκεια της εκπαίδευσης, καθώς ο αλγόριθμος κινείται πάνω στην επιφάνεια της απόδοσης του [65]. Έτσι έχουν προταθεί μέθοδοι για δυναμική τροποποίηση του βήματος εκπαίδευσης ώστε να βελτιώνουν τον ρυθμό σύγκλισης στην λύση, αλλά και τα χαρακτηριστικά της σύγκλισης όπως σταθερότητα στην εκπαίδευση και ευρωστία απέναντι στις ταλαντώσεις [66]. Από την άλλη μια στατική μέθοδος για την ταχύτερη σύγκλιση είναι να χρησιμοποιούνται διαφορετικά βήματα εκπαίδευσης για κάθε στρώμα του δικτύου ή ακόμα και για το κάθε συναπτικό βάρος. Την τελευταία μέθοδο την χρησιμοποιεί η εργασία στις εκπαιδεύσεις των ΤΝΔ που υλοποιεί.

Τα βήματα του αλγόριθμου back-propagation

Έχοντας γνωρίζει το μαθηματικό μοντέλο του back propagation με την χρήση ορμής, μπορούμε να περιγράψουμε τον αλγόριθμο συνοπτικά με έναν ιδιότυπο ψευδοκώδικα που εμπνέεται από την γλώσσα Pascal αλλά, τους τελεστές της C++ και ενσωματώνει αυτούσιους τους μαθηματικούς τύπους που γνωρίσαμε στα σημεία που χρησιμοποιούνται.

```
// Η επαναληπτική διαδικασία του αλγόριθμου Back-Propagation
Πρόγραμμα Εκπαίδευση_Back_Propagation()
έναρξη
    Αρχικοποίηση-Βαρών() θέτει τα w[] ;

    Πρέπει-να-τερματίσει := 'Όχι';
    Εποχή                 := 0;

    ενώ δεν Πρέπει-να-τερματιστεί επανέλαβε
    αρχή
        για κάθε (Πρότυπο από το Σύνολο-Προτύπων) επανέλαβε
        αρχή
            Ανάκληση(Πρότυπο.s, w), υπολογίζει τα y[], u[];
            Οπισθοδιάδοση-Σφάλματος(y, Πρότυπο.t, u), υπολογίζει τα Dw[];
            σε κάθε w[!] θέσε w[!] += Dw[!]; //Όπου ! τριπλή δεικτοδότηση
        τέλος.

        MSE := 0;
        για κάθε (Πρότυπο από το Σύνολο-Προτύπων) επανέλαβε
        αρχή
            Ανάκληση(Πρότυπο.s, w), υπολογίζει τα y[], u[];
            MSE += Μέσο-Τετραγωνικό-Σφάλμα(y, Πρότυπο.t);
        τέλος.

        Πρέπει-να-τερματιστεί := (MSE < e) είτε (Abs(MSE - PrevMSE) < a);

        PrevMSE := MSE;
        Εποχή   += 1;
    τέλος;
τερματισμός.
```

Τα δεδομένα που βλέπουμε να χρησιμοποιούνται παραπάνω με την σειρά εμφάνισης είναι:

- $w[L, N, D]$ όλα τα συναπτικά βάρη του δικτύου στα οποία γίνεται πρόσβαση με τρεις δείκτες όπου L ο αριθμός στρώματος, N η θέση του νευρώνα προορισμού στο στρώμα L , D ο αριθμός σύναψης του νευρώνα N ή η θέση του νευρώνα προέλευσης στο προηγούμενο στρώμα $L - 1$.
- **Πρέπει-να-τερματίσει**, μεταβλητή ελέγχου του κεντρικού βρόχου επανάληψης της εκπαίδευσης.
- **Εποχή**, δείχνει την εποχή στην οποία βρίσκεται η εκπαίδευση.
- **Πρότυπο**, μια δομή για το πρότυπο εκπαίδευσης, που περιλαμβάνει τα διανύσματα **Πρότυπο.s** με τις τιμές των σημάτων εισόδου του δικτύου και **Πρότυπο.t** με τις τιμές των στόχων για τους νευρώνες εξόδου.
- **Σύνολο-Προτύπων**, η συλλογή που περιλαμβάνει όλα τα πρότυπα εκπαίδευσης.
- $u[L, N]$ και $y[L, N]$ τα συναπτικά αθροίσματα και οι τιμές των συναρτήσεων ενεργοποίησης για κάθε νευρώνα N σε κάθε στρώμα L του δικτύου.
- $Dw[L, N, D]$ οι μεταβολές των συναπτικών βαρών που υπολογίζονται σε μια εποχή k .
- **MSE, PrevMSE** το μέσο τετραγωνικό σφάλμα όλου του συνόλου εκπαίδευσης σε μια σε μια εποχή k και στην αμέσως προηγούμενη $k-1$, αντίστοιχα.
- e το κατώφλι του κριτηρίου 1 τους τερματισμού της εκπαίδευσης.
- a η ανοχή του κριτηρίου 2 του τερματισμού της εκπαίδευσης.

Αναλυτικότερα έχουμε μέσα στο αλγόριθμο τις υπορουτίνες με σειρά εμφάνισης:

// 1. Αρχικοποίηση των συναπτικών βαρών με τυχαίες τιμές

```

διαδικασία Αρχικοποίηση-Βαρών() θέτει τα w[];
αρχή
για (L := 1 ως LayerCount) και για (N := 1 ως NeuronCount[L]) δράσε
αρχή
αν L = 1 τότε
    w[L, N, 1] := 1
    ειδάλλως
        για D := 1 ως NeuronCount[L - 1] δράσε
            w[L, N, D] := Τυχαία-Τιμή();
τέλος;
τέλος;

```

Τα δεδομένα που χρησιμοποιούνται :

- **LayerCount**, το πλήθος των στρωμάτων στο δίκτυο
- **NeuronCount[L]**: πίνακας [1 .. LayerCount] είναι το πλήθος των νευρώνων για κάθε στρώμα L . Υπάρχουν LayerCount στοιχεία στον πίνακα με $L = 1$ να δείχνει το στρώμα εισόδου και $L = LayerCount$ το στρώμα εξόδου.

// 2. Ανάκληση με υπολογισμό των συναπτικών αθροισμάτων και των τιμών των συναρτήσεων ενεργοποίησης όλων των νευρώνων του δικτύου

```

διαδικασία Ανάκληση(s, w) υπολογίζει τα y[], u[];
αρχή
  για N := 1 ως NeuronCount[1] θέσε
    x[1, N, 1] = s[N];

  για (L := 2 ως LayerCount) και για (N := 1 ως NeuronCount[L]) δράσε
  αρχή
    m = NeuronCount[L - 1];
    u[L,N] :=  $\sum_{i=1}^m w_i x_i$ ;
    y[L,N] := f(u[L,N]);
    αν L < LayerCount τότε
      για I := 1 ως NeuronCount[L + 1] θέσε
        x[L + 1, I, N] := y[L, N];
      τέλος;
  τέλος;
τέλος;

```

Τα δεδομένα που πρωτοεμφανίζονται εδώ :

- $x[L, N, D]$ η εισερχόμενη τιμή στην σύναψη D του νευρώνα N στο στρώμα L .

// 3. Διάδοση του σφάλματος της εξόδου προς τα πίσω, υπολογίζοντας τις μεταβολές των συναπτικών βαρών.

```

διαδικασία Οπισθοδιάδοση-Σφάλματος(y, t, u), υπολογίζει τα Dw[];
αρχή
  για (L := LayerCount κάτω στο 1) και για (N := 1 ως NeuronCount[L])
  δράσε
  αρχή
    αν L = LayerCount τότε
      d[L,N] := (t[N] - y[L,N]) f'(u[L,N]);
      ειδάλλως
      αρχή
        m := NeuronCount[L + 1];
        d[L,N] :=  $\sum_{i=1}^m w[m, i, N] d[m, i]$ ;
        για D := 1 ως m θέσε
          Dw[L, N, D] := b * d[L, N] * x[L, N, D] - m * DwPrev[L, N, D];
        τέλος;
      τέλος;
    Dw_Prev[!] := Dw[!]; //Όπου ! τριπλή δεικτοδότηση
  τέλος;
τέλος;

```

Τα δεδομένα που πρωτοεμφανίζονται εδώ :

- $d[L, N]$ Οι τιμές για του παράγοντες σφάλματος δ του υπολογισμού των τροποποιήσεων των συναπτικών βαρών. Μια για τον κάθε νευρώνα N του κάθε στρώματος L .
- $Dw_Prev[L, N, D]$ οι μεταβολές των συναπτικών βαρών που είχαν υπολογιστεί την αμέσως παρελθούσα εποχή $k-1$, σε σχέση με την εποχή k που βρίσκεται ο αλγόριθμος.
- b το βήμα της εκπαίδευσης.
- m ο συντελεστής ορμής.

// 4. Συνάρτηση υπολογισμού του μέσου τετραγωνικού σφάλματος της εξόδου του δικτύου σε σχέση με τους στόχους, για ένα πρότυπο

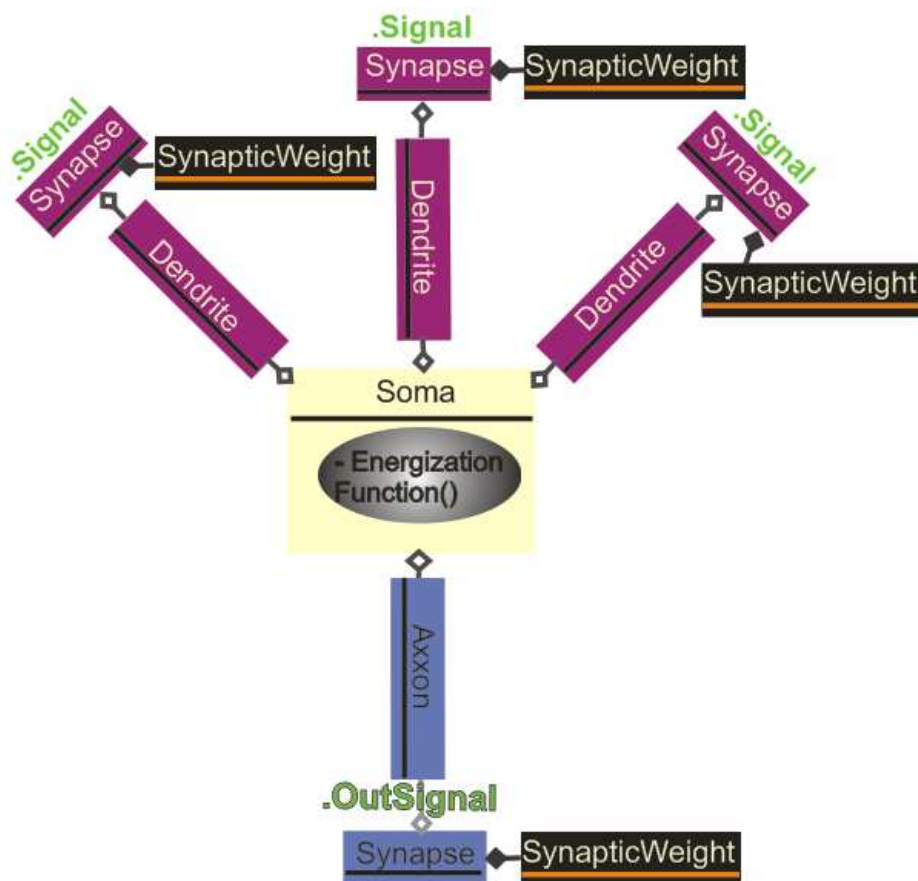
```
συνάρτηση Μέσο-Τετραγωνικό-Σφάλμα(y, t), αποτέλεσμα: πραγματικός;  
αρχή  
  m := NeuronCount[LayerCount];  
  αποτέλεσμα :=  $(\frac{1}{m}) \sum_{i=1}^m |t[i] - y[i]|^2$   
τέλος;
```

Η συνάρτηση επιστρέφει το αποτέλεσμα του τύπου υπολογισμού του μέσου τετραγωνικού σφάλματος, με δεδομένα τις τιμές εξόδων y των νευρώνων και τους στόχους t για το στρώμα εξόδου, που έχει δείκτη ίδιο με το πλήθος των στρωμάτων `LayerCount`.

4. Τα Αντικείμενα Τεχνητών Νευρώνων

Ο Αντικειμενοστραφής Τεχνητός Νευρώνας

Έχοντας σαν βάση το μαθηματικό και αλγοριθμικό μοντέλο των τεχνητών νευρώνων, αλλά επιστρέφοντας ξανά πίσω στο βιολογικό μοντέλο, σχεδιάστηκε ο αντικειμενοστραφής τεχνητός νευρώνας. Αποτελείται από ένα σύνολο κλάσεων που έχουν στην ονομασία τους το πρόθεμα T, σύμφωνα με την σύμβαση κώδικα (code convention) της γλώσσας Delphi με την οποία ξεχωρίζεται αμέσως οποιαδήποτε ονομασία κλάσης, ακολουθούμενο από το γράμμα G. Οποιαδήποτε κλάση θεωρείται επαναχρησιμοποιήσιμη ονοματοδοτείται με πρόθεμα TG και με αυτόν τον τρόπο θεωρείται μέρος του framework που στηρίζει την υλοποίηση του Horasis. Το ίδιο framework μπορεί να χρησιμοποιηθεί αυτόνομα για μελλοντική υλοποίηση άλλων τεχνητών νευρωνικών δικτύων



4-1: Σχηματική αναπαράσταση με στοιχεία UML, του αντικειμενοστραφούς τεχνητού νευρώνα ©2011 Horasis

Η αναφορά στις κλάσεις θα γίνει συνοπτικά προσπαθώντας να δοθεί η ουσία του σχεδιασμού του αντικειμενοστραφούς τεχνητού νευρώνα και όχι οι επιμέρους λεπτομέρειες. Η αρχιτεκτονική παρατίθεται ως παράδειγμα για μελλοντικές παρόμοιες υλοποιήσεις ή μεταφορά σε άλλες γλώσσες προγραμματισμού. Έτσι στο ιδιότυπο διάγραμμα κλάσεων UML [41] της εικόνας 4-1, οι κλάσεις ονοματίζονται χωρίς τα δύο γράμματα του προθέματος, για να παραθέτουν τον σχεδιασμό ανεξάρτητα από την παρούσα υλοποίηση.

Η κλάση του νευρώνα και τα συστατικά της

Η κλάση `TGNeuron` είναι αυτή που αντιστοιχεί στην οντότητα του τεχνητό νευρώνα. Όλα τα συστατικά αντικείμενα του αντικειμενοστραφούς νευρώνα δημιουργούνται και καταστρέφονται μαζί με τον `TGNeuron`. Συστατικά αποτελούν ένα αντικείμενο της κλάσης `TGSoma` που αντιστοιχεί στο σώμα, ένα αντικείμενο `TGAxon` για τον άξονα και ένα πλήθος από αντικείμενα-δενδρίτες `TGDendrite`. Η `TGAxon` και η `TGDendrite` κληρονομούν από την ίδια προγονική κλάση `TGNeuronFiber` καθώς είναι και οι δύο νευρωνικές ίνες που ενδέχεται να δημιουργήσουν μεταξύ τους συνάψεις. Η ίδια η σύναψη είναι ένα αντικείμενο της κλάσης `TGSynapse` και γι αυτό στην `TGNeuronFiber` υπάρχει μια λίστα από αντικείμενα `TGSynapse`, δημιουργώντας σχεδιαστικά συνάθροιση (aggregation) όλων των συνάψεων στις οποίες συμμετέχει η ίνα. Η έξοδος του νευρώνα είναι η ιδιότητα `OutSignal` που υπάρχει στο αντικείμενο-άξονα, μαζί με την `TargetOutSignal` που είναι ο στόχος για την έξοδο του νευρώνα. Έτσι κάθε νευρώνας μπορεί εν δυνάμει να έχει έναν δικό του στόχο εκπαίδευσης, ανεξάρτητα αν ανήκει στο στρώμα εξόδου. Δημιουργείται έτσι μια νέα προοπτική για έρευνα, αν οριστούν στόχοι για ενδιάμεσους νευρώνες στα κρυφά στρώματα σταματώντας την προς τα πίσω διάδοση του σφάλματος της εξόδου του δικτύου σε αυτούς. Για κάθε έναν από αυτούς θα ξεκινήσει ένα νέο back propagation, μια νέα κατάβαση δυναμικού για την ελαχιστοποίηση του σφάλματος πλέον στην μέση του δικτύου, δημιουργώντας ομάδες νευρώνων που εκπαιδεύονται ανεξάρτητα από όλο το δίκτυο.

Το αντικείμενο-σώμα `TGSoma` έχει τις ευθύνες της ανάκλησης του νευρώνα, του υπολογισμού του σφάλματος κατά την εκπαίδευση και της οπισθοδιάδοσης προς τους νευρώνες προέλευσης που συνάπτονται με τα αντικείμενα `TGDendrite`. Έτσι υλοποιείται εμφωλευμένη η συνάρτηση ενεργοποίησης και η παράγωγος της, αντιπροσωπεύοντας στην ουσία την νευρωνική μεμβράνη για την οποία δεν υπάρχει ξεχωριστό αντικείμενο. Επειδή η διαδικασία της νευρωνικής ολοκλήρωσης αφορά διαύλους ιόντων που βρίσκονται στο εσωτερικό του σώματος [52] σχεδιάστηκε σε αυτό το σημείο η αντίστοιχη τεχνητή διαδικασία που την προσομοιώνει.

Με την μέθοδο ανάθεση μιας ιδιότητας `DendritesCount` που κρατάει το πλήθος των δενδριτών του νευρώνα, δημιουργούνται αυτόματα όλα τα υποσυστατικά `TGDendrite`. Σε κάθε ένα από αυτά δίνεται αντίστοιχα και μια ιδιότητα `Index`, η οποία θα δείχνει την θέση του δενδρίτη στον νευρώνα υλοποιώντας τον δείκτη i του γνωστού γινομένου $W_i X_i$.

Η σύναψη ως αντικείμενο

Το αντικείμενο-σύναψη `TGSynapse` τοποθετείται ως αναφορά και στις δύο λίστες του μεν αντικειμένου-άξονα που προέρχεται και του δε αντικειμένου-δενδρίτη στον οποίο καταλήγει, υλοποιώντας μια αξονοδενδρική σύναψη. Η `TGSynapse` θα κρατήσει αναφορές για τα δύο αντικείμενα που συνδέονται, με την ιδιότητα (attribute) `SourceFiber: TGNeuronFiber` για την προέλευση και την ιδιότητα `DestFiber: TGNeuronFiber` για τον προορισμό. Σχεδιαστικά αφήνεται η δυνατότητα για την ύπαρξη και αξονοαξονικής σύναψης, χωρίς όμως να έχει υλοποιηθεί κάτι τέτοιο. Το αντικείμενο-σύναψη έχει ως συστατικό υποαντικείμενο το συναπτικό βάρος `TGSynapticWeight` που κρατάει την τιμή του και μια ιδιότητα `Signal` στην οποία δίδεται το εισερχόμενο στη σύναψη σήμα. Μέσα από μια μέθοδο ανάθεσης (setter method) όταν ανατίθεται τιμή στην `Signal` υπολογίζεται μια ιδιότητα:

Value := Signal * SynapticWeight.Value

με την οποία επιστρέφει το γνωστό γινόμενο που χρησιμοποιείται για την παραγωγή του συναπτικού αθροίσματος. Δηλαδή η Value επιστρέφει την ζυγισμένη τιμή της σύναψης. Εκτός αυτού του υπολογισμού η σύναψη δεν έχει άλλες μεθόδους καθώς στην πραγματικότητα είναι μια σχισμή που κατά την διαβίβαση της πληροφορίας την τροποποιεί.

Το συναπτικό βάρος ως ενεργό αντικείμενο.

Το συναπτικό βάρος είναι ένα αντικείμενο που κρατάει πληροφορίες και υλοποιεί εμφωλευμένα τον τρόπο που αρχικοποιείται και που αυτοτροποποιείται. Έχει μια αναφορά στην σύναψη που ανήκει με την ιδιότητα Synapse: TGSynapse. Οι πληροφορίες που κρατάει είναι η αρχική αριθμητική τιμή του βάρους InitValue που τίθεται κατά την μέθοδο Init(), η τρέχουσα τιμή του βάρους Value και η αμέσως προηγούμενη τιμή του βάρους OldValue. Για την τροποποίηση $\Delta w_{ij}(k)$ που υφίσταται στην τρέχουσα εποχή k , έχει την ιδιότητα WeightDifference ενώ για την προηγούμενη τροποποίηση $\Delta w_{ij}(k-1)$ υπάρχει η OldWeightDifference. Και τα δύο τίθενται και αξιοποιούνται από την μέθοδο BackPropagate() που υπάρχει στην κλάση TGSynapticWeight, όπου υλοποιείται η ενημέρωση ενός συγκεκριμένου συναπτικού βάρους με τον αλγόριθμο back propagation.

Για τον λόγο αυτό κάθε αντικείμενο TGSynapticWeight έχει δικές του οριζόμενες ιδιότητες για το βήμα εκπαίδευσης TrainingStep και τον συντελεστή ορμής MomentumFactor. Αυτές μπορεί να αντιγράφονται από μια πιο κεντρική ρύθμιση όπως για όλες τις συνάψεις του νευρώνα ή όλες τις συνάψεις ενός στρώματος νευρώνων, αλλά δίνουν παράλληλα την δυνατότητα ύπαρξης διαφορετικού βήματος εκπαίδευσης και ορμής σε κάθε σύναψη του νευρωνικού δικτύου. Με αυτόν τον τρόπο εμπλουτίζεται ο αλγόριθμος back propagation ανοίγοντας ένα νέο δρόμο για την ταχύτερη σύγκλιση και την ακρίβεια της λύσης. Πλέον μπορούν να δοκιμαστούν απευθείας σε υπάρχοντα αντικειμενοστραφή νευρωνικά δίκτυα, μέθοδοι για την επιλογή των βημάτων εκπαίδευσης και των ορμών, βγάζοντας άμεσα συμπεράσματα δια της δοκιμής και απόρριψης. Επίσης κατά την διάρκεια της εκπαίδευσης, οι πρώην σταθερές για όλες τις συνάψεις, είναι πλέον μεταβλητές ανά σύναψη, μπορούν να τροποποιούνται δυναμικά εμπλουτίζοντας είτε το αντικείμενο-νευρώνα ή το αντικείμενο-συναπτικού βάρους με νέες προσεγγίσεις εκπαίδευσης του ΤΝΔ. Έτσι μέσα στην εμφωλευμένη BackPropagate() υπάρχει ο υπολογισμός

```
WeightDifference := TrainingStep * SomaError * Synapse.Signal
                    + MomentumFactor * OldWeightDifference
```

που βρίσκεται σε αντιστοιχία με τον μαθηματικό τύπο 3.22 :

$$\Delta w_{ij}(k) = \beta \delta_i x_j - \mu \Delta w_{ij}(k-1)$$

Το SomaError περνιέται ως παράμετρος στην BackPropagate() καθώς ο υπολογισμός του βρίσκεται εντός των σχεδιαστικών ορίων της σύναψης και είναι Η παράμετρος δ_i του σφάλματος του νευρώνα προορισμού μιας σύναψης. Αυτή υπολογίζεται στο αντικείμενο-σώμα, που εμφωλεύει την συνάρτηση ενεργοποίησης και την παράγωγο της. Η BackPropagate() καλείται μέσα σε μια μέθοδο της TGSoma, κατά την εκπαίδευση και για κάθε δενδρίτη διαδίδεται προς τα πίσω ένα σφάλμα που ήδη έχει υπολογιστεί. Στο τέλος της μεθόδου κρατιέται το τρέχον συναπτικό βάρος για χρήση στην επόμενη εποχή με τον συντελεστή ορμής:

```
OldWeightDifference := WeightDifference
```

Συναπτικές ρυθμίσεις για την λειτουργία του αντικειμένου-βάρους

Οι ρυθμίσεις για την αρχικοποίηση και την λειτουργία του συναπτικού βάρους τοποθετούνται σε αντικείμενα της κλάσης `TGSynapticProperties`. Για κάθε τεχνητό νευρώνα υπάρχει ένα τέτοιο συστατικό ως ιδιότητα `TGNeuron.Properties`. Η κλάση υλοποιεί την μέθοδο ανάθεσης της ιδιότητας (setter method) έτσι μπορεί εύκολα να αντιγραφεί η ρύθμιση από ένα αντικείμενο σε άλλο. Η πρώτη από τις ιδιότητες που κρατάει είναι μια τιμή βάσης για την αρχικοποίηση του συναπτικού βάρους `SynapticWeightNorm`, η οποία χρησιμοποιείται κατά την `TGSynapticWeight.Init()` ως εξής: Χρησιμοποιώντας γεννήτρια ψευδοτυχαίων αριθμών και βάσει συγκεκριμένων πιθανοτήτων για την κάθε περίπτωση, παράγονται αρχικά βάρη που είναι συνάρτηση της `SynapticWeightNorm`. Με την αντίστοιχη πιθανότητα να συμβεί στο αριστερό τμήμα, επιλέγονται τα αρχικά βάρη:

- 16,667%: `InitValue := SynapticWeightNorm`
- 33,333%: `InitValue := SynapticWeightNorm / 2`
- 33,333%: `InitValue := -SynapticWeightNorm / 2`
- 16,667%: `InitValue := -SynapticWeightNorm`

Έχουμε δηλαδή έναν νέο τρόπο αρχικοποίησης των συναπτικών βαρών, όχι με εντελώς τυχαία αλλά με ομοιόμορφα βάρη που ακολουθούν κάποια κατανομή πιθανοτήτων. Έχοντας διαφορετικές υλοποιήσεις της `Init()` μπορούν να γίνουν εύκολα πειράματα για άλλου είδους αρχικοποιήσεις των συναπτικών βαρών. Αυτός ο τρόπος αρχικοποίησης φάνηκε στην πράξη ότι βελτιώνει την ταχύτητα σύγκλισης και την σταθερότητα της εκπαίδευσης του back propagation.

Οι άλλες δύο ιδιότητες είναι ο συντελεστής ορμής `MomentumFactor` που θα χρησιμοποιηθεί σε όλα τα `TGSynapticWeight` που λειτουργούν για έναν `TGNeuron` και η ακρίβεια της εκπαίδευσης `TrainingStepScale`, που καθορίζει το βήμα εκπαίδευσης:

$$TrainingStep = \left(\frac{1}{TrainingStepScale} \right) \quad (4.1)$$

Με μεγαλύτερη τιμή στην ακρίβεια `TrainingStepScale`, ορίζεται ένα μικρότερο βήμα εκπαίδευσης, που μπορεί να οριστεί για το κάθε αντικείμενο-συναπτικό βάρος μέσα στο ΤΝΔ.

Το Αντικειμενοστραφές Νευρωνικό Δίκτυο

Το δίκτυο και τα στρώματα των νευρώνων

Έχοντας γνωρίσει την κλάση `TGNeuron`, που έχει εμφωλευμένη όλη την λογική του νευρώνα, θα δούμε πως χρησιμοποιείται στην δημιουργία ενός τεχνητού νευρωνικού δικτύου από αντικείμενα. Η κλάση `TGNeuralNetwork` που αντιπροσωπεύει ένα MLP δίκτυο είναι μια συλλογή από `TGNeuralLayer` που υλοποιούν τα στρώματα των νευρώνων, με το κάθε ένα από αυτά να είναι μια συλλογή από αντικείμενα `TGNeuron`. Στο αντικείμενο-στρώμα υπάρχει ένας αριθμός `LayerIndex` που δείχνει την θέση του μέσα στο δίκτυο και δύο αναφορές σε αντικείμενα της ίδια κλάσης. Το `PreviousLayer: TGNeuralLayer` αναφέρεται στο προηγούμενο στρώμα πριν από αυτό και το `NextLayer: TGNeuralLayer` στο επόμενο, υλοποιώντας μια συνδεδεμένη λίστα. Επειδή η σχέση των κλάσεων εδώ είναι συνάθροιση (aggregation) υπάρχει η δυνατότητα σε ένα στρώμα, να αποσυνδεθεί από ένα νευρωνικό δίκτυο και να τοποθετηθεί σε κάποιο άλλο κατά την διάρκεια της ζωής του. Αυτό γίνεται με αφαίρεση από το πρώτο `TGNeuralNetwork`, προσθήκη στο δεύτερο και τροποποίηση των δύο αναφορών του ανάλογα με την νέα θέση που θα πάρει. Έτσι οι νευρώνες μεταφέρονται με τα συναπτικά βάρη τους χωρίς το κόστος επαναδημιουργίας του στρώματος και φόρτωσης των συναπτικών βαρών από την περιφερειακή στην κεντρική μνήμη του συστήματος.

Στο κάθε `TGNeuralLayer` τίθεται σε μια ιδιότητα `NeuronCount` το πλήθος των νευρώνων και μέσω της μεθόδου ανάθεσης (setter method), δημιουργείται την ίδια στιγμή που ορίζεται το αντίστοιχο πλήθος από αντικείμενα `TGNeuron`. Σε τυχόν δεύτερη ανάθεση της με μικρότερο μέγεθος νευρώνων αφαιρούνται τα αντικείμενα-νευρώνες που περισσεύουν, ενώ αν το πλήθος των νευρώνων του στρώματος αυξάνεται δημιουργούνται συμπληρωματικά μόνο οι νέοι νευρώνες που λείπουν. Έτσι μπορούμε να έχουμε ακόμα ένα πρωτοεμφανιζόμενο χαρακτηριστικό στο αντικείμενο-ΤΝΔ, ένα δυναμικά τροποποιούμενο σμήνος νευρώνων. Κατά την δημιουργία του κάθε `TGNeuron` του στρώματος, δίνεται αυτόματα τιμή σε μια ιδιότητα `Position`, η οποία δείχνει την θέση του αντικειμενοστραφούς νευρώνα μέσα στο αντικειμενοστραφές στρώμα. Επίσης με την ύπαρξη της ιδιότητα `Properties: TGSynapticProperties` ως συστατικό αντικείμενο του `TGNeuralLayer`, αντιγράφονται από αυτήν οι συναπτικές ρυθμίσεις στο αντίστοιχο `Properties: TGSynapticProperties` του `TGNeuron`, κατά την αρχικοποίηση του δικτύου.

Οι διασυνδέσεις των στρωμάτων

Υπάρχει μια μέθοδος `ConnectOutputTo()` με την οποία ένα στρώμα συνδέεται με ένα άλλο, υλοποιώντας την δομή διασυνδέσεων του MLP ενός προηγούμενου με το επόμενο στρώμα. Μέσα σε αυτήν για κάθε αντικείμενο-νευρώνα του στρώματος, δημιουργούνται συνάψεις με τα αντικείμενα νευρώνες του επομένου, κάνοντας χρήση των αντικειμένων `TGSynapse`. Σε μία σύναψη, ως νευρωνική ίνα προέλευσης `SourceFiber` ορίζεται ο `TGNeuron.Axon` του νευρώνα του στρώματος και ως προορισμός `DestFiber` ένας `TGDendrite` από το επόμενο στρώμα που έχει την ιδιότητα `Index` να ισούται με το `TGNeuron.Position` του νευρώνα προέλευσης. Επιπρόσθετη ευθύνη μέσα σε αυτήν την μέθοδο είναι η κλήση της `TGSynapticWeight.Init()` σε κάθε αντικείμενο-συναπτικού βάρος, δηλαδή να αρχικοποιήσει όλα τα συναπτικά βάρη για τις συνάψεις μεταξύ των δύο στρωμάτων.

Η έναρξη της χρήσης μιας σύναψης ξεκινάει από την στιγμή που συνδέει δύο άκρα, ενώ πριν ήταν **ουδέτερη** (null synapse) και σε αδράνεια. Γι αυτό και θεωρήθηκε η κατάλληλη στιγμή για την αρχικοποίηση των βαρών, η στιγμή κλήσης της μεθόδου `ConnectOutputTo` που συνδέει τα δύο στρώματα. Μηδενικές συνάψεις συνεχίζουν να υπάρχουν στα αντικείμενα-δενδρίτες του στρώματος εισόδου καθώς δεν υπάρχουν άξονες που συνδέονται σε αυτά, και τα συναπτικά τους βάρη αρχικοποιούνται στην τιμή 1. Με τον σχεδιασμό αυτό μπορούν εύκολα να συνδεθούν δύο διαφορετικά MLP δίκτυα, με τις ουδέτερες συνάψεις που υπάρχουν στην είσοδο του δεύτερου να γίνονται κανονικές συνδέοντας το με τις εξόδους του πρώτου. Έτσι σε ένα μελλοντικό πλαίσιο λογισμικού διαφορετικά αντικειμενοστραφή MLP, λειτουργώντας ως κατασκευές Lego, θα έχουν την δυνατότητα να συνδεθούν μεταξύ τους εύκολα με έναν και μοναδικό τρόπο διασύνδεσης για οποιαδήποτε περίπτωση.

Με την μέθοδο `Initialize()` της `TGNeuralNetwork` προετοιμάζεται για χρήση το αντικειμενοστραφές ΤΝΔ. Έχουν σχεδιαστεί δύο αναφορές η `InputLayer : TGNeuralLayer` και η `OutputLayer : TGNeuralLayer` για βοηθητική λειτουργία, οι οποίες συμπληρώνονται κατά την `Initialize()` με το στρώμα εισόδου και το στρώμα εξόδου αντίστοιχα. Το στρώμα εισόδου αυτοκαθορίζεται με την μη ύπαρξη αναφοράς σε `PreviousLayer` και το στρώμα εξόδου με την μη ύπαρξη αναφοράς σε `NextLayer`. Για την σωστή λειτουργία της πρέπει να βρεθούν στρώματα με αυτά τα χαρακτηριστικά στο δίκτυο. Επίσης συμπληρώνονται δύο λίστες αντικειμένων, η `NetworkInputs` που συναθροίζει όλες τις συνάψεις `TGSynapse` εισόδου του δικτύου και η `NetworkOutputs` που κρατάει όλους τους άξονες `TGΑχχο` της εξόδου του δικτύου. Έτσι έχουμε πλέον μια έτοιμη δομή από αντικείμενα στην οποία μπορεί να λειτουργήσει άμεσα η ανάκληση του δικτύου με κάποιες τιμές εισόδου, που όπως θα δούμε οργανώνονται και αυτές σε ένα νέο αντικείμενο.

Τα Πρότυπα Εκπαίδευσης

Το αντικείμενο ως ερέθισμα εισόδου και η συλλογή από πρότυπα εκπαίδευσης

Θα ορίσουμε ως **Αντικείμενο Νευρωνικού Μοτίβου (Neural Pattern Object – NPO)** ένα αντικείμενο το οποίο περιέχει τουλάχιστον έναν πίνακα με αριθμητικές τιμές που αντιπροσωπεύουν τα εισερχόμενα σήματα εισόδου σε ένα τεχνητό νευρωνικό δίκτυο. Επιπρόσθετα στο ίδιο αντικείμενο μπορεί να υπάρχει και ένας πίνακας με αριθμητικές τιμές που αντιπροσωπεύουν στόχους της εξόδου του δικτύου που προκύπτουν από αυτές τις τιμές εισόδου. Αυτό λειτουργεί ως ερέθισμα (stimuli) ξεκινώντας την τεχνητή νευρωνική επεξεργασία. Αυτό το αντικείμενο είναι και το πρότυπο για την εκπαίδευσης του αντικειμενοστραφούς ΤΝΔ.

Η κλάση `TGNeuralPattern` υλοποιεί αυτόν τον περιέκτη αριθμητικών τιμών. Έχει δύο ιδιότητες `InputSignalCount` και `OutputSignalCount` βάσει των οποίων δημιουργεί δυναμικούς πίνακες (dynamic arrays) δεσμεύοντας μνήμη μόνο κατά την φόρτωση των δεδομένων του μοτίβου. Μια συλλογή από αντικείμενα νευρωνικού μοτίβου είναι η `TGNeuralTrainingSet` η οποία έχει την ευθύνη να κρατάει, αποθηκεύει, φορτώνει και να διατρέχει αντικείμενα τύπου `TGNeuralPattern`. Είναι στην ουσία το σύνολο των προτύπων εκπαίδευσης του τεχνητού νευρωνικού δικτύου που υλοποιείται ως ένα αυτόνομο σύνολο δεδομένων (data set). Επίσης υπάρχουν οι ιδιότητες λίστας συμβολοσειρών `InputDescription` και `OutputDescription`, που περιγράφουν την κάθε πληροφορία εισόδου και εξόδου του προτύπου.

Ανάκληση του αντικειμενοστραφούς δικτύου

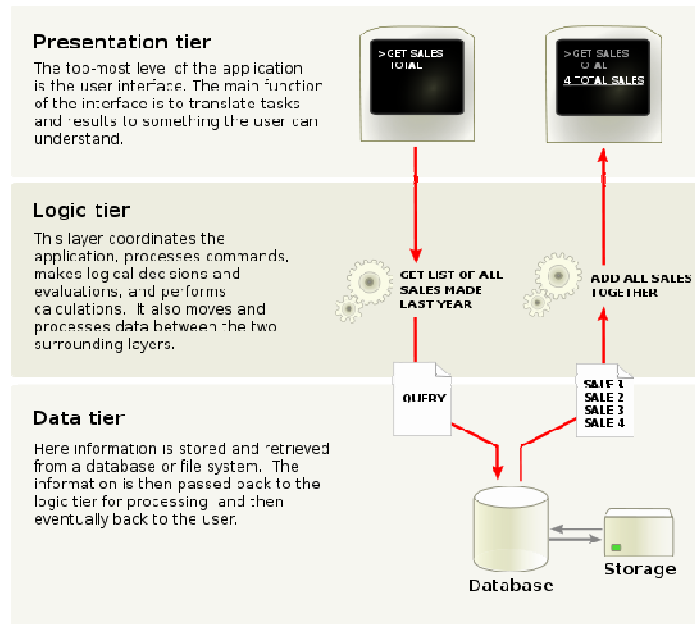
Η διαδικασία της ανάκλησης σε ένα `TGNeuralNetwork` βρίσκεται μέσα στην μέθοδο `Stimulate()` στην οποία δίδεται ως παράμετρος ένα φορτωμένο αντικείμενο `TGNeuralPattern`, αποτελώντας ερέθισμα στην είσοδο του δικτύου. Λόγω του σχεδιασμού που αναφέραμε η διαδικασία που ακολουθείται για την ανάκληση του δικτύου είναι απλούστατη.

- Σε κάθε μια σύναψη από τις `NetworkInputs`: Ανατίθεται στην η ιδιότητα `TGSynapse.Signal` η αντίστοιχη τιμή εισόδου από το αντικείμενο `TGNeuralPattern`.
- Σε κάθε έναν `TGAxxon` από τους `NetworkOutputs` : Ανατίθεται στην ιδιότητα `TGAxxon.TargetOutSignal` η αντίστοιχη τιμή στόχου από `TGNeuralPattern`, ασχέτως με το αν βρισκόμαστε μέσα σε εκπαίδευση.
- Για κάθε στρώμα του δικτύου και για κάθε νευρώνα μέσα στο στρώμα του δικτύου καλείται η μέθοδος `Process()` εκτελώντας την λειτουργία του τεχνητού νευρώνα

Η `Process()` υλοποιεί τις ευθύνες της ανάκλησης του νευρώνα και τελικά εμφωλεύεται μέσα στην κλάση `TGSoma` που όπως προαναφέραμε προσομοιώνει την νευρωνική ολοκλήρωση. Είναι ουσιαστικά η μεταφορά του μοντέλου `McCulloch-Pitts` στον αντικειμενοστραφή προγραμματισμό. Μέσα της αθροίζονται όλες οι ζυγισμένες τιμές των συνάψεων `TGSynapse.Value` και το με το άθροισμα αυτό καλείται μια συνάρτηση `EnergizationFunction` που επιστρέφει την τιμή στην ιδιότητα `TGAxxon.OutSignal`. Απόγονες κλάσεις της `TGSoma` μπορούν να αντικαταστήσουν (override) αυτήν την συνάρτηση υλοποιώντας διαφορετικές συναρτήσεις ενεργοποίησης, χωρίς καμία περαιτέρω αλλαγή στον σχεδιασμό και την λειτουργία του νευρωνικού δικτύου. Παράλληλα υπάρχει και η `EnergizationDerivative()` ως η παράγωγος της συνάρτησης ενεργοποίησης που χρησιμοποιείται για τον υπολογισμό του σφάλματος στην έξοδο σύμφωνα με τον αλγόριθμο `back propagation`, και που θα συναντήσουμε κατά την εκπαίδευση του αντικειμενοστραφούς ΤΝΔ.

Η τριών επιπέδων σχεδίαση και η διαχείριση των προτύπων ως δεδομένα

Βλέπουμε ότι τα δεδομένα εισόδου του νευρωνικού δικτύου έχουν απομονωθεί σε σχέση με την λογική της ανάκλησης ώστε να υπάρχει η ελάχιστη αλληλεπίδραση των αντικειμένων του νευρωνικού δικτύου με αυτά. Ο σχεδιασμός αυτός στηρίχθηκε σε μια κατεύθυνση της μηχανικής λογισμικού που υποστηρίζεται κυρίως από την Microsoft, την αρχιτεκτονική 3 επιπέδων (3-tier architecture) [67] [68] με την οποία διαχωρίζονται τα αντικείμενα σε στρώματα [69] ανάλογα με το αν σχετίζονται με τα δεδομένα (data), την λογική (logic) και την παρουσίαση τους (presentation).



4-2: Η αντικειμενοστραφής σχεδίαση τριών επιπέδων (Bartledan/Wikipedia Commons)

- Το Επίπεδο Παρουσίασης (Presentation Layer) περιλαμβάνει την διασύνδεση του χρήστη (UI) με την επιχειρησιακή λογική⁴⁵ (business logic) δηλαδή τις εσωτερικές λειτουργίες του συστήματος.
- Το Επίπεδο Λογικής (Logic Layer, Business Logic Layer ή Business Layer) υλοποιεί την καρδιά της λειτουργικότητας του συστήματος, περικλείοντας την λογική των ΤΝΔ, καθαρή και χωρίς εξαρτήσεις από το UI.
- Το Επίπεδο Δεδομένων παρέχει πρόσβαση στα δεδομένα που είναι αποθηκευμένα μέσα στα όρια του συστήματος και ίσως βρίσκονται φυσικά τοποθετημένα σε άλλους δικτυακούς υπολογιστές. Τα αντικείμενα σε αυτό το στρώμα έχουν μειωμένη σύζευξη με τα αντικείμενα του επιπέδου λογικής ως προς την πρόσβαση και την ανάκτηση των δεδομένων.

Αυτά τα στρώματα υπάρχουν σε ολόκληρο τον σχεδιασμό του αντικειμενοστραφούς νευρωνικού δικτύου και εκτός από την λογική της ανάκλησης του δικτύου ως Επίπεδο Λογικής, και τα δεδομένα του προτύπου ως Επίπεδο Δεδομένων, θα δούμε γνωρίσουμε παρακάτω κλάσεις που

⁴⁵ Ο όρος business logic χρησιμοποιείται κυρίως για να περιγράψει προγραμματιστικές διαδικασίες που σχετίζονται με βάσεις δεδομένων.

ανήκουν στο Επίπεδο Παρουσίασης. Αυτές οπτικοποιούν την εσωτερική λειτουργία του δικτύου εμφανίζοντας της στον χρήστη με έναν εύκολο και κατανοητό τρόπο.

Τοποθετώντας το `TGNeuralTrainingSet` στο επίπεδο δεδομένων της εφαρμογής `Horasis`, εμπλουτίζεται πλέον με μεθόδους για αποθήκευση και ανάκτηση των δεδομένων που αυτό κρατάει. Τέτοιες είναι η `SaveToFile()` και `LoadFromFile()` που αποθηκεύουν και φορτώνουν τα πρότυπα στον δίσκο και σε αρχεία της μορφής XML [70][71]. Πολύ εύκολα υλοποιούνται και οι `SaveToStream()` και `LoadFromStream()` για την αποθήκευση των προτύπων σε δυαδικές ροές (binary streams). Μέσω των ροών τα δεδομένα μπορούν να αποθηκευτούν σε πεδία βάσεων δεδομένων τύπου BLOB⁴⁶ ή να μεταφερθούν μέσω συνδέσεων TCP sockets, χρησιμοποιώντας για παράδειγμα σε μια υλοποίηση σε Delphi με την κλάση `TWinSocketStream` [72].

Επίσης τα μοτίβα `TGNeuralPattern` που περιέχονται σε ένα `TGNeuralTrainingSet` μπορούν να διατρεχθούν σειριακά με τις μεθόδους `First()`, `Next()` και τις ιδιότητες `Current` που έχει το τρέχον μοτίβο. Υπάρχει ακόμα η `EndOf` που επιστρέφει αν έχουμε φτάσει στο «τέλος αρχείου» των προτύπων. Όλες οι παραπάνω σχεδιάστηκαν και ονοματοδοτήθηκαν κατ' αντιστοιχία με το `TDataSet` της Delphi [3] που και αυτό λειτουργεί ως ένα σειριακό αρχείο, έχοντας κατά νου ένα υπερβολικά μεγάλο σύνολο προτύπων που δεν μπορεί να φορτωθεί ολόκληρο στην μνήμη, ή βρίσκεται σε κάποια δικτυακή συσκευή και πρέπει να μεταφερθεί τμηματικά από αυτήν.

⁴⁶ **BLOB** = Binary Large Object

Η Μορφή του Δικτύου

Ο γονότυπος του αντικειμενοστραφούς νευρωνικού δικτύου

Στους βιολογικούς οργανισμούς υπάρχει ένα ολοκληρωμένο σύνολο από οδηγίες για την κατασκευή τους, το γονιδίωμα (genome) το οποίο είναι κωδικοποιημένο με χημικά μόρια DNA ή RNA [73]. Αντίστοιχα θα μπορούσαμε να πούμε ότι σε ένα τεχνητό μοντέλο τον ρόλο του γονότυπου, δηλαδή της γενετικής πληροφορίας του ζωντανού οργανισμού, παίζει μια σειρά από δυαδικά ψηφία [74]. Αυτή η σειρά στην δική μας περίπτωση, προκύπτει από την αποθήκευση πολλών αριθμητικών δεδομένων που αφορούν την μορφή και τις ιδιότητες του ΤΝΔ. Για αυτό δημιουργήθηκε ένα αντικείμενο-γονότυπος το `TGNeuralNetworkConfig` που περιλαμβάνει κωδικοποιημένες όλες τις ρυθμίσεις που μπορούν να οδηγήσουν στην δημιουργία και την εκπαίδευση ενός αντικειμενοστραφούς τεχνητού νευρωνικού δικτύου.

Τα συστατικά της κλάσης είναι μια δομή `TGNNConfigHeaderRec` που περιλαμβάνει τις βασικότερες πληροφορίες όπως το είδος του δικτύου, δηλαδή MLP και το πλήθος των νευρώνων. Επίσης υπάρχει μια δομή `TGNNConfigTrainingRec` που κρατάει τις οδηγίες για την εκπαίδευση του δικτύου όπως το κατώφλι σφάλματος `ErrorTolerance` με το οποίο σταματάει ο `backpropagation`, τον σπόρο της γεννήτρια ψευδοτυχαίων αριθμών `RandomSeed` που τίθεται πριν την χρήση της κατά την αρχικοποίηση των αντικειμένων-βαρών `TGSynapticWeight` και την πληροφορία χρονισμού για την εκπαίδευση `ErrorCheckPeriodMsecs`. Η τελευταία, που όπως θα δούμε θα παίζει ρόλο στην επιτάχυνση της εκτέλεσης της εκπαίδευσης, καθορίζει τα `milliseconds` που πρέπει να έχουν περάσει από τον τελευταίο υπολογισμό του μέσου τετραγωνικού σφάλματος για να το επαναυπολογίσει. Για κάθε αντικείμενο-στρώμα του δικτύου υπάρχει αντίστοιχα ένα συστατικό `TGNNConfigLayerRec`, που κρατάει πληροφορίες για την θέση του στρώματος στο δίκτυο, το πλήθος νευρώνων και τις συναπτικές ρυθμίσεις του στρώματος

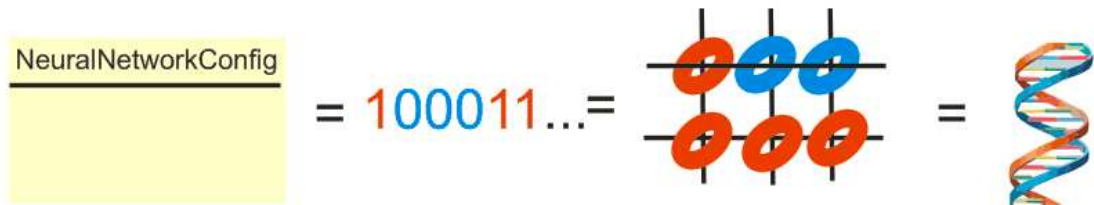
Δυνατότητες αποθήκευσης και επαναφοράς μια μορφής δικτύου

Από ένα κατασκευασμένο αντικειμενοστραφές νευρωνικό δίκτυο γραμμένο με *σκληροκώδικα* (`hard coded`⁴⁷) μπορεί να συγκεντρωθούν και να εγγραφούν όλες οι ρυθμίσεις σε ένα αντικείμενο `TGNeuralNetworkConfig`. Το πιο ενδιαφέρον όμως και αυτό που μιμείται την βιολογία είναι ότι ένα αντικείμενο-γονότυπου εφόσον έχει πάνω του φορτωμένες αυτές τις πληροφορίες μπορεί να κατασκευάσει ένα ολοκληρωμένο αντικειμενοστραφές νευρωνικό δίκτυο στην μνήμη.

Έτσι υπάρχει στην `TGNeuralNetworkConfig` η συναθροιστική (`aggregate`) ιδιότητα `NeuralNet` με την οποία γίνεται αναφορά σε ένα αντικείμενο-ΤΝΔ. Όταν αυτήν είναι κενή και κληθεί μια από τις μεθόδους `LoadFromFile()` ή `LoadFromStream()` δημιουργείται και επιστρέφεται ένα νέο στιγμιότυπο (`instance`) της κλάσης `TGNeuralNetwork`, με πλήρως δημιουργημένα υποσυστατικά αντικείμενα νευρώνων, στρωμάτων, δενδριτών, συνάψεων κ.α., έτοιμο προς αρχικοποίηση. Με την `LoadFromFile()` ως γενική τυποποίηση έχουμε φόρτωση δεδομένων από αρχείο δίσκου, ενώ με την `LoadFromStream()` φόρτωση από δυαδική ροή που χρησιμοποιείται συνήθως για φόρτωση δεδομένων που υπάρχουν σε ένα πεδίο `BLOB`. Αντίστοιχα υπάρχουν οι `SaveToFile()` και `SaveToStream()` για την αποθήκευση των γενετικών πληροφοριών του ΤΝΔ. Όσον αφορά την

⁴⁷ `Hard coded`, χαρακτηρίζεται ένας πηγαίος κώδικας στον οποία τα δεδομένα και οι ρυθμίσεις έχουν ενσωματωθεί σε αυτόν, αντί να διαβάζονται από εξωτερικές πηγές ή εναλλακτικά να εισάγονται ως είσοδος στο πρόγραμμα.

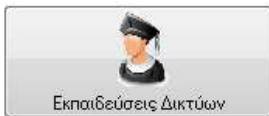
διαχείριση δεδομένων μέσω δυαδικών ροών λόγω των αναγκών της κλάσης TStream της Delphi χρησιμοποιήθηκαν records αντί για αντικείμενα, επειδή η ανάγνωση και εγγραφή σε αυτές γίνεται μέσω απομονωτών (buffers). Τα διαφορετικά δεδομένα σε μια δομή είναι στην ουσία μια συνεχόμενη σειρά από bytes, μεγέθους ίσου με το άθροισμα των μεγεθών των επιμέρους πεδίων της δομής. Στην Delphi αυτό το μέγεθος υπολογίζεται αυτόματα με την χρήση της SizeOf()[3]. Επιστρέφοντας ξανά στην αρχή της ενότητας έχουμε έτσι πάλι μπροστά μας την σειρά από δυαδικά ψηφία που αναφέραμε, και η οποία κωδικοποιεί τον γονότυπο του αντικειμενοστραφούς τεχνητού νευρωνικού δικτύου.



4-3: Αντικείμενο NeuralNetworkConfig, ο γονότυπος του νευρωνικού δικτύου ©2011 Horasis

5. Η Εκπαίδευση του Δικτύου

Η περιγραφή του πλαισίου λογισμικού (framework) για το αντικειμενοστραφές τεχνητό νευρωνικό δικτύου θα συνεχιστεί πλέον μέσα από την τελική του χρηστικότητα και ενσωμάτωση του στο περιβάλλον της εφαρμογής Horasis. Έτσι θα δούμε την εκπαίδευση που υλοποιείται στο framework, μέσα από την φόρμα οθόνης που έχει δημιουργηθεί για τον χειρισμό της με τον τίτλο «Εκπαιδεύσεις Δικτύων». Η φόρμα λειτουργεί κοινόχρηστα για την εκπαίδευση όλων των δικτύων της εφαρμογής. Ανάλογα με τον επιλεγμένο τύπο ΤΝΔ μέσα από όλους που υπάρχουν στο ΤΝΔ Horasis δημιουργείται και αρχικοποιείται ένα αντικείμενο TGNeuralNetwork προς εκπαίδευση. Όπως ήδη αναφέρθηκε ο σχεδιασμός του Horasis στηρίζεται στην αρχιτεκτονική

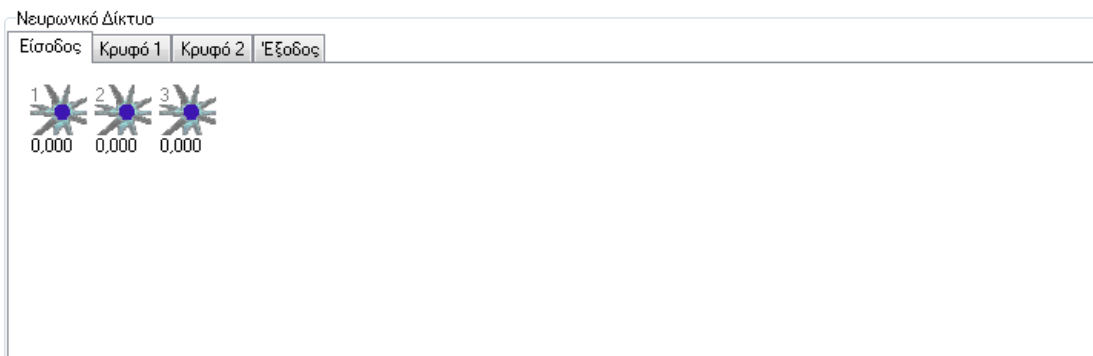


5-1: Πλήκτρο εντολής «Εκπαιδεύσεις Δικτύων» της εφαρμογής Horasis

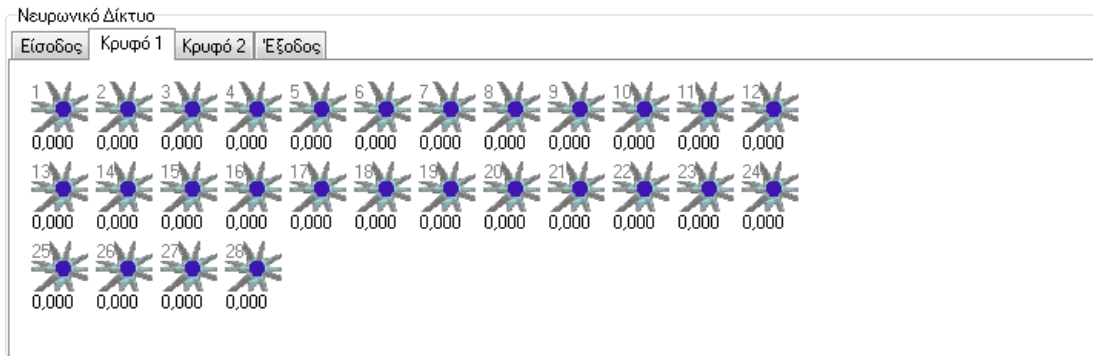
τριών επιπέδων και σε αυτό το κεφάλαιο θα δούμε για πρώτη φορά υλοποιήσεις που αφορούν το επίπεδο παρουσίασης (presentation layer). Επίσης θα συμπληρωθεί η εικόνα του αντικειμενοστραφούς νευρωνικού δικτύου μέσα από την έννοια της τρέχουσας κατάστασης (state) στην οποία βρίσκεται μετά από κάποια εποχή εκπαίδευσης.

Η Οπτικοποίηση της μορφής του δικτύου

Το δίκτυο μας, που όπως αναφέραμε στο κεφάλαιο 3, αποτελείται από 4 στρώματα οπτικοποιείται ως ένα οπτικό συστατικό σελίδων (tabs), TPageControl της Delphi, με μια σελίδα για κάθε ένα από τα στρώματα. Μέσα στην σελίδα δημιουργούνται οπτικά συστατικά με πλήθος αντίστοιχο με το πλήθος των νευρώνων του στρώματος. Το κάθε ένα που ανήκει στην κλάση TGVisualNeuron εμφανίζει ένα εικονίδιο νευρώνα συνοδευόμενο από αριθμητικές πληροφορίες και δημιουργείται ως συστατικό της φόρμας δηλαδή θα καταστραφεί μαζί της.



5-2: Σελίδα (Tab) που εμφανίζει τους νευρώνες του στρώματος εισόδου

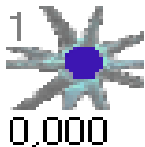


5-3: Σελίδα (Tab) που εμφανίζει τους νευρώνες του κρυφού στρώματος 1.

Μια άλλη κλάση `TGVisualNeuralLayer` η οποία υλοποιεί το πρότυπο σχεδίασης `Factory Pattern` [75][76], αναλαμβάνει την δημιουργία ενός πλήθους αντικειμένων `TGVisualNeuron` που αντιστοιχεί στο πλήθος των τεχνητών νευρώνων στο αντίστοιχο στρώμα. Στην `TGVisualNeuralLayer` υπάρχουν οι ιδιότητες:

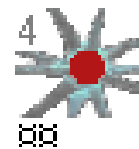
- **Layer: TGNeuralLayer.** Το αντικείμενο-στρώμα του νευρωνικού δικτύου που θα εμφανιστεί.
- **Parent: TWinControl.** Ένα οποιοδήποτε οπτικό συστατικό που έχει δείκτη χειρισμού παραθύρου windows (handle) [36], μέσα στο οποίο θα εμφανιστούν τα οπτικά αντικείμενα.
- **StartLeft, StartTop, NeuronsPerLine** που ρυθμίζουν την αρχική πάνω δεξιά γωνία της ανάπτυξης και το πλήθος οπτικών νευρώνων που αναδιπλώνονται ανά γραμμή

Έτσι μπορούμε να μετακινήσουμε όλη την οπτικοποίηση του στρώματος, σε οποιαδήποτε διάταξη φόρμας και αν σχεδιαστεί αλλάζοντας μόνο την ιδιότητα `Parent`. Το κάθε οπτικό αντικείμενο `TGVisualNeuron` αντιπροσωπεύει στο επίπεδο παρουσίασης τον τεχνητό νευρώνα που υπάρχει στο επίπεδο λογικής. Εμφανίζει πάνω δεξιά τον αριθμό του νευρώνα μέσα στο επίπεδο, δηλαδή το `TGNeuron.Position` και κάτω τον συντελεστή σφάλματος δ_i . Η τιμή που εμφανίζεται είναι αυτή που έχει υπολογιστεί για τον νευρώνα στην τρέχουσα εποχή και θα την δούμε να αλλάζει, κατά την διάρκεια της εκπαίδευσης με `back propagation`. Αν η τιμή του σφάλματος είναι μεγαλύτερη από το 10 τότε εμφανίζεται στην θέση της το σήμα του άπειρου ∞ . Όταν η τιμή είναι μικρότερη του 10 εμφανίζεται ένα



5-5:
`TGVisualNeuron`
με $\delta_i = 0$

δεκαδικό ψηφίο μέχρι το 2.2 και από εκεί και κάτω τρία δεκαδικά ψηφία. Παρατηρούμε επίσης τον διαφορετικό χρωματισμό τους ο οποίος έχει μια ενδιαφέρουσα σημαντική κατά την διάρκεια της εκπαίδευσης του δικτύου και θα ασχοληθούμε μαζί του αργότερα. Τα οπτικά συστατικά για τους νευρώνες του στρώματος εισόδου και εξόδου μπορεί να εμφανίσουν και ένα tooltip⁴⁸ με περιγραφή της σημαντικής τους στο ΤΝΔ. Έτσι αν σε ένα σύνολο προτύπων `TGNeuralTrainingSet` το οποίο τροφοδοτείται στο ΤΝΔ έχουν οριστεί λεζάντες στην ιδιότητα `InputDescription`, τότε όταν αφήνουμε τον δείκτη του ποντικιού πάνω σε έναν οπτικό νευρώνα εισόδου, θα εμφανιστεί η περιγραφή που μας λέει τι είναι η αριθμητική τιμή που εισάγεται ως σήμα σε αυτόν.



5-4:
`TGVisualNeuron` με
 $\delta_i > 10$

⁴⁸ Είναι το μικρό κίτρινο παραθυράκι με βοηθητικό κείμενο, που εμφανίζεται αν ο δείκτης του ποντικιού αφεθεί πολύ ώρα πάνω σε ένα οπτικό συστατικό

Διαδικασία της Παράλληλης Εκπαίδευσης

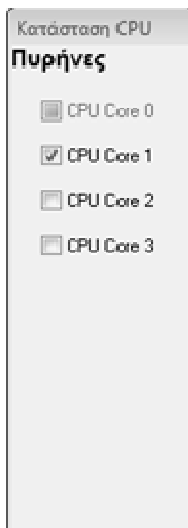
Αφήνοντας για λίγο το επίπεδο παρουσίας κατεβαίνουν προς τα κάτω στο επίπεδο λογικής όπου και θα δούμε τι γίνεται εσωτερικά από την στιγμή που ο χρήστης πατάει το πλήκτρο «Εκπαίδευση» μέσα στην φόρμα «Εκπαιδεύσεις Δικτύων». Η πρώτη και σχεδόν αυτονόητη παρατήρηση, είναι ότι δεν εκτελείται κάποιος βρόχος μέσα από αυτό, «παγώνοντας» τον οπτικό χειρισμό της εφαρμογής για κάποιες προκαθορισμένες εποχές εκτέλεσης. Αυτό που γίνεται είναι η δημιουργία ενός νήματος (thread) των Windows με την βοήθεια μιας απόγονης κλάσης του νήματος TThread [4]. Μετά την δημιουργία ενεργοποιείται το thread ξεκινώντας μια παράλληλη διαδικασία εκπαίδευσης, η οποία μπορεί να σταματήσει προσωρινά με το πλήκτρο «Παύση» και να συνεχίσει όταν ο χρήστης πατήσει ξανά το «Εκπαίδευση».

Μέγιστη ταχύτητα με πραγματικά παράλληλη λειτουργία σε πυρήνα επεξεργαστή.

Με την κλάση TGNeuralNetworkThread η οποία κληρονομεί από την TThread και στην μέθοδο DoThreadLoop() υλοποιείται η επαναληπτική διαδικασία του back propagation και ο βρόχος που γνωρίσαμε με τον ψευδοκώδικα του αλγορίθμου:

`ενώσω δεν Πρέπει-να-τερματιστεί επανέλαβε`

Δύο βασικές μέθοδοι καθορίζουν την ενεργοποίηση και την απενεργοποίηση της επαναληπτικής διαδικασίας. Με την StartTraining() ξεκινάει η διαδικασία εκτέλεσης του νήματος (thread function) [36] η οποία υλοποιείται μέσα σε αντικαταστάτρια (override) μέθοδο της TThread.Execute. Μέσα σε αυτήν καλείται η DoThreadLoop(). Εφόσον το νήμα είχε γίνει ανενεργό (suspend) καλώντας πάλι την ίδια μέθοδο συνεχίζεται ξανά η εκτέλεση του (resume). Με την δεύτερη μέθοδο PauseTraining(), σηματοδοτείται η απενεργοποίηση της εκτέλεσης του νήματος, και η οποία θα συμβεί αφού τελειώσει η τρέχουσα επανάληψη του βρόχου εκπαίδευσης.



5-6: Φόρμα Κατάσταση CPU

Η αντικαταστημένη TGNeuralNetworkThread.Execute() τρέχει όσο το δυνατόν γρηγορότερα μπορεί να επιτευχθεί στο λειτουργικό Windows, επειδή εκτός του ότι είναι εκτέλεση σε νήμα δεν έχει υλοποιηθεί καμία αναμονή μέσα σε αυτή, και έτσι όταν τρέχει καταναλώνει έτσι 100% της υπολογιστικής ισχύος ενός πυρήνα του επεξεργαστή. Γι αυτό το TGNeuralNetworkThread μετά την δημιουργία του φορτώνεται στον πρώτο διαθέσιμο για εκπαίδευση πυρήνα του επεξεργαστή. Ο πρώτος πυρήνας (Processor Affinity 0) δεσμεύτηκε για την αποκλειστική λειτουργία του UI της εφαρμογής και την σταθερότητα του λειτουργικού συστήματος και οι άλλοι 3 ήταν διαθέσιμοι για εκπαίδευση. Αυτό οπτικοποιήθηκε με την διπλανή φόρμα «Κατάσταση CPU» που όταν λειτουργεί ένα νήμα εκπαίδευσης σημειώνεται ο αντίστοιχος πυρήνας που είναι δεσμευμένος. Η διαχείριση αυτή γίνεται αυτόματα, δηλαδή αν ανοίξουν πολλαπλά στιγμιότυπα (instances) της φόρμας «Εκπαιδεύσεις Δικτύων» δεσμεύεται ο αμέσως επόμενος διαθέσιμος πυρήνας την στιγμή της εκκίνησης του back propagation.

Η DoThreadLoop() θα μπορούσε να μην έχει απολύτως καμία σχέση με το UI, σε μια εκπαίδευση που θα έτρεχε στο περιθώριο (background task) και δεν θα έδινε καμία ένδειξη στον χρήστη, όπως μέσα σε μια υπηρεσία (service) των Windows. Επειδή όμως η οπτική πληροφόρηση προς τον χρήστη που θα γνωρίσουμε στην επόμενη ενότητα έχει μεγάλη σημασία, τρέχει μια

ενημέρωση του UI ανά 1,5 δευτερόλεπτο ή και παραπάνω, προσθέτοντας μια μικρή καθυστέρηση. Καθώς η εκπαίδευση δεν καταναλώνει το πρώτο πυρήνα βλέπουμε ομαλή λειτουργία του UI της εφαρμογής Horasis, και παράλληλα υπάρχει η δυνατότητα να εκτελέσουμε άλλες εργασίες με τον υπολογιστή όσο περιμένουμε τον εξέλιξη της εκπαίδευσης ή τον τερματισμό της.

Το πιο σημαντικό όφελος από όλα που προκύπτει από αυτόν τον σχεδιασμό, είναι η δυνατότητα να τρέχουν πολλαπλές εκπαιδύσεις παράλληλα, μια σε κάθε πυρήνα. Δίνεται η δυνατότητα δια της δοκιμής και απόρριψης να κρατηθεί πιο εύρωστη εκπαίδευση, μέσω οπτικού ελέγχου της διαδικασίας εκπαίδευσης. Μέσα από την εμπειρία της εφαρμογής του Horasis, διαπιστώθηκαν τα σημαντικά οφέλη αφενός της παράλληλης εκπαίδευσης και αφετέρου του οπτικού ελέγχου. Κερδήθηκε σημαντικός χρόνος στην ανάπτυξη του Horasis, καθώς έτρεχαν τρεις εκπαιδύσεις παράλληλα και γινόταν ένας οπτικός έλεγχος για την σύγκλιση τους. Κάθε φορά η κρατιόταν η καλύτερη και οι άλλες δύο αντικαθίσταντο από άλλες ξανατρέχοντας την εκπαίδευση. Έτσι στον χρόνο που θα εκπαιδευόταν 2 ΤΝΔ, εκπαιδευόταν 5 δηλαδή μια αύξηση 250% στην ταχύτητα εκπαίδευσης των 5 δικτύων σε σχέση με τις κλασσικές υλοποιήσεις.

Το αντικείμενο κατάστασης του αντικειμενοστραφούς ΤΝΔ

Πριν αναφερθούμε αναλυτικότερα στην εκπαίδευση, πρέπει να αναφέρουμε ότι μέσα στο αντικείμενο παράλληλης εκπαίδευσης `TGNeuralNetworkThread` δημιουργείται κατά την έναρξη της ένα αντικείμενο της κλάσης `TGNeuralNetworkState`, η οποία ανήκει στο επίπεδο δεδομένων (data layer). Αυτό αντιπροσωπεύει την τρέχουσα κατάσταση της εκπαίδευσης ενός `TGNeuralNetwork`. Έχει τον ρόλο του περιέκτη που κρατάει τις τρέχουσες τιμές μεταβλητών της εκπαίδευσης και αλλά και της γέφυρα για την μεταφορά των αριθμητικών τιμών των συναπτικών βαρών σε κάποιο μέσο αποθήκευσης. Οπότε και εδώ έχουμε τις πλέον γνωστές `LoadFromFile()`, `SaveToFile()` και `LoadFromStream`, `SaveToStream()` που λειτουργούν με την ίδια λογική των αντιστοίχων που γνωρίσαμε στο κεφάλαιο 4.

Έτσι μπορούμε να σταματήσουμε την εκπαίδευση με το «Παύση», να αποθηκεύσουμε την κατάσταση σε ένα αρχείο στον δίσκο και να κλείσουμε την εφαρμογή. Σε κάποια άλλη χρονική στιγμή μπορούμε να φορτώσουμε την αποθηκευμένη κατάσταση εκπαίδευσης, παίρνοντας ακριβώς τα ίδια αντικείμενα και τις τιμές που υπήρχαν κατά την στιγμή που αποθηκεύτηκε. Έτσι η εκπαίδευση μπορεί να συνεχίσει σαν να μην έκλεισε ποτέ η εφαρμογή. Τέλος η σειρά με την οποία επανέρχεται ένα αντικειμενοστραφές ΤΝΔ στην ίδια κατάσταση είναι:

Φόρτωση Μορφής / Δημιουργία ΤΝΔ ➤ Αρχικοποίηση ΤΝΔ ➤ Φόρτωση Κατάστασης

Επιπρόσθετα μέσα στο αντικείμενο κατάστασης υπάρχει μια ιδιότητα συστατικού αντικειμένου `ErrorGraph` στο οποίο αποθηκεύονται οι διαφορετικές τιμές που παίρνει το μέσο τετραγωνικό σφάλμα σε κάποιες εποχές ώστε να μπορεί να παραχθεί ένα γράφημα. Θα δούμε αναλυτικότερα πως αυτό το γράφημα θα μας βοηθήσει στην λήψη αποφάσεων κατά την εκπαίδευση του ΤΝΔ.

Κάτω από το καπώ και μέσα στο νήμα εκπαίδευσης *back propagation*

Ανοίγοντας την η διαδικασία εκτέλεσης του νήματος Execute() βρίσκουμε μέσα την κλήση τριών εικονικών (virtual) μεθόδων DoCreateThread, DoThreadLoop και DoDestroyThread. Αρχικά καλείται η DoCreateThread() στην οποία δημιουργείται το αντικείμενο κατάστασης NeuralNetState: TGNeuralNetworkState και επίσης καλείται ένα συμβάν (event) μέσα στο οποίο πρέπει να ανατεθεί ένα σύνολο προτύπων εκπαίδευσης TGNeuralTrainingSet. Πρακτικά λόγω της υλοποίησης της DoThreadLoop δεν θα φτάσει ποτέ να κληθεί η DoDestroyThread(). Το NeuralNetState ενώ δημιουργείται ασύγχρονα σε σχέση με την κλήση της, θα καταστραφεί ντετερμινιστικά⁴⁹ όταν θα καταστραφεί και το αντικείμενο TGNeuralNetworkThread.

Μέσα στην DoThreadLoop() υπάρχει ένας βρόχος που επαναλαμβάνεται ενώσω το νήμα είναι ενεργό, ενώ σταματάει όταν αυτό απενεργοποιείται (suspend). Τα πιο βασικά σημεία του πηγαίου κώδικα εμφανίζονται στην επόμενη σελίδα και μέσα σε αυτό παρατηρούμε ότι την επανάληψη καθορίζει η μεταβλητή αληθείας Stopped. Όταν σηματοδοτηθεί για οποιονδήποτε λόγο Stopped := True τότε το νήμα απενεργοποιείται με εκτέλεση της Suspend(). Οι λόγοι είναι:

- **(FNeuralNet.MeanSquareError < FErrorTolerance):** Κριτήριο τερματισμού 1 του back propagation.
- **(FNeuralNet.OutputNeuronsErrorMean < FMinOutputNeuronsErrorMean):** Ένα νέο κριτήριο τερματισμού που εισάγει η εργασία Horasis και σχετίζεται με τον παράγοντα σφάλματος δ_i .
- **(FNeuralNet.Epoch > FMaxEpochs):** αν έχουν περάσει παραπάνω εποχές εκπαίδευσης από ένα όριο MaxEpochs που έχουμε ορίσει στις ρυθμίσεις εκπαίδευση του ΤΝΔ
- **FMustPauseTraining:** Ο χρήστης θέτει αυτήν την σημαία (flag) σε True ασύγχρονα όταν πατάει το πλήκτρο «Παύση».

⁴⁹ Ντετερμινιστική καταστροφή αντικειμένων είναι όταν η στιγμή της καταστροφής μπορεί να προβλεφθεί με ακρίβεια και θα είναι πάντα η ίδια όσες φορές και αν επαναληφθεί η διαδικασία.

```
//Απόσπασμα από πηγαίο κώδικα της DoThreadLoop()

FNeuralNet.FMeanSquareError := FNeuralNet.TrainingSetRMSE;
Stopped := False;
while not Stopped do
begin
...

FNeuralNet.TrainPatternSet;

...

if (FNeuralNet.Epoch mod FCheckPeriod = 0) then
begin
FNeuralNet.FMeanSquareError := FNeuralNet.TrainingSetMSE;
...
if FUserInterfaceEnabled then
Synchronize(CallUserInterfaceUpdate);
end;

...

Stopped := (FNeuralNet.MeanSquareError < FErrorTolerance)
or (FNeuralNet.OutputNeuronsErrorMean <
FMinOutputNeuronsErrorMean)
or (FNeuralNet.Epoch > FMaxEpochs)
or FMustPauseTraining;
if Stopped then
begin
FMustPauseTraining := False;
Suspend;
Stopped := False;
end;
end;
```

Πλέον το ενδιαφέρον μας στρέφεται προς τις μεθόδους TrainPatternSet() και TrainingSetMSE() του αντικειμένου FNeuralNet: TGNeuralNetwork που εμφανίζονται μέσα στην υλοποίηση του αλγορίθμου.

Η πρώτη μέθοδος έχει υλοποιηθεί με τον παρακάτω πηγαίο κώδικα Delphi καλώντας την ήδη γνωστή μέθοδος ανάκλησης Stimulate() ακολουθούμενη από την TrainPattern() για κάθε πρότυπο του συνόλου. Στο τέλος αυξάνεται ο δείκτης της τρέχουσας εποχής εκπαίδευσης.

```
{-----}

procedure TGNeuralNetwork.TrainPatternSet;
begin
  { Για κάθε πρότυπο του συνόλου εκπαίδευσης }
  FTrainingSet.First;
  while not FTrainingSet.EndOf do
  begin
    Stimulate;
    TrainPattern;
    FTrainingSet.Next;
  end;
  Inc(FEpoch);
end;
{-----}
```

Εξίσου απλή στην λειτουργία της και με υψηλή συνοχή είναι η `TrainPattern()`, καθώς ο υπολογισμός του σφάλματος γίνεται μέσα στο αντικείμενο-σώμα με την κλήση της `CalcError()`. Επίσης η διάδοση του σφάλματος προς τα πίσω γίνεται με την κλήση της `BackPropagateError()` που όπως είδαμε στο κεφάλαιο 4, ενημερώνει τα συναπτικά βάρη. Πρώτα υπολογίζονται όλοι οι συντελεστές δ_i στα αντικείμενα `TGSoma` στον πρώτο βρόχο `for`, ο οποίος τρέχει αντίστροφη διαδρομή των αντικειμένων-στρωμάτων του δικτύου. Στον δεύτερο βρόχο `for` χρησιμοποιούνται οι συντελεστές για την ενημέρωση των συναπτικών βαρών του δικτύου.

```

{-----}
procedure TGNeuralNetwork.TrainPattern;
var
  L, N      : Integer;
  CurLayer  : TGNeuralLayer; { object reference }
begin
  for L := FLayers.Count - 1 downto 0 do
  begin
    CurLayer := FLayers[L];
    for N := 0 to CurLayer.Neurons.Count - 1 do
      CurLayer.Neurons[N].Soma.CalcError;
    end;

    for L := FLayers.Count - 1 downto 0 do
    begin
      CurLayer := FLayers[L];
      for N := 0 to CurLayer.Neurons.Count - 1 do
        CurLayer.Neurons[N].Soma.BackPropagateError;
      end;
    end;
  end;
{-----}

```

Όπως γνωρίσαμε επίσης στο προηγούμενο κεφάλαιο η Stimulate() καλεί πολλές Process() μια για τον κάθε νευρώνα του δικτύου. Πριν να δούμε σε βάθος την CalcError(), παρατίθεται ο πηγαίος κώδικας της Process() το συναπτικό άθροισμα που παράγεται εδώ, ανατίθεται σε ένα ιδιωτικό (private) πεδίο της κλάσης FInputSum για μελλοντική χρήση από την CalcError()

```

{-----}
procedure TGSoma.Process;
var
    U          : Double;
    I          : Integer;
begin
    { U = S(wi * xi) }
    U := 0;
    for I := 0 to FNeuronDendrites.Count - 1 do
        U := U + FNeuronDendrites[I].Synapses[0].Value;

    FInputSum := U;
    FNeuron.FAxxon.OutSignal := EnergizationFunction(U);
end;
{-----}

```

Παρακάτω βλέπουμε αναλυτικά τον πηγαίο κώδικα του υπολογισμού του δ_i ενός νευρώνα που είναι εμφωλευμένος στο αντικείμενο-σώμα του. Το αποθηκευμένο FInputSum κατά την ανάκληση περνιέται ως παράμετρος στην παράγωγο της συνάρτησης ενεργοποίησης του νευρώνα και τοποθετείται σε προσωρινή μεταβλητή.

```

{-----}
procedure TGSoma.CalcError;
var
    I                : Integer;
    SumWeightedError : Double;
    Synapse          : TGSynapse; { object reference }
    EnergizationDerivFX : Double;
begin
    EnergizationDerivFX := EnergizationDerivative(FInputSum);

    { Για νευρώνες που ανήκουν στο τελευταίο στρώμα, υπολογίζεται το σφάλμα με
      την παράγωγο της συνάρτησης ενεργοποίησης }
    if FNeuron.FIsOutputNeuron then
        begin
            { E = f'(u) * ( t - f(u) ) }
            FError := EnergizationDerivFX *
                (FNeuron.FAxxon.TargetOutSignal - FNeuron.FAxxon.OutSignal);
        end
    else
        begin
            //Για νευρώνες που ανήκουν σε ενδιάμεσα στρώματα
            SumWeightedError := 0;

            { Για κάθε εξερχόμενη σύναψη από τον άξονα προς δενδρίτες άλλων
              νευρώνων, αθροίζονται τα γινόμενα λάθους νευρώνα προορισμού
              επί του συναπτικού βάρους σύναψης }
            for I := 0 to FNeuron.FAxxon.Synapses.Count - 1 do
                begin
                    Synapse := TGSynapse(FNeuron.FAxxon.Synapses[I]);
                    SumWeightedError := SumWeightedError
                        + Synapse.Weight.Value * Synapse.DestFiber.Soma.Error;
                end;

            FError := EnergizationDerivFX * SumWeightedError;
        end;
    end;
{-----}

```


Επίσης παρατίθεται και η ενημέρωση των συναπτικών βαρών σε ένα νευρώνα:

```

{-----}
procedure TGSoma.BackPropagateError;

begin
  for I := 0 to FNeuron.Dendrites.Count - 1 do
    FNeuron.Dendrites[I].Synapses[0].Weight.BackPropagate(FNeuron.Soma.Error);
end;
{-----}

```

Επιστρέφοντας στην DoThreadLoop() και στον βρόχο της εκπαίδευσης που εκτελείται από το νήμα βλέπουμε την συνάρτηση TrainingSetMSE() που επιστρέφει και αποθηκεύει σε πεδίο του αντικείμενου-TNΔ το μέσο τετραγωνικό σφάλμα. Αυτή είναι ανάμεσα στις μεταβλητές της τρέχουσας κατάστασης του δικτύου, που θα αναλάβει να αποθηκεύει το αντικείμενο κατάστασης TGNeuralNetworkState. Επειδή η TrainingSetMSE() κοστίζει σε χρόνο, καθώς πρέπει να γίνει ανάκληση ολόκληρου του συνόλου των προτύπων, βλέπουμε ότι δεν καλείται σε κάθε εποχή αλλά ανά FCheckPeriod εποχές. Η περίοδος αυτή που είναι ένας ακέραιος αριθμός εποχών υπολογίζεται από την ρύθμιση ErrorCheckPeriodMSEcs που βρίσκεται στην μορφή του δικτύου. Γίνεται μια χρονομέτρηση κατά την οποία μετρώνται πόσες εποχές έχουν περάσει κατά το χρονικό διάστημα ErrorCheckPeriodMSEcs, στρογγυλοποιούμενες στον κοντινότερο ακέραιο. Αν έχουν περάσει περισσότερα από ErrorCheckPeriodMSEcs για την εκτέλεση της πρώτης εποχής τότε το FCheckPeriod γίνεται ίσο με την μονάδα εκτελώντας υπολογισμό του μέσου τετραγωνικού σφάλματος σε κάθε εποχή.

Παράλληλα έχοντας την ίδια περιοδικότητα γίνεται συγχρονισμένη κλήση ⁵⁰ της μεθόδου CallUserInterfaceUpdate(), μέσω της οποίας ενημερώνεται το UI. Η συνηθισμένη τιμή της ErrorCheckPeriodMSEcs για όλες τις εκπαιδεύσεις που εκτελέστηκαν στο Horasis ήταν τα 1500ms, οπότε ανά 1,5 δευτερόλεπτο ανανεωνόταν η φόρμα «Εκπαιδεύσεις Δικτύων».

⁵⁰ Η συγχρονισμένη κλήση κρατάει την εκτέλεση του νήματος μέχρι να τελειώσει η ενημέρωση του UI, προκαλώντας καθυστέρηση.

Το νέο κριτήριο τερματισμού του *back propagation*

Σε κάθε αντικείμενο TGSoma υπάρχει μια ιδιότητα ErrorMean η οποία αντιπροσωπεύει τον μέσο όρο των συντελεστών σφάλματος δ_i που παράγεται από την ανάκληση όλων των προτύπων ενός συνόλου εκπαίδευσης. Για ολόκληρο το αντικειμενοστραφές ΤΝΔ, υπάρχει μια ιδιότητα OutputNeuronsErrorMean που είναι ο μέσος όρος όλων των ErrorMean των νευρώνων που ανήκουν στο στρώμα εξόδου. Το νέο κριτήριο που προτείνεται με την εργασία, είναι να τερματίζεται η εκπαίδευση του *back propagation* όταν αυτή η ποσότητα OutputNeuronsErrorMean είναι μικρότερη από ένα κατώφλι MinOutputNeuronsErrorMean. Μπορεί να διατυπωθεί ως εξής: Η εκπαίδευση του *back propagation* θα τερματιστεί όταν:

$$\left(\frac{1}{n}\right) \sum_{i=1}^n \left(\left(\frac{1}{c}\right) \sum_{p=1}^c \delta[p]_i \right) < \varepsilon \quad (5.1)$$

όπου:

- ε ένας μικρός δεκαδικός αριθμός,
- n το πλήθος των νευρώνων του στρώματος εξόδου,
- c το πλήθος των προτύπων στο σύνολο των προτύπων εκπαίδευσης,
- $\delta[p]_i$ ο παράγοντας σφάλματος του νευρώνα, από την σχέση 3.22

$$\Delta w_{ij}(k) = \beta \delta_i x_j - \mu \Delta w_{ij}(k-1)$$

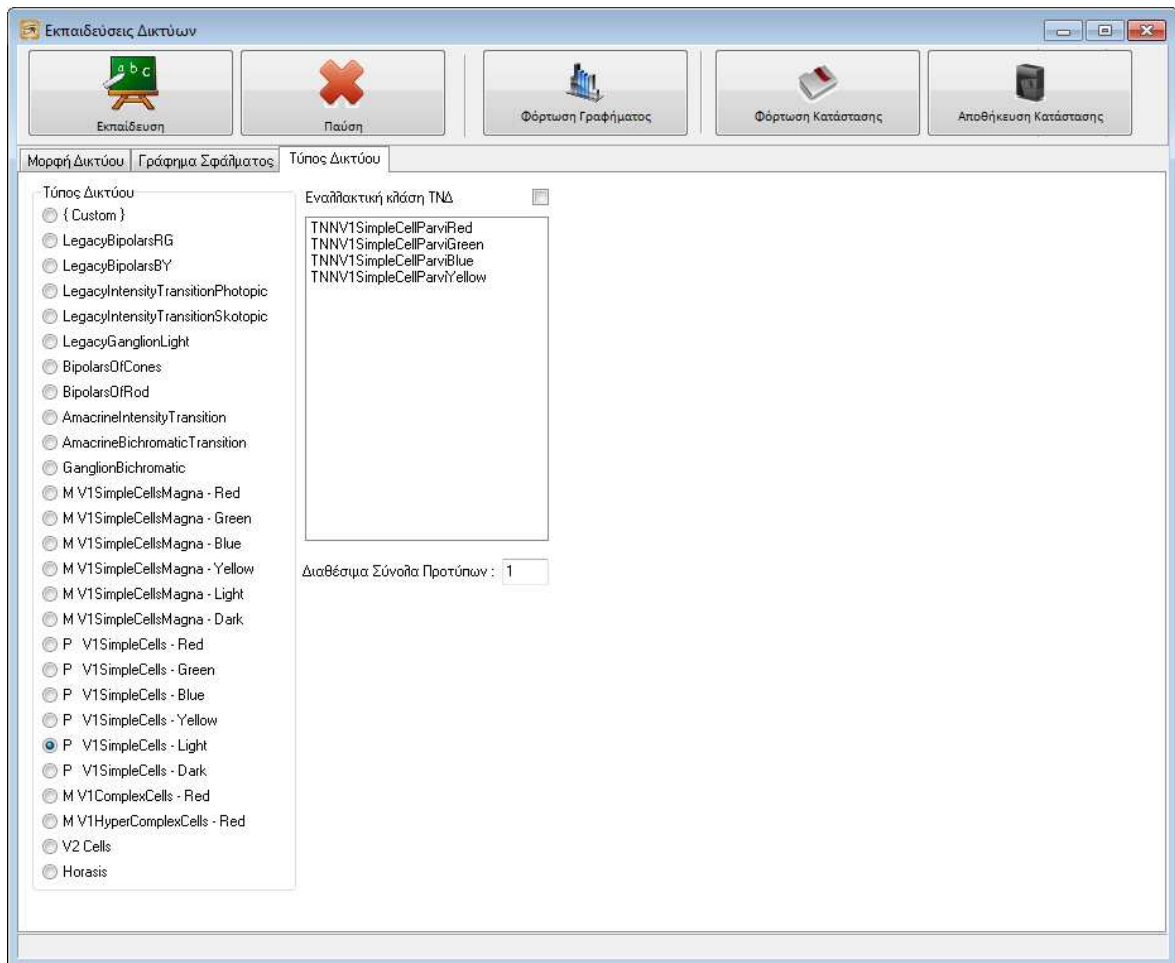
ο οποίος υπολογίζεται κατά την ανάκληση ενός προτύπου p για τον νευρώνα i του στρώματος εξόδου.

Με αυτό το κριτήριο αν ένας από τους νευρώνες της εξόδου δεν εκπαιδεύεται σωστά με τους άλλους να έχουν χαμηλό σφάλμα, ο τερματισμός της εκπαίδευσης δεν θα είναι εφικτός καθώς ο μέσος όρος δεν θα πέφτει ποτέ κάτω από το επιθυμητό κατώφλι. Έστω τέσσερις νευρώνες εξόδου με αντίστοιχα ErrorMean 0,1 0,15 0,12 και 1,1. Αν το κατώφλι είναι η τιμή 0,3 δεν θα μπορέσει να γίνει τερματισμός της εκπαίδευσης αν ο τέταρτος νευρώνας δεν ρίξει το σφάλμα του κάτω του 1.

Οπτικοποίηση της Εκπαίδευσης

Το επίπεδο εμφάνισης (presentation layer) της εκπαίδευσης

Γνωρίζοντας από τις προηγούμενες ενότητες για την οπτική ενημέρωση που παρέχεται προς τον χρήστη κατά την διαδικασία της εκπαίδευσης, θα δούμε αναλυτικά από ποιες πληροφορίες συνίσταται. Καταρχήν βλέπουμε παρακάτω την πλήρη εικόνα της φόρμας «Εκπαιδεύσεις Δικτύων», μέσα από την οποία γίνονται οι εκπαιδεύσεις, η οποία αποτελείται από μια λωρίδα εντολών και από τρεις σελίδες ενός οπτικού συστατικού σελίδων TPageControl:



5-7: Ρυθμίσεις εκπαίδευσης των ΤΝΔ του Horasis

Στην λωρίδα υπάρχουν οι εντολές «Εκπαίδευση» και «Παύση» που ξεκινούν και σταματούν το νήμα της εκπαίδευσης, οι «Αποθήκευση Κατάστασης» και «Φόρτωση Κατάστασης» με τις οποίες αποθηκεύεται η κατάσταση του δικτύου στον δίσκο και επαναφέρεται από αυτό αντίστοιχα και η εντολή «Φόρτωση Γραφήματος». Με την τελευταία φορτώνονται μόνο οι τιμές του μέσου τετραγωνικού σφάλματος που είχαν προκύψει κατά την εκπαίδευση, μέσα από μια αποθηκευμένη κατάσταση δικτύου. Από αυτές παράγεται μια γραφική παράσταση στην σελίδα «Γράφημα Σφάλματος».

Αρχικά σε μια εκπαίδευση χρησιμοποιείται η σελίδα «Τύπος Δικτύου» για την επιλογή μεταξύ των προκαθορισμένων διαφορετικών τύπων ΤΝΔ που χρησιμοποιούνται στο Horasis. Οι προκαθορισμένοι τύποι εμφανίζονται στην ομώνυμη ομάδα επιλογής «Τύπος Δικτύου». Με την επιλογή της ομάδας εμφανίζεται το πλήθος των διαθέσιμων συνόλων εκπαίδευσης για τον

συγκεκριμένο τύπο δικτύου. Ο κάθε τύπος ΤΝΔ έχει και την αντίστοιχη κλάση του που κληρονομεί από την TGNeuralNetwork, οπότε πριν την έναρξη της εκπαίδευσης του δικτύου θα δημιουργηθεί δυναμικά ένα αντίστοιχο αντικείμενο. Εφόσον στην θέση της προκαθορισμένης κλάσης θέλουμε να δοκιμάσουμε μια παραλλαγή της ή και ακόμα την κλάση ενός άλλου τύπου ΤΝΔ με τον ίδιο αριθμό εισόδων και εξόδων, χρησιμοποιείται η λίστα «Εναλλακτική κλάση ΤΝΔ». Αυτή η λίστα εμφανίζει ονόματα των κλάσεων που έχουν δηλωθεί σε ένα μητρώο στον πηγαίο κώδικα της εφαρμογής. Έτσι αν επιλέξουμε ένα τύπο για εκπαίδευση και μια εναλλακτική κλάση ΤΝΔ, η τελευταία χρησιμοποιείται για την δημιουργία του αντικειμένου ΤΝΔ, αλλά φορτώνονται σε αυτό τα πρότυπα εκπαίδευσης του συγκεκριμένου τύπου. Μπορούμε κατ' αυτόν τον τρόπο να δοκιμάσουμε αν μια από της μορφές δικτύου που έχουν εκπαιδευτεί με επιτυχία, ταιριάζει για την επίλυση ενός διαφορετικού προβλήματος.

Το παράθυρο παρακολούθησης εποχών και μέσου τετραγωνικού σφάλματος

Ανοίγοντας την σελίδα «Μορφή δικτύου» βλέπουμε εκτός από οπτικοποίηση της μορφής του αντικειμενοστραφούς ΤΝΔ ένα πλαίσιο καταγραφής (log) στο οποίο παρέχονται πληροφορίες για την εποχή στην οποία βρίσκεται η εκπαίδευση, το μέσο τετραγωνικό σφάλμα και τα millisecond που πέρασαν από την τελευταία ενημέρωση του UI μέσα σε αγκύλες. Ο χρόνος αυτός όπως είδαμε ταυτίζεται και με τον χρόνο που πέρασε από τον τελευταίο υπολογισμό του μέσου τετραγωνικού σφάλματος.

Εκπαιδύσεις Δικτύων

Εκπαίδευση Παύση Φόρτωση Γραφήματος Φόρτωση Κατάστασης Αποθήκευση Κατάστασης

Μορφή Δικτύου Γράφημα Σφάλματος Τύπος Δικτύου

Νευρωνικό Δίκτυο

Είσοδος Κρυφό 1 Κρυφό 2 Εξόδος

Μέσος Όρος Σφάλματος Νευρώνων Εξόδου: 1,157374

1 2 3 4 5 6
0,954 0,953 1,107 1,456

000015 | 0,89264020 {0}
000030 | 0,89259478 {1482}
000045 | 0,89256640 {1482}
000060 | 0,89249377 {1482}
000075 | 0,89237777 {1466}
000090 | 0,89186181 {1482}
000105 | 0,88905469 {1451}
000120 | 0,87678126 {1451}
000135 | 0,83707535 {1451}
000150 | 0,75023312 {1466}
000165 | 0,62156339 {1466}
000180 | 0,49448407 {1467}
000195 | 0,38953567 {1451}
000210 | 0,31783194 {1466}
000225 | 0,26810454 {1451}
000240 | 0,23198937 {1451}
000255 | 0,20539685 {1466}
000270 | 0,18481413 {1482}

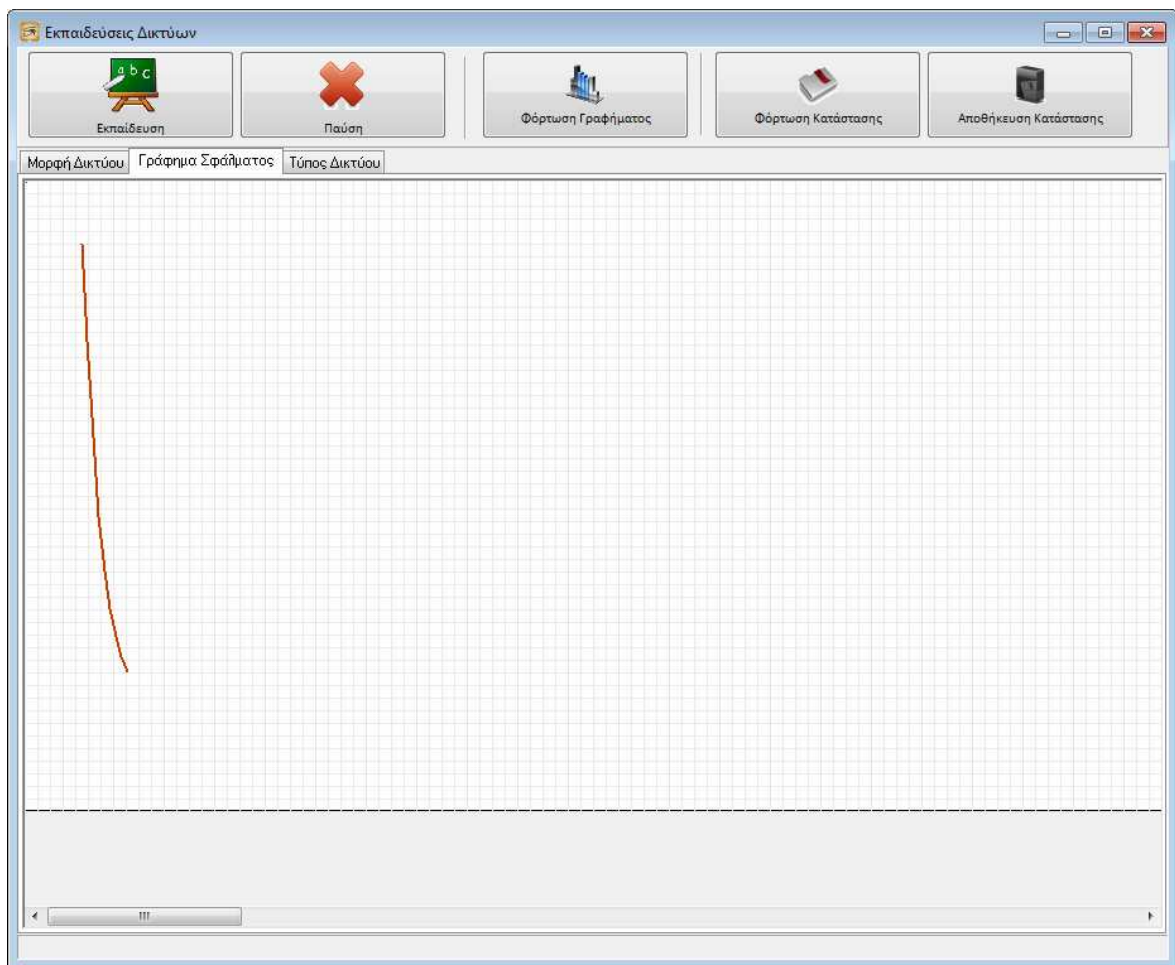
5-8: Εκπαίδευση ενός ΤΝΔ μέσα στο Horasis

Έτσι πληροφορούμαστε από το log ότι στην παραπάνω εκπαίδευση βρισκόμαστε στην εποχή 270 της εκπαίδευσης του δικτύου στην οποία το σφάλμα είναι περίπου 0,18 και πέρασαν 1482 msec από την προηγούμενη δειγματοληψία. Με αυτόν τον τρόπο παρατηρούμε πόσο γρήγορα πέφτει το σφάλμα και το κυριότερο αν έχει «κολλήσει», δηλαδή οι μεταβολές των συναπτικών βαρών μετά από ένα πλήθος εποχών δεν καταφέρνουν να το μειώσουν. Επίσης παρατηρούμε την απόδοση του νήματος εκπαίδευσης αφού έχοντας ορίσει 1500msec την περίοδο υπολογισμού και ενημέρωση του UI, περιμένουμε το ίδιο χρόνο ή λιγότερο να εμφανίζεται μέσα στις αγκύλες. Τυχόν καθυστέρηση στην λειτουργία της εφαρμογής ή κατανάλωση πόρων επεξεργασίας του υπολογιστή, θα μας δώσει ένα νούμερο μεγαλύτερο από αυτό που περιμένουμε. Τέλος από την στιγμή που θα παρατηρήσουμε ότι το μέσο τετραγωνικό σφάλμα είναι κοντά ή έχει πέσει κάτω

από το 0,8 μπορούμε να γυρίσουμε στην σελίδα «Γράφημα Σφάλματος» και να παρατηρήσουμε πλέον την εξέλιξη της εκπαίδευσης μέσω της γραφικής παράστασης που υπάρχει.

Η σελίδα γραφικής παράστασης του μέσου τετραγωνικού σφάλματος

Η σελίδα αυτή αποτελείται από ένα παράθυρο με μια δυναμικά ενημερούμενη γραφική παράσταση. Σε αυτήν στον οριζόντιο άξονα των τετμημένων (x) έχουμε τις εποχές εκπαίδευσης και στον κάθετο άξονα των τεταγμένων (y) έχουμε την τιμή του μέσου τετραγωνικού σφάλματος. Έτσι βλέπουμε την φθίνουσα πορεία της συνάρτησης του σφάλματος κατά την εκπαίδευση του back propagation, βλέποντας την να «κατηφορίζει» προς την λύση ακριβώς όπως στο παράδειγμα της μπάλας που κατηφορίζει την πλαγιά προς κάποιο ελάχιστο, που ελπίζουμε να μην είναι τοπικό.

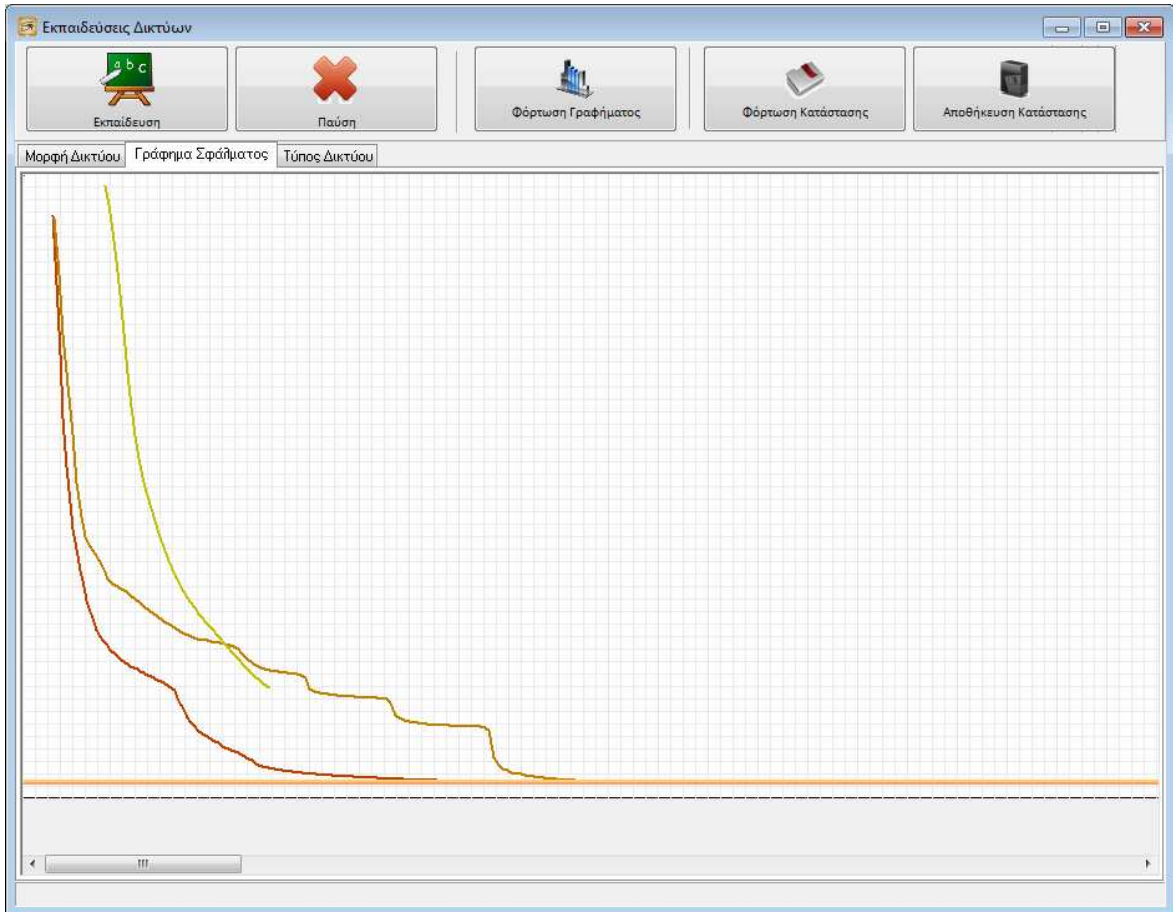


5-9: Γράφημα μέσου τετραγωνικού σφάλματος της εκπαίδευσης

Η γραφική παράσταση του σφάλματος ίσως είναι το σημαντικότερο οπτικό βοήθημα στην πειραματική επιλογή της καλύτερης μορφής του δικτύου. Δείχνει την σταθερότητα ή αντίθετα της διακυμάνσεις τις εκπαίδευσης. Επίσης δείχνει τα τοπικά ελάχιστα όταν η γραμμή τείνει να γίνει ευθεία δηλαδή όσο βλέπουμε το ρυθμό που αυτή μετακινείται προς μικρότερες κάθετες συντεταγμένες της οθόνης να μειώνεται. Επίσης δείχνει τις περιπτώσεις που το δίκτυο ταλαντώνεται, έχοντας πολύ βίαιες αυξομειώσεις του σφάλματος προς τα πάνω και προς τα κάτω σαν μια πριονωτή κυματομορφή.

Στο Hogasis εισάγεται μια περαιτέρω βελτίωση για το γράφημα του σφάλματος. Με την χρήση της «Φόρτωσης Γραφήματος» φορτώνονται από μια αποθηκευμένη κατάσταση δικτύου, οι τιμές

του γραφήματος που είχε η εκπαίδευση του. Αυτές αναπαριστούνται ως γραμμές διαφορετικών χρωμάτων για να ξεχωρίσουν στον χρήστη. Έτσι βλέπουμε παρακάτω την κόκκινη γραμμή που αναπαριστά την εκπαίδευση με την προκαθορισμένη μορφή του επιλεγμένου τύπου ΤΝΔ, με πορτοκαλί έναν εντελώς διαφορετικό τύπο ΤΝΔ του Horasis και με πράσινο την εκπαίδευση του επιλεγμένου τύπου με μια εναλλακτική μορφή. Με αυτόν τον τρόπο μπορεί να γίνει απευθείας οπτική σύγκριση των τριών διαφορετικών εκπαιδεύσεων.

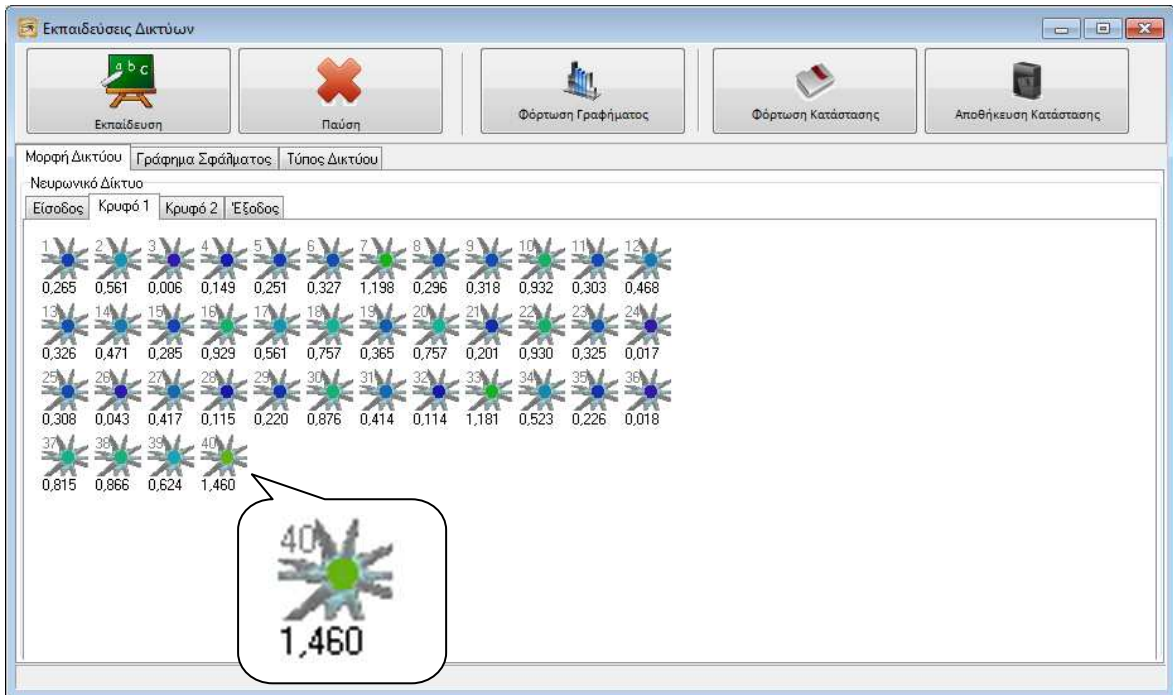


5-10: Σύγκριση γραφημάτων διαφορετικών εκπαιδεύσεων

Ακόμα ένα βοήθημα που παρατηρούμε είναι ότι στην συντεταγμένη οθόνης γ , στην οποία αντιστοιχεί η τελευταία τιμή που μετρήθηκε για το μέσο τετραγωνικό σφάλμα τραβιέται μια οριζόντια γραμμή. Η γραμμή αυτή που την ονομάζουμε **Γραμμή Πυθμένα Σφάλματος (Error Bottom Line)** αντιπροσωπεύει το τοπικό ελάχιστο στο οποίο έχει φτάσει η εκπαίδευση με back propagation για το ΤΝΔ και αποτελεί μέτρο σύγκρισης για μια άλλη εκπαίδευση του ίδιου δικτύου. Αν σε μια νέα εκπαίδευση, όπως είναι εδώ η πράσινη γραμμή, πλησιάζει πιο γρήγορα ή πιο ομαλά την γραμμή πυθμένα σφάλματος μια παλαιότερης, τότε εμπειρικά διαφαίνεται ότι ίσως την ξεπεράσει. Αν ξεπεραστεί η γραμμή του πυθμένα σφάλματος κατά πολύ τότε η νέα εκπαίδευση είναι σημαντικά καλύτερη από την παλαιότερη. Στην περίπτωση που αυτό γίνει τέμνοντας την γραμμή της παλιάς, έχουμε και γρηγορότερη σύγκλιση του αλγορίθμου back propagation. Με τα παραπάνω εργαλεία μπορούμε να λάβουμε γρήγορες και έγκαιρες αποφάσεις για την αλλαγή της μορφή του δικτύου που θα μας οδηγήσει στην πιο εύρωστη εκπαίδευση.

Η χρωματική αναπαράσταση του σφάλματος των νευρώνων

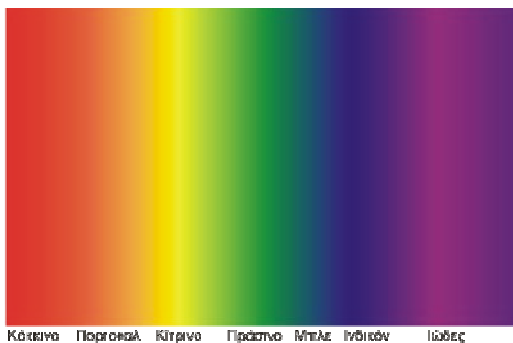
Ήδη γνωρίσαμε την αναπαράσταση της μορφής του δικτύου με οπτικά συστατικά και την αναγραφή σε κάθε νευρώνα που αναπαρίσταται, του συντελεστή σφάλματος δ_i από τον τύπο υπολογισμού 3.22 της τροποποίησης των συναπτικών βαρών.



5-11: Εμφάνιση σφαλμάτων των νευρώνων του κρυφού στρώματος 1, κατά την εκπαίδευση του Horasis

Ως ένας τρόπος για να φανεί διαισθητικά στον χρήστη αυτό σφάλμα, χρησιμοποιείται ο χρωματισμός ενός κύκλου στο κέντρο του οπτικού συστατικού που αντιπροσωπεύει τον νευρώνα. Οι χρωματισμοί αυτοί κινούνται στο φάσμα του ορατού φωτός ξεκινώντας από το κόκκινο περνώντας σταδιακά στο πορτοκαλί όταν το σφάλμα του νευρώνα κατεβαίνει το όριο του 2.2, κατόπιν στο πράσινο, κυανό, καταλήγοντας στο ινδικό⁵¹(indigo) χρώμα εφόσον το

Ορατό Φάσμα

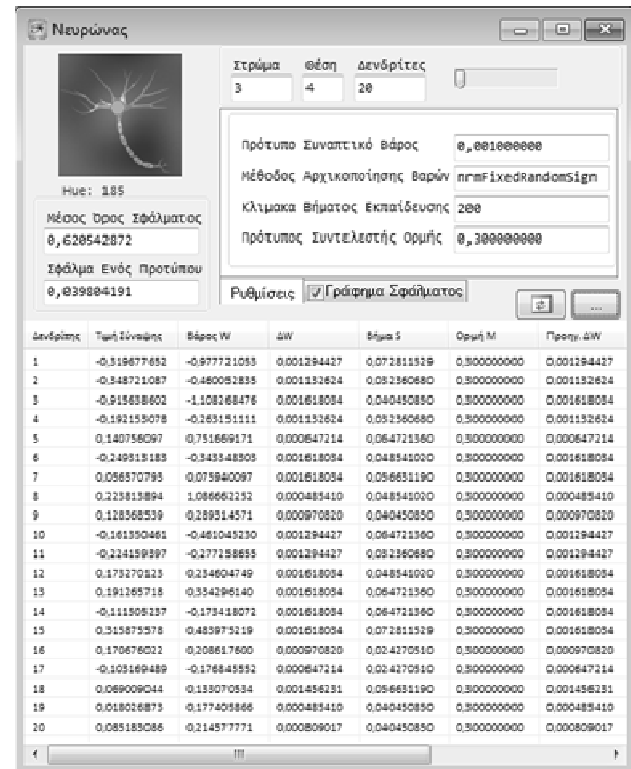
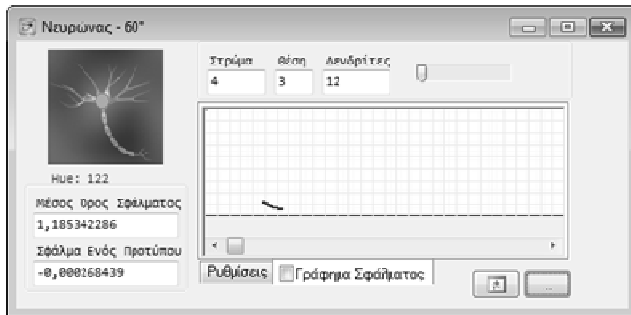
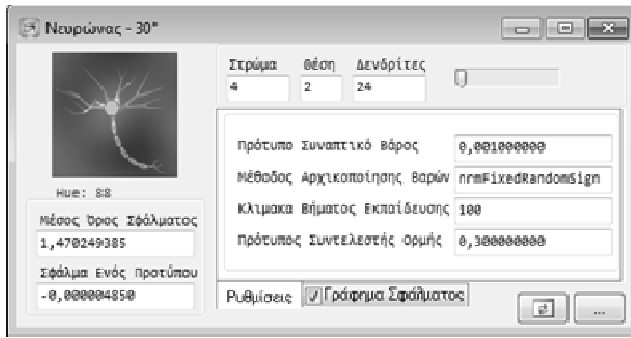


5-12: 7 βασικά χρώματα του ορατού φάσματος

σφάλμα είναι μικρότερο από 0,001. Μια ακόμη ενδιαφέρουσα αντιστοιχία είναι αν θεωρήσουμε το χρώμα του πυρωμένου σιδήρου καθώς αυτό κρυνώνει. Έτσι μπορούμε να φανταστούμε «ζεστούς» τους νευρώνες που είναι χρωματισμένοι κόκκινοι και «κρύους» όταν είναι πλήρως σχηματισμένοι ως προς τα συναπτικά τους βάρη και εμφανίζονται ως μπλε. Ο χρωματισμός όπως φαίνεται και στην παραπάνω οθόνη έχει μεγάλο ενδιαφέρον καθώς δίνει κατευθείαν πληροφόρηση για την διαφορετική κατανομή του σφάλματος της εξόδου του δικτύου στα πιο πίσω στρώματα.

⁵¹ Η αγγλική λέξη για το χρώμα indigo προέρχεται από την λατινική «indicum» η οποία αντιστοιχούσε στην ελληνική «ινδικόν», λόγω του φυτού από το οποίο παραγόταν το χρώμα.

Φόρμα επισκόπησης νευρώνα

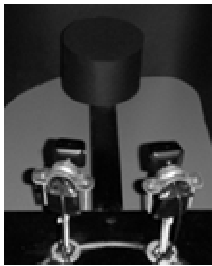


Κάτω από το κάθε οπτικό συστατικό νευρώνα κρύβεται μια φόρμα που δείχνει όλες τις εσωτερικές του πληροφορίες. Η φόρμα εμφανίζεται με κλικ πάνω στο εικονίδιο του νευρώνα και δείχνει αρχικά τις πληροφορίες για την θέση του, το πλήθος των δενδριτών και τις συναπτικές ιδιότητες που έχουν οριστεί για αυτόν. Υπάρχει επίσης η δυνατότητα εμφάνισης ενός μίνι γραφήματος της τιμής μέσου όρου των συντελεστών σφάλματος ErrorMean για τον κάθε νευρώνα του δικτύου, δηλαδή του $\left(\frac{1}{c}\right) \sum_{p=1}^c \delta[p]$ όπου c το πλήθος των προτύπων εκπαίδευσης. Η φόρμα μπορεί να επεκταθεί με την χρήση του πλήκτρου με τις τρεις τελείες και να εμφανίσει μια λίστα με όλους τους δενδρίτες του νευρώνα. Για κάθε έναν από αυτούς εμφανίζει μια πληθώρα δεδομένων όπως το συναπτικό βάρος της εισερχόμενης σύναψης, την τροποποίηση του κατά ΔW μέσα στην εποχή, το βήμα εκπαίδευσης, την ορμή, το ΔW της προηγούμενης εποχής και άλλα. Η φόρμα αυτή είναι το «ηλεκτρονικό μικροσκόπιο» που έχουμε για μικρόκοσμο της εκπαίδευσης back propagation. Δύο σημαντικές πληροφορίες που μπορούν να εξαχθούν είναι αν υπάρχουν πολύ μικρά ΔW σε έναν νευρώνα για την εποχή και το πρόσημο του. Η φόρμα αυτή ενημερώνεται δυναμικά όπως όλο το UI εφόσον έχουμε ανοικτή την σελίδα του στρώματος στο οποίο ανήκει ο νευρώνας, ειδάλλως χρησιμοποιούμε το πλήκτρο ανανέωσης.

5-13: Τρεις διαφορετικές προβολές της φόρμας επισκόπησης ενός νευρώνα

6. Τα Οπτικά Πρότυπα

Με το σύστημα Horasis επιχειρείται αναγνώριση πραγματικών αντικειμένων στον χώρο μέσα από τα οπτικά του πρότυπα. Τα πρότυπα αυτά θα μπορούσαν να έχουν συλλεχθεί από έτοιμες φωτογραφίες ή να είναι σκίτσα των αντικειμένων. Στην δική μας περίπτωση τα πρότυπα δημιουργούνται από το ίδιο το σύστημα, καθώς τα αντικείμενα φωτογραφίζονται με την χρήση ψηφιακής φωτογράφισης τοποθετούμενα σε μια σκηνή. Το αρχικό πρότυπο για το αντικείμενο του χώρο θεωρηθεί η ψηφιοποιημένη μορφή της εικόνας, που αποτελείται από ένα σύνολο



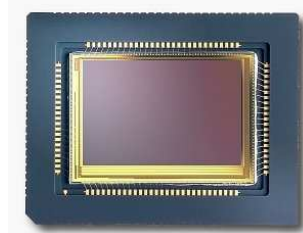
6-1: Σκηνή
φωτογράφισης
Horasis

ψηφίδων (raster) που χαρακτηρίζονται από τρεις χρωματικές συνιστώσες R, G, B για το κόκκινο, πράσινο και μπλε χρώμα. Αυτό το πρότυπο μετασχηματίζεται σε μια νέα μορφή πληροφορίας από το Horasis, με την οποία τελικά θα εκπαιδευτεί το τελικό ΤΝΔ που κατηγοριοποιείται τα πραγματικά αντικείμενα. Θα ξεκινήσουμε από την αρχή βλέποντας τον τρόπο που τα πραγματικά αντικείμενα μετατρέπονται σε σήματα εισόδου στους πρώτους νευρώνες του Horasis. Για να γίνει αυτό αξιοποιείται το υλικό της ψηφιακής κάμερας web (web camera), το υποσύστημα πολυμέσων του λειτουργικού συστήματος Windows και οι προγραμματιστικές δυνατότητες της RAD γλώσσας προγραμματισμού Delphi.

Η εικόνα στον υπολογιστή

Η τεχνολογία της ψηφιακής κάμερας

Μια ψηφιακή φωτογραφική μηχανή ή κάμερα βιντεοσκόπησης μετατρέπει το φως σε ηλεκτρικά σήματα μέσα από έναν αισθητήρα που διαθέτει [77]. Ένας συνηθισμένος αισθητήρας που υπάρχει σε πολλές συσκευές είναι ο CCD (Charge Coupled Device) που αποτελείται από μια συστοιχία κυψελών από πυκνωτές MOS⁵²[78]. Ο δεύτερος συνηθισμένος αισθητήρας που υπάρχει είναι ο CMOS (Complementary Metal Oxide Semiconductor) που όπως και ο CCD είναι και αυτός μια συστοιχία κυψελών. Η διαφορά έγκειται στον τρόπο που μετατρέπουν το φως σε ηλεκτρικό σήμα καθώς κάθε ριχέλ της CMOS εμπεριέχει μια φωτοδίοδο, έναν πυκνωτή και έως τρία τρανζίστορ [79]. Οι βασικές διαφορές μεταξύ των δύο τύπων αισθητήρων είναι:



6-2: Αισθητήρας CMOS
Foveon X3

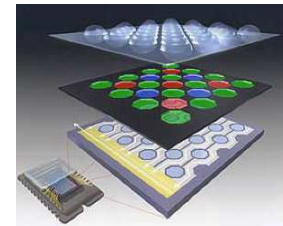
- Οι CCD αισθητήρες δημιουργούν εικόνες υψηλής ποιότητας καθώς έχουν λιγότερο θόρυβο (noise) σε σχέση με τις CMOS.
- Η ευαισθησία του αισθητήρα CMOS στο φως είναι μικρότερη καθώς όλα τα φωτόνια στα χωρικά όρια μια κυψέλης, δεν πέφτουν πάνω στις φωτοδιόδους.
- Ένας αισθητήρα CMOS καταναλώνει μόνο 1% της ενέργειας ενός αντίστοιχου CCD άρα μπορεί να λαμβάνει φωτογραφίες με την ίδια ενέργεια για πολύ περισσότερο χρόνο.

⁵² MOS = Metal Oxide Semiconductor

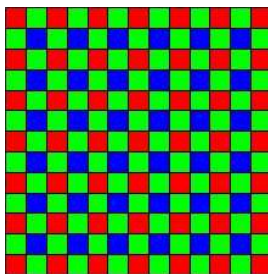
- Οι CCD είναι πιο ώριμα προϊόντα καθώς η τεχνολογία CMOS ξεκίνησε το 1993 με την έρευνα στο κέντρο JPL ⁵³της NASA.

Η ποσότητα των λεπτομερειών που μπορεί να καταγράψει μια ψηφιακή κάμερα λέγεται ανάλυση και μετρείται σε Megarixels δηλαδή εκατομμύρια rixels. Αυτά Όσο μεγαλύτερη είναι τόσο η εικόνα έχει καλύτερες λεπτομέρειες και λιγότερο θόλωμα. Σήμερα οι καλύτερες καταναλωτικές συσκευές έχουν 12 Megarixels, κάποιες επαγγελματικές 16 έως 20 ενώ για να συγκρίνουμε το αναλογικό φιλμ 35 χιλιοστών υπολογίζεται από την εταιρία Hewlett Packard να αντιστοιχεί στην ποιότητα των 20 Megarixels. Έτσι σε μια ψηφιακή φωτογραφική μηχανή των 11.1 Megarixels η ανάλυση είναι η 4064x2704 δηλαδή υπάρχουν τόσες κυψέλες στον αισθητήρα της όσο το γινόμενο των δύο διαστάσεων της ανάλυσης.

Η κάθε κυψέλη του αισθητήρα δεν βλέπει απευθείας χρώμα αλλά μετράει μια ποσότητα ηλεκτρισμού για τα φωτόνια που πέφτουν πάνω της. Έτσι για να καθοριστεί η τιμή των τριών χρωμάτων κόκκινου, πράσινου και κίτρινου χρησιμοποιείται φιλτράρισμα τοποθετώντας μπροστά από τον αισθητήρα το φίλτρο που θα επιτρέψει μόνο το επιθυμητό χρώμα να περάσει προς αυτόν. Μια υλοποίηση είναι να υπάρχουν τρεις διαφορετικοί αισθητήρες με το χρωματικό φίλτρο τους αντίστοιχο με το χρώμα και ένας διαχωριστής δέσμης (beam splitter) να οδηγεί ίση ποσότητα φωτός και προς τους τρεις αισθητήρες ταυτόχρονα. Μια άλλη υλοποίηση είναι η ταχύτητα εναλλαγής τριών διαφορετικών φίλτρων μπροστά από έναν κοινόχρηστο αισθητήρα σε διαδοχικές χρονικές στιγμές.



6-3: Διαστρωμάτωση αισθητήρα CMOS (<http://www.steves-digicams.com>)



6-4: Μωσαϊκό χρωμάτων του φίλτρου Bayer

Μια τρίτη υλοποίηση είναι η χρήση μωσαϊκού χρωμάτων έχοντας ένα σύνθετο φίλτρο που επιτρέπει διαφορετικό χρώμα να περάσει σε κάθε κυψέλη. Με αυτόν τον τρόπο φτιάχνονται φθηνότερες συσκευές με μόνο έναν αισθητήρα, και αρκετά αποδοτικές για το κόστος τους. Έτσι η κάθε κυψέλη μετράει μόνο μια από τις χρωματικές συνιστώσες και για τις άλλες δύο χρησιμοποιεί παρεμβολή⁵⁴ (interpolation) κοιτώντας τα γειτονικά της rixel. Ένα τέτοιο φίλτρο είναι το φίλτρο με μοτίβο Bayer ⁵⁵το οποίο στηρίζεται στην μεγαλύτερη ευαισθησία που έχει το ανθρώπινο μάτι ως προς τα χαμηλά μήκη κύματος έχοντας παραπάνω πράσινες κυψέλες σε σχέση με τις κόκκινές και τις ακόμα λιγότερες μπλε. Η παρεμβολή των τιμών που αφορούν μόνο μια συνιστώσα σε τριπλέτες τιμών R,G,B για την κάθε κυψέλη μωσαϊκού, γίνεται με αλγορίθμους απομωσαϊκοποίησης (demosaicing) [80].

⁵³ **JPL** = Jet Propulsion Laboratory είναι το ερευνητικό κέντρο της Nasa στην Pasadena της California το οποίο συντονίζει την αποστολή μη επανδρωμένων εξερευνητικών σκαφών όπως το Cassini-Huygens.

⁵⁴ Πηγή μετάφρασης Ε.Μ.Π. Ορολογία Μαθηματικής Λογικής [158]

⁵⁵ Ονομάστηκε έτσι από τον εφευρέτη Bryce E. Bayer της εταιρίας Kodak.

Η πρόοδος που επιτελέστηκε πάνω σε αυτήν την τεχνολογία υλικού, που οφείλεται και στην μείωση του κόστους που την κάνει προσιτή σε όλους, είναι πολύ σημαντική χωρίς βέβαια να έχει αλλάξει η θεμελιώδης φυσική που την διέπει. Έτσι από τον πρώτο αισθητήρα CCD το 1974 της εταιρίας Fairchild, που είχε ανάλυση 100x100 pixel και χρησιμοποιήθηκε ερευνητικά σε αστρονομικό τηλεσκόπιο, φτάσαμε να έχουμε στον αισθητήρα μιας απλούστατης κάμερας του εμπορίου 1280x720 pixel, η οποία χρησιμοποιείται απλά για επικοινωνία μέσω του internet με προγράμματα τηλεδιάσκεψης όπως το Skype⁵⁶. Ακριβώς μια τέτοια ψηφιακή συσκευή απεικόνισης είναι η Microsoft LifeCam HD-5000 που έχει έναν αισθητήρα CMOS και ανάλυση 1280x720. Με αυτήν ελήφθησαν οι φωτογραφίες των προτύπων του Horasis, από καρέ της ροής βίντεο που παράγει, έχοντας κόστος κατά την στιγμή της



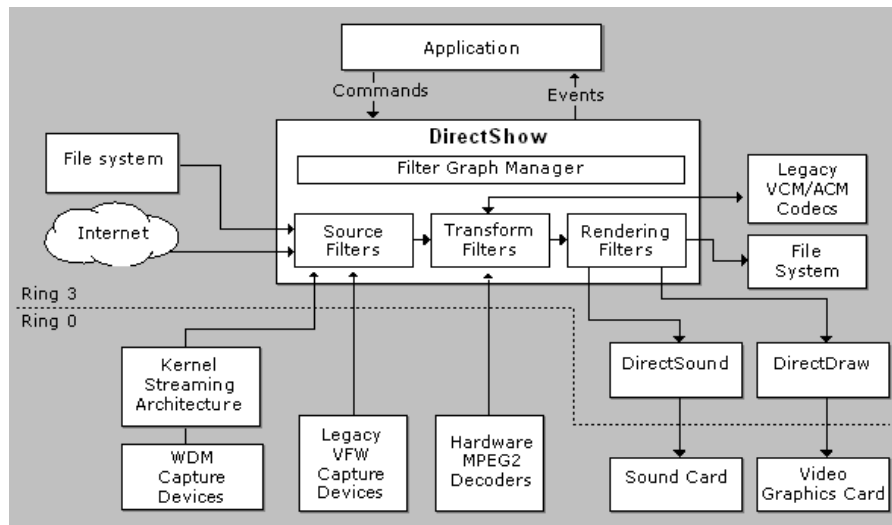
6-5: Η απλή web κάμερα του horasis, Microsoft LifeCam HD-5000

αγοράς της μόλις 40€. Η πιστότητα της εικόνας είναι ίσως αρκετά κατώτερη από μια πιο εξειδικευμένη συσκευή απεικόνισης. Παρόλαταυτα επιλέχθηκε επίτηδες με την σκέψη ότι αν πετύχαινε η μέθοδος του Horasis, αυτομάτως δημιουργείται μεγάλη προοπτική με την εφαρμογή της πάνω σε καλύτερο υλικό. Αν κοιτάξουμε λίγο στο παρόν για το τι δυνατότητες μπορεί να δώσει ένα εξειδικευμένο υλικό βλέπουμε το εντυπωσιακό αποτέλεσμα που παρήγαγε ο φωτογράφος David Bergman, κατορθώνοντας να δημιουργήσει μια φωτογραφία 1474 Megapixel από την αναγόρευση του 44^{ου} προέδρου των Η.Π.Α. Barack Obama [81].

⁵⁶ Το Skype είναι ένα δωρεάν σύστημα τηλεφωνικής και κλήσης βίντεο μέσω του inter

Η διασύνδεση με Microsoft DirectShow

Από τον ένα πόλο του hardware θα μεταπηδήσουμε στον άλλο πόλο του στο λειτουργικό σύστημα (operating system) [82], ή για συντομία ΛΣ, Windows. Ένα λειτουργικό πλέον δεν εξυπηρετεί μόνο τους βασικούς στόχους, όπως η εκτέλεση των εφαρμογών στο υλικό του υπολογιστή και η επίβλεψη του και η διαχείριση των πόρων του υπολογιστή και των διάφορων συσκευών, ανάλογα με την περίπτωση που αντιμετωπίζει [83]. Η παροχή κοινόχρηστων λειτουργιών προς όλες τις εφαρμογές είναι ένα από τα βασικά χαρακτηριστικά όλων των σύγχρονων ΛΣ. Μέσα σε αυτές τις λειτουργίες των Windows βρίσκουμε το Microsoft DirectShow. Είναι ένα ενδιάμεσο υποσύστημα ή στα αγγλικά middleware της πλατφόρμας των Windows, που υλοποιεί μια αρχιτεκτονική πολυμέσων, όπως η ροή βίντεο που μας ενδιαφέρει και στηρίζεται στο COM⁵⁷ μοντέλο [84]. Οι εφαρμογές που μπορούν να γραφτούν στηριζόμενες σε αυτό είναι αναπαραγωγή βίντεο, τηλεόρασης, DVD, προγράμματα επεξεργασίας βίντεο, λογισμικό μετατροπής διαμορφώσεων (converters) και εφαρμογές καταγραφής ήχου και εικόνας.



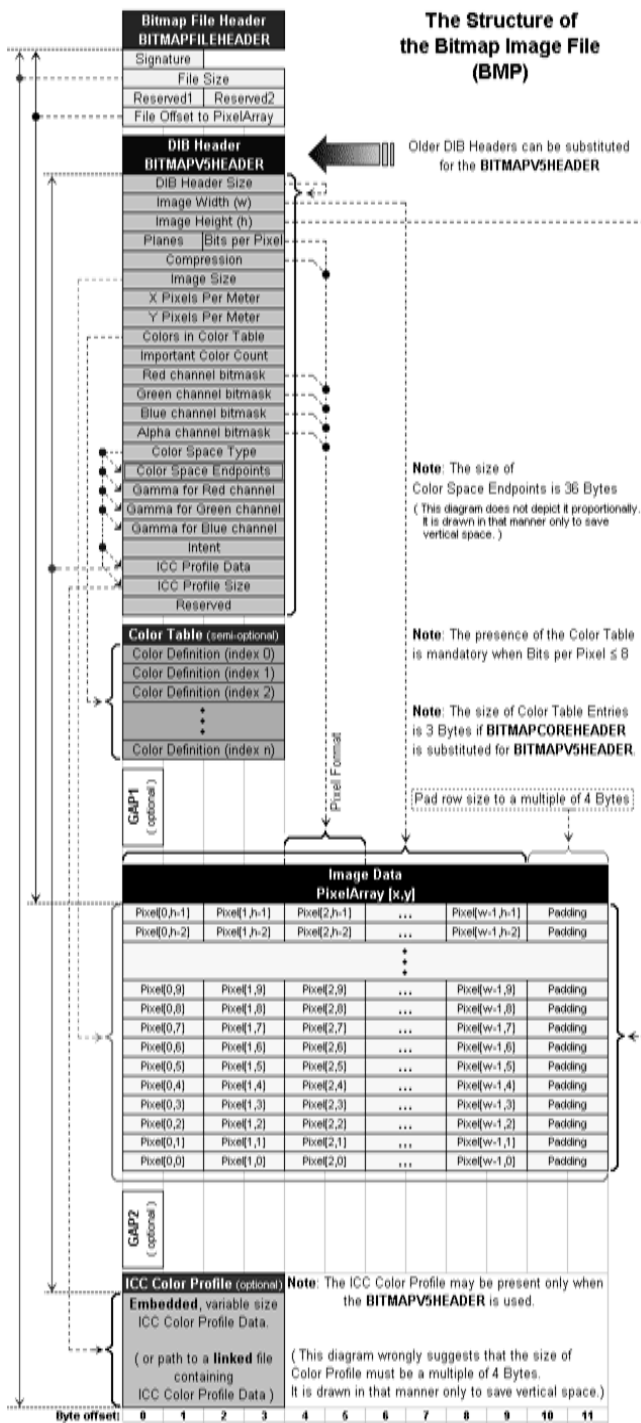
6-6: Σχηματική αναπαράσταση της αρχιτεκτονικής DirectShow (MSDN)

Στην εφαρμογή Horasis αξιοποιείται η λήψη καρτέ από μια ροή βίντεο του DirectShow. Από το παραπάνω σχεδιάγραμμα μας ενδιαφέρει η χρήση των “WDM Capture Devices” στις οποίες ανήκει η web κάμερα. Οι λεπτομέρειες της υλοποίησης με την χρήση του DirectShow API⁵⁸ [85][86], δεν κρίνεται σκόπιμο να αναφερθούν καθώς αποκρύπτονται και εμφωλεύονται μέσα σε μια κλάση Delphi που κατασκευάστηκε για τις ανάγκες του Horasis την TGVideoCaptureAgent. Αυτή έχει απλές μεθόδους όπως Init() για αρχικοποίηση, PreviewStart() και PreviewStop() για έναρξη και τερματισμό προεπισκόπησης και SnapShot() για την λήψη ενός καρτέ από την ροή βίντεο της προεπισκόπησης. Αξίζει να αναφερθεί σύστημα DirectShow λειτουργήσει χωρίς προβλήματα και σε 64bit λειτουργικό σύστημα, καθώς το περιβάλλον ανάπτυξης του Horasis ήταν Windows 7 x64. Το αποτέλεσμα της SnapShot() είναι η δημιουργία ενός αντικειμένου TBitmap [4] στην μνήμη που κρατάει τα δεδομένα σε μια διαμόρφωση εικόνας Bitmap που θα γνωρίσουμε παρακάτω.

⁵⁷ COM = Component Object Model

⁵⁸ API = Application Programming Interface είναι μια σύμβαση από εντολές και δομές δεδομένων που χρησιμοποιείται για την επικοινωνία δύο εφαρμογών χωρίς να έχει η μια ορατότητα της υλοποίησης που έχει η άλλη.

Η μορφή μνήμης της εικόνας BMP



6-7: Η ψηφιδωτή διαμόρφωση εικόνας BMP (Verpies/Wikimedia Commons)

Η διαμόρφωση Bitmap (BMP) που αναπαριστά την εικόνα, αρχικά στην μνήμη και στην συνέχεια σε κάποιο μέσο αποθήκευσης, είναι ψηφιδωτή (raster). Δηλαδή χρησιμοποιείται ένα ψηφιδωτό από εικονοστοιχεία (pixels), στο οποίο αποθηκεύεται για κάθε ένα από τα ψηφία η τριπλέτα χρωμάτων κόκκινο, πράσινο και μπλε. Αυτές αντιπροσωπεύουν τις φασματικές υπογραφές του στοιχείου της εικόνας [87]. Με απλά λόγια η εικόνα τεμαχίζεται σε μικρά τετραγωνίδια-ψηφίδες που το κάθε ένα από αυτά έχει συγκεκριμένη θέση στην εικόνα με συντεταγμένες δύο διαστάσεων, και συγκεκριμένες τιμές για κάθε ένα από τα τρία βασικά χρώματα. Όταν αυτές τρεις τιμές αυτές συνδυαστούν, παράγουν ένα πραγματικό χρώμα του φάσματος και γι αυτό ονομάζονται χρωματικές συνιστώσες.

Ονομάζεται χάρτης των bit (bit-map) επειδή έχουμε στην ουσία bits τοποθετημένα σε έναν πίνακα με θέσεις αντίστοιχες με τις συντεταγμένες της ψηφίδας. Η Microsoft έχει ορίσει την συγκεκριμένη διαμόρφωση γενικευμένα με σκοπό να προάγει την ανταλλαγή εικόνων διαφορετικών χρωματικών χαρακτηριστικών και μεταξύ διαφορετικών συσκευών υλικού και εφαρμογών λογισμικού. Έτσι η διαμόρφωση BMP καλείται και DIB (Device Independent Bitmaps) [88], δηλαδή χάρτες bit ανεξάρτητοι από συσκευή. Η συγκεκριμένη διαμόρφωση

BMP που χρησιμοποιείται είναι 24bits, δηλαδή τιμή της κάθε μιας χρωματικής συνιστώσας έχει μέγεθος 8bits με εύρος ακεραίων τιμών από 0 έως 255. Συνδυαζόμενες παράγουν 16 εκατομμύρια διαφορετικά πραγματικά χρώματα το λεγόμενο βάθος χρωμάτων (color depth) που έχει η εικόνα. Παρακάτω θα δούμε πως από τα bits που αναπαριστούν το πραγματικό χρώμα μια ψηφίδας της πραγματικής εικόνας, επανερχόμαστε στην απόχρωση δηλαδή στο αρχικό φάσμα του φωτός που μετρήθηκε από την κυψέλη του αισθητήρα της κάμερας.

Χρωματικά Μοντέλα

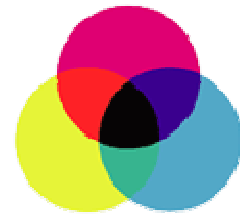
Η σύγχρονη χρωματική θεωρία και το προσθετικό μοντέλο RGB

Γνωρίσαμε ήδη ότι καταγράφονται τρία βασικά χρώματα κόκκινο, πράσινο και μπλε για κάθε χρώμα ενός σημείου της εικόνας, χωρίς να έχουμε αναφέρει ακόμα γιατί. Η απάντηση



6-9: Το προσθετικό σύστημα χρωμάτων

βρίσκεται στην σύγχρονη χρωματική θεωρία [89] ένα σώμα από πρακτικές οδηγίες για την ανάμιξη των χρωμάτων και τον οπτικό αντίκτυπο των συνδυασμών τους. Το σύστημα RGB είναι ένα προσθετικό σύστημα χρωμάτων (additive color) στο οποίο όσο ανεβαίνουν οι τιμές των χρωματικών συνιστωσών τότε το χρώμα γίνεται πιο φωτεινό καταλήγοντας στο άσπρο. Από την άλλη υπάρχει το αφαιρετικό σύστημα χρωμάτων (subtractive color) είναι όταν η ανάμιξη χρωμάτων παράγει ολοένα και πιο σκούρο χρώμα τείνοντας να γίνει μαύρο. Πρακτικά το αφαιρετικό μοντέλο είναι όταν ο ζωγράφος αναμιγνύει στην παλέτα του



6-8: Το αφαιρετικό σύστημα χρωμάτων

όταν ο ζωγράφος αναμιγνύει στην παλέτα του μπλε και κίτρινο χρώμα για να φτιάξει το πράσινο. Η σύγχυση γίνεται συχνά επειδή ως αφαίρεση είναι η απορρόφηση συγκεκριμένων συχνοτήτων από το φάσμα του φωτός, από ένα υλικό. Έτσι το νέο υλικό που προκύπτει από την ανάμιξη μπλε και κίτρινης μπογιιάς απορροφάει όλα τα φάσματα του φωτός εκτός από το πράσινο που αντανακλάται πίσω. Ένα τέτοιο μοντέλο είναι το CMYK των εκτυπωτών όπου τα βασικά χρώματα είναι κυανό C, μαγεντιανό κόκκινο (magenta) M, κίτρινο Y και μαύρο K.

Στο προσθετικό μοντέλο αντίστοιχα προστίθενται φάσματα για την παραγωγή μιας απόχρωσης και είναι ακριβώς ο τρόπος με τον οποίο λειτουργούν οι οθόνες των υπολογιστών. Γι αυτόν ακριβώς τον λόγο η καταγραφή των εικόνων γίνεται με το προσθετικό μοντέλο. Η επιλογή του κόκκινου, πράσινου και μπλε ως βασικών χρωμάτων ανάγεται στο James Clerk Maxwell, που παρήγαγε την πρώτη τριχρωματική έγχρωμη φωτογραφία με απώτερο σκοπό να επιδείξει την τριπλή φύση της χρωματικής όρασης (color vision). Κυρίως όμως οφείλεται στο τρόπο που λειτουργεί το ανθρώπινο μάτι, που όπως θα γνωρίζουμε αναλυτικότερα στο επόμενο κεφάλαιο, αντιλαμβάνεται το φως έχοντας τριών ειδών διαφορετικά κύτταρα, ευαίσθητα στο κόκκινο, πράσινο και μπλε [90]. Η βιομηχανική δηλαδή ήδη έχει εφαρμοστεί στους υπολογιστές με την χρήση του χρωματικού μοντέλου RGB για την καταγραφή και απεικόνιση της εικόνας.

Το μοντέλο HSL και ο τροχός των χρωμάτων



6-10: Τροχός των χρωμάτων

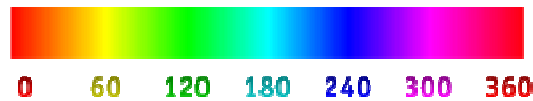
Πηγαίνοντας παρακάτω στα προαπαιτούμενα για την υλοποίηση του Horasis, βλέπουμε το προσθετικό σύστημα χρωμάτων HSL (Hue, Saturation, Lightness) που έχει ως χρωματικές συνιστώσες τις την απόχρωση H, τον κορεσμό S και την φωτεινότητα L [91]. Η απόχρωση που στην ουσία ταυτίζεται διαισθητικά με το φάσμα του χρώματος, είναι αυτή που θα χρησιμοποιηθεί για την παραγωγή των στόχων εκπαίδευσης στους πρώτους εκπαιδευόμενους τεχνητούς νευρώνες του συστήματος. Με την απόχρωση τοποθετείται το χρώμα μέσα σε ένα κύκλο διαφορετικών αποχρώσεων, τον

τροχό των χρωμάτων (color wheel). Τα χρώματα ακολουθούν αυτά του φάσματος έτσι ξεκινώντας από το κόκκινο και με αντίστροφη φορά από αυτήν των δεικτών του ρολογιού (counter-clockwise) πηγαίνουμε προς το ιώδες κλείνοντας τον κύκλο ξανά στο κόκκινο.



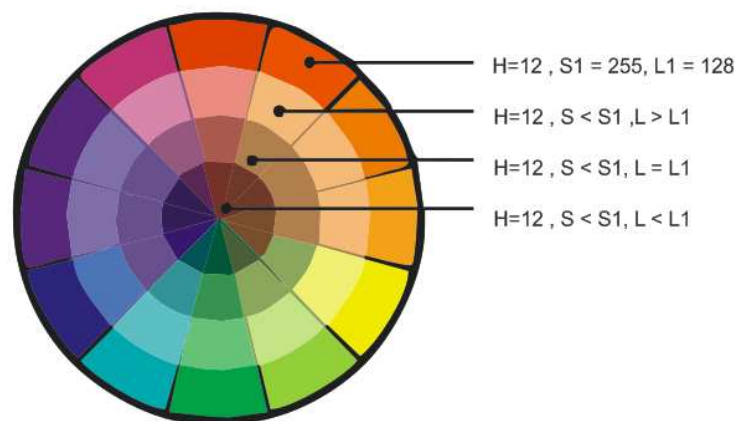
6-11: Τα χρώματα του ορατού φάσματος από το υψηλότερο φάσμα προς το χαμηλότερο

Μπορούμε να δώσουμε στην απόχρωση μια τιμή από το 0 έως το 360 που θα αντιπροσωπεύει τις μοίρες της στον τροχό των χρωμάτων, πάντα σε αντίθετη φορά ρολογιού. Για ευκολία όμως θα θεωρήσουμε ότι και οι τις τρεις χρωματικές συνιστώσες παίρνουν ακέραιες τιμές από 0 έως 255 που μπορούν εύκολα να αναπαρασταθούν ως πληροφορία 8bit. Αυτό μειώνει μεν κλίμακα τιμών της απόχρωσης αλλά αυξάνει δε την ευκολία στην απεικόνιση της στο δυαδικό σύστημα. Παράλληλα η χρήση 8bit τιμών μειώνει τον απαιτούμενο χώρο για την αποθήκευση της.



6-12: Τιμές απόχρωσης H από 0 έως 360 στο χρωματικό μοντέλο HSL

Ο κορεσμός είναι μια ποσότητα που όσο μικραίνει τόσο το χρώμα γίνεται πιο μουντό και πλησιάζει προς μια γκρι απόχρωση. Η φωτεινότητα που συνδέεται με το πόσο φωτεινό είναι ένα χρώμα, είναι μια τιμή που όσο ανεβαίνει τόσο το χρώμα ξανοίγει προς το άσπρο, ενώ όσο κατεβαίνει σβήνει προς το μαύρο. Παρακάτω μπορούμε να δούμε ένα παράδειγμα με χρώματα της ίδιας απόχρωσης 12 σε κλίμακα 255, με μείωση του κορεσμού σε κάθε περίπτωση και αυξομείωση της φωτεινότητας [92].



6-13: Μετατόπιση κορεσμού και φωτεινότητας στο μοντέλο HSL

Το HSL μοντέλο είναι ένα πολύ ενδιαφέρον μοντέλο καθώς η σημασία του είναι εντελώς διαφορετική από αυτήν του RGB. Σε αυτήν την εργασία μας ενδιαφέρει μόνο η απόχρωση H, αλλά στο μέλλον θα ήταν ενδιαφέρον να εξεταστεί περαιτέρω και ο κορεσμός S καθώς εμφανίζεται στην φύση. Ένα χαρακτηριστικό παράδειγμα είναι από τα φυτά, όπου το άνθος εμφανίζει χρώματα με μεγαλύτερο κορεσμό από τα υπόλοιπα μέρη του



6-14: Τιμές κορεσμού σε ασπρόμαυρο από μια φωτογραφία ενός τριαντάφυλλου. Φαίνεται ως φωτεινότερη η περιοχή του άνθους έχοντας μεγαλύτερο κορεσμό από τα φύλλα και το κοτσάνι.

Μετατροπές μεταξύ μοντέλων RGB και HSL

Για τα δύο μοντέλα RGB και HSL έχουν δημιουργηθεί αλγόριθμοι που επιτρέπουν την μετατροπή των χρωματικών συνιστωσών του ενός σε αυτές του άλλου [93]. Θα θεωρήσουμε ότι υπάρχουν έξι συναρτήσεις σε δύο ομάδες, που η κάθε μια από αυτές θα υπολογίζει μια χρωματική συνιστώσα του ενός μοντέλου ως συνάρτηση όλων των υπολοίπων του άλλου:

- Για μετατροπή HSL σε RGB: $R = f_R(H, S, L)$, $G = f_G(H, S, L)$, $B = f_B(H, S, L)$
- Για μετατροπή RGB σε HSL: $H = f_H(R, G, B)$, $S = f_S(R, G, B)$, $L = f_L(R, G, B)$

Για τις συναρτήσεις μετατροπής θεωρούμε ότι $R, G, B, H, S, L \in [0, 1] \subseteq \mathbb{R}$ δηλαδή ότι το πεδίο ορισμού και το σύνολο τιμών είναι το ίδιο διάστημα. Γι αυτό πρέπει να μετατρέψουμε την κλίμακα των τιμών διαιρώντας δια της μέγιστης τιμής 255 (ή 360 για το H) έχοντας έναν δεκαδικό αριθμό από το 0 έως το 1.

Ο αλγόριθμος μετατροπής από HSL σε RGB [94] που παρατίθεται παρακάτω ως ψευδοκώδικας, υλοποιεί και τις τρεις επιμέρους συναρτήσεις f_R, f_G, f_B σε μια διαδικασία:

```
// Μετατροπή από HSL σε RGB χρωματικές συνιστώσες
διαδικασία HSL-Σε-RGB (H, S, L) υπολογίζει R, G, B
έναρξη
  αν S = 0 τότε
    αρχή
      R := L; G := L; B := L;
    τέλος
  ειδάλλως
    αρχή
      αν L <= 0.5 τότε
        B := L * ( 1 + S )
      ειδάλλως
        B := (L + S) - (L * S);

      A := (2 * L) - B;
      R := Απόχρωση-Σε-Χρωματική-Συνιστώσα(H + (1/3), A, B);
      G := Απόχρωση-Σε-Χρωματική-Συνιστώσα(H, A, B);
      B := Απόχρωση-Σε-Χρωματική-Συνιστώσα(H - (1/3), A, B);
    τέλος;
  τέλος;
//-----
συνάρτηση Απόχρωση-Σε-Χρωματική-Συνιστώσα(NewH, A, B)
, αποτέλεσμα: πραγματικός;
έναρξη
  αν NewH < 0 τότε
    NewH := NewH + 1
  ειδάλλως αν NewH > 1 τότε
    NewH := NewH - 1;

  αν (6 * NewH) < 1 τότε
    αποτέλεσμα := A + (B - A) * NewH * 6
  ειδάλλως αν (2 * NewH) < 1 τότε
    αποτέλεσμα := B
  ειδάλλως αν (3 * NewH) < 2 τότε
    αποτέλεσμα := A + (B - A) * (2/3 - NewH) * 6
  ειδάλλως
    αποτέλεσμα := A;
  τέλος;
```

Ο αλγόριθμος μετατροπής από RGB σε HSL που χρησιμοποιείται στο Horasis υπολογίζει και αυτός τις συναρτήσεις f_H, f_S, f_L μέσα στην ίδια διαδικασία:

```
// Μετατροπή από RGB σε HSL χρωματικές συνιστώσες
διαδικασία RGB-Σε-HSL (R,G,B) υπολογίζει H,S,L
έναρξη
  Max := μέγιστο των {R, G, B};
  Min := ελάχιστο των {R, G, B});

  L := (Max + Min) / 2;

  αν Max = Min τότε
  αρχή
    H := 0; S := 0;
  τέλος
  ειδάλλως
  αρχή
    A := Max - Min;

    αν L < 0.5 τότε
      S := A / (Max + Min)
    ειδάλλως
      S := A / (2 - Max - Min);

    αν R = Max τότε
      H := (G - B) / A
    ειδάλλως αν G = Max τότε
      H := 2 + (B - R) / A
    ειδάλλως
      H := 4 + (B - G) / A;

    H := H / 6;
    αν H < 0 τότε
      H := H + 1;
  τέλος;
τέλος;
```

Πειραματικά διαπιστώθηκε ότι η διαδικασία μετατροπής από RGB σε HSL δεν είναι αντιστρέψιμη, δηλαδή από τις τιμές H,S,L που παράγονται δεν μπορούμε να πάρουμε ξανά τις αρχικές R,G,B. Αυτό μπορεί να γραφτεί ως εξής:

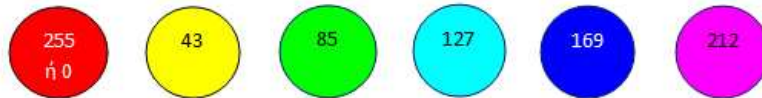
$$f_R(f_H(R, G, B), f_S(R, G, B), f_L(R, G, B)) \langle \rangle R$$

$$f_G(f_H(R, G, B), f_S(R, G, B), f_L(R, G, B)) \langle \rangle G$$

$$f_B(f_H(R, G, B), f_S(R, G, B), f_L(R, G, B)) \langle \rangle B$$

Τα αμιγή χρώματα και οι τιμές της απόχρωσης τους

Ως **Αμιγές Χρώμα (Pure Color)** θα ονομάζουμε ένα χρώμα που έχει μια ή δύο από τις χρωματικές συνιστώσες R,G,B στην μέγιστη τιμή της και όλες τις υπόλοιπες στην ελάχιστη, θεωρώντας το ως το «τέλειο» πρότυπο για την έννοια χρώματος που αντιπροσωπεύει. Έτσι για την έννοια «κόκκινο» που χαρακτηρίζει μια μεγάλη περιοχή του φάσματος, θεωρούμε «πιο κόκκινο» από όλα το απόλυτο χρώμα με R,G,B = [255, 0, 0]. Έτσι θα έχουμε 6 τέτοια χρώματα που θα υπολογίσουμε τις τιμές των αποχρώσεων (hue) για αυτά, με μετατροπή RGB σε HSL, σε μια κλίμακα από 0 έως 255:



6-15: Τα έξι αμιγή χρώματα

Αν βέβαια στηριζόμασταν στο τι θεωρούν οι περισσότεροι ως το «τέλειο» πρότυπο χρώματος που συνδέεται με μια έννοια, τότε οι τιμές RGB μπορεί να διαφέρουν. Ένα τέτοιο ιδιότυπο πείραμα μαζικής ψυχολογίας έχει γίνει στο internet [95] και αν δούμε τα αποτελέσματα οι τιμές RGB που οι άνθρωποι θεωρούν απλώς «κόκκινο», «πράσινο», «μπλε» και «κίτρινο» έχουν ομοιότητες και διαφορές.

purple (#7e1e9c)	green (#15b01a)	blue (#0343df)	pink (#ff81c0)	brown (#653700)	red (#e50000)
light blue (#95d0fc)	teal (#029386)	orange (#f97306)	light green (#96f97b)	magenta (#c20078)	yellow (#ffff14)
sky blue (#75bbfd)	grey (#929591)	lime green (#89fe05)	light purple (#bf77f6)	violet (#9a0eea)	dark green (#033500)
turquoise (#06c2ac)	lavender (#c79fef)	dark blue (#00035b)	tan (#d1b26f)	cyan (#00ffff)	aqua (#13eac9)
forest green (#06470c)	mauve (#ae7181)	dark purple (#35063e)	bright green (#01ff07)	maroon (#650021)	olive (#6e750e)
salmon (#ff796c)	beige (#e6daa6)	royal blue (#0504aa)	navy blue (#001146)	lilac (#cea2fd)	black (#000000)
hot pink (#ff028d)	light brown (#ad8150)	pale green (#c7fdb5)	peach (#ffb07c)	olive green (#677a04)	dark pink (#cb416b)
periwinkle (#8e82fe)	sea green (#53fca1)	lime (#aaff32)	indigo (#380282)	mustard (#ceb301)	light pink (#ffd1df)

6-16: Χρώματα από έρευνα XKCD. Κάθε χρώμα έχει τον αντίστοιχο τίτλο που ανέφεραν οι συμμετέχοντες στην έρευνα και τον αντίστοιχο κωδικό χρώματος HTML⁵⁹ (xkcd.com)

⁵⁹ Ο κωδικός χρώματος HTML είναι στην ουσία η δεκαεξαδική αναπαράσταση των τριών αριθμών R,G,B 0..255

Συγκρίνοντας την προηγούμενη ταξινόμηση με την θεώρηση των αμιγών χρωμάτων παρατηρούμε ομοιότητα στην απόχρωση για το κόκκινο και κίτρινο, μια μικρή απόκλιση κατά 1 στο πράσινο και μια αρκετά μεγάλη απόκλιση κατά 9 μονάδες στην κλίμακα των 255 για το μπλε. Τα τέσσερα αμιγή χρώματα που αναφέραμε και οι αποχρώσεις τους μας ενδιαφέρουν ιδιαίτερα στο Horasis, καθώς η κάθε μια από αυτές τις αποχρώσεις, όπως θα δούμε παρακάτω, δίνει την μέγιστη τιμή στον στόχο εκπαίδευσης ενός νεурώνα.

Μετατροπή από χρώμα RGB σε μετρήσιμη ποσότητα φωτός

Για να ολοκληρώσουμε τα απαιτούμενα στοιχεία που αντλούνται από την χρωματική θεωρία για την υλοποίηση του Horasis, θα δούμε πως από τις τρεις τιμές του κόκκινου, πράσινου και μπλε για κάθε εικονοστοιχείο, υπολογίζουμε μια ποσότητα σε μια τιμές τις σκάλας του γκρι (grayscale). Η τιμή αυτή που αντιπροσωπεύει την ποσότητα του συνολικού άσπρου φωτός και διαφέρει από την φωτεινότητα L του μοντέλου HSL που είδαμε πριν. Αυτή η πληροφορία ονομάζεται luma, που στα ελληνικά μπορούμε να μεταφράσουμε απλώς ως φως και είναι αναλογική με την γάμμα-διορθωμένη (gamma corrected) φωτεινότητα ενός χρώματος [96], διαφοροποιούμενη έτσι από την φωτεινότητα lightness του μοντέλου HSL.

Αυτή η ποσότητα χρησιμοποιήθηκε για την αναλογική μετάδοση της ασπρόμαυρης τηλεόρασης στην Βόρεια Αμερική, ως η συνιστώσα Y του χρωματικού συστήματος YIQ NTSC. Ο τύπος υπολογισμού του Y σε αυτό το σύστημα, αλλά και σε άλλα συστήματα που χρησιμοποιούν luma [96] είναι

$$Y = 0.29889531 * R + 0.58662247 * G + 0.11448223 * B \quad (6.1)$$

ή στρογγυλοποιημένα

$$Y = 0.299 * R + 0.587 * G + 0.114 * B \quad (6.2)$$

όπου $R, G, B \in [0, 255] \subseteq \mathbb{N}$

Έτσι μπορούμε κατά έναν τρόπο να επιστρέψουμε κατά προσέγγιση στην αρχική πληροφορία που είναι η ποσότητα του φωτός, δηλαδή πόσα φωτόνια, έφτασαν πάνω στον αισθητήρα CCD/CMOS του ψηφιακού μέσου απεικόνισης, πριν τον διαχωρισμό τους σε κόκκινο, πράσινο και μπλε και πολύ πριν την ψηφιοποίηση σε 24bits. Έχοντας γνωρίσει τα φίλτρα Bayer, βλέπουμε ότι υπάρχει διαφορά στην κατανομή των χρωμάτων σε σχέση με τον παραπάνω τύπο, όπου υπάρχει περίπου 30% κόκκινο, 59% πράσινο και 11% μπλε, καθώς υπάρχουν 50% πράσινες 25% κόκκινες και 25% μπλε κυψέλες σε κάθε επανάληψη του μοτίβου. Παρόλαταυτα οι σταθεροί παράγοντες του τύπου υπολογισμού του Y, βολεύουν για την εφαρμογή του Horasis, καθώς βρίσκονται πιο κοντά στην βιολογική λειτουργία του ματιού, όπως θα γνωρίσουμε στο επόμενο κεφάλαιο που ασχολείται με την φυσιολογία της ανθρώπινης όρασης.

Δειγματοληψία των εικόνων στο Horasis

Η ρύθμιση των δύο καμερών στην σκηνής φωτογράφισης.



6-17: Δειγματοληψία οπτικών προτύπων για ένα αντικείμενο στο Horasis

Μετά την θεωρία και τεχνολογία που κρύβεται πίσω από την ψηφιοποίηση της εικόνας, μπορούμε να δούμε την διαδικασία παραγωγής οπτικών προτύπων για τα πραγματικά αντικείμενα μέσα στο σύστημα Horasis. Καταρχήν το πραγματικό αντικείμενο, όπως μια κούπα του καφέ που βλέπουμε αριστερά, τοποθετείται στο βάθος της σκηνής φωτογράφισης στο οποίο στοχεύουν ήδη οι δύο web κάμερες, έχοντας μια σταθερή ρύθμιση. Η ρύθμιση περιλαμβάνει τις συντεταγμένες που έχουν οι κάμερες πάνω στο οριζόντιο επίπεδο, οι οποίες

καθορίζονται με την βοήθεια ενός καναλιού-οδηγού. Αυτό προσομοιώνει την σύγκλιση και απόκλιση των ματιών για την εστίαση σε κοντινά ή μακρινά αντικείμενα. Επιπρόσθετα υπάρχει μια περιστροφή των δύο καθέτων αξόνων κατά κάποιες μοίρες, για την μεν αριστερή κάμερα δεξιόστροφα και για την δε δεξιά αριστερόστροφα, ώστε η κάθε μια από τις δύο κάμερες να στοχεύει το αντικείμενο. Η όλη ρύθμιση έγινε εντελώς εμπειρικά με μοναδικό βοήθημα έναν στυλό κατάδειξης laser ώστε να εμφανίζεται ένα σταθερό σημείο και στις δύο φωτογραφίες. Προσπαθώντας να τοποθετηθεί η κουκίδα του laser στο κέντρο και των δύο λήψεων της αριστερής και δεξιάς κάμερας, γινόνταν οι αντίστοιχες μετακινήσεις και περιστροφές. Μελλοντικά θα μπορούσε η διαδικασία να υλοποιείται με έναν επιστημονικό τρόπο, εφόσον η κατασκευή επαναληφθεί με την ακρίβεια χιλιοστού της μηχανολογίας, ώστε να προϋπάρχει η απαιτούμενη ακρίβεια και για τις ρυθμίσεις.

Το όλο σύστημα στήριξης των καμερών θα μπορούσε να υλοποιηθεί μελλοντικά ως αυτοματισμός με την χρήση αξόνων, συνδέσμων, βηματικών κινητήρων και γρاناζιών για την μεταφορά της κίνησης [97]. Για την εκκίνηση ενός τέτοιου έργου, δεν απαιτείται κατ' ανάγκην πρόσβαση σε ακριβά εργαλεία μηχανολογικών κατασκευών ή χρονοβόρα παραγγελιά κομματιών σε μηχανουργείο καθώς το πρωτότυπο μπορεί κάλλιστα να υλοποιηθεί με έτοιμα κομμάτια Lego Mindstorms⁶⁰ [98]. Θα ήταν πραγματικά ένα ενδιαφέρον πεδίο συνεργασίας μεταξύ Τεχνολόγων Μηχανικών Υπολογιστών και Τεχνολόγων Μηχανικών Αυτοματισμού, η υλοποίηση ενός τέτοιου συστήματος αυτομάτου ελέγχου.

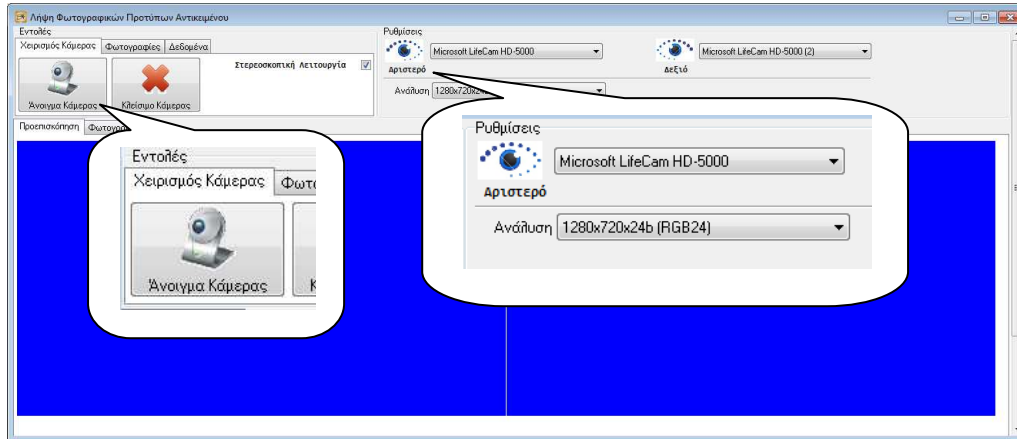


6-18: Λεπτομέρεια στήριξης καμερών στην σκηνή φωτογράφισης του Horasis

⁶⁰ Τα Lego Mindstorms είναι κατασκευές της γνωστής από τα ομώνυμα παιχνίδια εταιρίας, μόνο που πλέον εκτός από τα κλασικά «τουβλάκια» τα κομμάτια που συνδυάζονται είναι γρανάζια, ρόδες, αισθητήρες, σερβομηχανισμοί μέχρι και ένα τουβλάκι-μικροϋπολογιστής το NXT.

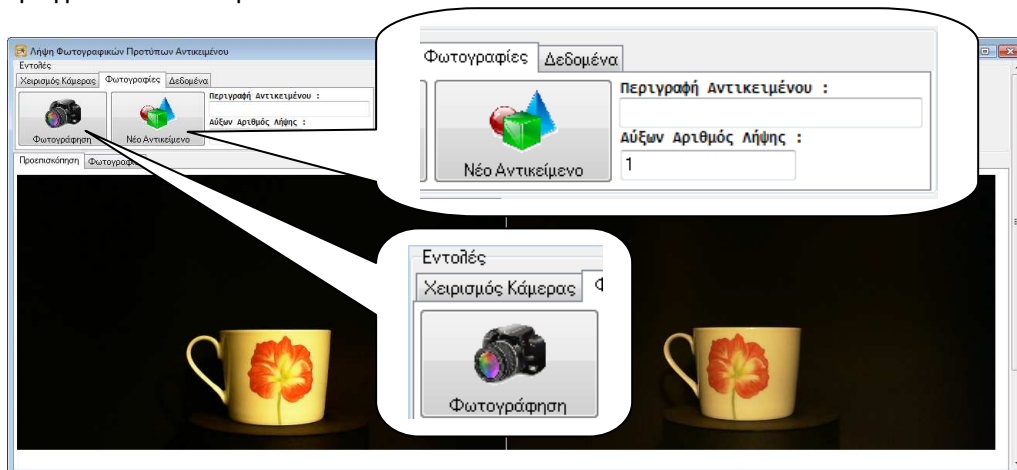
Η φόρμα της λήψης φωτογραφιών

Στο προγραμματιστικό μέρος της εργασίας και μέσα στην εφαρμογή Horasis, υπάρχει η φόρμα «Λήψη Φωτογραφική Προτύπων Αντικειμένου» που ανοίγει με την εντολή «Λήψη Εικόνας» της ομάδας «Δειγματοληψία» της κεντρικής λωρίδας εντολών. Σε αυτήν υπάρχουν δύο παράθυρα για την προεπισκόπηση της λήψης της φωτογραφίας μέσω ζωντανής ροής βίντεο του DirectShow.



6-19: Η φόρμα λήψης φωτογραφιών στην εφαρμογή Horasis

Με την εντολή «Άνοιγμα Κάμερας» ξεκινάει η προεπισκόπηση ενώ με το «Κλείσιμο Κάμερας» σταματάει. Δίπλα υπάρχουν δύο ρυθμίσεις για επιλογή μιας συσκευής απεικόνισης των Windows ως αριστερό ή δεξί «μάτι» του συστημάτων, δίνοντας την δυνατότητα της σωστής αντιστοίχισης κάμερας με θέση και της ανάλυσης στην οποία θα ληφθούν οι εικόνες. Αρχίζοντας την προεπισκόπηση οι κάμερες ξεκινούν την αυτόματη εστίαση (focus) πάνω στο πραγματικό αντικείμενο.



6-20: Προεπισκόπηση αντικειμένου πριν την φωτογράφιση, στην φόρμα φωτογράφισης

Εφόσον κουνηθεί, αντικατασταθεί ή αλλάξει ο φωτισμός οι κάμερες μπορεί να εστιάσουν ξανά. Όπως βλέπουμε έχει ενδιαφέρον ότι τα δύο πρότυπα του πραγματικού αντικειμένου διαφέρουν μεταξύ τους, εμφανίζοντας και οι δύο συνδυαστικά μεγαλύτερο τμήμα της τρισδιάστατης επιφάνειας του αντικειμένου από ότι κάθε μια ξεχωριστά. Πρακτικά έχουμε την χρήση στερεοσκοπίας, αφού η χρησιμοποιούμενη τεχνική ενισχύει την αντίληψη του βάθους της εικόνας χρησιμοποιώντας δύο λήψεις υπό διαφορετική γωνία μια για το αριστερό και μια για το δεξί τεχνητό μάτι [99].

Με την χρήση της εντολής «Νέο Αντικείμενο» επανεκκινεί ο αύξων αριθμός λήψης και καθαρίζει η περιγραφή αντικείμενου ώστε να γράψουμε τον τίτλο του πραγματικό αντικείμενο που φωτογραφίζεται. Πλέον με την εντολή «Φωτογράφιση» καταγράφεται για κάθε μια κάμερα ένα σταθερό καρέ (still frame) της ροής βίντεο DirectShow και αποθηκεύεται αρχικά στην μνήμη σε διαμόρφωση BMP 24bit.



6-21: Ζεύγος φωτογραφιών, αριστερής και δεξιάς κάμερας ως πρότυπο του αντικείμενου.

Η σχεσιακή βάση δεδομένων των προτύπων

Αμέσως μετά την λήψη των εικόνων στην μνήμη σε ψηφιδωτή μορφή, αποθηκεύονται σε μια σχεσιακή βάση δεδομένων [8] της εφαρμογής, ως μια εγγραφή σε έναν πίνακα PHOTO. Στα BLOB πεδία της εγγραφής LEFT_PHOTO και RIGHT_PHOTO μεταφέρονται αυτούσιες οι διαμορφώσεις BMP, της αριστερής και της δεξιάς εικόνας αντίστοιχα. Η υλοποίηση αυτή σε Delphi, όπως και όλες οι παρακάτω υλοποιήσεις διαχείρισης δεδομένων, στηρίζεται πάνω στην δυνατότητα που δίνει η βιβλιοθήκη VCL για διασύνδεση με το υποσύστημα ADO⁶¹/OleDb [10] των Windows μέσω κλάσεων όπως TADOConnection, TADOQuery[100]. Μέσω του ADO συνδεόμαστε με έναν Microsoft SQL Server⁶² [9], έκδοσης 2008 Express, όπου μπορούμε να εκτελέσουμε εντολές SQL⁶³[101] στέλνοντας και ανακτώντας δεδομένα από αυτόν.

Έχοντας την αποθήκευση των φωτογραφιών ανεξάρτητη, τοποθετώντας την αποκλειστικά στο επίπεδο δεδομένων (data layer), ο σχεδιασμός παρέχει μια ανεξαρτησία από το σύστημα βάσης δεδομένων που φιλοξενεί τα δεδομένα του Horasis. Έτσι μπορεί να χρησιμοποιηθεί στην θέση του MSSQL οποιοδήποτε σύστημα διαχείρισης βάσεων δεδομένων RDBMS παρέχει παρόχους (providers) για τις τεχνολογίες ADO/ OleDb ή ακόμα και για την παλαιότερη ODBC⁶⁴. Επίσης το RDBMS μπορεί να βρίσκεται σε κάποιον φυσικά διαφορετικό υπολογιστή που συνδέεται μέσω δικτύου ή σε κάποιον κεντρικό εξυπηρετητή (database server).

⁶¹ ADO = Active Data Objects

⁶² Ο Microsoft SQL Server αναφέρεται και για συντομία ως MSSQL

⁶³ SQL = Structure Query Language

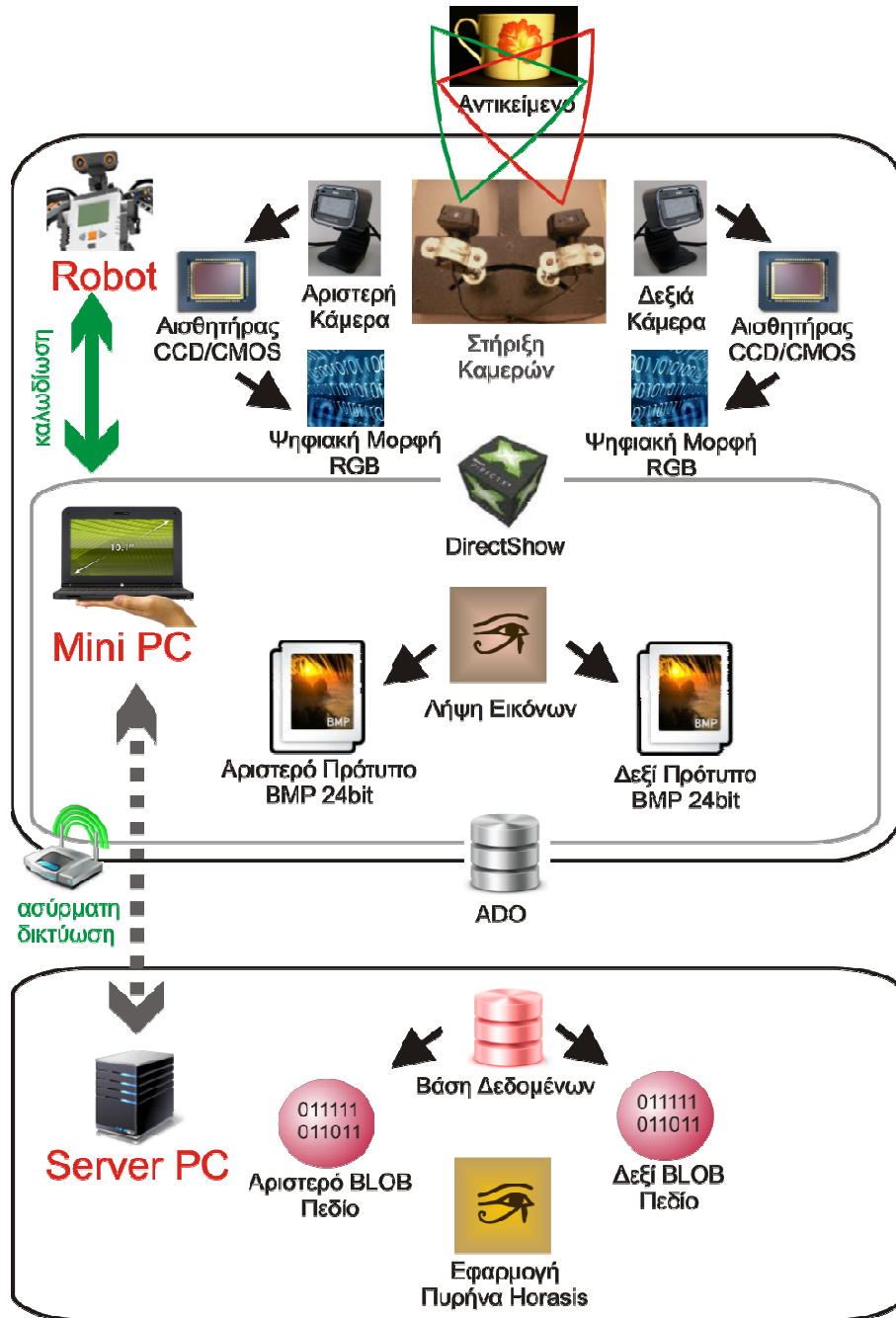
⁶⁴ ODBC = Open DataBase Connectivity

Με αυτόν τον τρόπο δημιουργείται η δυνατότητα να αποσπαστεί το τμήμα του λογισμικού που αφορά την φωτογράφιση και να τοποθετηθεί σε διαφορετικό υλικό από την υπόλοιπη υλοποίηση. Αυτό το υλικό μπορεί να είναι υψηλότερης φορητότητας και χαμηλότερης κατανάλωσης ρεύματος, αλλά παράλληλα συνήθως είναι μικρότερης επεξεργαστικής ισχύος⁶⁵. Καθώς δεν χρειάζεται να τρέχει το υπόλοιπο σύστημα, που απαιτεί μεγάλη επεξεργαστική ισχύ, η φωτογράφιση ως αυτόνομο λογισμικό μπορεί να χαρακτηριστεί λύση thin-client [102]. Έτσι θα μπορούσε να τρέξει μελλοντικά πάνω σε έναν μικροϋπολογιστή PC που κουβαλάει πάνω του ένα ρομπότ, στο οποίο θα περιλαμβανόταν και ο αυτοματισμός για την κίνηση των δύο καμερών. Για την επεξεργασία των λήψεων θα υπάρχει ένα Horasis server, στον οποίο θα τρέχει ο πυρήνας του συστήματος.

⁶⁵ Ένα τέτοιο σύστημα που μπορεί να τρέξει Windows, μπορεί να βασίζεται μητρική κάρτα Intel D410PT με επεξεργαστή Intel Atom, κόστος 60€ και κατανάλωση μόλις 40W σε πλήρη λειτουργία [160].

Σχηματική αναπαράσταση του συστήματος λήψης οπτικών προτύπων του Horasis

Συμπληρώνοντας την αρχιτεκτονική που γνωρίσαμε σε αυτό το κεφάλαιο, που αποτελεί το πρώτο και αυτόνομο τμήμα του λογισμικού Horasis, μπορούμε να δημιουργήσουμε το παρακάτω σχεδιάγραμμα. Υπάρχουν δύο φυσικά διαχωρισμένοι υπολογιστές, ο mini pc τοποθετημένος πάνω σε ένα ιδεατό ρομπότ τροφοδοτούμενος από τις μπαταρίες του και ο server pc που τρέχει τον πυρήνα της εφαρμογής Horasis και το RDBMS. Τα δύο PC συνδέονται μέσω ασυρμάτου δικτύου παρέχοντας ανεξαρτησία στην κίνηση του ρομπότ στον χώρο. Ο πρώτος υπολογιστής είναι χαμηλής κατανάλωσης ρεύματος, συνδέεται καλωδιακά με τις κάμερες και τον αυτοματισμό με τον οποίο κινούνται.



6-22: Μελλοντική υλοποίηση Horasis ως thin-client/server και η θέση του τμήματος I της αρχιτεκτονικής Horasis που υλοποιεί την λήψη φωτογραφιών

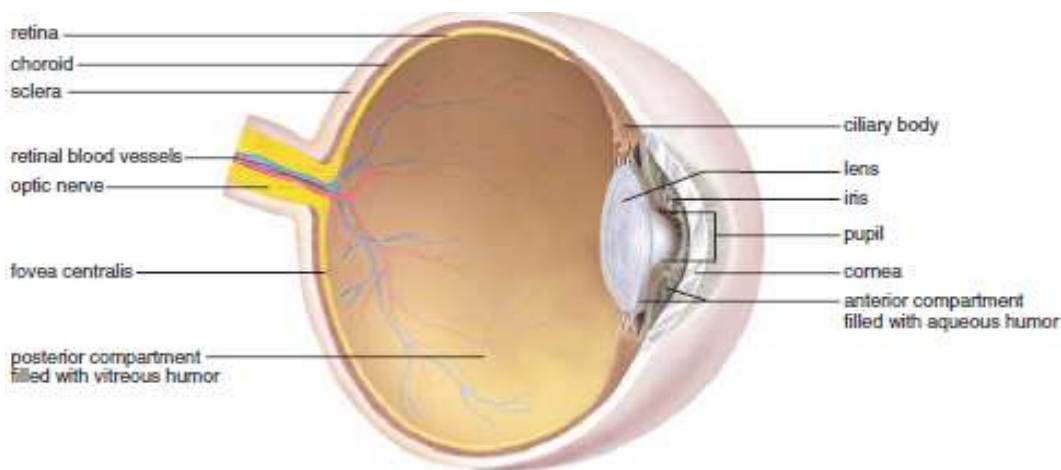
7. Η Ανθρώπινη Όραση

Όπως ήδη αναφέραμε το Hovasis στηρίζεται στην ιδέα της βιομμητικής, με απλά λόγια προσπαθεί να μοντελοποιήσει με τον σχεδιασμό του αυτό που συμβαίνει στην φύση. Γι αυτό προηγήθηκε της υλοποίησης, μια επισταμένη έρευνα πάνω στην βιολογία, φυσιολογία [103], και νευροεπιστήμη με την απλοϊκή ματιά ενός επιστήμονα υπολογιστών. Σίγουρα οι πληροφορίες που ανακαλύφθηκαν μέσω της έρευνας, θα μπορούσαν να είναι πολύ πιο ακριβείς αν στην εργασία συμμετείχαν ειδικοί των παραπάνω επιστημών. Με μια προσέγγιση όσο το δυνατόν πιο κοντά στην πραγματικότητα και την πραγμάτωση της ως ένα λογισμικό που τρέχει, γίνεται παράλληλα και μια προσπάθεια να επιδειχθεί η προοπτική μια τέτοιας συνεργασίας. Παρακάτω θα δούμε κάποια στοιχεία της έρευνας που ενδιαφέρουν στην υλοποίηση του Hovasis και είναι απαραίτητα ως γνώσεις για την κατανόηση της εργασίας. Γνωρίζοντας το βιολογικό σύστημα της ανθρώπινης όρασης θα δούμε τα τρία μέρη που συνεργάζονται για την λειτουργία της αίσθησης: Τους οφθαλμούς, τα δίκτυα των νευρώνων του ματιού και ένα τμήμα του εγκεφάλου που λέγεται και πρωτοταγής οπτικός φλοιός.

Το Ανθρώπινο Μάτι

Γενική περιγραφή του οφθαλμού

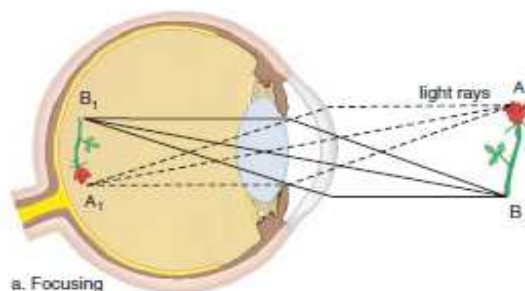
Τα μάτια είναι το αισθητήρια όργανα της ανθρώπινης όρασης. Ένας οφθαλμός δέχεται ερεθίσματα φωτός ενός φάσματος της ηλεκτρομαγνητικής ακτινοβολίας που βρίσκεται σε μήκη κύματος μεταξύ 400 και 780 nm. Ως όργανα δέχονται μια είσοδο από φωτόνια και παράγουν ως έξοδο έναν ηλεκτρικό νευρωνικό παλμό[104]. Βρίσκονται τοποθετημένοι σε μια κοιλότητα του κρανίου την οφθαλμική κόγχη και έχουν σχεδόν σφαιρικό σχήμα με διάμετρο περίπου 25mm, εξ' ου και η ονομασία βολβός που περιγράφει το κυρίως μέρος του οργάνου. Ο κάθε βολβός περιστρέφεται με την βοήθεια μυών που λειτουργούν συντονισμένα και για τα δύο μάτια συγκλίνοντας και αποκλίνοντας ανάλογα με την απόσταση ενός αντικειμένου. Αποτελείται από τρία ομόκεντρα στρώματα ιστού ή χιτώνες:



7-1: Ο βολβός του ανθρώπινου οφθαλμού (Mader) [105]

- τον ινώδη χιτώνα που είναι στο εξώτερο τμήμα του ματιού και αποτελείται από:
 - το κεντρικό τμήμα, τον κερατοειδή χιτώνα (cornea) που είναι διαφανές
 - το υπόλοιπο τμήμα του βολβού, τον σκληρό χιτώνα (sclera) που είναι άσπρος
- τον χοριοειδή χιτώνα (choroid) που είναι αδιαφανής
- τον αμφιβληστροειδή χιτώνα (retina) που βρίσκεται στο εσωτερικό του ματιού

Στον χοριοειδή χιτώνα που βρίσκεται στο ενδιάμεσο ανήκει η ίριδα (iris), ένας μυς με όμορφους χρωματισμούς που φαίνονται εξωτερικά και δίνουν το χαρακτηριστικό χρώμα των ματιών. Στο κέντρο της υπάρχει μια κυκλική σχισμή η κόρη (pupil), η οποία διαστέλλεται και συστέλλεται μεταβάλλοντας την διάμετρο της επιτρέποντας περισσότερο ή λιγότερο φως να περάσει προς το εσωτερικό του ματιού. Πίσω από την σχισμή υπάρχει ο κρυσταλλοειδής φακός (lens) με τον οποίο επιτυγχάνεται η εστίαση με την τροποποίηση του σχήματος του, που γίνεται με την βοήθεια ενός ακτινωτού μυ⁶⁶ (ciliary muscle). Η λειτουργία των τμημάτων στα δύο πρώτα στρώματα του ματιού έχει ως σκοπό την εστίαση της εικόνας στον αμφιβληστροειδή χιτώνα και περιγράφεται σε μία πρόταση: Το φως ξεκινάει από τον κερατοειδή, περνάει από την κόρη κατά ένα ποσοστό, διαθλάται από τον φακό και οδηγείται ώστε να φωτίσει τα πρώτα κύτταρα του αμφιβληστροειδή. Η εικόνα κατά αυτήν την διαδικασία σχηματίζεται σε ένα πολύ μικρότερο μέγεθος από το πραγματικό αντικείμενο, αναποδογυρισμένη και ανεστραμμένη από τα αριστερά προς τα δεξιά [105]. Αυτό είναι το ανεστραμμένο είδωλο του πραγματικού αντικειμένου.



7-2: Αναστροφή της πραγματικής εικόνας ενός τριαντάφυλλου (Mader) [105]

Στον αμφιβληστροειδή χιτώνα, όπου σχηματίζεται αυτή η εικόνα, υπάρχουν δύο σημεία που έχουν ενδιαφέρον το κεντρικό βοθρίο (fovea centralis) και η θηλή του οπτικού νεύρου. Το κεντρικό βοθρίο είναι το κέντρο του αμφιβληστροειδούς χιτώνα στο οποίο όπως θα δούμε έχουμε την καθαρότερη όραση. Έτσι μπορούμε να το ταυτίσουμε με το κέντρο της όρασης για κάθε ένα από δύο μάτια. Η θηλή του οπτικού νεύρου είναι το σημείο από το οποίο ξεκινάει το οπτικό νεύρο και λέγεται και τυφλό σημείο του οφθαλμού. Το κομμάτι της εικόνας που πέφτει στο τυφλό σημείο δεν μπορεί να το δει ένα μάτι μας από μόνο του και υπάρχουν απλά πειράματα με το οποίο ο καθένας μας μπορεί να διαπιστώσει την ύπαρξη του [106]. Ο συνδυασμός ενός ζευγαριού ματιών, εξαφανίζει αυτόν τον βιολογικό περιορισμό της όρασης.

⁶⁶ Η μυωπία είναι μια συνήθης παραμόρφωση αυτού του φακού που οφείλεται και στην παρατεταμένη εστίαση σε κοντινά αντικείμενα με υπερβολική χρήση αυτού του ακτινωτού μυ, όπως η εργασία με ηλεκτρονικούς υπολογιστές.

Αντιστοιχία της ψηφιακής κάμερας με τον οφθαλμό.

Θεωρούμε στην εργασία μας ότι την λειτουργία των πρώτων επιπέδων του οφθαλμού για τον σχηματισμό της εικόνας στον αμφιβληστροειδή την προσομοιώνει η ψηφιακή κάμερα. Έτσι για να γίνει εστίαση και παραγωγή της εικόνας, θα χρησιμοποιείται ένα υλικό που θα αποτελείται από

- Την ψηφιακή κάμερα ως τεχνητός βολβός (ίριδα, κόρη, κρυσταλλικός φακός)
- Αυτοματισμούς κίνησης της ψηφιακής κάμερας ως τεχνητοί μύες κίνησης του βολβού

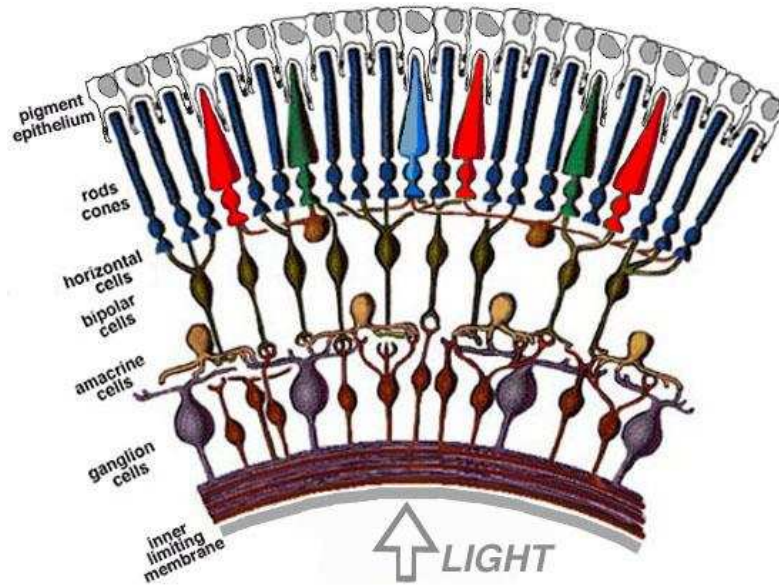
Στην εργασία μας έχουμε διαθέσιμο υλικό για το πρώτο μέρος, ενώ απλοποιείται η λειτουργία του δεύτερου μέρους με την χρήση της στατικής ρύθμισης των καμερών στην σκηνή φωτογράφισης. Οι κάμερες που χρησιμοποιήθηκαν διαθέτουν αυτόματη εστίαση (auto focus) και καθώς είναι στραμμένες προς το βάθρο εστιάζουν σε οποιοδήποτε πραγματικό αντικείμενο τοποθετηθεί πάνω σε αυτό. Έχοντας εξετάσει ήδη την λειτουργία της δειγματοληψίας εικόνα θεωρούμε ότι οι μορφές BMP, που αποθηκεύονται μέσα στα δύο πεδία της βάσης δεδομένων αντιστοιχούν στο ανεστραμμένο είδωλο του αντικειμένου για κάθε ένα από τα δύο μάτια.

Έτσι φτάνουμε επιτέλους στο σημείο που έχουμε ορίσει μια αντιστοιχία για όλα τα εξωτερικά χαρακτηριστικά της όρασης, αλλά και για το ίδιο το οπτικό ερέθισμα. Πλέον μπορούμε αφού γνωρίσουμε την λειτουργία του βιολογικού αμφιβληστροειδούς χιτώνα, να δούμε πως το Horasis προσπαθεί να τον μοντελοποιήσει, υλοποιώντας ένα **Αντικειμενοστραφή Τεχνητό Αμφιβληστροειδή (Object Oriented Artificial Retina - OOAR)**. Θα δώσουμε έναν πρώτο ορισμό ως τον σχεδιασμό αντικειμένων λογισμικού, που μπορεί να περιέχει τεχνητά νευρωνικά δίκτυα, κατά τον οποίο τα αντικείμενα οργανώνονται και λειτουργούν αντίστοιχα με τα κύτταρα που υπάρχουν σε έναν βιολογικό αμφιβληστροειδή χιτώνα. Η υλοποίηση αυτού του σχεδιασμού, μέσα στα πλαίσια του Horasis, παρατίθενται ως ένα πρώτο παράδειγμα λογισμικού τεχνητής όρασης που στηρίζεται στην χρήση του.

Αμφιβληστροειδής Χιτώνας

Τα συστατικά και η γενική συνδεσμολογία του αμφιβληστροειδούς

Ο αμφιβληστροειδής χιτώνας βρίσκεται όπως είδαμε στο εσωτερικό μέρος του βολβού και ως προς την φυσιολογία του χωρίζεται σε στρώσεις κυττάρων, ή στοιβάδες, ξεκινώντας από την πιο απομακρυσμένη μέχρι αυτήν που βρίσκεται πλησιέστερα στον κρυσταλλοειδή φακό[107] [52]. Το φως κατευθυνόμενο από τον φακό καταλήγει στο πίσω μέρος του αμφιβληστροειδούς διερχόμενο από τον χιτώνα. Τα τμήματα του είναι:



7-3: Σχηματική αναπαράσταση του αμφιβληστροειδούς χιτώνα. (Webvision) [107]

- Το μελάγχρουν επιθήλιο (pigment epithelium), που βρίσκεται πίσω από τον αμφιβληστροειδή και καλύπτει τον άσπρο σκληρό χιτώνα. Αποτελείται από κύτταρα που περιέχουν την μαύρη χρωστική ουσία μελανίνη⁶⁷ και υπάρχουν ώστε να απορροφούν το φως που δεν δεσμεύεται από τον αμφιβληστροειδή, ώστε να μην υπάρχουν αντανακλάσεις. Είναι η πιο απομακρυσμένη από τον φακό στοιβάδα (outer).
- Τους φωτοϋποδοχείς (photoreceptors) που είναι κύτταρα ευαίσθητα στο φως και παράγουν ανάλογα με τα φωτόνια που τα χτυπούν νευρικές ώσεις, δηλαδή σήματα προς τους παρακάτω από νευρώνες. Τα κυτταρικά τους σώματα είναι συγκεντρωμένα σε ένα στρώμα, την εξωτερική κοκκώδη στοιβάδα [52] (outer nuclear layer).
- Τους δίπολους (bipolar), οριζόντιους (horizontal) και βραχύνους ή αμακρίνους⁶⁸ (amacrine) νευρώνες που βρίσκονται στην εσωτερική κοκκώδη στοιβάδα (inner nuclear layer)
- Τους γαγγλιακούς (ganglion) νευρώνες που σχηματίζουν την στοιβάδα γαγγλιακών κυττάρων (ganglionic layer), η οποία αποτελεί την έξοδο του βιολογικού νευρωνικού δικτύου του αμφιβληστροειδούς.

⁶⁷ Η μελανίνη είναι μια πολύ συνηθισμένη χρωστική ουσία και παράγεται στο δέρμα μας, όταν το καλοκαίρι μαυρίζουμε στην παραλία.

⁶⁸ Α-μακρύ-ινος είναι ένας νευρώνας που δεν έχει μεγάλες σε μήκος νευρωνικές ίνες δηλαδή βραχείες.

Εκτός από τα κύτταρα οριοθετούνται δύο ακόμη στοιβάδες που περικλείουν τις συνάψεις μεταξύ των στρωμάτων:

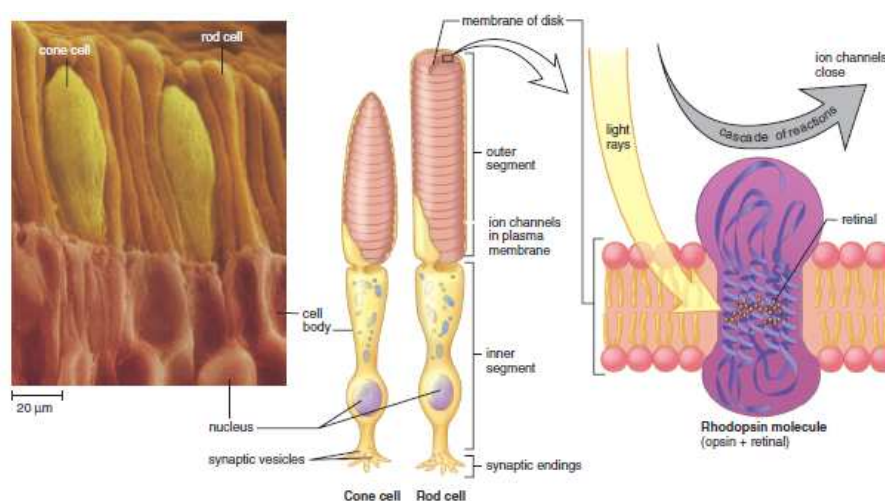
- Η εξωτερική δικτυακή στοιβάδα OPL (outer plexiform layer) στην οποία υπάρχουν συνάψεις:
 - Μεταξύ φωτοϋποδοχέων και της εισόδου των δίπολων νευρώνων.
 - Μεταξύ φωτοϋποδοχέων και οριζοντίων νευρώνων.
 - Μεταξύ δίπολων και οριζόντιων νευρώνων.
 - Μεταξύ δύο διαφορετικών οριζόντιων νευρώνων [108].
- Η εσωτερική δικτυακή στοιβάδα IPL (inner plexiform layer) όπου υπάρχουν συνάψεις :
 - Μεταξύ της εξόδου ενός δίπολου και αμακρύινων νευρώνων
 - Μεταξύ της εξόδου ενός δίπολου νευρώνα και της εισόδου ενός γαγγλιακού νευρώνα
 - Μεταξύ ενός βραχύινου νευρώνα και της εισόδου γαγγλιακών νευρώνων.

Η έξοδος του παραπάνω νευρωνικού δικτύου είναι το οπτικό νεύρο, το οποίο αποτελείται από μεγάλους σε μήκος άξονες των γαγγλιακών νευρώνων οι οποίοι επεκτείνονται μέχρι το πίσω μέρος του εγκεφάλου. Εδώ βλέπουμε και την πρακτική λειτουργία του μυελώδους έλυτρου καθώς όταν εξέρχονται από το μάτι, μετά την θηλή του οπτικού νεύρου, οι άξονες περικλείονται από αυτό [52]. Οι άξονες του οπτικού νεύρου μεταφέρουν τον ηλεκτρικό παλμό προς ένα τμήμα του εγκεφάλου που λέγεται οπτικός φλοιός και θα γνωρίσουμε στις επόμενες ενότητες.

Μας ενδιαφέρει ιδιαίτερα η λειτουργία του κάθε ενός από τους νευρώνες του αμφιβληστροειδούς και θα αναφερθούμε σε κάποιες λεπτομέρειες, ξεχωριστά για τον κάθε ένα, οι οποίες ελήφθησαν υπόψη για τον σχεδιασμό του Horasis.

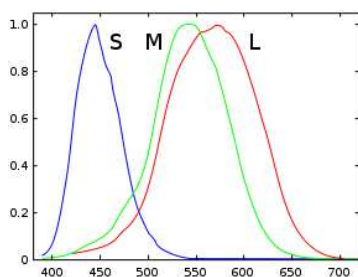
Φωτοϋποδοχείς

Έχουμε ήδη γνωρίσει ότι η επιλογή του κόκκινου, πράσινου και μπλε ως χρωματικές συνιστώσες στο μοντέλο RGB έγινε στηριζόμενη στην λειτουργία του ανθρώπινου ματιού. Εδώ μπορούμε να καταλάβουμε ακριβώς τους λόγους αυτής της επιλογής εξετάζοντας τους φωτοϋποδοχείς (photoreceptors). Οι φωτοϋποδοχείς είναι κύτταρα που δεσμεύουν φως χάρις στις οπτικές χρωστικές ουσίες που περιέχουν όπως ή ροδοψίνη (rhodopsin). Κατόπιν απελευθερώνουν αντίστοιχα ιόντα δημιουργώντας στην ουσία έναν ηλεκτρικό παλμό στις συνάψεις της εξωτερικής δικτυακής στοιβάδας. Έτσι φαίνεται άμεσα μια αντιστοιχία των φωτοϋποδοχέων με τις κυψέλες ενός αισθητήρα CCD/CMOD που γνωρίσαμε, δηλαδή λήψη φωτός και παραγωγή ηλεκτρισμού. Οι φωτοϋποδοχείς είναι δύο ειδών τα κωνία⁶⁹ (cones) ή κωνιοφόρα⁷⁰ κύτταρα και τα ραβδία (rods) ή ραβδιοφόρα κύτταρα ονομασμένα αντίστοιχα με το σχήμα τους.



7-4: Κωνίο και ραβδίο (Madder) [105]

Τα ραβδία είναι υπερευαίσθητα στην παραμικρή ύπαρξη φωτός και έχει αποδειχθεί ότι μπορούν να διεγερθούν από ένα μοναδικό φωτόνιο [52].. Έτσι είναι ευαίσθητα στο αμυδρό φως και στην κίνηση καθώς ανιχνεύουν μικρές διαφορές της φωτεινής έντασης λειτουργώντας όμως μονόχρωμα (άσπρο – μαύρο) και χωρίς να μπορούν να δουν με μεγάλη λεπτομέρεια. Επίσης είναι αυτά που βοηθούν περισσότερο την όραση σε σκοτεινές συνθήκες φωτισμού όπως την νύχτα, δηλαδή την σκοτοπική όραση [104].



7-5: Ευαισθησίες τριών κωνίων S,M,L σε περιοχές του ορατού φάσματος

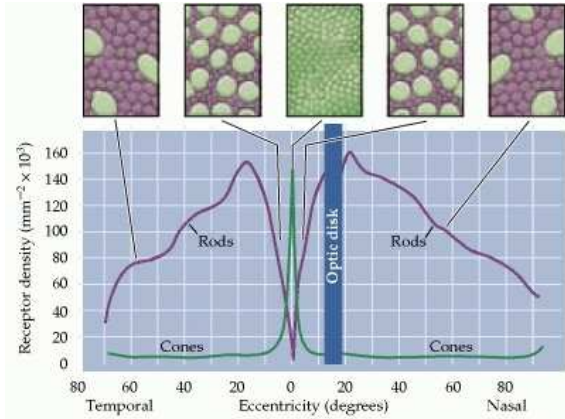
Από την άλλη χάρις στα κωνία είναι εφικτή η έγχρωμη όραση στον άνθρωπο και η αναγνώριση των λεπτομερειών ενός αντικειμένου. Υπάρχουν τριών ειδών κωνία ευαίσθητα σε τρία διαφορετικά μήκη κύματος φωτός. Τα μεγάλου (long) μήκους κύματος L, μέσου M (medium) και μικρού S (short) [109] και αντιστοιχούν στο κόκκινο, πράσινο και μπλε φάσμα του φωτός. Ανάλογα με το χρώμα του αντικειμένου που βλέπουμε ενεργοποιούνται διαφορετικά τα κωνία στην κάθε μια από τις τρεις κλάσεις έχοντας μέγιστες τιμές ενεργοποίησης που εμφανίζονται στην διπλανή εικόνα. Έτσι διαπιστώνουμε ακριβώς τον λόγο που οδήγησε στην δημιουργία του χρωματικό μοντέλου RGB και την χρήση του στους υπολογιστές, κλείνοντας τον

⁶⁹ Πηγή ελληνικού όρου [104]

⁷⁰ Πηγή ελληνικού όρου [52]

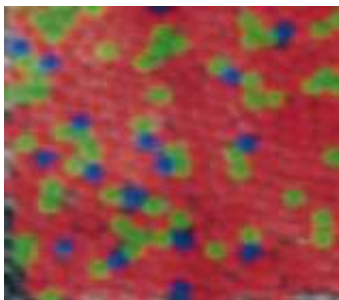
κύκλο που ανοίξαμε. Ο λόγος είναι ότι στον αμφιβληστροειδή η εικόνα αναλύεται σε τρία βασικά χρώματα κόκκινο, πράσινο και μπλε.

Ένα πολύ ενδιαφέρον χαρακτηριστικό που επηρεάζει την όραση είναι κατανομή των κωνίων και ραβδίων πάνω στην σφαιρική επιφάνεια που καλύπτει ο αμφιβληστροειδής. Δηλαδή το πλήθος των κωνίων και ραβδίων που υπάρχουν σε μια περιοχή του αμφιβληστροειδή, εξαρτάται από την εκκεντρότητα (eccentricity) που έχει η περιοχή σε σχέση με το κεντρικό βοθρίο (fovea). Σύμφωνα με την μελέτη του Osterberg [110] υπάρχει η μεγαλύτερη συγκέντρωση κωνίων στο κεντρικό βοθρίο με ταυτόχρονη απουσία ραβδίων ενώ μόλις 6° έκκεντρα σε σχέση με αυτό η οξύτητα της όρασης πέφτει κατά 75% [111]. Δηλαδή στην άκρη του ματιού μας βλέπουμε μονόχρωμα και θολά ενώ στο κέντρο έγχρωμα και με μεγάλη λεπτομέρεια. Τεκμηριώνοντας ακόμα ένα θέμα που προαναφέραμε, οι σακκαδικές κινήσεις των οφθαλμών στοχεύουν, μέσα από διαδοχικές μετακινήσεις του βολβού, να ξεθολώσουν τις λεπτομέρειες του αντικειμένου στοχεύοντας



7-6: Κατανομή κωνίων και ραβδίων στον ανθρώπινο αμφιβληστροειδή κατά Osterberg. Με πράσινο φαίνονται τα κωνία και με ιώδες τα ραβδία (Purves et al) [111]

τις στο κεντρικό βοθρίο [27].



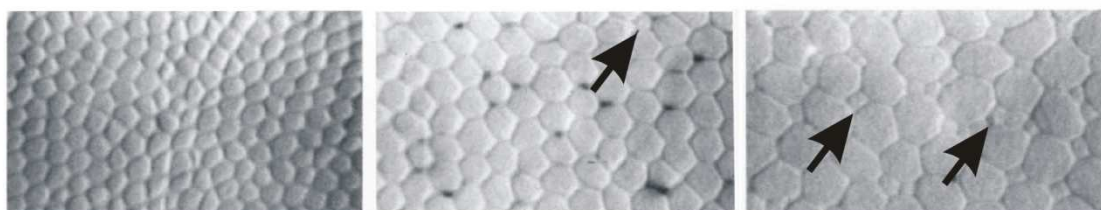
7-9: Χρωματισμένη φωτογραφία αμφιβληστροειδούς ως προς τους τρεις διαφορετικούς τύπους κωνίων

εκκεντρότητα [113] ενώ ακριβώς πάνω στο κεντρικό βοθρίο παίρνει μορφή κυψέλης της μέλισσας (honeycomb) δηλαδή μια καθαρά εξαγωνική ψηφίδωση, που αποτέλεσε πηγή έμπνευσης για την υλοποίηση του Horasis. Τέλος θα αναφέρουμε, ως σύνδεση με την εργασία μας, ότι θα μοντελοποιηθεί μια απλή αντιστοίχιση με τους τέσσερις διαφορετικούς τύπους φωτοϋποδοχέων. Οι τιμές R,G,B θα αντιστοιχούν σε αποκρίσεις κωνίων L,M,S αντίστοιχα και η τιμή Y του φωτός (luma) ως απόκριση ραβδίου.

Table 1 Cone numbers for two subjects with normal colour vision

Subject	Number of cones	L cones (%)	M cones (%)	S cones (%)	Error (%)	L:M ratio
JW	1462	75.8	20.0	4.2	2.1	3.79
AN	522	50.6	44.2	5.2	5.6	1.15

7-7: Κατανομή L,M,S κωνίων σε δύο υγιείς ανθρώπους (Roorda & Williams) [109]



7-8: Ψηφιδώσεις κωνίων στον αμφιβληστροειδή. Με βέλη εμφανίζεται η παρουσία ραβδίων (Curcio et al.) [113]

Οριζόντια κύτταρα

Η πληροφορία ξεκινώντας από τους φωτοϋποδοχείς φτάνει στα γαγγλιακά κύτταρα, ακολουθώντας δύο διαφορετικές οδούς τις πλάγιες και τις κατακόρυφες. Τις πλάγιες (lateral) οδούς τις δημιουργούν τα οριζόντια κύτταρα αμφιβληστροειδούς (horizontal cells), συντετμημένα HC, μεταδίδοντας πληροφορίες από πολλά διαφορετικά κωνία προς τα δίπολα κύτταρα, τα οποία δημιουργούν τις κατακόρυφες οδούς [52]. Οι συνάψεις των οριζοντίων κυττάρων, βρίσκονται στην εξωτερική δικτυακή στοιβάδα OPL και υλοποιούνται από τριών ειδών διαφορετικών ειδών κύτταρα που έχουν ανακαλυφθεί στον ανθρώπινο αμφιβληστροειδή τα HI, HII και σχετικά πρόσφατα τα HIII [114].



7-10: Τα δύο πρώτα είδη των οριζοντίων κυττάρων (Masland) [115]

Τα κύτταρα HC έχουν δενδρίτες που αναπτύσσονται κατά μήκος της OPL, δηλαδή κατά την οριζόντια διάσταση, εξ' ου και η ονομασία τους, δημιουργώντας συνάψεις με από διαφορετικούς φωτοϋποδοχείς του υποδεκτικού τους πεδίου. Το υποδεκτικό πεδίο (receptive field) ενός νευρώνα γενικά, αντιπροσωπεύει την περιοχή του χώρου της εικόνας, που παράγει ερεθίσματα εισόδου από τα οποία εξαρτάται η έξοδος του νευρώνα [116].

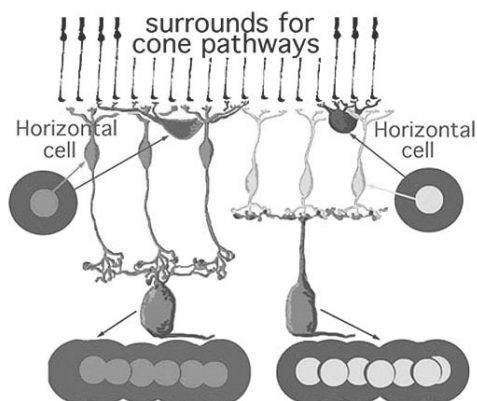


Fig. 12. Diagram of the organization of center-surround responses using horizontal cell circuitry to provide the antagonistic surround.

7-11

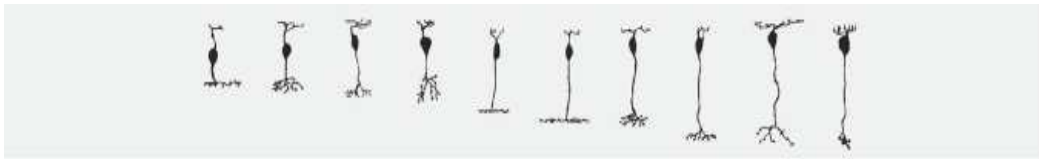
Η πρώτη και βασική λειτουργία των οριζοντίων νευρώνων είναι η γραμμική ολοκλήρωση (linear integration) των σημάτων των φωτοϋποδοχέων, δηλαδή η γραμμική τους άθροιση που έχει ως αποτέλεσμα την παραγωγή στην έξοδο τους μιας γραμμική ενεργοποίησης [108]. Έτσι αυτό το αποτέλεσμα έχει εμπρόσθια τροφοδότηση (feedforward) νέων σημάτων προς τους δίπολους νευρώνες.

Η δεύτερη λειτουργία τους επιγραμματικά είναι η επανατροφοδοσία (feedback) των φωτοϋποδοχέων με τους οποίους συνδέονται ανάλογα με την οπτική πληροφορία στην είσοδο τους [108]. Έτσι τα HC μπορούν να μπλοκάρουν το σήμα κάποιων φωτοϋποδοχέων ενισχύοντας παράλληλα το σήμα άλλων, υλοποιώντας έναν ανταγωνισμό μεταξύ τους που ονομάζεται πλευρική αναστολή (lateral inhibition). Αυτή η οπίσθια τροφοδότηση (feedback) βοηθάει στον τονισμό των ακμών σε μια εικόνα [117] [118].

Μέσα στα πλαίσια της πτυχιακής εργασίας θα προσομοιωθεί η εμπρόσθια τροφοδότηση και η γραμμική άθροιση που επιτελούν τα οριζόντια κύτταρα. Έτσι θα μοντελοποιηθούν ως αντικειμενοστραφείς τεχνητοί νευρώνες McCulloch-Pitts που θα αθροίζουν τιμές φωτοϋποδοχέων. Θα χρησιμοποιηθεί η συνάρτηση ενεργοποίησης την $f(u) = u$, όπου u το συναπτικό άθροισμα, δηλαδή το άθροισμα των γινομένων των τιμών των φωτοϋποδοχέων επί τα συναπτικά βάρη για τους δενδρίτες του τεχνητού οριζοντίου κυττάρου. Ο καθορισμός των τιμών αυτών των συναπτικών βαρών είναι ένα ζήτημα που μπορεί να αποτελέσει μελλοντικά, ένα νέο θέμα έρευνας.

Δίπολα κύτταρα και η πρώτη παρουσία του χρωματικού ανταγωνισμού

Τα δίπολα κύτταρα (bipolar), συντετμημένα BC, δημιουργούν τις κατακόρυφες οδούς ώστε να φτάσουν τα σήματα από τους υποδοχείς προς τα γαγγλιακά κύτταρα, συντετμημένα GC [119]. Είναι κύτταρα με ένα σώμα και δύο νευρικές ίνες εκατέρωθεν του μια για την είσοδο και μια για την έξοδο, έχοντας στην ουσία δύο πόλους. Διαβιβάζουν το σήμα που λαμβάνουν από συνάψεις με κωνίο, με ραβδίο, ή με οριζόντιο κύτταρο δημιουργώντας μέχρι και πολύπλοκους συνδυασμούς που περιέχουν και τους τρεις τύπους συνάψεων. Η συνδεσμολογία τους γίνεται και αυτή στην εξωτερική δικτυακή στοιβάδα OPL και σε μια απλουστευμένη θεώρηση αυτής, έχοντας με χωρική οργάνωση, διαβιβάζουν το σήμα από πολλούς φωτοϋποδοχείς άμεσα ή έμμεσα προς τα γαγγλιακά κύτταρα.

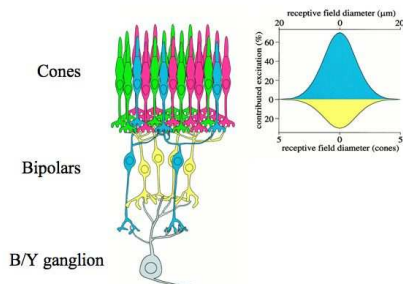


7-12: Διαφορετικοί τύποι δίπολων κυττάρων (Masland) [115]

Η ανθρώπινη όραση λειτουργεί με το προσθετικό μοντέλο των χρωμάτων, καθώς κάθε χρώμα γίνεται αντιληπτό από τις τιμές τριών διαφορετικών κωνίων L,M,S που όπως είδαμε μετρούν το φάσμα του. Το βασικό χαρακτηριστικό της τριχρωματικής όρασης, το οποίο βλέπουμε να εμφανίζεται για πρώτη φορά μέσα οργάνωση των BC νευρώνων, είναι ο χρωματικός ανταγωνισμός. Αναφέρεται στο γεγονός ότι ο άνθρωπος δεν μπορεί να αντιληφθεί κάποια ζευγάρια χρωμάτων σε συνδυασμό, όπως για παράδειγμα δεν μπορεί να θεωρήσει ένα αντικείμενο είναι και μαύρο και άσπρο ταυτόχρονα. Σύμφωνα με την θεωρία της ανταγωνιστικής διεργασίας που διατυπώθηκε από τον Ewald Herding το 1877 υπάρχουν έξι χρώματα, των οποίων η επεξεργασία γίνεται σε τρία ανταγωνιστικά ζεύγη [52]:

- κόκκινο – πράσινο
- μπλε – κίτρινο
- λευκό – μαύρο

Blue/yellow pathway



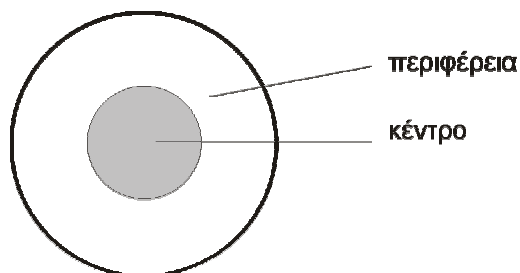
7-13: Σχηματική αναπαράσταση διαύλου ανταγωνισμού μπλε-κίτρινου (<http://www.cns.nyu.edu> [162])

Έτσι δεν μπορούμε να χαρακτηρίσουμε ένα μοναδικό χρώμα ως κοκκινοπράσινο ή μπλεκίτρινο ενώ μπορεί να χαρακτηριστεί μπλεκόκκινο (ματζέντα), κοκκινοκίτρινο (πορτοκαλί), μπλεκόκκινο (κυανό ή πετρόλ) και πρασινοκίτρινο (λαδί). Αυτόν τον ανταγωνισμό εμφανίζουν οι διαυλοι χρωματικού ανταγωνισμού (color opponency pathways) πάνω στους οποίους κινείται η οπτική πληροφορία, η οποία παραγόμενη από τα κωνία και τα ραβδία, περνάει μέσα από τους BC νευρώνες και καταλήγει στους GC νευρώνες. Έτσι βλέπουμε να υπάρχουν συνάψεις BC νευρώνων που ενεργοποιούνται αντίστοιχα για κάθε ένα από τα έξι ανταγωνιστικά χρώματα κόκκινο, πράσινο, μπλε, κίτρινο, λευκό, μαύρο.

Η ακριβείς συνδέσεις τους με τα χρωματικά κωνία, ή τα μονόχρωμα ραβδία δεν μας ενδιαφέρουν στην παρούσα εργασία και στο απλοποιημένο μοντέλο μας θα θεωρήσουμε ως

είσοδος στους τεχνητούς δίπολους νευρώνες, τις τρεις συνιστώσες του χρωματικού μοντέλου RGB. Για τα κύτταρα BC του άσπρου και του μαύρου που λέγονται και δίπολα των ραβδίων (rod bipolar cells), επειδή στην πραγματικότητα συνδέονται κυρίως με ραβδία [120], θα εκμεταλλευθούμε την τιμή του φωτός Y, που θα θεωρείται πλέον η τιμή ενός τεχνητού ραβδίου. Οι υπόλοιποι BC νευρώνες που αντιπροσωπεύουν την έγχρωμη όραση ονομάζονται και δίπολα των κωνίων (cone bipolar cells).

Κυκλικά υποδεκτικά πεδία

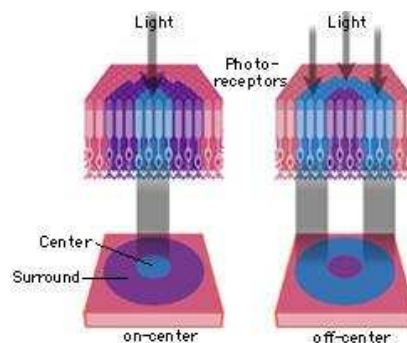


7-14: Οργάνωση κυκλικού υποδεκτικού πεδίου με κέντρο και περιφέρεια.

Ως προς την χωρική οργάνωση τα υποδεκτικά πεδία των BC νευρώνων είναι κυκλικά εμφανίζουν μια οργάνωση κέντρου και περιφέρειας δηλαδή δύο ομόκεντρων κυκλικών δίσκων που ο μικρότερος από τους δύο θεωρείται το κέντρο του υποδεκτικού πεδίου [121]. Η υπόλοιπη επιφάνεια που περισσεύει από τον μεγάλο δίσκο είναι η περιφέρεια του υποδεκτικού πεδίου. Αυτή η οργάνωση εμφανίζεται παρακάτω και σε άλλους

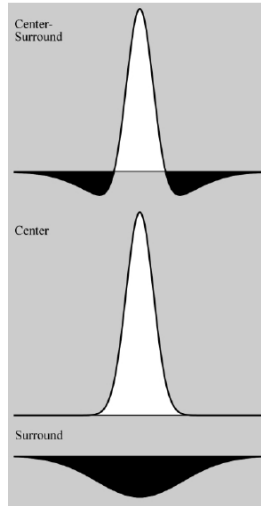
νευρώνες, γι αυτό αξίζει να αναφερθούμε σε αυτήν ως ξεχωριστή παράγραφο.

Βασιζόμενη στον τρόπο που αποκρίνονται οι νευρώνες στις τιμές που έρχονται ως είσοδος στα κυκλικά υποδεκτικά τους πεδία, μπορούμε να τους διαχωρίσουμε σε αναμμένου-κέντρου (on center) και σβησμένου-κέντρου (off center) [52]. Ως αναμμένο εννοούμε ένα κέντρο στο οποίο πέφτει πάνω το φως και ως σβησμένο ένα κέντρο στο οποίο υπάρχει απόλυτο σκοτάδι. Τα αναμμένου-κέντρου υποδεκτικά πεδία λέγονται έτσι γιατί οι νευρώνες διεγείρονται περισσότερο όταν το φως χτυπάει στο κέντρο τους και αναστέλλονται όταν το φως χτυπάει στην περιφέρεια τους. Έτσι έχουν το «μεγαλύτερο» σήμα στην έξοδο τους όταν το υπάρχει ένας κυκλικός δίσκος φωτός ακριβώς στο κέντρο τους χωρίς να πέφτει καθόλου φως γύρω από αυτόν. Αντίστροφα στους νευρώνες με σβησμένου κέντρου υποδεκτικά πεδία, παράγεται το «μεγαλύτερο» σήμα στην έξοδο τους παράγεται όταν το φως διαχέεται μόνο πάνω στην περιφέρεια, αφήνοντας μια σκοτεινή κηλίδα ακριβώς στο κέντρο της. Κάθε μετακίνηση της κηλίδας μακριά από το κέντρο μειώνει την απόκριση του νευρώνα.



7-15: Αναμμένα και σβησμένα κέντρα σε κυκλικά υποδεκτικά πεδία (Brainconnection.com) [161]

Στο σύστημα Horasis μοντελοποιείται η οργάνωση κέντρου και περιφέρειας στις εισόδους των τεχνητών γαγγλιακών νευρώνων, αλλά δεν χρησιμοποιείται στους τεχνητούς δίπολους νευρώνες για απλοποίηση. Κυκλικά υποδεκτικά πεδία προσομοιώνονται επίσης στους τεχνητούς νευρώνες που αντιστοιχούν στον οπτικό φλοιό του εγκεφάλου, που θα γνωρίσουμε λίγο παρακάτω. Οι δύο διεγέρσεις ενεργοποίηση και αναστολή μπορούν να θεωρηθούν ότι ακολουθούν



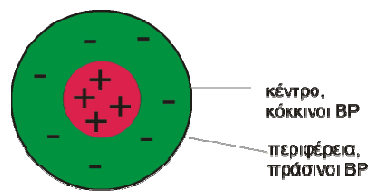
7-16: Γκαουσιανές Αποκρίσεις κέντρου –περιφέρειας (Rowe) [125]

χαρακτηριστικά κανονικής (gauss) κατανομής [122] εξαρτώμενες από την οργάνωση κέντρου – περιφέρειας των υποδεκτικών πεδίων[123] [124]. Ενοποιούμενες δημιουργούν μια κοινή απόκριση κωδικοποιώντας την κυματομορφή που παράγεται στην έξοδο του βιολογικού νευρώνα [125]. Αυτήν την σημαντική θεώρηση θα την χρησιμοποιήσουμε για την παραγωγή στόχων εκπαίδευσης που θα προσομοιώνουν την λειτουργία κέντρου – περιφέρειας.

Γαγγλιακά κύτταρα

Ήδη έχουμε δει ότι ο τελικός στόχος της επεξεργασίας από τους νευρώνες της εσωτερικής κοκκώδους στοιβάδας είναι να διασπαστεί η οπτική πληροφορία σε γραμμές χρωμάτων που καταλήγουν στα γαγγλιακά κύτταρα (ganglion cells) ή γάγγλια, συντεταγμένα GC. Λαμβάνουν είσοδο μόνο από τα δίπολα και βραχύινα κύτταρα με συνάψεις που γίνονται στην εσωτερική δικτυακή στοιβάδα IPL (Inner Plexiform Layer) [52] [126]. Οι νευρώνες GC παράγουν την έξοδο του αμφιβληστροειδούς, δηλαδή μια νέα κωδικοποίηση της πληροφορίας η οποία συναρτάται από την αρχική τριχρωματική πληροφορία που παράγουν οι φωτοϋποδοχείς. Η στοιβάδα γαγγλιακών κυττάρων που περιλαμβάνει μόνο GC νευρώνες, είναι το στρώμα εξόδου ολόκληρου του νευρωνικού δικτύου του αμφιβληστροειδούς. Μπορούμε να θεωρήσουμε δηλαδή τον αμφιβληστροειδή χιτώνα του ματιού ως ένα ενιαίο νευρωνικό δίκτυο που δέχεται σαν είσοδο το φως μιας εικόνας, το επεξεργάζεται και παράγει στην έξοδο του μια συστοιχία ηλεκτρικών σημάτων. Το κάθε ένα σήμα που είναι και φορέας της οπτικής πληροφορίας μπορεί να εκφραστεί ως μια **Κωδικοποίηση Γαγγλιακού Κυττάρου (Ganglion Cell Encoding - GCE)** δηλαδή μια μαθηματική αναπαράσταση της μορφής του ηλεκτρικού παλμού που παράγει στην έξοδο του ένα βιολογικό γαγγλιακό κύτταρο. Το σύνολο αυτών των σημάτων θα αποτελεί την **Νευρωνική Διαμόρφωση Αμφιβληστροειδούς (Retina Neural Format – RNF)** δηλαδή την αναπαράσταση μιας εικόνας αφού έχει επεξεργαστεί από τον αμφιβληστροειδή, η οποία συνιστά την πληροφορία που μεταφέρεται μέσω του οπτικού νεύρου.

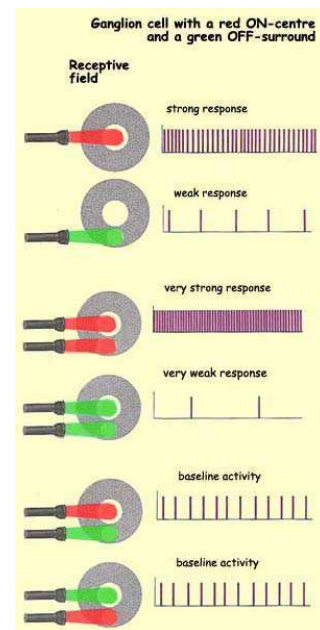
Η έξοδος αυτή θα εξαρτάται από την είσοδο του GC νευρώνα και το υποδεκτικό πεδίο του, το οποίο είναι κυκλικό και οργανώνεται με κέντρο και περιφέρεια [125]. Η οργάνωση αυτή εμπεριέχει και χρωματικό ανταγωνισμό καθώς ο κάθε νευρώνας αποκρίνεται καλύτερα όταν ένα χρώμα είναι αναμμένο (basic on) στο κέντρο του και το ανταγωνιστικό του φωτίζει (basic off) παράλληλα την περιφέρεια του. Το πρώτο χρώμα για ευκολία θα το ονομάσουμε βασικό χρώμα και το δεύτερο το ανταγωνιστικό χρώμα και μπορούμε να ονομάσουμε ένα γαγγλιακό κύτταρο χρωματικού ανταγωνισμού κόκκινου και πράσινου όπως το παρακάτω, απλά ως γάγγλιο αναμμένου-κόκκινου (red-on ganglion), ή συντετμημένα R-GC. Αντίστοιχα θα υπάρχουν για τους άλλους δύο ανταγωνισμούς για το μεν μπλε/κίτρινο η σύντηξη B-GC και τον άσπρο/μαύρο ή φως/σκότος η L- GC. Τα αντίστροφα όπως πράσινου κέντρου και κόκκινης περιφέρειας θα αναφέρονται πάλι με τα αντίστοιχα αρχικά του βασικού χρώματος δηλαδή G-GC, Y-GC και D-GC.



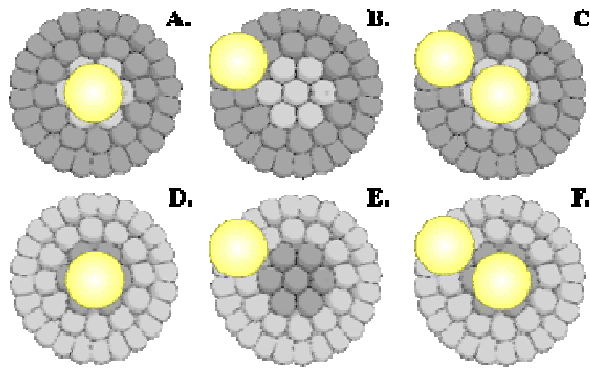
7-17: Οργάνωση κέντρου περιφέρειας γαγγλίου με χρωματικό ανταγωνισμό κόκκινου με πράσινο, R-GC

Ένα R/G γάγγλιο συνδέεται με κόκκινους και πράσινους δίπολους νευρώνες, αλλά γενικά ένας GC νευρώνας μπορεί να συνδέεται με ένα πλήθος διαφορετικών BC νευρώνων. Όταν στο υποδεκτικό του πεδίο, οι κόκκινοι BC νευρώνες του κέντρου είναι διεγερμένοι και οι αντίστοιχοι κόκκινοι της περιφέρειας αναστέλλονται, τότε η έξοδος του GC είναι και αυτή διεγερτική και η ισχυρότερη που μπορεί να υπάρξει [52]. Αυτό μπορεί να εκφραστεί ισοδύναμα, αν πούμε ότι έχουμε αναμμένους τους κόκκινους BC νευρώνες του κέντρου και αναμμένους τους πράσινους νευρώνες BC της περιφέρειας. Η πραγματική συνδεσμολογία τους και εδώ δεν θα μας απασχολήσει στα πλαίσια της εργασίας, ώστε να υλοποιηθούν απλές συνδεσμολογίες όπως η παραπάνω.

Ως προς την λειτουργία τους μπορεί επίσης να ειπωθεί ότι είναι πολύ ευαίσθητα στην μετακίνηση του φωτισμού του βασικού χρώματος μακριά από το κέντρο. Έτσι αν το βασικό χρώμα χτυπήσει έκκεντρα τους φωτοϋποδοχείς σε μια περιοχή του αμφιβληστροειδούς η οποία αντιστοιχίζεται σε έναν GC νευρώνα, θα διεγείρει το κέντρο και θα αναστείλει την περιφέρεια του νευρώνα παράγοντας μια σύνθετη απόκριση. Ας δούμε τις περιπτώσεις θεωρώντας τα υποδεκτικά πεδία δύο γαγγλιακών νευρώνων που συμμετέχουν στον δίαυλο χρωματικού ανταγωνισμού άσπρου-μαύρου. Το πρώτο που είναι αναμμένου-άσπρου L/D GC εμφανίζεται με ανοικτό γκρι στο κέντρο του στα σχήματα A,B και C της εικόνας 7-19 και το δεύτερο αναμμένου-μαύρου D/L GC και εμφανίζεται με σκούρο γκρι κέντρο στα σχήματα D,E,F. Τρεις διαφορετικές δέσμες λευκού φωτός, που εμφανίζονται ως κίτρινοι δίσκοι, χτυπάνε τους φωτοϋποδοχείς που σχετίζονται με αυτά.



7-18: Συμπεριφορά R-GC



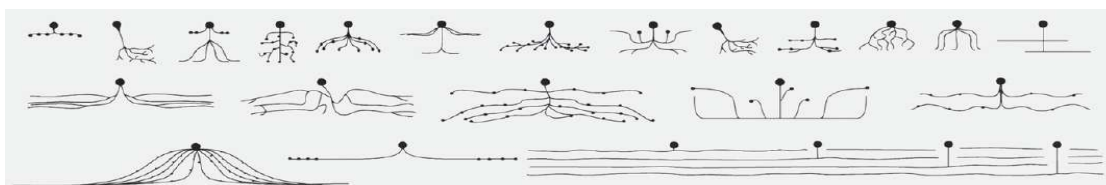
7-19: Αναπαράσταση μετατόπισης δέσμης φωτός μέσα σε υποδοκτικά πεδία γαγγλίων
(Northern Kentucky University)[127]

- Πρώτη δέσμη με μια κηλίδα που πέφτει στο κέντρο
 - A. Το λευκό φως πέφτει ακριβώς στο άσπρο κέντρο διεγείροντας το L-GC και η μαύρη περιφέρεια προσθέτει στην διέγερση καθώς δεν φωτίζεται καθόλου. Αυτή είναι και η καλύτερη απόκριση του L-GC
 - D. Το λευκό φως πέφτει ακριβώς στο μαύρο κέντρο αναστέλλοντας το D- GC με την άσπρη περιφέρεια να προσθέτει στην αναστολή καθώς δεν φωτίζεται καθόλου. Το D-GC αποκρίνεται εντελώς αντίστροφα από το L-GC με την χειρότερη απόκριση του.
- Δεύτερη δέσμη με μια κηλίδα που πέφτει περιφερειακά
 - B. Το λευκό φως πέφτει πάνω στην μαύρη περιφέρεια του L-GC αναστέλλοντας το, με το άσπρο κέντρο που δεν φωτίζεται να ενισχύει την αναστολή.
 - E. Το λευκό φως πέφτει πάνω στην λευκή περιφέρεια του D-GC διεγείροντας το εν μέρει, και το κέντρο επειδή δεν φωτίζεται ενισχύει την διέγερση.
- Τρίτη δέσμη που φωτίζει και κέντρο και περιφέρεια
 - C. Το λευκό φως πέφτει και στο άσπρο κέντρο διεγείροντας και στην μαύρη περιφέρεια αναστέλλοντας το L/D GC παράγοντας μια σύνθετη απόκριση
 - F. Το λευκό φως πέφτει και στο μαύρο κέντρο αναστέλλοντας και στην άσπρη περιφέρεια διεγείροντας το D/L GC παράγοντας μια σύνθετη απόκριση.

Όπως θα δούμε στην περιγραφή της μοντελοποίησης του Horasis, γίνεται η προσπάθεια να καλυφθούν όλες αυτές οι περιπτώσεις, με την εκπαίδευση των τεχνητών γαγγλιακών νευρώνων χρησιμοποιώντας στόχους που αναπαριστούν αυτές τις διεγέρσεις και αναστολές. Αυτοί οι δύο αντικρουόμενοι μεταξύ τους GC νευρώνες, αντιδρούν παράλληλα στο ίδιο ερέθισμα επειδή είναι όλοι προσανατολισμένοι ώστε να ελέγχουν το ίδιο χωρικό σημείο του αμφιβληστροειδούς. Έτσι σε ένα συγκεκριμένο σημείο του αμφιβληστροειδούς θεωρούμε για τις ανάγκες της εργασίας ότι υπάρχουν 6 GC νευρώνες, που ο κάθε ένας έχει ως βασικό χρώμα για το κέντρο του υποδοκτικού του πεδίου ένα από τα 6 ανταγωνιστικά χρώματα. Όπως στην πραγματικότητα ο αμφιβληστροειδής χωρίζεται κάτω από τους φωτοϋποδοχείς σε περιοχές δράσεις των GC νευρώνων, έτσι και εδώ θα χωρίσουμε τον αντικειμενοστραφή τεχνητό αμφιβληστροειδή σε πολλά αντικείμενα, ένα για κάθε εξάδα GC νευρώνων.

Βραχύινα κύτταρα

Η τέταρτη κλάση νευρώνων που υπάρχουν στον αμφιβληστροειδή είναι τα βραχύινα ή αμακρίνα (amacrine) κύτταρα, συντετμημένα AC, που όπως είπαμε ονομάζονται έτσι λόγω τους σχήματος του και όχι λόγω της κατεύθυνσης με την οποία δημιουργούν συνάψεις. Συνδέονται και με δίπολα κύτταρα και με γαγγλιακά κύτταρα με συνάψεις που ανήκουν στην εσωτερική δικτυακή στοιβάδα IPL. Οι συνάψεις ενός AC είναι αρκετά σύνθετες και μπορεί να συνδυάζουν εισόδους από πολλά BC και GC αλλά και ανατροφοδότηση (feedback) προς αυτές [128]. Υπάρχουν 29 διαφορετικοί τύπου AC νευρώνων [115] που εξυπηρετούν διάφορες λειτουργίες, με μερικές από αυτές να μην είναι πλήρως γνωστές ως σήμερα [129]. Είναι ένα από τα πεδία έρευνας της νευροεπιστήμης που αναμένεται να παράγει πολλά χρήσιμα δεδομένα που μπορούν στο μέλλον να εκμεταλλευθούν από εργασίες σαν το Horasis.



7-20: Διαφορετικοί τύποι βραχύνων κυττάρων (Masland) [115]

Έτσι ένα παράδειγμα από λειτουργίες που ανακαλύφθηκαν είναι η διέγερση των ακτινωτών (starburst) AC νευρώνων σε κινούμενα αντικείμενα. Ένα άλλο είναι η δημιουργία μιας διασταυρούμενης συσχέτισης μεταξύ των GC νευρώνων ώστε δύο GC νευρώνες που μοιράζονται είσοδο από το ίδιο AC νευρώνα να παράγουν έξοδο «πυροβολώντας» ταυτόχρονα [115]. Μια άλλη λειτουργία των AC που είναι η ρύθμιση της απόκρισης του αμφιβληστροειδούς κάτω από ισχυρό ή ασθενή φωτισμό [115], έχει προσομοιωθεί με έναν στοιχειώδη τρόπο από την εργασία. Περιγράφοντας την μέσα από το πρίσμα ενός μη ειδικού οι AC νευρώνες δημιουργούν ανασταλτικές συνάψεις πάνω στον άξονα των BC νευρώνων, καταπιέζοντας την έξοδο των τελευταίων πριν αυτή τροφοδοτηθεί προς τα εμπρός, στους GC νευρώνες [130]. Καθώς λαμβάνουν είσοδο και από ραβδία τα οποία μετρούν την συνολική ένταση του φωτός, έγινε η θεώρηση ότι οι AC νευρώνες ρυθμίζουν την έξοδο των BC που σχετίζονται με το χρώμα. Έτσι για κάθε ένα από τα κόκκινο R, πράσινο G, μπλε B και κίτρινο Y χρώματα μπορούμε να θεωρήσουμε τις ρυθμισμένες εξόδους για τα αντίστοιχα BC των κωνίων ως συναρτήσεις

$$y_R = f_R(R, L), y_G = f_G(G, L), y_B = f_B(B, L), y_Y = f_Y(Y, L)$$

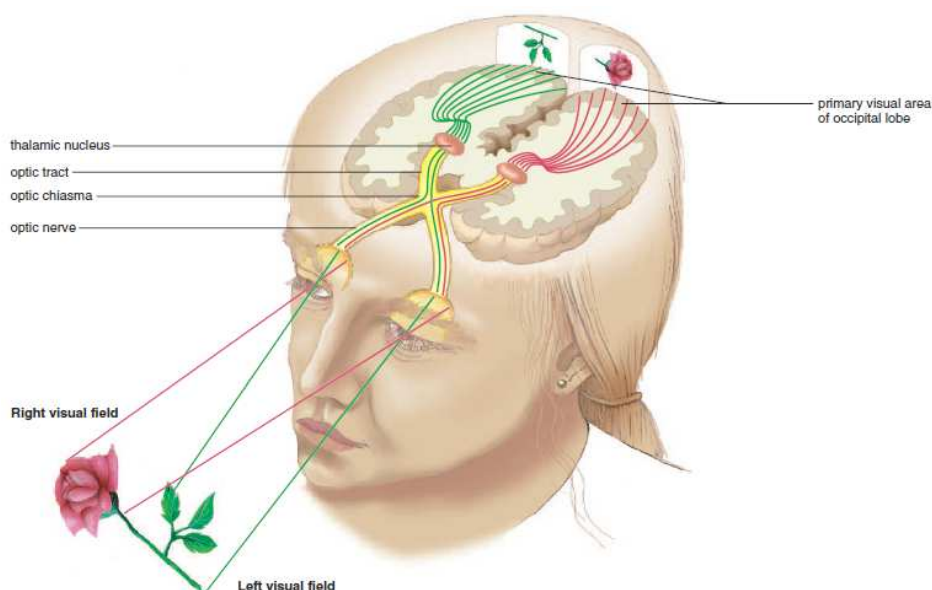
όπου R, G, B, L οι τιμές εξόδου των νευρώνων BP πριν την ρύθμιση και L η συνολική ποσότητα του φωτός της περιοχής που ανήκουν τα κωνία από τα οποία τροφοδοτούνται με πληροφορίες οι νευρώνες. Το μοντέλο Horasis έχει μια απλοϊκή αντιμετώπιση για μια αρκετά σύνθετη, δυσνόητη και σχετικά άγνωστη λειτουργία, καθώς θεωρεί ότι οι συνθήκες φωτισμού είναι σταθερές, έχοντας πάντα επαρκή φωτισμό στα οπτικά πρότυπα που θα δημιουργούνται από την δειγματοληψία. Έτσι δεν χρειάζεται να ασχοληθούμε με κάποιον τρόπο δυναμικής προσαρμογής της εξόδου των συναρτήσεων f_R, f_G, f_B, f_Y που θα προσομοίωνε καλύτερα την λειτουργία προσαρμογής στον φωτισμό που επιτελούν τα βραχύινα κύτταρα.

Ο τελικός στόχος των τεχνητών βραχύνων κυττάρων είναι να αγνοούνται οι χρωματικές τιμές, σε μια πολύ σκοτεινή περιοχή μιας εικόνας, η οποία όμως λαμβάνεται από το Horasis σε συνθήκες ισχυρού φωτισμού. Μια τέτοια περιοχή είναι το μαύρο φόντο της εικόνας, που δημιουργείται λόγω του μαύρου παραβάν που υπάρχει στην σκηνή φωτογράφησης.

Ο Οπτικός Φλοιός V1

Η έξοδος του αμφιβληστροειδούς ως είσοδος στον εγκέφαλο

Θα αφήσουμε πλέον τον μικρόκοσμο του αμφιβληστροειδούς και του ματιού και ακολουθώντας το οπτικό νεύρο θα φτάσουμε στο τμήμα του εγκεφάλου που είναι υπεύθυνο για την πρώτη επεξεργασία της οπτικής πληροφορίας, πολύ πριν λειτουργήσουν οι ανώτερες λειτουργίες του που συνδέονται με αυτήν. Η αρχική εικόνα, μετά την παραμόρφωση της με σμίκρυνση και αναστροφή ως είδωλο στον αμφιβληστροειδή, και την επεξεργασία από αυτόν έχει εκφραστεί ως μια συστοιχία κωδικοποιημένων σημάτων, με κωδικοποίηση εξόδου γαγγλίου. Αυτή αποτελεί την νευρωνική διαμόρφωση αμφιβληστροειδούς RNF, δηλαδή το πρότυπο του αντικειμένου για τον εγκέφαλο.



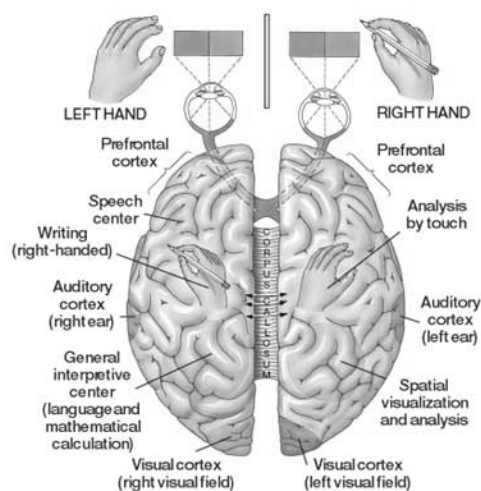
7-21: Διαδρομή του οπτικού νεύρου (Madder) [105]

Έτσι μπορούμε να πούμε με απλά λόγια ότι το οπτικό νεύρο είναι απλά το «καλώδιο» που κουβαλάει ένα πακέτο πληροφοριών από το μάτι προς τον εγκέφαλο. Οι άξονες των γαγγλίων του αριστερού και δεξιού ματιού διασταυρώνονται χιαστί στο λεγόμενο οπτικό χίασμα (optic chiasma) που είναι και ο πρώτος σταθμός της οπτικής πληροφορίας. Κατόπιν οι ίνες του αριστερού οφθαλμού συγκλίνουν και η συνεχίζουν μαζί στην αριστερή οπτική οδό (optic tract) και αντίστοιχα του δεξιού οφθαλμού στην δεξιά οπτική οδό [105].

Η πληροφορία κινούμενη από αυτές τις οδούς περνάει τον υποθάλαμο του εγκεφάλου και καταλήγει σε μια περιοχή στο πίσω τμήμα του εγκεφάλου που λέγεται οπτικός φλοιός. Ο οπτικός φλοιός εκτείνεται και στα δύο ημισφαίρια του εγκεφάλου με το κάθε μάτι να απεικονίζεται στο ακριβώς αντίθετο ημισφαίριο και χωρίζεται σε πέντε στρώματα τα V1, V2, V3, V4 και V5 [131].

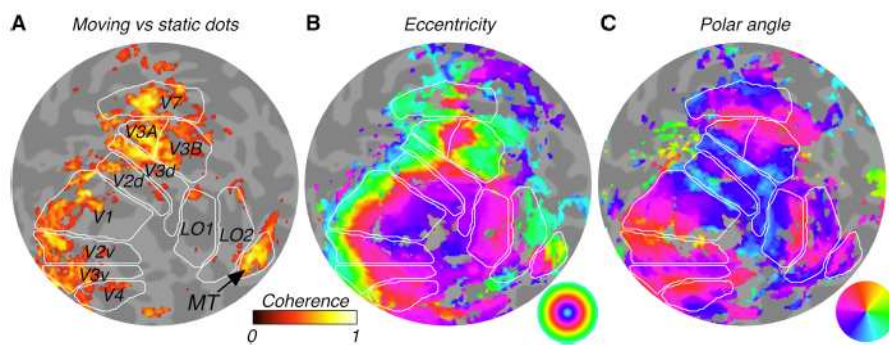


7-22: Θέση του οπτικού φλοιού στο εγκέφαλο, με πράσινο χρώμα



7-23: Μέσω του χιάσματος, το δεξί μάτι πηγαίνει στο αριστερό ημισφαίριο και το αριστερό μάτι στο δεξί ημισφαίριο

Το τμήμα V1 του οπτικού φλοιού λέγεται και πρωτοταγές οπτικός φλοιός (primary visual cortex) και είναι μια περιοχή για την οποία υπάρχουν σαφώς ορισμένα ανατομικά όρια από παλαιότερες και νεότερες μελέτες. Ο ρόλος του στην επεξεργασία της οπτικής πληροφορίας που έρχεται από τον αμφιβληστροειδή είναι σημαντικός καθώς η πληροφορία οδηγείται μέσω αυτού για να φτάσει στα υπόλοιπα τμήματα του οπτικού φλοιού [132]. Σε αυτόν δημιουργείται ένας χάρτης του αμφιβληστροειδούς ή ρετινοτοπικός χάρτης (retinotopic maps) δηλαδή μια χωρική οργάνωση ανάλογη με τα διάφορα χωρία του αμφιβληστροειδούς που καλύπτουν τα γάγγλια. Αυτό σημαίνει ότι κοντινά σημεία στον V1 φλοιό του εγκεφάλου αναπαριστούν και αντιστοιχούν σε κοντινά σημεία της οπτικής σκηνής, όπως αυτή αποτυπώνεται στον αμφιβληστροειδή [133]. Με την χρήση της τεχνικής fMRI ή μέσω των αγγείων του αμφιβληστροειδούς [134] έχουν ήδη αποτυπωθεί οι πρώτοι χάρτες στους οποίους ενδιαφέρουν δύο χωρικές ιδιότητες της σφαίρας του, η εκκεντρότητα δηλαδή η απόσταση από το κεντρικό βοθρίο και το αζιμούθιο (polar angle) [135].



7-24: Ρετινοτοπικός Χάρτης με Α. Ενεργοποιήσεις τμημάτων του οπτικού φλοιού σε σχέση με κινούμενα ή σταθερά σημεία. Β. Εκκεντρότητα Γ. Αζιμούθιο (Gardner et al.) [136]

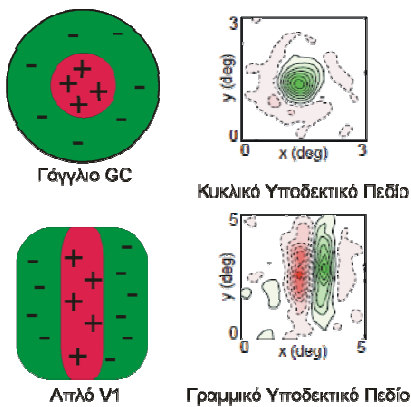
Έτσι η έξοδος των γαγγλίων γίνεται είσοδος σε κύτταρα του V1 αντίστοιχα με την θέση της εικόνας που καλύπτουν τα πρώτα. Δύο διαφορετικοί τύποι κυττάρων του V1 ανταποκρίνονται σε οπτικά ερεθίσματα που έχουν μορφή γραμμής ή ράβδου και κατατάχθηκαν από τους Hubel και Wiesel σε δύο ομάδες, τα απλά και τα σύνθετα [137]. Επειδή τα κύτταρα έχουν όπως θα δούμε υποδεκτικά πεδία ευαίσθητα σε σχήμα και κατεύθυνση, μετά την επεξεργασία από τον V1 θα εξαχθούν μορφολογικά χαρακτηριστικά από την αρχική εικόνα τα οποία θα βρίσκονται σε σαφώς

ορισμένες θέσεις μέσα σε αυτήν. Βασιζόμενη σε αυτήν την πραγματικότητα του εγκεφάλου θα υλοποιηθούν στην εργασία οι τεχνητοί V1 νευρώνες οι οποίοι θα μετασχηματίζουν τις εισόδους που λαμβάνουν υπό μορφή κωδικοποίησης εξόδου γαγγλίου GCE, σε μια νέα πληροφορία που περιέχει σχήμα. Η νέα αυτή διαμόρφωση είναι η εσωτερική αναπαράσταση του πραγματικού αντικειμένου για το σύστημα μας, που αποτελεί το πρότυπο εκπαίδευσης για ένα τελικό ΤΝΔ ταξινόμησης.

Ακόμη ένα χαρακτηριστικό της πραγματικής όρασης που και αυτό θα υλοποιηθεί στο επίπεδο της προσομοίωσης των V1 κυττάρων, είναι η ύπαρξη δύο διαφορετικών κυκλωμάτων του μικρού P (parvi) και του μεγάλου (magna) που χαρακτηρίζουν την διαφορετική λειτουργία τους [52]. Το πρώτο κύκλωμα P με απλά λόγια επεξεργάζεται μικρότερη περιοχή της εικόνας ενώ το δεύτερο M μια αρκετά μεγαλύτερη. Τα δύο μεγέθη ξεκινάν να εμφανίζονται στον αμφιβληστροειδή και συνεχίζουν και στον V1. Στην εργασία μας θα απλοποιήσουμε το μοντέλο υλοποιώντας διαφορετικά μεγέθη υποδεκτικών πεδίων μόνο για τους τεχνητούς V1 νευρώνες.

Απλοί V1 νευρώνες και υπερστήλες νευρώνων.

Η πρώτη ομάδα νευρώνων, του οπτικού φλοιού λέγονται απλά κύτταρα V1 (V1 simple cells), συντετημημένα V1S και έχουν υποδεκτικά πεδία ειδικά διαμορφωμένα με μια τουλάχιστον επιμήκη και ευθύγραμμη ζώνη διέγερσης και μια ζώνη αναστολής που την περιβάλλει [138]. Το

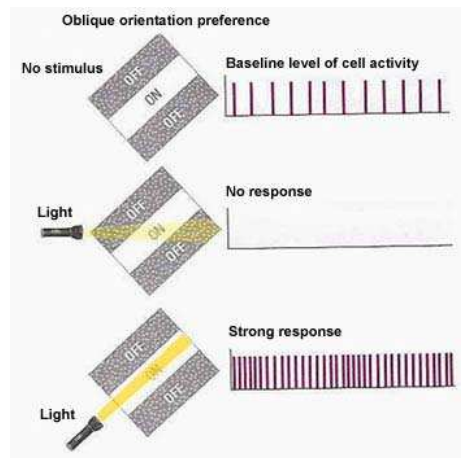


ερέθισμα για να έχει την άριστη αποτελεσματικότητα ως προς την έξοδο του V1S, πρέπει να έχει ακριβώς τον ίδιο προσανατολισμό με το υποδεκτικό πεδίο του. Δηλαδή το ερέθισμα πρέπει να βρίσκεται ακριβώς μέσα στην ζώνη διέγερσης χωρίς να εκτείνεται στην ανασταλτική περιοχή. Λόγω αυτών των χαρακτηριστικών κάθε απλό κύτταρο είναι προσανατολισμένο στην ανίχνευση μιας μπάρας φωτός σε μια περιοχή των 10° [52]. Καθώς ένα ερέθισμα περιστρέφεται κατά 180° προσανατολιζόμενο σε όλες τις πιθανές διευθύνσεις του επιπέδου, θα διεγείρει περίπου 20

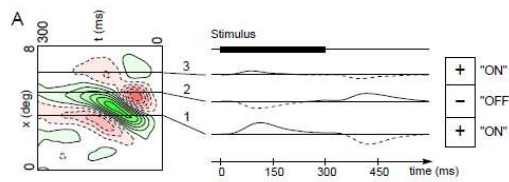
7-26: Δομή υποδεκτικών πεδίων GC και V1S. Διέγερση με + και αναστολή με - (βασισμένο σε DeAngelis et al.) [138]

και σύνθετων κυττάρων. Δηλαδή υπάρχει ένα κύτταρο V1S για κάθε δυνατή περιστροφή μιας γραμμής στο επίπεδο για έναν συγκεκριμένο χρώμα του ανταγωνισμού. Έτσι θα θεωρήσουμε έξι διαφορετικές ομάδες V1S για τα έξι ανταγωνιστικά χρώματα που θα χρησιμοποιήσουμε συντημήσεις ανάλογες με αυτές των GC κυττάρων R-V1S για κόκκινη διέγερση και πράσινη αναστολή, G-V1S για το αντίστροφο, B-V1S για μπλε/κίτρινο, Y-V1S για κίτρινο/μπλε L-V1S για φως/σκοτάδι και D-V1S για διέγερση σκοτεινής γραμμής πάνω σε φωτεινό φόντο.

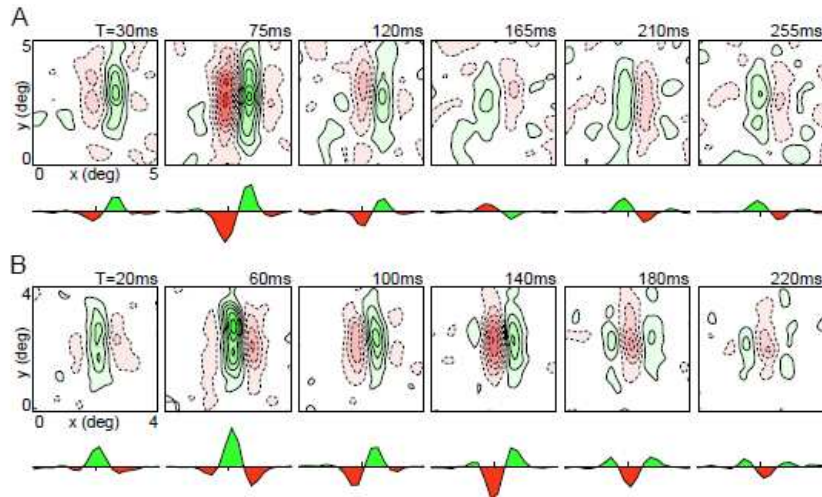
διαφορετικούς πληθυσμούς απλών, αλλά όπως θα δούμε



7-25: Διέγερση V1S με φωτεινή ράβδο (The Brain From Top to Bottom) [164]

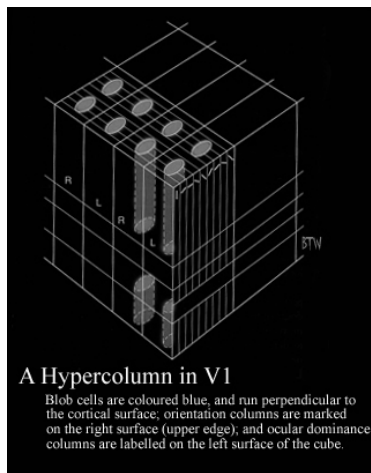


7-27: Σταθερό ανασταλτικό ερέθισμα σε ένα R-V1 (DeAngelis et al.) [138]



7-28: Συμπεριφορά του υποδεκτικού πεδίου των απλών V1 κυττάρων σε διαφορετικές χρονικές στιγμές από την έναρξη του ερεθίσματος (DeAngelis et al.) [138]

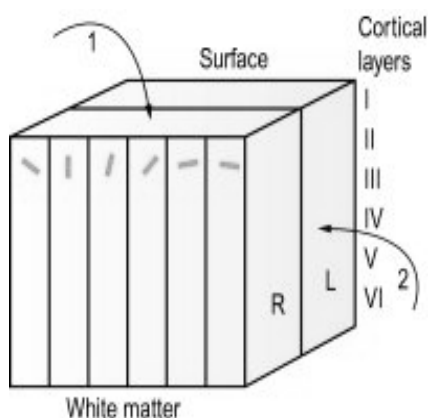
Και στην πραγματικότητα υπάρχει ένα το πλήθος κυττάρων του V1 φλοιού, που υλοποιούν χρωματικό ανταγωνισμό [139], ώντας επικεντρωμένα σχεδόν στις ίδιες χωρικές θέσεις του αμφιβληστροειδούς με ορισμένα να έχουν μεγαλύτερο υποδεκτικό πεδίο από τα άλλα, δημιουργώντας υπερστήλες (hypercolumns) [140]. Τα μεγαλύτερα ανήκουν στο μεγάλο (magna) κύκλωμα της όρασης ενώ τα μικρότερα στο μικρό (parvi) συνεχίζοντας το από τον αμφιβληστροειδή.



A Hypercolumn in V1

Blob cells are coloured blue, and run perpendicular to the cortical surface; orientation columns are marked on the right surface (upper edge), and ocular dominance columns are labelled on the left surface of the cube.

7-29: 7-30: Υπερστήλες (VAL - University Of Calgary) [139]



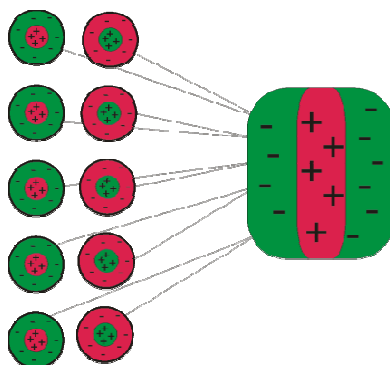
7-31: Υπερστήλες ως προς τον προσανατολισμό του βασικού χρώματος στο υποδεκτικό πεδίο των απλών νευρώνων (Millodot) ([140])

Οι νευρώνες σε κάθε υπερστήλη αποκρίνονται ταυτόχρονα σε έναν συγκεκριμένο άξονα προσανατολισμού [52] και υπάρχουν γειτονικές στήλες για όλες τις κατευθύνσεις. Άρα μπορούμε να πούμε ότι από μια περιοχή της εικόνας, που εισάγεται στον V1 ως κωδικοποιήσεις GCE, εξάγονται παράλληλα πολλές αποκρίσεις για προσανατολισμένες μπάρες φωτός που περιέχονται σε αυτήν. Θα δημιουργούνται παράλληλες αποκρίσεις για κάθε ένα από τα ανταγωνιστικά χρώματα σε κάθε προσανατολισμό, παράγοντας κατά αυτόν τον τρόπο μια ν-διάστατη πληροφορία για κάθε τμήμα της εικόνας που περιλαμβάνει χρώμα και σχήμα. Σκοπός της επεξεργασίας σε αυτό το επίπεδο είναι να μετασχηματίσει τις πληροφορίες που έρχονται από τον αμφιβληστροειδή με χρώμα, προσθέτοντας τις επιπλέον διαστάσεις, του σχήματος.

Αυτή η θεώρηση που περιγράψαμε για την μοντελοποίηση και λειτουργία των κυττάρων του V1 θα αποτελέσει βάση για την δημιουργία των τελικών προτύπων αναγνώρισης εικόνων, μέσα στο σύστημα Horasis. Αν εκφράσουμε το πλήθος των απλών κυττάρων V1S του μοντέλου με έναν απλό μαθηματικό τύπο:

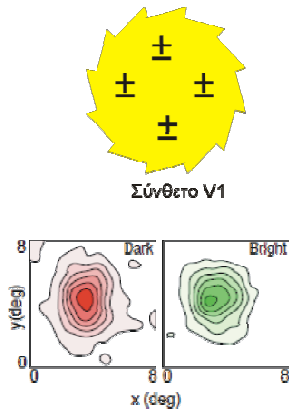
$$v = c \left(\frac{180}{s} \right) \quad (7.1)$$

όπου s είναι η περιοχή ευθύνης σε μοίρες του προσανατολισμού μιας υπερστήλης, 180 οι μοίρες για όλες τις διαφορετικές κατευθύνσεις του επιπέδου της εικόνας, πάνω στο οποίο υπάρχουν οι γραμμές και c οι χρωματικές διαστάσεις που επεξεργάζεται η κάθε υπερστήλη. Όλο αυτοί οι νευρώνες έχουν στα υποδεκτικά πεδία συνάψεις με γαγγλιακά κύτταρα μιας συγκεκριμένης περιοχής του αμφιβληστροειδούς οπότε αποκρίνονται παράλληλα στο οπτικό ερέθισμα που πέφτει πάνω της.



7-32: Γενική θεώρηση συνδέσεων GC με V1S ©2011 Horasis

Σύνθετοι V1 νευρώνες

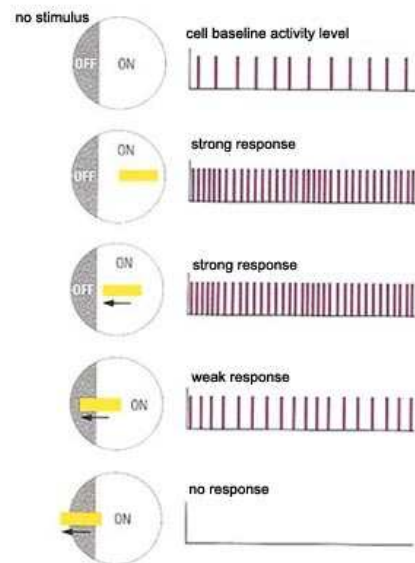


7-33: Δομή υποδεκτικών πεδίων H1C (βασισμένο DeAngelis et al.) [138]

Μια ακόμη ομάδα νευρώνων που έχει ανακαλυφθεί είναι τα υπερσύνθετα κύτταρα (V1 hypercomplex cells), συντετημημένα V1HC, που θεωρούνται πλέον ως υποομάδα των V1S ικανή να ανιχνεύει μήκος και γενικότερα χαρακτηριστικά ημιτερματισμού γραμμών (end-stopped features) [142]. Δηλαδή η λειτουργία τους μοιάζει με τα V1C, αλλά σε αντίθεση με αυτά ανιχνεύουν μια γραμμή μέσα σε μια περιοχή της εικόνας, η οποία μπορεί να κόβεται αντί να εκτείνεται από το ένα όριο του υποδεκτικού πεδίου ως κάποιο άλλο.

Κατά την διάρκεια της ανάπτυξης του συστήματος Horasis, δοκιμάστηκε με επιτυχία η εκπαίδευση τεχνητών νευρωνικών δικτύων που μιμούνται την συμπεριφορά των σύνθετων και υπερσύνθετων νευρώνων του πρωτοταγούς οπτικού φλοιού. Για να μην προστεθεί επιπλέον πολυπλοκότητα στο ήδη σύνθετο μοντέλο της εργασίας, δεν συμπεριλήφθηκαν στην τελική δικτύωση που αναπαριστά τον τεχνητό πρωτοταγή οπτικό φλοιό. Έτσι όπως θα δούμε στην υλοποίηση του περιέχονται μόνο απλά κύτταρα V1, αλλά σε δύο διαφορετικά μεγέθη ώστε να μοντελοποιηθεί η ύπαρξη των P και M κυκλωμάτων της όρασης.

Για να ολοκληρώσουμε τις γενικές γνώσεις της νευροεπιστήμης για την όραση μέχρι το επίπεδο V1, εκτός από τους απλούς νευρώνες του V1 συναντούμε και μια δεύτερη ομάδα κυττάρων τα σύνθετα V1 κύτταρα (V1 complex cells), ή συντετημημένα V1C. Τα κύτταρα αυτά συνειπάρχουν μέσα σε μια προσανατολισμένη υπεστήλη μαζί με τα V1S [141]. Και αυτά τα κύτταρα έχουν προσανατολισμένο υποδεκτικό πεδίο ανά μοίρες, με την διαφορά ότι δεν έχουν σαφώς καθορισμένες ζώνες διέγερσης και αναστολής. Έτσι η ακριβής θέση του ερεθίσματος δεν έχει τόσο μεγάλη σημασία όπως στα V1S αλλά μόνο ο προσανατολισμός με τον οποίο γίνεται η μετάβαση από το ένα ανταγωνιστικό χρώμα στο άλλο [52].



7-34: Διέγερση V1HC με ένα μικρό φωτεινό ευθύγραμμο τμήμα. (The Brain From Top to Bottom) [164]

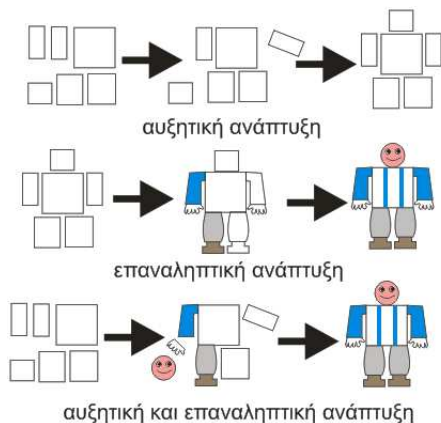
8. Η Αρχιτεκτονική του Horasis

Μετά από την πλεύση στα άγνωστα νερά της νευροεπιστήμης, μέσω των οποίων αποκτήθηκαν οι απαραίτητες εμπειρίες για τον τελικό σχεδιασμό του συστήματος, επιστρέφουμε στην βάση μας που είναι η επιστήμη των υπολογιστών. Έχουμε ήδη δει κάποια στοιχεία μέσα από την υλοποίηση του Horasis και κάποιες θεωρήσεις που έγιναν κατά την διάρκεια της έρευνας, ώστε να βοηθήσουν στον τελικό σχεδιασμό και υλοποίηση του συστήματος. Επίσης έχουμε αναφερθεί σε όλες τις πρώτες ύλες που χρησιμοποιεί το οικοδόμημα του. Σε αυτό το κεφάλαιο πλέον, θα δούμε την ολοκληρωμένη αρχιτεκτονική του, ενώ θα αρχίσουμε να ασχολούμαστε με τον σχεδιασμό και την υλοποίηση του. Έτσι θα γνωρίσουμε γενικά την αρχιτεκτονική του Horasis και θα δούμε συγκεντρωτικά τις αντιστοιχίες που εισάγει για την υλοποίηση της τεχνητής όρασης. Επίσης θα δούμε ειδικότερα πως θεωρούνται οι βιολογικές μονάδες μέσα από το πρίσμα της επιστήμης των υπολογιστών ως τμήματα ενός πληροφοριακού συστήματος λογισμικού και ως υποσυστήματα αυτού.

Η Διεργασία Ανάπτυξης Λογισμικού του Horasis

Η αυξητική και επαναληπτική ανάπτυξη του Horasis

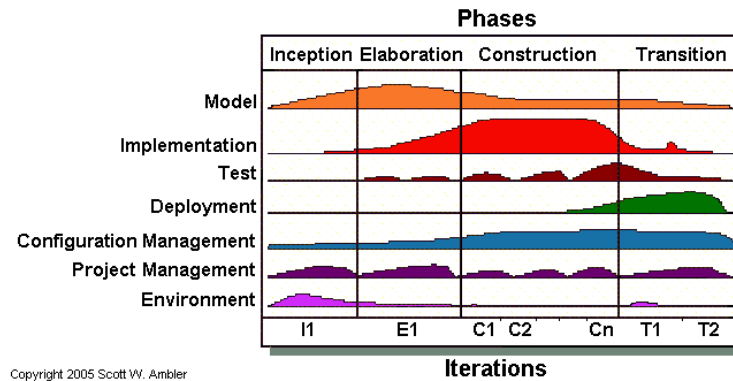
Για να μπορέσει να γίνει πραγματικότητα μια σειρά από ιδέες και σκέψεις, αξιοποιήθηκαν συγκεκριμένες αρχές που δίνει η μηχανική λογισμικού, οριοθετώντας την δουλειά όλης της εργασίας ως μια διεργασία ανάπτυξης λογισμικού (software development process - SDP). Ως διεργασία ορίζουμε ένα σύνολο διατεταγμένων εργασιών, μια σειρά δηλαδή βημάτων που σχετίζονται με δραστηριότητες, πόρους και περιορισμούς και αποσκοπούν στην παραγωγή κάποιου αποτελέσματος που έχει τεθεί ως αρχικός στόχος [143]. Αυτή συχνά ταυτίζεται με τον κύκλο ζωής (life cycle) του προϊόντος, στην δική μας περίπτωση τον κύκλο ζωής του Horasis, επειδή τον περιγράφει από την στιγμή σύλληψης της ιδέας μέχρι την γέννηση του πρωτοτύπου, την χρήση και ωρίμανση του μέσα από βελτιώσεις. Υπάρχουν πολλές προσεγγίσεις για διεργασίες ανάπτυξης λογισμικού στην θεωρία, που αναφέρονται σε ομάδες πολλών ατόμων. Στην εργασία του Horasis, παρόλο που ήταν ένα ατομικό έργο, έγινε μεταφορά κάποιων αρχών της ενοποιημένης διεργασίας (unified process) ανάπτυξης λογισμικού [144] ή συντετημημένα UP.



8-1: Παραστατική σύγκριση, της αυξητικής, της επαναληπτικής ανάπτυξης και του συνδυασμού και των δύο ©2011 Horasis

Η διεργασία αυτή στηρίζεται στην αυξητική (incremental) και επαναληπτική (iterative) ανάπτυξη. Κατά την αυξητική ανάπτυξη το σύστημα κατακερματίζεται σε μικρότερα υποσυστήματα με βάση τις λειτουργίες του ώστε να υλοποιηθούν σταδιακά και ανεξάρτητα. Κατά την επαναληπτική σε κάθε νέα έκδοση του συστήματος οι λειτουργίες του υποσυστήματος τροποποιούνται και εμπλουτίζονται έχοντας αρχικά ένα γρήγορο αλλά ημιτελές σύστημα, αλλά με την πάροδο των εκδόσεων του να τελειοποιείται [143]. Θέτοντας ένα σημείο αναφοράς μπορούμε να πούμε ότι στηρίχθηκε στην Agile

Unified Process (Agile UP) μια προσέγγιση της IBM Rational Unified Process (RUP) που είναι διαθέσιμη ελεύθερα στο διαδίκτυο [145]. Έτσι για να γίνει πραγματικότητα η σκέψη ενός συστήματος τεχνητής όρασης αξιοποιήθηκαν οι αρχές (disciplines) της UP που οριοθετούν τις διατεταγμένες εργασίες, δημιουργώντας ένα είδος πειθαρχίας στην σκέψη και οργάνωση τους. Από όλες χρησιμοποιήθηκαν κυρίως οι αρχές της μοντελοποίησης (model), υλοποίησης (implementation), δοκιμής (test) συνεπικουρούμενες από μια διαχείριση αλλαγών και σύνθεσης (change and configuration) και ένα περιβάλλον ανάπτυξης (environment). Οι αρχές στην UP. Την στιγμή της συγγραφής της εργασίας, την οποία θα επακολουθήσει η παρουσίαση του Horasis, μπορεί να θεωρηθεί ότι ως έργο βρίσκεται στο στάδιο της παράταξης (deployment).



8-2: Οι αρχές της Agile Unified Process (Ambler) [145]

Στο προηγούμενο χρονικό διάστημα και με αναρίθμητες επαναλήψεις (iterations) γινόταν μοντελοποίηση, υλοποίηση και δοκιμή με διαφορετική βαρύτητα ανάλογα με την πρόοδο του έργου μέσα από τέσσερις βασικές φάσεις (phases). Το πρώτο είναι η σύλληψη (inception) του έργου, το δεύτερο η επεξεργασία (elaboration), τρίτο η κατασκευή (construction) και τέταρτο η μετάβαση (transition). Κατά την σύλληψη του έργου ορίστηκε ο σκοπός του, δηλαδή η αναγνώριση από τον υπολογιστή, πραγματικών αντικειμένων μέσα από έγχρωμες εικόνες τους. Επίσης εξετάστηκαν τα ρίσκα και η βιωσιμότητα της λύσης και έγινε προετοιμασία για το περιβάλλον ανάπτυξης που περιλάμβανε:

- Την προμήθεια όλου του απαιτούμενου υλικού όπως νέος υπολογιστής, κάμερες, υλικά για την κατασκευή της σκηνής κ.α.
- Την εγκατάσταση όλων των απαιτούμενων εφαρμογών λογισμικού.
- Την δημιουργία ενός πλαισίου λογισμικού (framework) αντικειμενοστραφών νευρωνικών δικτύων.
- Την απόκτηση ενός πλήθους γνώσεων πάνω στην νευροεπιστήμη ως βάση για την δημιουργία της προσομοίωσης της όρασης.

Κατά τις φάσεις της επεξεργασίας και κατασκευής εμπλουτιζόταν το μοντέλο του συστήματος με προσθήκη επιπλέον κομματιών του σχεδιασμού και βελτίωση των υπαρκτών. Κύρια εργασία ήταν η συγγραφή του κώδικα σε γλώσσα Object Pascal μέσα στο RAD περιβάλλον της Delphi. Μετά την κάθε συγγραφή, γινόταν γρήγορες δοκιμές μονάδας (unit testing) [146] [147] που έφταναν από την λειτουργία του κάθε αντικειμένου μέχρι εσωτερικά στην κάθε κρίσιμη μέθοδο γι αυτήν. Τέλος γινόταν δοκιμές ενοποίησης (integration testing) [148] για την σωστή λειτουργία των νέων αντικειμένων μέσα στο όλο πλαίσιο της εφαρμογής. Η εξέλιξη κάποιου σταδίου της εκπόνησης της εργασίας, μπορεί να οδηγήσει ξανά στην υλοποίηση ή τον σχεδιασμό για την

διόρθωση των τυχών προβλημάτων. Έτσι αυξητικά το σύστημα σχηματιζόταν στην τελική σταθερή μορφή του.

Η αρχή UP της μοντελοποίησης το όραμα και η τεκμηρίωση της αρχιτεκτονικής

Στην άλλη διάσταση της UP τον άξονα των αρχών και συγκεκριμένα της μοντελοποίησης έγινε αρχικά ο καθορισμός του οράματος (vision) της εφαρμογής. Ως προς το τελικό αποτέλεσμα τέθηκε ως όραμα από τον επιβλέπων καθηγητή η κατηγοριοποίηση μερικών απλών πραγματικών αντικειμένων. Έτσι μπορούμε να θεωρήσουμε ότι έχουμε και εδώ δύο ενδιαφερόμενα μέρη (stakeholders) όπως ορίζει και η θεωρία της ενοποιημένης διεργασίας [143] [145] με το πρώτο να είναι ο φοιτητής και το δεύτερο ο καθηγητής. Το όραμα αυτό εμπλουτίστηκε μέσα από την έρευνα πάνω στην νευροεπιστήμη ώστε να παρουσιαστεί μια νέα πρόταση πάνω στην αρχική απαίτηση της εργασίας. Μετά από πολλές επαναλήψεις σε όλο τον κύκλο ζωής του Horasis, αποφασίστηκε και σταθεροποιήθηκε το τελικό όραμα. Έτσι επιπρόσθετα στις ανάγκες της πτυχιακής εργασίας καθορίστηκε ότι έπρεπε να λειτουργεί με έγχρωμα πρότυπα εικόνων και παράλληλα ότι έπρεπε να μοντελοποιεί αρκετές λειτουργίες των νευρώνων του αμφιβληστροειδούς και του πρωτοταγή οπτικού φλοιού ως αντικείμενα τεχνητών νευρωνικών δικτύων. Για να γίνει πιο εφικτή η επίτευξη αυτού του οράματος αποφασίστηκε να χρησιμοποιηθούν απλά καθημερινά αντικείμενα, πάνω σε ένα απλοποιημένο μαύρο φόντο, ως πρότυπα για την πρώτη πλήρως λειτουργική αναγνώριση.

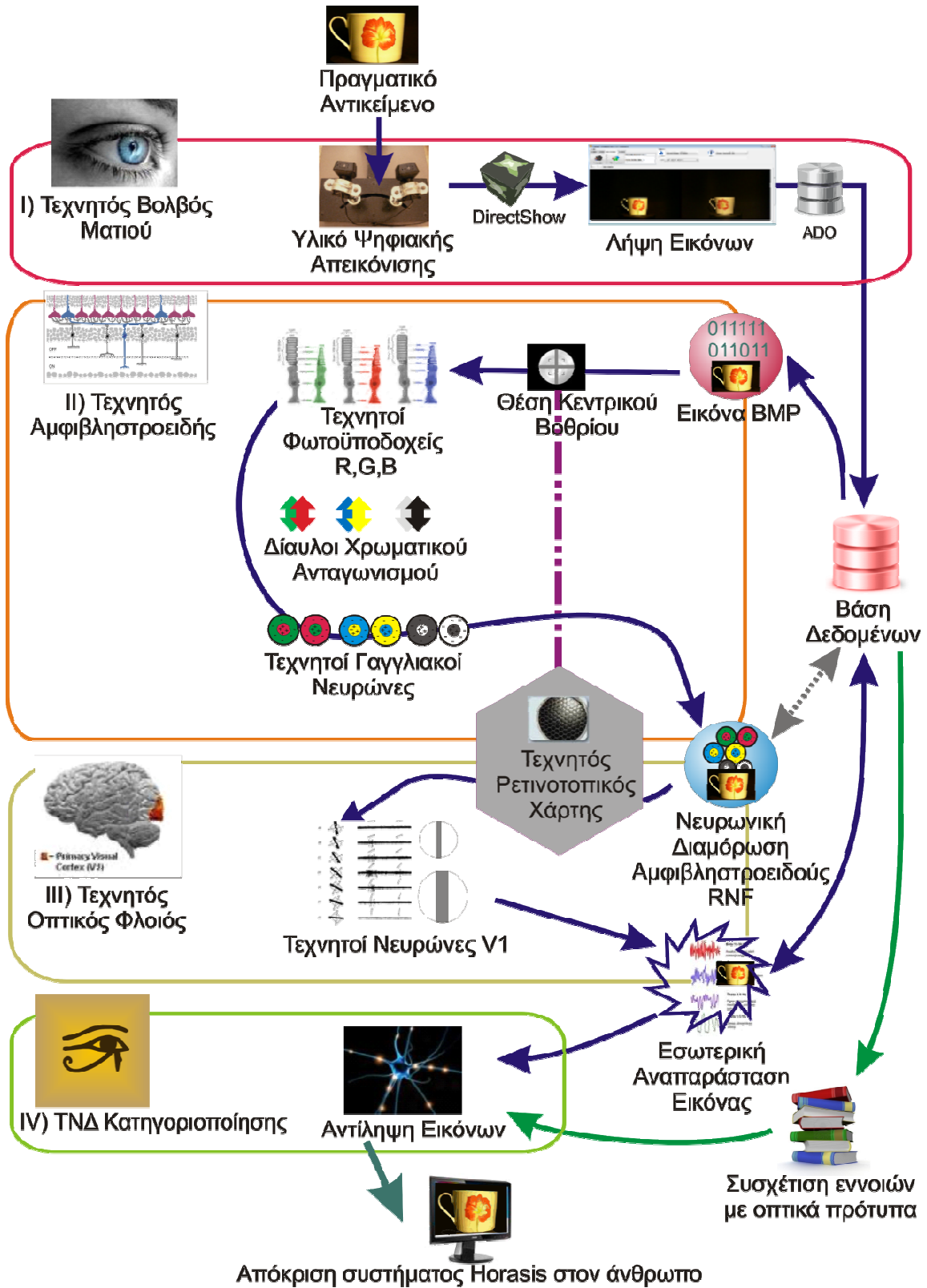
Επίσης κατά το στάδιο της μοντελοποίησης έγινε η χρήση ενός κοινού λεξιλογίου με την νευροεπιστήμη μέσα στο μοντέλο. Παράδειγμα του τελευταίου ήδη έχουμε δει με την ύπαρξη κλάσεων όπως TGSoma, TGAxon, TGDendrite μέσα στον αντικειμενοστραφή νευρώνα. Κατά την μοντελοποίηση και χρονικά πολύ πρώιμα μέσα στην φάση της σύλληψης οι πιο σημαντικές αποφάσεις που ελήφθησαν ήταν:

- Η χρήση αντικειμενοστραφούς προγραμματισμού για βιωσιμότητα απέναντι στην πολυπλοκότητα του εγχειρήματος
- Η χρήση RAD εργαλείων για υλοποίηση του μέσα σε λογικά χρονικά πλαίσια μερικών μηνών
- Η χρήση σχεδιασμού τριών επιπέδων για σαφή καθορισμό της διάδραστικότητας, της λειτουργικότητας και των δεδομένων της εφαρμογής.
- Η χρήση βάσεων δεδομένων ως de facto αρχιτεκτονική για την διαχείριση μεγάλου όγκου δεδομένων.
- Η χρήση του back propagation ως αλγόριθμο για την εκπαίδευση όλων των τεχνητών νευρωνικών δικτύων, για αύξηση των πιθανοτήτων επιτυχών εκπαιδεύσεων.
- Η χρήση απλού και φθηνού υλικού για την επίδειξη ενός μειωμένου κόστους σε μια λύση υψηλής τεχνολογίας.
- Η χρήση όλων των δυνατοτήτων που βοηθούν και προέρχονται από την μηχανική λογισμικού, αξιοποιούμενες συνήθως μέσα σε επιχειρηματικά και εμπορικά πλαίσια, στην υπηρεσία της έρευνας της τεχνητής νοημοσύνης

Όλες οι παραπάνω αποφάσεις καθόρισαν την υποψήφια αρχιτεκτονική που θα ακολουθούσε το σύστημα, ενώ έχουμε γνωρίσει ήδη πτυχές της εργασίας που υλοποιούν, μέχρι αυτό το σημείο. Εδώ κλείνουμε έναν ανοικτό κύκλο, που αφορά τις τεχνολογίες και επιστημονικές προσεγγίσεις που αναφέρθηκαν, συνδέοντας τις με τις αντίστοιχες αποφάσεις που ελήφθησαν.

Η Αρχιτεκτονική του Horasis

Το γενικό αρχιτεκτονικό πλάνο του συστήματος Horasis



8-3: Παραστατική απεικόνιση της αρχιτεκτονικής του συστήματος «Όρασις»

Με τα μπλε βέλη κατεύθυνσης φαίνεται η ροή της οπτικής πληροφορίας και με τα πράσινα της εννοιολογικής. Τα τέσσερα χρωματιστά πλαίσια αντιπροσωπεύουν τα όρια των τεσσάρων τμημάτων του συστήματος.

Όπως φαίνεται στο σχεδιάγραμμα 8-3 το σύστημα έχει τέσσερα βασικά τμήματα στα οποία περικλείονται τα υποσυστήματα του:

- **I) Τμήμα Τεχνητού Βολβού**
- **II) Τμήμα Τεχνητού Αμφιβληστροειδούς**
- **III) Τμήμα Τεχνητού Οπτικού Φλοιού**
- **IV) Τμήμα Αντίληψης**

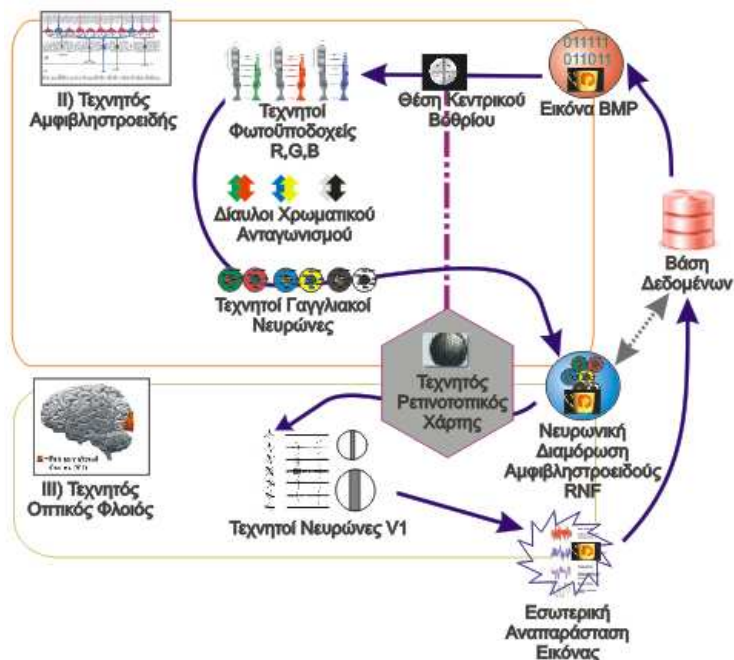
Η δράση του οριοθετείται σε τρεις διαφορετικές ροές εργασίας (workflows):

- 1) Την φωτογράφιση ενός πραγματικού αντικείμενου και την αποθήκευση της εικόνας BMP στην βάση δεδομένων.



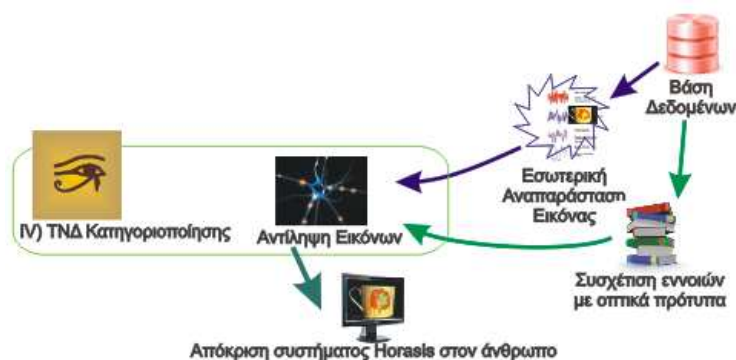
8-4: Ροή εργασίας για την φωτογράφιση ενός πραγματικού αντικείμενου.

- 2) Την ανάκληση την αποθηκευμένης εικόνας BMP μέσω του ΤΝΔ των τμημάτων II και III, που την μετατρέπουν σε εσωτερική αναπαράσταση εικόνας. Η διαδικασία επαναλαμβάνεται για κάθε πραγματικό αντικείμενο και όλες οι αναπαραστάσεις αποθηκεύονται στην βάση δεδομένων.



8-5: Ροή εργασίας για την μετατροπή της εικόνας σε εσωτερική αναπαράσταση

- 3) Η εκπαίδευση του τελικού ΤΝΔ που κατηγοριοποιεί τα πραγματικά αντικείμενα με ένα σύνολο εκπαίδευσης από αποθηκευμένες αναπαραστάσεις.



8-6: Ροή εργασίας για την εκπαίδευση του ΤΝΔ κατηγοριοποίησης

Με τον πράσινο βέλος που ξεκινάει από την βάση δεδομένων αναπαρίσταται μια ροή εννοιολογικής πληροφορίας σε αντιδιαστολή με το κόκκινο βέλος που αναπαριστά την ροή της οπτικής πληροφορίας.

Όπως φαίνεται ότι δεν μας ενδιαφέρει ο τρόπος με τον οποίο παράγεται η συσχέτιση των οπτικών προτύπων με τις έννοιες. Καθώς η εργασία είναι επικεντρωμένη πάνω στην λειτουργία της τεχνητής όρασης δεν κρίθηκε σκόπιμο να γίνει μια περαιτέρω σχεδιασμός που θα αφορούσε αυτήν την συσχέτιση. Οι πληροφορίες τοποθετούνται στην βάση δεδομένων δίπλα σε κάθε οπτικό πρότυπο χειροκίνητα, και θα μπορούσε να κατασκευαστεί εύκολα κάποιο υποσύστημα λογισμικού για τον καθορισμό τους.

Τα τμήματα της αρχιτεκτονικής του Horasis

Το πρώτο τμήμα το έχουμε ήδη γνωρίσει στο κεφάλαιο 5, και μπορεί να αυτονομηθεί από τα υπόλοιπα ώστε να τοποθετηθεί σε έναν μικρών δυνατοτήτων υπολογιστή, που θα μπορεί να αποτελεί μέρος ενός αυτοματισμού όπως ένα ρομπότ. Θα το ονομάσουμε **Τεχνητό Βολβό Μαιού (Artificial Eyeball - AEB)** καθώς είναι αυτό που περιλαμβάνει υλικό και λογισμικό που μιμούνται την λειτουργία του βιολογικού βολβού. Το φως εισάγεται στην είσοδο ενός τέτοιου συστήματος και στην έξοδο του παράγεται μια ψηφιακή μορφή της εικόνας, ως αντιστοιχία με την διαδικασία αναστροφής και σμίκρυνσης της εικόνας που συμβαίνει στο ανθρώπινο μάτι.

Το δεύτερο τμήμα που είναι ο **Αντικειμενοστραφής Τεχνητός Αμφιβληστροειδής (Object-Oriented Artificial Retina - OOAR)**, είναι ένα τμήμα του λογισμικού με αντικειμενοστραφή σχεδίαση, που μοντελοποιεί και εμφωλεύει όλη την επεξεργασία του βιολογικού αμφιβληστροειδή. Δέχεται ως είσοδο μια εικόνα σε ψηφιδωτή μορφή και παράγει ως έξοδο μια νέα πληροφορία νευρωνικής διαμόρφωσης αμφιβληστροειδούς RNF, η οποία έχει προκαθορισμένα χαρακτηριστικά τοποθετημένα σε χωρική αντιστοιχία με την εικόνα της εισόδου. Η μορφή αυτή μπορεί να τροφοδοτείται προς τα εμπρός στο επόμενο τμήμα άμεσα, ή να αποθηκεύεται στην βάση δεδομένων και να ανακτάται σε κάποια άλλη χρονική στιγμή.

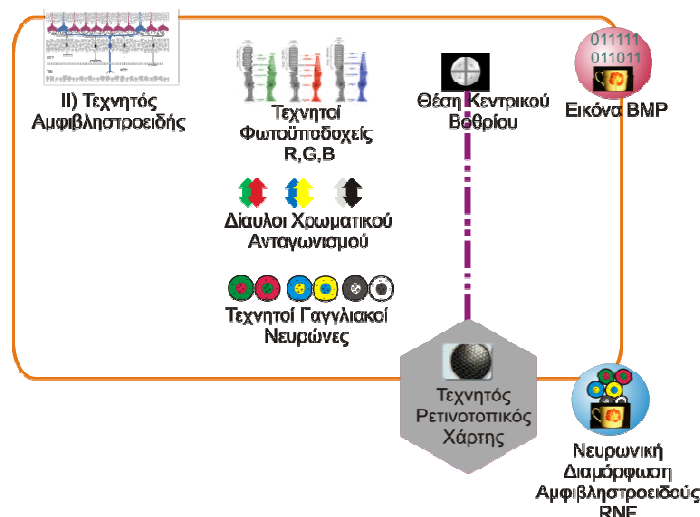
Για να υλοποιηθεί αυτή η χωρική αντιστοιχία υπάρχει ένα κοινόχρηστο υποσύστημα μεταξύ των τμημάτων II και III που ονομάζεται **Τεχνητός Ρετινοτοπικός Χάρτης (Artificial Retinotopic Map - ARTM)** και στο οποίο υλοποιείται μια αντιστοίχιση μεταξύ των εικονοστοιχείων (pixels) της εικόνας με τα υποδεκτικά πεδία των τεχνητών νευρώνων του αμφιβληστροειδούς. Επίσης βοηθάει στην διασύνδεση των δύο συστημάτων με αντιστοίχιση των τεχνητών νευρώνων εξόδου του αμφιβληστροειδούς για τον κάθε νευρώνα του οπτικού φλοιού. Έτσι τα εικονοστοιχεία της αρχικής εικόνας θα απεικονίζονται στην τελική ανάλυση ως είσοδος σε νευρώνες του οπτικού φλοιού, όπως ακριβώς με τον βιολογικό ρετινοτοπικό χάρτη.

Στο τρίτο τμήμα του τεχνητού οπτικού φλοιού τα δεδομένα εξόδου του δεύτερου τμήματος ΟΟΑΡ, θα εισάγονται ακολουθώντας τον χάρτη ARTM σε Τεχνητούς Νευρώνες V1 (Artificial V1 Cells – **AV1S**) . Αυτό θα αποτελεί τον **Τεχνητό Οπτικό Φλοιό (Artificial Visual Cortex – AVC)** δηλαδή ένα λογισμικό το οποίο παίρνει είσοδο από έναν τεχνητό αμφιβληστροειδή και παράγει στην έξοδο του θα παράγει μια νέα πληροφορία στην οποία προσαρτώνται νέα χαρακτηριστικά σχήματος και μορφής. Ένα τέτοιο λογισμικό θα υλοποιεί μια λειτουργία αναγνώρισης σχημάτων και μορφής, που προσθέτει νέες διαστάσεις στην οπτική πληροφορία. Η έξοδος αυτή θα αποτελεί την εσωτερική αναπαράσταση της εικόνας του αντικειμένου, για το σύστημα τεχνητής νοημοσύνης που το περιέχει. Δηλαδή το τμήμα III σε συνεργασία με το τμήμα II χρησιμοποιούνται για την παραγωγή των προτύπων, υπό την μορφή εσωτερικής αναπαράστασης. Αυτό γίνεται με ανάκληση της αρχικής ψηφιδωτής εικόνας μέσα από το ενιαίο ΤΝΔ που συνιστούν τα δύο τμήματα. Η εσωτερική αναπαράσταση αυτή μπορεί να εκπαιδεύσει τεχνητά νευρωνικά δίκτυα που την συνδέουν με κάποια έννοια ή θα την κατηγοριοποιούν σε κάποια κλάση. Στην αρχιτεκτονική του Horasis αποθηκεύεται η εσωτερική αναπαράσταση σε ένα blob πεδίο της βάσης δεδομένων ώστε να ανακτηθεί σε άλλη χρονική στιγμή ως πρότυπο εκπαίδευσης για το ΤΝΔ κατηγοριοποίησης.

Το ΤΝΔ αυτό ανήκει στο τέταρτο τμήμα, όπου υλοποιείται η αντίληψη από το σύστημα τεχνητής νοημοσύνης της εικόνας ενός πραγματικού αντικειμένου. Αυτή ακριβώς είναι και η λειτουργία του υποσυστήματος αντίληψης εικόνων. Ανάλογα με τις τιμές των νευρώνων του στρώματος εξόδου στο ΤΝΔ κατηγοριοποίησης του Horasis μπορούν να ληφθούν αποφάσεις, όπως η πληροφόρηση του χρήστη μέσω της οθόνης για την κατηγορία στην οποία ανήκει το αντικείμενο. Το τμήμα IV είναι αυτόνομο και δεν βρίσκεται σε αντιστοίχιση με την βιολογική πραγματικότητα, καθώς ακόμα δεν υπάρχει σαφής περιγραφή από την νευροεπιστήμη. Είναι ένα απλό ΤΝΔ κατηγοριοποίησης που θα εκπαιδευτεί με πρότυπα εσωτερικής αναπαράστασης των πραγματικών αντικειμένων. Μπορούν να σχεδιαστούν νέα εντελώς συστήματα τεχνητών νευρωνικών δικτύων τα οποία να προσθέτουν στην λειτουργικότητα του ή να αλληλεπιδρούν με αυτό, ως εντελώς αυτόνομες εργασίες.

Έχοντας ολοκληρώσει την περιήγηση στην αρχιτεκτονική του Horasis βλέπουμε ότι έχουν εγκιβωτίσει σε αυτήν τα βασικά χαρακτηριστικά ενός πρότυπου συστήματος τεχνητής όρασης. Θα μπορούσαν να προκύψουν νέες ιδέες και διαφορετικές υλοποιήσεις κρατώντας την αρχιτεκτονική αυτή ως υπόδειγμα και την πρώτη υλοποίηση σε Object Pascal ως παράδειγμα (demo). Αν τυχόν υπάρξουν τέτοιες περιπτώσεις τότε θα έχει υλοποιηθεί ένας ακόμη δευτερεύοντας στόχος της εργασίας, που είναι η δημιουργία ενός νέου κύματος τεχνητών νευρωνικών δικτύων με πρακτική εφαρμογή στην τεχνητή όραση.

Τα υποσυστήματα του τεχνητού αμφιβληστροειδή



Εικόνα 8-7: Τμήμα II της αρχιτεκτονικής Horasis: Τεχνητός Αμφιβληστροειδής
©2011 Horasis

Όπως είδαμε στο γενικό πλάνο, στο κέντρο της αρχιτεκτονικής βρίσκεται το τμήμα II, που αντιπροσωπεύει τον τεχνητό αμφιβληστροειδή. Αποτελείται από τα παρακάτω υποσυστήματα:

- Τεχνητών Φωτοϋποδοχέων **APC** (αρχικά των Artificial Photoreceptor Cell)
- Διαύλων Χρωματικού Ανταγωνισμού **INL** (αρχικά των Inner Nuclear Layer)
- Τεχνητών Γαγγλιακών Νευρώνων **AGC** (αρχικά των Artificial Ganglion Cell)
- Τεχνητού Ρετινοτοπικού Χάρτη **ARTM** (αρχικά των Artificial Retino-Topic Map)
- Τοποθέτηση Κεντρικού Βοθρίου **FCP** (αρχικά των Fovea Centralis Placement)

Στο πρώτο υποσύστημα APC, περιλαμβάνονται αντικείμενα που προσομοιώνουν την λειτουργία των φωτοϋποδοχέων του ματιού, τροφοδοτώντας τιμές κόκκινου R, πράσινου G και μπλε B στους επόμενους νευρώνες. Για κάθε εικονοστοιχείο της εικόνας μπορούμε να θεωρήσουμε τρία ιδεατά κωνία R,B,G από το χρώμα του, και ένα ιδεατό ραβδί W που λειτουργεί με την μετατροπή των R,G,B σε κλίμακα του γκρι, έχοντας ως τιμή του το φώς (luma).

Στο υποσύστημα διαύλων χρωματικού ανταγωνισμού INL, υπάρχουν τεχνητοί νευρώνες αντιστοιχεί με τα οριζόντια κύτταρα HC, βραχύινα κύτταρα AC και κυρίως τεχνητά νευρωνικά δίκτυα με νευρώνες στα στρώματα εξόδων που αντιπροσωπεύουν τα δίπολα κύτταρα BPC. Με όρους της μεθοδολογίας του Ευθυνοκεντρικού Σχεδιασμού RDD⁷¹ (Responsibility-Driven Design) [149] [150] το υποσύστημα είναι επιφορτισμένο με τις παρακάτω ευθύνες:

- ✓ Να διαβάζει τις τιμές των φωτοϋποδοχέων από το υποσύστημα APC
- ✓ Να τροφοδοτεί τα R,G,B,W σε τεχνητά νευρωνικά δίκτυα που έχουν στην έξοδο τους έναν τεχνητό νευρώνα για κάθε ανταγωνιστικό χρώμα.
- ✓ Να τροφοδοτεί την έξοδο των παραπάνω νευρώνων στο υποσύστημα AGC, ως είσοδο σε ένα κυκλικό υποδεκτικό πεδίο που έχει οργάνωση κέντρου-περιφέρειας.
- ✓ Να ρυθμίσει την έξοδο των παραπάνω νευρώνων, ώστε αυτές να απενεργοποιείται όταν το φως W πέσει κάτω από ένα επίπεδο.

⁷¹ Είναι μια μεθοδολογία σχεδίασης του αντικειμενοστραφούς προγραμματισμού που προτάθηκε ως όρος από την Rebecca Wirfs-Brock και τον Brian Wilkerson, σχετιζόμενος την πρόταση των Ward Cunningham και Kent Beck και τα δύο προϊόντα του συνεδρίου OOPSLA το 1989.

Έτσι στο υποσύστημα INL τα τρία βασικά χρώματα κόκκινο, πράσινο, μπλε θα διαχωριστούν σε έξι γραμμές κόκκινο R, πράσινο G, μπλε B, κίτρινο Y, άσπρο (φως) L, και μαύρο (σκοτάδι) D.

Το παρακάτω υποσύστημα AGC που υλοποιεί την αντιστοιχία με τα βιολογικά γάγγλια έχει τις ακόλουθες ευθύνες.

- ✓ Να έχει ένα ΤΝΔ, που θα αντιπροσωπεύει ένα τεχνητό γαγγλιακό νευρώνα.
- ✓ Να παράγει στην έξοδο ενός τέτοιου ΤΝΔ ένα διάνυσμα, που αντιπροσωπεύει την κωδικοποίηση γαγγλιακού κυττάρου GCE.
- ✓ Να οργανώνει το υποδεκτικό πεδίο του παραπάνω ΤΝΔ με κέντρο και περιφέρεια, έχοντας χωρική συσχέτιση με μια περιοχή της εικόνας.
- ✓ Να χρησιμοποιεί τον τεχνητό ρετινοτοπικό χάρτη, που οργανώνει τις χωρικές συσχετίσεις, όλων των τεχνητών γαγγλίων.
- ✓ Να έχει για την ίδια περιοχή της εικόνας, τόσα γάγγλια όσα τα ανταγωνιστικά χρώματα, έχοντας διέγερση του κέντρου τους με το αντίστοιχο χρώμα.
- ✓ Να συγκεντρώνει όλα τα διανύσματα GCE σε μια ενιαία πληροφορία, που αποτελεί την νευρωνική διαμόρφωση αμφιβληστροειδούς RNF

Βλέπουμε ότι η έξοδος του υποσυστήματος θα είναι μια πληροφορία RNF που θα περιλαμβάνει για κάθε τμήμα της εικόνας, βάσει του ρετινοτοπικού χάρτη, ένα σύνολο από αποκρίσεις νευρώνων που υλοποιούν χρωματικό ανταγωνισμό με αντίθεση κέντρου και περιφέρειας. Οι τιμές των εξόδων είναι ανάλογες με τους χρωματισμούς που υπάρχουν σε αυτό το τμήμα της εικόνας

Ο ρετινοτοπικός χάρτης που αποτελεί το υποσύστημα ARTM τοποθετείται σε αυτό το τμήμα της αρχιτεκτονικής, παρόλο που είναι κοινόχρηστος με το τμήμα III του τεχνητού οπτικού φλοιού. Αυτό οφείλεται στο γεγονός ότι, ο ρετινοτοπικός χάρτης δημιουργείται μέσα στο τμήμα II που περιγράφουμε, δημιουργώντας στην ουσία ένα κάρναβο (grid) και χωρίζοντας την εικόνα σε περιοχές. Η εικόνα στην ψηφιδωτή της μορφή έχει ήδη ένα κάρναβο με τετραγωνίδια ή αλλιώς μια τετραγωνική ψηφίδωση (tesselation) [112]. Αυτό σημαίνει ότι στο υποσύστημα APC το μωσαϊκό των τεχνητών φωτουποδοχέων δεν ταιριάζει με το βιολογικό ψηφιδωτό που έχει σχεδόν εξαγωνική ψηφίδωση. Έχοντας ως πηγή έμπνευσης το τελευταίο η εξαγωνική ψηφίδωση μεταφέρθηκε στο υποσύστημα ARTM, όπου υλοποιήθηκε ένας αντίστοιχος εξαγωνικός κάρναβος.

Το τελευταίο υποσύστημα FCP είναι ένα ιδεατό υποσύστημα που ορίστηκε ως μελλοντική επέκταση του σχεδιασμού για την μετακίνηση του κέντρου της τεχνητής όρασης. Το απόλυτο κέντρο της όρασης αντιπροσωπεύει η κεντρικότερη εξαγωνική κυψελίδα μέσα στην κυψέλη του ARTM. Το υποσύστημα αυτό έχει μια στοιχειώδη υλοποίηση καθώς δεν γίνεται καμία μετακίνηση μέσα στο Horasis, αλλά απλώς ορίζονται στατικές συντεταγμένες για το κέντρο της όρασης. Το υποσύστημα παίρνει την ονομασία του από το κέντρο της βιολογικής όρασης που είναι το κεντρικό βοθρίο (fovea centralis).

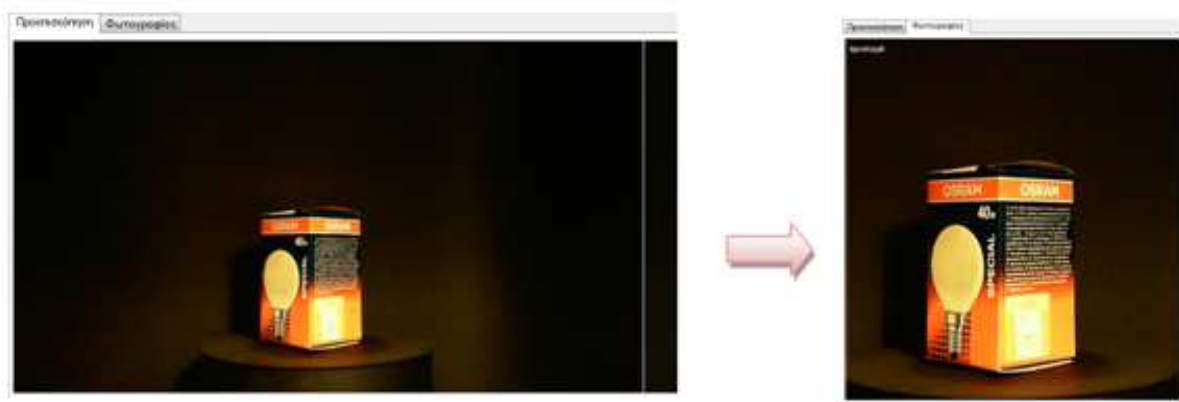
9. Η Μοντελοποίηση της Όρασης

Έχοντας γνωρίσει το γενικό πλάνο της αρχιτεκτονικής του συστήματος, είμαστε πλέον σε θέση να ανοίξουμε τον σχεδιασμό των δύο σημαντικότερων τμημάτων της, του τεχνητού αμφιβληστροειδούς και του τεχνητού φλοιού. Με ανάκληση των αντικειμενοστραφών ΤΝΔ παράγεται η εσωτερική αναπαράσταση του πραγματικού αντικειμένου στο Horasis. Εδώ υπάρχει ο κυριότερος όγκος της εργασίας καθώς γίνεται η προσπάθεια της μοντελοποίησης της ανθρώπινης όραση. Για το κάθε τμήμα θα δούμε πιο αναλυτικά τα υποσυστήματα μαζί με κάποια χαρακτηριστικά της υλοποίησής τους. Σύμφωνα με την διεργασία ανάπτυξης του Horasis, που όπως είπαμε είναι UP, στο κεφάλαιο αυτό θα δούμε ένα σημαντικό τεχνούργημα της UP (UP artifact) [151] το μοντέλο του σχεδιασμού (design model). Αυτό δημιουργήθηκε αρχικά στην φάση της επεξεργασίας (elaboration) ,μετά από πολλές επαναλήψεις (iterations) βελτιώθηκε και τελειοποιήθηκε τελικά στην φάση της κατασκευής (construction) .

Τεχνητοί Φωτοϋποδοχείς

Περικοπή της εικόνας σε τετράγωνο

Όπως έχουμε αναφέρει το αντικείμενο φωτογραφίζεται και μετατρέπεται σε ψηφιδωτή διαμόρφωση τύπου BMP. Αρχικά οι δύο εικόνες για την αριστερή και για την δεξιά κάμερα, έρχονται στην μνήμη με μέγεθος 1280x720 pixels, μια ανάλυση HD⁷² με αναλογία 16:9. Στο πρώιμο στάδιο της σύλληψης μέσα στον κύκλο ζωής του Horasis, θεωρήθηκε ότι 921.600 pixels επί τρία χρώματα R,G,B δηλαδή 2.764.800 διαφορετικές τιμές εισόδου σε ένα νευρωνικό δίκτυο είναι ένα πολύ μεγάλο πλήθος για να γίνει γρήγορη επεξεργασία του και εκτός των στόχων της εργασίας. Γι αυτό αποφασίστηκε η περικοπή του ορθογωνίου παραλληλογράμμου σε ένα τετράγωνο, που θα μπορούσε να καλύπτει οποιοδήποτε μικρό αντικείμενο φωτογραφιζόταν πάνω στο βάθρο της σκηνής φωτογράφισης.

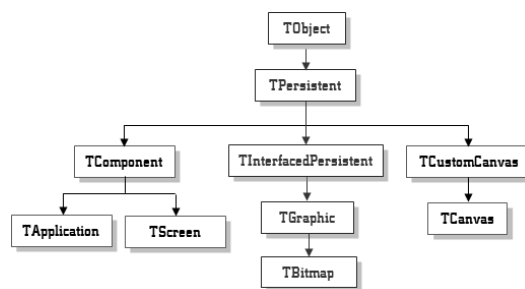


9-1: Αριστερά η προεπισκόπηση με ανάλυση 1280x720 και δεξιά η τελική εικόνα 500x500 pixels

⁷² HD – High Definition

Η κάθε εικόνα περικόπτεται σε ένα μέγεθος 500x500 pixels, έχοντας 750.000 διαφορετικά δεδομένα RGB προς είσοδο στο ΤΝΔ, μια μείωση στο 27% από τον αρχικό όγκο της HD εικόνας. Κατά τις δοκιμές ανάπτυξης, τροποποιήθηκε εμπειρικά η θέση της κάθε μιας από τις κάμερες, ώστε πάνω μέρος του βάθρου να εμφανίζεται στο κάτω μέρος της εικόνας και το αντικείμενο να τοποθετείται σχεδόν στο κέντρο της. Όταν ένα τετράγωνο, το βαρύκεντρο της εικόνας υπολογίζεται εύκολα στις συντεταγμένες {250,250} και αυτό θα θεωρείται το κέντρο της όρασης.

Η τελική εικόνα σε μορφή BMP 24bit, αποθηκεύεται στην βάση δεδομένων και μετά την φόρτωση της επαναφέρεται ακριβώς όπως λήφθηκε, καθώς το BMP είναι μια διαμόρφωση στην οποία δεν γίνεται συμπίεση και δεν έχουμε απώλεια πληροφορίας (lossless encoding⁷³). Κάθε εικόνα στην Delphi περιέχεται σε ένα αντικείμενο της κλάσης TBitmap και γίνεται πρόσβαση σε αυτήν μέσω των μεθόδων και ιδιοτήτων της κλάσης.



9-2: Ιεραρχία κλάσης TBitmap

Πρόσβαση στις χρωματικές συνιστώσες ενός εικονοστοιχείου μέσω scanlines

Κάθε εικονοστοιχείο έχει δύο συντεταγμένες στήλη X, και γραμμή Y που ορίζουν την θέση του στην εικόνα. Εφόσον έχουμε μια εικόνα BMP 24bit φορτωμένη σε ένα αντικείμενο TBitmap, μπορούμε να βρούμε τις τιμές των χρωματικών συνιστωσών κόκκινου R, πράσινου G και μπλε B για μια θέση {X,Y}. Με την ιδιότητα πίνακα (array property⁷⁴) TBitmap.ScanLines[] επιστρέφεται ένας δείκτης, μέσα στα δεδομένα του bit-map της εικόνας, ανάλογα με την γραμμή Y. Ο δείκτης είναι τύπου PGCOLORARRAY και δείχνει έναν πίνακα (array) που έχει ως στοιχεία εγγραφές TGColorStruct.

```

TGColorStruct = record
    Blue      : Byte;
    Green     : Byte;
    Red       : Byte;
end;
TGColorArray = array [0..65535] of TGColorStruct;
PGColorArray = ^TGColorArray;
  
```

⁷³ Διαμορφώσεις εικόνας όπως η JPEG (JPG) όταν φορτώνονται δεν αντιστοιχούν 100% στο αρχικό δείγμα, λόγω της συμπίεσης που γίνεται κατά την αποθήκευσή τους.

⁷⁴ Οι array properties της Delphi αντιστοιχεί στις indexed properties της C# και είναι ιδιότητες που έχουν δείκτη όπως τα απλά arrays.

Οπότε η πρόσβαση στο εικονοστοιχείο στις συντεταγμένες {X, Y} γίνεται ως εξής:

```
RowPointer      := Bitmap.ScanLine[Y];
PixelPointer    := @RowPointer^[X];
```

Εφόσον βρεθεί ο τελικός δείκτης μέσα στα δεδομένα της εικόνας μπορούμε να διαβάσουμε τις χρωματικές συνιστώσες του εικονοστοιχείο στο σύστημα R,G,B. Έτσι μπορούμε να ζητήσουμε με χρήση του δείκτη PixelPointer, την επιστροφή μια τιμής από το 0 έως το 255 για κάθε ένα από αυτά:

```
PixelPointer    ^.Red
PixelPointer    ^.Green
PixelPointer    ^.Blue
```

Το αντικείμενο φωτοϋποδοχέων σημείου

Επειδή όπως είδαμε παραπάνω έχουμε χρήση δεικτών, ένα είδος προγραμματισμού που πρέπει να αποφεύγεται σε αντικειμενοστραφείς υλοποιήσεις. Γι αυτό η παραπάνω λογική εμφωλεύεται στην κλάση TPhotoReceptors, για να αποκρύψει τις λεπτομέρειες της απαιτούμενης υλοποίησης με δείκτες. Η κλάση αυτή αντιπροσωπεύει τους τεχνητούς φωτοϋποδοχείς που υπάρχουν σε ένα τετραγωνισμένο σημείο του αμφιβληστροειδούς.

Σε τυχόν αλλαγή του χαμηλού επιπέδου (low-level) τρόπου πρόσβασης ο υπόλοιπος σχεδιασμός θα μείνει ανέπαφος, καθώς το αντικείμενο επιστρέφει δύο ιδιότητες την .Rod και την ιδιότητα πίνακα .Cone[]. Βλέπουμε στην πράξη μηδενική σύζευξη του υπόλοιπου κώδικα με τα scanlines και υψηλή συνοχή καθώς οι τιμές των τεχνητών φωτοϋποδοχέων ενός σημείου της εικόνας ταυτίζονται με τις τιμές των R,G,B. Έτσι επιστρέφονται οι τιμές για τα τεχνητά κωνία με χρήση των ιδιοτήτων

```
.Cone[ctLConeRed]
.Cone[ctMConeGreen]
.Cone[ctSConeBlue]
```

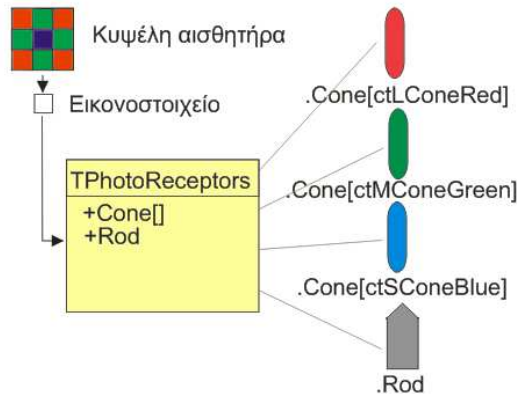
που στην μέθοδο ανάκτησης (getter method) υλοποιούν scanlines. Για να επιστραφεί η τιμή του τεχνητού ραβδίου, χρησιμοποιείται ο τύπος υπολογισμού 6.2 του φωτός (luma) στο σύστημα YIU που είναι:

Rod = 0.299 * Cone[ctLConeRed] + 0.587 * Cone[ctMConeGreen] + 0.114 * Cone[ctSConeBlue]

Η θεώρηση αυτή στηρίζεται προσπαθώντας να ανασυνθέσουμε την αρχική πληροφορία δηλαδή πόσο φως υπήρχε στην συγκεκριμένη κυψέλη του αισθητήρα CCD/CMOS πριν σπάσει σε τρία μέρη. Σύμφωνα με το φίλτρο Bayer το μοτίβο επανάληψης περιλαμβάνει 25% κόκκινες κυψέλες, 50% πράσινες κυψέλες και 25% μπλε κυψέλες, αλλά λόγω της απομωσαϊκοποίησης (demosaicing) αυτά τα ποσοστά δεν ισχύουν πλέον για ένα συγκεκριμένο pixel. Αν σκεφτούμε

την νευροεπιστήμη σύμφωνα με τους Roorda και Williams [109], η οποία λέει ότι το πλήθος των μπλε κωνίων, είναι λιγότερο από 10% (εικόνα 7-6) σε ένα κβαντισμένο κομμάτι του αμφιβληστροειδή, ο τύπος υπολογισμού 6.2 θεωρείται αποδεκτός.

Έτσι θεωρούμε ότι το pixel αντιπροσωπεύει ένα κομμάτι αμφιβληστροειδή στο οποίο υπάρχουν τεχνητά κωνία για τις τρεις χρωματικές συνιστώσες και ένα τεχνητό ραβδί το οποίο μετράει φως. Μια σχεδιαστική απόφαση ήταν να μην υπάρξουν ξεχωριστά αντικείμενα για κάθε φωτοϋποδοχέα ενός pixel, αλλά ένα που θα τα περικλείει καθώς έτσι το πλήθος τους θα είναι ανάλογο της ανάλυσης, δηλαδή εδώ $500 \times 500 = 250000$. Κάθε αντικείμενο θα δίνει τιμές σε εισόδους τεσσάρων διαφορετικών τεχνητών νευρώνων, που θα δούμε στην επόμενη ενότητα.



9-3: Η θέση και η λειτουργία της κλάσης TPhotoReceptors ©2011 Horasis

Σε μια μελλοντική υλοποίηση με ίσως πιο εξειδικευμένο υλικό ψηφιακής απεικόνισης θα μπορούσαμε να έχουμε μια πιο ακριβή προσομοίωση της βιολογικής πραγματικότητας των φωτοϋποδοχέων. Σε αυτήν το υλικό ίσως μπορούσε να επιστρέψει την μέτρηση του φωτός που πέφτει πάνω στον αισθητήρα χωρίς χρήση φίλτρου, δηλαδή την ποσότητα λευκού φωτός όπως ακριβώς κάνει και ένα βιολογικό ραβδί. Και σε αυτήν την περίπτωση, θα υπήρχε η ίδια κλάση TPhotoReceptors στο λογισμικό μας που θα επέστρεφε τιμές για τεχνητούς φωτοϋποδοχείς, απλώς θα άλλαζε η εσωτερική υλοποίηση της.

Τεχνητός Αμφιβληστροειδής

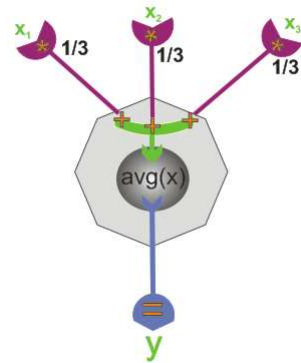
Τεχνητά οριζόντια κύτταρα και ο χάρτης υποδεκτικού πεδίου

Αφού διαβαστούν οι τιμές των τεχνητών κωνίων R,G,B και υπολογιστεί η τιμή του τεχνητού ραβδίου W, τροφοδοτούνται σε αντικείμενα τεχνητών νευρώνων που παίζουν το ρόλο των οριζοντίων κυττάρων HC του αμφιβληστροειδούς. Αυτά συνιστούν την είσοδο του υποσυστήματος INL, των διαύλων χρωματικού ανταγωνισμού. Οργανώνονται σε τέσσερις τύπους έναν για το κάθε είδος φωτοϋποδοχέα, οπότε θα έχουμε αντίστοιχα R-HC για το κόκκινο G-HC για το πράσινο, B-HC για το μπλε και W-HC για το άσπρο (φως). Η λειτουργία τους μοντελοποιεί την εμπρόσθια τροφοδότηση και την γραμμική άθροιση που κάνουν οι βιολογικοί νευρώνες HC, η οποίες αναφέρθηκαν στο κεφάλαιο 7.

Το απλοποιημένο μοντέλο θεωρεί ότι ένας τεχνητός HC νευρώνας είναι ένας απλός McCulloch-Pitts με συνάρτηση ενεργοποίησης την γραμμική $f(u) = u$ που δέχεται εισόδους μόνο από τους αντίστοιχους φωτοϋποδοχείς της ομάδας που ανήκει. Έτσι αν είναι R-HC θα λαμβάνει τιμές εισόδου 0..255 διαβάζοντας την ιδιότητα Cone[ctLConeRed] σε ένα πλήθος n αντικειμένων TPhotoReceptors. Στην απλούστατη περίπτωση τα αντίστοιχα βάρη των εισόδων ορίζονται εκ των προτέρων σε:

$$w_i = \frac{1}{n}, \quad i = 1..n$$

οπού n είναι το πλήθος των εισόδων του νευρώνα. Έτσι ο νευρώνας θα παράγει στην έξοδο του τον μέσο όρο των εισόδων του. Βλέπουμε δηλαδή ότι αντί να υλοποιηθεί ένας απλός υπολογισμός μέσου όρου, προτιμήθηκε να εκφραστεί ως ένας τεχνητός νευρώνας ώστε να μπορούν να τρέξουν μελλοντικά διάφορα πειράματα. Θα μπορούσαν να υλοποιηθούν νέοι τρόποι για τον καθορισμό αυτών των βαρών είτε στατικά βάσει κάποιου τρόπου υπολογισμού, ή δυναμικά. Επίσης θα μπορούσε να χρησιμοποιηθούν πειραματικά διαφορετικές συναρτήσεις ενεργοποίησης, με κανονικοποίηση των εισόδων στο [0..1] και προβολή των εξόδων στο [0,255].



9-4: Τεχνητός οριζόντιος νευρώνας, μέσο όρο, τριών εισόδων.



9-5: Χάρτης 13 υποδεκτικών πεδίων τεχνητών HC
©2011 Horasis

Το πλήθος εισόδων του τεχνητού HC, αντιστοιχεί σε ένα πλήθος εικονοστοιχείων κάποιας περιοχής της εικόνας, υλοποιώντας το υποδεκτικό του πεδίο. Οι συντεταγμένες των εικονοστοιχείων που ανήκουν σε κάθε περιοχή, καθορίζονται με μια πρακτική χωρική αντιστοίχιση τον **Χάρτη Υποδεκτικού Πεδίου (Receptive Field Map – RFM)**. Αυτός είναι μια ψηφιδωτή εικόνα στην οποία χρωματίζονται με συγκεκριμένο χρώμα όσα εικονοστοιχεία θα χρησιμοποιηθούν στο υποδεκτικό πεδίο ενός νευρώνα και με μαύρο όσα θα αγνοηθούν. Σε έναν τέτοιο χάρτη μπορούν να υπάρχουν πολλά χρώματα, όπου το κάθε ένα αντιστοιχεί σε ένα υποδεκτικό πεδίο διαφορετικού νευρώνα. Όπως βλέπουμε και στην διπλανή εικόνα 9-5 ταξινομούμε χωρικά κάποια εικονοστοιχεία από ένα τετραγωνικό χωρίο της εικόνας σε 13 διαφορετικά υποδεκτικά πεδία, με το κάθε ένα να

αποτελεί τις εισόδους ενός τεχνητού ΗΣ.

Στο ίδιο παράδειγμα υπάρχουν 19 pixels στην περιοχή του χάρτη που είναι βαμμένη με σκούρο κόκκινο χρώμα και 29 pixels στην περιοχή με το κίτρινο χρώμα, για ένα τετράγωνο 17 x 17 pixels. Τα 29 κίτρινα pixels στο κέντρο θα δημιουργήσουν 29 δενδρίτες εισόδου σε 4 διαφορετικούς τεχνητούς ΗΣ, και πλέον η φέτα (slice) αυτή θα χαρακτηρίζεται με τον δείκτη 0. Οι τιμές των 29 τεχνητών φωτοϋποδοχών R θα εισάγονται σε ένα R-ΗΣ της φέτας και στην έξοδο του θα παράγεται ο μέσος όρος. Τα 19 κόκκινα pixels στο κομμάτι της περιφέρειας αποτελούν μια δεύτερη φέτα που θα αριθμείται με δείκτη 1 και ακολουθώντας την φορά του ρολογιού θα αριθμήσουμε αντίστοιχα με δείκτες και τις υπόλοιπες φέτες. Για όλη την περιοχή θα έχουμε 52 τεχνητούς ΗΣ νευρώνες, 13 για κάθε έναν από τους 4 τύπους R-ΗΣ, G-ΗΣ, B-ΗΣ και W-ΗΣ.



9-6 Φέτες μιας περιοχής εικόνας
©2011 Horasis

Αν υπολογίσουμε στην κάθε μια από τις τέσσερις γωνίες του χάρτη της εικόνας 9-5, βρίσκονται 8 μαύρα pixel τότε 32 pixel δεν θα χρησιμοποιηθούν για είσοδο στους νευρώνες. Έτσι έχουμε $17 * 17 - 4 * 8 = 289 - 32 = 257$ pixel που αρχικά χρησιμοποιούνταν στην περιοχή. Στην θέση τους έχουμε 13 εξόδους, άρα έχουμε 5% του αρχικού πλήθους των πληροφοριών. Για όλη την εικόνα τα 250.000 pixel συμπιέζονται σε περίπου 13.000 φέτες, αν λάβουμε υπόψη και τις πιθανές επικαλύψεις (overlapping). Το πλήθος των διαφορετικών τιμών που θα εξαγονται από αυτό το πρώτο στρώμα νευρώνων θα είναι περίπου 52.000, που είναι μεν ακόμα μεγάλο αλλά σημαντικά μειωμένο σε σχέση με την αρχική εικόνα. Βέβαια πλέον έχουμε την πρώτη απώλεια πληροφορίας καθώς η επεξεργασία είναι ένα θόλωμα (blur) όπως στην κλασική επεξεργασία εικόνας [152]. Για να ανασυνθέσουμε την εικόνα βάλουμε κάθε φέτα, χρησιμοποιώντας τις εξόδους των αντίστοιχων R,G,B τεχνητών ΗΣ νευρώνων και βλέπουμε στην παρακάτω εικόνα την αλλαγή που έχει υποστεί σε αυτό το στάδιο.



9-7: Η εικόνα Μετά την επεξεργασία από του τεχνητούς νευρώνες ΗΣ

Μια σημαντική παρατήρηση πάνω στον χάρτη υποδεκτικού πεδίου που είδαμε είναι ότι υλοποιεί ταυτόχρονα και μια οργάνωση κέντρου περιφέρειας, που θα μας χρειαστεί παρακάτω στους τεχνητούς γαγγλιακούς νευρώνες. Η κίτρινη περιοχή αντιπροσωπεύει το κέντρο του υποδεκτικού



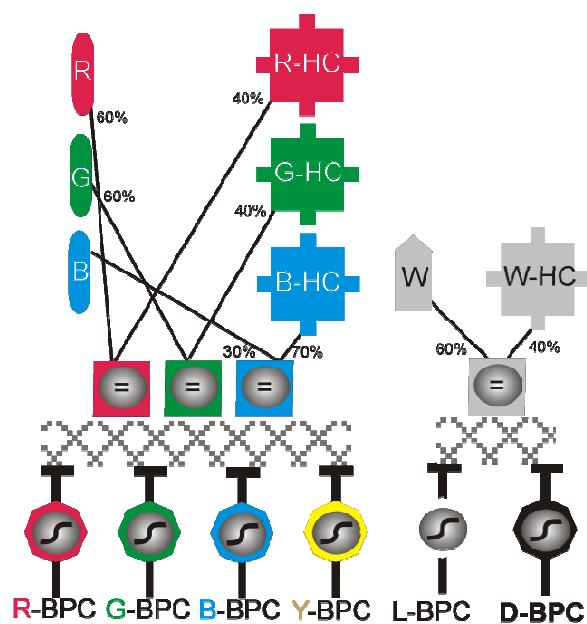
9-8 Φέτα κέντρου 0 και φέτες περιφέρειας 1-12
©2011 Horasis

πεδίου και όλα τα υπόλοιπα χρώματα την περιφέρεια του. Γι αυτό όλη η χρωματισμένη περιοχή είναι ένας κύκλος εγγεγραμμένος μέσα σε μαύρο φόντο, ο οποίος έχει διάμετρο ίση με την πλευρά του τετραγώνου. Στο κέντρο υπάρχει ακόμα ένας μικροσκοπικός κύκλος μικρότερης διαμέτρου. Επιλέχθηκε ένας μονός αριθμός, όπως το 17, για την διάμετρο ώστε να υπάρχει ένα εικονοστοιχείο ακριβώς στο κέντρο και των δύο ομόκεντρων κύκλων.

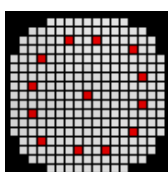
Η έναρξη των διαύλων χρωματικού ανταγωνισμού και τα τεχνητά δίπολα κύτταρα

Προχωρώντας μέσα στο υποσύστημα INL, βλέπουμε το επόμενο στρώμα αντικειμένων που αντιπροσωπεύουν τα βιολογικά δίπολα κύτταρα BPC. Πλέον όμως τα αντικείμενα αυτά δεν είναι απλοί τεχνητοί νευρώνες αλλά έχουν σιγμοειδή συνάρτηση ενεργοποίησης, αφού ανήκουν σε στρώματα εξόδου τεχνητών νευρωνικών δικτύων. Αυτά τα ΤΝΔ παίρνουν είσοδο από την έξοδο των τεχνητών HC αλλά και απευθείας από τεχνητούς φωτουπόδοχείς. Πιο αναλυτικά υπάρχουν δύο τύπων ΤΝΔ, ένα για BPC κωνίων και ένα για BPC ραβδίου.

Ένα ΤΝΔ κωνίων δέχεται τρεις εισόδους από τα αντίστοιχα βασικά χρώματα R,G,B και παράγει 4 εξόδους κόκκινο R, πράσινο G, μπλε B, κίτρινο Y υλοποιώντας τους δύο πρώτους διαύλους του χρωματικού ανταγωνισμού, δηλαδή **κόκκινο/πράσινο** και **μπλε/κίτρινο**. Από την άλλη το ΤΝΔ ραβδίων δέχεται μία είσοδο W και παράγει 2 εξόδους λευκό (φώς) L και μαύρο (σκοτάδι) D συμπληρώνοντας και την τρίτη γραμμή **λευκό/μαύρο**. Οι νευρώνες εξόδου αυτών των δικτύων είναι τα τεχνητά BPC και με αυτά ξεκινάν οι διάυλοι χρωματικού ανταγωνισμού ,που λειτουργούν παράλληλα για κάθε περιοχή της εικόνας και επεκτείνονται μέχρι και την δημιουργία της εσωτερικής αναπαράστασης της οπτικής πληροφορίας. Σε κάθε τμήμα της εικόνας που περιέχει 13 φέτες, θα υπάρχουν για ανά φέτα 6 τεχνητοί BPC νευρώνες αντίστοιχοι για τα 6 ανταγωνιστικά χρώματα R-BPC, G-PC, B-BPC,Y-BPC,L-BPC,D-BPC. Έτσι για ένα τμήμα θα υπάρχουν 78 BPC νευρώνες με τους 6 να δίνουν τις τιμές του κέντρου και τους υπόλοιπους 72 να δίνουν τις τιμές της περιφέρειας.



9-9: Τα ΤΝΔ κωνίων (αριστερά) και ραβδίου (δεξιά) με τα 6 τεχνητά δίπολα κύτταρα στο στρώμα εξόδου και των δύο. ©2011 Horasis



9-10: Κεντρικά κωνία
©2011 Horasis

Οι εισοδοί των δύο διαφορετικών τύπων ΤΝΔ δέχονται τιμές από αντίστοιχους τεχνητούς HC νευρώνες του προηγούμενου επιπέδου, αλλά και κατευθείαν από έναν τεχνητό φωτουπόδοχέα. Αυτό βρίσκεται σε αντιστοιχία με την βιολογική πραγματικότητα καθώς ένα δίπολο κύτταρο δέχεται είσοδο απευθείας από έναν ή και περισσότερους φωτουπόδοχείς, αλλά δημιουργεί και συνάψεις με οριζόντια κύτταρα. Το ποσοστό των δύο τιμών ορίζεται από την σχέση:

$$X_R = (1 - \alpha)Y_{R-HC} + \alpha R \quad (9.1)$$

όπου X_R η είσοδος που αντιστοιχεί στο κόκκινο σε ένα ΤΝΔ κωνίων, Y_{R-HC} ή έξοδος ενός τεχνητού R-HC που διαβάζει τεχνητά κωνία τύπου LConeRed και R η τιμή του τεχνητού κωνίου τύπου LConeRed που βρίσκεται στο κέντρο του υποδεκτικού πεδίου του R-HC. Το α είναι το ποσοστό συμμετοχής του κεντρικού κωνίου, το οποίο ορίζεται σε αυτήν την πρώτη υλοποίηση του Horasis σε 0.6 για τα R,G,W και σε 0.3 για τα B. Στην ουσία η είσοδος αποτελείται από η τιμή του κεντρικού κωνίου R κατά 60% και τον μέσο όρο μιας περιοχής γύρω από αυτό κατά 40%.

Η εκπαίδευση του ΤΝΔ που περιέχει τεχνητούς δίπολους νευρώνες των τριών κωνίων ανάλογα με την απόχρωση.

Οι δύο τύποι τεχνητών νευρωνικών δικτύων όπως είπαμε έχουν είσοδο τιμές 0..255 που δίνουν οι τεχνητοί φωτοϋποδοχείς. Και τα δύο δίκτυα εκπαιδεύονται ώστε στην έξοδο τους να παράγουν θετικές τιμές για διέγερση αλλά και αρνητικές τιμές για αναστολή.

Το ΤΝΔ για BPC των κωνίων εκπαιδεύεται ώστε ανάλογα με το χρώμα που αντιπροσωπεύουν οι τρεις χρωματικές συνιστώσες R,G,B της εισόδου του να παράγει παράλληλα τέσσερις διαφορετικές αποκρίσεις. Η απόκριση δεν αντιπροσωπεύει πλέον την ποσότητα για παράδειγμα του κόκκινου μέσα στο χρώμα όπως η χρωματική συνιστώσα R στην είσοδο. Πλέον ο τεχνητός R-BPC νευρώνας θα επιστρέψει το πόσο κοντά ή μακριά από αμιγές κόκκινο βρίσκεται το χρώμα που αντιπροσωπεύει το διάνυσμα {R,G,B} της εισόδου. Έχουμε ορίσει στο κεφάλαιο 6 τα αμιγή χρώματα και έχουμε υπολογίσει για τα 4 ανταγωνιστικά που μας ενδιαφέρουν τις τιμές της απόχρωσης H στο χρωματικό σύστημα HSL, σε κλίμακα 0..255. Έτσι έχουμε με την σειρά που εμφανίζονται στον κύκλο των χρωμάτων

- Κόκκινο H = 0 ή H = 255, καθώς ανοίγει και κλείνει τον κύκλο των χρωμάτων
- Κίτρινο H = 43
- Πράσινο H = 85
- Μπλε H = 169

Θα δημιουργήσουμε τις συναρτήσεις $f_{R-BPC}(H)$, $f_{G-BPC}(H)$, $f_{B-BPC}(H)$, $f_{Y-BPC}(H)$ οι οποίες βοηθούν στην παραγωγή των στόχων εκπαίδευσης του ΤΝΔ των κωνίων, οι οποίοι είναι συνάρτηση της απόχρωσης H. Ειδικότερα για την πρώτη θα την ορίσουμε ως σύνθεση δύο συναρτήσεων $f_{R-BPC}(H) = f_{l_{R-BPC}}(H) + f_{h_{R-BPC}}(H)$.

Οι συναρτήσεις έχουν ως βάση την συνάρτησης κανονικής κατανομής ή αλλιώς κατανομής gauss [153] η οποία μπορεί να εκφράσει το πόσο απέχει μια απόχρωση από την απόχρωση του αμιγούς χρώματος.

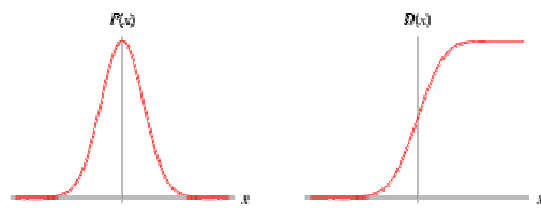
$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (9.2)$$

που θα την γράψουμε πλέον ως

$$f_{nd}_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (9.3)$$

όπου σ είναι η απόκλιση (standard deviation) και μ ο μέσος (mean).

Με αυτόν τον τρόπο δημιουργούμε αρχικά συναρτήσεις κανονικής κατανομής με το μ να ισούται με την απόχρωση του αμιγούς χρώματος και ένα σ που στενεύει ή φαρδαίνει την καμπάνα της κατανομής ανάλογα με την περίπτωση έτσι έχουμε:



9-11: Γραφική παράσταση κανονικής κατανομής και αθροιστικής κανονικής κατανομής (cumulative distribution function)

- Για το κόκκινο χρώμα:

$$\begin{aligned} g_{R-BPC}(H) &= gl_{1R-BPC}(H) + gh_{R-BPC}(H) \\ &= fnd_{0,22}(x) / fnd_{0,22}(0) + fnd_{255,22}(x) / fnd_{255,22}(255) \end{aligned} \quad (9.4)$$

από αντικατάσταση των επιμέρους:

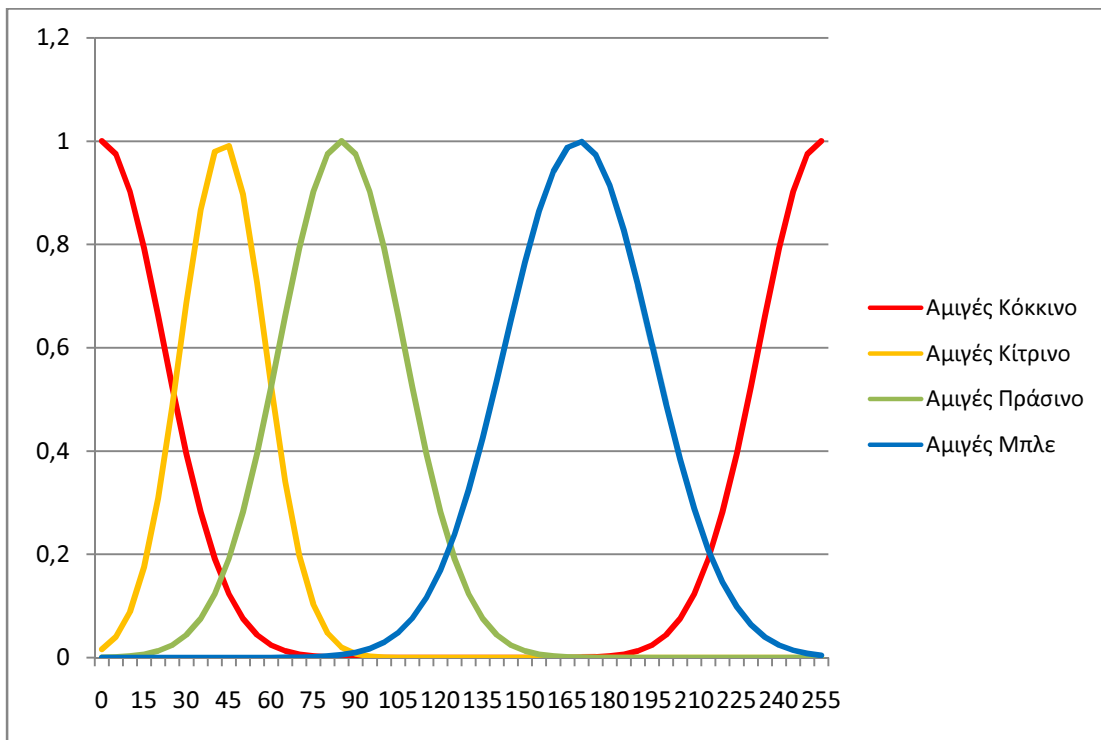
- $gl_{1R-BPC}(H) = fnd_{0,22}(x) / fnd_{0,22}(0)$
- $gh_{R-BPC}(H) = fnd_{255,22}(x) / fnd_{255,22}(255)$

- Για το πράσινο χρώμα: $g_{G-BPC}(H) = fnd_{85,22}(x) / fnd_{85,22}(85)$ (9.5)

- Για το κίτρινο χρώμα: $g_{Y-BPC}(H) = fnd_{43,15}(x) / fnd_{43,15}(43)$ (9.6)

- Για το μπλε χρώμα: $g_{B-BPC}(H) = fnd_{169,26}(x) / fnd_{169,26}(169)$ (9.7)

Έτσι το πεδίο τιμών για όλες τις συναρτήσεις είναι το $[0,1] \in \mathbb{R}$. Βλέπουμε παρακάτω τις τέσσερις γραφικές παραστάσεις για τα αντίστοιχα χρώματα.



9-12: Γραφική παράσταση των συναρτήσεων απόστασης από αμιγές χρώμα ©2011 Horasis

Στα μέγιστα των συναρτήσεων βρίσκεται η απόχρωση του αντίστοιχου αμιγούς χρώματος. Το κόκκινο όπως βλέπουμε κόβεται στα δύο στην αρχή του κύκλου των χρωμάτων μειούμενο και στο τέλος αυξανόμενο. Βλέπουμε για το κόκκινο και το πράσινο που είναι ανταγωνιστικά μεταξύ τους ότι όσο ισαπέχει μια απόχρωση και από τα δύο, τόσο μεγιστοποιείται το κίτρινο. Επίσης οι αποχρώσεις του κίτρινου φθίνουν μέχρι να συναντήσουν τις πρώτες αδύναμες αποχρώσεις του μπλε, έχοντας ένα σημείο σχεδόν πάνω στην απόχρωση του αμιγούς πράσινου στο οποίο οι δύο αποκρίσεις ισούνται.

Αν και το παραπάνω μοντέλο εκφράζει τον χρωματικό ανταγωνισμό οι τιμές που μας δίνει αφορούν μόνο θετική διέγερση καθώς σε μια απόχρωση μπορούμε να έχουμε διεγέρσεις δύο χρωμάτων. Για να αποφευχθεί αυτό και παράλληλα να εισαχθεί η λειτουργία της αναστολής σε

κάθε τεχνητό νευρώνα όταν εισάγεται μια ανταγωνιστική ως προς την λειτουργία του απόχρωση, ορίστηκαν οι τελικές συναρτήσεις χρωματικού ανταγωνισμού για τους στόχους ως εξής:

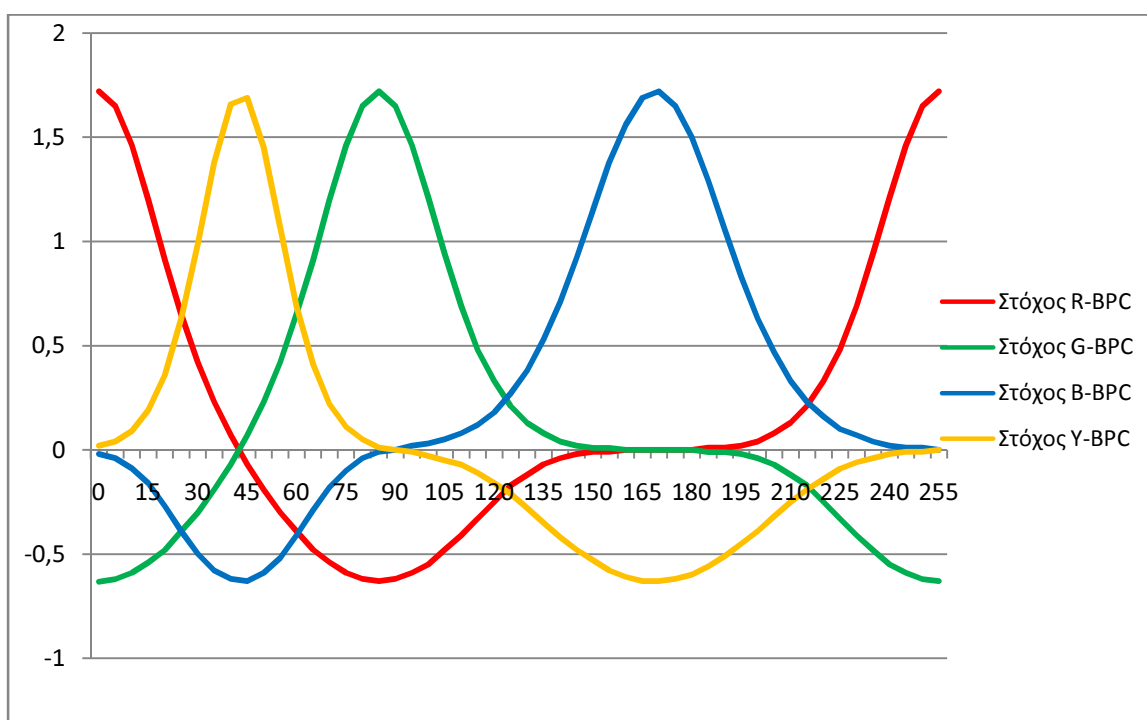
- Για το κόκκινο χρώμα: $f_{R-BPC}(H) = \frac{g_{R-BPC}(H)}{g_{G-BPC}(H)}$ (9.8)

- Για το πράσινο χρώμα: $f_{G-BPC}(H) = \frac{g_{G-BPC}(H)}{g_{R-BPC}(H)}$ (9.9)

- Για το κίτρινο χρώμα: $f_{B-BPC}(H) = \frac{g_{B-BPC}(H)}{g_{Y-BPC}(H)}$ (9.10)

- Για το μπλε χρώμα: $f_{Y-BPC}(H) = \frac{g_{Y-BPC}(H)}{g_{B-BPC}(H)}$ (9.11)

Έτσι το πεδίο τιμών για όλες τις συναρτήσεις είναι το $[-0.63, 1.72] \in \mathbb{R}$. Βλέπουμε παρακάτω τις τέσσερις γραφικές παραστάσεις για τα αντίστοιχα χρώματα.



9-13: Γραφική παράσταση των συναρτήσεων παραγωγής στόχων για ΤΝΔ BPC των κωνίων ©2011 Horasis

Πλέον έχουμε ένα μοντέλο τιμών που εκφράσει ακριβώς τον χρωματικό ανταγωνισμό. Ας πάρουμε το παράδειγμα ενός τεχνητού δίπολου νευρώνα του G-BPC. Θα δούμε ότι όταν η απόχρωση του χρώματος, που εκφράζουν οι τιμές {R,G,B} στην είσοδο του ΤΝΔ που τον περιέχει, είναι αμιγές κόκκινο τότε το G-BPC θα αναστέλλεται πλήρως με τιμή -0.63. Από την άλλη στο αμιγές πράσινο θα διεγείρεται πλήρως με τιμή 1.72. Το ίδιο γίνεται για όλα τα BPC του ΤΝΔ των κωνίων. Πλέον μπορούμε να κανονικοποιήσουμε τις εισόδους στο $[0..1]$ διαιρώντας με το 255 και να προβάσουμε τους στόχους στο $[0..1]$ με τον τύπο $t = \frac{f(H)+0.63}{2.35}$ (9.12) φτιάχνοντας ένα πρότυπο εκπαίδευσης. Το εκπαιδευμένο δίκτυο καταφέρνει να πετύχει προσέγγιση και των τεσσάρων συναρτήσεων (function approximation) δείχνοντας ότι η θεώρηση των αμιγών χρωμάτων αλλά και η επιλογή των διαφορετικών σ ανά χρώμα είναι πετυχημένη. Επίσης η εκπαίδευση κατάφερε να προσπεράσει τα τοπικά ελάχιστα, παρόλο που οι στόχοι δεν ήταν δυαδικοί αλλά δεκαδικοί μεταξύ 0 και 1.

Η εκπαίδευση του ΤΝΔ που περιέχει τεχνητούς δίπολους νευρώνες ενός ραβδίου ανάλογα με το φως.

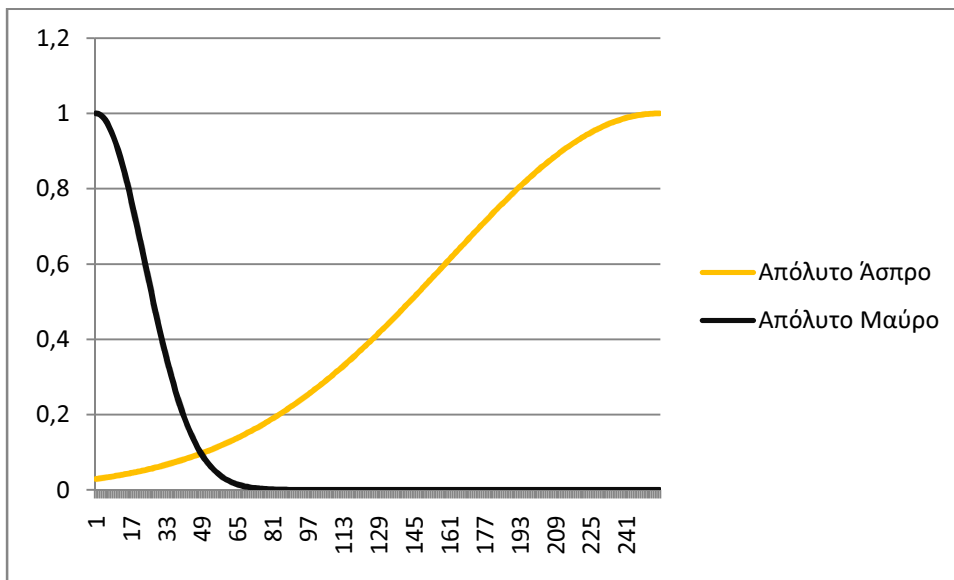
Το ΤΝΔ για BPC ραβδίου είναι όπως είπαμε δίκτυα με μια τιμή εισόδου και δύο τιμές εξόδου. Εδώ οι αποκρίσεις είναι πάλι θετικές ή αρνητικές αλλά πλέον δεν έχουμε την έννοια της απόχρωσης καθώς το μόνο που έχουμε ως είσοδο είναι μια τιμή από το 0 έως το 255 για το φως που παράγεται από τα {R,G,B} του ίδιου εικονοστοιχείου.

Παρόλαταυτα η αντιμετώπιση και εδώ είναι ακριβώς η ίδια χρησιμοποιώντας δύο συναρτήσεις για τους στόχους, που και αυτές περιλαμβάνουν κανονικές κατανομές. Ξεκινώντας πρώτα από τις συναρτήσεις που δείχνουν την απόσταση μια τιμής φωτός από το απόλυτο μαύρο ή το απόλυτο λευκό έχουμε:

- Για το μαύρο χρώμα: $g_{D-BPC}(H) = fnd_{0,22}(x) / fnd_{0,22}(0)$ (9.13)

- Για το λευκό χρώμα: $g_{L-BPC}(H) = fnd_{255,96}(x) / fnd_{255,96}(255)$ (9.14)

Το πεδίο τιμών για τις παραπάνω είναι πάλι το $[0,1] \in \mathbb{R}$.



9-14: Γραφική παράσταση των συναρτήσεων απόστασης από απόλυτο μαύρο και απόλυτο λευκό ©2011 Horasis

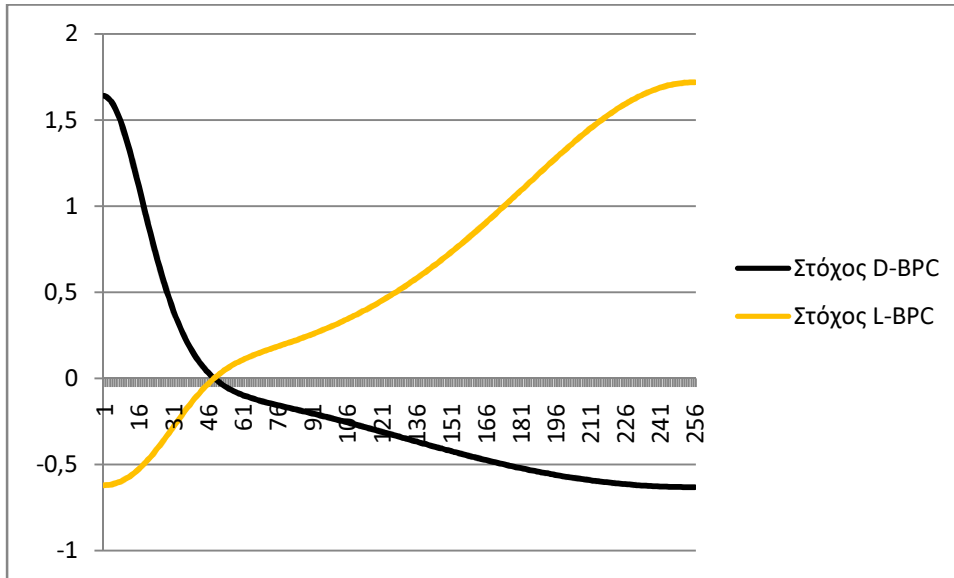
Με αυτόν τον τρόπο χωρίζουμε όλες τις αποχρώσεις του γκρι σε αυτές που είναι πιο κοντά στο μαύρο και στις υπόλοιπες που το λευκό υπερσχύει του μαύρου. Αν και υπάρχουν συγκεκριμένες αποχρώσεις για το μαύρο μετά τις οποίες η συνάρτηση 9.13 επιστρέφει 0, για το λευκό έχουμε επιστροφή τιμής από την 9.14 σε κάθε περίπτωση, που μετατρέπει την γραμμική συμπεριφορά των τιμών στην κλίμακα του γκρι σε μια καμπύλη.

Οι αντίστοιχες συναρτήσεις χρωματικού ανταγωνισμού για τους στόχους των τεχνητών BPC στην έξοδο του ΤΝΔ είναι:

- Για το μαύρο χρώμα: $f_{D-BPC}(H) = \frac{g_{D-BPC}(H)}{g_{L-BPC}(H)}$ (9.15)

- Για το λευκό χρώμα: $f_{L-BPC}(H) = \frac{g_{L-BPC}(H)}{g_{D-BPC}(H)}$ (9.16)

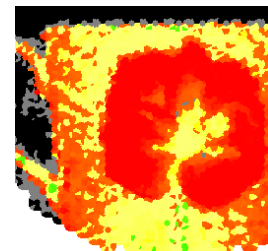
με το πεδίο τιμών να είναι και σε αυτήν την περίπτωση το $[-0.63, 1.72] \in \mathbb{R}$.



9-15 : Γραφική παράσταση των συναρτήσεων παραγωγής στόχων για ΤΝΔ BPC ενός ραβδίου ©2011 Horasis

Εδώ εμφανίζεται μια έκφανση του διαύλου χρωματικού ανταγωνισμού λευκό/μαύρο, μέσα από μια μόνο τιμή που υπήρχε πριν την θεώρηση αυτή η οποία αντιπροσώπευε μόνο το λευκό. Θεωρούμε ότι υπάρχει περισσότερο μαύρο χρώμα από την τιμή γκρι 0 μέχρι και την τιμή 45 οπότε η συνάρτηση που αντιπροσωπεύει τον στόχο του λευκού L-BPC τεχνητού νευρώνα, αρχίζει να επιστρέφει θετικές τιμές. Δημιουργήθηκαν πρότυπα εκπαίδευσης έχοντας όλες τις τιμές εισόδου και στόχων στο $[0..1]$. Η εκπαίδευση αυτού του περιέργου δικτύου πέτυχε μια πολύ μικρή τιμή μέσου τετραγωνικού σφάλματος στο 0,0053, καταφέροντας να παράγει από μια τιμή εισόδου δύο διαφορετικές τιμές εξόδου. Και εδώ δεν υπήρξαν προβλήματα με τοπικά ελάχιστα και η προσέγγιση των δύο συναρτήσεων λειτούργησε κανονικά.

Πλέον με τα δύο ΤΝΔ κωνίων και ραβδίου εκπαιδευμένα η έξοδος του παράγει 6 διαφορετικά κανάλια χρωμάτων για την αρχική εικόνα, τα οποία μπορούμε να τα παραστήσουμε στην οθόνη χρωματίζοντας πάλι τις φέτες μιας περιοχής. Επίσης έχουν δημιουργηθεί δύο χρώματα τα οποία δεν υπάρχουν στους τεχνητούς φωτοϋποδοχείς και πολύ περισσότερο στις χρωματικές συνιστώσες R,G,B των εικονοστοιχείων. Οι συνιστώσες αυτές που δεν προσδιόριζαν από μόνες τους το χρώμα, πλέον έχουν αντικατασταθεί από τις εξόδους των R-BPC, G-BPC και B-BPC οι οποίες έχουν η κάθε μια ευαισθησία ανάλογη με το πόσο κοντά είναι μια απόχρωση στο χρώμα των τεχνητών νευρώνων.



9-16 Εμφάνιση καναλιών πράσινου και κόκκινου. Με κίτρινο και γκρι φαίνεται η απροσδιόριστα, ενώ με μαύρο χαμηλή φωτεινότητα

Γαγγλιακά Κύτταρα

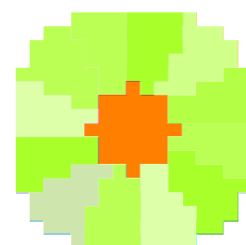
Προχωρώντας στο τέλος της υλοποίησης του τμήματος λογισμικού του αντικειμενοστραφούς τεχνητού αμφιβληστροειδή ή ΟΟΑΡ, μετά το υποσύστημα INL υπάρχουν τα τελευταία υποσυστήματα AGC των τεχνητών γαγγλιακών κυττάρων ή τεχνητών GC, του τεχνητού ρετινοτοπικού χάρτη ARTN και μια απλή θεώρηση για την τοποθέτηση του κεντρικού βοθρίου ως υποσύστημα FCP. Στο πρώτο από αυτά υπάρχουν αντίστοιχα αντικείμενα που προσομοιώνουν την λειτουργία του βιολογικού GC. Μόνο που σε αυτήν την περίπτωση το αντικείμενο είναι ένα ολόκληρο τεχνητό νευρωνικό δίκτυο. Στο στρώμα εισόδου υπάρχουν 13 νευρώνες που λαμβάνουν τις τιμές της εισόδου τους, από τις εξόδους των τεχνητών BPC. Στο στρώμα εξόδου του ΤΝΔ-GC υπάρχουν 4 νευρώνες και ο συνδυασμός των τιμών τους παράγει ένα διάνυσμα το οποίο αντιπροσωπεύει την κωδικοποίηση γαγγλιακού κυττάρου GCE.

Αυτό το διάνυσμα GCE κωδικοποιεί την σύνθετη απόκριση ενός τεχνητού GC σε τέσσερις διαστάσεις, αποτελούμενο από τέσσερις τιμές αντίστοιχα:

1. διέγερση
2. αναστολή
3. διάχυτο φως
4. ώρα μετάβασης προς το φώς

Αυτό το διάνυσμα προσπαθεί να μοντελοποιήσει την σύνθετη απόκριση που υπάρχει στην κυματομορφή του νευρωνικού σήματος στην έξοδο ενός βιολογικού GC, η οποία εξαρτάται από την θέση που πέφτει το φως πάνω στο υποδεκτικό του πεδίο. Στο Horasis κρίθηκε σκόπιμο να υπάρχουν δύο διαστάσεις μια για την διέγερση και μια για την αναστολή στην θέση των θετικών και αρνητικών τιμών που αναφέρονται σε μια διάσταση. Αυτό οφείλεται στο γεγονός ότι το τεχνητό GC κύτταρο μοντελοποιεί την συμπεριφορά του βιολογικού GC κυττάρου, το οποίο μπορεί να έχει ταυτόχρονα και διέγερση και αναστολή στο σύνθετο σήμα που παράγει, έχοντας κυκλικό υποδεκτικό πεδίο με οργάνωση κέντρου και περιφέρειας. Επίσης προστέθηκε μια επιπλέον τρίτη διάσταση όταν υπάρχει διάχυτο φως που καλύπτει κέντρο και περιφέρεια, δηλαδή όταν ένα ερέθισμα διεγείρει τον τεχνητό νευρώνα πλήρως στο κέντρο του και τον αναστέλλει πλήρως στην περιφέρεια του. Σε αυτήν την περίπτωση υπάρχει μια απροσδιοριστία ως προς τις δύο πρώτες διαστάσεις γι αυτό θεωρήθηκε ότι πρέπει να υπάρξει μια ακόμα τιμή που αυξάνεται όταν δεν μπορεί να γίνει σαφής διαχωρισμός μεταξύ διέγερσης και αναστολής. Η τέταρτη διάσταση που θα δούμε, προστέθηκε επιπλέον για να χαρακτηρίσει την κατεύθυνση από την οποία έρχεται ο φωτισμός στο κυκλικό υποδεκτικό πεδίο.

Όλες οι διαστάσεις υλοποιούνται μέσω του χάρτη υποδεκτικού πεδίου που είδαμε να χρησιμοποιείται αρχικά για να ορίσει τα υποδεκτικά πεδία των τεχνητών νευρώνων HC. Η τιμή για το κέντρο ενός ΤΝΔ-GC έρχεται από τον τεχνητό BPC νευρώνα του βασικού χρώματος που διεγείρει το κέντρο, και που αντιστοιχεί στην κεντρική φέτα του χάρτη με δείκτη 0. Όλα τα τεχνητά BPC του ανταγωνιστικού χρώματος στις υπόλοιπες φέτες από 1 έως 12 συνδέονται σε εισόδους που αποτελούν την περιφέρεια του τεχνητού GC. Έτσι για παράδειγμα για την υλοποίηση ενός ΤΝΔ R-GC, δηλαδή αναμμένου-κόκκινου (red on), έχουμε την πρώτη είσοδο του από το R-BPC της κεντρικής φέτας και τις υπόλοιπες από τα G-BPC στις περιφερειακές φέτες που υπάρχουν στον



9-17: Κόκκινο κέντρο και πράσινη περιφέρεια σε έναν χάρτη κυκλικού υποδεκτικού πεδίου τεχνητού GC

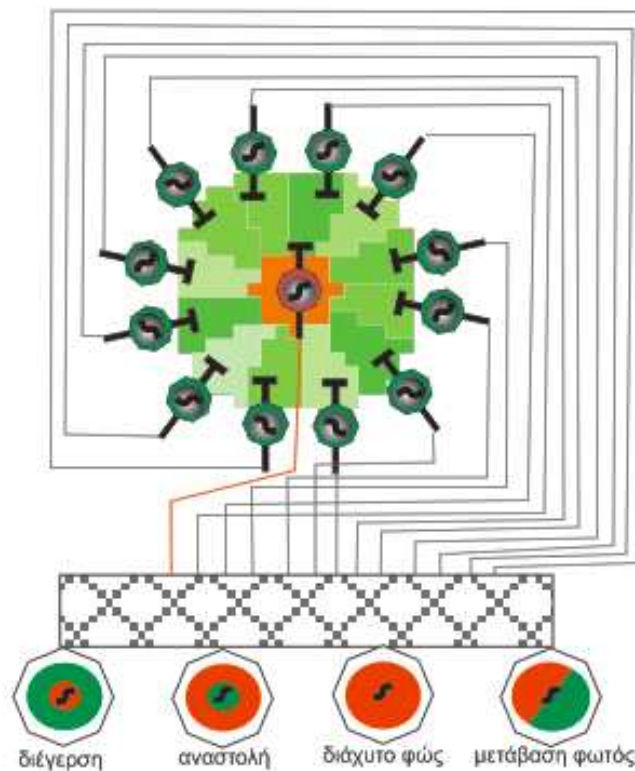
©2011 Horasis

χάρτη. Ως τελικό αποτέλεσμα έχουμε την απεικόνιση των τεχνητών φωτοϋποδοχέων της περιοχής της εικόνας στο υποδεκτικό πεδίο των τεχνητών γαγγλιακών κυττάρων.

Τα ΤΝΔ-GC που δημιουργούνται για μία περιοχή της εικόνας, έχοντας 4 εξόδους είναι 6:

- R-GC Αναμμένου κόκκινου κέντρου (Red-On)
- G-GC Αναμμένου πράσινου κέντρου (Green-On)
- B-GC Αναμμένου μπλε κέντρου (Blue-On)
- Y-GC Αναμμένου κίτρινου κέντρου (Yellow-On)
- L-GC Αναμμένου φωτισμού κέντρο ή φωτεινού κέντρου (Light-On)
- D-GC Σβηστού φωτισμού κέντρου ή σκοτεινού κέντρου(Dark-On)

Βλέπουμε παρακάτω την συνδεσμολογία για το παράδειγμα ενός R-GC σε σχέση με την περιοχή της εικόνας.

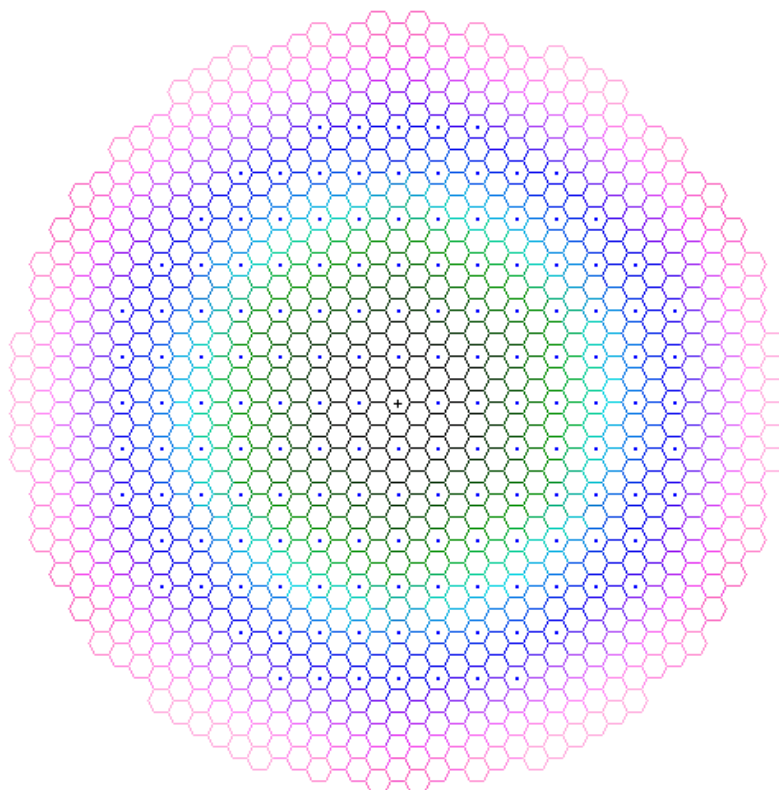


9-18: Το ΤΝΔ τεχνητού γαγγλίου και οι συνάψεις του με τα τεχνητά δίπολα ©2011 Horasis

Υπολογίζουμε για κάθε χωρίο των 17x17 pixels με κυκλικό υποδεκτικό πεδίο διαμέτρου 17pixels $6 * 4 = 24$ τιμές που την χαρακτηρίζουν, σε αντιδιαστολή με τις $13 * 6 = 78$ τιμές που υπήρχαν για τις 13 φέτες έχοντας 6 τεχνητά BPC στην κάθε μια. Το πιο σημαντικό όμως δεν είναι η μείωση του όγκου των δεδομένων στο 35% σε σχέση με αυτά που υπήρχαν στους διαύλους χρωματικού ανταγωνισμού, ούτε η μείωση στο 1,8% σε σχέση με τα εικονοστοιχεία της εικόνας 500x500 pixels. Με το υποσύστημα αυτό προστίθεται η πληροφορία αντίθεσης μεταξύ των 3 ζευγών ανταγωνιστικών χρωμάτων και παράλληλα ένα είδος εντοπισμού ακμών που λειτουργεί σε 6 διαφορετικά χρωματικά κανάλια.

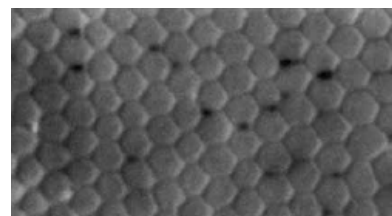
Τεχνητός ρετινοτοπικός χάρτης και η προκαθορισμένη τοποθέτηση του κέντρου του.

Έχουμε δει ακριβώς πως χρησιμοποιείται ένα τμήμα της εικόνας ως υποδεκτικό πεδίο των τεχνητών GC. Ο χωρισμός της σε πολλά κομμάτια γίνεται με την βοήθεια του υποσυστήματος ARTN που, όπως αναφέραμε και στο κεφάλαιο 8, δημιουργεί ένα σχηματισμό κυψέλης μέλισσας, δηλαδή μια εξαγωνική ψηφίδωση (tessellation) [112] που μοντελοποιεί τον βιολογικό ρετινοτοπικό χάρτη [136]. Με την ψηφίδωση αυτή καλύπτεται χωρίς κενά μια περιοχή της εικόνας και όπως θα δούμε το κάθε κανονικό εξάγωνο αντιστοιχεί σε έναν χάρτη υποδεκτικού πεδίου.



9-19: Τεχνητός ρετινοτοπικός χάρτης με εξαγωνική ψηφίδωση. ©2011 Horasis

Η επιλογή αυτού του τύπου της ψηφίδωσης δεν είναι τυχαία και σίγουρα δημιούργησε επιπρόσθετο φόρτο στην υλοποίηση του συστήματος σε σχέση με την απλή τετραγωνική ψηφίδωση που έχει ήδη ένα bitmap. Στηρίζεται στην βιολογική πραγματικότητα που υπάρχει για τους φωτούποδοχείς, προσπαθώντας να την μοντελοποιήσει σε κάποιο μεταγενέστερο επίπεδο, αφού δεν είναι δυνατό να παραχθεί ως πρωτογενή πληροφορία εικονοστοιχείων με χρήση συνηθισμένων αισθητήρων CCD/CMOS που γνωρίσαμε στο κεφάλαιο 6. Η χρήση διαφορετικού εξειδικευμένου υλικού, είναι μια μελλοντική προοπτική που δεν έχει διερευνηθεί μέσω της εργασίας. Ίσως υπάρχουν αισθητήρες σε θεωρητικό ή πειραματικό στάδιο που μπορούν να ψηφιοποιήσουν την εικόνα σε εξαγωνικό κάρναβο. Περιγράφοντας τα χαρακτηριστικά που ρυθμίζουν την παραγωγή της ψηφίδωσης μέσα στο υποσύστημα ARTN ξεκινούμε από το κεντρικό εξάγωνο ή την κεντρική κυψέλη. Το κέντρο του



9-20: Βιολογική εξαγωνική ψηφίδωση κωνίων στον ανθρώπινο αμφιβληστροειδή (Curcio et. a) [113]

εξαγώνου, έχει συντεταγμένες που υπολογίζονται απλά από το βαρύκεντρο του τετραγώνου της ψηφιακής εικόνας 500x500 pixels, δηλαδή {250,250}. Στην εικόνα 9-19 υπάρχει ένας σταυρός ακριβώς πάνω σε αυτές τις συντεταγμένες, που στιγματίζει την κεντρική κυψέλη. Το σημείο αυτό αντιπροσωπεύει το κέντρο της τεχνητής όρασης του Horasis, έχοντας στην ουσία μοντελοποίηση του κεντρικού βοθρίου (fovea centralis) που γνωρίσαμε στο κεφάλαιο 7. Εδώ έρχεται και η θεώρηση ενός ιδεατού υποσυστήματος τοποθέτησης του κεντρικού βοθρίου, του υποσυστήματος FCP. Θεωρείται ως ένα αυτόνομο υποσύστημα, ώστε σε μελλοντικές υλοποιήσεις οι συντεταγμένες του κεντρικού εξαγώνου να υπολογίζονται με διαφορετικούς τρόπους. Έτσι μπορεί μια μελλοντική υλοποίηση του Horasis να συνυπολογίζει τις μαύρες περιοχές που υπάρχουν στον φόντο της εικόνας ώστε να κεντράρεται όλη η εξαγωνική κυψέλη ακριβώς πάνω στο αντικείμενο, σε οποιαδήποτε περίπτωση.

Το δεύτερο χαρακτηριστικό είναι το πλήθος των επιπέδων στα οποία αναπτύσσεται η εξαγωνική ψηφίδωση από το κέντρο προς τα άκρα της τεχνητής όρασης. Έτσι στην εικόνα 9-19 έχουμε 25 επίπεδα να αναπτύσσονται στην εξαγωνική ψηφίδωση με τις κυψέλες που ανήκουν στο ίδιο επίπεδο να είναι χρωματισμένες με το ίδιο χρώμα. Κάθε επίπεδο έχει διαφορετικό χρώμα έτσι αναγνωρίζονται οπτικά τα διαφορετικά επίπεδα.

Το τρίτο χαρακτηριστικό είναι η διάμετρος του κύκλου στον οποίο εγγράφεται το κανονικό εξαγώνο της κυψέλης. Δηλαδή είναι η απόσταση δύο ακμών του εξαγώνου που είναι τοποθετημένες εκατέρωθεν του κέντρου, σχηματίζοντας ένα ευθύγραμμο τμήμα που περνάει από αυτό. Με αυτό το χαρακτηριστικό ορίζουμε το μέγεθος της κυψέλης.

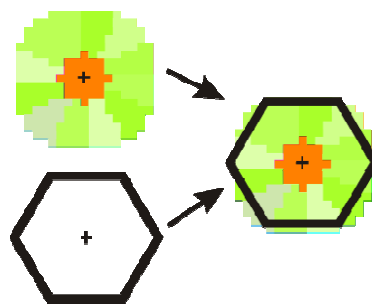
Το τέταρτο και ίσως πιο ενδιαφέρον χαρακτηριστικό είναι η μέγιστη εκκεντρότητα που επιτρέπεται να έχει μια κυψέλη που ανήκει στην ψηφίδωση. Αν αυτή υπερβεί το προκαθορισμένο όριο η κυψέλη απορρίπτεται έχοντας ως αποτέλεσμα την δημιουργία ενός κατά προσέγγιση κυκλικού δίσκου από κυψέλες, που είναι εμφανής στην εικόνα 9-19. Στην ψηφίδωση που απεικονίζεται έχει επιλεχθεί ως μέγιστη απόσταση του κέντρου μιας κυψέλης από το κέντρο της όρασης, δηλαδή εκκεντρότητα, τα 230 pixels. Έτσι δεν καλύπτεται ολόκληρο το τετράγωνο της εικόνας αλλά μόνο μια κυκλική περιοχή. Οι λόγοι που έγινε αυτό είναι δύο:

- Αφενός η πληρέστερη προσομοίωση του βιολογικού αμφιβληστροειδούς που είναι το ανάπτυγμα μιας σφαιρικής επιφάνειας.
- Αφετέρου η περαιτέρω μείωση του όγκου των δεδομένων προς επεξεργασία.

Στην τελική υλοποίηση ο τεχνητός ρετινοτοπικός χάρτης με 25 επίπεδα, διάμετρο κυψέλης 16 pixel και μέγιστη εκκεντρότητα 230pixels, αποτελείται από 1017 κυψέλες. Για την οπτική πληροφορία που περιέχεται σε μια κυψέλη είδαμε να υπάρχουν 24 τιμές, 6 διανύσματα των 4 πραγματικών τιμών. Έτσι αν η αρχική εικόνα με ανάλυση 500 x 500 pixels και 3 χρώματα/pixels είχε 750.000 τιμές, η έξοδος του τεχνητού αμφιβληστροειδούς είναι μια πληροφορία με $1017 * 24 = 24408$ τιμές δηλαδή είναι μόνο το 3,25% του αρχικού όγκου. Η πληροφορία αυτή είναι διατεταγμένη ανά κυψέλες και αποτελεί την νευρωνική διαμόρφωση αμφιβληστροειδούς RNF (Retinal Neural Format). Μετά την παραγωγή της όπως είδαμε αποθηκεύεται στην βάση δεδομένων έχοντας μια συσχέτιση με την αρχική φωτογραφία. Μέσα στην RNF τηρείται η χωροταξία των κυψελών που υπάρχει στην εξαγωνική ψηφίδωση οπότε πρώτα έχουμε την πρώτη κυψέλη με τους 24 πραγματικούς αριθμούς διπλής ακριβείας μεγέθους 192 bytes. Ακολουθούν πακέτα των 192bytes για τις κυψέλες του πρώτου επιπέδου, κατόπιν του δεύτερου και ούτω καθεξής.

Χρήση του ρετινοτοπικού χάρτη.

Όπως είπαμε στην αρχή της προηγούμενης ενότητας κάθε κυψέλη του τεχνητού αμφιβληστροειδούς χρησιμοποιείται για την εφαρμογή ενός χάρτη υποδεκτικού πεδίου. Το κέντρο του κανονικού εξαγώνου της κυψέλης ταυτίζεται με το κέντρο του κυκλικού υποδεκτικού πεδίου του χάρτη. Η πρώτη χρήση του παρατηρείται, όταν δημιουργούνται τα υποδεκτικά πεδία των τεχνητών ΗC, συγκεντρώνοντας τα εικονοστοιχεία της κάθε φέτας. Κατόπιν για κάθε μια από τις 13 φέτες συνδέονται οι τεχνητοί BPC με τους αντίστοιχους ΗC, δίνοντας και σε αυτούς αντιστοιχία με τον χώρο. Στο τέλος αξιοποιείται η οργάνωση κέντρου περιφέρειας που έχουν οι φέτες, συνδέοντας τους τεχνητούς BPC με τα τεχνητά GC. Όλη αυτή η συνδεσμολογία έχει ως τελικό αποτέλεσμα για κάθε κυψέλη της εξαγωνικής ψηφίδωσης την ύπαρξη ενός συνόλου από αντικείμενα που έχουν ως αποκλειστικό στόχο την επεξεργασία αυτού του τμήματος της εικόνας.



9-21: Τοποθέτηση χάρτη υποδεκτικού πεδίου στη κυψέλη του τεχνητού αμφιβληστροειδούς
©2011 Horasis

Φτάνοντας σε αυτό το σημείο μπορούμε να ορίσουμε μια νέα ορολογία για να χαρακτηρίσουμε το παραπάνω, έχοντας στο νου ότι το κβάντο ⁷⁵(quantum) της ψηφιδωτής εικόνας είναι στα αγγλικά το pixel από τις λέξεις **picture element** και στα ελληνικά το εικονοστοιχείο. Ορίζουμε μέσα σε ένα σύστημα τεχνητής νοημοσύνης που υλοποιεί τεχνητό αμφιβληστροειδή ως **rexel**, από τις αγγλικές λέξεις **retinal element**, μια κυψέλη οποιασδήποτε μορφής, που αντιστοιχεί σε συγκεκριμένο χώρο της εικόνας, αρκεί η ψηφίδωση (tessellation) αυτής της μορφής να μπορεί να καλύψει ομοιόμορφα όλη την εικόνα. Αυτή η κυψέλη θα αντιστοιχεί την οπτική πληροφορία αυτού του χώρου σε εισόδους του συστήματος TN που πλέον θα χαρακτηρίζονται λόγω αυτής ως προς την θέση τους στην εικόνα. Μια ελληνική μετάφραση του όρου είναι **ρετινοστοιχείο** σε αντιστοιχία με το εικονοστοιχείο.

Έτσι σε όλον τον κυκλικό δίσκο που καλύπτει ένας τεχνητός ρετινοτοπικός χάρτης με 25 επίπεδα και 230 μέγιστη εκκεντρότητα, μπορούμε πλέον να λέμε ότι υπάρχουν 1017 rexel ή ρετινοστοιχεία. Η εικόνα που παράγεται από την έξοδο όλων των τεχνητών ΗC στις 13321 φέτες αυτών των ρετινοστοιχείων σχηματίζει την διπλανή εικόνα. Παρατηρούμε, ότι με την απλή θεώρηση έχοντας προκαθορισμένο το κέντρο της όρασης, δεν κεντράρεται με ακρίβεια η εικόνα του αντικειμένου. Σε μια μελλοντική έκδοση, υλοποιώντας το υποσύστημα FCP θα είχαμε σαφώς καλύτερη ακρίβεια.



9-22: Κυκλικός δίσκος τεχνητού αμφιβληστροειδή που εμφανίζει την έξοδο των τεχνητών ΗC που βρίσκονται σε

οπτικού φλοιού, όπου ο χάρτης αυτός αποτελεί βάση για την δημιουργία των υποδεκτικών πεδίων που υπάρχουν εκεί.

⁷⁵ κβάντο είναι η μικρότερη ποσότητα μιας φυσικής οντότητας που εμπλέκεται σε μια διάδραση.

Η μαθηματική έκφραση κυκλικού υποδεκτικού πεδίου με ανταγωνισμό κέντρου και περιφέρειας.

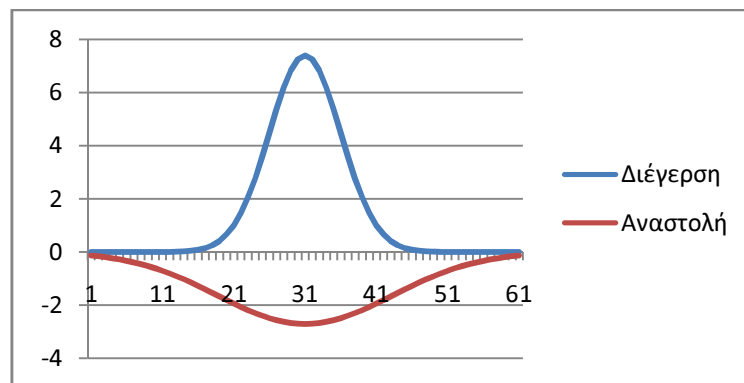
Με τελικό στόχο να εκπαιδύσουμε το ΤΝΔ που μοντελοποιεί ένα τεχνητό GC, έγιναν κάποιες θεωρήσεις, στην προσπάθεια να προσομοιώσουμε την οργάνωση κέντρου-περιφέρειας σε ένα κυκλικό υποδεκτικό πεδίο.

Το μοντέλο που χρησιμοποιείται συνήθως για την αναπαράσταση των αποκρίσεων που έχουν κυκλικά υποδεκτικά πεδία με οργάνωση κέντρου-περιφέρειας στα GC, στηρίζεται στην συναρτήσεις κατανομής gauss [123] και η διατύπωση του έγινε το 1965 από τον Rodieck [122]. Προσπαθώντας να εκφράσουμε κάτι αντίστοιχο, χωρίς να μπούμε στο μαθηματικό βάθος που υπάρχει πίσω του, φτιάχνουμε δύο συναρτήσεις που χρησιμοποιούν συναρτήσεις gauss $fnd_{\mu,\sigma}(x)$ με μ μέσο και σ απόκλιση, για την διέγερση και την αναστολή αντίστοιχα. Θεωρούμε ότι έχουμε ένα κυκλικό υποδεκτικό πεδίο διαμέτρου D , όπου ο δίσκος του κέντρου έχει διάμετρο C . Για την διέγερση που παράγεται από την ύπαρξη του βασικού χρώματος κατά μήκος μιας διαμέτρου του κυκλικού υποδεκτικού πεδίου, θα ορίσουμε μια συνάρτηση σε σχέση με ένα σημείο X πάνω σε αυτήν:

$$gex_{C,D} = e^s \left(\frac{fnd_{\frac{D}{2}, C}(x)}{fnd_{\frac{D}{2}, C}(D/2)} \right) \quad (9.17)$$

Το s είναι ένας συντελεστής που εκφράζει το πόσες φορές μεγαλύτερη είναι η τιμή της μέγιστης δυνατής διέγερσης του νευρώνα σε σχέση με την απόλυτη τιμή της μέγιστης αναστολής του. Αντίστοιχα για την αναστολή που παράγεται από την ύπαρξη ανταγωνιστικού χρώματος ξεκινώντας από την περιφέρεια με $X = 0$, περνώντας προς στο κέντρο μετά το $X = \frac{D-C}{2}$ και καταλήγοντας στο τέλος της διαμέτρου με $X = D$ ορίζουμε την συνάρτηση:

$$gin_{C,D} = -e \left(\frac{fnd_{\frac{D}{2}, (\frac{12}{5})C}(x)}{fnd_{\frac{D}{2}, (\frac{12}{5})C}(D/2)} \right) \quad (9.18)$$



9-23: Συναρτήσεις υπολογισμού διέγερσης και αναστολής τεχνητού GC, κατά μήκος διαμέτρου κυκλικού υποδεκτικού πεδίου $s = 2$, $D = 61$ με διάμετρο κέντρου $C = 5$. ©2011 Horasis

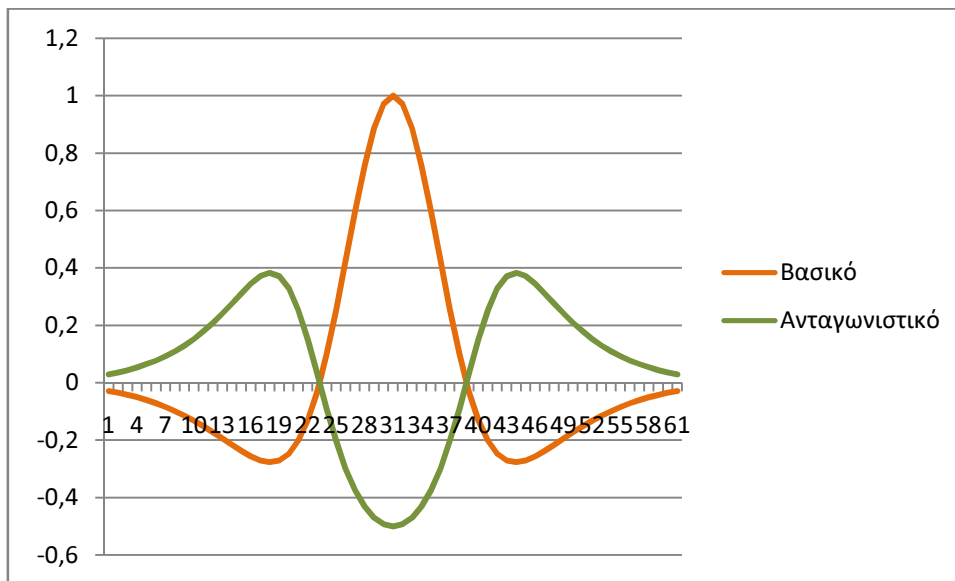
Όπως υποδεικνύουν οι τύποι, και φαίνεται παραστατικά στις γραφικές παραστάσεις η απόκλιση σ της καμπάνας gauss για το βασικό χρώμα, είναι $\frac{12}{5}$ ή **2,4** επί της απόκλισης της καμπάνας gauss για το ανασταλτικό χρώμα

Προσπαθώντας να ενοποιήσουμε τις δύο παραπάνω συναρτήσεις, φτιάχτηκαν άλλες δύο που εκφράζουν την ποσότητα διέγερσης που προκαλείται και στις δύο περιπτώσεις των χρωμάτων, με θετικούς αριθμούς να δείχνουν μεγαλύτερη ποσότητα και αρνητικούς μικρότερη:

$$fbc_{C,D} = \frac{(gex_{C,D} + gin_{C,D})}{e^s - e}, \text{ για το βασικό χρώμα (basic color)} \quad (9.19),$$

$$foc_{C,D} = \frac{1}{(fex_{C,D} + 1)} - 1, \text{ για το ανταγωνιστικό χρώμα (opponent color)} \quad (9.20)$$

Οι παραπάνω συναρτήσεις εκφράζουν μια ενιαία απόκριση για διέγερση που περιέχει και χρωματικό ανταγωνισμό και οργάνωση κέντρου περιφέρειας, σε ένα τεχνητό GC. Για να γίνει αυτό καλύτερα κατανοητό, στην παρακάτω γραφική παράσταση εμφανίζονται η διέγερση που προκαλείται από κόκκινο χρώμα ενός R-GC και η διέγερση που προκαλείται από το πράσινο χρώμα στο ίδιο τεχνητό κύτταρο. Βλέπουμε ότι στην περιφέρεια το πράσινο διεγείρει και το κόκκινο αναστέλλει, ενός στο κέντρο το κόκκινο διεγείρει και το πράσινο αναστέλλει. Οι ποσότητες φτιάχτηκαν με το παραπάνω μαθηματικό μοντέλο ώστε να είναι ασύμμετρες. Με το καλύτερο σήμα στο R-GC κόκκινο κέντρο και πράσινη περιφέρεια, η συνολική διέγερση για ολόκληρη την διάμετρο έστω ότι είναι A. Με το χειρότερο σήμα πράσινο κέντρο και κόκκινη περιφέρεια, η συνολική αρνητική ποσότητα διέγερσης B είναι μικρότερη κατά απόλυτη τιμή από την A.



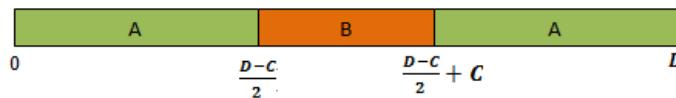
9-24: Συναρτήσεις ποσότητας ενεργοποίησης για βασικό και ανταγωνιστικό χρώμα, κατά μήκος διαμέτρου κυκλικού υποδεκτικού πεδίου $s = 2$, $D = 61$ με διάμετρο κέντρου $C = 5$. ©2011 Horasis

Η εκπαίδευση του ΤΝΔ που αντιπροσωπεύει ένα τεχνητό γαγγλιακό κύτταρο

Η κωδικοποίηση γαγγλιακού κυττάρου GCE που είναι η έξοδος ενός ΤΝΔ που μοντελοποιεί το τεχνητό GC, όπως είδαμε έχει τέσσερις τιμές στο το διάνυσμα της.

1. διέγερση
2. αναστολή
3. διάχυτο φως
4. ώρα μετάβασης προς το φώς

Με τις συναρτήσεις ποσότητας ενεργοποίησης μπορούμε να πάρουμε την ποσότητα ενεργοποίησης για συγκεκριμένο σημείο X πάνω στην διάμετρο ενός κυκλικού υποδεκτικού πεδίου. Έχουμε δει όμως ότι στο δικό μας τεχνητό GC, υπάρχει μόνο μια είσοδος για την τιμή του κέντρου, που προέρχεται από το τεχνητό BPC της φέτας 0μ ενώ μόνο 12 εισόδοι για τις τιμές της περιφέρειας από τις αντίστοιχες φέτες. Θα θεωρήσουμε γι αυτόν τον λόγο την ποσότητα ενεργοποίησης που προκαλείται από το βασικό χρώμα στην είσοδο 0 στο τεχνητό GC, να είναι ανάλογη με το άθροισμα όλων των ποσοτήτων στο ευθύγραμμο τμήμα B του κέντρου $[\frac{D-C}{2}, \frac{D-C}{2} + C]$. Επίσης η ποσότητα ενεργοποίησης για μια από τις εισόδους 1 έως 12 θα αναλογεί με το άθροισμα όλων των ποσοτήτων σε ένα ευθύγραμμο τμήμα A $[0, \frac{D-C}{2}]$ της περιφέρειας.



9-25: Διάμετρος κυκλικού υποδεκτικού πεδίου σε ευθύγραμμο τμήματα

Τα αθροίσματα υπολογίζονται ως εξής:

- κόκκινο στο κέντρο: $RC = \sum_{x=\frac{D-C}{2}}^{\frac{D-C}{2}+C} (fbc_{C,D})$ (9.21)

- κόκκινο στην περιφέρεια: $RP = \sum_{x=0}^{\frac{D-C}{2}} (fbc_{C,D})$ (9.22)

- πράσινο στο κέντρο: $GC = \sum_{x=\frac{D-C}{2}}^{\frac{D-C}{2}+C} (foc_{C,D})$ (9.23)

- πράσινο στο κέντρο: $GP = \sum_{x=0}^{\frac{D-C}{2}} (foc_{C,D})$ (9.24)

Στο παράδειγμα που εμφανίζει η γραφική παράσταση 9-24 μοντελοποιείται ένα R-GC με αναλογία διέγερσης-αναστολής $s = 2$, διάμετρο κυκλικού πεδίου $D = 61$ και διάμετρο κέντρου $C = 5$. Υπολογίζονται οι παραπάνω τιμές κατά προσέγγιση:

- $RC = 8,996$
- $RP = -3,241$
- $GC = 4,04$
- $GP = -5,227$

Έπειτα κάνουμε αναγωγή, έχοντας 1 για την ποσότητα για κόκκινο χρώμα στο κέντρο και $-1/s = -1/2 = -0.5$ για την ποσότητα του πράσινου χρώματος στο κέντρο. Η τελευταία θα γίνει αρνητική καθώς θέλουμε αναστολή όταν υπάρχει πράσινο χρώμα στο κέντρο. Οι νέες τιμές υπολογίζονται:

- $RC2 = 1$
- $RP2 = -0,36$
- $GC2 = -0.5$
- $GP2 = 0,386$

Με τα παραπάνω νούμερα και έχοντας για κανονικοποίηση δύο αριθμούς EN^{76} και IN^{77}

$$EN = e - 1, \quad IN = \frac{1}{e} - 1 \quad (9.25)$$

υπολογίζουμε τους συντελεστές των στόχων εκπαίδευσης που αποτελούν και σταθερές μέσα στο σύστημα Horasis:

- $CenterON = 1 / EN = 0.582$, ο συντελεστής διέγερσης κέντρου από βασικό χρώμα.
- $SurroundOFF = 0,386 / EN = 0.225$, ο συντελεστής διέγερσης περιφέρειας από ανταγωνιστικό χρώμα.
- $CenterOFF = -0,36 / IN = 0.791$, ο συντελεστής αναστολής κέντρου από ανταγωνιστικό χρώμα.
- $SurroundON = -0.5 / IN = 0.569$, ο συντελεστής αναστολής περιφέρειας από βασικό χρώμα.

Τελικά οι στόχοι εκπαίδευσης για τους δύο τεχνητούς νευρώνες εξόδου του ΤΝΔ ενός R-GC, που αντιστοιχούν σε ενεργοποίηση και αναστολή υπολογίζονται βάσει των παραπάνω συντελεστών και με αλγοριθμικό τρόπο ως εξής:

```
// Υπολογισμός στόχου για τεχνητό νευρώνα διέγερσης σε ΤΝΔ-GC
διαδικασία Υπολόγισε-Διέγερση (C_BPC, P_BPC[1..12]) υπολογίζει SumExc
έναρξη
  αν C_BPC >= 0 τότε
    CExc := C_BPC * CenterON
  ειδάλλως
    CExc := C_BPC * CenterOFF;

  PExc := 0
  για I := 1 έως 12 συνάθροισε
  αρχή
    αν P_BPC[I] >= 0 τότε
      PExc := PExc + P_BPC[I] * SurroundOFF
    ειδάλλως
      PExc := PExc + P_BPC[I] * SurroundON;
  τέλος

  SumExc := CExc + (PExc / 12);
τέλος;
```

⁷⁶ EN = Excitation Normalizer

⁷⁷ IN = Inhibition Normalizer

```

// Υπολογισμός στόχου για τεχνητό νευρώνα αναστολής σε ΤΝΔ-GC
διαδικασία Υπολόγισε-Αναστολή (C_BPC, P_BPC[1..12]) υπολογίζει SumInh
έναρξη
  αν C_BPC >= 0 τότε
    CInh := 0
  ειδάλλως
    CInh := C_BPC * CenterOFF;

  PInh := 0
  για I := 1 έως 12 συνάθροισε
  αρχή
    αν P_BPC[I] < 0 τότε
      PInh := PInh + P_BPC[I] * SurroundON;
  τέλος

  SumInh := - ( CInh + (PInh / 12) );
τέλος;

```

Και στους δύο ψευδοκώδικες χρωματίζονται οι μεταβλητές και οι σταθερές ανάλογα με τον χρωματικό ανταγωνισμό κόκκινου και πράσινου για ένα τεχνητό R-GC με εισόδους έναν R-BPC και 12 G-BPC. Μια σταθερά όπως η *SurroundON* εμφανίζεται στον ψευδοκώδικα χρωματισμένη ώστε να δίνει την χρωματική της πληροφορία, δηλαδή ότι το κόκκινο χρώμα ανάβει στην περιφέρεια.

Μια πρώτη παρατήρηση είναι ότι για την περιφέρεια και στους δύο στόχους υπολογίζεται ο μέσος όρος των BPC της περιφέρειας στις φέτες 1 έως 12, ενώ για το κέντρο υπολογίζεται άμεσα η τιμή από τον BPC του βασικού χρώματος της φέτας 0. Μια δεύτερη παρατήρηση είναι ότι η συνάθροιση της αναστολής στον δεύτερο αλγόριθμο, δίνει πάντα αρνητικό αριθμό οπότε γίνεται αντιστροφή του προσήμου για τον στόχο της αναστολής.

Πλέον μπορούμε να δούμε και το τρίτο είδος στόχου για την εκπαίδευση του τεχνητού νευρώνα που αντιστοιχεί στο διάχυτο φως. Εδώ, αν και δεν είναι είσοδοι στο ΤΝΔ-GC, θα χρησιμοποιηθούν οι τεχνητοί BPC της περιφέρειας του βασικού χρώματος για να δημιουργηθεί ένας μέσος όρος των τιμών που έχουν στις εξόδους τους. Για το γνωστό παράδειγμα του R-GC υπολογίζουμε το ποσό του διάχυτου φωτισμού *Diffuse* ως:

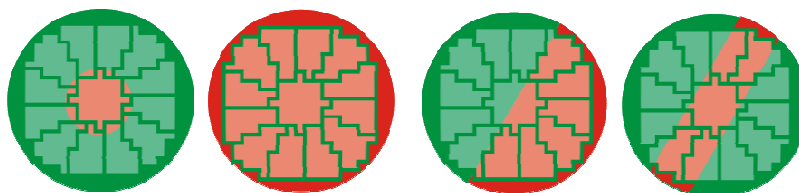
$$P = \left(\frac{1}{12}\right) \sum_{i=1}^{12} y_{R-BPC}[i] \text{ και } C = y_{R-BPC}[0] \quad (9.26)$$

- Αν $C > P$ τότε $Diffuse = \frac{P}{C}$ με $C <> 0$ ή $Diffuse = 0$ με $C = 0$
- Αν $C = P$ τότε $Diffuse = 1$ με $C <> 0$
- Αν $P < C$ τότε $Diffuse = \frac{C}{P}$ με $P <> 0$ ή $Diffuse = 0$ με $P = 0$

Παρατηρούμε ότι ο διάχυτος φωτισμός γίνεται 1 όταν φωτίζεται με την ίδια ακριβώς βασική απόχρωση Η και το κέντρο και η περιφέρεια, $Diffuse = -1$ όταν φωτίζονται με την ανταγωνιστική και $Diffuse \rightarrow 0$ όταν η απόκριση στην μια από τις δύο περιοχές μεγαλώνει και στην άλλη να μικραίνει.

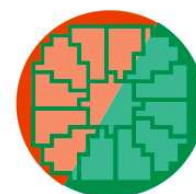
Η τέταρτη τιμή στο διάνυσμα της GCE, δηλαδή η ώρα μετάβασης προς το φως δεν υπολογίζεται μαθηματικά αλλά τίθεται ανάλογα με τον σχεδιασμό των προτύπων. Έτσι κατά την δημιουργία των προτύπων σχεδιάζονται μοτίβα διαφόρων τύπων που ανήκουν σε τέσσερις κλάσεις.

- i. Μικρός κυκλικός φωτισμός του κέντρου με το βασικό χρώμα και της περιφέρειας με το ανταγωνιστικό.
- ii. Διάχυτος φωτισμός με το βασικό χρώμα και αργότερα διάχυτος με το ανταγωνιστικό του.
- iii. Φωτισμός ενός ημικυκλίου με βασικό χρώμα και του υπολοίπου με το ανταγωνιστικό.
- iv. Μπάρα βασικού χρώματος που διασχίζει το κέντρο πάνω σε φόντο του ανταγωνιστικού χρώματος.



9-26: Τέσσερις κλάσεις προτύπων εκπαίδευσης ενός ΤΝΔ R-GC ©2011 Horasis

Μόνο στην τρίτη κλάση που έχουμε ημικύκλια μπορεί να οριστεί μια ώρα μετάβασης προς το φως θεωρώντας έναν ωροδείκτη του ρολογιού πάνω στο σημείο εναλλαγής των χρωμάτων. Στην παραπάνω εικόνα που έχουμε ως βασικό χρώμα το κόκκινο στην δεξιά πλευρά θεωρούμε ότι έχουμε 1 η ώρα (01:00) μετάβαση από το πράσινο στο κόκκινο. Η σειρά ανάγνωσης είναι από τα αριστερά προς τα δεξιά έτσι στο σχήμα 9-27 που τα χρώματα εμφανίζονται αντίστροφα, θεωρείται 7 η ώρα (07:00) για την μετάβαση από το πράσινο στο κόκκινο. Συνεχίζοντας με αυτήν την λογική περιστρέφουμε το ημικύκλιο του βασικού φωτισμού ανά 30° σε 12 θέσεις μια για κάθε ώρα. Στις υπόλοιπες κλάσεις προτύπων που δεν υπάρχει αυτή η έννοια θεωρούμε την τιμή της ώρας μετάβασης του φωτός σε 0. Τέλος για να έχουμε έναν στόχο από 0 έως 1 διαιρούμε την τιμή την ώρας δια του 12, και έχουμε έναν δεκαδικό αριθμό για στόχο εκπαίδευσης του τέταρτου τεχνητού νευρώνα εξόδου.

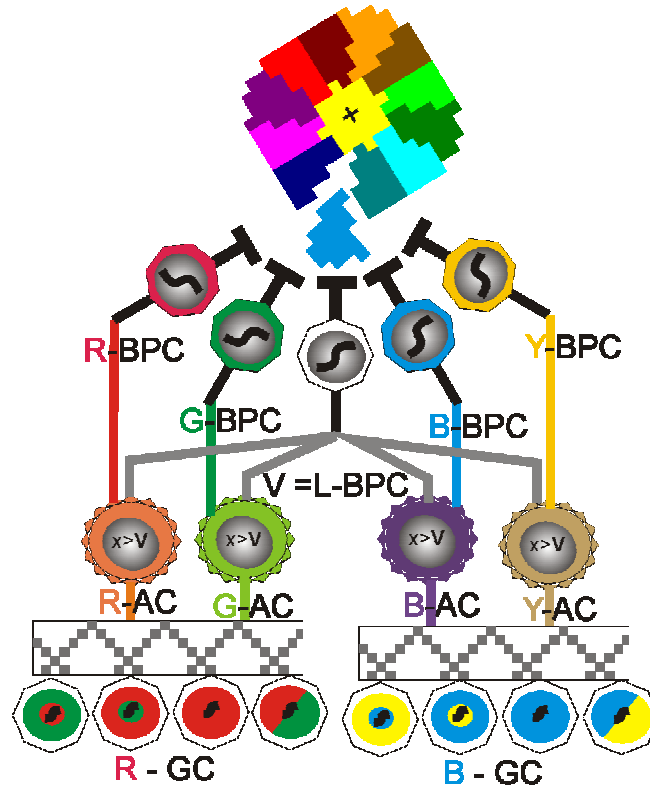


9-27: Μετάβαση από το πράσινο στο κόκκινο μεγαλύτερη της 06:00 ©2011 Horasis

Με αυτό το τέχνασμα ξεπερνούμε μια απροσδιοριστία που γεννούσε ο διαφορετικός χρωματισμός δύο ίσων ημικυκλίων, μοντελοποιώντας παράλληλα και μια ευαισθησία στην κατεύθυνση που έχει παρατηρηθεί από την νευροεπιστήμη και στα βιολογικά GC [154]. Έχει δημιουργηθεί πλέον μια απόκριση τεσσάρων διαστάσεων για το τεχνητό GC, που καλύπτει πολλά διαφορετικά σενάρια φωτισμού του υποδεκτικού πεδίου. Τα δεδομένα της κωδικοποίησης GCE, όπως θα δούμε πετυχαίνουν στον στόχο τους, δηλαδή να χρησιμοποιηθούν ως είσοδος για την εκπαίδευση των επόμενων ΤΝΔ του τμήματος του τεχνητού φλοιού. Οι αποκρίσεις ολόκληρης της κυψέλης του ρετινοτοπικού χάρτη συνθέτουν την νευρωνική διαμόρφωση αμφιβληστροειδούς RNF για την εικόνα, η οποία αποθηκεύεται τελικά και στην βάση δεδομένων.

Τεχνητά βραχύινα κύτταρα

Αν και φτάσαμε πρώτα στην τελική έξοδο του τεχνητού αμφιβληστροειδή θα επιστρέψουμε λίγο πίσω, για να δούμε μια επιπρόσθετη υλοποίηση που παρεμβάλλεται μεταξύ της εξόδου των τεχνητών BPC και των τεχνητών GC. Κρατώντας ότι έχουμε πει για την συνδεσμολογία μεταξύ τους, μπορούμε να θεωρήσουμε ότι παρεμβάλλονται μεταξύ των δύο τα τεχνητά βραχύινα κύτταρα. Σε αυτά ως είσοδοι έχουμε την έξοδο x από ένα τεχνητό BPC χρώματος (R,G,B ή Y) και την έξοδο L του τεχνητού L-PC, που μετράει το επίπεδο του λευκού φωτός στην ίδια φέτα.



9-28: Δύο ΤΝΔ-GC και η συνδεσμολογία τους με AC και τεχνητά δίπολα ©2011 Horasis

Αν το επίπεδο στην είσοδο φωτεινότητας ξεπεράσει ένα κατώφλι V^{78} τότε η έξοδος του τεχνητού AC νευρώνα επιστρέφει αυτούσια την τιμή της χρωματικής εισόδου x . Αν είναι παρακάτω από το κατώφλι V τότε επιστρέφει 0. Και στις δύο περιπτώσεις συνδέεται η έξοδος από το τεχνητό AC ως είσοδος στο τεχνητό GC, παρεμβαίνοντας στην μέση της αρχικής συνδεσμολογίας. Μπορούμε να θεωρήσουμε γενικά ότι υπάρχει μια συνάρτηση ενεργοποίησης για τον τεχνητό AC νευρώνα της μορφής $y = f_{AC}(R, G, B, L)$, η οποία μελλοντικά μπορεί να είναι διαφορετική από την συνάρτηση κατωφλίου που υλοποιείται εδώ. Για παράδειγμα σε έναν R-AC έχουμε $L, R = x, G = 0, B = 0$ και συνάρτηση ενεργοποίησης $y = x, x \geq V, y = 0, x < V$. Μπορούν μελλοντικά να γίνουν πειραματισμοί σε σχέση με την λειτουργία του τεχνητού AC και την επίδραση του στην λειτουργία όλου του τεχνητού αμφιβληστροειδούς. Με την συνάρτηση κατωφλίου και την παρεμβολή των νευρώνων αυτών, αγνοούνται χρωματικές αποκρίσεις για μια πολύ σκοτεινή περιοχή της εικόνας. Οι κυψέλες που βλέπουν ένα τμήμα του μαύρου φόντου της εικόνας δεν θα έχουν αποκρίσεις στα αντίστοιχα χρωματικά τεχνητά GC τους.

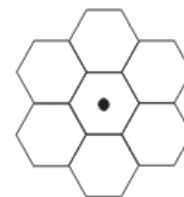
⁷⁸ V από την λέξη visible, δηλαδή αν η φέτα είναι ορατή ή σκοτεινή

Τεχνητός Οπτικός Φλοιός

Μικρά Απλά Κύτταρα V1

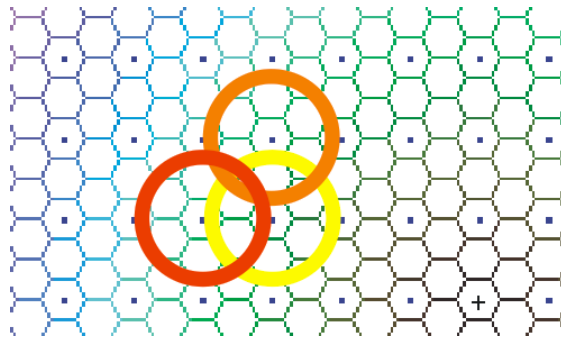
Είδαμε τον τρόπο με τον οποίο μια αρχική διαμόρφωση εικόνας BMP, εισαγόμενη σε ένα πολύπλοκο τεχνητό νευρωνικό δίκτυο, που αποτελεί τον τεχνητό αμφιβληστροειδή, μετατρέπεται σε νευρωνική διαμόρφωση αμφιβληστροειδούς RNF. Πλέον στο τμήμα III του Horasis, τον τεχνητό οπτικό φλοιό V1 ή AVC για συντομία, θα δούμε πως η είσοδος RNF μετατρέπεται στην τελική εσωτερική αναπαράσταση του αντικειμένου. Υπάρχει η δυνατότητα να αποθηκεύσουμε στην βάση δεδομένων την διαμόρφωση RNF σε συσχέτιση με την αρχική εικόνα, ώστε να την χρησιμοποιήσουμε κατευθείαν στο AVC χωρίς να χρειάζεται να στήσουμε όλα τα αντικείμενα τεχνητών νευρώνων του ΟΟΑΡ. Σίγουρα όμως χρειάζεται να υπάρχει ήδη προετοιμασμένος ο τεχνητός ρετινοτοπικός χάρτης ARTM, καθώς βάσει αυτού δημιουργούνται τα υποδεκτικά πεδία των τεχνητών νευρώνων του AVC.

Οι τεχνητοί νευρώνες που έχουν υλοποιηθεί στο Horasis για να προσομοιώσουν την λειτουργία των βιολογικών απλών κυττάρων V1 λέγονται τεχνητά κύτταρα V1S και διακρίνονται σε μικρά (parvi) PV1S και μεγάλα (magna) MV1S. Και τα δύο έχουν υποδεκτικά πεδία ευαίσθητα στην κατεύθυνση του οπτικού ερεθίσματος και συγκεκριμένα διεγείρονται από μια μπάρα βασικού χρώματος σε συγκεκριμένη κατεύθυνση πάνω σε ένα φόντο από ανταγωνιστικό χρώμα. Υπάρχει ένα πλήθος τεχνητών MV1S που έχει παράλληλη απόκριση πάνω στην ίδια περιοχή της εικόνας, δηλαδή έχοντας εισόδους από τα ίδια ρετινοστοιχεία ή rexels. Για κάθε rexel γίνονται συνάψεις και με τους δύο τεχνητούς GC, του βασικού και του ανταγωνιστικού χρώματος του. Για ένα R-MV1S έχουμε δηλαδή την είσοδο να προέρχεται από όλα τα R-GC και G-GC της περιοχής.



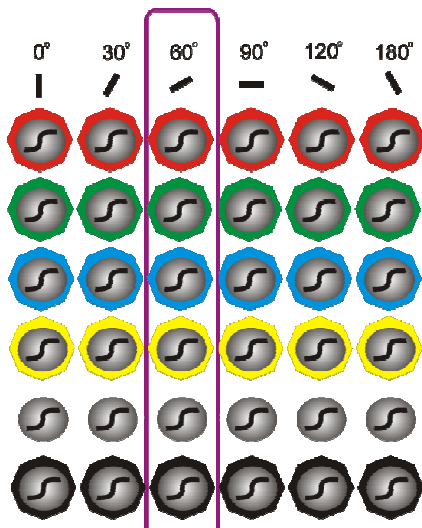
9-29: Υποδεκτικό πεδίο τεχνητού PV1S με 7 rexels

Το πλήθος των rexel της περιοχής, δηλαδή του υποδεκτικού πεδίου των τεχνητών PV1S, είναι 7 και στηρίζεται στον ARTM, όπου έχουμε ένα κεντρικό εξάγωνο και τα 6 εξάγωνα γύρω του. Κατά αυτόν τον τρόπο επιλέγονται συγκεκριμένα εξάγωνα ως κεντρικά του υποδεκτικού πεδίου των τεχνητών V1S, με την επιλογή να γίνεται ανά δύο κυψέλες κατά την οριζόντια διεύθυνση και ανά τέσσερις κατά την κάθετη. Από τα 1017 rexels στον τεχνητό ρετινοτοπικό χάρτη της πρώτης έκδοσης του συστήματος Horasis που εμφανίζεται στην εικόνα 9-19, επιλέγονται 155 ως κεντρικά και σημειώνονται με μια τελεία στο κέντρο τους. Λόγω του τρόπου επιλογής του κέντρου υπάρχει επικάλυψη (overlapping) ώστε ένα rexel της περιφέρειας να χρησιμοποιείται σε υποδεκτικά πεδία δύο διαφορετικών ομάδων τεχνητών V1S.



9-30: Επιλογή κεντρικών rexel για υποδεκτικά πεδία τεχνητών PV1S. Εμφανίζεται η επικάλυψη μεταξύ τριών ομάδων PV1S ©2011 Horasis

Το πλήθος τους για κάθε ομάδα εξαρτάται από τις μοίρες στις οποίες θέλουμε να μοντελοποιήσουμε την ευαισθησία του υποδεκτικού πεδίου. Οπότε αν θέλουμε να έχουμε μια διαφορετική απόκριση για μια κόκκινη μπάρα που πέφτει ανά s μοίρες τότε το πλήθος των τεχνητών R-V1S για μια περιοχή της εικόνας είναι $(\frac{180}{s})$. Επειδή έχουμε ένα πλήθος $C = 6$ ανταγωνιστικών χρωμάτων και στην παρούσα υλοποίηση του Horasis και θέλουμε ευαισθησία ανά 30° μπορούμε να υπολογίσουμε $n = 6 \left(\frac{180}{30}\right) = 6 * 6 = 36$ τεχνητούς νευρώνες PV1S που

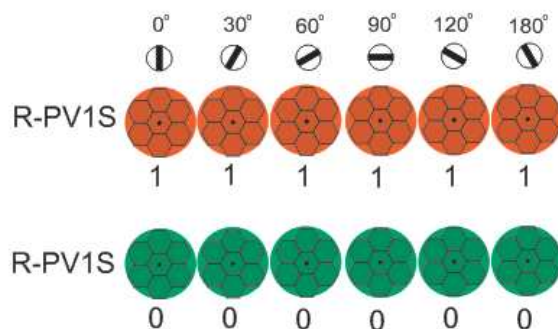


9-31: Στρώματα εξόδου διαφορετικών ΤΝΔ ανά χρώμα, και τεχνητή υπερστήλη των 60° (σε μωβ πλαίσιο) ©2011 Horasis

αποκρίνονται παράλληλα για την περιοχή των 7 rexel. Ανά 6 που ανήκουν στις ίδιες μοίρες σχηματίζεται μια τεχνητή υπερστήλη (hypercolumn) που έχει R-PV1S, G-PV1S, B-PV1S, Y-PV1S, L-PV1S και D-PV1S για τα αντίστοιχα ανταγωνιστικά χρώματα κόκκινο R, πράσινο G, μπλε B, κίτρινο Y, λευκό L και μαύρο D. Το συνολικό πλήθος των τεχνητών νευρώνων PV1S για όλα τα χρώματα, και όλες τις κατευθύνσεις για τον γνωστό ARTN που αναφέραμε είναι 5580.

Όλοι οι νευρώνες V1S ανήκουν σε στρώματα εξόδου 6 διαφορετικών ΤΝΔ που εκπαιδεύονται ανά χρώμα, έχοντας 6 εξόδους, μια για κάθε κατεύθυνση $0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ$ και 180° . Δηλαδή η τεχνητή υπερστήλη συντίθεται από τον νευρώνα σε μια συγκεκριμένη θέση στα στρώματα εξόδου των 6 χρωματικών ΤΝΔ.

Μετά την εκπαίδευση αυτών των ΤΝΔ, παράγουν τιμές στην έξοδο τους ανάλογη με το σχήμα το βασικού χρώματος που φωτίζει την περιοχή της εικόνας που δραστηριοποιούνται. Παίρνουμε ως παράδειγμα μια μπάρα κόκκινου χρώματος σε κατεύθυνση 60° σε σχέση με τον κάθετο άξονα ή αντίστοιχα δείκτη ρολογιού στις 2 η ώρα. Στο ΤΝΔ του κόκκινου-πράσινου θα έχουμε ως έξοδο το διάνυσμα $[0,0,1,0,0,0]$ δηλαδή θα ενεργοποιηθεί ο τρίτος τεχνητός νευρώνας R-PV1S και όλοι οι άλλοι θα απενεργοποιηθούν. Θεωρούμε για την περίπτωση που έχουμε ερέθισμα διάχυτου κόκκινου φωτός, ότι διεγείρονται παράλληλα όλοι οι τεχνητοί

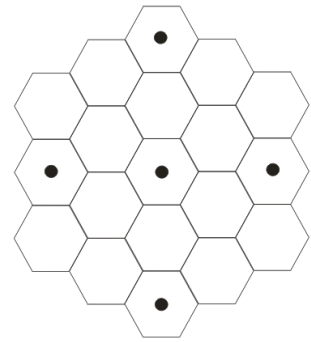


9-32: Διεγερση και αναστολή ενός ΤΝΔ R-PV1S με διάχυτο κόκκινο και πράσινο φως. ©2011 Horasis

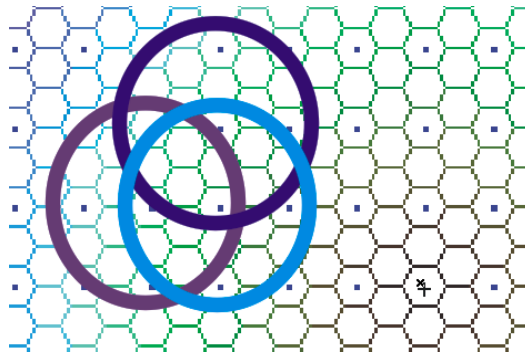
νευρώνες δίνοντας διάνυσμα εξόδου [1,1,1,1,1] ενώ αν έχουμε διάχυτο πράσινο φως αναστέλλονται όλοι με [0,0,0,0,0] στην έξοδο του ΤΝΔ των R-PV1S.

Μεγάλα Απλά Κύτταρα V1

Τα μεγάλα τεχνητά κύτταρα V1 ή MV1S οργανώνονται με ακριβώς τον ίδιο τρόπο με τα PV1S με διαφορά ότι έχουν μεγαλύτερο υποδεκτικό πεδίο που αποτελείται από 19 hexel. Δηλαδή έχουμε το ίδιο κεντρικό hexel, όλα τα διπλανά του και όλα τα διπλανά αυτών. Η επιλογή των κεντρικών εξαγώνων γίνεται με τον ίδιο τρόπο ανά δύο κυψέλες κατά την οριζόντια διεύθυνση και ανά τέσσερις κατά την κάθετη. Έχουμε σε αυτήν την περίπτωση μεγαλύτερη επικάλυψη των υποδεκτικών πεδίων τριών γειτονικών ομάδων τεχνητών MV1S. Σε αυτήν την επικάλυψη θα δούμε σε ένα υποδεκτικό πεδίο την ύπαρξη τουλάχιστον τεσσάρων κεντρικών κυψελών που σχηματίζουν έναν σταυρό.



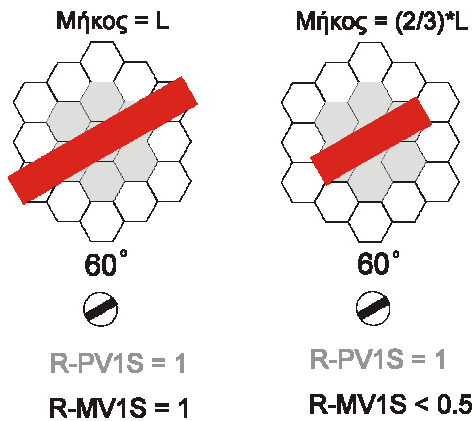
9-33: Υποδεκτικό πεδίο τεχνητού MV1S με 19 hexels



9-34: Επικάλυψη μεταξύ τριών ομάδων MV1S

©2011 Horasis

Πλέον μπορούμε να θεωρήσουμε ότι γύρω από το ίδιο κεντρικό εξαγώνο στον ARTM λειτουργεί μια ομάδα MV1S παράλληλα με μια ομάδα PV1S. Με αυτόν τον τρόπο μπορούμε να εντοπίσουμε



9-35: Δύο κόκκινες μπάρες διαφορετικού μήκους που ενεργοποιούν παράλληλα ένα μικρό και ένα μεγάλο τεχνητό V1S. Κάτω εμφανίζονται οι τιμές ενεργοποίησης.

©2011 Horasis

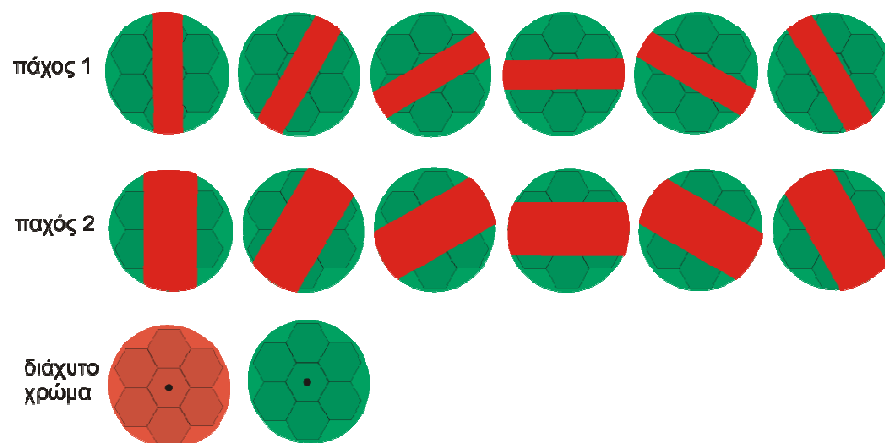
στην ίδια περιοχή της εικόνας δύο διαφορετικά μήκη σε μια μπάρα χρώματος συγκεκριμένης κατεύθυνσης αν συνδυάσουμε τις αποκρίσεις ενός PV1S και ενός MV1S. Στην οπτική πληροφορία που υπήρχε μέσα στην διαμόρφωση RNF που είχε μόνο χρωματικές διαστάσεις, μετά την επεξεργασία από τα μικρά και μεγάλα τεχνητά V1S, προσθέτονται οι διαστάσεις του σχήματος, του μήκους και της κατεύθυνσης μέσω της ανίχνευση γραμμών που γίνεται εδώ.

Στον τεχνητό οπτικό φλοιό θα μπορούσαν να δημιουργηθούν επιπρόσθετα τεχνητά σύνθετα και υπερσύνθετα κύτταρα για τον περαιτέρω εμπλουτισμό της οπτικής πληροφορίας. Στην πρώτη περίπτωση θα ανιχνεύονται γενικά ακμές αντί για γραμμές και στην δεύτερη περίπτωση γραμμές που τερματίζονται μέσα στο υποδεκτικό πεδίο αντί να το διασχίζουν από την μια άκρη στην άλλη. Χάρην απλότητας και κρίνοντας εκ του τελικού αποτελέσματος ότι δεν χρειάζεται επιπλέον ανίχνευση σχηματικών χαρακτηριστικών, υπάρχουν μόνο τεχνητά απλά κύτταρα V1 μικρού και μεγάλου μεγέθους στην τελική υλοποίηση του Horasis.

Η εκπαίδευση του ΤΝΔ που αντιπροσωπεύει ένα τεχνητό απλό κύτταρο V1.

Η δημιουργία των προτύπων εκπαίδευσης των ΤΝΔ των PV1S και MV1S έγινε με έναν πολύ απλό τρόπο. Πάνω σε μια περιοχή εικόνας BMP αντίστοιχη με 7 rexels ή 19 rexels αντίστοιχα, σχεδιάζοταν πρότυπα δύο κλάσεων:

- i. Μπάρα βασικού χρώματος σε 6 κατευθύνσεις επί δύο μεγέθη, πάνω σε φόντο του ανταγωνιστικού χρώματος.
- ii. Διάχυτος φωτισμός με το βασικό χρώμα ή το ανταγωνιστικό του.



9-36: Πρότυπα εκπαίδευσης ΤΝΔ τεχνητών PV1S ©2011 Horasis

Το σχέδιο που δημιουργήθηκε πάνω σε μια εικόνα BMP στην μνήμη, εισάγεται για την περίπτωση των PV1S σε έναν τεχνητό αμφιβληστροειδή με 2 επίπεδα, μέγιστη εκκεντρότητα 230pixels (που δεν ενεργοποιείται ποτέ) και διάμετρο κυψέλης 16pixels. Αυτός παράγει τις εξόδους των τεχνητών GC για τα 7 rexels. Οι έξοδοι αυτοί αποτελούν τις εισόδους ενός πρότυπου εκπαίδευσης που συνοδεύονται από ένα διάνυσμα στόχων με 6 τιμές. Οι στόχοι τίθενται ανάλογα με την περίπτωση:

- Μια τιμή στο 1 και οι υπόλοιπες στο 0, για μπάρα βασικού φωτός σε συγκεκριμένη κατεύθυνση.
- Όλες οι τιμές στο 1 για διάχυτο φωτισμό με το βασικό χρώμα.
- Όλες οι τιμές στο 0 για διάχυτο φωτισμό με το ανταγωνιστικό χρώμα.

Στην περίπτωση των MV1S ακολουθείται η ίδια ακριβώς λογική σε μεγαλύτερο καμβά σχεδίασης, με τον τεχνητό αμφιβληστροειδή να έχει 3 επίπεδα αντί για 2 και να παράγει εξόδους τεχνητών GC για 19 rexels. Έτσι στα ΤΝΔ των τεχνητών V1S με 6 εξόδους κατεύθυνσης, το πλήθος των εισόδων υπολογίζεται ανάλογα με το μέγεθος τους:

- μικρά (parvi) : 7 rexel * 2 τεχνητά GC * 4 τιμές της GCE = 56 είσοδοι
- μεγάλα (magna): 7 rexel * 2 τεχνητά GC * 4 τιμές της GCE = 152 είσοδοι

Εσωτερική αναπαράσταση της εικόνας του αντικειμένου

Σε ένα τεχνητό ρετινοτοπικό χάρτη με 25 επίπεδα, διάμετρο κυψέλης 16 pixels και μέγιστη εκκεντρότητα 230pixels είδαμε ότι υπάρχουν 155 κεντρικά rexel. Σε κάθε κεντρικό rexel αντιστοιχούν 36 PV1S και 36 MV1S. Έτσι έχουμε $155 * 36 * 2 = 11160$ διαφορετικά τεχνητά κύτταρα οπτικού φλοιού V1. Αν θυμηθούμε τις 750.000 τιμές RGB μιας εικόνας 500x500 pixel πλέον το πλήθος των εισόδων σε ένα ΤΝΔ ταξινόμησης οπτικών προτύπων έχει μειωθεί στο 1,5% των αρχικών. Το πιο σημαντικό όμως είναι ότι η τιμή σε κάθε μια από τις νέες εισόδους έχει ιδιαίτερη σημασία καθώς αντιπροσωπεύει ένα χρώμα από τα έξι ανταγωνιστικά, ανίχνευση αντίθεσης κέντρου περιφέρειας σε συνδυασμό με ανίχνευση γραμμών και κατεύθυνσης. Επίσης οι τιμές αυτές οργανώνονται σε ομάδες που λειτουργούν παράλληλα για την ίδια περιοχή της εικόνας και περιπλέκονται μεταξύ τους, λόγω της επικάλυψης που υπάρχει στον τεχνητό οπτικό φλοιό.

Το μόνο που απομένει για την δημιουργία της εσωτερικής αναπαράστασης της εικόνας του πραγματικού αντικειμένου είναι να συμπιέσουμε αυτό το πλήθος εισόδων ακόμα περισσότερο, συνδυάζοντας τα ανταγωνιστικά μεταξύ τους τεχνητά V1S ως εισόδους σε ένα τελευταίο ΤΝΔ. Γι αυτό θα πάρουμε για ένα βασικό χρώμα, έστω το κόκκινο, τα R-PV1S και R-MV1S και επίσης τα δύο ανταγωνιστικά τους G-PV1S και G-MV1S. Από τις τιμές εξόδου που ανήκουν στο $[0,1] \in \mathbb{R}$ θα υπολογίσουμε τις απόλυτες τιμές:

- $D_{Parvi} = |R_{PV1S} - G_{PV1S}|$
- $D_{Magna} = |R_{MV1S} - G_{MV1S}|$

και θα έχουμε τις εξής περιπτώσεις στόχων για το ένα μέγεθος, εδώ παράδειγμα το μικρό (parvi).

- 1) Αν $D_{Parvi} < F$ τότε $t_{Parvi} = 0$ όπου F ένας πολύ μικρός αριθμός, που καθορίζει το επίπεδο κάτω από το οποίο οι δύο τιμές θεωρούνται ίσες.
- 2) Αν $R_{PV1S} > G_{PV1S}$ και $R_{PV1S} > E$ τότε $t_{Parvi} = 1$, όπου $E \in [0,1] \in \mathbb{R}$, το κατώφλι για την παραγωγή διεγερτικής (θετικής) εξόδου.
- 3) Αν $R_{PV1S} < G_{PV1S}$ και $G_{PV1S} > I$ τότε $t_{Parvi} = -1$, όπου $I \in [0,1] \in \mathbb{R}$, το κατώφλι για την παραγωγή ανασταλτικής (αρνητικής) εξόδου.

Ομοίως υπολογίζεται και ο δεύτερος στόχος της εκπαίδευσης για το μεγάλο (magna).

- 1) Αν $D_{Magna} < F$ τότε $t_{Magna} = 0$ όπου F ένας πολύ μικρός αριθμός, που καθορίζει το επίπεδο κάτω από το οποίο οι δύο τιμές θεωρούνται ίσες.
- 2) Αν $R_{MV1S} > G_{MV1S}$ και $R_{MV1S} > E$ τότε $t_{Magna} = 1$, όπου $E \in [0,1] \in \mathbb{R}$, το κατώφλι για την παραγωγή διεγερτικής (θετικής) εξόδου
- 3) Αν $R_{MV1S} < G_{MV1S}$ και $G_{MV1S} > I$ τότε $t_{Magna} = -1$, όπου $I \in [0,1] \in \mathbb{R}$, το κατώφλι για την παραγωγή ανασταλτικής (αρνητικής) εξόδου

Οι στόχοι για να λειτουργήσει η εκπαίδευση προβάλλονται στο $[0,1] \in \mathbb{R}$. Στο νέο ΤΝΔ δηλαδή έχουμε 4 εισόδους και 2 εξόδους με την κάθε έξοδο να αντιστοιχεί σε μέγεθος αλλά με την τιμή της να δείχνει θετικό ή αρνητικό ανάλογα με το αν ενεργοποιείται το πρώτο ή το δεύτερο χρώμα σε ένα ζευγάρι χρωματικού ανταγωνισμού. Στην έξοδο του ΤΝΔ για μια συγκεκριμένη κατεύθυνση θα έχουμε έξι τιμές,

- RG-PV1S: Κόκκινη μπάρα σε πράσινο φόντο = 1, Πράσινη μπάρα σε κόκκινο φόντο = -1
- BY-PV1S: Μπλε μπάρα σε κίτρινο φόντο = 1, Κίτρινη μπάρα σε μπλε φόντο = -1
- LD-PV1S: Λευκή μπάρα σε μαύρο φόντο = 1, Μαύρη μπάρα σε λευκό φόντο = -1
- RG-MV1S: Κόκκινη μπάρα σε πράσινο φόντο = 1, Πράσινη μπάρα σε κόκκινο φόντο = -1
- BY-MV1S: Μπλε μπάρα σε κίτρινο φόντο = 1, Κίτρινη μπάρα σε μπλε φόντο = -1
- LD-MV1S: Λευκή μπάρα σε μαύρο φόντο = 1, Μαύρη μπάρα σε λευκό φόντο = -1

και στις υπόλοιπες περιπτώσεις που το πρότυπο δεν ταιριάζει με το υποδεκτικό πεδίο θα επιστρέφεται 0. Για κάθε χρωματικό ανταγωνισμό τώρα θα δημιουργήσουμε μια ενιαία απόκριση για μικρές και μεγάλες μπάρες από το -1 έως το 1 με την εξής λογική:

Αν οι τιμές εξόδου P και M των PV1S και MV1S αντίστοιχα, έχουν ίδια πρόσημα τότε η απόκριση είναι $R = 0.4 * M + 0.6 P$ (9.27). Αν τα πρόσημα διαφέρουν η απόκριση γίνεται 0. Για ίδια πρόσημα και την ενεργοποίηση, έχουμε αναλυτικότερα:

- RG-V1S:
 - Μεγάλη κόκκινη μπάρα σε πράσινο φόντο = 1,
 - Μικρή κόκκινη μπάρα σε πράσινο φόντο = 0.6,
 - Μικρή πράσινη μπάρα σε κόκκινο φόντο = -0.6,
 - Μεγάλη πράσινη μπάρα σε κόκκινο φόντο = -1
- BY-V1S:
 - Μεγάλη μπλε μπάρα σε κίτρινο φόντο = 1,
 - Μικρή μπλε μπάρα σε κίτρινο φόντο = 0.6,
 - Μικρή κίτρινη μπάρα σε μπλε φόντο = -0.6,
 - Μεγάλη κίτρινη μπάρα σε μπλε φόντο = -1
- LD-V1S:
 - Μεγάλη λευκή μπάρα σε μαύρο φόντο = 1,
 - Μικρή λευκή μπάρα σε μαύρο φόντο = 0.6,
 - Μικρή μαύρη μπάρα σε λευκό φόντο = -0.6,
 - Μεγάλη μαύρη μπάρα σε λευκό φόντο = -1

Με αυτόν τον τρόπο οι τέσσερις αρχικές είσοδοι στο ΤΝΔ παράγουν μια μόνο τιμή και το πλήθος των δεδομένων από 11160 γίνεται 2790. Αυτό το διάγραμμα αποτελεί την εσωτερική αναπαράσταση της εικόνας ενός αντικειμένου στο Horasis και θα χρησιμοποιηθεί ως πρότυπο για την εκπαίδευση του ΤΝΔ κατηγοριοποίησης. Γι αυτό κάθε εικόνα αφού μετατραπεί σε εσωτερική αναπαράσταση αποθηκεύεται στην βάση, ώστε να φορτωθεί γρήγορα όταν χρειαστεί να δημιουργηθούν τα πρότυπα εκπαίδευσης.

Φθάνοντας στο τέλος βλέπουμε ότι όλες οι εκπαιδεύσεις των ΤΝΔ γίνονται με αυτόνομο τρόπο χωρίς να υπάρχει κάποια οπτική πληροφορία από ένα πρότυπο εκπαίδευσης. Και εδώ έχουμε μια ομοιότητα με την βιολογία καθώς ο άνθρωπος γεννιέται με τις λειτουργίες της όρασης πολύ πριν να αρχίσει να λαμβάνει καθαρά οπτικά πρότυπα. Αυτές αναπτύσσονται εξελικτικά στα πρώτα χρόνια της ζωής του μέχρι να τελειοποιηθούν. Σε εμάς τα τμήματα II και III υπάρχουν ώστε να μπορεί το σύστημα να λαμβάνει και να αποθηκεύει οπτικές πληροφορίες για τον πραγματικό κόσμο. Το πως θα μπορέσει να αξιοποιήσει αυτές τις εσωτερικές αναπαραστάσεις παρακάτω, δεν είναι ένα θέμα που αφορά ειδικά την τεχνητή όραση. Αφορά γενικά την τεχνητή νοημοσύνη και γι αυτό υπάρχει στο τέλος μόνο ένα ΤΝΔ που κατηγοριοποιεί αυτές τις εσωτερικές αναπαραστάσεις των πραγματικών αντικειμένων.

Συνδεσμολογία ΤΝΔ στο Horasis

Πριν δούμε στο επόμενο κεφάλαιο την κατηγοριοποίηση στο τελευταίο ΤΝΔ, θα κάνουμε μια σύνοψη όλων των τεχνητών νευρώνων και νευρωνικών δικτύων που υλοποιήθηκαν ώστε να δημιουργηθεί η εσωτερική αναπαράσταση του αντικειμένου. Στην παρακάτω εικόνα εμφανίζεται η ροή της οπτικής πληροφορίας ξεκινώντας από την μορφή BMP.



9-37: Τα τεχνητά νευρωνικά δίκτυα του Horasis που μοντελοποιούν την όραση, από την BMP μορφή της εικόνας έως την εσωτερική αναπαράσταση της εικόνας. ©2011 Horasis

10. Οργάνωση και Υλοποίηση

Βασικό χαρακτηριστικό στον κύκλο ζωής του Horasis, βάσει της ενοποιημένης διεργασίας ανάπτυξης λογισμικού UP, είναι η αρχή (discipline) της υλοποίησης. Σε όλες τις φάσεις της διεργασίας, η υλοποίηση εμφανίζεται όπως είδαμε στην εικόνα 8-2 έστω και σε μικρό βαθμό. Ο κύριος όγκος της εκτελείται στις επαναλήψεις που γίνονται μέσα στις φάσεις UP επεξεργασία (elaboration) και κατασκευή (construction). Έχοντας γνωρίζει μέσα από τα προηγούμενα κεφάλαια το επιστημονικό πλαίσιο, τις θεωρητικές βάσεις, το ερευνητικό υπόβαθρο, την αρχιτεκτονική και τον σχεδιασμό μπορούμε πλέον να αναφερθούμε στο τελικό αποτέλεσμα της εφαρμογής λογισμικού Horasis. Θα δούμε ορισμένες λεπτομέρειες που αφορούν την πρώτη έκδοση της υλοποίησης η οποία δημιουργήθηκε για τις ανάγκες της πτυχιακής εργασίας, την τελική εφαρμογή της ιδέας και όλες τις τιμές που την παραμετροποιούν.

Η αυξητική και επαναληπτική υλοποίηση, συνοπτικά

Η υλοποίηση όπως είπαμε έγινε αυξητικά και επαναληπτικά, έτσι στο στάδιο της σύλληψης (inception) και κατά τις πρώτες επαναλήψεις (iterations) υλοποιήθηκαν σε ξεχωριστές παράλληλες και αυτόνομες γραμμές:

- I. Η υποδομή αντικειμενοστραφών νευρώνων και αντικειμενοστραφών ΤΝΔ και η υποδομή εκπαίδευσης τους χωρίς νήματα. Δημιουργήθηκαν κλάσεις όπως TGNeuron, TGNeuralNetwork.
- II. Η κλάση αντικειμένων **TGImage** για την επεξεργασία ψηφιδωτών μορφών εικόνας BMP που στηρίζεται στην κλάση TBitmap [4] της Delphi.
- III. Αντικείμενο **TGColor** για μετατροπές χρωματικών συστημάτων RGB σε HSL, και RGB σε κλίμακα του γκρι (Grayscale).
- IV. Οι κλάσεις και οι φόρμες για οπτικοποίηση των νευρωνικών δικτύων, και η φόρμα εκπαίδευσης. Δημιουργήθηκαν οι TGVisualNeuron, TGVisualNeuralLayer και η πρώτη μορφή της φόρμας εκπαίδευσης TfrmTraining.
- V. Η υποδομή επικοινωνίας με την βάση δεδομένων MSSQL μέσω ADO, που στηρίχθηκε κυρίως στις κλάσεις της Delphi TADOConnection και TADOCommand [3].
- VI. Αντικείμενο-συνάρτηση κανονικής (gauss) κατανομής **TGNormalDist** με ιδιότητες όπως απόκλιση TGNormalDist.StdDev και μέσος TGNormalDist.Mean.

Όλα τα υποσυστήματα δοκιμάστηκαν αυτόνομα χωρίς να ενοποιηθούν. Κατόπιν σε επόμενες επαναλήψεις και κυρίως στην φάση της επεξεργασίας, η υλοποίηση των παραπάνω αυτόνομων γραμμών συνεχίστηκε αυξάνοντας την λειτουργικότητα του μοντέλου της υλοποίησης. Έτσι κατ'αντιστοιχία για τα πρώτα τέσσερα υλοποιήθηκαν επιπρόσθετα σε δεύτερο χρόνο.

- I. Πολυνηματική επεξεργασία (multithreading) και χρήση των πυρήνων του επεξεργαστή για την εκπαίδευση των αντικειμενοστραφών ΤΝΔ. Δημιουργήθηκε η TGNeuralNetworkThread και τελειοποιήθηκε η TGNeuralTrainingSet.
- II. Αντικείμενα για την εμφάνιση στην οθόνη και σχεδίαση πάνω σε bitmap με τις κλάσεις **TGViewPort** και **TGDrawing** αντίστοιχα. Εδώ κατασκευάστηκε και η γραφική παράσταση του σφάλματος.

- III. Αντικείμενο **TGPixel**, για την πρόσβαση σε pixel ενός bitmap μέσω scanlines, ώστε να διαβαστούν οι χρωματικές συνιστώσες RGB και να γίνει εμφώλευση (encapsulation) των μετατροπών RGB σε HSL και RGL σε κλίμακα του γκρι (grayscale).
- IV. Βελτιώθηκε η οπτικοποίηση και προσαρμόστηκε ώστε να χρησιμοποιείται μέσα από νήμα. Εδώ δημιουργήθηκε η φόρμα επισκόπησης ενός νευρώνα

Μπαίνοντας μέσα στην φάση της κατασκευής οι διάφορες γραμμές άρχισαν να συνεργάζονται μεταξύ τους, καθώς πλέον η υλοποίηση αφορούσε ολόένα και περισσότερα στοιχεία της αρχιτεκτονικής. Όλα τα αντικείμενα που αφορούσαν υποδομή βελτιωνόταν αντίστοιχα με τους σκοπούς της εκάστοτε επανάληψης, δημιουργώντας σταδιακά τις τελικές και σταθερές εκδόσεις τους. Έτσι στην αρχή της υλοποίησης ο στόχος ήταν η δημιουργία του τμήματος του τεχνητού αμφιβληστροειδούς, και υλοποιήθηκαν επιγραμματικά:

- Τεχνητά οριζόντια κύτταρα HC, ως αντικείμενα της κλάσης **THorizontalCell**
- Πρότυπα εκπαίδευσης τεχνητών δίπολων BPC νευρώνων, εκπαίδευση με αυτά και δημιουργία κλάσης **TBipolarCell** για χρήση του εκπαιδευμένου ΤΝΔ.
- Πρότυπα εκπαίδευσης τεχνητών γαγγλιακών κυττάρων GC, εκπαίδευση και δημιουργία της κλάσης **TGanglionCell**.
- Χάρτης υποδεκτικών πεδίων που είναι στην ουσία και το αντικείμενο **rexel**, που ονομάστηκε **TGRetinalElement**.
- Τεχνητός ρετινοτοπικός χάρτης ARTN που υλοποιεί εξαγωνική ψηφίδωση. Δημιουργήθηκε η κλάση **THoneyComb** που είναι εργοστάσιο (factory) κατασκευής αντικείμενων **TGRetinalElement**, σύμφωνα με τα πρότυπα σχεδίασης (design patterns) [75][76].
- Τεχνητός αμφιβληστροειδής **TRetina** που έχει ως κομμάτια τα αντικείμενα **TRetinaSlice** και ένα συστατικό αντικείμενο **THoneyComb**, που βοηθάει στον τεμαχισμό σε αυτά.
- Προσθήκη στην κλάση **TBipolarCell** της συνάρτησης f_{ac} που αντιστοιχεί με την ύπαρξη ενός τεχνητού βραχύινου κυττάρου AC.
- Οπτικοποίηση για την εμφάνιση των σταδιακών μετατροπών της εικόνας με την βοήθεια των **TGViewPort** και **TGDrawing**.
- Τελειοποίηση της φόρμας εκπαίδευσης των ΤΝΔ
- Αποθήκευση καταστάσεων εκπαιδευμένων ΤΝΔ στην βάση δεδομένων και στον πίνακα **NEURALNET**.

Στην μέση της φάση κατασκευής χτίστηκαν επιπλέον στην υλοποίηση:

- Η φόρμα της λήψης φωτογραφικών του Horasis, ως αυτόνομο τμήμα.
- Η αποθήκευση των φωτογραφιών πραγματικών αντικειμένων στην βάση δεδομένων και στον πίνακα **PHOTO**.
- Η μετατροπή των φωτογραφιών μέσα από την επεξεργασία που γίνεται από την **TRetina**, σε μορφή RNF και η αποθήκευση τους στην βάση δεδομένων και στον πίνακα **RNFPATTERN**
- Η οπτικοποίηση της εξόδου της **TRetina**, σε μια νέα κλάση **TRetinaVisualizer** ώστε να εμφανίζεται η μορφή RNF στην οθόνη
- Μια υλοποίηση αντικειμενοστραφούς ΤΝΔ που χρησιμοποιείται αποκλειστικά για ανάκληση, λειτουργώντας πιο γρήγορα και με μικρότερο κόστος σε μνήμη το **TGFastNeuralNetwork**.

Το κλείσιμο της υλοποίησης της πρώτης έκδοσης του Horasis περιλάμβανε:

- Πρότυπα εκπαίδευσης PV1S τεχνητών μικρών απλών κυττάρων V1, και εκπαίδευση τους για κάθε ένα από τα 6 χρώματα.
- Πρότυπα εκπαίδευσης MV1S τεχνητών μεγάλων απλών κυττάρων V1, και εκπαίδευση τους για κάθε ένα από τα 6 χρώματα.
- Δημιουργία αντικειμένου **TVisualCortexSlice** που υλοποιεί το υποδεκτικό πεδίο μικρού ή μεγάλου τεχνητού V1S. Το κάθε ένα περιέχει ως συστατικά τα 12 TV2Block
- Σύστημα συντεταγμένων για την εξαγωγική ψηφίδωση του THoneyComb που βοηθάει στον καθορισμό των κεντρικών εξαγώνων. Η ονομασία της κλάσης είναι **THoneyCombCoordinateSystem**.
- Δημιουργία τεχνητού οπτικού φλοιού **TVisualCortex** που περιλαμβάνει αντικείμενα **TVisualCortexSlice** αντίστοιχα με τα κεντρικά rexel που καθορίζονται με την βοήθεια του THoneyCombCoordinateSystem και του THoneyComb.
- Πρότυπα εκπαίδευσης του ΤΝΔ ενοποίησης των αποκρίσεων των V1S, εκπαίδευση τους και χρήση τους ως αντικείμενο **TV2Block**.
- Παραγωγή εσωτερικής αναπαράστασης εικόνας και αποθήκευση της στον πίνακα **BRAINWAVE**.
- Χρήση εσωτερικής αναπαράστασης εικόνων ως πρότυπα εκπαίδευσης του ΤΝΔ κατηγοριοποίησης του Horasis, εκπαίδευση του και ανάκληση του μέσα από την φόρμα των προτύπων.
- Εμφάνιση απόκρισης στον χρήση ανάλογα με την έξοδο του Horasis.

Από όλα τα παραπάνω θα αναφερθούμε αναλυτικότερα στον τρόπο που δημιουργήθηκαν τα πρότυπα και εκπαιδεύτηκαν τα ΤΝΔ του Horasis, στο τελικό αποτέλεσμα για την λειτουργία του που είναι η οπτική αναγνώριση των πραγματικών αντικείμενων και σε όλες τις παραμέτρους που την καθορίζουν, μέσα σε αυτήν την πρώτη έκδοση.

Η πλήρης τεκμηρίωση (documentation) όλων των κλάσεων και των αντικειμένων δεν παρατίθεται σε αυτό το σύγγραμμα, καθώς τα αντικείμενα δύναται να οργανωθούν με διαφορετικό τρόπο για την επίτευξη του ιδίου αποτελέσματος. Ο πηγαίος κώδικας της εφαρμογής, όπως ήδη φάνηκε με την ονοματοδοσία κεντρικών κλάσεων, έχει συγγραφεί με αυτοπεριγραφικές (self-documenting) [155] αγγλικές ονομασίες για τα διάφορα αντικείμενα. Επίσης σε όσα σημεία του κώδικα κρίθηκε σκόπιμο έχουν τοποθετηθεί σχόλια, χωρίς όμως την υπερβολική τους χρήση, καθώς οι ίδιες οι ονομασίες των μεθόδων και ιδιοτήτων είναι σχόλια από μόνες τους.

Τα Τεχνητά Νευρωνικά Δίκτυα του Horasis και η Εκπαίδευσή τους

Το αντικείμενο διαχείρισης των διαφορετικών τύπων τεχνητών νευρωνικών δικτύων.

Έχουμε δει στον σχεδιασμό, πολλά διαφορετικά ΤΝΔ να συνεργάζονται μέσα στο Horasis. Στο Κεφάλαιο 5 όταν γνωρίσαμε την φόρμα εκπαίδευσης αναφέραμε την ύπαρξη διαφορετικών τύπων ΤΝΔ. Πλέον μπορούμε να δούμε τους διαφορετικούς τύπους ΤΝΔ που χρησιμοποιήθηκαν στην πρώτη έκδοση της εφαρμογής, σε αντιστοιχία με τον σχεδιασμό της εφαρμογής. Για αυτούς κατασκευάστηκε στον κώδικα ένας τύπος απαρίθμησης (enumeration type) με το όνομα `TNeuralNetKind` και όλοι οι τύποι έχουν το πρόθεμα `nnk` από τα αρχικά των λέξεων `neural network kind`:

- Για τεχνητά BPC χρησιμοποιούνται δύο διαφορετικοί τύποι ΤΝΔ:
 - **nnkBipolarsOfCones** για το ΤΝΔ των τεχνητών κωνίων που έχει 3 εισόδους και 4 εξόδους και εμπεριέχει στο στρώμα εξόδου τα:
 - R-BPC
 - G-BPC
 - B-BPC
 - Y-BPC
 - **nnkBipolarsOfRod** για το ΤΝΔ του τεχνητού ραβδίου που έχει 1 είσοδο και 2 εξόδους και εμπεριέχει στο στρώμα εξόδου τα:
 - L-BPC
 - D-BPC
- Για τα τεχνητά GC χρησιμοποιείται σε κάθε περίπτωση ο ίδιος τύπος ΤΝΔ:
 - **nnkGanglionBichromatic** για τα έξι διαφορετικά ΤΝΔ των ανταγωνιστικών χρωμάτων που έχουν 14 εισόδους και 4 εξόδους και αντιστοιχούν με τα:
 - R-GC
 - G-GC
 - B-GC
 - Y-GC
 - L-GC
 - D-GC
- Για τα τεχνητά PV1S έχουμε έναν τύπο ΤΝΔ ανά χρώμα, ενώ σε όλα τα δίκτυα υπάρχουν 56 είσοδοι και 6 εξόδοι
 - **nnkV1SimpleCellsParviRed** για ανίχνευση κόκκινων ράβδων σε πράσινο φόντο
 - R-PV1S
 - **nnkV1SimpleCellsParviGreen** για ανίχνευση πράσινων ράβδων σε κόκκινο φόντο
 - G-PV1S
 - **nnkV1SimpleCellsParviBlue** για μπλε / κίτρινο
 - B-PV1S
 - **nnkV1SimpleCellsParviYellow** για κίτρινο / μπλε
 - Y-PV1S

- **nnkV1SimpleCellsParviLight** για λευκό / μαύρο
 - L-PV1S
- **nnkV1SimpleCellsParviDark** για μαύρο / λευκό
 - D-PV1S
- Για τα τεχνητά MV1S έχουμε έναν τύπο ΤΝΔ ανά χρώμα, ενώ σε όλα τα δίκτυα υπάρχουν 152 είσοδοι και 6 έξοδοι
 - **nnkV1SimpleCellsMagnaRed** για κόκκινο / πράσινο
 - R-MV1S
 - **nnkV1SimpleCellsMagnaGreen** για πράσινο / κόκκινο
 - G-MV1S
 - **nnkV1SimpleCellsMagnaBlue** για μπλε / κίτρινο
 - B-MV1S
 - **nnkV1SimpleCellsMagnaYellow** για κίτρινο / μπλε
 - Y-MV1S
 - **nnkV1SimpleCellsMagnaLight** για λευκό / μαύρο
 - L-MV1S
 - **nnkV1SimpleCellsMagnaDark** για μαύρο / λευκό
 - D-MV1S
- Για το ΤΝΔ ενοποίησης των V1S ,το οποίο θεωρείται ως τμήμα του οπτικού φλοιού V2, υπάρχει ένας τύπος για όλες τις περιπτώσεις:
 - **nnkV2CellPatterns** για ανίχνευση ράβδων διαφορετικού μηκών στους τρεις χρωματικούς ανταγωνισμούς:
 - RG-V1S
 - BY-V1S
 - LD-V1S
- Για το ΤΝΔ εξόδου του Horasis υπάρχει ένας αντίστοιχος τύπος:
 - **nnkHorasis**

Μόνο αυτοί οι τύποι χρησιμοποιήθηκαν στην τελική μορφή της εφαρμογής και γι αυτούς θα παρατεθούν περαιτέρω λεπτομέρειες.

Στο κέντρο της υλοποίησης του Horasis βρίσκεται ένα αντικείμενο **TNeuralNetManager**, χαρακτηριζόμενο ως singleton σύμφωνα με τα πρότυπα σχεδίασης (design patterns) [75][76], με το οποίο γίνεται η διαχείριση των διαφορετικών τύπων ΤΝΔ του Horasis. Ο TNeuralNetManager δίνοντας του τύπο ΤΝΔ μπορεί να επιστρέψει είτε ένα πλήρες αντικειμενοστραφές ΤΝΔ της κλάσης TGNesuralNetwork για εκπαίδευση ή ένα TGFastNeuralNetwork μόνο για ανάκληση. Λειτουργεί με το πρότυπο σχεδίασης lazy initialization [75] έτσι όταν χρειαστεί για πρώτη φορά το TGNesuralNetwork ενός τύπου, δημιουργείται ως αντικείμενο και φορτώνεται εκείνη την στιγμή με την αποθηκευμένη κατάσταση που έχει στην βάση δεδομένων. Πλέον μετά την φόρτωση του πλήρους ΤΝΔ μπορούμε να παράγουμε όσα νέα TGFastNeuralNetwork έχοντας το ως πρωτότυπο για την αντιγραφή της δομής του και των συναπτικών βαρών του.

Μια δεύτερη ευθύνη του TNeuralNetManager είναι να κρατάει για κάθε τύπο ΤΝΔ μια στοιβή που περιέχει τα σύνολα προτύπων εκπαίδευσης. Αυτή υλοποιείται με την κλάση **TNeuralTrainingSetStack** και έχει δύο μεθόδους Push και Pop. Με Push γίνεται προσθήκη ενός

δημιουργημένου αντικειμένου προτύπων `TGNeuralTrainingSet` στην στοίβα. Με `Pop` αφαιρείται ώστε να καταναλωθεί μέσα σε μια διαδικασία παράλληλης εκπαίδευσης. Έτσι μπορούμε να έχουμε πολλαπλά αντίγραφα του ίδιου συνόλου στην στοίβα ώστε να τρέξουν με τα ίδια δεδομένα έως και 3 παράλληλες εκπαιδεύσεις στον τετραπύρρηνο επεξεργαστή. Παρακάτω θα δούμε με ποιον τρόπο γεμίζουν αυτές οι στοίβες.

Οργάνωση των προτύπων σε συλλογές.

Όλα τα πρότυπα εκπαίδευσης για τους διαφορετικούς τύπους ΤΝΔ έχουν κοινά γνωρίσματα ως προς την οργάνωση τους σε σύνολα και την χρήση τους στην εφαρμογή μέσω της στοίβας. Για την δημιουργία τους χρησιμοποιείται μια αντίστοιχη ιεραρχία κλάσεων με προγονικές κλάσεις τις **TPattern** που αντιστοιχεί ένα προς ένα με το πρότυπο `TGNeuralPattern` και **TPatternCollection** που είναι μια συλλογή από πρότυπα και αντιστοιχεί στο σύνολο προτύπων `TGNeuralTrainingSet`. Οι `TPattern` και `TPatternCollection` υλοποιούν την λογική με την οποία δημιουργούνται τα πρότυπα ενώ οι `TGNeuralPattern` και `TGNeuralTrainingSet` περιέχουν απλώς τα δεδομένα. Για τις μεν πρώτες υπάρχει ένα αντίγραφο, ενώ για τις μεν δεύτερες μπορεί να υπάρχουν πολλά διαφορετικά αντίγραφα τα οποία είναι αυτά που τοποθετούνται στην αντίστοιχη στοίβα του τύπου με `Push`.

Έχοντας απογόνους των δύο κλάσεων `TPattern` και `TPatternCollection`, υλοποιείται μέσα σε αυτούς η διαδικασία παραγωγής των προτύπων. Η διαφορετική υλοποίηση ανάλογα με τον κάθε τύπο ΤΝΔ, γίνεται επικαλύπτοντας (`override`) την μέθοδο `TPatternCollection.Generate()` στην οποία δημιουργούνται τα αντικείμενα του κάθε προτύπου. Στην `Generate` υπάρχουν βρόχοι που δημιουργούν τους συνδυασμούς των διαφορετικών προτύπων. Σε κάθε βήμα ελέγχεται αν ο συγκεκριμένος συνδυασμός υπάρχει ήδη ως δημιουργημένο αντικείμενο `TPattern` και αν δεν υπάρχει τότε μόνο δημιουργείται. Τα χαρακτηριστικά γνωρίσματα του συνδυασμού τίθενται ως ιδιότητες στο νέο αντικείμενο. Για παράδειγμα χαρακτηριστικά γνωρίσματα ενός συνδυασμού σε ένα πρότυπο εκπαίδευσης για το ΤΝΔ τύπου `nhkBipolarsOfCones`, είναι οι τρεις διαφορετικές τιμές των `R,G,B` όπως θα δούμε παρακάτω.

Στο επίπεδο του προτύπου επικαλύπτεται η μέθοδος `TPattern.AddToSet` που καλείται όταν έρθει η στιγμή αντιγραφής των δεδομένων από την συλλογή σε ένα νέο `TGNeuralTrainingSet`. Εκεί μέσα καθορίζεται ανάλογα με τις ιδιότητες του αντικειμένου, που αντιστοιχούν σε χαρακτηριστικά του συνδυασμού, οι τιμές των εισόδων και των αντίστοιχων στόχων που είναι πραγματικοί αριθμοί από 0 έως 1. Έτσι γίνονται απαιτούμενη κανονικοποίηση ή προβολή τιμών ώστε τα δεδομένα να είναι κατάλληλα για την εκπαίδευση ενός αντικειμενοστραφούς `MLP` με σιγμοειδείς συναρτήσεις.

Βλέπουμε δηλαδή ότι για την συγγραφή της συνδυαστικής λογικής που παράγει το σύνολο προτύπων έχουμε κατανοητές τιμές ιδιοτήτων. Μέσα στο ίδιο το αντικείμενο του προτύπου υλοποιείται εμφωλευμένη και αποκρυφθείσα η αναπαράσταση των ιδιοτήτων ως τιμές εισόδου και εξόδου στο νευρωνικό δίκτυο. Ακόμα η συλλογή δεν ενδιαφέρεται για αυτήν την μετατροπή, καθώς το μόνο που πρέπει να κάνει είναι να παράγει όλους τους συνδυασμούς μέσω της `Generate`. Ο σχεδιασμός γίνεται έτσι αρκετά εύελικτος, έχοντας την μικρή σύζευξη και υψηλή συνοχή που ευαγγελίζεται μέσα στην εργασία, ώστε αν θελήσουμε να αλλάξουμε τον τρόπο που παράγονται τα πρότυπα δεν χρειάζεται να ασχοληθούμε με πως το κάθε πρότυπο αναπαρίσταται

με δεκαδικούς αριθμούς. Συν τοις άλλοις υπάρχει και πολύ καθαρή συγγραφή του πηγαίου κώδικα με πολύ μικρές και ευανάγνωστες μεθόδους.

Για να παραχθούν οι συλλογές προτύπων υπάρχουν αντίστοιχες φόρμες που συνδέονται με έναν ή παραπάνω τύπους ΤΝΔ. Ανοίγοντας και χρησιμοποιώντας μια από αυτές, τροφοδοτείται με πρότυπα εκπαίδευσης μια στοίβα προτύπων εκπαίδευσης. Κατόπιν ανοίγουμε την φόρμα εκπαίδευσης που καταναλώνει τα πρότυπα από την στοίβα, κλείνοντας έτσι τον κύκλο της διαδικασίας εκπαίδευσης. Η φόρμα ως πρωτότυπο διασύνδεσης χρήστη (UI prototype) [156] περιλαμβάνει ένα πλήκτρο για την εκτέλεση της παραγωγής του συνόλου προτύπων, έναν οπτικό χειρισμό για μετακίνηση μέσα στα διαφορετικά πρότυπα του συνόλου και οπτικά συστατικά που εμφανίζουν τις τιμές ή οπτικοποιούν τα πρότυπα.

Οι τιμές που φαίνονται στις φόρμες αφορούν τις εισόδους και τους στόχους και εκτός από τα κατανοητά χαρακτηριστικά του προτύπου εμφανίζονται και οι αντίστοιχοι πραγματικοί αριθμοί που θα χρησιμοποιηθούν. Μέσα στην ίδια φόρμα χρησιμοποιείται το αντίστοιχο εκπαιδευμένο ΤΝΔ από τον TNeuralNetManager ώστε να εκτελέσει ανάκληση πάνω σε ένα συγκεκριμένο πρότυπο της συλλογής. Έτσι παράλληλα με τις εισόδους και τους στόχους, εμφανίζεται και η αντίστοιχη έξοδος για το πρότυπο, δίνοντας την δυνατότητα απευθείας σύγκριση των στόχων με την τελική έξοδο του εκπαιδευμένου ΤΝΔ.

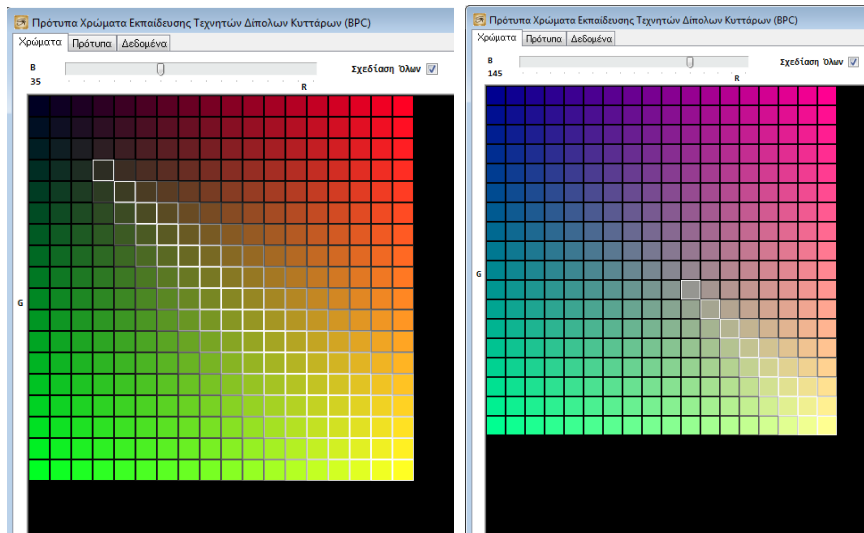
Θα εξετάσουμε τέσσερις φόρμες που περικλείουν την δημιουργία προτύπων εκπαίδευσης και την δοκιμή λειτουργίας των ΤΝΔ με ανάκληση, τα οποία εμπλέκονται στην δημιουργία της εσωτερικής αναπαράστασης του πραγματικού αντικειμένου:

1. Πρότυπα Χρώματα Εκπαίδευσης Τεχνητών Δίπολων Κυττάρων (BPC)
2. Πρότυπα Εκπαίδευσης Τεχνητών Γαγγλιακών Κυττάρων (GC)
3. Πρότυπα Εκπαίδευσης Απλών Κυττάρων V1 (PV1S, MV1S)
4. Πρότυπα Εκπαίδευσης V2

Για κάθε μια από αυτές θα περιγραφούν τα χαρακτηριστικά και η συνδυαστική διαδικασία παραγωγής των προτύπων.

Φόρμα προτύπων εκπαίδευσης τεχνητών δίπολων κυττάρων

Στην φόρμα με τον τίτλο «Πρότυπα Εκπαίδευσης Τεχνητών Δίπολων Κυττάρων (BPC)» παράγονται τα αντίστοιχα πρότυπα για την εκπαίδευση των ΤΝΔ των τύπων `nnkBipolarsOfCones` και `nnkBipolarsOfRod`. Στην φόρμα αρχικά εμφανίζεται μια δειγματοληπτική απεικόνιση των χρωμάτων μιας συγκεκριμένη τιμή του μπλε Β στην οποία σημειώνονται βάσει της απόχρωσης Η, τα χρώμα που είναι πλησιέστερα στο κίτρινο. Υπάρχει η δυνατότητα αυξομείωσης της τιμής του μπλε με τα δείγματα των χρωμάτων να τροποποιούνται αντίστοιχα. Η οθόνη αυτή χρησιμοποιήθηκε πριν δημιουργηθούν τα πρότυπα εκπαίδευσης των τεχνητών BPC, για να οπτικοποιηθεί ο χρωματικός ανταγωνισμός μεταξύ μπλε και κίτρινου που μας ενδιαφέρει καθώς μετά την εκπαίδευση θα αποκτούσαμε ακόμα ένα χρώμα το κίτρινο, που δεν υπάρχει στις τρεις χρωματικές συνιστώσες του RGB.



10-1: Εμφάνιση δείγματος χρωμάτων συγκεκριμένης τιμής του μπλε, με τα χρώματα κοντά στο κίτρινο να έχουν γκρι ή λευκό περίγραμμα. Όσο πλησιάζει προς το λευκό το περίγραμμα τόσο πιο κοντά βρίσκεται το χρώμα στην απόχρωση του αμιγούς κίτρινου.

Με την χρήση του πλήκτρου «Παραγωγή» δημιουργούνται δύο συλλογές προτύπων και προστίθενται στις αντίστοιχες στοίβες των τύπων `nhkBipolarsOfCones` και `nhkBipolarsOfRod`. Το κάθε αντικείμενο προτύπου έχει τρεις χαρακτηριστικές ιδιότητες ανάλογα με τις τρεις χρωματικές συνιστώσες του συστήματος RGB:

- **Red**, κόκκινο 0..255
- **Green**, πράσινο 0..255
- **Blue**, μπλε 0..255

Οι συνδυασμοί τιμών για αυτές παράγονται στην `Generate` και στον βρόχο διασφαλίζεται ότι δεν θα υπάρχουν δύο πρότυπα με τα ίδια χαρακτηριστικά R,G,B:

Συνδυασμοί προτύπων εκπαίδευσης τεχνητών BPC

Για κάθε **BI** από 0 έως 16 υπολόγισε το μπλε $B = BI^2$
 Για κάθε **RI** από 0 έως 17 υπολόγισε το κόκκινο $R = RI * 15$
 Για κάθε **GI** από 0 έως 17 υπολόγισε το πράσινο $G = GI * 15$

Από τις συναρτήσεις παραγωγής στόχων που αναφέρθηκαν στο Κεφάλαιο 9, σε κάθε πρότυπο υπολογίζονται οι 3 στόχοι για το ΤΝΔ του τύπου `nhkBipolarsOfCones`

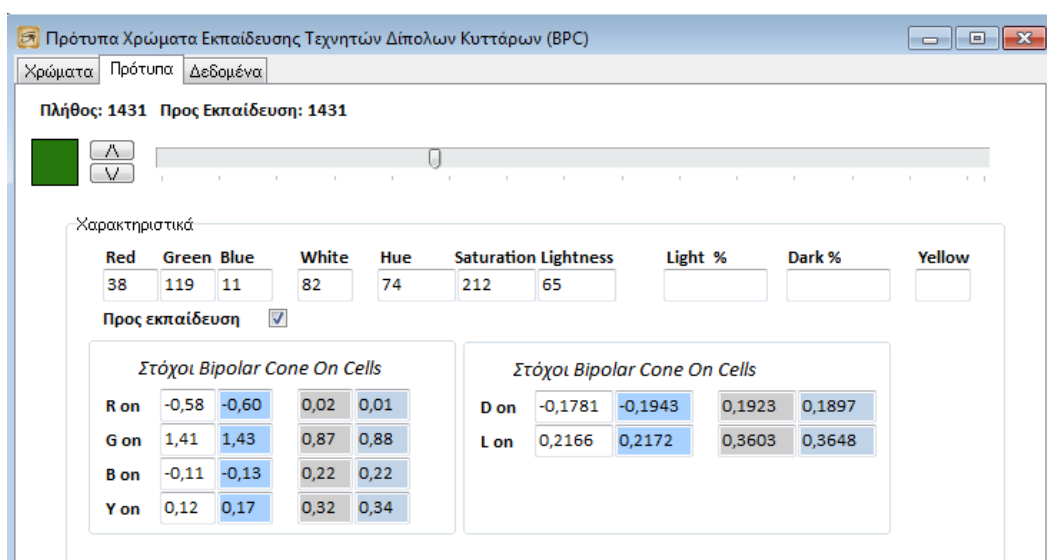
- **TargetRed**, με πιθανές τιμές από -0.63 έως 1.72
- **TargetGreen**, με πιθανές τιμές από -0.63 έως 1.72
- **TargetBlue**, με πιθανές τιμές από -0.63 έως 1.72

Αντίστοιχα στο ΤΝΔ του τύπου `nhkBipolarsOfRod` υπάρχουν οι ιδιότητες για τους 2 στόχους:

- **TargetLight**, με πιθανές τιμές από -0.63 έως 1.72
- **TargetDark**, με πιθανές τιμές από -0.63 έως 1.72

Εσωτερικά υπολογίζεται η τιμή του φωτός (luma) από τα R,G,B και αυτή τοποθετείται ως η μοναδική τιμή εισόδου για το πρότυπο τύπου nhkBipolarsOfRod. Όλες οι τιμές χαρακτηριστικών και στόχων προβάλλονται στο $[0,1] \in \mathbb{R}$.

Το πλήθος των προτύπων για σύνολα των τύπων nhkBipolarsOfCones και nhkBipolarsOfRod είναι 1431. Στην φόρμα εμφανίζονται οι ιδιότητες για τα πρότυπα σε πλαίσια κειμένου με άσπρο φόντο και οι αντίστοιχες πραγματικές τιμές των εισόδων από 0 έως 1 σε γκρι φόντο. Με έντονο γαλάζιο είναι η έξοδος από ανάκληση του εκπαιδευμένου ΤΝΔ μεταφρασμένη στην αντίστοιχη ιδιότητα. Με θαμπό μπλε είναι η έξοδος όπως παράγεται από την ανάκληση μεταξύ 0 και 1.



10-2: Φόρμα που δημιουργεί, εμφανίζει και ανακαλεί ένα πρότυπο εκπαίδευσης των τεχνητών BPC

Φόρμα προτύπων εκπαίδευσης τεχνητών γαγγλιακών κυττάρων

Στην φόρμα με τον τίτλο «Πρότυπα Εκπαίδευσης Τεχνητών Γαγγλιακών Κυττάρων (GC)» παράγονται τα αντίστοιχα πρότυπα για την εκπαίδευση του ΤΝΔ για τον τύπο nhkGanglionBichromatic. Με την χρήση του πλήκτρου «Παραγωγή» δημιουργείται μια συλλογή προτύπων και προστίθεται στην στοιβή του τύπου nhkGanglionBichromatic. Το κάθε αντικείμενο προτύπου έχει τέσσερις χαρακτηριστικές ιδιότητες

- **CenterValue**, τιμή φωτισμού βασικού χρώματος στο κέντρο από -0,63 έως 1,72
- **PeripheryValue**, τιμή φωτισμού ανταγωνιστικού χρώματος στην περιφέρεια από -0,63 έως 1,72
- **IlluminationType**, το είδος φωτισμού που πέφτει πάνω στο υποδεκτικό πεδίο από τέσσερα διαφορετικά είδη:
 - gitCenterSpot, μια κηλίδα στο κέντρο και το ανταγωνιστικό χρώμα στην περιφέρεια
 - gitDiffuse, διάχυτος φωτισμός
 - gitHalf, μισός φωτισμός με βασικό και μισός με ανταγωνιστικό χρώμα
 - gitBar, μια μπάρα βασικού χρώματος που περνάει από το κέντρο, φωτίζοντας παράλληλα και τμήμα της περιφέρειας, με φόντο το ανταγωνιστικό χρώμα.
- **LightTransHour**, ώρα (διεύθυνση) μετάβασης φωτισμού από το ένα χρώμα στο άλλο με 0 να ισοδυναμεί με απροσδιοριστία, ενώ 1 έως 12 για την αντίστοιχη ώρα.

Οι συνδυασμοί τιμών για αυτές παράγονται στην Generate και στον βρόχο διασφαλίζεται ότι δεν θα υπάρχουν δύο πρότυπα με τις ίδιες τιμές χαρακτηριστικών ιδιοτήτων.

Συνδυασμοί προτύπων εκπαίδευσης τεχνητών GC

Καθορίζονται 14 δείγματα τιμών από -0.632 έως 1.718 = $\{-0.632, -0.485, -0.338, -0.191, -0.0445, 0, 0.102, 0.249, 0.543, 0.836, 1.131, 1.278, 1.571, 1.718\}$

Για κάθε **C** σε 14 διαφορετικά δείγματα υπολόγισε **CenterValue = C**

Για κάθε **P** σε 14 διαφορετικά δείγματα υπολόγισε **PeripheryValue = P**

Αν **C <> P** τότε

Θέσε **LightTransHour = 0** και **IlluminationType = gitCenterSpot**

Για κάθε διαφορετικό **L** από 1..12 θέσε

LightTransHour = L και **IlluminationType = gitHalf**

LightTransHour = L και **IlluminationType = gitBar**

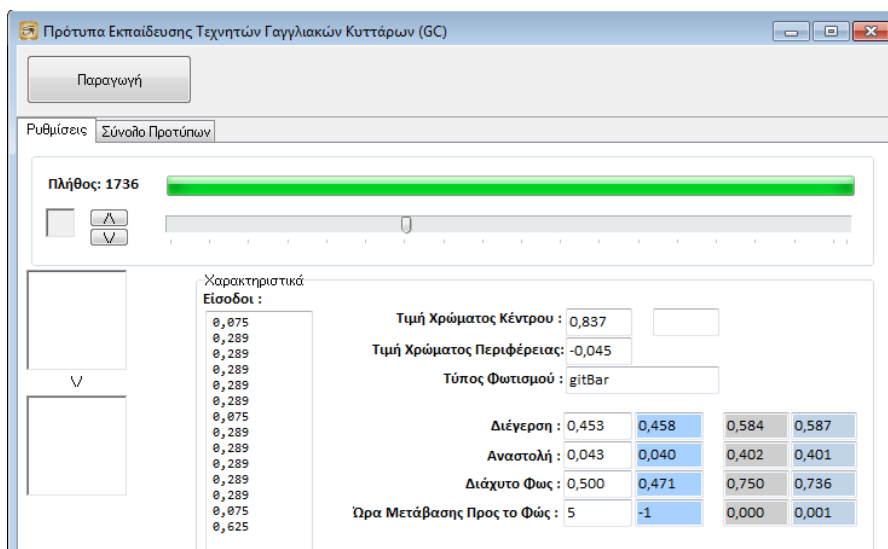
ειδάλλως αν $\text{Abs}(C - P) = 3$ τότε θέσε

LightTransHour = 0 και **IlluminationType = gitDiffuse**

Σε κάθε πρότυπο υπολογίζονται οι 4 στόχοι για το ΤΝΔ του τύπου nnkGanglionBichromatic με την λογική που περιγράφηκε στο κεφάλαιο 9.

- **Excitation**, με πιθανές τιμές από -0.86 έως 1.386
- **Inhibition**, με πιθανές τιμές από -0.86 έως 1.386
- **Diffuse**, με πιθανές τιμές από -1 έως 1
- **LightTrans**, με πιθανές ακέραιες τιμές από 0 έως 12

Όλες οι τιμές των ιδιοτήτων χαρακτηριστικών και στόχων, προβάλλονται στο $[0,1] \in \mathbb{R}$ κατά την αντιγραφή τους σε αντικείμενα TGNeuralPattern. Το πλήθος των προτύπων που παράγονται για την συλλογή τύπου nnkGanglionBichromatic είναι 1736. Στην φόρμα εμφανίζεται μια λίστα με τις τιμές εισόδου του προτύπου κανονικοποιημένες από 0 έως 1 καθώς και οι υπόλοιπες ιδιότητες για εισόδους αλλά και για στόχους. Δίπλα στους στόχους εμφανίζονται με τα γνωστά γκρι πλαίσια οι κανονικοποιημένοι στόχοι και με τα γαλάζια τα αποτελέσματα της ανάκλησης.



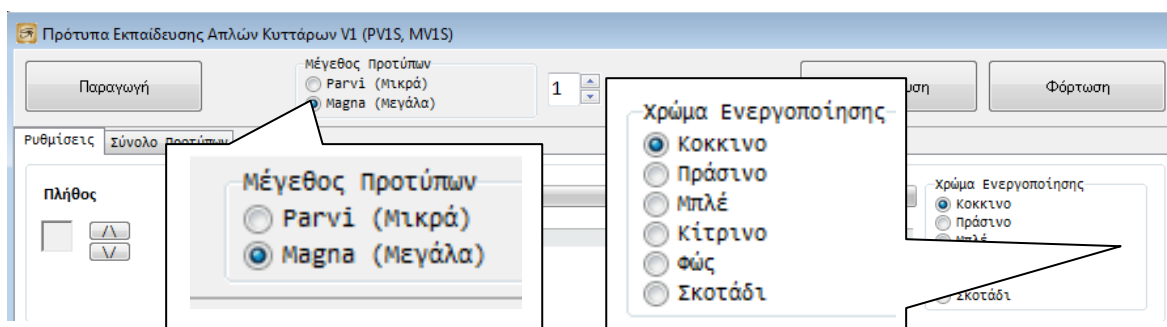
10-3: Φόρμα που δημιουργεί, εμφανίζει και ανακαλεί ένα πρότυπο εκπαίδευσης των τεχνητών GC

Φόρμα προτύπων εκπαίδευσης τεχνητών απλών κυττάρων V1

Η φόρμα «Πρότυπα Εκπαίδευσης Απλών Κυττάρων V1 (PV1S, MV1S)» διαφέρει από τις προηγούμενες δύο καθώς μέσω αυτής μπορούν να παραχθούν πρότυπα για 12 διαφορετικούς τύπους ΤΝΔ του Horasis:

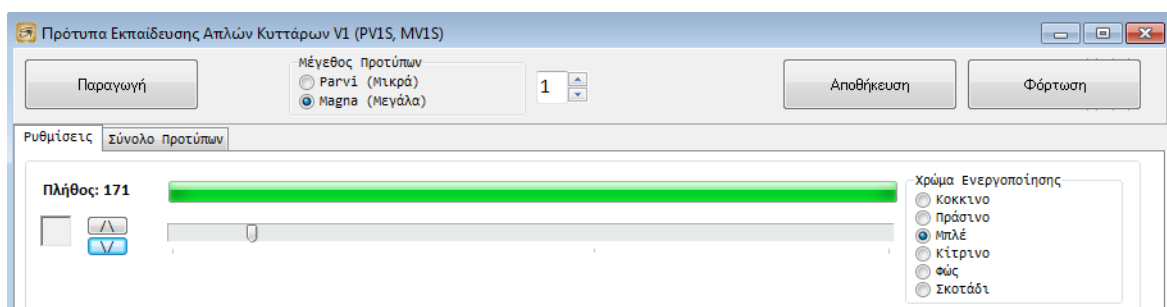
- nnkV1SimpleCellsParviRed
- nnkV1SimpleCellsParviGreen
- nnkV1SimpleCellsParviBlue
- nnkV1SimpleCellsParviYellow
- nnkV1SimpleCellsParviLight
- nnkV1SimpleCellsParviDark
- nnkV1SimpleCellsMagnaRed
- nnkV1SimpleCellsMagnaGreen
- nnkV1SimpleCellsMagnaBlue
- nnkV1SimpleCellsMagnaYellow
- nnkV1SimpleCellsMagnaLight
- nnkV1SimpleCellsMagnaDark

Ο τύπος καθορίζεται από την επιλογή κουμπιών ραδιοφώνου (radio button) «Μέγεθος Προτύπων» σε συνδυασμό με την επιλογή «Χρώμα Ενεργοποίησης».



10-4: Καθορισμός τύπου τεχνητών V1S ως συνδυασμός δύο επιλογών

Για παράδειγμα επιλέγοντας «Magna (Μεγάλα)» και «Μπλε» με το πάτημα του πλήκτρου «Παραγωγή», έχουμε την δημιουργία μια συλλογής προτύπων εκπαίδευσης του τύπου nnkV1SimpleCellsMagnaBlue. Με την τροποποίηση του αριθμού που υπάρχει δίπλα στην επιλογή μεγέθους θα έχουμε την δημιουργία και προσθήκη στην στοίβα αντίστοιχου πλήθους από αντίγραφα TGNeuralTrainingSet του συνόλου προτύπων που παρήχθηκε.



10-5: Φόρμα που δημιουργεί, εμφανίζει και ανακαλεί πρότυπα 12 διαφορετικών τύπων τεχνητών V1S

Η οπτικοποίηση του ενός προτύπου παρουσιάζει ιδιαίτερο ενδιαφέρον καθώς εκτός από κείμενο στα γνωστά πλαίσια και τις λίστες για τις εισόδους, εμφανίζονται παράθυρα σχεδίασης γραφικών. Στο αριστερό παράθυρο εμφανίζεται η έξοδος των τεχνητών ΗC μετά από την σχεδίαση μιας οριζόντιας μπάρας μπλε πάνω σε κίτρινο φόντο. Κατόπιν μέσα στο περίγραμμα (frame) «Χαρακτηριστικά» υπάρχουν έξι παράθυρα που εμφανίζουν την απόκριση τεχνητών GC των έξι ανταγωνιστικών χρωμάτων της περιοχής που σχεδιάστηκε η μπάρα. Το υποδεκτικό πεδίο που εμφανίζεται εξαρτάται από την περίπτωση και στο παράδειγμα μας του τύπου nnkV1SimpleCellsMagnaBlue υπάρχουν 19 rhexel επειδή έχουμε μεγάλα (magna) κύτταρα. Επίσης τα τεχνητά GC που θα χρησιμοποιηθούν είναι μόνο αυτά που λειτουργούν για τον χρωματικό ανταγωνισμό μπλε με κίτρινο, οπότε τα αντίστοιχα παράθυρα είναι και αυτά που μας ενδιαφέρουν.

Χαρακτηριστικά

Είσοδοι :

Κόκκινο - Πράσινο

Μπλέ - Κίτρινο

Φως - Σκοτάδι

Χρώματα :

Γωνία Μπάρας : 0

Πάχος Μπάρας : 24

Είσοδοι : Γάγγλια: 19 Πλήθος Εισόδων: 152

Στόχοι :	Εξοδος :
1,000	0,975
0,000	0,000
0,000	0,000
0,000	0,000
0,000	0,000
0,000	0,000
0,000	0,000

10-6: Οπτικοποίηση ενός προτύπου εκπαίδευσης τεχνητού B-MV1S, με τύπο ΤΝΔ nnkV1SimpleCellsMagnaBlue. Ο συνδυασμός χρωμάτων είναι αμιγές μπλε και αμιγές κίτρινο, η γωνία 0° και το πάχος της μπάρας 24 pixels. Βλέπουμε 19 rhexel που θα δώσουν τιμές σε 152 εισόδους.

Στην φόρμα επίσης παρατηρούμε να αναγράφεται η γωνία με την οποία σχεδιάστηκε η μπάρα πάνω στον καμβά και το πάχος της μπάρας σε pixels. Στην λίστα των στόχων βλέπουμε να υπάρχει 1 στην αντίστοιχη θέση των μοιρών της γωνίας της μπάρας, έχοντας 6 θέσεις για τον διαχωρισμό όλων των κατευθύνσεων ανά 30° . Στις υπόλοιπες θέσεις τοποθετείται το 0. Σε πρότυπα με διάχυτο φωτισμό κίτρινου και μπλε έχουμε σε όλους τους στόχους 0 και σε όλους τους στόχους 1 αντίστοιχα:

Χαρακτηριστικά

Είσοδοι :

Κόκκινο - Πράσινο

Μπλέ - Κίτρινο

Χρώματα :

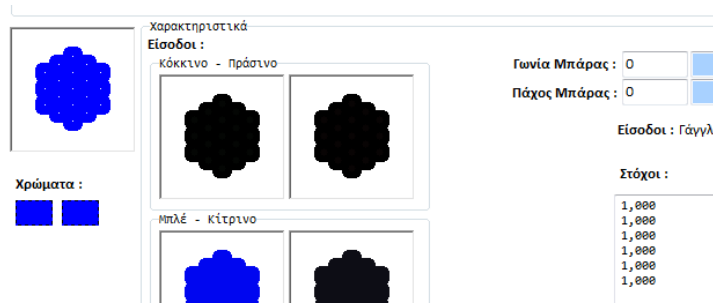
Γωνία Μπάρας : 0

Πάχος Μπάρας : 0

Είσοδοι : Γάγγλια

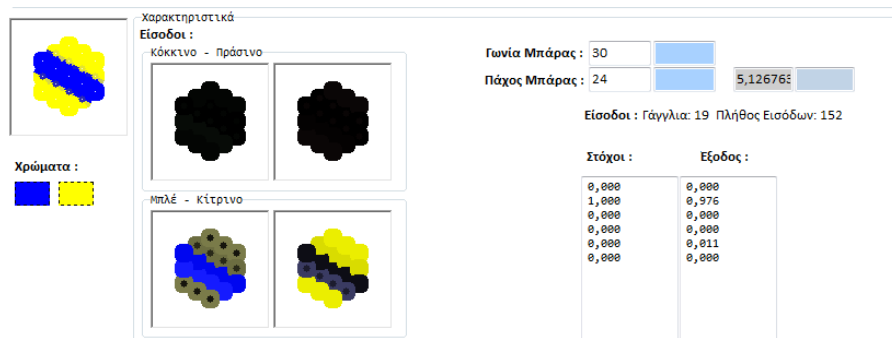
Στόχοι :	Εξοδος :
0,000	0,000
0,000	0,000
0,000	0,000
0,000	0,000
0,000	0,000
0,000	0,000
0,000	0,000

10-7: Οπτικοποίηση ενός προτύπου εκπαίδευσης τεχνητού B-MV1S με διάχυτο φωτισμό ανταγωνιστικού χρώματος.

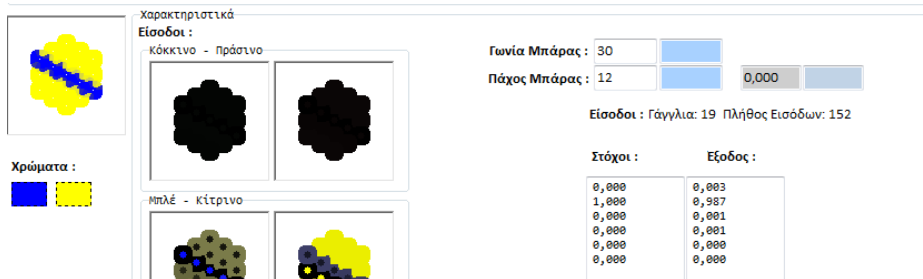


10-8: Οπτικοποίηση ενός προτύπου εκπαίδευσης τεχνητού B-MV1S με διάχυτο φωτισμό βασικού χρώματος.

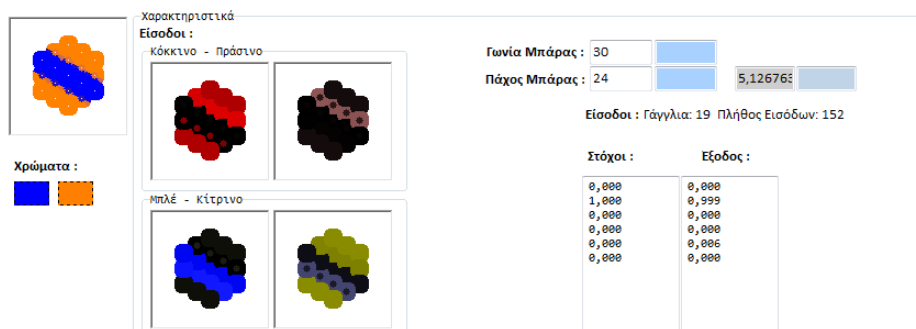
Επίσης αν διατρέξουμε τα πρότυπα βλέπουμε τις διαφορετικές γωνίες στον ίδιο συνδυασμό χρωμάτων, τα διαφορετικά πάχη στην ίδια γωνία του ίδιου συνδυασμού χρωμάτων, αλλά και εντελώς διαφορετικούς συνδυασμούς χρωμάτων.



10-9: Οπτικοποίηση προτύπου για B-MV1S με γωνία 30° και πάχος μπάρας 24 pixels, αμιγές μπλε πάνω σε αμιγές κίτρινο



10-10: Οπτικοποίηση προτύπου για B-MV1S με γωνία 30° και πάχος μπάρας 12 pixels, αμιγές μπλε πάνω σε αμιγές κίτρινο



10-11: Οπτικοποίηση προτύπου για B-MV1S με γωνία 30° και πάχος μπάρας 24 pixels, αμιγές μπλε πάνω σε πορτοκαλί

Έτσι οι χαρακτηριστικές ιδιότητες για ένα πρότυπο που καθορίζουν και τον τρόπο που θα σχεδιαστεί είναι:

- **Kind**, το είδος της σχεδίασης από πέντε διαφορετικά είδη,
 - v1sBar, μπάρα βασικού χρώματος πάνω σε φόντο ανταγωνιστικού χρώματος
 - v1sNonBar, μπάρα ανταγωνιστικού χρώματος πάνω σε φόντο βασικού χρώματος
 - v1sDiffuseStrong, διάχυτο βασικό χρώμα με ισχυρή διέγερση
 - v1sDiffuseWeak, διάχυτο βασικό χρώμα με ασθενή διέγερση
 - v1sNonDiffuse, διάχυτο ανταγωνιστικό χρώμα
- **BarColor**, το χρώμα με το οποίο σχεδιάζεται η μπάρα καθοριζόμενο από {R,G,B},
- **BackgroundColor**, το χρώμα του φόντου σχεδίασης καθοριζόμενο με {R,G,B},
- **BarAngle**, η γωνία με την οποία σχεδιάζεται η μπάρα σε μοίρες,
- **BarThicknessInPixels**, το πάχος σε pixels με το οποίο σχεδιάζεται η μπάρα.

Οι συνδυασμοί αυτών των χαρακτηριστικών παράγονται στην Generate και όπως πάντα διασφαλίζεται ότι δεν θα υπάρχουν δύο πρότυπα με τις ίδιες τιμές χαρακτηριστικών ιδιοτήτων. Επίσης κατά το βήμα του βρόχου εκτελείται σχεδίαση πάνω σε έναν κοινόχρηστο καμβά bitmap, τον οποίο «βλέπει» έναν τεχνητός αμφιβληστροειδής. Μετά την σχεδίαση εκτελείται ανάκληση των pixels του καμβά μέσα στον τεχνητό αμφιβληστροειδή ώστε να παραχθούν οι τιμές του διανύσματος GCE για όλα τα τεχνητά GC, τις οποίες είδαμε παραπάνω να οπτικοποιούνται στα έξι παράθρα. Οι συνδυασμοί παράγονται ως εξής:

Συνδυασμοί προτύπων εκπαίδευσης τεχνητών V1S

Για κάθε ζεύγος χρωμάτων **BarColor**, **BackgroundColor** από 8 διαφορετικούς συνδυασμούς
 Για κάθε **WI** σε 3 διαφορετικά πάχη θέσε **BarThicknessInPixels = Πάχος[WI]**
 Για κάθε **AI** από το 0 έως το 6 υπολόγισε **BarAngle = AI * 30**
 Για το βασικό χρώμα θέσε **Kind = v1sBar**
 Για το ανταγωνιστικό χρώμα θέσε **Kind = v1sNonBar**

Πρόσθεσε για το βασικό χρώμα

Kind = v1sDiffuseStrong, BarThicknessInPixels = 0, BarAngle = 0

Πρόσθετε για κάθε χρώμα κοντά στο βασικό χρώμα

Kind = v1sDiffuseWeak, BarThicknessInPixels = 0, BarAngle = 0

Πρόσθεσε για το ανταγωνιστικό χρώμα

Kind = v1sNonDiffuse, BarThicknessInPixels = 0, BarAngle = 0

Τα ζεύγη χρωμάτων έχουν ως πρώτο χρώμα BarColor το βασικό χρώμα και ως BackGroundColor πολλά διαφορετικά χρώματα ξεκινώντας στο πρώτο ζεύγος με το αμιγώς ανταγωνιστικό του. Το όγδοο και τελευταίο ζεύγος είναι το ακριβώς αντίστροφο του πρώτου έχοντας ως BarColor το ανταγωνιστικό και ως BackGroundColor το βασικό. Με το τελευταίο παράγονται πρότυπα που αναστέλλουν την λειτουργία του τεχνητού κυττάρου V1S .

Όπως ήδη αναφέραμε υπάρχουν 6 στόχοι αντίστοιχοι με τις εξόδους οποιουδήποτε ΤΝΔ. Έτσι υπάρχει σε ένα πρότυπο μια ιδιότητα πίνακα με τις τιμές στόχων:

- **AngleResponse** [1..6], με τιμές από 0 έως 1

Η ιδιότητα αυτή παίρνει πάντα τιμές 0 ή 1 σε περιπτώσεις που σχεδιάζεται μπάρα. Στην περίπτωση όμως που υπάρχει διάχυτος φωτισμός μπορεί να έχει και ενδιάμεσες τιμές για να εκφράσει αβεβαιότητα. Για παράδειγμα όταν έχουμε ένα V1S ευαίσθητο σε βασικό χρώμα πράσινο με χρήση διάχυτου φωτισμού με αμιγές πράσινο {0,255,0} θα παραχθούν οι στόχοι [1, 1, 1, 1, 1, 1]. Αν όμως χρησιμοποιηθεί αμιγές κυανό {0, 255, 255} οι στόχοι πλέον θα είναι [0.85, 0.85, 0.85, 0.85, 0.85, 0.85] καθώς θεωρείται ότι ενεργοποιεί κατά 85% τον κάθε τεχνητό V1S νευρώνα.

Όλα τα παραπάνω ισχύουν και στις δύο περιπτώσεις μικρών και μεγάλων τεχνητών V1S. Απλώς στα μικρά χρησιμοποιούνται διαφορετικά πάχη για τις μπάρες, έχοντας τις ίδιες 6 περιστροφές. Το πλήθος των προτύπων κυμαίνεται από 100 έως και 400 γιατί διαφοροποιούνται εκτός του πάχους και οι χρωματικοί συνδυασμοί ανάλογα με το βασικό χρώμα του κάθε τύπου ΤΝΔ, για τον οποίο παράγονται τα πρότυπα.

Χαρακτηριστικά Είσοδοι :

Κόκκινο - Πράσινο

Μπλέ - Κίτρινο

Χρώματα :

Γωνία Μπάρας : 30

Πάχος Μπάρας : 12

Είσοδοι : Γάγγλια: 7 Πλήθος Εισόδων: 56

Στόχοι :	Εξοδος :
0,000	0,007
1,000	0,988
0,000	0,000
0,000	0,000
0,000	0,000
0,000	0,000
0,000	0,002

10-12: Οπτικοποίηση προτύπου για R-PV1S με γωνία 30° και πάχος μπάρας 12 pixels, αμιγές κόκκινο πάνω σε αμιγές πράσινο. Βλέπουμε την ύπαρξη 7 rexels που θα χρησιμοποιηθούν σε 56 εισόδους.

Φόρμα προτύπων εκπαίδευσης ενοποιημένης απόκρισης V1

Στην φόρμα με τον τίτλο «Πρότυπα Εκπαίδευσης V2» δημιουργείται μια συλλογή προτύπων για την εκπαίδευση του ΤΝΔ τύπου nnkV2CellPatterns. Αυτό το ΤΝΔ αναλαμβάνει να δημιουργήσει μια κοινή απόκριση για δύο ζεύγη μικρών και μεγάλων V1S. Το κάθε ζεύγος περιλαμβάνει τεχνητά PV1S ή MV1S δύο ανταγωνιστικών χρωμάτων, για παράδειγμα κόκκινου και πράσινου {R-PV1S, G-PV1S} και {R-MV1S, G-MV1S}. Για οποιοδήποτε ζεύγος χρησιμοποιείται ένα ΤΝΔ και με την χρήση του του πλήκτρου «Παραγωγή» στην φόρμα δημιουργείται μια συλλογή προτύπων για την εκπαίδευση του. Ένα πρότυπο έχει 4 χαρακτηριστικές ιδιότητες

- **ParviOnValue**, τιμή από 0 έως 1 για το μικρό V1S του πρώτου χρώματος
- **ParviOffValue**, τιμή από 0 έως 1 για το μικρό V1S του ανταγωνιστικού χρώματος
- **MagnaOnValue**, τιμή από 0 έως 1 για το μεγάλο V1S του πρώτου χρώματος
- **MagnaOffValue**, τιμή από 0 έως 1 για το μεγάλο V1 του ανταγωνιστικού χρώματος

Στην Generate παράγονται συνδυασμοί των τεσσάρων παρακάτω ιδιοτήτων, ασχέτως αν ο συνδυασμός μπορεί να εμφανιστεί ή όχι στην έξοδο των ΤΝΔ των V1S:

Συνδυασμοί προτύπων εκπαίδευσης V2

Καθορίζονται 7 δείγματα τιμών από 0 έως 1 = {0, 0.25, 0.8, 0.85, 0.9, 0.95, 1}

Για κάθε τιμή **V1** από τα 7 δείγματα θέσε **ParviOnValue = V1**

Για κάθε τιμή **V2** από τα 7 δείγματα θέσε **ParviOnValue = V2**

Για κάθε τιμή **V3** από τα 7 δείγματα θέσε **ParviOnValue = V3**

Για κάθε τιμή **V4** από τα 7 δείγματα θέσε **ParviOnValue = V4**

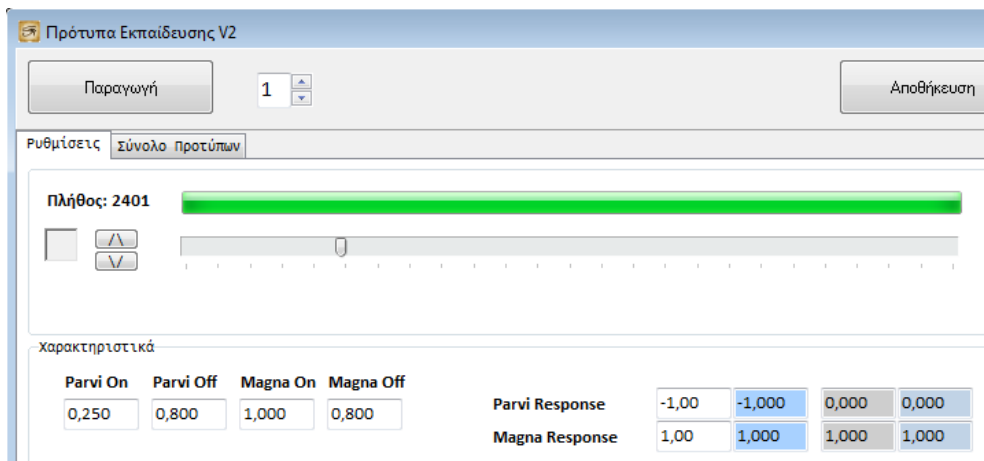
Με αυτόν τον τρόπο καλύπτονται πολλές διαφορετικές περιπτώσεις παράλληλη ενεργοποίησης των τεσσάρων διαφορετικών V1S που δραστηριοποιούνται για έναν συγκεκριμένο χρωματικό ανταγωνισμό από τους τρεις πάνω στην ίδια περιοχή της εικόνας. Οι δύο έξοδοι υπολογίζονται σύμφωνα με την λογική που περιγράφηκε στο κεφάλαιο 9 και είναι

- **ParviResponse**, με τρεις πιθανές τιμές -1,0, 1 για αναστολή, απροσδιοριστία και διέγερση αντίστοιχα
- **MagnaRespons**, με τρεις πιθανές τιμές -1,0, 1 για αναστολή, απροσδιοριστία και διέγερση αντίστοιχα

Οι δύο έξοδοι συνδυάζονται σε μια κοινή απόκριση με τον απλό μαθηματικό τύπο 9.27, που περιγράψαμε στο κεφάλαιο 9, και τον οποίο αντιπροσωπεύει η ιδιότητα

- **Response**, από -1 έως 1 με πιθανές ενδιάμεσες τιμές.

Το πλήθος των διαφορετικών προτύπων που παράγεται είναι 2401 όπως βλέπουμε και στην παρακάτω αποτύπωση οθόνης (screenshot).



10-13: 10-14: Φόρμα που δημιουργεί, εμφανίζει και ανακαλεί πρότυπα εκπαίδευσης του δικτύου ενοποίησης της απόκρισης των τεχνητών V1S

Οπτική Αναγνώριση των Πραγματικών Αντικειμένων

Ανάπτυξη του τεχνητού αμφιβληστροειδούς

Το αντικείμενο Retina: TRetina αποτελεί την πρόσοψη του τμήματος II του Horasis, του τεχνητού αμφιβληστροειδούς η για συντομία ΟΟΑΡ. Θεωρούμε ότι έχουμε μια εικόνα σε ψηφιδωτή μορφή BMP που μπορεί να έχει ληφθεί από μια web κάμερα ή να έχει σχεδιαστεί προγραμματιστικά όπως γίνεται στα πρότυπα εκπαίδευσης. Σε κάθε περίπτωση θεωρούμε ένα αντικείμενο της κλάσης TGImage που περιέχει την εικόνα και ανατίθεται στην ιδιότητα Retina.SourceImage. Επίσης γράφουμε μια για τον χειρισμό του συμβάντος (event) **Retina.OnProcessEvent**. Τέλος δημιουργούμε το αντικείμενο StartPoint: TGCoords στο οποίο δίνουμε ως ιδιότητες X και Y τις συντεταγμένες {250, 250} του κέντρου της εικόνας.

Με την κλήση της μεθόδου **Retina.GenerateFoveaCentralis(StartPoint)** που έχει ως όρισμα αυτό το σημείο, παράγεται ένας ΟΟΑΡ έτοιμος με όλα τα ΤΝΔ που υπάρχουν σε αυτόν φορτωμένα με τις τιμές συναπτικών βαρών που προέκυψαν από τις εκπαίδευσεις. Πρώτα από όλα δημιουργείται ο τεχνητός ρετινοτοπικός χάρτης ARTN στην ιδιότητα Retina.Honeycomb με βάση τα χαρακτηριστικά της ψηφίδωσης που καθορίζονται στις ιδιότητες, έχοντας τις αντίστοιχες αρχικές τιμές:

- Retina.HexagonalDiameter = 16
- Retina.TesselationDepth = 25
- Retina.MaxEccentricity = 230

Έπειτα για κάθε εξάγωνο μέσα στον χάρτη δημιουργείται ένα rexel της κλάσης TRetinalElement και αποστέλλεται ως παράμετρος Sender μέσω του συμβάντος OnProcessEvent, στην μέθοδο χειρισμού του συμβάντος (event handler). Η δημιουργία ενός NewRexel: TRetinalElement περιλαμβάνει

- Την φόρτωση ενός χάρτη υποδεκτικού πεδίου από μια ειδικά φτιαγμένη εικόνα BMP FHorizontalCellMappings με κλήση της μεθόδου **NewRexel.BuildAreasFromImage(FHorizontalCellMappings)**
- Την ανάθεση προς τα κάτω της ιδιότητας Retina.SourceImage στην NewRexel.SourceImage
- Την αρχικοποίηση του ρετινοστοιχείου με **NewRexel.Initialize()** στην οποία έχουμε:
 - Την δημιουργία μιας νέας φέτας NewSlice: TRetinalSlice για κάθε μια περιοχή του χάρτη υποδεκτικού πεδίου
 - Το καθορισμό της θέσης της φέτας μέσα στον τεχνητό αμφιβληστροειδή με ανάθεση των ιδιοτήτων NewSlice.SlicePosition και NewSlice.ElementNumber
 - Την εύρεση και σύνδεση με τα pixels της αρχικής εικόνας που χαρτογραφεί η περιοχή του χάρτη υποδεκτικού πεδίου.
 - Την αρχικοποίηση της φέτας με **NewSlice.Initialize()** κατά την οποία:
 - Δημιουργούνται 4 τεχνητά HC για την φέτα
 - Δημιουργούνται 6 τεχνητά BPC για την φέτα που συνδέονται με 2 ΤΝΔ
- Την δημιουργία 6 τεχνητών ΤΝΔ-GC NewGanglion: TGanglionCell για το ρετινοστοιχείο NewRexel, ένα για κάθε ανταγωνιστικό χρώμα, όπου
 - Φορτώνεται η δομή του δικτύου από το TGNeuralNet του κάθε τύπου μέσω του singleton TNeuralNetManager

- Συνδέονται τα υποδεκτικά πεδία των ΤΝΔ-GC με τα τεχνητά BPC των χρωμάτων που χρησιμοποιούνται από όλες τις φέτες κέντρου και περιφέρειας

Πλέον έχει αναπτυχθεί ο ΟΟΑΡ και είναι έτοιμος για την ανάκληση της εικόνας. Η ανάκληση γίνεται διατρέχοντας κάθε Rexel: TRetinalElement που περιέχεται στην λίστα **Retina.Elements** και καλώντας από αυτό την **Rexel.Process()** όπου:

- Για κάθε φέτα Slice: TRetinalSlice που υπάρχει στο rexel καλείται η **Slice.Stimulate()** που:
 - Φορτώνει τις τιμές από τους τεχνητούς φωτοϋποδοχείς στα τεχνητά HC
 - Παράγει την έξοδο των τεσσάρων τεχνητών HC
- Για κάθε φέτα Slice: TRetinalSlice που υπάρχει στο rexel καλείται η Slice.Stimulate που:
 - Μεταφέρει την έξοδο των HC ως είσοδο στα ΤΝΔ των τεχνητών BPC
 - Εκτελεί ανάκληση στα ΤΝΔ των τεχνητών BPC, υπολογίζονται την τιμή εξόδου για κάθε ένα από τα έξι.
- Για κάθε ένα από τα έξι τεχνητά GC που υπάρχουν για το rexel και συμβολίζεται Ganglion: TGanglionCell καλείται η **Ganglion.StimulateCell()** κατά την οποία
 - Μεταφέρονται οι τιμές εξόδου που έχουν υπολογιστεί για τα τεχνητά BPC
 - Εκτελείται ανάκληση του ΤΝΔ και παράγονται οι τέσσερις τιμές της εξόδου που αποτελούν την κωδικοποίηση GCE.

Καλώντας την Process μέσα από την υλοποίηση του συμβάντος Retina.OnProcessElement μπορούμε κατά την ανάπτυξη του ΟΟΑΡ, να δούμε το σχηματίζεται σταδιακά η οπτικοποίηση της εξόδου του. Αυτό ακριβώς γίνεται σε μια φόρμα που οπτικοποιεί την εσωτερική λειτουργία των ΤΝΔ του horasis που θα γνωρίσουμε παρακάτω.

Ανάπτυξη του τεχνητού οπτικού φλοιού

Το αντικείμενο `VisualCortex`: `TVisualCortex` αποτελεί την πρόσοψη του τμήματος III του `Horasis`, του τεχνητού οπτικού φλοιού ή για συντομία `AVC`. Θεωρούμε ότι έχουμε ήδη ανεπτυγμένο έναν `OOAR` στο αντικείμενο `Retina`: `TRetina`, ο οποίος περιέχει την πηγαία εικόνα και τον `ARTN`. Αναθέτουμε την ιδιότητα `VisualCortex.Retina := Retina`. Με την κλήση της μεθόδου `VisualCortex.Initialize()` αναπτύσσεται ο `AVC` ως εξής:

- Για κάθε εξάγωνο `Cell`: `THoneyComb` στον `ARTN` και εφόσον αυτό ανήκει στα κεντρικά
 - Δημιουργείται ένα νέο κομμάτι οπτικού φλοιού `NewSlice`: `TVisualCortexSlice`.
 - Δίδεται σε αυτό το αντικείμενο `OOAR`.
 - Δίνεται σε αυτό το κεντρικό εξάγωνο.
 - Αρχικοποιείται το κομμάτι με κλήση της `NewSlice.Initialize()`
 - Βρίσκει τα 7 `rexel` του υποδεκτικού πεδίου των τεχνητών `PV1S` βασιζόμενο στο κεντρικό εξάγωνο
 - Βρίσκει τα 19 `rexel` του υποδεκτικού πεδίου των τεχνητών `MV1S` βασιζόμενο στο κεντρικό εξάγωνο
 - Για κάθε ένα από τα 6 ανταγωνιστικά χρώματα δημιουργείται ένα νέο αντικείμενο `NewV2Block`: `TV2Block` και αρχικοποιείται με `NewV2Block.Initialize()`
 - Δημιουργούνται τόσα 6 αντικείμενα `TV2Cell` όσο το πλήθος των διαφορετικών κατευθύνσεων που υπάρχουν, για ευαισθησία ανά 30° . Σε κάθε `TV2Cell` υπάρχει ένα τεχνητό `PV1S` και ένα `MV1S`.

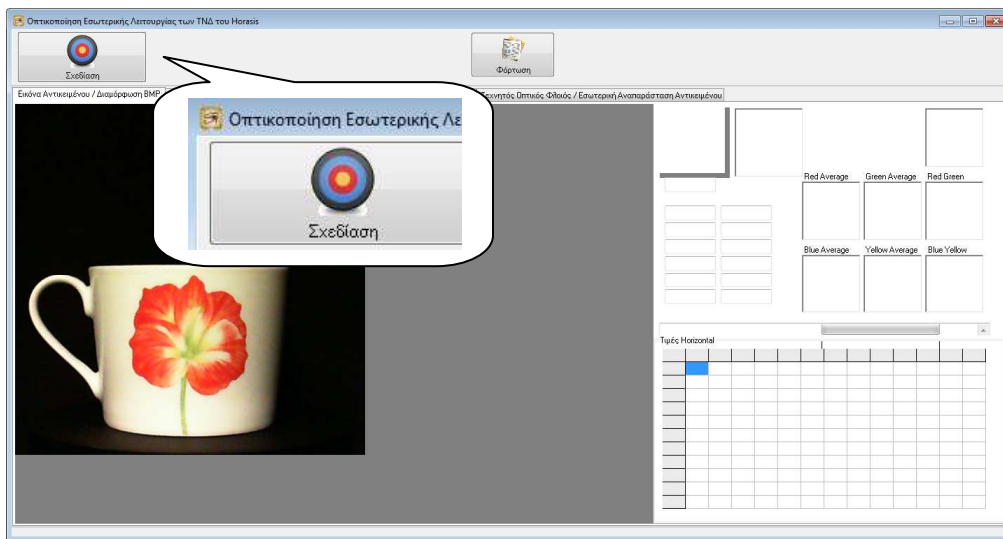
Μετά την ανάπτυξη του `AVC` μπορούμε να εκτελέσουμε μια πλήρη ανάκληση της εικόνας που την μετατρέπει στην έξοδο του `OOAR` σε εσωτερική αναπαράσταση για το `Horasis`. Καλώντας για κάθε `VCSlice`: `TVisualCortexSlice` που περιέχεται στην λίστα `VisualCortex.Slices` την μέθοδο `VCSlice.Process()` έχουμε:

- Ανάκληση όλων των `rexel` που υπάρχουν στα υποδεκτικά πεδία των `PV1S` και `MV1S` με χρήση της δικής τους `Process`
- Για κάθε ένα από τα τεχνητά `PV1S` που υπάρχουν σε `TV2Block` όλων των χρωμάτων
 - Φορτώνει τις τιμές από τα 7 τεχνητά `GC` των `rexel` του υποδεκτικού τους πεδίου.
 - Υπολογίζει την έξοδο των 6 `PV1S`.
- Για κάθε ένα από τα τεχνητά `PV1S` που υπάρχουν σε `TV2Block` όλων των χρωμάτων:
 - Φορτώνει τις τιμές από τα 19 τεχνητά `GC` των `rexel` του υποδεκτικού τους πεδίου.
 - Υπολογίζει την έξοδο των 6 `MV1S`.
- Για κάθε ζεύγος `TV2Block` κόκκινο/πράσινο, μπλε/κίτρινο ή λευκό/μαύρο
 - Για κάθε μια από τις 6 κατευθύνσεις
 - συγχωνεύει τις αποκρίσεις `PV1S` και `MV1S` και των δύο ανταγωνιστικών χρωμάτων, σύνολο 4 τεχνητά κύτταρα να παράγουν 1 ενιαία απόκριση

Μετά την κλήση της `Process` μπορούμε να οπτικοποιήσουμε τα αποτελέσματα 3 ενιαίων αποκρίσεων που υπάρχουν σε μια κατεύθυνση ενός `TVisualCortexSlice` μέσα σε τρία διαφορετικά παράθυρα, ένα για κάθε χρωματικό ανταγωνισμό. Αυτό ακριβώς θα δούμε στην φόρμα οπτικοποίησης της εσωτερικής λειτουργίας του `Horasis`

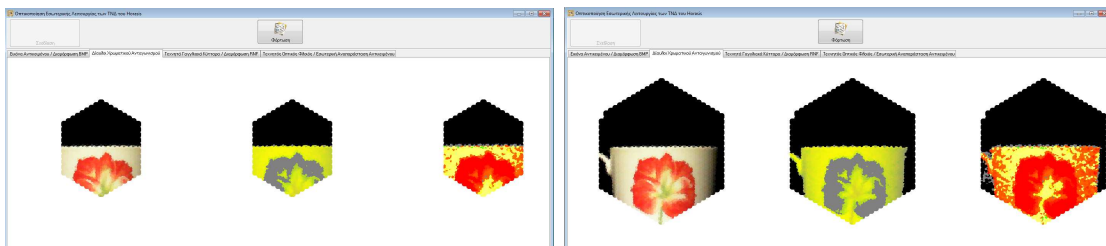
Φόρμα οπτικοποίησης εσωτερικής λειτουργίας του Horasis

Κατά την υλοποίηση του Horasis θεωρήθηκε ως ένας εύκολος τρόπος για την αξιολόγηση της λειτουργίας της εφαρμογής να περιέχει οπτική πληροφορία στην θέση της αριθμητικής. Έτσι με μια ματιά θα μπορούσε να γίνει οπτική επιβεβαίωση για την ορθότητα της λειτουργίας ενός από τα πολλά διαφορετικά στρώματα επεξεργασίας που υπάρχουν στην εφαρμογή. Όλες οι τιμές εξόδου των ΤΝΔ αυτών των στρωμάτων παρήγαγαν μια οπτικοποίηση, η οποία τελικά βοήθησε σημαντικά στην επιτυχία της υλοποίησης. Η κάθε αριθμητική τιμή συνδέθηκε με μια χρωματική αναπαράσταση η οποία εμφανιζόταν στο κάθε τμήμα της εικόνας, που διαχωρίστηκε από την χρήση του τεχνητού ρετινοτοπικού χάρτη ARTN. Όλα τα τμήματα της εικόνας χρωματίζονται ανάλογα με τις τιμές και συνθέτουν μια νέα εικόνα που οπτικοποιεί την εσωτερική λειτουργία και μπορεί να χρησιμοποιηθεί για πολλαπλή πληροφόρηση στο ίδιο παράθυρο. Τέτοια παράθυρα υπάρχουν στην φόρμα «Οπτικοποίηση Εσωτερικής Λειτουργίας των ΤΝΔ του Horasis».



10-15: Η φόρμα οπτικοποίησης εσωτερικής λειτουργίας του Horasis, εμφανίζοντας την αρχική εικόνα BMP

Με την χρήση του πλήκτρου «Σχεδίαση» ξεκινά η ανάπτυξη του τεχνητού αμφιβληστροειδούς πάνω σε μια αρχική εικόνα BMP ενός αντικειμένου. Ο ARTN αναπτύσσεται σταδιακά από μέσα προς τα έξω, φτάνοντας τελικά στο σχεδόν κυκλικό σχήμα του. Κατά την διάρκεια της ανάπτυξης σχεδιάζεται για κάθε rexel η οπτικοποίηση παράλληλα, σε πολλά παράθυρα, με το κάθε ένα να εμφανίζει την έξοδο των διαφορετικών ομάδων τεχνητών κυττάρων.



10-16: Δύο διαφορετικά ενδιάμεσα στάδια κατά την οπτικοποίηση της ανάπτυξης του τεχνητού ρετινοτοπικού χάρτη

Όταν τελειώνει η ανάπτυξη μπορούμε να μετακινηθούμε στις διαφορετικές σελίδες με μεγαλύτερη ταχύτητα και να δούμε τα παράθυρα των αντίστοιχων οπτικοποιήσεων:

1. Σελίδα «Δίαυλοι Χρωματικού Ανταγωνισμού»

- **Οπτικοποίηση τεχνητών HC**

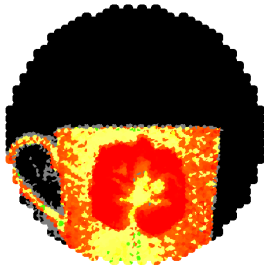


Χρωματισμός των όλων των pixel μιας φέτας με το χρώμα που προκύπτει από τους μέσους όρους των R,G,B που παράγονται από τα αρχικά pixel της περιοχής της εικόνας BMP



10-17: Λεπτομέρεια οπτικοποίησης τεχνητών HC στην οποία εμφανίζεται η παρουσία του χάρτη υποδεκτικού πεδίου το κέντρο του και οι φέτες της περιφέρειας του.

- **Οπτικοποίηση τεχνητών BPC κόκκινου/πράσινου,**



Εμφανίζεται με κόκκινο και πράσινο οι ενεργοποιήσεις R-BPC και G-BPC με κίτρινο η απροσδιοριστία και με μαύρο η ύπαρξη χαμηλού φωτισμού που ισοδυναμεί με την λειτουργία του φίλτρου των τεχνητών βραχύινων κυττάρων AC.

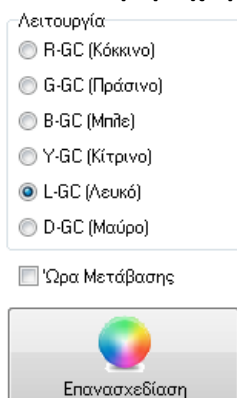
- **Οπτικοποίηση τεχνητών BPC μπλε/κίτρινου,**



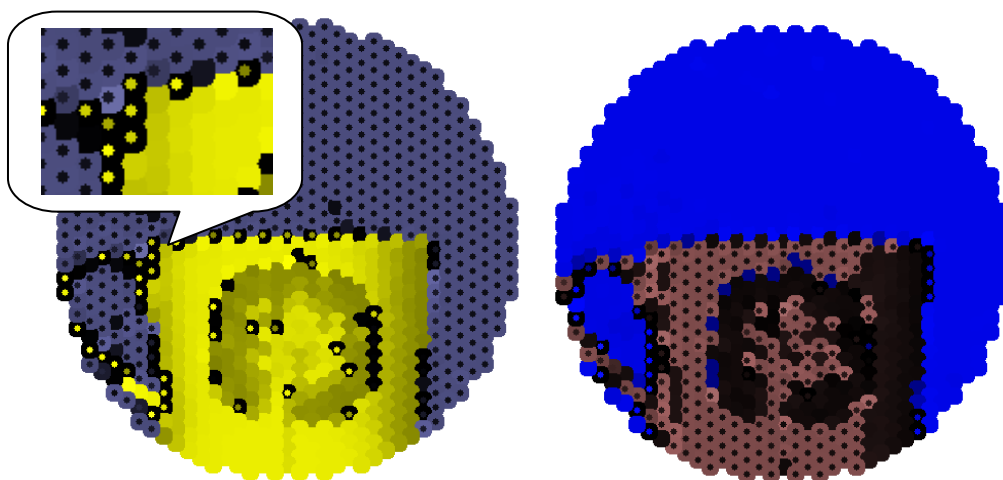
Εμφανίζεται με μπλε η ενεργοποίηση B-BPC και με κίτρινο των Y-BPC με γκρι η απροσδιοριστία και με μαύρο ο χαμηλός φωτισμός και η λειτουργία του φίλτρου των τεχνητών AC

2. Σελίδα «Τεχνητά Γαγγλιακά Κύτταρα / Διαμόρφωση RNF»

- **Οπτικοποίηση τεχνητών GC κόκκινου, πράσινου, μπλε, κίτρινου, λευκού, μαύρου,**



Επιλέγεται αρχικά το είδος του ΤΝΔ-GC που θα οπτικοποιηθεί, έτσι αν επιλέξουμε διαδοχικά το λευκό και μαύρο χρώμα πατώντας «Επανασχεδίαση», στο ίδιο παράδειγμα:

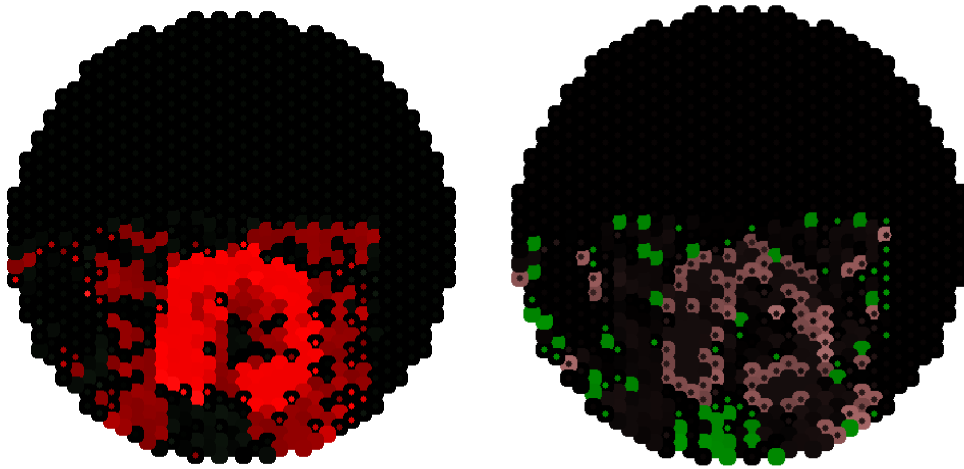


10-18: Οπτικοποίηση εξόδου τεχνητών L-GC αριστερά και των ανταγωνιστικών D-GC δεξιά

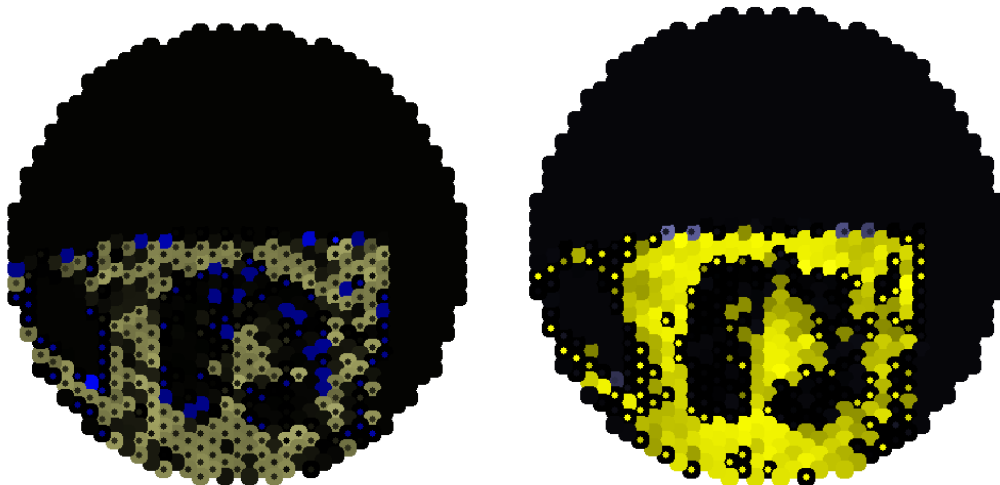
Για το τεχνητό L-GC του λευκού εμφανίζονται με μωβ χρώμα τα σκοτεινά σημεία της εικόνας και με κίτρινο τα φωτεινά. Όταν έχουμε σε ένα rexel έναν κίτρινο μικρό κύκλο σε έναν μαύρο κυκλικό δίσκο σημαίνει ότι έχουμε διέγερση του κέντρου με το βασικό χρώμα και διέγερση της περιφέρειας με το ανταγωνιστικό του. Με μαύρο κύκλο πάνω σε έναν μωβ κυκλικό δίσκο έχουμε ακριβώς το αντίστροφο με πλήρη αναστολή και από το κέντρο και από την περιφέρεια. Με γεμάτο κυκλικό δίσκο έχουμε διάχυτο φωτισμό, κίτρινο για διεγερτικό και μωβ για ανασταλτικό. Με μαύρο έχουμε πλήρη απενεργοποίηση του νευρώνα ή απροσδιοριστία. Όσο μεγαλύτερη η φωτεινότητα του κάθε χρώματος τόσο μεγαλύτερη η τιμή απόκρισης του τεχνητού νευρώνα της διέγερσης ή της αναστολής. Άρα φωτεινό κίτρινο σημαίνει μέγιστη διέγερση ενώ σκούρο κίτρινο ασθενή διέγερση. Με αυτόν τον τρόπο απεικονίζεται οπτικά μια πληροφορία που αποτελείται από τέσσερις διαστάσεις που είναι το διάνυσμα GCE της εξόδου κάθε ΤΝΔ-GC.

Για το τεχνητό D-GC του μαύρου εμφανίζονται με την ίδια φιλοσοφία με μπλε χρώμα τα σκοτεινά σημεία που ενεργοποιούνται και με κόκκινο τα φωτεινά που απενεργοποιούνται. Με αυτόν τον τρόπο έχουμε την πρώτη οπτική πληροφόρηση για την παράλληλη απόκριση δύο ανταγωνιστικών τεχνητών κυττάρων που λειτουργούν πάνω στα ίδια δεδομένα.

Ομοίως έχουμε τέσσερις ακόμα οπτικοποιήσεις για R-GC, G-GC, B-GC και Y-GC. Και εδώ η φωτεινότητα στα τέσσερα χρώματα κόκκινο, πράσινο, μπλε και κίτρινο δείχνει ισχυρή διέγερση.



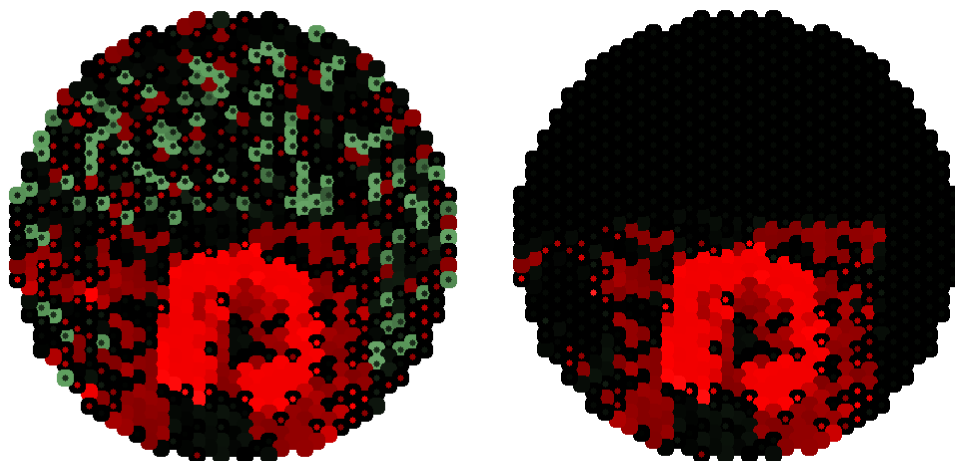
10-19: Οπτικοποίηση εξόδου τεχνητών R-GC αριστερά και των ανταγωνιστικών G-GC δεξιά



10-20 Οπτικοποίηση εξόδου τεχνητών B-GC αριστερά και των ανταγωνιστικών Y-GC δεξιά

Πρώτη παρατήρηση που προκύπτει από την οπτική εποπτεία της λειτουργίας είναι ότι ανά δύο τα ανταγωνιστικά χρώματα δεν παράγουν εντελώς αντίθετες αποκρίσεις. Δεν είναι δηλαδή η μια απόκριση το «αρνητικό» της άλλης. Εμφανίζεται ένας πολυμορφισμός που στο παράδειγμα μας φαίνεται στην εικόνα 10-18, όταν το κόκκινο λουλούδι ενεργοποιεί ισχυρά ως διάχυτο φως τα R-GC του τεχνητού αμφιβληστροειδούς χωρίς όμως να παράγει την ίδια ισχυρή και διάχυτη αναστολή στα G-GC. Επίσης βλέπουμε ότι στην λαβή της κούπας ενεργοποιούνται με τέλει φωτισμό κέντρου περιφέρειας κάποια κίτρινα Y-GC ενώ τέσσερα από αυτά εμφανίζουν διάχυτο φωτισμό. Στα ανταγωνιστικά B-GC και στα ίδια ακριβώς ρετινοστοιχεία, δεν παράγεται στην έξοδο διάχυτος ανασταλτικός φωτισμός που θα φαινόταν ως ένας θαμπός γεμάτος κίτρινος κύκλος. Με αυτήν την παρατήρηση αξιολογήθηκε στην μέση της υλοποίησης ότι η προσέγγιση με την χρήση διαφορετικών διαύλων χρωματικών ανταγωνισμών ήταν πετυχημένη, καθώς ο κάθε δίαυλος λειτουργούσε με διαφορετικό τρόπο.

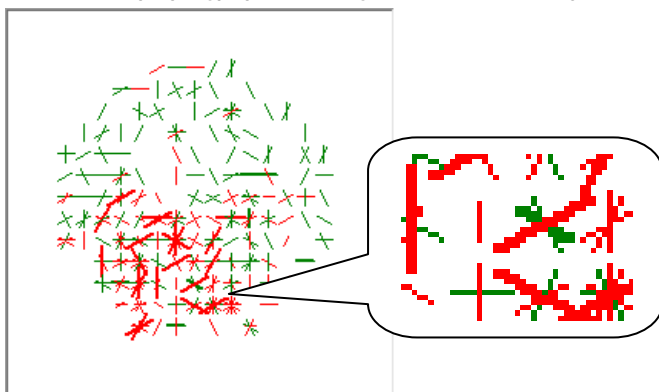
Μια δεύτερη παρατήρηση που οδήγησε στον καθορισμό της λειτουργίας των τεχνητών βραχύνων κυττάρων AC AC ήταν η παρατήρηση χρωματικών «σκουπιδιών» στο μαύρο φόντο της εικόνας. Επειδή ένα τεχνητό BPC αγνοεί τον φωτισμό και λειτουργεί αποκλειστικά βάσει της απόχρωσης, εμφανιζόταν αποκρίσεις σε τεχνητά GC για σκοτεινές περιοχές του φόντου. Έχοντας άμεση οπτική πληροφόρηση έγιναν δοκιμές διαφορετικών τιμών κατωφλίου για την συνάρτηση f_{ac} , που γνωρίσαμε στο κεφάλαιο 9, ώστε τελικά να επιλεγθεί η καλύτερη που καθάρισε το φόντο του αντικειμένου χωρίς να αφαιρεί χρήσιμες πληροφορίες από το ίδιο το αντικείμενο.



10-21: Οπτικοποίηση τεχνητών R-GC χωρίς φιλτράρισμα εισόδων μέσω τεχνητών AC αριστερά, και με χρήση φίλτρου τεχνητών AC δεξιά

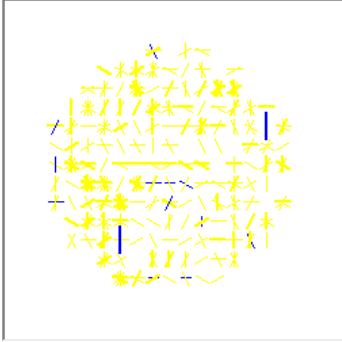
3. Σελίδα «Τεχνητός Οπτικός φλοιός / Εσωτερική Αναπαράσταση Αντικειμένου»

- **Οπτικοποίηση τεχνητών V1S για κόκκινο και πράσινο**



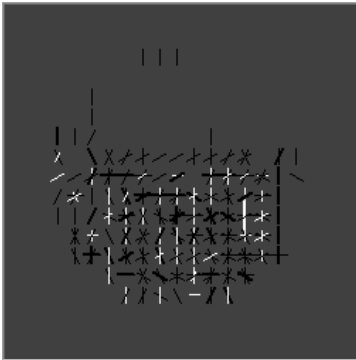
Στο κάθε ένα από τα δύο χρώματα εμφανίζεται μια λεπτή και μικρή γραμμή για κάθε τεχνητό PV1S που ενεργοποιείται. Η περιστροφή της γραμμής αντιπροσωπεύει την κατεύθυνση στην οποία είναι προσανατολισμένος ο τεχνητός νευρώνας PV1S. Αν η γραμμή είναι λίγο μεγαλύτερη και παχύτερη τότε σε αυτήν την κατεύθυνση ενεργοποιείται έναν τεχνητός νευρώνας MV1S με μεγαλύτερο υποδεκτικό πεδίο. Αυτή είναι μια όψη της εσωτερικής αναπαράστασης του αντικειμένου που αφορά τον πρώτο από τους τρεις χρωματικούς ανταγωνισμούς.

- **Οπτικοποίηση τεχνητών V1S για μπλε και κίτρινο**



Και εδώ έχουμε μπλε και κίτρινες, μικρές και μεγάλες γραμμές ανάλογα με την ενεργοποίηση συγκεκριμένων κατευθύνσεων. Είναι ο δεύτερος χρωματικός ανταγωνισμός μέσα στην εσωτερική αναπαράσταση

- **Οπτικοποίηση τεχνητών V1S για λευκό και μαύρο**



Με ύπαρξη γκρι φόντου για να εμφανίζονται σωστά, έχουμε λευκές και μαύρες, μικρές και μεγάλες γραμμές συγκεκριμένης κατεύθυνσης. Είναι ο τρίτος χρωματικός ανταγωνισμός που ολοκληρώνει την εσωτερική αναπαράσταση του αντικειμένου. Κοιτώντας το πλαίσιο που εμφανίζει και τα τρία παράθυρα έχουμε με μια ματιά πληροφόρηση για την εσωτερική λειτουργία του τμήματος του τεχνητού οπτικού φλοιού.

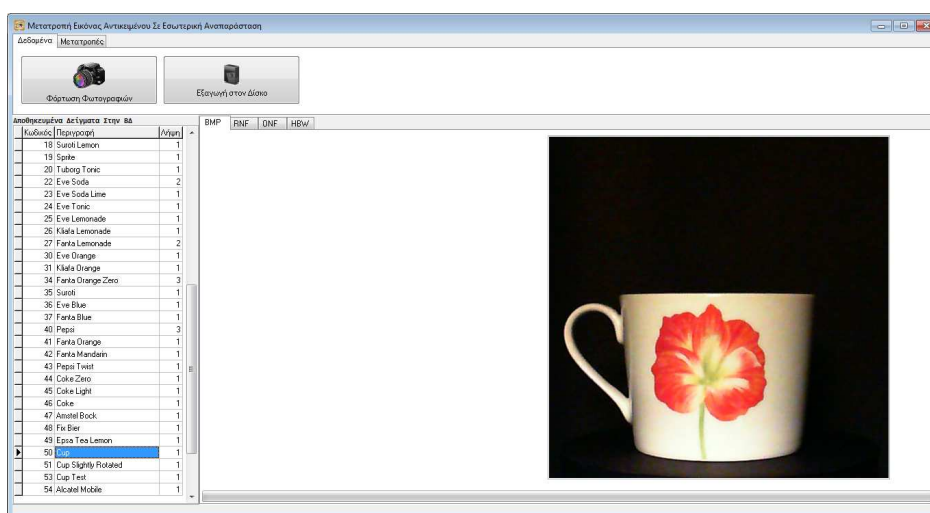
Η φόρμα αυτή αναπτύχθηκε αυξητικά καθώς μετά από κάθε πετυχημένη εκπαίδευση ενός από τα ΤΝΔ του Horasis, υλοποιούνταν η εμφάνιση του αποτελέσματος μέσα σε αυτήν. Έχοντας την ως βασικό εργαλείο για τον έλεγχο της προόδου της συνολικής υλοποίησης, ελήφθησαν σημαντικές αποφάσεις που οδήγησαν ακόμα και σε αλλαγή του σχεδιασμού των ΤΝΔ ή επανεκπαίδευση τους με καλύτερα πρότυπα. Σε τυχόν μελλοντικές εκδόσεις του Horasis μπορεί να βοηθήσει σημαντικά καθώς μπορεί να γίνει μια γρήγορη οπτική σύγκριση του αποτελέσματος της λειτουργίας της νέας έκδοσης σε σχέση με την παλιά.

Για να εμφανιστούν στην φόρμα τα αποτελέσματα όλης της λειτουργίας του Horasis, δημιουργούνται εκατομμύρια αντικειμένων, αφού στήνεται ένα τεράστιο ΤΝΔ που συντίθεται από όλα τα διαφορετικά ΤΝΔ που απαιτούνται για κάθε τμήμα της εικόνας. Γι αυτό εμφανίζει την μεγαλύτερη κατανάλωση μνήμης που υπάρχει στο Horasis. Ως αποτέλεσμα δεν μπορεί να χρησιμοποιηθεί ένα δεύτερο στιγμιότυπο της φόρμας, γιατί θα εξαντλήσει της μνήμη του συστήματος. Λόγω αυτού του περιορισμού τα οπτικά πρότυπα, όπως θα δούμε στην αμέσως επόμενη ενότητα, οι αρχικές εικόνες BMP μετατρέπονται στην τελική εσωτερική αναπαράσταση με ανάκληση αυτού του τεράστιου ΤΝΔ. Κατόπιν αποθηκεύονται ώστε να μην χρειάζεται η ύπαρξη του για την εκπαίδευση και λειτουργία του τελικού ΤΝΔ εξόδου του Horasis.

Φόρμα μετατροπών οπτικών προτύπων του πραγματικού αντικειμένου

Η σύσταση όλων των ΤΝΔ για την μετατροπή της αρχικής εικόνας του αντικειμένου σε μια εσωτερική αναπαράσταση, μπορούμε να πούμε ότι αποτελεί ένα υπερμέγεθες ΤΝΔ. Για την δημιουργία μιας πληθώρας από αντικείμενα και ΤΝΔ για όλα τα υποσυστήματα των τμημάτων II και III απαιτείται σημαντικό μέγεθος μνήμης που μετριέται 2 με 3GB. Για αυτό κρίθηκε σκόπιμο να μην υπάρχει ολόκληρη αυτή η συνάθροιση αντικειμένων σε κάθε στιγμή της εκτέλεσης της εφαρμογής, αλλά να δημιουργείται μόνο για τον μετασχηματισμό των οπτικών προτύπων του πραγματικού αντικειμένου.

Προκειμένου να τηρηθεί το παραπάνω δημιουργήθηκε η φόρμα με τίτλο «Μετατροπές Οπτικών Προτύπων», στην οποία δημιουργείται και επαναχρησιμοποιείται το συνολικό ΤΝΔ. Στην πρώτη σελίδα εμφανίζεται μια λίστα που γεμίζει με τις εγγραφές που υπάρχουν στην βάση δεδομένων και στον πίνακα PHOTO, οι οποίες δημιουργήθηκαν κατά την λήψη φωτογραφιών στο τμήμα I. Στην λίστα υπάρχει η περιγραφή του αντικειμένου και με την επιλογή της εμφανίζεται η αντίστοιχη εικόνα BMP στην οθόνη.



10-22: Φόρμα που χρησιμοποιείται για μετατροπή των εικόνων BMP ενός αντικειμένου σε εσωτερική αναπαράσταση της εικόνας.

Στην δεύτερη σελίδα που έχει τον τίτλο «Μετατροπές» υπάρχουν πλήκτρα με τα οποία η εικόνα μετατρέπεται από BMP σε διαφορετικές διαμορφώσεις που προκύπτουν μετά από την νευρωνική επεξεργασία του Horasis. Με το διπλανό πλήκτρο «Μετατροπή σε διαμόρφωση αμφιβληστροειδούς RNF» εκτελείται ανάκληση των ΤΝΔ μέχρι και τα τεχνητά GC, με αρχική είσοδο την ψηφιδωτή εικόνα BMP του αντικειμένου και τελική έξοδο μια νευρωνική διαμόρφωση αμφιβληστροειδούς RNF. Κατόπιν η μορφή RNF αποθηκεύεται στον πίνακα RNF_PATTERN και συνδέεται με την αρχική εγγραφή που περιέχει την εικόνα BMP.





Συνεχίζοντας πατώντας και το πλήκτρο «Μετατροπή σε Εσωτερική Αναπαράσταση HBW» γίνεται ανάκληση και στα τελευταία ΤΝΔ με είσοδο την μορφή RNF και έξοδο την διαμόρφωση HBW⁷⁹ της εσωτερικής αναπαράστασης

της εικόνας του αντικειμένου. Η μορφή αυτή αποθηκεύεται στον πίνακα BRAINWAVE της βάσης δεδομένων και επίσης συσχετίζεται με την υπάρχουσα εγγραφή του πίνακα PHOTO.

Μετά την μετατροπή όλων των φωτογραφιών BMP σε αντίστοιχες μορφές HBW, έχουμε έναν πίνακα που περιέχει τα πρότυπα εκπαίδευσης για το ΤΝΔ εξόδου του Horasis. Αυτά θα χρησιμοποιηθούν στο τμήμα IV απλά ως πίνακες με αριθμητικές τιμές, χωρίς να υπάρχει η ανάγκη να παραχθεί το συνολικό ΤΝΔ. Δηλαδή η βασική λειτουργία του Horasis που συνδέεται με την προσομοίωση της όρασης, εκτελείται σε αυτήν την φόρμα με την μετατροπή των οπτικών προτύπων. Από εδώ και στο εξής η εσωτερική αναπαράσταση είναι έτοιμη για να χρησιμοποιηθεί στο τελικό ΤΝΔ κατηγοριοποίησης.

⁷⁹ **HBW = Horasis Brain Wave**

Φόρμα προτύπων εκπαίδευσης τεχνητού νευρωνικού δικτύου κατηγοριοποίησης

Το γνωστό πρωτότυπο διασύνδεσης χρήστη που είδαμε σε όλες τις φόρμες παραγωγής πρότυπων εκπαίδευσης, χρησιμοποιείται και για την εκπαίδευση του ΤΝΔ που βρίσκεται στην τελική έξοδο του Horasis σε μια φόρμα με τον τίτλο «Πρότυπα Πραγματικών Αντικειμένων». Η διαφορά σε αυτήν την περίπτωση είναι ότι δεν έχουμε παραγωγή προτύπων από συνδυασμό χαρακτηριστικών για τα πρότυπα, αλλά απλώς φόρτωση εγγραφών από την βάση δεδομένων. Οι εγγραφές που θα φορτωθούν, πρέπει να έχουν εγγραφές στον πίνακα PHOTO όπου η σημαία (flag) IS_TRAINING_PATTERN να έχει τιμή 1. Εκτελώντας μέσα στην Generate της συλλογής προτύπων το ακόλουθο ερώτημα SQL

```
select ID, DESCR from PHOTO where IS_TRAINING_PATTERN = 1
```

έρχονται όλα τα κλειδιά του πίνακα PHOTO για τα οποία πρέπει να φορτωθούν οι εγγραφές προτύπων από τον BRAINWAVE. Για κάθε μια εγγραφή δημιουργείται και ένα αντικείμενο προτύπου στην συλλογή. Μέσα σε αυτό το αντικείμενο φορτώνεται η εσωτερική αναπαράσταση HBW με χρήση της SQL

```
select HBW from BRAINWAVE where PHOTO_ID = :PHOTOID
```

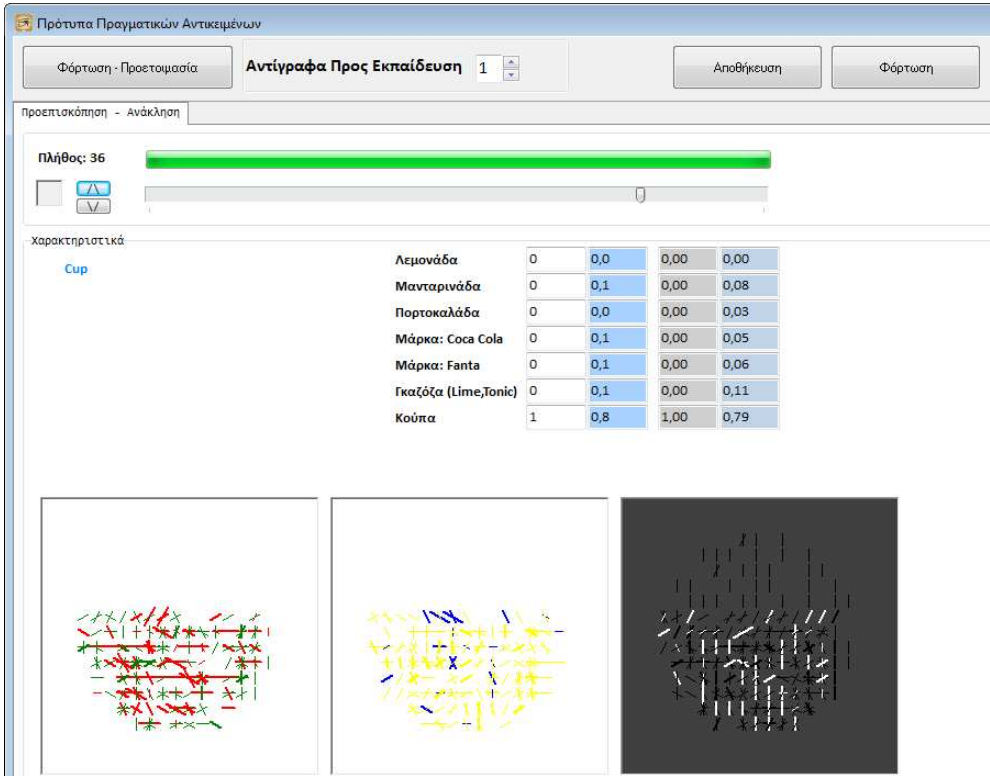
όπου στην παράμετρο :PHOTOID δίδεται η τιμή κλειδιού του πίνακα PHOTO. Η HBW που αποτελείται από 2790 πραγματικούς αριθμούς είναι η είσοδος του προτύπου. Από την άλλη οι στόχοι του προτύπου καθορίζονται σε πεδία μέσα στην εγγραφή του πίνακα PHOTO.

Για την πρώτη έκδοση του Horasis έχουν ληφθεί πρότυπα για πραγματικά αντικείμενα δύο διαφορετικών όγκων, μερικά αλουμινένια κουτιά αναψυκτικών και μια κούπα καφέ. Τα αλουμινένια κουτιά αναψυκτικών είναι διαφορετικά ως προς τα χρώματα και τα σχήματα καθώς ανήκουν σε πολλές διαφορετικές μάρκες. Έτσι έχουμε τρεις διαφορετικές μάρκες πορτοκαλάδας που έχουν παρόμοια χρώματα με εντελώς διαφορετικά σχήματα. Επίσης έχουμε τρία διαφορετικά αναψυκτικά της ίδιας μάρκας με παρόμοια σχήματα και διαφορετικά χρώματα. Μάλιστα τα δύο από τα τρία χρώματα μοιάζουν μεταξύ τους καθώς χρησιμοποιήθηκε πορτοκαλάδα και μανταρινάδα της ίδιας μάρκας που διαφέρουν ελάχιστα. Τέλος χρησιμοποιήθηκαν αναψυκτικά της ίδιας μάρκας με εντελώς διαφορετικά χρώματα μεταξύ τους με σκοπό να κατηγοριοποιηθούν ως προς την μάρκα.

Από αυτά τα πρότυπα θεωρήθηκαν 7 διαφορετικές κλάσεις στην οποία το Horasis θα έπρεπε να τα αναγνωρίσει και να ταξινομήσει τα αντικείμενα. Γι αυτό δημιουργήθηκαν 7 αντίστοιχα πεδία στον πίνακα PHOTO. Θέτοντας τις τιμές των παρακάτω πεδίων σε 1 ή 0 ορίζουμε τους στόχους εκπαίδευσης του κάθε προτύπου θεωρώντας παράλληλα ότι το πραγματικό αντικείμενο ανήκει ή δεν ανήκει στις παρακάτω κλάσεις:

- IS_LEMON, αν το αντικείμενο είναι αναψυκτικό και λεμονάδα
- IS_MANDARIN, αν το αντικείμενο είναι αναψυκτικό και μανταρινάδα
- IS_ORANGE, αν το αντικείμενο είναι αναψυκτικό και πορτοκαλάδα
- IS_COKE, αν το αντικείμενο είναι αναψυκτικό μάρκας Coca-Cola
- IS_FANTA, αν το αντικείμενο είναι αναψυκτικό μάρκας Fanta
- IS_LIME, αν το αντικείμενο είναι αναψυκτικό και γκαζόζα
- IS_CUP, αν το αντικείμενο είναι κούπα και όχι αναψυκτικό.

Στην φόρμα εκπαίδευσης εμφανίζονται όλες οι αριθμητικές τιμές των στόχων στα αντίστοιχα πλαίσια κειμένου. Επίσης οπτικοποιείται η εσωτερική αναπαράσταση της εικόνας του αντικειμένου που έχει τελειοποιηθεί κατόπιν της επεξεργασία από το ΤΝΔ της ενοποίησης των αποκρίσεων των τεχνητών V1S. Οι γραμμές «σκουπίδια» που εμφανιζόταν στην φόρμα οπτικοποίησης της εσωτερικής λειτουργίας δεν εμφανίζονται πλέον στο τελικό πρότυπο.



Χαρακτηριστικά	Λεμονάδα	Μανταρινάδα	Πορτοκαλάδα	Μάρκα: Coca Cola	Μάρκα: Fanta	Γκαζόζα (Lime,Tonic)	Κούπα
0	0,0	0,00	0,00	0,00	0,00	0,00	0,00
0	0,1	0,00	0,00	0,08	0,00	0,00	0,08
0	0,0	0,00	0,00	0,03	0,00	0,00	0,03
0	0,1	0,00	0,00	0,05	0,00	0,00	0,05
0	0,1	0,00	0,00	0,06	0,00	0,00	0,06
0	0,1	0,00	0,00	0,11	0,00	0,00	0,11
1	0,8	1,00	0,79	0,00	0,00	0,00	0,79

10-23: Φόρμα προτύπων πραγματικών αντικειμένων και ανάκλησης ΤΝΔ εξόδου του Horasis.

Μετά την εκπαίδευση του ΤΝΔ εξόδου η φόρμα πλέον μπορεί να εκτελεί ανάκληση πάνω σε οπτικά πρότυπα και εμφανίζει τις τιμές των εξόδων του δικτύου μέσα στα γαλάζια πλαίσια. Στην πρώτη επιτυχημένη εκπαίδευση του ΤΝΔ εξόδου της πρώτης έκδοσης του Horasis, χρησιμοποιήθηκαν 12 οπτικά πρότυπα αντικειμένων με κάποια να ανήκουν σε παραπάνω από μια κλάση. Έτσι για παράδειγμα μια πορτοκαλάδα FANTA είχε τους στόχους IS_ORANGE = 1 (Είναι πορτοκαλάδα), IS_FANTA = 1 (Είναι μάρκα Fanta) και όλους τους υπόλοιπους στόχους στο 0. Η εκπαίδευση λειτούργησε χωρίς να χρειαστεί δεύτερη προσπάθεια με άλλη μορφή ΤΝΔ. Τα 12 πραγματικά αντικείμενα ταξινομήθηκαν στις 7 κλάσεις που υποδείχθηκαν στο σύστημα με σχεδόν μηδαμινό σφάλμα. Ακόμα και στην περίπτωση της κούπας που είχε διαφορετικό όγκο σε σχέση με όλα τα υπόλοιπα ταξινομήθηκε και αυτή με σιγουριά 80% από το σύστημα.

Για την επανάληψη της εκπαίδευσης με περισσότερα πρότυπα αρκεί να διαχωριστούν επιπλέον εγγραφές του PHOTO με IS_TRAINING_PATTERN = 1 και να δοθούν σε αυτές οι τιμές για τις 7 κλάσεις. Έχοντας ολοκληρώσει τον σκοπό του το Horasis δεν χρειάζεται να επεκταθεί παραπάνω σε εννοιολογικές συσχετίσεις και σε πολλά διαφορετικά οπτικά πρότυπα αντικειμένων. Ακόμα ο τρόπος καθορισμού των κλάσεων στις οποίες ανήκει το αντικείμενο αφορά μια εντελώς διαφορετική υλοποίηση λογισμικού και γι αυτό τοποθετούνται οι τιμές στα πεδία IS_xxx χειροκίνητα στην βάση δεδομένων. Οι κλάσεις και οι ονομασίες των αντικειμένων θα μπορούσαν να βρίσκονται ήδη στην βάση από ένα τρίτο λογισμικό και απλώς με το τμήμα I του Horasis θα γινόταν λήψη φωτογραφιών για συσχέτιση τους με υπαρκτές έννοιες.

Παραμετροποίηση της Λειτουργίας

Κλείνοντας την περιγραφή της τελικής λειτουργίας του συστήματος αξίζει να συγκεντρωθούν και να περιγραφούν επιγραμματικά όλες οι παράμετροι που αν τροποποιηθούν, μπορούμε να έχουμε εντελώς διαφορετική λειτουργία του συστήματος και διαφορετικά αποτελέσματα. Έτσι έχουμε

- Οπτικές παραμέτρους, που κωδικοποιούνται έχοντας ως πρόθεμα το γράμμα Ο
- Γεωμετρικές παραμέτρους, που κωδικοποιούνται έχοντας ως πρόθεμα το γράμμα Γ
- Αλγεβρικές παραμέτρους, που κωδικοποιούνται έχοντας ως πρόθεμα το γράμμα Α
- Νευρωνικές παραμέτρους, που κωδικοποιούνται έχοντας ως πρόθεμα το γράμμα Ν

Οπτικές παράμετροι

Ο1. Ρυθμίσεις ατμοσφαιρικού φωτισμού

ΠΡΟΕΠΙΛΟΓΗ: Δυνατός ατμοσφαιρικός φωτισμός με μικρό φωτιστικό σώμα, πίσω από τις κάμερες. Και τα δύο θερμού φωτισμού.

Με αλλαγή του ατμοσφαιρικού φωτισμού μπορούμε να έχουμε φωτεινότερες ή σκοτεινότερες φωτογραφίες του πραγματικού αντικειμένου. Με ύπαρξη παραπάνω φωτισμού στον φόντο μπορεί να επανεμφανιστούν χρωματικά σκουπίδια, οπότε πρέπει να γίνει προσαρμογή του κατωφλίου στην συνάρτηση των τεχνητών ΑC. Με θερμό φωτισμό έχουμε παραπάνω ενεργοποίηση της γραμμής του κίτρινου χρώματος σε σχέση με τον ψυχρό φωτισμό.

Ο2. Ρυθμίσεις web κάμερας Microsoft Lifecam HD-5000

ΠΡΟΕΠΙΛΟΓΗ: Στο συνοδευτικό πρόγραμμα οδήγησης LifeCam Dashboard, χρησιμοποιήθηκαν οι παρακάτω ρυθμίσεις

- TrueColor: ΟΧΙ
- AutoFocus: ΝΑΙ
- Zoom = 0
- Pan = 0
- Tilt = 0
- Brightness = 30
- White Balance = Auto
- Saturation = 80
- Exposure = Auto
- Contrast = 9
- PowerLine Frequency (Anti Flicker) = 60Hz

Η τροποποίηση τους έχει ποικίλες επιδράσεις στην ποιότητα λήψης των οπτικών προτύπων και μπορεί να προκύψει αύξηση των σκουπιδιών της εικόνας και χειρότερος διαχωρισμός μεταξύ φόντου και αντικειμένου.

03. Ανάλυση λήψης φωτογραφιών

ΠΡΟΕΠΙΛΟΓΗ: 1280 x 720 pixel με 24bit χρώματα ως μεγαλύτερη δυνατή για την Microsoft LifeCam HD-5000

Η χρήση καλύτερων συσκευών απεικόνισης θα μπορούσε να συνοδευτεί με ακόμα μεγαλύτερες αναλύσεις. Στην συγκεκριμένη περίπτωση η μείωση της HD ανάλυσης σε κάποια μικρότερη θα οδηγούσε σε απώλεια λεπτομερειών της εικόνας πριν καν την παραγωγή του μέσου όρου από τα τεχνητά ΗC.

04. Αποκοπή τετραγώνου της εικόνας

ΠΡΟΕΠΙΛΟΓΗ: 500 x 500 pixel

Με την χρήση μεγαλύτερου τμήματος της αρχικής εικόνας θα μπορούσε να γίνει λήψη οπτικών προτύπων για μεγαλύτερα αντικείμενα. Για μικρότερο τετράγωνο από 500 x 500 θα χωρούσαν μόνο αντικείμενα μικρότερα σε ύψος από τα αλουμινένια κουτιά αναψυκτικών που χρησιμοποιήθηκαν.

Χρωματικές Παράμετροι

Χ1. Συντελεστές μετατροπής RGB σε κλίμακα του γκρι

ΠΡΟΕΠΙΛΟΓΗ: Τιμές από σύστημα YIQ NTSCQ

- 0.299 * Κόκκινο
- 0.587 * Πράσινο
- 0.114 * Μπλε

Η χρήση άλλων τιμών θα μπορούσε να προσομοιώσει χειρότερα ή καλύτερα την ποσότητα φωτός που έπεσε πάνω στον αισθητήρα του ψηφιακού μέσου απεικόνισης.

Χ2. Αμιγή ανταγωνιστικά χρώματα

ΠΡΟΕΠΙΛΟΓΗ: Με όλες τις τιμές στην κλίμακα από 0 έως 255

- Κόκκινο: $R,G,B = \{255, 0, 0\}$, $H = 255$ ή $H = 0$
- Πράσινο: $R,G,B = \{0, 255, 0\}$, $H = 85$
- Μπλε : $R,G,B = \{0, 255, 0\}$, $H = 169$
- Κίτρινο: $R,G,B = \{0, 255, 0\}$, $H = 43$

Με την αλλαγή των αμιγών χρωμάτων θα τροποποιηθεί άμεσα η απόκριση των τεχνητών BPC καθώς θα μετατοπιστεί πάνω στον τροχό των χρωμάτων το σημείο απόχρωσης, για τα οποία θα εμφανίζουν την μέγιστη διέγερση και την μέγιστη αναστολή.

Γεωμετρικές Παράμετροι

Γ1. Διαστάσεις χάρτη υποδεκτικών πεδίων

ΠΡΟΕΠΙΛΟΓΗ: 17 x 17 pixels

Με μικρότερες διαστάσεις του χάρτη υποδεκτικών πεδίων θα έχουμε μεγαλύτερη λεπτομέρεια στην χρήση της εικόνας. Σε κάθε περίπτωση θα πρέπει να έχουν μονό αριθμό για τις διαστάσεις του ώστε να υπάρχει ένα κεντρικό pixel. Αν αυτές είναι μικρότερες από 13 x 13 pixels, τότε θα αρχίζει να χάνεται η κυκλική μορφή του και δεν θα μπορεί να χωριστεί η περιφέρειας του σε πολλές φέτες. Επίσης με μικρότερες διαστάσεις θα έχουμε περισσότερα ρετινοστοιχεία στον τεχνητό αμφιβληστροειδή και έτσι μεγαλύτερη απαίτηση σε μνήμη. Με μεγαλύτερες διαστάσεις θα χάνεται η σταδιακά λεπτομέρεια της εικόνας.

Γ2. Φέτες χάρτη υποδεκτικών πεδίων



ΠΡΟΕΠΙΛΟΓΗ: Μια κεντρική με διάμετρο 5 pixels και 12 περιφερειακές που χωρίζουν τον υπόλοιπο κυκλικό δίσκο εκτός του κέντρου.

Με λιγότερες φέτες στην περιφέρεια μειώνεται και η λεπτομέρεια της εικόνας. Αυξάνεται όμως η ταχύτητα επεξεργασίας από τα τεχνητά GC καθώς μειώνονται έτσι οι είσοδοι τους. Με περισσότερες φέτες αυξάνεται η λεπτομέρεια της εικόνας, αλλά μειώνεται η ταχύτητα επεξεργασίας και αυξάνεται η απαίτηση σε μνήμη. Θα απαιτηθεί επανεκπαίδευση του ΤΝΔ τύπου bnnkGanglionBichromatic με οποιοδήποτε νέο πλήθος εισόδων.

Γ3. Γεωμετρικό σχήμα ψηφίδας

ΠΡΟΕΠΙΛΟΓΗ: Κανονικό εξάγωνο ένα σχήμα στο οποίο οι ακμές ισαπέχουν από το κέντρο του

Θα μπορούσαν να επιλεγούν και άλλα σχήματα που καλύπτουν το επίπεδο χωρίς να αφήνουν κενά σε μια τέλεια ψηφίδωση. Τα μόνα σχήματα όμως που χωρίς συνδυασμό με άλλα καλύπτουν πλήρως την ψηφίδωση είναι το τρίγωνο το τετράγωνο και το εξάγωνο [112]. Σε κάθε περίπτωση η δημιουργία μιας διαφορετικής ψηφίδωσης θα απαιτούσε έναν νέο αλγόριθμο για την παραγωγή του ARTN, ίσως και αρκετά δυσκολότερο.

Γ4. Παράμετροι ανάπτυξης τεχνητού ρετινοτοπικού χάρτη

ΠΡΟΕΠΙΛΟΓΗ:

- Εξαγωνική ψηφίδα διαμέτρου 16 pixels
- 25 Επίπεδα
- 230 pixels μέγιστη εκκεντρότητα

Αλλάζοντας το μέγεθος της εξαγωνικής ψηφίδας μπορούμε να καλύψουμε την ίδια περιοχή της εικόνας με περισσότερα ή λιγότερα rixel . Κράτωντας σταθερή την παράμετρο Γ1 στα 17x17 pixels έχουμε είτε μεγαλύτερη επικάλυψη με ψηφίδα διαμέτρου μικρότερης των 16 pixels, είτε δημιουργία κενών στην κάλυψη με διάμετρο μεγαλύτερη από 18 pixels. Επίσης σε πολλές περιπτώσεις είναι εφικτή η δημιουργία μια τέλειας ψηφίδωσης με όλες τις διαμέτρους ψηφίδας. Αλλάζοντας τα επίπεδα αυξάνεται η μειώνεται η περιοχή κάλυψης της εικόνας. Αλλάζοντας την μέγιστη εκκεντρότητα μπορούμε να μειώσουμε την διάμετρο του κύκλου του τεχνητού αμφιβληστροειδούς, εφόσον υπάρχουν και τα απαραίτητα επίπεδα για την παραγωγή κυκλικού σχήματος. Αν λόγω των επιπέδων δεν λειτουργεί η ρύθμιση της μέγιστης εκκεντρότητας, τότε έχουμε εξαγωνικό σχήμα για τον τεχνητό αμφιβληστροειδή.

Γ5. Τοποθέτηση κέντρου όρασης

ΠΡΟΕΠΙΛΟΓΗ: Βαρύκεντρο του τετραγώνου εικόνας 500 x 500 pixels στο σημείο $X,Y = \{250,250\}$

Με την τοποθέτηση του κέντρου όρασης σε διαφορετικές συντεταγμένες μπορούμε να φέρουμε ένα αντικείμενο κεντραρισμένο ακριβώς στον τεχνητό αμφιβληστροειδή παράγοντας καλύτερα πρότυπα εκπαίδευσης.

Γ6. Επίπεδα υποδεκτικού πεδίου μικρών τεχνητών κυττάρων V1

ΠΡΟΕΠΙΛΟΓΗ: 1 επίπεδο που συγκεντρώνει 7 rixel

Σε συνδυασμό με μικρότερο μέγεθος rixel θα μπορούσαν να οριστούν περισσότερα επίπεδα αυξάνοντας το πλήθος των rixel. Σε κάθε περίπτωση θα γεννιόταν η απαίτηση επανεκπαίδευσης νέων και πιο σύνθετων ΤΝΔ για PV1S, καθώς θα αυξανόταν το πλήθος των εισόδων τους.

Γ7. Επίπεδα υποδεκτικού πεδίου μεγάλων τεχνητών κυττάρων V1

ΠΡΟΕΠΙΛΟΓΗ: 2 επίπεδα που συγκεντρώνει 19 rixel

Και εδώ ισχύουν αυτά που αναφέρθηκαν στην παράμετρο Γ6 μόνο που τα ΤΝΔ για MV1S θα ήταν πλέον και αργά στην ανάκληση καθώς θα υπήρχαν πάνω από 200 εισοδοί στο κάθε ένα.

Γ8. Μοίρες ευαισθησίας σε κατεύθυνση για τεχνητά κύτταρα V1S

ΠΡΟΕΠΙΛΟΓΗ: Ευαισθησία ανά 30° με 6 τεχνητά κύτταρα V1S για την κάλυψη όλων των πιθανών διευθύνσεων

Με την μείωση των μοιρών μπορούμε να έχουμε μεγαλύτερη λεπτομέρεια ως προς την διεύθυνση μιας μπάρας, αλλά παράλληλα δημιουργούνται περισσότερα τεχνητά κύτταρα V1S στο στρώμα εξόδου των ΤΝΔ. Και εδώ σε κάθε περίπτωση τροποποίησης της ρύθμισης απαιτείται επανεκπαίδευση για 12 τύπους ΤΝΔ που σχετίζονται με την λειτουργία των τεχνητών V1S.

Αλγεβρικές ΠαράμετροιA1. Συνάρτηση που περιγράφει την λειτουργία των τεχνητών ΗC

ΠΡΟΕΠΙΛΟΓΗ: Η συνάρτηση παραγωγής μέσου όρου για n pixels x_1, x_2, \dots, x_n

$$f(X) = \frac{1}{n} \sum_{i=1}^n x_i$$

Αλλάζοντας αυτήν την συνάρτηση θα πρέπει να δημιουργηθεί ένα εντελώς νέο είδος τεχνητών ΗC, ίσως ως στρώμα εξόδου σε ένα ΤΝΔ που θα εκπαιδευτεί, ώστε να λειτουργεί με την νέα συνάρτηση.

A2. Βαρύτητα α της τιμής τεχνητού ΗC σε σχέση με τιμή κεντρικού τεχνητού φωτοϋποδοχέα φέτας

ΠΡΟΕΠΙΛΟΓΗ: $\alpha = 0.6$ οπότε συμμετέχει τιμή του κεντρικού φωτοϋποδοχέα κατά 60% και η τιμή του τεχνητού ΗC (μέσος όρος φέτας) κατά 40%.

Αλλάζοντας την αναλογία εις βάρος του τεχνητού ΗC, κάνουμε όλο το μοντέλο να στηρίζεται στις τιμές συγκεκριμένων pixel περισσότερο από ότι στον μέσο όρο της περιοχής που τα περιβάλλει. Από την άλλη αν $\alpha = 1$ τότε κάνουμε το μοντέλο να λειτουργεί αποκλειστικά με τους κεντρικούς φωτοϋποδοχείς της κάθε φέτας, αγνοώντας τους μέσους όρους των pixels που ανήκουν σε αυτές.

A3. Χαρακτηριστικά συνάρτησης απόστασης από αμιγές χρώμα, για το κόκκινο

ΠΡΟΕΠΙΛΟΓΗ: Για το πρώτο μισό $\mu = 255$ και $\sigma = 22$ και για το δεύτερο μισό $\mu = 0$ και $\sigma = 22$, όπου μ ο μέσος και σ η απόκλιση στην συνάρτηση κανονικής (gauss) κατανομής

Ο μέσος παίρνει την τιμή του από την απόχρωση H (σε κλιμακα 0 έως 255) του αμιγούς κόκκινου και αλλαγή των μ ισοδυναμεί με αλλαγή της παράμετρου X_2 για το αμιγές κόκκινο χρώμα. Αν τροποποιηθεί το σ τότε μειώνεται ή αυξάνεται το εύρος πάνω στον τροχό των χρωμάτων για το οποίο θα διεγείρεται το R-BPC.

A4. Χαρακτηριστικά συνάρτησης απόστασης από αμιγές χρώμα, για το πράσινο

ΠΡΟΕΠΙΛΟΓΗ: $\mu = 85$ και $\sigma = 22$, όπου μ ο μέσος και σ η απόκλιση στην συνάρτηση κανονικής (gauss) κατανομής

Ο μέσος παίρνει την τιμή του από την απόχρωση H (σε κλίμακα 0 έως 255) του αμιγούς πράσινου και αλλαγή του μ ισοδυναμεί με αλλαγή της παράμετρου X_2 για το αμιγές πράσινο χρώμα. Αν τροποποιηθεί το σ τότε μειώνεται ή αυξάνεται το εύρος πάνω στον τροχό των χρωμάτων για το οποίο θα διεγείρεται το B-BPC. Αν το σ για το πράσινο διαφέρει από το σ για το κόκκινο τότε γέρνουμε την πλάστιγα του χρωματικού ανταγωνισμού προς το ένα από τα δύο χρώματα, καθώς θα παράγονται περισσότερες αποκρίσεις γι αυτό με το μεγαλύτερο σ .

A5. Χαρακτηριστικά συνάρτησης απόστασης από αμιγές χρώμα, για το κίτρινο

ΠΡΟΕΠΙΛΟΓΗ: $\mu = 43$ και $\sigma = 15$, όπου μ ο μέσος και σ η απόκλιση στην συνάρτηση κανονικής (gauss) κατανομής

Ο μέσος παίρνει την τιμή του από την απόχρωση H (σε κλίμακα 0 έως 255) του αμιγούς κίτρινου και αλλαγή του μ ισοδυναμεί με αλλαγή της παράμετρου X_2 για το αμιγές κίτρινο χρώμα. Αν τροποποιηθεί το σ τότε μειώνεται ή αυξάνεται το εύρος πάνω στον τροχό των χρωμάτων για το οποίο θα διεγείρεται το Y-BPC. Μεγαλύτερο σ από 15 αυξάνει και το όριο της απροσδιοριστίας για αποχρώσεις μεταξύ του κόκκινου και πράσινου καθώς χρώματα που πριν δημιουργούσαν ενεργοποιήσεις στα R-BPC και G-BPC τώρα θα δημιουργούν και για Y-BPC

A6. Χαρακτηριστικά συνάρτησης απόστασης από αμιγές χρώμα, για το μπλε

ΠΡΟΕΠΙΛΟΓΗ: $\mu = 169$ και $\sigma = 26$, όπου μ ο μέσος και σ η απόκλιση στην συνάρτηση κανονικής (gauss) κατανομής

Ο μέσος παίρνει την τιμή του από την απόχρωση H (σε κλίμακα 0 έως 255) του αμιγούς μπλε και αλλαγή του μ ισοδυναμεί με αλλαγή της παράμετρου X_2 για το αμιγές μπλε χρώμα. Αν τροποποιηθεί το σ τότε μειώνεται ή αυξάνεται το εύρος πάνω στον τροχό των χρωμάτων για το οποίο θα διεγείρεται το B-BPC.

A7. Χαρακτηριστικά συνάρτησης απόστασης από απόλυτο λευκό

ΠΡΟΕΠΙΛΟΓΗ: $\mu = 255$ και $\sigma = 96$, όπου μ ο μέσος και σ η απόκλιση στην συνάρτηση κανονικής (gauss) κατανομής

Ο μέσος παίρνει την τιμή του από την μέγιστη τιμή της κλίμακας του γκρι 0 έως 255 που αντιστοιχεί στο απόλυτα λευκό χρώμα. Αν τροποποιηθεί το σ τότε αυξάνεται ή μειώνεται η διέγερση του L-BPC στις διάφορες αποχρώσεις του γκρι. Το σ δεν μπορεί να γίνει πολύ μικρό καθώς μπορεί να προκύψουν αποχρώσεις του γκρι για τις οποίες δεν θα ενεργοποιείται ούτε το L-BPC ούτε το D-BPC.

A8. Χαρακτηριστικά συνάρτησης απόστασης από απόλυτο μαύρο

ΠΡΟΕΠΙΛΟΓΗ: $\mu = 0$ και $\sigma = 22$, όπου μ ο μέσος και σ η απόκλιση στην συνάρτηση κανονικής (gauss) κατανομής

Ο μέσος παίρνει την τιμή του από την ελάχιστη τιμή της κλίμακας του γκρι 0 έως 255 που αντιστοιχεί στο απόλυτα μαύρο χρώμα. Αν τροποποιηθεί το σ τότε αυξάνεται η μειώνεται η διέγερση του D-BPC στις διάφορες αποχρώσεις του γκρι. Το σ δεν μπορεί να γίνει πολύ μεγάλο καθώς αποχρώσεις του γκρι που είναι ευδιάκριτα φωτεινές θα θεωρούνται κοντά στο μαύρο. Η θέση του σ σε αυτήν την ρύθμιση καθορίζει το σημείο πάνω στην κλίμακα του γκρι, στο οποίο χωρίζονται κατά έναν ιδιότυπο τρόπο, οι αποχρώσεις του μαύρου από τις αποχρώσεις του λευκού.

A9. Διάμετροι κέντρου και περιφέρειας για υπολογισμούς στόχων τεχνητού GC

ΠΡΟΕΠΙΛΟΓΗ: Διάμετρος περιφέρειας $D = 61$ και διάμετρος κέντρου $C = 5$

Αλλαγή των διαμέτρων οδηγεί σε διαφορετικές σταθερές EN και IN που χρησιμοποιούνται στους στόχους εκπαίδευσης των τεχνητών GC και απαιτεί επανεκπαίδευση των ΤΝΔ-GC.

A10. Χαρακτηριστικά συνάρτησης διέγερσης κέντρου τεχνητού GC

ΠΡΟΕΠΙΛΟΓΗ: $D = 61$, $\mu \approx D/2 = 30$ και $\sigma = C = 5$, όπου μ ο μέσος και σ η απόκλιση στην συνάρτηση κανονικής (gauss) κατανομής

Μπορούμε να αλλάξουμε ελεύθερα το D καθώς η μορφή της συνάρτησης δεν θα επηρεαστεί. Αλλάζοντας την διάμετρο C του κέντρου αλλάζουμε την ευαισθησία των τεχνητών GC σε μικρές ή μεγάλες φωτεινές κηλίδες. Αλλαγή αυτής της ρύθμισης οδηγεί σε διαφορετικές σταθερές EN και IN για τους στόχους εκπαίδευσης των τεχνητών GC και απαιτεί επανεκπαίδευση των ΤΝΔ-GC.

A11. Χαρακτηριστικά συνάρτησης αναστολής περιφέρειας τεχνητού GC

ΠΡΟΕΠΙΛΟΓΗ: $D = 61$, $\mu \approx D/2 = 30$ και $\sigma \approx (12/5) C = 12$

Το D πρέπει να ισούται πάνω με το D της ρύθμισης A9. Αν αλλάξουμε το σ η την αναλογία του με την διάμετρο του κέντρου του κυκλικού δίσκου, μπορούμε να πειραματιστούμε πάνω στο μοντέλο κέντρου-περιφέρειας των τεχνητών GC. Αλλαγή αυτής της ρύθμισης οδηγεί σε διαφορετικές σταθερές EN και IN για τους στόχους εκπαίδευσης των τεχνητών GC και απαιτεί επανεκπαίδευση των ΤΝΔ-GC.

A12. Αναλογία μέγιστης διέγερσης σε σχέση με μέγιστη αναστολή

ΠΡΟΕΠΙΛΟΓΗ: $s = 2$

Αλλάζοντας αυτήν την αναλογία προς τα κάτω μειώνουμε την διέγερση σε σχέση με την αναστολή. Από την άλλη αν αυξηθεί το s τότε η διέγερση που παράγει το κέντρο για ένα βασικό χρώμα, είναι πολύ μεγαλύτερη από την αντίστοιχη αναστολή που παράγει στο κέντρο για το ανταγωνιστικό του.

A13. Συνάρτηση ενεργοποίησης των τεχνητών AC

ΠΡΟΕΠΙΛΟΓΗ: Η συνάρτηση κατωφλίου $f(x) = x$, αν $L \geq V$, $f(x) = 0$, αν $L < V$

Με την εισαγωγή διαφορετικού τύπου συνάρτησης μπορεί να αλλαχθεί ο τρόπος που μηδενίζεται το σήμα των τεχνητών BPC, πριν την είσοδο του στα τεχνητά GC. Μπορούμε να αλλάξουμε δηλαδή τον τρόπο που ρυθμίζεται η ροή της οπτικής πληροφορίας από το υποσύστημα των διαύλων χρωματικού ανταγωνισμού, προς το υποσύστημα των τεχνητών GC.

A14. Τιμή κατωφλίου συνάρτησης ενεργοποίησης των τεχνητών AC

ΠΡΟΕΠΙΛΟΓΗ: $V = 0.05$ ή επίπεδο φωτισμού φέτας στο 5%.

Αν αυξηθεί η φωτεινότητα της φωτογραφίας κατόπιν τροποποίησης μιας από τις οπτικές ρυθμίσεις, τότε το μαύρο φόντο της εικόνας θα αρχίσει πάλι να εμφανίζει αποκρίσεις στα χρωματικά τεχνητά BPC. Ανεβάζοντας το κατώφλι V μπορούμε να θέσουμε το νέο ποσοστό φωτισμού κάτω από το οποίο θα μηδενίζεται η είσοδος στα τεχνητά GC, προερχόμενη από τεχνητά BPC μιας φέτας του υποδεκτικού τους πεδίου. Δεν απαιτείται επανεκπαίδευση των τεχνητών GC με αλλαγή αυτής της ρύθμισης

A15. Ποσοστά συμμετοχής στην ενοποιημένη απόκριση για μικρά και μεγάλα V1S

ΠΡΟΕΠΙΛΟΓΗ: 40% από την απόκριση των μεγάλων MV1S και 60% απόκριση μικρού PV1S.

Αλλαγή των ποσοστών τροποποιεί την εσωτερική αναπαράσταση της εικόνας του αντικειμένου και απαιτεί επανεκπαίδευση του ΤΝΔ εξόδου του Horasis με τα νέο πρότυπα. Επίσης για την συγκεκριμένη περιοχή δεν πρέπει κάποιο από τα δύο να δίνει ισχυρότερη απόκριση καθώς η ύπαρξη μια μακριάς μπάρας ενεργοποιεί και τον MV1S και τον PV1S τεχνητό νευρώνα και αντιστοιχεί στο 100% της απόκρισης. Δεν μπορεί η ενεργοποίηση μόνο ενός από τους δύο να παράγει απόκριση μεγαλύτερη έστω από 70% καθώς η μπάρα ή δεν είναι μακριά γιατί δεν ενεργοποιεί τον MV1S ή δεν είναι εντελώς ευθύγραμμη αν δεν ενεργοποιεί τον μικρό PV1S.

A16. Κατώφλια παραγωγής διεγερτικής εξόδου και ανασταλτικής εξόδου σε τεχνητούς V1S

ΠΡΟΕΠΙΛΟΓΗ: $E = 0.85$ κατωφλι για παραγωγή διεγερσης και $I = 0.85$ κατώφλι για παραγωγή αναστολής

Έχοντας διαφορετική τιμή για τα E και I , ρυθμίζουμε την μια απόκριση να λειτουργεί σε βάρος της άλλης αγνοώντας κάποιες αποκρίσεις. Κατ' αυτόν τον τρόπο γέρνει η πλάστιγγα του χρωματικού ανταγωνισμού προς ένα από τα δύο χρώματα. Με μικρά E και I έχουμε την εμφάνιση περισσότερων σκουπιδιών στην εσωτερική αναπαράσταση της εικόνας του αντικειμένου, ενώ με μεγάλα έχουμε απώλεια πληροφορίας. Απαιτείται και εδώ επανεκπαίδευση του ΤΝΔ κατηγοριοποίησης του Horasis, αν τροποποιηθεί η ρύθμιση καθώς και πάλι έχουμε νέα εσωτερική παράσταση της πληροφορίας

Νευρωνικές Παράμετροι

N1. Συνάρτηση ενεργοποίησης τεχνητών νευρώνων στα MLP δίκτυα

ΠΡΟΕΠΙΛΟΓΗ:

$$f(u) = a_L + \left(\frac{\alpha_H - \alpha_L}{1 + be^{-d(u-c)}} \right)$$

$\alpha_L = 0$, $\alpha_H = 1$, $b = 1$, $c = 0$, $d = 1$, οπότε απλοποιείται σε

$$f(u) = \left(\frac{1}{1 + e^{-u}} \right),$$

Μπορούμε να αλλάξουμε την παράμετρο d και να έχουμε τα αποτελέσματα που περιγράφησαν στο κεφάλαιο 3. Γενικά οι αλλαγή των παραμέτρων της γενικής λογιστικής συνάρτησης, θα δημιουργήσει διαφορετικές ενεργοποιήσεις στην έξοδο των νευρώνων, αρκεί να κρατηθούν οι δύο ασύμπτωτες $\alpha_L = 0$, $\alpha_H = 1$ για να συγκρατηθεί το σύνολο τιμών στο $[0,1] \in \mathbb{R}$. Σε κάθε περίπτωση πρέπει να υπολογιστεί η παράγωγος της νέας συνάρτησης και βέβαια να επανεκπαιδευτούν όλοι οι τύποι ΤΝΔ.

N2. Βήμα εκπαίδευσης και ορμή

ΠΡΟΕΠΙΛΟΓΗ: Το βήμα είναι σε κάθε περίπτωση $0 < \beta < 0,1$ και η χρησιμοποιείται πάντα και η ορμή με $0 < \mu < 0.8$

Το βήμα εκπαίδευσης και η ορμή διαφοροποιούνται σε κάθε στρώμα τεχνητών νευρώνων στα εκπαιδευόμενα ΤΝΔ του Horasis κατά περίπτωση. Καθορίστηκαν στις τελικές τιμές μετά από δοκιμή και απόρριψη (trial and error)

N3. Αριθμητική ακρίβεια τιμών και συναπτικών βαρών

ΠΡΟΕΠΙΛΟΓΗ: Πραγματικοί αριθμοί διπλής ακριβείας, τύπος δεδομένων Delphi **Double**, των 64bits με 15-16 σημαντικά ψηφία και εύρος προσημασμένων τιμών από $5.0 \times 10^{-324} \dots 1.7 \times 10^{308}$

Τα 15 σημαντικά ψηφία καλύπτουν την εφαρμογή μας χωρίς να δημιουργούν επιπρόσθετη ανάγκη σε μνήμη όπως ο αμέσως μεγαλύτερος τύπος Extended που έχει ακρίβεια 20 σημαντικών ψηφίων και 80bits. Η χρήση τύπου single με ακρίβεια 7-8 δεκαδικών ψηφίων δεν ενδείκνυται ενώ θα μπορούσε να γίνει χρήση Extended μόνο σε ένα σύστημα με μεγαλύτερη επεξεργαστική ισχύ, ώστε να αντισταθμιστεί η καθυστέρηση που θα προέκυπτε.

N4. Βάρη τεχνητών ΗC

ΠΡΟΕΠΙΛΟΓΗ: Για n pixels το κάθε βάρος ορίζεται σε $W_i = \frac{1}{n}$.

Αλλάζοντας τα προκαθορισμένα βάρη που παράγουν τον μέσο όρο, στην ουσία αλλάζουμε την προεπιλεγμένη συνάρτηση της ρύθμισης A1, έχοντας πάλι τους ίδιους νευρώνες. Αν κάποια βάρη είναι μεγαλύτερα από τα άλλα τηρώντας το άθροισμα $\mu = \sum_{i=1}^n W_i = 1$, τότε κάποια pixels θα έχουν μεγαλύτερη σημασία από τα άλλα μέσα σε μια φέτα του τεχνητού αμφιβληστροειδούς. Αν το σύνολο των βαρών $\mu < 1$ για μια χρωματική συνιστώσα, τότε θα έχουμε μικρότερη συμμετοχή αυτής της συνιστώσας για την φέτα του τεχνητού αμφιβληστροειδούς.

11. Συμπεράσματα και Προτάσεις

Επιτεύγματα Του Horasis

Χρήση αντικειμενοστραφούς σχεδιασμού και προγραμματισμού στα τεχνητά νευρωνικά δίκτυα.

Ξεκινώντας την αναφορά σε αυτά που επιτεύχθηκαν μέσα από την παρούσα πτυχιακή εργασία, θα αναφερθούμε πρωτίστως στα θετικά που δημιούργησε η χρήση του αντικειμενοστραφούς προγραμματισμού (object oriented programming) για την υλοποίηση των τεχνητών νευρωνικών δικτύων. Το κυριότερο σημείο είναι ότι ένα πάρα πολύ σύνθετο έργο, όπως το λογισμικό Horasis, μπόρεσε να τεμαχιστεί σε μια πληθώρα από μικρά και διαχειρίσιμα κομμάτια κώδικα.

Με αυτήν την λογική πρώτα από όλα κατασκευάστηκε όλη η υποδομή για αντικειμενοστραφείς τεχνητούς νευρώνες και αντικειμενοστραφή ΤΝΔ και κατόπιν όλη η υποδομή για την εκπαίδευση τους με τον αλγόριθμο back-propagation. Αφήνοντας στην άκρη την υπόλοιπη υλοποίηση, κατασκευάστηκε και δοκιμάστηκε αυτόνομα επιλύοντας ένα γνωστό πρόβλημα ταξινόμησης που είχε λυθεί στο παρελθόν, στα πλαίσια του εργαστηριακού μαθήματος των τεχνητών νευρωνικών δικτύων. Το πρόβλημα αυτό που λεγόταν Quadrupeds είχε λυθεί στα πλαίσια μιας εργασίας εξαμήνου με μια πρώιμη μορφή των αντικειμενοστραφών ΤΝΔ του Horasis, των οποίων ο κεντρικός σχεδιασμός επαναχρησιμοποιήθηκε εν μέρει. Στην νέα τελική υλοποίηση, επιλέγοντας ακριβώς τις ίδιες συνθήκες εκπαίδευσης, δοκιμάστηκε αν το νέο αντικειμενοστραφές ΤΝΔ θα έδινε ακριβώς τα ίδια αποτελέσματα με το λυμένο πρόβλημα. Όταν όλα λειτούργησαν σωστά το κομμάτι αυτό του πηγαιού κώδικα θεωρήθηκε ένα κλειστό κουτί που λειτουργεί και μπορεί να επαναχρησιμοποιηθεί μέσω μιας διασύνδεσης μεθόδων και ιδιοτήτων που έχει. Δηλαδή όλη η λογική ήταν εμψυλευμένη (encapsulated) και δεν χρειαζόταν να υπάρχει διπλότυπη σε κανένα άλλο σημείο του Horasis, ούτε καν ελαφρώς τροποποιημένη.

Στην συνέχεια και με την ίδια φιλοσοφία κάθε κομμάτι λογισμικού αντιμετωπιζόταν αυτόνομα και επαναχρησιμοποιούνταν μέσα στα πλαίσια της υποδομής του Horasis με αποτέλεσμα να υπάρξει μικρή σύζευξη (coupling). Κάθε αναγκαία βελτίωση και τροποποίηση των κομματιών του κώδικα δεν χρειαζόταν να επαναληφθεί σε πολλά σημεία. Η όλη μέθοδος είχε τρία βασικά οφέλη. Πρώτο η αύξηση της ποιότητας καθώς μόνο μετά από δοκιμασμένα «κλειστά κουτιά» κώδικα προχωρούσε η υλοποίηση παρακάτω, αφού ολοκληρωνόταν οι τυχόν βελτιώσεις πάνω σε αυτά. Δεύτερον υπήρξε μείωση της πολυπλοκότητας του κώδικα οπότε η διαχείριση των αναγκιών μεταβολών σε αυτόν ήταν πιο εύκολη. Τρίτον που συνδέεται και με το προηγούμενο είναι η μείωση του απαιτούμενου χρόνου για την παραγωγή της τελικής μορφής του κώδικα και της πρώτης έκδοσης της εφαρμογής Horasis.

Επιπρόσθετα η χρήση του αντικειμενοστραφούς σχεδιασμού στην πτυχιακή εργασία, βοήθησε στην δημιουργία της βιώσιμης και ευέλικτης αρχιτεκτονικής του Horasis. Με αυτήν υλοποιήθηκε υψηλή συνοχή (cohesion) μεταξύ των υποσυστημάτων της εφαρμογής, τα οποία περιέχουν συναφείς ομάδες από αντικείμενα. Κάθε τμήμα λογισμικού έχει μόνο τα αντικείμενα που στην βιολογική πραγματικότητα βρίσκονται σε στενή συνεργασία μεταξύ τους. Αυτό με την σειρά του ένα αποτελεί κλειστό κουτί από την πλευρά των υπολοίπων τμημάτων. Για παράδειγμα ο τεχνητός οπτικός φλοιός δεν έχει καμία ορατότητα πάνω στο αρχικό bitmap και τις χρωματικές συνιστώσες R,G,B της εικόνας. Ένα δεύτερο παράδειγμα είναι ότι το αρχικό bitmap και τα R,G,B

που διαβάζουν τα αντικείμενα του υποσυστήματος των τεχνητών φωτοϋποδοχέων, δεν είναι ορατά στο υποσύστημα των τεχνητών γαγγλιακών κυττάρων άσχετα με το αν βρίσκονται μέσα στο ίδιο τμήμα λογισμικού. Το τελευταίο αντικατοπτρίζει την πραγματικότητα που υπάρχει στην οργάνωση των νευρώνων του αμφιβληστροειδούς στο ανθρώπινο μάτι. Εκτός από την ποιότητα, την μείωση πολυπλοκότητας και την ταχύτητα στην υλοποίηση υπάρχει και ένα τέταρτο όφελος που είναι η επιτυχημένη μεταφορά του βιολογικού μοντέλου της όρασης ως προσομοίωση.

Με όλα τα παραπάνω μπορούμε να πούμε ότι έχει επιτευχθεί ένας από τους τέσσερις στόχους της εργασίας, να επιδειχθεί ότι η χρήση αντικειμενοστραφούς τεχνολογίας με τεχνητά νευρωνικά δίκτυα έχει ουσιαστικά και απτά οφέλη.

Μεταφορά του βιολογικού μοντέλου της όρασης σε λογισμικό τεχνητών νευρωνικών δικτύων.

Όπως είπαμε με τον αντικειμενοστραφή σχεδιασμό, μπόρεσε να μεταφερθεί το μοντέλο της βιολογικής όρασης. Αυτός ήταν εξ αρχής ακόμα ένας στόχος της πτυχιακής εργασίας, καθώς η σκέψη ήταν να διερευνήσει την προοπτική χρήσης των ΤΝΔ με τρόπο που θυμίζει την βιομηχανική. Δηλαδή όχι απλώς να χρησιμοποιούνται ως εργαλείο για την παραγωγή λύσης σε θεωρητικά μαθηματικά μοντέλα, αλλά να αποτελεί το ίδιο το δίκτυο την μεταφορά του βιολογικού μοντέλου.

Στην πρώτη περίπτωση για να προσομοιώσουμε μια λειτουργία για μια συμπεριφορά του πραγματικού κόσμου δημιουργείται πρώτα το μοντέλο με μαθηματική περιγραφή. Κατόπιν το ΤΝΔ είναι το εργαλείο με το οποίο τρέχει το μοντέλο και παράγει τα επιθυμητά αποτελέσματα. Αυτή η μέθοδος είναι πιο άρτια από επιστημονική πλευρά, αλλά απαιτεί ισχυρή θεωρητική διατύπωση και απόδειξη μέσα στα πλαίσια που ορίζουν τα μαθηματικά. Καθώς η εργασία ανήκει στην επιστήμη των υπολογιστών μια τέτοια προσπάθεια θα ήταν πολύ δύσκολο και ίσως και εκτός θέματος.

Ως εναλλακτική και απλούστερη προσέγγιση χρησιμοποιούμε ως μοντέλο την περιγραφή με όρους βιολογίας, που προέκυψε από πειράματα που κατέγραψαν μια λειτουργία ή συμπεριφορά στον πραγματικό κόσμο. Με την χρήση της επιστήμης των υπολογιστών και μέσα στο πλαίσιο του συνδεσιασμού (connectionism), δημιουργούνται ένα προς ένα διαφορετικά ΤΝΔ, που προσπαθούν να προσομοιώσουν αυτό το μοντέλο. Δεν θα μπορούσε και εδώ να μην χρησιμοποιηθούν τα μαθηματικά, αλλά πλέον χρειάζονται σαφώς απλούστερα τα οποία είδαμε μέσα στην εργασία. Αυτά ενώνονται μεταξύ τους φτιάχνοντας στην ουσία ένα μεγάλο ΤΝΔ το οποίο πλέον, αποτελεί το μαθηματικό μοντέλο. Η παραγωγή της λύσης είναι μια εντελώς καινούργια διαδικασία που δεν μπορούσε να περιγραφεί εξ' αρχής με έναν σαφώς ντετερμινιστικό τρόπο. Παρόλατα από την πειραματική λύση που παράγει το ΤΝΔ, ίσως μπορεί να εξαχθεί και μια τυπική μαθηματική περιγραφή.

Με τον παραπάνω τρόπο η εργασία έφτασε τον στόχο της δηλαδή να φτάσει σε ένα λειτουργικό μοντέλο λογισμικού, αντίστοιχο ως ένα βαθμό με την πραγματική όραση. Η σκέψη που ήδη αναφέρθηκε «Παρατηρώ κάτι πως γίνεται στην φύση, ώστε να φτιάξω κάτι τεχνητό που λειτουργεί με τον ίδιο τρόπο, προσπαθώντας να έχω το ίδιο αποτέλεσμα» είναι ο βασικός πυλώνας που δημιούργησε αυτόν τον στόχο. Η φύση μέσα από εκατομμύρια χρόνια εξέλιξης, που περιλάμβανε μεταλλάξεις και φυσική επιλογή, δημιούργησε ένα λειτουργικό μοντέλο

όρασης τέλεια προσαρμοσμένο για ένα νοήμον όν, τον άνθρωπο. Αυτό σημαίνει ότι έχει δημιουργήσει τα μαθηματικά με τα οποία επεξεργάζεται μια πληροφορία εισόδου από φωτόνια παράγοντας στην έξοδο μια πληροφορία ικανή να αξιοποιηθεί από μια ανώτερη σκέψη και νοημοσύνη. Με την απλοϊκή προσέγγιση σε αυτήν την πτυχιακή εργασία, απλώς σημειώνεται πόσο σημαντική πρόκληση για πολλές επιστήμες είναι να φτάσουμε στην κατανόηση και αναπαραγωγή αυτών των αποτελεσμάτων.

Ολοκληρωμένος σχεδιασμός τεχνητής όρασης και μεταφορά του σε ένα πλήρως λειτουργικό λογισμικό.

Η πτυχιακή εργασία και το σύστημα Horasis που προέκυψε από αυτήν, αποτελούν έναν ολοκληρωμένο σχεδιασμό τεχνητής όρασης. Σίγουρα θα μπορούσε να εμπλουτιστεί στο μέλλον στηριζόμενος πάντα στην προσπάθεια της βιομηχανικής. Το σημαντικό είναι ότι παρέχεται για τον σχεδιασμό αυτό η πρώτη πλήρως λειτουργική έκδοση ενός συστήματος λογισμικού, φτιαγμένο με απλά συστατικά όπως συνηθισμένες γλώσσες προγραμματισμού και υλικό προσωπικών υπολογιστών.

Το ίδιο το λογισμικό δεν αποτελεί το επίτευγμα, αλλά το παρακάτω γεγονός. Με την σωστή χρήση των εργαλείων της επιστήμης των υπολογιστών, είναι εφικτό μια σύνθετη ιδέα τεχνητής όρασης, με άγνωστο εκ των προτέρων αποτέλεσμα, να οδηγηθεί γρήγορα σε μια λειτουργική υλοποίηση. Από την πρώτη έκδοση που υλοποιήθηκε μέσα στα χρονικά πλαίσια μιας πτυχιακής εργασίας και ως ατομικό έργο, φάνηκε ότι το σύστημα τεχνητής όρασης Horasis, που πρότεινε η εργασία έχει κάποια απτά αποτελέσματα. Το γεγονός αυτό δίνει ελπίδες ότι μπορεί να αποτελέσει στο μέλλον, ένα θέμα για πιο επισταμένη έρευνα.

Λειτουργία αναγνώρισης οπτικών αντικειμένων που στηρίζεται στο χρώμα.

Στην πτυχιακή εργασία και σχετικά με τον στόχο της μίμησης της βιολογικής πραγματικότητας, αποφασίστηκε να χρησιμοποιηθεί χρώμα για την αναγνώριση των πραγματικών αντικειμένων. Όπως η όραση του ανθρώπου είναι τριχρωματική, έτσι και στο Horasis θα έπρεπε να αξιοποιηθούν τα R,G,B και όχι μόνο η κλίμακα του γκρι που θα έκανε τα πράγματα αρκετά πιο εύκολα.

Η επιλογή αυτή οδήγησε μέχρι το τέλος την εργασία στον σωστό δρόμο, καθώς στην πορεία πολλά πράγματα προσαρμόστηκαν με γνώμονα την ύπαρξη χρωμάτων. Σε κάθε προσαρμογή το μοντέλο πλησίαζε όλο και πιο κοντά σε αυτό της πραγματικής όρασης, στην οποία λειτουργούν οι δίαυλοι χρωματικού ανταγωνισμού. Γι αυτό στο τέλος η εσωτερική αναπαράσταση της οπτικής πληροφορίας περιγράφει το πραγματικό αντικείμενο ως διάνυσμα τιμών που αναπαριστούν τους τρεις χρωματικούς ανταγωνισμούς. Ως προς την εκπαίδευση του δικτύου με back propagation και παράλληλα την ύπαρξη έγχρωμων φωτογραφιών για τα αρχικά πρότυπα, το Horasis διαφοροποιείται σε σχέση με άλλες υλοποιήσεις. Στην τελική εκπαίδευση χρησιμοποιήθηκαν παρόμοια πρότυπα που διέφεραν ελάχιστα ως προς το χρώμα, όπως το κουτί

της μανταρινάδας και το κουτί της πορτοκαλάδας της ίδιας μάρκας, τα οποία διέφεραν ως προς την απόχρωση του πορτοκαλί.

Το γεγονός ότι ακόμα και από την πρώτη απλή χρήση του συστήματος ήταν ικανό να αναγνωρίσει έγχρωμες φωτογραφίες των πραγματικών αντικειμένων, είναι ακόμα μια κορυφή στην οποία μπόρεσε να φτάσει. Με αυτόν τον τρόπο εκπληρώνεται και ο τελικός στόχος της πτυχιακής εργασίας που είναι η αναγνώριση των οπτικών προτύπων ορισμένων πραγματικών αντικειμένων του χώρου.

Μη αλγοριθμικός τρόπος εξαγωγής πληροφοριών σε σχέση με σχήμα.

Ένα δεύτερο χαρακτηριστικό εκτός από το χρώμα με το οποίο λειτούργησε και αναγνώρισε τα πραγματικά αντικείμενα το Horasis, είναι και το σχήμα. Με αυτόν τον τρόπο πρότυπα στα οποία επικρατούσαν τα ίδια χρώματα έχοντας διαφορετικά λογότυπα και σχηματισμούς, δεν είχαν την ίδια εσωτερική αναπαράσταση. Δεν υπάρχει δηλαδή αποκλειστικά στήριξη στην ποσότητα των χρωμάτων που αποτελούσαν την εικόνα, κατά την οποία απλώς θα δημιουργούνταν ένα ιστόγραμμα των χρωμάτων. Πλέον έχει σημασία η θέση του χρώματος μέσα στην φωτογραφία και οι σχηματισμοί που αυτό δημιουργεί σε ένα κβαντισμένο κομμάτι της.

Υπάρχουν πολλοί τρόποι για εξαγωγή σχηματικών πληροφοριών στην θεωρία της επιστήμης των υπολογιστών, με τον πιο απλό την ανίχνευση ακμών (edge detection). Τέτοιοι τρόποι στηρίζονται στην αλγοριθμική και τα μαθηματικά, εξάγοντας σχήμα από την εικόνα με μια διαδικαστική (procedural) λειτουργία. Στην εργασία και στο σύστημα Horasis τα χαρακτηριστικά σχήματος είναι μια πληροφορία που παράγεται μέσω ανάκλησης, σε ήδη εκπαιδευμένα ΤΝΔ. Θα λέγαμε ότι σε αντίθεση με τον αλγοριθμικό-μαθηματικό τρόπο, εδώ έχουμε έναν καθαρά «νευρωνικό» τρόπο για εξαγωγή των σχημάτων της εικόνας.

Με την οργάνωση κέντρου-περιφέρειας των τεχνητών γαγγλιακών κυττάρων και τον προσανατολισμό κατευθύνσεων των τεχνητών V1 κυττάρων έχουμε μια ιδιόμορφη ανίχνευση ακμών. Με τον τεχνητό ρετινοτοπικό χάρτη προσδίδονται χαρακτηριστικά θέσης στο κάθε σχήμα που ανιχνεύεται στην εικόνα. Όλα αυτά γίνονται με την παράλληλη λειτουργία ενός μεγάλου πλήθους από απλούς τεχνητούς νευρώνες McCulloch-Pits που αντικαθιστούν τις αλγοριθμικές διαδικασίες.

Δημιουργία οπτικών προτύπων από καρτέ ροής βίντεο σε μη ιδανικές συνθήκες.

Όπως γνωρίσαμε στο τμήμα του τεχνητού βολβού, ο σχεδιασμός λειτουργεί με πολύ απλό υλικό όπως κάμερες web, και με λήψη μέσω DirectShow ενός καρτέ από μια ροή βίντεο. Βλέπουμε δηλαδή ότι οι συνθήκες παραγωγής των οπτικών προτύπων δεν είναι «αποστειρωμένες», όπου υπάρχει ιδανική φωτογράφιση των αντικειμένων, χωρίς θόρυβο και με την υψηλότερη δυνατή ευκρίνεια.

Θα μπορούσε να είχαν χρησιμοποιηθεί ψηφιακές φωτογραφικές μηχανές που θα παρήγαγαν μεγάλες αναλύσεις της τάξεων των πολλών MegaPixel. Παράλληλα μια ψηφιακή μηχανή θα τελειοποιούσε τον φωτισμό της φωτογραφίας και θα αναπαριστούσε καλύτερα τα χρώματα, κάνοντας την δουλειά της επεξεργασίας πολύ ευκολότερη. Το σημαντικό όμως είναι ότι θα εξαφανιζόταν ο θόρυβος που υπάρχει κατά την λήψη ροής βίντεο, ο οποίος προσθέτει τυχαία χρώματα σε διάφορες περιοχές του κάθε καρτέ, τα οποία δεν υπάρχουν στο πραγματικό αντικείμενο.

Το τελευταίο όπως είδαμε δημιούργησε αποκρίσεις σε τεχνητά δίπολα κύτταρα διαφορετικών χρωμάτων στον μαύρο φόντο της φωτογραφίας, εκεί που κανονικά θα έπρεπε να υπήρχε ο ελάχιστος κορεσμός (saturation), το S στο σύστημα HSL. Δηλαδή κανονικά αν δεν υπήρχε ο θόρυβος, στον μαύρο φόντο της εικόνας τα χρώματα έπρεπε να είναι πιο κοντά στο γκρι. Αντ' αυτού ξεπερνούσαν ένα κατώφλι παράγοντας χρωματική απόκριση. Εδώ βοήθησε το φίλτρο των τεχνητών βραχύνων κυττάρων που σβήνει τα χρώματα που υπήρχαν ως θόρυβος σε πολύ χαμηλές φωτεινότητες.

Το όφελος που προκύπτει από την πειραματική εφαρμογή του Horasis σε καρτέ ροής βίντεο με θόρυβο, είναι ότι μπορούν να ληφθούν φωτογραφίες σε πραγματικές συνθήκες. Η λειτουργία του ΤΝΔ μπορεί να γίνει υπό κίνηση ή ακόμα χειρότερα με αντανάκλασεις και αλλοιώσεις που προκύπτουν από τον φωτισμό. Αν και αυτές τις δύσκολες συνθήκες δεν μπορεί να τις καλύψει μια πτυχιακή εργασία, το Horasis προσπαθεί να επιδείξει ότι πρέπει να επιδιωχθεί η χρήση ατελών οπτικών προτύπων από την αρχή σε κάθε ερευνητική προσπάθεια τεχνητής όρασης.

Συμπύεση οπτικής πληροφορίας χωρίς να χάνει τα βασικά χαρακτηριστικά της.

Είδαμε ότι από μια αρχική οπτική πληροφορία ενός αντικειμένου που διατηρεί για κάθε μια από τις 500x500 τετραγωνικές ψηφίδες 3 χρώματα, φτάνουμε σε μια εσωτερική αναπαράσταση του αντικειμένου. Σε αυτήν διατηρούνται για κάθε ένα από τα 155 κεντρικά εξάγωνα και για κάθε έναν από τους 3 χρωματικούς ανταγωνισμούς, 6 κατευθύνσεις γραμμών. Υπολογίζοντας:

- $500 * 500 * 3 = 750.000$ αρχικές τιμές
- $155 * 3 * 6 = 2.790$ τελικές τιμές

Η νέα πληροφορία είναι μόνο το 0.37% της αρχικής, διατηρώντας ένα απαραίτητο τμήμα της αρχικής, με το οποίο είναι ικανό να αναγνωριστεί το αντικείμενο. Η μείωση αυτή από μόνη της συνιστά συμπύεση της εικόνας με απώλεια πληροφορίας. Αν και δεν μπορεί να αναπαραχθεί από την συμπιεσμένη μορφή η αρχική εικόνα, οι πληροφορίες που περιέχονται μπορούν να επεξεργαστούν με πολύ μεγάλη ταχύτητα λόγω του μικρού όγκου τους.

Με αυτόν τον τρόπο επιτυγχάνεται μια πολύ μικρή απαίτηση σε χώρο για αποθήκευση του κάθε προτύπου, ενώ παράλληλα έχοντας αυξημένη ταχύτητα, ανοίγει ο δρόμος για επεξεργασία μιας ακολουθίας από καρτέ εικόνες στην θέση της μιας στατικής φωτογραφίας. Η νέα διαμόρφωση RNF μπορεί να χρησιμοποιηθεί αντί του BMP σε μελλοντικές υλοποιήσεις ΤΝΔ τεχνητής όρασης.

Χρήση παράλληλης επεξεργασίας

Η κατασκευή του Hogasis με αντικειμενοστραφή τεχνολογία, άνοιξε τον δρόμο για την χρήση παράλληλης επεξεργασίας. Λόγω της σωστής οργάνωσης των αντικειμένων ένα ολόκληρο ΤΝΔ μπορεί να λειτουργεί σε ένα ξεχωριστό νήμα. Αυτό οφείλεται σε δύο λόγους, πρώτον ότι η εκπαίδευση είναι αποπολυπλεγμένη (decoupled) από το UI, το οποίο τρέχει πάντα στο κεντρικό νήμα (main thread) της εφαρμογής και κατά δεύτερον λόγο ότι κάθε αντικειμενοστραφές ΤΝΔ περιλαμβάνει τον κώδικα λειτουργίας μαζί με τα δικά του δεδομένα. Ένα άλλο στιγμιότυπο (instance) ενός αντικειμενοστραφούς ΤΝΔ έχει διαφορετική ιδιωτική περιοχή μνήμης, στην οποία κανένα άλλο νήμα δεν χρειάζεται να αποκτήσει πρόσβαση.

Η χρήση πολλών παράλληλων εκπαιδεύσεων μείωσε σημαντικά τον χρόνο που χρειάστηκε για να εκπαιδευτούν όλοι οι διαφορετικοί τύποι ΤΝΔ στο Hogasis. Γι αυτό θα λέγαμε ότι μείωσε και ως ένα βαθμό την πολυπλοκότητα της εύρεσης της τελικής συνταγής, για την λύση του προβλήματος. Αν συνυπολογιστεί ότι με αυτήν την τεχνολογική προσέγγιση, αξιοποιήθηκαν νέες δυνατότητες του υλικού προσωπικών υπολογιστών που παρουσιάστηκαν τα τελευταία χρόνια, θα λέγαμε ότι φτάνει και σε ακόμα έναν από τους αρχικούς στόχους της εργασίας.

Ο στόχος που επετεύχθη ήταν να εκμεταλλευθούμε καλλίτερα τις νέες δυνατότητες του υλικού, προς όφελος των ΤΝΔ. Αυτό εκτός από την υλοποίηση του αποτέλεσε σημαντικό παράγοντα επιτυχίας, βοηθώντας την εργασία να ολοκληρωθεί σε λογικό χρόνο.

Εισαγωγή νέων προτάσεων και παραγωγή νέων όρων στην τεχνητή όραση

Παρότι το σύστημα Hogasis είναι απλώς μια πτυχιακή εργασία, που ο αυτοσκοπός της είναι η απόκτηση ενός πτυχίου στην επιστήμη των υπολογιστών, έγινε προσπάθεια για την εισαγωγή νέων ιδεών ως προτάσεις για την δημιουργία τεχνητής όρασης. Οι προτάσεις αυτές παρατίθενται για να αξιολογηθούν από οποιονδήποτε αναγνώστη, και να χαρακτηριστούν αν είναι η όχι καινοτόμες.

Με ίσως λίγο υπερβατικό τρόπο έγινε μια προσπάθεια για την παραγωγή νέων όρων που συνοδεύουν τις παραπάνω προτάσεις. Αυτό χρησιμεύει κυρίως για τον ορισμό ενός λεξιλογίου με το οποίο κάποιος μπορεί να κατανοήσει σημεία της εργασίας. Το αν οι όροι αυτοί φανούν χρήσιμοι γενικότερα στον τομέα της τεχνητής όρασης, απομένει να αποδειχθεί. Επιπρόσθετα τοποθετήθηκαν Ελληνικές μεταφράσεις όρων που εμφανίζονται στην ξένη βιβλιογραφία και το διαδίκτυο, πάλι με το σκεπτικό ότι κατόπιν αξιολόγησης ίσως φανούν δόκιμοι και χρήσιμοι στους κύκλους της Ελληνικής Πληροφορικής.

Στην επόμενη ενότητα θα γίνει μια σύνοψη όλων των παραπάνω, συγκεντρώνοντας προτάσεις, όρους, μεταφράσεις όρων αλλά και συμβάσεις που έγιναν κατά μήκος όλου του συγγράμματος.

Προτάσεις, Ονοματολογικές Συμβάσεις, Νέες Έννοιες,

Προτάσεις και καινοτομίες πάνω στα τεχνητά νευρωνικά δίκτυα

Επεξεργαστές Νευρωνικών Δικτύων (Κεφ. 1)

Θα μπορούσαν να φτιαχτούν ψηφιακά ολοκληρωμένα κυκλώματα που υλοποιούν ή επιταχύνουν υπολογισμούς, μεταφορά πληροφοριών και διαδικασίες εκπαίδευσης τεχνητών νευρωνικών δικτύων. Είναι δηλαδή συμβατικό hardware που υποβοηθάει τα ΤΝΔ, το οποίο μπορεί να είναι ενσωματωμένο σε μια CPU, κατ' αντιστοιχία με τα PPU που αποτελούν πλέον μέρος των εξελιγμένων GPU.

Αντικείμενο Νευρωνικού Μοτίβου = Neural Pattern Object = NPO (Κεφ. 5)

Ορίζεται ως ένα αντικείμενο λογισμικού που περιέχει τουλάχιστον έναν πίνακα με αριθμητικές τιμές που αντιπροσωπεύουν τα εισερχόμενα σήματα εισόδου σε ένα τεχνητό νευρωνικό δίκτυο. Επιπρόσθετα μπορεί στο ίδιο αντικείμενο να τοποθετηθεί και έναν πίνακα με αριθμητικές τιμές που αντιπροσωπεύουν τους στόχους στις εξόδους του δικτύου για τις συγκεκριμένες τιμές εισόδου. Είναι το ερέθισμα στην νευρωνική επεξεργασία και το πρότυπο για την εκπαίδευση του ΤΝΔ.

Νέο κριτήριο τερματισμού του back propagation (Κεφ. 5)

$$\left(\frac{1}{n}\right) \sum_{i=1}^n \left(\left(\frac{1}{c}\right) \sum_{p=1}^c \delta[p]_i \right) < \varepsilon$$

όπου:

- ε ένας μικρός δεκαδικός αριθμός,
- n το πλήθος των νευρώνων του στρώματος εξόδου,
- c το πλήθος των προτύπων στο σύνολο των προτύπων εκπαίδευση,
- $\delta[p]_i$ ο παράγοντας σφάλματος του νευρώνα, από την σχέση 3.22

$$\Delta w_{ij}(k) = \beta \delta_i x_j - \mu \Delta w_{ij}(k-1)$$

ο οποίος υπολογίζεται κατά την ανάκληση ενός προτύπου p για τον νευρώνα i του στρώματος εξόδου.

Μεταφράσεις ξένων όρων



Συνδεδασισμός (Κεφ.1)

Χρήση μετάφρασης του όρου connectionism. Χρησιμοποιήθηκε η λέξη «συνδεδασισμός» καθώς εμφανιζόταν 41 αποτελέσματα στην αναζήτηση του όρου μέσω Google σε σχέση με 3 αποτελέσματα του «συνδεδασμός».



Ανθρώπινο Συνδεσίωμα (Κεφ. 2)

Δημιουργία νέου όρου για το human connectome. Από την ρίζα της λέξης σύνδεσι-ς και την ίδια κατάληξη με την λέξη γονιδ-ίωμα που έχει επικρατήσει ως μετάφραση του όρου genome. Η νέα μετάφραση του όρου παρατίθεται προς έλεγχο, ώστε να κριθεί αν είναι δόκιμη.



Σύζευξη και Συνοχή (Κεφ.2)

Χρήση μετάφρασης των όρων coupling και cohesion αντίστοιχα. Και οι δύο όροι είναι μεταφράσεις από τα Αγγλικά που εμφανίζονται ως πρώτες επιλογές στο translate.google.com. Και οι δύο όροι στα Ελληνικά περιγράφουν ακριβώς το θέμα που αναφέρει η Μηχανική Λογισμικού.



Σύσταση και Συναθροίση (Κεφ.2)

Χρήση μετάφρασης «σύσταση» για τον όρο composition, καθώς σε ένα αντικείμενο τα αντικείμενα που διατηρούν την σχέση composition με αυτό, είναι στην ουσία τα συστατικά από τα οποία αποτελείται. Για την σχέση aggregation μεταξύ αντικειμένων χρησιμοποιήθηκε ο όρος «συναθροίση» καθώς τα αντικείμενα δεν συνδέονται μεταξύ τους με μόνιμη σχέση. Απλώς βρίσκονται συναθροισμένα σε μια χρονική στιγμή της εκτέλεσης του προγράμματος, όπως οι άνθρωποι συναθροίζονται προσωρινά χωρίς να έχουν κάποια μόνιμη εξάρτηση μεταξύ τους.



Σκληροκώδικας (Κεφ. 4)

Δημιουργία νέου όρου για τον προγραμματιστικό όρο hardcoded. Οι τιμές ιδιοτήτων αντικειμένων που εγγράφονται στο κώδικα, προκαθορίζοντας για παράδειγμα μια ρύθμιση λέγονται hardcoded. Αυτές οι τιμές θα λέγαμε ότι είναι «χαράγματα με σκληρό αντικείμενο» πάνω στην «επιφάνεια» του προγράμματος και όχι «γραμμένες με μολύβι» ώστε να μπορούν αν σβηστούν και να αλλαχθούν σε κάποια στιγμή της εκτέλεσης του. Εξ' ου και ο όρος σκληροκώδικας που εισάγεται γι αυτές.

Ονομασίες κατά σύμβαση που αφορούν τεχνητά νευρωνικά δίκτυα

Συναπτικό Άθροισμα = Synaptic Summation (Κεφ. 3)

Κατά σύμβαση περιγράφουμε το άθροισμα u των γινομένων συναπτικών βαρών w_i επί των μεταβλητών εισόδου x_i σε έναν τεχνητό νευρώνα McCulloch – Pits.

Ουδέτερη Σύναψη = Null Synapse (Κεφ. 4)

Κατά σύμβαση ονομάζεται ένα αντικείμενο σύναψης που ανήκει σε αντικειμενοστραφή τεχνητό νευρώνα, το οποίο αν και έχει δημιουργηθεί δεν αναφέρεται σε κάποιον άξονα από τον οποίο προέρχεται η εισερχόμενη τιμή της σύναψης.

Γραμμή Πυθμένα Σφάλματος (Κεφ.5)

Κατά σύμβαση ονομάζεται μια οριζόντια παράλληλη γραμμή στο γράφημα του μέσου τετραγωνικού σφάλματος που αντιπροσωπεύει το τοπικό ελάχιστο στο οποίο έχει φτάσει μια εκπαίδευση, σε οποιαδήποτε περίπτωση που αυτή διακόπτεται.

Νέες έννοιες και καινοτομίες πάνω στην τεχνητή όραση

Αμιγή Χρώματα = Pure Colors (Κεφ. 6)

Ονομάζεται ένα χρώμα που έχει μια η δύο από τις χρωματικές συνιστώσες R,G,B στην μέγιστη τιμή της και όλες τις υπόλοιπες στην ελάχιστη, και θεωρείται ως το «τέλειο» πρότυπο για την έννοια χρώματος που αντιπροσωπεύει.

Αντικειμενοστραφής Τεχνητός Αμφιβληστροειδής – Object Oriented Artificial Retina = OOAR (Κεφ. 7, 8)

Ορίζεται ως ένας σχεδιασμός αντικειμένων λογισμικού, που μπορεί να περιέχει τεχνητά νευρωνικά δίκτυα, κατά τον οποίο τα αντικείμενα οργανώνονται και λειτουργούν αντίστοιχα με τα κύτταρα που υπάρχουν σε έναν βιολογικό αμφιβληστροειδή χιτώνα. Μέσα στα πλαίσια του Horasis υλοποιείται έναν πρώτος αντικειμενοστραφής τεχνητός αμφιβληστροειδής.


Κωδικοποίηση Γαγγλιακού Κυττάρου = Ganglion Cell Encoding = GCE (Κεφ. 7)

Ορίζεται ως μια μαθηματική αναπαράσταση της μορφής του ηλεκτρικού παλμού που παράγει στην έξοδο του ένα βιολογικό γαγγλιακό κύτταρο. Η κωδικοποίηση αυτή χρησιμοποιείται σε ένα

τεχνητό γαγγλιακό κύτταρο και μπορεί να είναι ένα διάλυμα πολλών διαστάσεων, που εξαρτάται από το τεχνητό μοντέλο.

 **Νευρωνική Διαμόρφωση Αμφιβληστροειδούς = Retina Neural Format = RNF (Κεφ.7)**

Είναι ένα σύνολο από επιμέρους σήματα GCE που αποτελούν την αναπαράσταση μια εικόνας αφού έχει επεξεργαστεί από έναν βιολογικό αμφιβληστροειδή και η οποία πληροφορία μεταφέρεται μέσω του οπτικού νεύρου. Σε έναν τεχνητό αμφιβληστροειδή αποτελεί την διαμόρφωση της πληροφορίας στην έξοδο του και παράλληλα την πληροφορία που εισάγεται σε έναν τεχνητό οπτικό φλοιό.

 **Τεχνητός Βολβός Ματιού = Artificial EyeBall = AEB (Κεφ. 8)**


Ορίζεται ως ένα πληροφοριακό σύστημα που περιλαμβάνει υλικό και λογισμικό που μιμούνται την λειτουργία του βιολογικού βολβού του ματιού.

 **Τεχνητός Ρετινοτοπικός Χάρτης = Artificial Retinotopic Map = ARTM (Κεφ. 8)**

Είναι ένα υποσύστημα λογισμικού με το οποίο υλοποιείται μια χωρική αντιστοίχιση των εικονοστοιχείων της εικόνας, με τα υποδεκτικά πεδία των τεχνητών νευρώνων στα επόμενα υποσυστήματα.

 **Τεχνητός Οπτικός Φλοιός = Artificial Visual Cortex = AVC (Κεφ. 8)**

Ορίζεται ως ένα λογισμικό που θα παίρνει είσοδο από έναν τεχνητό αμφιβληστροειδή και στην έξοδο του θα παράγει μια νέα πληροφορία στην οποία προσαρτώνται νέα χαρακτηριστικά σχήματος και μορφής στην οπτική πληροφορία. Η έξοδος αυτή θα αποτελεί την εσωτερική αναπαράσταση της εικόνας, για το σύστημα τεχνητής νοημοσύνης που το περιέχει.

 **Χάρτης Υποδεκτικού Πεδίου = Receptive Field Map = RFM (Κεφ.9)**

Είναι μια χωρική αντιστοίχιση που υλοποιείται ως μια ψηφιδωτή εικόνα στην οποία χρωματίζονται με συγκεκριμένο χρώμα όσα εικονοστοιχεία θα χρησιμοποιηθούν στο υποδεκτικό πεδίο ενός νευρώνα και με μαύρο όσα θα αγνοηθούν. Σε έναν τέτοιο χάρτη μπορούν να υπάρχουν πολλά χρώματα, που αντιστοιχούν σε υποδεκτικά πεδία διαφορετικών τεχνητών νευρώνων.

 **Ρετινοστοιχείο = Retinal Element = Rexel (Κεφ. 9)**

Ορίζεται μέσα σε ένα σύστημα τεχνητής νοημοσύνης που υλοποιεί τεχνητό αμφιβληστροειδή, ως μια κυψέλη οποιασδήποτε μορφής που αντιστοιχεί σε συγκεκριμένο χώρο της εικόνας, αρκεί η ψηφίδωση αυτής της μορφής να μπορεί να καλύψει ομοιόμορφα όλη την εικόνα.

Βιβλιογραφία και Παραπομπές URL

1. **Lindholm, Tim and Yellin, Frank.** The Java Virtual Machine Specification. *Oracle (Sun Microsystems)*. [Online] 199. [Cited: 10 19, 2011.]
http://java.sun.com/docs/books/jvms/second_edition/html/VMSpecTOC.doc.html.
2. **ECMA International.** *Standard ECMA-335 5th Edition: Common Language Infrastructure (CLI)*. s.l. : ECMA International, 2010.
3. **Teixeira, Steve and Pacheco, Xavier.** *Borland Delphi 6 Developers Guide*. Indianapolis, USA : Sams Publishing, 2002.
4. **Borland Software Corporation.** *Delphi 7 Developer's Guide*. s.l. : Borland Software Corporation, 2002. ch.13 / pp. 13.1-13.16.
5. **Wikipedia.** Anders Hejlsberg. *Wikipedia, The Free Encyclopedia*. [Online] 10 18, 2011. [Cited: 10 19, 2011.] http://en.wikipedia.org/wiki/Anders_Hejlsberg.
6. **Embarcadero Technologies.** What's new in Delphi XE2. *Embarcadero Technologies*. [Online] 2011. [Cited: 10 19, 2011.] <http://edn.embarcadero.com/article/41599>.
7. **Δέρβος, Δημήτρης.** *Μαθήματα βάσεων δεδομένων, 1η έκδοση*. Θεσσαλονίκη : Τζιόλα, 1995.
8. **Oppel, Andy.** *Databases Demystified - A Self-Teaching Guide*. s.l. : McGraw-Hill/Osborne, 2004.
9. **Dewson, Robert.** *Beginning SQL Server 2005 for Developers*. s.l. : APres, 2006.
10. **Roff, Jason T.** *ADO: ActiveX Data Objects*. s.l. : O'Reilly, 2001. pp. 12-55.
11. **Angelini, Chris.** AMD Phenom II X4 965 Black Edition Review. *Tom's Hardware*. [Online] 08 13, 2009. [Cited: 10 20, 2011.] <http://www.tomshardware.com/reviews/phenom-x4-965,2389.html>.
12. **Thagard, Paul.** Stanford Encyclopedia of Philosophy - Cognitive Science. *Stanford Encyclopedia of Philosophy*. [Online] 09 23, 1996. [Cited: 09 17, 2011.]
<http://plato.stanford.edu/entries/cognitive-science/>.
13. **MuCulloch, W. and Pitts, W.** A logical calculus and the ideas immanent in the nervous activity. *Bulletin of Mathematical Biophysics*. 1943, Vol. 5, pp. 115-113.
14. **I.R.I.S.** *Readiris Pro 9 User's Guide*. s.l. : I.R.I.S, 2003. pp.3-4.
15. **Singh, Raghuraj, et al.** Optical Character Recognition (OCR) for Printed Devnagari Script Using Artificial Neural Network . *International Journal of Computer Science & Communication*. 2010, Vols. 1, No.1, pp. 91-95.
16. **Durah, Kheder, et al.** Future Biometric Passports and Neural Networks. *Global Journal of Computer Science and Technology*. 2011 , Vols. 11, No.7, pp. 20-26.

17. **Ελληνική Αστυνομία - Διεύθυνση Διαβατηρίων.** Βιομετρικά Χαρακτηριστικά. *Διεύθυνση Διαβατηρίων*. [Online] [Cited: 10 11, 2011.] <http://www.passport.gov.gr/npc-perioxomeno/npc-perioxomeno/viometrika-xarakteristika-diavatiriou.html>.
18. **Garson, James.** Standform Encyclopedia of Philosophy - Connectionism. *Standform Encyclopedia of Philosophy*. [Online] 05 18, 1997. [Cited: 10 11, 2011.] <http://plato.stanford.edu/entries/connectionism/>.
19. **Benyus, Janine.** *Biomimicry - Innovation Inspired by Nature*. New York : Perennial, 2002.
20. **Pedretti, Carlo.** *A Chronology of Leonardo Da Vinci's Architectural Studies after 1500*. Geneva : E. Droz, 1962.
21. **Hamilton, Kim and Miles, Russell.** *Learning UML 2.0*. s.l. : O' Reilly, 2006.
22. **Rorres, Chris.** The Turn Of The Screw - Optimal Design Of An Archimedes Screw. *Journal Of Hydraulic Engineering*. 2000, Vol. 126. No 1, pp. 72-80.
23. **brainz.org.** The 15 Coolest Cases of Biomimicry. [Online] [Cited: 10 11, 2011.] <http://brainz.org/15-coolest-cases-biomimicry/>.
24. **Popular Science.** Best Of What's New 2008 - Speedo LZR. [Online] [Cited: 10 11, 2011.] <http://www.popsci.com/bown/2008/product/speedo-lzr>.
25. *Speedo, Like Michael Phelps, Goes For World Domination in an LZR Suit.* **Thurow, Roger and Rhoads, Christopher.** s.l. : Wall Street Journal, 14/09/2008. pp. A14.
26. **Windhorst, Uwe and Johansson, Hakan.** *Modern Techniques in Neuroscience Research*. Berlin : Springer-Verlag, 1999.
27. **Βρεττός, Γιάννης.** *Μελέτη Οφθαλμικών Κινήσεων Προσήλωσης*. Ηράκλειο : Πανεπιστήμιο Κρήτης - Διατμηματικό Μεταπτυχιακό Πρόγραμμα Σπουδών: Οπτική και Όραση, 2006. pp. 20-24.
28. **Simon, Brian and Simon, Joan.** *Educational Psychology in the U.S.S.R.* s.l. : Stanford University Press, 1963. PP. 192-198.
29. **Stroustrup, Bjarne.** *What is "Object-Oriented Programming"? (1991 revised version)*. New Jersey : AT&T Bell Laboratories, 1991.
30. **O' Docherty, Mike.** *Object-Oriented Analysis & Design - Understanding System Development with UML 2.0*. s.l. : John Wiley & Sons, 2005. pp. 20-22.
31. **Troelsen, Andrew.** *Pro C3 2005 and the .NET 2.0 Platform, 3rd Edition*. s.l. : apress, 2005. pp. 347-349.
32. **Microsoft Corporation.** Singularity - Microsoft Research. [Online] [Cited: 10 11, 2011.] <http://research.microsoft.com/en-us/projects/singularity/>.
33. **Hunt, Galen C and Larus, James R.** Singularity: Rethinking the Software Stack. *ACM SIGOPS Operating Systems Review*. Association for Computing Machinery, Inc, 2007, Vols. 41, No. 2, pp. 37-49.

34. **Hunt, Galen, et al.** *Sealing OS Processes to Improve Dependability and Safety*. Association for Computing Machinery, Inc. Lisbon, Portugal : Association for Computing Machinery, Inc., 2007.
35. **James, Martin.** *RAD, Rapid Application Development*. New York : MacMillan Publishing Co., 1990.
36. **Petzold, Charles.** *Programming Windows, 5th edition*. s.l. : Microsoft Press, 1998. ch. 3, ch. 20.
37. **Halabi, Mohamad.** Introducing .NET 4.0 Parallel Programming. [Online] 10 07, 2010. [Cited: 10 11, 2011.] http://www.codeproject.com/KB/Parallel_Programming/NET4ParallelIntro.aspx.
38. **Sanoulla, Mohamed.** Cohesion and Coupling: Two OO Design Rules. [Online] 06 26, 2008. [Cited: 10 11, 2011.] <http://blog.sanoulla.info/2008/06/26/cohesion-and-coupling-two-oo-design-principles/>.
39. **Hitz, Martin and Montazeri, Behzad.** *Measuring Coupling and Cohesion in Object Oriented Systems*. Vienna : Institut für Angewandte Informatik und Systemanalyse, University of Vienna, 1995.
40. **Lanza, Michele, Marinescu, Radu and Ducasse, Stephane.** *Object-Oriented Metrics in Practice*. Berlin : Springer, 2006. pp. 73-137.
41. **Ambler, Scott W.** *The Elements of UML 2.0 Style*. s.l. : Cambridge University Press, 2005. pp. 70-72.
42. **Passmark Software.** CPU Benchmarks - High End CPUs. [Online] Passmark Software, 10 15, 2011. [Cited: 10 15, 2011.] http://www.cpubenchmark.net/high_end_cpus.html.
43. **Faiedh, H., Gafsi, Z., Besbes, K.** *Digital hardware implementation of sigmoid function and its derivative for artificial neural networks*. Rabat, Morocco : Proc. 13th Int. Conf. on Microelectronics, Rabat, 2001. pp. 189-192.
44. **Tommiska, M. T.** Efficient digital implementation of the sigmoid. *Reprinted with permission from IEE Proceedings – Computers and Digital Techniques*. 2003, Vols. 150, No 6, pp. 403-411.
45. **NVidia Corporation.** Physx FAQ. [Online] [Cited: 10 12, 2011.] http://www.nvidia.com/object/physx_faq.html.
46. **Time-Life.** *Understanding Computers (Ελληνική έκδοση)*. Αθήνα : Αλκυών, 1990. pp. 45-67.
47. **Wozniak, Steve.** *iWoz: From Computer Geek to Cult Icon*. s.l. : W. W. Norton & Company, 2006.
48. **Internet World Stats.** World Internet Users and Population Stats. [Online] Miniwatts Marketing Group, 10 06, 2011. [Cited: 10 15, 2011.] <http://www.internetworldstats.com/stats.htm>.
49. **McElheny, Victor K.** *Drawing the map of life: Inside the Human Genome Project*. New York : Basic Books, 2010.

50. **US Dept. of Health and Human Services - National Institute of Health (NIH).** NIH Launches the Human Connectome Project to Unravel the Brain's Connections. *News Release*. [Online] 07 15, 2009. [Cited: 10 11, 2011.] <http://www.nih.gov/news/health/jul2009/ninds-15.htm>.
51. **Raj, A, Hess, C and Mukherjee, P.** Spatial HARDI: improved visualization of complex white matter architecture with Bayesian spatial regularization. *Neuroimage*. 2011, Vols. 54, No 1 Epub, pp. 396-409.
52. **Kandel, Eric R., Schwarz, James H and Jessel, Thomas M.** *Νευροεπιστήμη και Συμπεριφορά (Essentials of Neural Science and Behaviour)*. Ηράκλειο : Πανεπιστημιακές Εκδόσεις Κρήτης, 2009. pp.24-45 , pp. 231-240, pp. 429-470, pp. 487-493.
53. **Kuffler, Stephen W.** Discharge Patterns And Functional Organization Of Mammalian Retina. *The Journal of Neurophysiology*. 1953, Vols. 16, Vol. 1, pp. 37-68.
54. **Διαμαντάρας, Κωνσταντίνος.** *Τεχνητά Νευρωνικά Δίκτυα*. s.l. : Κλειδάριθμος, 2007. pp. 16-19, 44-69.
55. **Poirazi, Panayiota, Brannon, Terrence and Mel, Bartlett W.** Arithmetic of Subthreshold Synaptic Summation in a Model CA1 Pyramidal Cell. *Neuron*, 2003, Vol. 37, pp. 977-978.
56. **Anthony, Martin.** *Discrete Mathematics Of Neural Networks Selected Topics*. London, UK. : Society for Industrial and Applied Mathematics, 2001. pp. 21.
57. *Artificial neural networks - ICANN 2002.* **Chen, Hsin and Murray, Alan.** Madrid, Spain : Springer, 2002. pp. 359.
58. **Βλαχάβας, Ι., et al.** *Τεχνητή Νοημοσύνη*. s.l. : Γαρταγάνης, 2002. pp. 172.
59. **Rojas, Raul.** *Neural Networks - A Systematic Introduction*. s.l. : Springer-Verlag, 1996. pp. 151-173.
60. **Valle, Maurizio.** *Smart adaptive systems on silicon*. Dordrecht : Kluwer Academic Publishers, 2004. pp. 177-177.
61. **Bryson, Arthur Earl and Ho, Yu-Chi.** *Applied optimal control: optimization, estimation, and control*. s.l. : Blaisdell Publishing Company or Xerox College Publishing, 1969. pp. 481.
62. **Russel, Stuart and Norvig, Peter.** *Artificial Intelligence: A Modern Approach*. s.l. : Prentice Hall, 1995.
63. **Werbos, Paul J.** *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. PhD thesis*. s.l. : Harvard University. 1974.
64. **Rumelhard, David E., Hinton, Geoffrey E. and Williams, Ronald J.** Learning representations by back-propagating errors. *Nature* 323. 1986, pp. 533-536.
65. **Amini, J.** Optimum Learning Rate in Back-Propagation Neural Network for Classification of Satellite Images (IRS-1D). *Scientia Iranica*. 2008, Vols. 15, No.6, pp. 559-559.

66. **Magoulas, G. D., Vrahatis, M. N. and Androulakis, G. S.** Improving the Convergence of the Backpropagation Algorithm Using Learning Rate Adaptation Methods. *Neural Computation*. 1999, Vol. 11. No. 7, pp. 1769-1796.
67. **Tarhini, Ali.** Concepts of Three-Tier Architecture. [Online] [Cited: 11 13, 2011.] <http://alitarhini.wordpress.com/2011/01/22/concepts-of-three-tier-architecture/>.
68. **Microsoft Corporation.** *Microsoft Application Architecture Guide, 2nd Edition - Patterns & Practices*. s.l. : Microsoft Corporation, 2009. pp. 55-60.
69. **Eckerson, Wayne W.** Three Tier Client/Server Architecture: Achieving Scalability, Performance, and Efficiency in Client Server Applications. *Open Information Systems*. 1995, Vols. 10, No. 1, 3(20).
70. **Evjenet, Bill et al.** *Professional XML*. s.l. : Wrox Press, 2007. Ch. 1.
71. **Wood, Keith.** *Delphi Developer's Guide to XML*. s.l. : Wordware Publishing, Inc, 2001. pp. 108-156.
72. **Gajic, Zarko and Claessens, Dirk.** Streams, streams, ... TStream. [Online] [Cited: 10 13, 2011.] <http://delphi.about.com/od/vclusing/l/aa110803a.htm>.
73. **Pierce, Benjamin A.** *Genetics a Conceptual Approach*. s.l. : W. H. Freeman, 2007. pp. 16-27.
74. **Καπλάνογλου, Νικόλαος Ι. and Γεροντίδης, Κωνσταντίνος.** *Ξενοφάνης: Αναγνώριση Δυναμικών Συστημάτων Με Χρήση Αντικειμενοστραφούς Γενετικού Αλγορίθμου*. Θεσσαλονίκη : Α.Τ.Ε.Ι. Θεσσαλονίκης, 2007. pp. 41-41.
75. **Wikipedia, the free encyclopedia.** Design pattern (computer science). [Online] 10 12, 2011. [Cited: 10 13, 2011.] [http://en.wikipedia.org/wiki/Design_pattern_\(computer_science\)](http://en.wikipedia.org/wiki/Design_pattern_(computer_science)).
76. **Alexandrescu, Andrei.** *Modern C++ Design: Generic Programming and Design Patterns Applied*. s.l. : Addison Wesley, 2001. pp. 174-180.
77. **Nice, Karim, Wilson, Tracy V and Gurevich, Gerald.** HowStuffWorks: "How Digital Cameras Work". [Online] HowStuffWorks. [Cited: 10 14, 2011.] <http://electronics.howstuffworks.com/cameras-photography/digital/digital-camera2.htm>.
78. **Felber, Philip.** *Charge-Coupled Devices - A literature study as a project for ECE 575*. s.l. : Illinois Institute of Technology, 2002.
79. **Taylor, Stuart A.** *CCD and CMOS Imaging Array Technologies - Technology Review*. Cambridge : Xerox Research Centre Europe, 1998.
80. **Li, Xin, Gunturk, Bahadir and Zhang, Lei.** Image demosaicing: a systematic survey. *Visual Communications and Image Processing*. Proceedings of the SPIE, 2008, Vol. 6822, pp. 68221J-68221J-15 .
81. **Bergman, David.** How I Made a 1,474-Megapixel Photo During President Obama's Inaugural Address. [Online] David Bergman. [Cited: 10 14, 2011.] <http://www.davidbergman.net/blog/how-i-made-a-1474-megapixel-photo-during-president-obamas-inaugural-address/>.

82. **Silberscharz, Abraham and Galvin, Peter.** *Operating Systems Concepts 8th edition.* s.l. : John Wiley & Sons, 2008.
83. **Silberscharz, Abraham and Galvin, Pete.** *Student's Manual to Accompany: Operating System Concepts 7th Edition.* s.l. : John Willey & Sons, 2007. pp.1-1.
84. **Pattison, Ted.** *Programming Distributed Application with COM and Microsoft Visual Basic 6.0.* Redmond : Microsoft Press, 1998. pp. 1-18.
85. **MSDN .** About DirectShow. [Online] Microsoft Corporation. [Cited: 10 14, 2011.] [http://msdn.microsoft.com/en-us/library/windows/desktop/dd373389\(v=VS.85\).aspx](http://msdn.microsoft.com/en-us/library/windows/desktop/dd373389(v=VS.85).aspx).
86. **MSDN.** DirectShow. [Online] Microsoft Corporation. [Cited: 10 14, 2011.] <http://msdn.microsoft.com/en-us/library/ms783323.aspx>.
87. **Στυλιάδης, Αθανάσιος Δ.** *Γραφικά με ηλεκτρονικό υπολογιστή.* s.l. : Εκδόσεις Ζήτη, 1999. pp. 68-68.
88. **Microsoft Corporation.** DIBs and Their Uses. [Online] Microsoft Corporation, 02 11, 2005. [Cited: 10 14, 2011.] <http://support.microsoft.com/kb/q81498/>.
89. **Ford, Janet Lynn.** Color Theory. [Online] Worqx.com. [Cited: 10 14, 2011.] <http://www.worqx.com/color/>.
90. **Hunt, R.W.G.** *The Reproduction of Colour.* s.l. : John Wiley & Sons, 2004. pp. 9-16.
91. **Wikipedia.** HSL and HSV. *Wikipedia, the free encyclopedia.* [Online] 10 09, 2011. [Cited: 10 14, 2011.] http://en.wikipedia.org/wiki/HSL_and_HSV.
92. **Lauren, Marie.** The Color Wheel and Color Theory. [Online] 05 16, 2008. [Cited: 10 14, 2011.] <http://creativecurio.com/2008/05/the-color-wheel-and-color-theory/>.
93. **Foley, James D. and Van Dam, Andries.** *Fundamentals of Interactive Computer Graphics.* s.l. : Addison-Wesley, 1982. ch. 17.
94. **Monger, Patricia.** RGB - HSL , HSL - RGB. *McMaster University.* [Online] [Cited: 07 15, 2011.] <http://130.113.54.154/~monger/hsl-rgb.html>.
95. **Munroe, Randall.** Color Survey Results. [Online] 05 03, 2010. [Cited: 10 15, 2011.] <http://blog.xkcd.com/2010/05/03/color-survey-results/>.
96. **Pratt, William K.** *Digital Image Processing, 3rd Edition.* s.l. : John Willey & Sons, 2001. pp.71-81.
97. **Wise, Edwin.** *Robotics Demystified - A Self Teaching Guide.* s.l. : McGraw Hill, 2005. pp.124-154.
98. **Miguel, Agullo.** The Shape-Shifting Camera Tank. [book auth.] Agullo et al. Miguel. *LEGO Mindstorms Masterpieces.* s.l. : Syngress Publishing, 2003.
99. **Wikipedia.** Stereoscopy. *Wikipedia, the free encyclopedia.* [Online] 10 15, 2010. [Cited: 10 15, 2010.] <http://en.wikipedia.org/wiki/Stereoscopy>.

100. **Borland Software Corporation.** *Delphi 7 Developer's Guide*. s.l. : Borland Software Corporation, 2002. ch. 27/pp. 27.1 - 27.20.
101. **Sheldon, Robert.** *SQL A Begginers Guide, 2nd Edition*. s.l. : McGraw-Hill/Osborne, 2003.
102. **Kanter, Joel P.** *Understanding Thin Client/Server Computing*. s.l. : Microsoft Press, 1998. pp. 3-7.
103. **Science Clarified.** Physiology. [Online] Advameg, Inc., 2011. [Cited: 10 15, 2011.] <http://www.scienceclarified.com/Ph-Py/Physiology.html>.
104. **Μακροπούλου, Μ.** *Σημειώσεις Βιο-Οπτικής*. Αθήνα : Μεταπτυχιακό Πρόγραμμα "Ιατρική Φυσική - Ακτινοφυσική", Ενότητα: Η Φυσική στην Ιατρική, 2007.
105. **Mader, Sylvia S.** *Understanding Human Anatomy and Physiology, 5th Edition*. s.l. : McGraw-Hill , 2004. pp. 168 - 175.
106. **Chudler, Eric H.** The Blind Spot. *Neuroscience for Kids - University of Washington*. [Online] 10 11, 2011. [Cited: 10 16, 2011.] <http://faculty.washington.edu/chudler/chvision.html>.
107. **Kolb, Helga, et al.** Simple Anatomy Of The Retina. *Webvision*. [Online] [Cited: 10 15, 2011.] <http://webvision.med.utah.edu/book/part-i-foundations/simple-anatomy-of-the-retina/>.
108. **Fernandez, Eduardo, et al.** Webvision. *Physiology of Horizontal Cells - Miguel Hernández University of Elche*. [Online] 01 2001. [Cited: 10 15, 2001.] <http://retina.umh.es/webvision/HC1.html>.
109. **Roorda, Austin and Williams, David R.** The arrangement of the three cone classes in the living human eye. *Nature*. 1999, Vol. 397, pp. 520-523.
110. **Osterberg, G.** Topography of the layer of rods and cones in the human retina. *Acta Ophthalmology*. Suppl. VI, 1935, pp. 1-103.
111. **Purves D, Augustine GJ, Fitzpatrick D, et al., editors.** Neuroscience - 2nd edition. *NCBI Bookshelf*. [Online] 2001. [Cited: 10 16, 2011.] <http://www.ncbi.nlm.nih.gov/books/NBK10848/>.
112. **Edkins, Jo.** Shapes that tessellate. [Online] 2007. [Cited: 10 19, 2011.] <http://gwydir.demon.co.uk/jo/tess/grids.htm>.
113. **Curcio, Christina A., et al.** Human Photoreceptor Topography. *The Journal of Comparative Neurology*. 1992, Vol. 292, pp. 497-523.
114. **Kolb, H., et al.** Are there three types of horizontal cell in the human retina? *The Journal Of Comparative Neurology*. 1994, Vols. 353, No 3, pp. 370-386.
115. **Masland, Richard H.** The fundamental plan of the retina. *Nature Neuroscience*. 2011, Vols. 4, No. 9, pp. 877 - 886.
116. **Wikipedia.** Receptive Field. *Wikipedia, The Free Encyclopedia*. [Online] 09 31, 2011. [Cited: 10 17, 2011.] http://en.wikipedia.org/wiki/Receptive_field.

117. **Samuel, Leslie.** 034 How Lateral Inhibition Enhances Visual Edges. *Interactive Biology - Making Biology Fun*. [Online] [Cited: 10 17, 2011.] <http://www.interactive-biology.com/1886/how-lateral-inhibition-enhances-visual-edges—episode-34/>.
118. **Arkachar, Pradeep and Wagh, Meghanad D.** Criticality of lateral inhibition for edge enhancement in neural systems. *Neurocomputing*. 2007, Vol. 70, pp. 991-999.
119. **Kolb, Helga, et al.** Bipolar cell pathways in the vertebrate retina. *Webvision*. [Online] University of Utah. [Cited: 10 17, 2011.] <http://webvision.med.utah.edu/book/part-v-phototransduction-in-rods-and-cones/bipolar-cell-pathways-in-the-vertebrate-retina/>.
120. **Kolb, Helga, et al.** Circuitry for Rod Signals Through The Retina. *Webvision*. [Online] [Cited: 10 17, 2011.] <http://webvision.med.utah.edu/book/part-iii-retinal-circuits/circuitry-for-rod-cells-through-the-retina/>.
121. **Dacey, Dennis, et al.** Center surround receptive field structure of cone bipolar cells in primate retina. *Vision Research*. 2000, Vol. 40, pp. 1801-1811.
122. **Rodieck, R. W.** Quantitative analysis of cat retina ganglion cell response to visual stimuli. *Vision Research*. 1965, Vol. 5, pp. 583-601.
123. **Carandini, Matteo.** Receptive fields and suppressive fields in the early visual system. *The Cognitive Neurosciences (MIT Press)*. 3d ed., MS Gazzaniga, 2004, 3d ed., MS Gazzaniga, ed., pp 313-326.
124. **Devries, SH and Baylor, DA.** Mosaic arrangement of ganglion cell receptive fields in rabbit retina. *Journal of Neurophysiology*. 1997 Oct, Vols. 78, No. 4, pp. 2048-2060.
125. **Rowe, Michael H.** Trichromatic Color Vision in Primates. *News In Physiological Sciences*. 2002, Vols. 17, No. 3, pp. 94-98.
126. **Nelson, Ralph, et al.** Ganglion Cell Physiology. *Webvision*. [Online] University of Utah. [Cited: 10 17, 2011.] <http://webvision.med.utah.edu/book/part-ii-anatomy-and-physiology-of-the-retina/ganglion-cell-physiology/>.
127. **Douglas, Kevin.** What Is Lateral Inhibition? *Northern University Kentucky*. [Online] 10 14, 2004. [Cited: 10 17, 2011.] <http://www.nku.edu/~issues/illusions/LateralInhibition.htm>.
128. **Kolb, Helga, et al.** Feedback Loops. *Webvision*. [Online] [Cited: 10 17, 2011.] <http://webvision.med.utah.edu/book/part-iii-retinal-circuits/feedback-loops/>.
129. **Hubel, David.** Eye, Brain, and Vision - Amacrine Cells. *Eye, Brain, and Vision*. [Online] [Cited: 10 17, 2011.] <http://hubel.med.harvard.edu/book/b13.htm>.
130. **Eggers, Erika D. and Lukasiewicz, Peter D.** Interneuron Circuits Tune Inhibition in Retinal Bipolar Cells. *Journal of Neurophysiology*. 2010, Vols. 103, No. 1, pp. 25-37.
131. **Wikipedia.** Visual Cortex. *Wikipedia, The Free Encyclopedia*. [Online] 09 27, 2011. [Cited: 10 18, 2011.] http://en.wikipedia.org/wiki/Visual_cortex.
132. **Tootell, Roger B, et al.** Functional analysis of primary visual cortex (V1) in humans. *Proceedings of the National Academy of Sciences of the USA*. 1998, Vol. 95, pp. 811-817.

133. **Freeman, Jeremy and Merriam, Elisha.** Retinotopic Mapping. [Online] [Cited: 10 18, 2011.] http://jeremyfreeman.net/research/retinotopic_mapping.
134. **Adams, Daniel L. and Horton, Jonathan C.** A Precise Retinotopic Map of Primate Striate Cortex. *The Journal of Neuroscience*. 2003, Vols. 23, No. 9, pp. 3771-3789.
135. **Freeman, James A and Skapura, David M.** *Neural Networks Algorithms, Applications and Programming Techniques*. s.l. : Addison Wesley, 1991. pp. 89-106.
136. **Gardner, Justin L., et al.** Maps of Visual Space in Human Occipital Cortex Are Retinotopic, Not Spatiotopic. *The Journal of Neuroscience*. 2008, Vols. 28, No. 15, pp. 3988 - 3999.
137. **Hubel, D. H. and Wiesel, T.N.** Brain Mechanism of Vision. *Scientific American*. 1979, Vol. 241, pp. 150-162.
138. **DeAngelis, Gregory C., Ohzawa, Izumu and Freeman, Ralph D.** Receptive-field dynamics in the central visual pathways. *ScienceDirect - Trends in Neurosciences*. 1995, Vol. 18, pp. 451-458.
139. **Wagner, Brian Tomas and Kline, Donald.** The Sensorineural Basis of the Colour Experience. *Vision and Aging Lab (VAL) - University of Calgary*. [Online] [Cited: 10 18, 2011.] <http://psych.ucalgary.ca/pace/va-lab/Brian/neuralbases.htm>.
140. **Millodot, Michel.** *Dictionary of Optometry and Visual Science, 7th edition*. Edinburgh, New York : Butterworth-Heinemann Elsevier, 2009.
141. **Dowling, John E.** *Neurons and networks: an introduction to behavioral neuroscience*. s.l. : Harvard University Press, 2001. pp. 375 - 375.
142. **Gilbert, Cole.** Visual Neuroscience: Hypercomplex Cells in the Arthropod Visual System. *Current Biology* . 2007, Vols. 17, No. 11, pp. 412-414.
143. **Pfleeger, Shari Lawrence.** *Τεχνολογία Λογισμικού: Θεωρία και Πράξη (Software Engineering: Theory and Practice)*. s.l. : Κλειδάριθμος (Prentice Hall), 2001. pp.74-92.
144. **Jacobson, Ivar, Booch, Grady and Rumbaugh, James.** *The unified software development process*. s.l. : Addison-Wesley , 1999.
145. **Ambler, Scott W.** The Agile Unified Process v1.1. *Ambysoft Inc.* [Online] 05 13, 2006. [Cited: 10 18, 2011.] <http://www.ambysoft.com/downloads/agileUP.zip>.
146. **Geotechnical Software Services.** Unit Testing Guidelines. *Geotechnical Software Services*. [Online] 12 2010. [Cited: 10 18, 2011.] <http://geosoft.no/development/unittesting.html>.
147. **MSDN.** Unit Testing - Visual Studio.NET 2003. *Microsoft Corporation*. [Online] [Cited: 10 18, 2011.] [http://msdn.microsoft.com/en-us/library/aa292128\(v=vs.71\).aspx](http://msdn.microsoft.com/en-us/library/aa292128(v=vs.71).aspx).
148. **MSDN .** Integration Testing. *Microsoft Corporation*. [Online] [Cited: 10 18, 2011.] [http://msdn.microsoft.com/en-us/library/aa292128\(v=vs.71\).aspx](http://msdn.microsoft.com/en-us/library/aa292128(v=vs.71).aspx).
149. **Wirfs-Brock, Rebecca and Wilkerson, Brian.** Object-Oriented Design: A Responsibility-Driven Approach. *OOPSLA '89 Proceedings*. 1989, Vols. 24, No.10, pp. 71-75.

150. **Beck, Kent and Cunningham, Ward.** A Laboratory For Teaching Object-Oriented Thinking. *OOPSLA '89 Proceedings*. 1989, Vols. 24, No.10, pp. 1-6.
151. **Eeles, Peter, Houston, Kelli and Kozacsynski, Wojtek.** *Building J2EE Applications with the Rational Unified Process*. s.l. : Addison-Wesley, 2002. pp. 31-41.
152. **Jähne, Bernd.** *Digital Image Processing, 5th revised and extended edition*. Heidelberg : Springer, 2001. pp.283-314.
153. **Weisstein, Eric W.** Normal Distribution. *MathWorld - A Wolfram Web Resource*. [Online] [Cited: 10 21, 2011.] <http://mathworld.wolfram.com/NormalDistribution.html>.
154. **Minggang, Chen, et al.** Physiological properties of direction-selective ganglion cells in early postnatal and adult mouse retina. *The Journal Of Neurophysiology*. 2009, Vols. 587, No.4, pp. 819–828.
155. **Schach, Stephen P.** *Object-Oriented & Classical Software Engineering, 7th Edition*. s.l. : McGraw-Hill, 2007. pp. 469-471.
156. **Ambler, Scott W.** User Interface Prototypes. *Agile Modeling*. [Online] Ambyssoft Inc, 2009. [Cited: 10 29, 2011.] <http://www.agilemodeling.com/artifacts/uiPrototype.htm>.
157. **Da Vinci, Leonardo.** *Codex on the Flight of Birds*. 1505.
158. **Hagmann, Patric, et al.** Mapping the Structural Core of Human Cerebral Cortex. *PLoS Biology*. 2008, Vols. 6, No 7, e159, pp. 1499-1493.
159. **Εθνικό Μετσόβιο Πολυτεχνείο.** Ορολογία Μαθηματικής Λογικής Αγγλο-Ελληνικό λεξικό. [Online] [Cited: 10 14, 2011.] <http://www.mathphys.ntua.gr/logic/terms/dict.htm>.
160. **Anderson, Ryan.** How to Calculate Grayscale. [Online] 07 01, 2010. [Cited: 07 16, 2011.]
161. **Intel Corporation.** *Intel Desktop Board D410PT Technical Product Specification*. 2009. pp. 55-56.
162. **BrainConnection.com.** Overview of Receptive Fields. [Online] Posit Science Corporation. [Cited: 10 17, 2011.] <http://brainconnection.positscience.com/topics/?main=anat/receptive>.
163. **Heeger, David.** Perception Lecture Notes: Color. *Department of Psychology, New York University*. [Online] 2006. [Cited: 10 17, 2011.] <http://www.cns.nyu.edu/~david/courses/perception/lecturenotes/color/color.html>.
164. **Kolb, Helga; Nelson, Ralph; Fernandez, Eduardo; Bryan, Jones.** Roles of Amacrine Cells. *Webvision*. [Online] University of Utah. [Cited: 10 17, 2011.] <http://webvision.med.utah.edu/book/part-iii-retinal-circuits/roles-of-amacrine-cells/>.
165. **The Brain From Top To Bottom.** The Retina. [Online] [Cited: 10 18, 2011.] http://thebrain.mcgill.ca/flash/i/i_02/i_02_cl/i_02_cl_vis/i_02_cl_vis.html#3.
166. **Liu, C. - S.J., et al.** Magnocellular and Parvocellular Visual Pathways Have Different Blood Oxygen Level–Dependent Signal Time Courses in Human Primary Visual Cortex. *American Journal of Neuroradiology*. 2006, Vol. 27, pp. 1628-1634.

167. **Gottesdiener, Ellen.** RAD Realities: Beyond the hype to how RAD really works. *Application Development Trends*. 1995, Vols. 2, No. 8, pp. 28-38.

Πίνακας Εικόνων

1-2: Πρότυπα λογότυπα για το σύστημα στα Αγγλικά και στα Ελληνικά Σχεδίαση: Παντελής Ι. Καπλάνογλου	15
1-1: Έμβλημα εφαρμογής horasis.....	15
1-3: Βασικά στοιχεία του IDE της Delphi 7, στο επάνω μέρος της οθόνης.....	18
1-4: SQL Server Management Studio: Ένα ολοκληρωμένο περιβάλλον διαχείρισης και εκτέλεσης ερωτημάτων SQL.....	19
1-5: Η σκηνή φωτογράφισης του Horasis ©2011 Horasis	20
1-6: Βάση στήριξης δύο καμερών web με κανάλι-οδηγό ©2011 Horasis	20
2-1: Έντυπο του Εθνικού Κτηματολογίου στο οποίο τα στοιχεία του δικαιούχου συμπληρώνονται μέσα σε τετραγωνίδια.....	22
2-2: Το σύμβολο του βιομετρικού διαβατηρίου που εμφανίζεται στο εξώφυλλο των νέων διαβατηρίων.....	22
2-3: Φυσικοί γάντζοι (Zephyris/wikimedia commons)	23
2-4: Φυσικές σπείρες (arbyreed/Flickr)	23
2-5: Μια απλή εικόνα για την σύνθετη αίσθηση της όρασης	25
2-6: Σύγκριση κλασσικής ανάπτυξης με την RAD	27
2-7: Διάγραμμα Κλάσεων UML (Fowler & Scott).....	29
2-8: Υψηλή σύζευξη και χαμηλή συνοχή ©2011 Horasis.....	30
2-9: Χαμηλή σύζευξη και υψηλή συνοχή ©2011 Horasis	30
2-10: Πνεύμονες (Kurimoto, Fielding ,Musani).....	31
2-11: Σχηματική αναπαράσταση τετραπύρηνου CPU (www.newtechnology.co.in).....	33
2-12: Χάρτης λευκής ουσίας εγκεφάλου μέσω DSI (http://www.humanconnectomeproject.org/gallery/) Οπτικοποίηση: Vaughan Green, Λογισμικό TrackVis, Δεδομένα: Randy Buckner Ph.D	35
2-13: Αποτελέσματα fMRI (http://www.pbs.org/wgbh/nova/)	36
2-14: Στεφανιαία όψη με HARDI (http://www.humanconnectomeproject.org/gallery/) Οπτικοποίηση David Shattuck, Ph.D	36
3-1: Σχηματική αναπαράσταση βιολογικού νευρώνα ©2011 Horasis	37
3-2: Σχηματική αναπαράσταση σύναψης (http://www.niaaa.nih.gov) Τα μικρά σφαιρίδια αντιπροσωπεύουν μόρια νευροδιαβιβαστών	38
3-3: Ηλεκτρικά σήματα που παράγουν νευρώνες του ματιού (Kuffler) [53]	39
3-4: Οι διαφορετικοί τύποι συνάψεων. ©2011 Horasis	40
3-5: Σχηματική αναπαράσταση τεχνητού νευρώνα ©2011 Horasis.....	41
3-6: Γραφική παράσταση της βηματικής συνάρτησης	42
3-7: Διάφορες συναρτήσεις σιγμοειδούς μορφής με σύνολο τιμών το [-1,1] (Georg-Johann/Wikimedia Commons).....	42
3-8: Γραφική παράσταση της λογιστικής συνάρτησης με $\alpha_l = 0$, $\alpha_H = 1$, $b = 1$, $c = 0$	43
3-9: Γραφική παράσταση της σιγμοειδούς συνάρτησης (λογιστικής με $d = 1$).....	44

3-10: Σχηματική αναπαράσταση τεχνητού νευρώνα με σιγμοειδή ενεργοποίηση ©2011 Horasis	44
3-11: Ένα δίκτυο MLP του Horasis: Το στρώμα εισόδου με τρία διαφορετικά δεδομένα, χρωματισμένα με κόκκινο, ιώδες (μωβ) και γαλάζιο. Το κρυφό στρώμα 1 με πορτοκαλί, το κρυφό στρώμα 2 με κίτρινο και το στρώμα εξόδου με πράσινο. ©2011 Horasis	45
3-12: Σχηματική αναπαράσταση τεχνητού νευρώνα εισόδου σε δίκτυο MLP ©2011 Horasis.....	46
3-13: Διασπορά της τιμής εισόδου $\times 3$ σε όλους τους νευρώνες του κρυφού στρώματος 1 ©2011 Horasis.....	46
3-14: Κατάβαση Δυναμικού	49
3-15: Τοπικά Ελάχιστα	52
4-1: Σχηματική αναπαράσταση με στοιχεία UML, του αντικειμενοστραφούς τεχνητού νευρώνα ©2011 Horasis.....	57
4-2: Η αντικειμενοστραφής σχεδίαση τριών επιπέδων (Bartledan/Wikipedia Commons)	64
4-3: Αντικείμενο NeuralNetworkConfig, ο γονότυπος του νευρωνικού δικτύου ©2011 Horasis ..	67
5-2: Σελίδα (Tab) που εμφανίζει τους νευρώνες του στρώματος εισόδου	69
5-1: Πλήκτρο εντολής «Εκπαιδεύσεις Δικτύων» της εφαρμογής Horasis	69
5-3: Σελίδα (Tab) που εμφανίζει τους νευρώνες του κρυφού στρώματος 1.	70
5-4: TGVisualNeuron με $\delta_i > 10$	70
5-5: TGVisualNeuron με $\delta_i = 0$	70
5-6: Φόρμα Κατάσταση CPU	71
5-7: Ρυθμίσεις εκπαίδευσης των ΤΝΔ του Horasis.....	79
5-8: Εκπαίδευση ενός ΤΝΔ μέσα στο Horasis	80
5-9: Γράφημα μέσου τετραγωνικού σφάλματος της εκπαίδευσης.....	81
5-10: Σύγκριση γραφημάτων διαφορετικών εκπαιδεύσεων.....	82
5-11: Εμφάνιση σφαλμάτων των νευρώνων του κρυφού στρώματος 1, κατά την εκπαίδευση του Horasis.....	83
5-12: 7 βασικά χρώματα του ορατού φάσματος.....	83
5-13: Τρεις διαφορετικές προβολές της φόρμας επισκόπησης ενός νευρώνα	84
6-1: Σκηνή φωτογράφισης Horasis	85
6-2: Αισθητήρας CMOS Foveon X3	85
6-3: Διαστρωμάτωση αισθητήρα CMOS (http://www.steves-digicams.com)	86
6-4: Μωσαϊκό χρωμάτων του φίλτρου Bayer.....	86
6-5: Η απλή web κάμερα του horasis, Microsoft LifeCam HD-5000.....	87
6-6: Σχηματική αναπαράσταση της αρχιτεκτονικής DirectShow (MSDN)	88
6-7: Η ψηφιδωτή διαμόρφωση εικόνας BMP (Verpies/Wikimedia Commons)	89
6-8: Το αφαιρετικό σύστημα χρωμάτων.....	90
6-9: Το προσθετικό σύστημα χρωμάτων	90
6-10: Τροχός των χρωμάτων	90
6-11: Τα χρώματα του ορατού φάσματος από το υψηλότερο φάσμα προς το χαμηλότερο	91
6-12: Τιμές απόχρωσης H από 0 έως 360 στο χρωματικό μοντέλο HSL.....	91
6-13: Μετατόπιση κορεσμού και φωτεινότητας στο μοντέλο HSL	91
6-14: Τιμές κορεσμού σε ασπρόμαυρο από μια φωτογραφία ενός τριαντάφυλλου. Φαίνεται ως φωτεινότερη η περιοχή του άνθους έχοντας μεγαλύτερο κορεσμό από τα φύλλα και το κοτσάνι.	92

6-15: Τα έξι αμιγή χρώματα	95
6-16: Χρώματα από έρευνα ΧΚCD. Κάθε χρώμα έχει τον αντίστοιχο τίτλο που ανέφεραν οι συμμετέχοντες στην έρευνα και τον αντίστοιχο κωδικό χρώματος HTML (xkcd.com).....	95
6-17: Δειγματοληψία οπτικών προτύπων για ένα αντικείμενο στο Horasis.....	97
6-18: Λεπτομέρεια στήριξης καμερών στην σκηνή φωτογράφισης του Horasis	97
6-19: Η φόρμα λήψης φωτογραφιών στην εφαρμογή Horasis	98
6-20: Προεπισκόπηση αντικειμένου πριν την φωτογράφιση, στην φόρμα φωτογράφισης	98
6-21: Ζεύγος φωτογραφιών, αριστερής και δεξιάς κάμερας ως πρότυπο του αντικειμένου.	99
6-22: Μελλοντική υλοποίηση Horasis ως thin-client/server και η θέση του τμήματος Ι της αρχιτεκτονικής Horasis που υλοποιεί την λήψη φωτογραφιών ©2011 Horasis.....	101
7-1: Ο βολβός του ανθρώπινου οφθαλμού (Mader) [105]	103
7-2: Αναστροφή της πραγματικής εικόνας ενός τριαντάφυλλου (Mader) [105]	104
7-3: Σχηματική αναπαράσταση του αμφιβληστροειδούς χιτώνα. (Webvision) [107]	106
7-4: Κωνίο και ραβδίο (Madder) [105]	108
7-5: Ευαισθησίες τριών κωνίων S,M,L σε περιοχές του ορατού φάσματος.....	108
7-6: Κατανομή κωνίων και ραβδίων στον ανθρώπινο αμφιβληστροειδή κατά Osterberg. Με πράσινο φαίνονται τα κωνία και με ιώδες τα ραβδία (Purves et. al) [111].....	109
7-7: Κατανομή L,M,S κωνίων σε δύο υγιείς ανθρώπους (Roorda & Williams) [109].....	109
7-8: Ψηφιδώσεις κωνίων στον αμφιβληστροειδή. Με βέλη εμφανίζεται η παρουσία ραβδίων (Curcio et al.) [113].....	109
7-9: Χρωματισμένη φωτογραφία αμφιβληστροειδούς ως προς τους τρεις διαφορετικούς τύπους κωνίων.....	109
7-10: Τα δύο πρώτα είδη των οριζοντίων κυττάρων (Masland) [115]	110
7-11	110
7-12: Διαφορετικοί τύποι δίπολων κυττάρων (Masland) [115]	111
7-13: Σχηματική αναπαράσταση διαύλου ανταγωνισμού μπλε-κίτρινου (http://www.cns.nyu.edu [162]).....	111
7-14: Οργάνωση κυκλικού υποδεκτικού πεδίου με κέντρο και περιφέρεια.....	112
7-15: Αναμμένα και σβησμένα κέντρα σε κυκλικά υποδεκτικά πεδία (Brainconnection.com) [161]	112
7-16: Γκαουσιανές Αποκρίσεις κέντρου –περιφερείας (Rowe) [125]	113
7-17: Οργάνωση κέντρου περιφερείας γαγγλίου με χρωματικό ανταγωνισμό κόκκινου με πράσινο, R-GC	114
7-18: Συμπεριφορά R-GC	114
7-19: Αναπαράσταση μετατόπισης δέσμης φωτός μέσα σε υποδεκτικά πεδία γαγγλίων (Northern Kentucky University)[127].....	115
7-20: Διαφορετικοί τύποι βραχύνων κυττάρων (Masland) [115].....	116
7-21: Διαδρομή του οπτικού νεύρου (Madder) [105]	117
7-22: Θέση του οπτικού φλοιού στο εγκέφαλο, με πράσινο χρώμα	117
7-23: Μέσω του χιάσματος, το δεξί μάτι πηγαίνει στο αριστερό ημισφαίριο και το αριστερό μάτι στο δεξί ημισφαίριο	118
7-24: Ρετινοτοπικός Χάρτης με Α. Ενεργοποιήσεις τμημάτων του οπτικού φλοιού σε σχέση με κινούμενα ή σταθερά σημεία. Β. Εκκεντρότητα Γ. Αζιμούθιο (Gardner et al.) [136].....	118
7-25: Διέγερση V1S με φωτεινή ράβδο (The Brain From Top to Bottom) [164]	119
7-26: Δομή υποδεκτικών πεδίων GC και V1S. Διέγερση με + και αναστολή με – (βασισμένο σε DeAngelis et al.) [138].....	119

7-27: Σταθερό ανασταλτικό ερέθισμα σε ένα R-V1S (DeAngelis et al.) [138]	120
7-28: Συμπεριφορά του υποδεκτικού πεδίου των απλών V1 κυττάρων σε διαφορετικές χρονικές στιγμές από την έναρξη του ερεθίσματος (DeAngelis et al.) [138]	120
7-29: 7-30: Υπερστήλες (VAL - University Of Calgary) [139]	120
7-32: Γενική θεώρηση συνδέσεων GC με V1S ©2011 Horasis.....	121
7-31: Υπερστήλες ως προς τον προσανατολισμό του βασικού χρώματος στο υποδεκτικό πεδίο των απλών νευρώνων (Millodot) ([140]	121
7-33: Δομή υποδεκτικών πεδίων H1C (βασισμένο DeAngelis et al.) [138]	122
7-34: Διέγερση V1HC με ένα μικρό φωτεινό ευθύγραμμο τμήμα. (The Brain From Top to Bottom) [164]	122
8-1: Παραστατική σύγκριση, της αυξητικής ,της επαναληπτικής ανάπτυξης και του συνδυασμού και των δύο ©2011 Horasis.....	123
8-2: Οι αρχές της Agile Unified Process (Ambler) [145].....	124
8-3: Παραστατική απεικόνιση της αρχιτεκτονικής του συστήματος «Όρασις» Με τα μπλε βέλη κατεύθυνσης φαίνεται η ροή της οπτικής πληροφορίας και με τα πράσινα της εννοιολογικής. Τα τέσσερα χρωματιστά πλαίσια αντιπροσωπεύουν τα όρια των τεσσάρων τμημάτων του συστήματος. ©2011 Horasis	126
8-4: Ροή εργασίας για την φωτογράφιση ενός πραγματικού αντικειμένου.....	127
8-5: Ροή εργασίας για την μετατροπή της εικόνας σε εσωτερική αναπαράσταση	127
8-6: Ροή εργασίας για την εκπαίδευση του ΤΝΔ κατηγοριοποίησης.....	128
Εικόνα 8-7: Τμήμα II της αρχιτεκτονικής Horasis: Τεχνητός Αμφιβληστροειδής ©2011 Horasis. 130	
9-1: Αριστερά η προεπισκόπηση με ανάλυση 1280x720 και δεξιά η τελική εικόνα 500x500 pixels	133
9-2: Ιεραρχία κλάσης TBitmap	134
9-3: Η θέση και η λειτουργία της κλάσης TPhotoReceptors ©2011 Horasis.....	136
9-4: Τεχνητός οριζόντιος νευρώνας, μέσου όρου, τριών εισόδων.....	137
9-5: Χάρτης 13 υποδεκτικών πεδίων τεχνητών HC ©2011 Horasis	137
9-7: Η εικόνα Μετά την επεξεργασία από του τεχνητούς νευρώνες HC.....	138
9-6 Φέτες μιας περιοχής εικόνας ©2011 Horasis	138
9-8 Φέτα κέντρου 0 και φέτες περιφέρειας 1-12 ©2011 Horasis.....	138
9-9: Τα ΤΝΔ κωνίων (αριστερά) και ραβδίου (δεξιά) με τα 6 τεχνητά δίπολα κύτταρα στο στρώμα εξόδου και των δύο. ©2011 Horasis	139
9-10: Κεντρικά κωνία ©2011 Horasis.....	139
9-11: Γραφική παράσταση κανονικής κατανομής και αθροιστικής κανονικής κατανομής (cumulative distribution function)	140
9-12: Γραφική παράσταση των συναρτήσεων απόστασης από αμιγές χρώμα ©2011 Horasis..	141
9-13: Γραφική παράσταση των συναρτήσεων παραγωγής στόχων για ΤΝΔ BPC των κωνίων ©2011 Horasis.....	142
9-14: Γραφική παράσταση των συναρτήσεων απόστασης από απόλυτο μαύρο και απόλυτο λευκό ©2011 Horasis	143

9-15 : Γραφική παράσταση των συναρτήσεων παραγωγής στόχων για ΤΝΔ BPC ενός ραβδίου ©2011 Horasis	144
9-16 Εμφάνιση καναλιών πράσινου και κόκκινου. Με κίτρινο και γκρι φαίνεται η απροσδιοριστία, ενώ με μαύρο χαμηλή φωτεινότητα	144
9-17: Κόκκινο κέντρο και πράσινη περιφέρεια σε έναν χάρτη κυκλικού υποδεκτικού πεδίου τεχνητού GC ©2011 Horasis	145
9-18: Το ΤΝΔ τεχνητού γαγγλίου και οι συνάψεις του με τα τεχνητά δίπολα ©2011 Horasis....	146
9-19: Τεχνητός ρετινοτοπικός χάρτης με εξαγωνική ψηφίδωση. ©2011 Horasis.....	147
9-20: Βιολογική εξαγωνική ψηφίδωση κωνίων στον ανθρώπινο αμφιβληστροειδή (Curcio et. a) [113]	147
9-21: Τοποθέτηση χάρτη υποδεκτικού πεδίου στη κυψέλη του τεχνητού αμφιβληστροειδούς ©2011 Horasis	149
9-22: Κυκλικός δίσκος τεχνητού αμφιβληστροειδή που εμφανίζει την εξόδου των τεχνητών HC που βρίσκονται σε αυτόν.....	149
9-23: Συναρτήσεις υπολογισμού διέγερσης και αναστολής τεχνητού GC, κατά μήκος διαμέτρου κυκλικού υποδεκτικού πεδίου $s = 2$, $D = 61$ με διάμετρο κέντρου $C = 5$. ©2011 Horasis.....	150
9-24: Συναρτήσεις ποσότητας ενεργοποίησης για βασικό και ανταγωνιστικό χρώμα, κατά μήκος διαμέτρου κυκλικού υποδεκτικού πεδίου $s = 2$, $D = 61$ με διάμετρο κέντρου $C = 5$. ©2011 Horasis	151
9-25: Διάμετρος κυκλικού υποδεκτικού πεδίου σε ευθύγραμμα τμήματα.....	152
9-26: Τέσσερις κλάσεις προτύπων εκπαίδευσης ενός ΤΝΔ R-GC ©2011 Horasis	155
9-27: Μετάβαση από το πράσινο στο κόκκινο μεγαλύτερη της 06:00 ©2011 Horasis.....	155
9-28: Δύο ΤΝΔ-GC και η συνδεσμολογία τους με AC και τεχνητά δίπολα ©2011 Horasis.....	156
9-29: Υποδεκτικό πεδίο τεχνητού PV1S με 7 rexels.....	157
9-30: Επιλογή κεντρικών rexel για υποδεκτικά πεδία τεχνητών PV1S. Εμφανίζεται η επικάλυψη μεταξύ τριών ομάδων PV1S ©2011 Horasis	158
9-31: Στρώματα εξόδου διαφορετικών ΤΝΔ ανά χρώμα, και τεχνητή υπερστήλη των 60° (σε μωβ πλαίσιο) ©2011 Horasis	158
9-32: Διεγερση και αναστολή ενός ΤΝΔ R-PV1S με διάχυτο κόκκινο και πράσινο φως. ©2011 Horasis.....	158
9-34: Επικάλυψη μεταξύ τριών ομάδων MV1S ©2011 Horasis.....	160
9-33: Υποδεκτικό πεδίο τεχνητού MV1S με 19 rexels	160
9-35: Δύο κόκκινες μπάρες διαφορετικού μήκους που ενεργοποιούν παράλληλα ένα μικρό και ένα μεγάλο τεχνητό V1S. Κάτω εμφανίζονται οι τιμές ενεργοποίησης. ©2011 Horasis	160
9-36: Πρότυπα εκπαίδευσης ΤΝΔ τεχνητών PV1S ©2011 Horasis	161
9-37: Τα τεχνητά νευρωνικά δίκτυα του Horasis που μοντελοποιούν την όραση, από την BMP μορφή της εικόνας έως την εσωτερική αναπαράσταση της εικόνας. ©2011 Horasis	164
10-1: Εμφάνιση δείγματος χρωμάτων συγκεκριμένης τιμή του μπλε, με τα χρώματα κοντά στο κίτρινο να έχουν γκρι ή λευκό περίγραμμα. Όσο πληθαίνει προς το λευκό το περίγραμμα τόσο πιο κοντά βρίσκεται το χρώμα στην απόχρωση του αμιγούς κίτρινου.....	173

10-2: Φόρμα που δημιουργεί, εμφανίζει και ανακαλεί ένα πρότυπο εκπαίδευσης των τεχνητών BPC	174
10-3: Φόρμα που δημιουργεί, εμφανίζει και ανακαλεί ένα πρότυπο εκπαίδευσης των τεχνητών GC	175
10-4: Καθορισμός τύπου τεχνητών V1S ως συνδυασμός δύο επιλογών	176
10-5: Φόρμα που δημιουργεί, εμφανίζει και ανακαλεί πρότυπα 12 διαφορετικών τύπων τεχνητών V1S.....	176
10-6: Οπτικοποίηση ενός προτύπου εκπαίδευσης τεχνητού B-MV1S, με τύπο ΤΝΔ ηnkV1SimpleCellsMagnaBlue. Ο συνδυασμός χρωμάτων είναι αμιγές μπλε και αμιγές κίτρινο, η γωνία 0° και το πάχος της μπάρας 24 pixels. Βλέπουμε 19 rexel που θα δώσουν τιμές σε 152 εισόδους.....	177
10-7: Οπτικοποίηση ενός προτύπου εκπαίδευσης τεχνητού B-MV1S με διάχυτο φωτισμό ανταγωνιστικού χρώματος.	177
10-8: Οπτικοποίηση ενός προτύπου εκπαίδευσης τεχνητού B-MV1S με διάχυτο φωτισμό βασικού χρώματος.	178
10-9: Οπτικοποίηση προτύπου για B-MV1S με γωνία 30° και πάχος μπάρας 24 pixels, αμιγές μπλε πάνω σε αμιγές κίτρινο.....	178
10-10: Οπτικοποίηση προτύπου για B-MV1S με γωνία 30° και πάχος μπάρας 12 pixels, αμιγές μπλε πάνω σε αμιγές κίτρινο.....	178
10-11: Οπτικοποίηση προτύπου για B-MV1S με γωνία 30° και πάχος μπάρας 24 pixels, αμιγές μπλε πάνω σε πορτοκαλί.....	178
10-12: Οπτικοποίηση προτύπου για R-PV1S με γωνία 30° και πάχος μπάρας 12 pixels, αμιγές κόκκινο πάνω σε αμιγές πράσινο. Βλέπουμε την ύπαρξη 7 rexels που θα χρησιμοποιηθούν σε 56 εισόδους.....	180
10-13: 10-14: Φόρμα που δημιουργεί, εμφανίζει και ανακαλεί πρότυπα εκπαίδευσης του δικτύου ενοποίησης της απόκρισης των τεχνητών V1S	181
10-15: Η φόρμα οπτικοποίησης εσωτερικής λειτουργίας του Horasis, εμφανίζοντας την αρχική εικόνα BMP	185
10-16: Δύο διαφορετικά ενδιάμεσα στάδια κατά την οπτικοποίηση της ανάπτυξης του τεχνητού ρετινοτοπικού χάρτη.....	185
10-17: Λεπτομέρεια οπτικοποίησης τεχνητών HC στην οποία εμφανίζεται η παρουσία του χάρτη υποδεκτικού πεδίου το κέντρο του και οι φέτες της περιφέρειας του.	186
10-18: Οπτικοποίηση εξόδου τεχνητών L-GC αριστερά και των ανταγωνιστικών D-GC δεξιά.....	187
10-19: Οπτικοποίηση εξόδου τεχνητών R-GC αριστερά και των ανταγωνιστικών G-GC δεξιά	188
10-20 Οπτικοποίηση εξόδου τεχνητών B-GC αριστερά και των ανταγωνιστικών Y-GC δεξιά.....	188
10-21: Οπτικοποίηση τεχνητών R-GC χωρίς φιλτράρισμα εισόδων μέσω τεχνητών AC αριστερά, και με χρήση φίλτρου τεχνητών AC δεξιά	189
10-22: Φόρμα που χρησιμοποιείται για μετατροπή των εικόνων BMP ενός αντικειμένου σε εσωτερική αναπαράσταση της εικόνας.....	191
10-23: Φόρμα προτύπων πραγματικών αντικειμένων και ανάκλησης ΤΝΔ εξόδου του Horasis.	194