

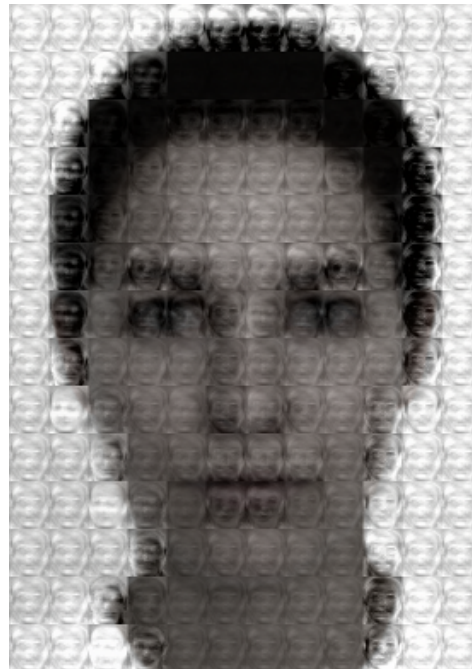


ΑΛΕΞΑΝΔΡΕΙΟ Τ.Ε.Ι. ΘΕΣΣΑΛΟΝΙΚΗΣ
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ



ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΑΝΑΓΝΩΡΙΣΗ ΠΡΟΣΩΠΩΝ ΜΕ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ



Του φοιτητή

Καραγιαννίδη Σάββα
Αρ. Μητρώου 410

Επιβλέπων καθηγητής

Διαμαντάρας Κωνσταντίνος

Θεσσαλονίκη, 2009

Πτυχιακή εργασία του φοιτητή Καραγιαννίδη Σάββα

Πτυχιακή εργασία του φοιτητή Καραγιαννίδη Σάββα

Αναγνώριση προσώπων με νευρωνικά δίκτυα

Καραγιαννίδης Σάββας

Πτυχιακή εργασία

Τμήμα Πληροφορικής

Τεχνολογικό Εκπαιδευτικό Ίδρυμα Θεσσαλονίκης

Περίληψη

Η αναγνώριση των ανθρωπίνων προσώπων αποτελεί ένα σύστημα βιομετρικής ταυτοποίησης το οποίο παρουσιάζει ιδιαίτερο ενδιαφέρον κυρίως λόγω της ευκολίας εφαρμογής ενός τέτοιου συστήματος. Τα συστήματα αναγνώρισης προσώπων αναγνωρίζουν και ταυτοποιούν αυτόματα ένα άτομο από μία ψηφιακή εικόνα ή από ένα καρέ μίας πηγής video. Για να το επιτύχει αυτό ένα τέτοιο σύστημα χρησιμοποιεί συνήθως μία μέθοδο εξαγωγής των χαρακτηριστικών του προσώπου και στη συνέχεια τη σύγκριση των χαρακτηριστικών αυτών με μία βάση αντίστοιχων δεδομένων συγκεκριμένων ατόμων. Στην εργασία αυτή παρουσιάζουμε τις διάφορες μεθόδους αναγνώρισης προσώπων και αναλύουμε και υλοποιούμε ένα σύστημα αναγνώρισης προσώπων χρησιμοποιώντας την τεχνική ανάλυσης κυρίων συνιστωσών και υλοποιώντας την με τεχνητά νευρωνικά δίκτυα APEX.

Για την αναγνώριση υπάρχει ένα στάδιο προεπεξεργασίας της εικόνας, με σκοπό την εξαγωγή της περιοχής του προσώπου από το σκηνικό. Στη συνέχεια κάθε άτομο περιγράφεται με ένα σύνολο εικόνων που προκύπτουν από την ανάλυση των κυρίων συνιστωσών του συνόλου των εικόνων προσώπου του συγκεκριμένου ατόμου. Χρησιμοποιώντας ένα σύνολο στατιστικών μετρήσεων υπολογίζουμε τα διανύσματα των χαρακτηριστικών από τις εικόνες της ανάλυσης. Κατόπιν,

χρησιμοποιώντας ένα μέτρο διανυσματικής απόστασης, υπολογίζονται οι αποστάσεις μεταξύ του διανύσματος του προσώπου εισόδου και όλων των διανυσμάτων των προσώπων της βάσης. Η μικρότερη απόσταση ορίζει την κοντινότερη κλάση προσώπου που μοιάζει με το πρόσωπο υπό ταξινόμηση.

Πτυχιακή εργασία του φοιτητή Καραγιαννίδη Σάββα

Face recognition using neural networks

Karagiannidis Savvas

Dissertation

Department of Informatics

Technological Educational Institute of Thessaloniki

Abstract

The person recognition of human faces is a biometric identification system which is of particular interest mainly because of the ease of implementing such a system. The facial recognition systems automatically recognize and identify a person from a digital image or a frame from a digital video source. To achieve this one such system uses a method of extraction of characteristics and then compares these characteristics with a corresponding database of individuals. This paper presents the various methods of face recognition and analyzes and implements a recognition system using a technique of principal components analysis and implementing with the APEX artificial neural networks.

The identification requires an image preprocessing stage in order to extract the face region from the scene and the background. Then each person is described by a set of images obtained from the analysis of principal components of the facial images of that person. Using a set of statistical measurements we calculate the vectors of features from the images analysis. Then, using a vector distance measure, we calculate the distances between the input vector and all persons in the database. The shortest distance from a class provides the nearest person who looks like the person to be classified.

Ευχαριστίες

Έχουν περάσει αρκετά χρόνια από τότε που ξεκίνησε η συνεργασία μου με τον επόπτη καθηγητή μου κ. Διαμαντάρα, ο οποίος όλο αυτό το διάστημα ήταν πάντα πρόθυμος και ανοιχτός στις συζητήσεις. Με την καθοδήγηση και τις συμβουλές του κατάφερα να ξεπεράσω όποια προβλήματα παρουσιάστηκαν κατά τη διάρκεια της εργασίας μου αυτής. Δίπλα του κατάφερα να αποκτήσω πολλές και ποικίλες γνώσεις τόσο σε επιστημονικό, όσο και σε ερευνητικό επίπεδο. Έτσι ευχαριστώ θερμά τον κ. Κωνσταντίνο Διαμαντάρα που με στήριξε σε αυτή μου την προσπάθεια.

Επίσης ευχαριστώ όλους τους καθηγητές του τμήματος Πληροφορικής του Αλεξάνδρειου Τεχνολογικού Εκπαιδευτικού Ιδρύματος οι οποίοι με στήριξαν και μου πρόσφεραν κατά τη διάρκεια των σπουδών μου τις απαραίτητες γνώσεις για μία επιτυχημένη καριέρα.

Κλείνοντας θα ήθελα να ευχαριστήσω τη μητέρα μου, η οποία με θυσίες και ατέλειωτη υπομονή με στήριξε όλα αυτά τα χρόνια.

Καραγιαννίδης Σάββας

Πίνακας Περιεχομένων

Περίληψη.....	3
Abstract	5
Ευχαριστίες.....	6
Πίνακας Περιεχομένων.....	7
1. Εισαγωγή.....	10
1.1. Περιγραφή του προβλήματος.....	10
1.1.1. Ανίχνευση προσώπου.....	11
1.1.2. Κανονικοποίηση προσώπου.....	12
1.1.3. Αναγνώριση προσώπων με μία μόνο εικόνα.....	15
1.2. Εφαρμογές.....	16
1.2.1. Συστήματα ασφαλείας και παρακολούθησης.....	16
1.2.2. Έλεγχος πρόσβασης	16
1.2.3. Έξυπνες κάρτες	17
1.2.4. Συστήματα επιβολής του νόμου.....	18
1.2.5. Αλληλεπίδραση ανθρώπου - υπολογιστή.....	18
1.2.6. Ευρετηριασμός βάσεων δεδομένων πολυμέσων	19
2. Επιλογές για τη λύση του προβλήματος.....	20
2.1. Ολιστικές Μέθοδοι.....	20
2.1.1. Μέθοδος Eigenface.....	20
2.1.2. Μέθοδος Fisherface.....	21
2.1.3. Μέθοδος Μηχανών διανυσμάτων υποστήριξης.....	23
2.1.4. Ανάλυση ανεξάρτητων συνιστωσών.....	24
2.1.5. Γραμμική διακριτική ανάλυση.....	24
2.1.6. Πιθανοθεωρητικά Νευρωνικά δίκτυα βασισμένα σε αποφάσεις.....	25
2.2. Μέθοδοι βασισμένες σε χαρακτηριστικά ή Αναλυτικές μέθοδοι.....	25
2.2.1. Καθαρά γεωμετρικές μέθοδοι 2Δ ή 3Δ.....	26
2.2.2. Αντιστοίχιση δυναμικών συνδέσμων (Dynamic Link Matching).....	27
2.2.3. Συνελκτικά νευρωνικά δίκτυα.....	27
2.3. Υβριδικές μέθοδοι.....	28

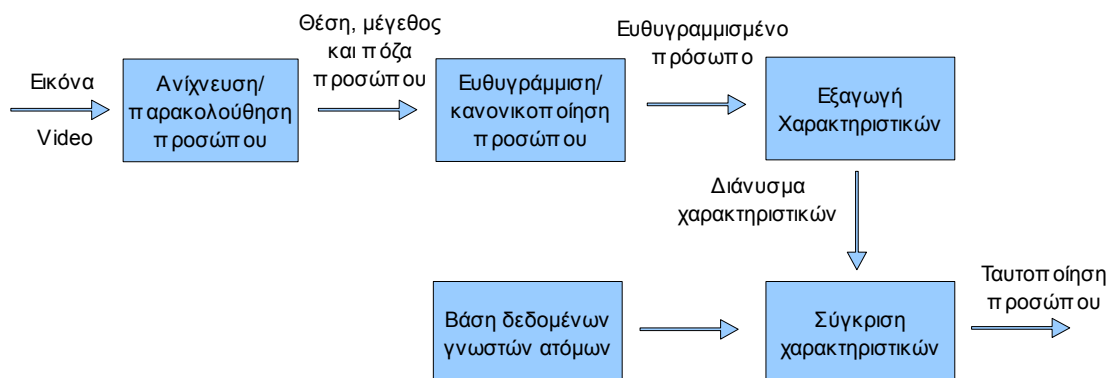
2.3.1. Modular eigenfaces, eigenmodules.....	28
2.3.2. Υβριδική ανάλυση τοπικών χαρακτηριστικών (Hybrid LFA).....	29
2.3.3. Αναγνώριση προσώπων με ευέλικτα εμφανισιακά μοντέλα	30
3. Η ανάλυση κυρίων συνιστωσών ως μέθοδος αναγνώρισης προσώπων.....	32
3.1. Μέθοδος ανάλυσης κυρίων συνιστωσών.....	32
3.1.1. Ιστορικό και θεωρία της ανάλυσης κυρίων συνιστωσών.....	34
3.1.2. Ο κανόνας του Hebb.....	37
3.2. Βασικές αρχές της αυτοοργάνωσης.....	40
3.3. Γραμμικός νευρώνας ως υλοποίηση εξαγωγής της πρώτης ιδιοτιμής....	42
3.4. Αυτοοργανούμενη Ανάλυση Κυρίων Συνιστωσών.....	45
3.5. Προσαρμοστική Ανάλυση Κυρίων Συνιστωσών με Πλευρική Αναστολή.	51
3.6. Η βέλτιστη παράμετρος εκπαίδευσης β	55
3.7. Άλλες μέθοδοι ανάλυσης κυρίων συνιστωσών.....	56
3.7.1. Ο κανόνας του Földiák.....	56
3.7.2. Ο κανόνας του υποχώρου (subspace rule).....	57
3.8. Αναγνώριση - ταξινόμηση νέων προτύπων.....	58
3.9. Άλλες εφαρμογές της ανάλυσης κυρίων συνιστωσών.....	59
3.9.1. Συμπύεση εικόνας.....	60
3.9.2. Ανίχνευση επιθέσεων σε δίκτυα υπολογιστών.....	63
4. Υλοποίηση του συστήματος αναγνώρισης προσώπων.....	65
4.1. Τεχνικά στοιχεία υλοποίησης της εφαρμογής.....	65
4.2. Περιγραφή των εννοιών/αντικειμένων της εφαρμογής.....	66
4.2.1. Εικόνες.....	66
4.2.2. Λίστες εικόνων.....	68
4.2.3. Νευρωνικό δίκτυο.....	70
4.2.4. Εκπαίδευση του νευρωνικού δικτύου.....	72
4.2.5. Σύστημα αναγνώρισης - ταξινόμησης προτύπων.....	74
4.2.6. Γεννήτρια αναφορών / εξαγωγή δεδομένων.....	74
4.3. Ροή λειτουργίας της εφαρμογής.....	74
4.3.1. Εισαγωγή προτύπων εκπαίδευσης.....	74
4.3.3. Οδηγός (wizard) δημιουργίας συστήματος αναγνώρισης.....	75

4.3.4. Εκπαίδευση των δικτύων.....	78
4.3.5. Αναγνώριση άγνωστων προτύπων.....	80
5. Πειραματικά δεδομένα.....	83
5.1. Βάσεις Προσώπων.....	83
5.2. Υλοποίηση συστήματος αναγνώρισης της βάσης προσώπων AT&T (πρώην Olivetti Research Labs).....	87
5.3. Αναγνώριση αλλοιωμένων προτύπων.....	90
5.4. Εκπαίδευση των δικτύων με τμήμα μόνο της βάσης AT&T.....	92
5.5. Υλοποίηση συστήματος αναγνώρισης της βάσης προσώπων Yale A....	93
5.6. Υλοποίηση συστήματος αναγνώρισης της βάσης προσώπων Essex (Grimaces, Faces95).....	94
5.6.1. Grimaces.....	95
5.6.2. Faces95.....	95
6. Συμπεράσματα.....	98
7. Βιβλιογραφία.....	102

1. Εισαγωγή

1.1. Περιγραφή του προβλήματος

Στη σύγχρονη ζωή μας η ανάγκη για προσωπική ασφάλεια και έλεγχο πρόσβασης κατέχει πλέον ιδιαίτερα σημαντικό ρόλο. Η βιομετρία είναι η τεχνολογία η οποία αναμένεται να αντικαταστήσει τις παραδοσιακές μεθόδους αυθεντικοποίησης τα διαπιστευτήρια των οποίων μπορούν πολύ εύκολα να υποκλαπούν, να αντιγραφούν ή να ξεχαστούν. Τα δαχτυλικά αποτυπώματα, το πρόσωπο, η ίριδα του ματιού και το φωνητικό αποτύπωμα είναι κάποια από τα χαρακτηριστικά τα οποία χρησιμοποιούνται από τη βιομετρία. Από τις παραπάνω μεθόδους η αναγνώριση προσώπου αποτελεί τον πλέον άμεσο, φιλικό προς το χρήστη και βολικό τρόπο αυθεντικοποίησης, μια που δεν απαιτεί ούτε ιδιαίτερο εξοπλισμό, ούτε άμεση επαφή και συνεργασία του χρήστη με το σύστημα. Γι' αυτό η αναγνώριση ανθρωπίνων προσώπων αποτελεί μαζί με την ανίχνευση προσώπου, ιδιαίτερα τα τελευταία χρόνια, αντικείμενο μελέτης για πολλές εφαρμογές στον τομέα της ανάλυσης εικόνων και των πολυμέσων. Η διαδικασία της αναγνώρισης προσώπων για να μπορεί να χρησιμοποιηθεί με τρόπο χρήσιμο και αξιόπιστο προϋποθέτει μία σειρά από στάδια επεξεργασίας των δεδομένων εισόδου όπως το στάδιο της ανίχνευσης του προσώπου σε ψηφιακή εικόνα και της κανονικοποίησης της εικόνας του προσώπου.



Σχήμα 1: Διάγραμμα ροής της διαδικασίας αναγνώρισης προσώπου

1.1.1. Ανίχνευση προσώπου

Το στάδιο της ανίχνευσης του προσώπου ως τμήμα μίας εικόνας αποτελεί θεμελιώδες βήμα και βασική προϋπόθεση για την αυτόματη αναγνώριση προσώπων. Υπάρχουν διάφορες μέθοδοι ανίχνευσης προσώπου είτε σε ψηφιακή εικόνα είτε σε ακολουθία video. Ειδικά για τη δεύτερη περίπτωση όπου τα δεδομένα είναι video της μορφής MPEG έχουν αναπτυχθεί αρκετά αποδοτικοί αλγόριθμοι οι οποίοι εκμεταλλεύονται την ήδη υπάρχουσα συμπίεση των δεδομένων και μπορούν να επιτύχουν εξαιρετικές επιδόσεις στον τομέα της ταχύτητας με ανίχνευση προσώπων σε πραγματικό χρόνο (Wang and Chang 1997; Acosta et al. 2002) . Οι μέχρι τώρα προσεγγίσεις στο πρόβλημα της ανίχνευσης μπορούν να ταξινομηθούν σε τρεις κατηγορίες:

Φόρμες συσχετίσεων (Correlation Templates): Υπολογίζεται ένα μέτρο διαφοράς μεταξύ μιας φόρμας και ενός υποψήφιου τμήματος της εικόνας. Με χρήση κάποιων κατωφλίων προκύπτει η τελική απόφαση. Η ίδια ιδέα έχει εφαρμοστεί και αντί για ανίχνευση ολόκληρου του προσώπου, για ανίχνευση μερικών μόνο χαρακτηριστικών του.

Φόρμες με δυνατότητα παραμόρφωσης (Deformable Templates): Μέθοδος παρόμοια με την προηγούμενη αλλά επιπλέον υπάρχει η δυνατότητα δυναμικής παραμόρφωσης της φόρμας ώστε αυτή να πετυχαίνει το καλύτερο δυνατό ταίριασμα πάνω στο αντικείμενο το οποίο ανιχνεύει. Η αποτελεσματικότητά της έγκειται στη χρήση παραμετρικών καμπυλών και επιφανειών για τη μοντελοποίηση του προσώπου ή και χαρακτηριστικών του. Οι καμπύλες και οι επιφάνειες προσαρμόζονται ελαστικά σε ένα γενικό μοντέλο, στο οποίο επιτρέπονται μικρές μεταβολές στη θέση των χαρακτηριστικών. Κατά τη διαδικασία της ανίχνευσης γίνεται αντιστοίχιση των καμπυλών και των επιφανειών με τα χαρακτηριστικά της εικόνας με στόχο την εύρεση της θέσης, όπου η παραπάνω μεταβολή ελαχιστοποιείται.

Αναλλοίωτα χαρακτηριστικά εικόνων (Image Invariants): Στις μεθόδους αυτής της κατηγορίας γίνεται η υπόθεση ότι υπάρχει ένα πλήθος χωρικών σχέσεων κοινό και πιθανότατα μοναδικό μεταξύ όλων των πρότυπων προσώπων.

Οι σχέσεις αυτές διατηρούνται ακόμη και κάτω από διαφορετικές συνθήκες φωτισμού. Για την ανίχνευση πρέπει πρώτα να εξαχθούν κάποια πρότυπα χαρακτηριστικά από εικόνες προσώπων της εκπαίδευσης. Έπειτα, η ανίχνευση γίνεται σε όλες τις δυνατές θέσεις της εικόνας. Οι μέθοδοι αυτής της κατηγορίας βασίζονται αποκλειστικά στον υπολογισμό κάποιων γεωμετρικών ποσοτήτων (αποστάσεις, στατιστικές ποσότητες κ.τ.λ.) από την εικόνα δημιουργώντας έτσι ένα διάνυσμα χαρακτηριστικών, βάση του οποίου λαμβάνεται η απόφαση για την ύπαρξη ή όχι προσώπου.

Η ανίχνευση προσώπου από μόνη της έχει πολλές εφαρμογές σήμερα. Χρησιμοποιείται ευρέως στις ψηφιακές φωτογραφικές μηχανές, σε συστήματα παρακολούθησης αλλά και σε εφαρμογές ψηφιακής επεξεργασίας εικόνας και video (Ken Burns effect).



Εικόνα 1: Παραδείγματα ανίχνευσης προσώπων σε διάφορες εικόνες.

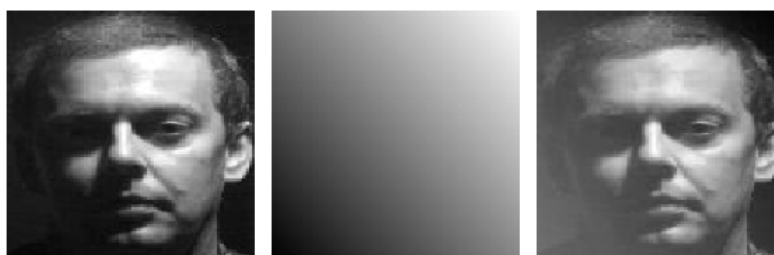
1.1.2. Κανονικοποίηση προσώπου

Από τη στιγμή που έχει εντοπιστεί το τμήμα της εικόνας το οποίο αποτελεί το πρόσωπο που μας ενδιαφέρει, και πριν εφαρμόσουμε τη μέθοδο της

αναγνώρισης του προσώπου ή εξαγωγής των χαρακτηριστικών του, θα πρέπει η εικόνα του προσώπου να κανονικοποιηθεί ως προς:

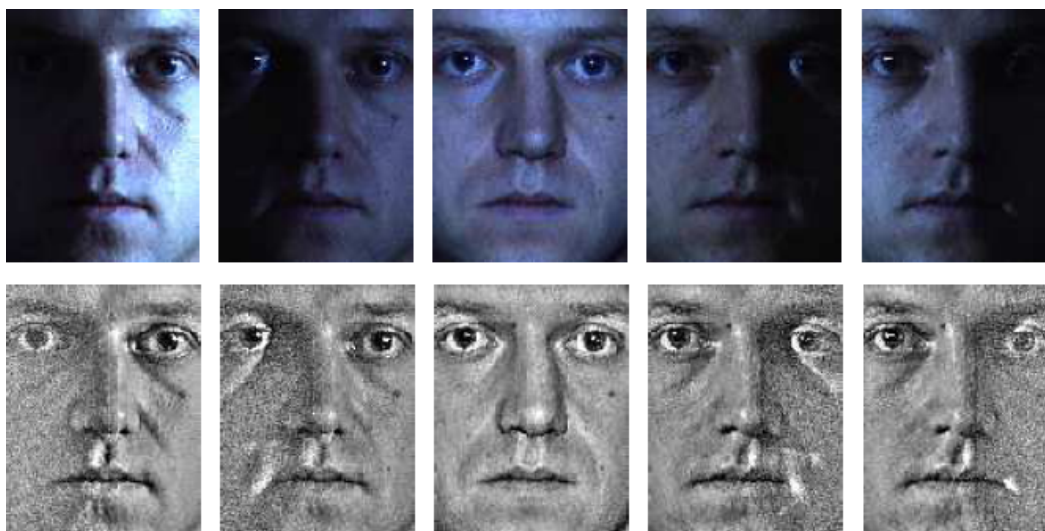
Το μέγεθος της εικόνας. Συνήθως το μέγεθος της εικόνας το οποίο απαιτείται από τη μέθοδο αναγνώρισης είναι σταθερό. Η εικόνα του προσώπου θα πρέπει να σμικρυνθεί ή να μεγεθυνθεί ανάλογα. Στην περίπτωση που η ανάλυση της εικόνας είναι χαμηλότερη από αυτή που αναμένεται πιθανό να έχουμε κάποια μείωση της αξιοπιστίας του αποτελέσματος.

Τη φωτεινότητα και το φωτισμό της εικόνας. Το πρόσωπο θα πρέπει να εμφανίζεται φωτιζόμενο ομοιόμορφα από μπροστά και όχι από το πλάι ή από πίσω χωρίς επίσης έντονο ή πολύ χαμηλό φωτισμό. Η διόρθωση της εικόνας από έντονο ή πολύ χαμηλό φωτισμό επιτυγχάνεται είτε εφαρμόζοντας μία απλή γραμμική προσαρμογή είτε με εξίσωση της εικόνας με βάση το ιστόγραμμα της. Για τη διόρθωση του φωτισμού μπορεί να χρησιμοποιηθεί ένα επίπεδο με γραμμική διαβάθμιση της μέσης τετραγωνικής φωτεινότητας της εικόνας. Συνδυάζοντας τις δύο εικόνες το αποτέλεσμα έχει ομοιόμορφα κατανομημένο φωτισμό.



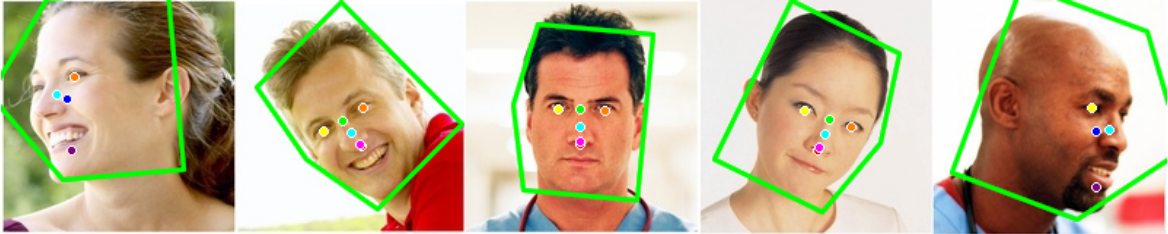
(α) (β) (γ)
Εικόνα 2: Διόρθωση φωτεινότητας εικόνας. (α) Αρχική εικόνα (β) Επίπεδο βέλτιστης φωτεινότητας (γ) Διαφορά (α)-(β)

Ακόμα καλύτερα αποτελέσματα επιτυγχάνονται χρησιμοποιώντας ακόμα πιο πολύπλοκες μεθόδους οι οποίες λαμβάνουν υπόψιν και τη συμμετρία του προσώπου, βελτιώνοντας την αποτελεσματικότητα του συστήματος αναγνώρισης.



Εικόνα 3: Επάνω γραμμή: Εικόνες ενός προσώπου κάτω από διαφορετικές συνθήκες φωτισμού (δείγμα από PIE face database). Κάτω γραμμή: Οι ίδιες εικόνες αφού έχουν υποστεί κατάλληλη επεξεργασία για την εξάλειψη των μεταβολών λόγω φωτισμού.

Γωνία και πόζα του προσώπου. Αν το επίπεδο του προσώπου παραμένει ευθυγραμμισμένο, η εικόνα μπορεί να διορθωθεί με απλή περιστροφή. Αν το επίπεδο του προσώπου έχει μικρή κλίση σε σχέση με το επίπεδο της εικόνας, τότε με τεχνικές morphing και λαμβάνοντας υπόψη τη γεωμετρία του προσώπου μπορούμε να διορθώσουμε την εικόνα και να λάβουμε ένα πρόσωπο το οποίο κοιτάει ευθεία μπροστά. Συνήθως η μέθοδος που ακολουθείται είναι αρχικά να εντοπιστεί η κλίση του προσώπου σε σχέση με την κάμερα και στη συνέχεια, ανάλογα με την κλίση να εντοπιστούν τα κατάλληλα εκείνα χαρακτηριστικά του προσώπου, τα οποία, ανάλογα με τη γωνία, μπορεί να μην είναι όλα ορατά. Σε πολλές περιπτώσεις όπου η γωνία του προσώπου είναι πολύ μεγάλη η ανακατασκευή με τον τρόπο αυτό είναι αδύνατη.



Εικόνα 4: Παράδειγμα εντοπισμού της κλίσης του προσώπου και χαρακτηριστικών σημείων αυτού. Παρατηρούμε ότι στην πρώτη και την τελευταία εικόνα, όπου η γωνία κλίσης του κεφαλιού είναι πολύ μεγάλη, τα χαρακτηριστικά που εντοπίζονται είναι διαφορετικά από ότι στις υπόλοιπες εικόνες.

Έκφραση και σχήμα του προσώπου. Μπορούμε να κανονικοποιήσουμε περαιτέρω την εικόνα του προσώπου, παραμορφώνοντάς την χρησιμοποιώντας κάποια σημεία αναφοράς του προσώπου και προσαρμόζοντας το σχήμα και την έκφραση του προσώπου σε αναμενόμενα πρότυπα ώστε η αναγνώριση να έχει μεγαλύτερη πιθανότητα επιτυχίας. Η μέθοδος αυτή θα πρέπει να χρησιμοποιηθεί με ιδιαίτερη προσοχή ώστε να διατηρηθούν αναλλοίωτες σημαντικές πληροφορίες που μπορούν να συμβάλουν σημαντικά στη μέθοδο αναγνώρισης όπως είναι το σχήμα του προσώπου.

1.1.3. Αναγνώριση προσώπων με μία μόνο εικόνα

Ένας πολύ σημαντικός παράγοντας για την αποτελεσματικότητα ενός συστήματος αναγνώρισης προσώπων είναι το πλήθος του συνόλου των εικόνων με βάση τις οποίες το σύστημα έχει εκπαιδευτεί. Σε ορισμένες όμως περιπτώσεις, δεν έχουμε στη διάθεσή μας αρκετά μεγάλο πλήθος από εικόνες ή μπορεί ακόμα και να έχουμε μία μόνο εικόνα για κάθε άτομο.

Στην περίπτωση αυτή έχουν αναπτυχθεί ειδικές τεχνικές, με τις οποίες μπορούμε από τη μία αυτή εικόνα να παράγουμε ένα σύνολο από εικόνες, το οποίο να πληρεί τις προϋποθέσεις για την εκπαίδευση του συστήματος αναγνώρισης. Για να παραχθούν αυτές οι εικόνες, χρησιμοποιούνται ειδικές τεχνικές επεξεργασίας της αρχικής εικόνας (μικρή τυχαία περιστροφή, προσομοίωση μεταβολής του φωτισμού, προσομοίωση μικρής αλλαγής της πόζας του προσώπου, προσομοίωση μεταβολής της έκφρασης).

1.2. Εφαρμογές

Οι κυριότεροι λόγοι για τους οποίους το πρόσωπο αποτελεί σήμερα σημαντικό στοιχείο επεξεργασίας μπορούν να χωριστούν σε τέσσερις κυρίως κατηγορίες. Η πρώτη είναι τα συστήματα ασφαλείας, η δεύτερη η ταυτοποίηση με χρήση έξυπνων καρτών, η τρίτη τα συστήματα παρακολούθησης και επιβολής του νόμου και τέλος ο τομέας της ψυχαγωγίας.

1.2.1. Συστήματα ασφαλείας και παρακολούθησης

Η κυριότερη ίσως εφαρμογή η οποία και ώθησε ιδιαίτερα την έρευνα στον τομέα της αναγνώρισης του ανθρωπίνου προσώπου και η οποία παρουσιάζει και τις περισσότερες προκλήσεις είναι η εφαρμογή σε συστήματα ασφαλείας. Σήμερα περισσότερο από ποτέ η ασφάλεια των επιβατών και του προσωπικού είναι η κύρια ανησυχία σε αεροδρόμια σε όλο τον κόσμο. Σε αυτά τα συστήματα ασφαλείας επιλέγονται χώροι όπου ο φωτισμός και ο προσανατολισμός των προσώπων είναι ελεγχόμενος, όπως μονές λωρίδες διάβασης των επιβατών. Το πρόσωπο καταγράφεται και γίνεται απόπειρα ταύτισης με μία βάση γνωστών τρομοκρατών και υπόπτων. Η ίδια φιλοσοφία ισχύει και για τα συστήματα παρακολούθησης με τη διαφορά ότι εκεί οι συνθήκες είναι λιγότερο ελεγχόμενες και μπορεί σε μία εικόνα να υπάρχει ένας μεγάλος αριθμός από πρόσωπα με αποτέλεσμα η αξιοπιστία να είναι ακόμα πιο μειωμένη. Ανασταλτικό παράγοντα αποτελεί επίσης ότι για μια τέτοια εφαρμογή, η επεξεργασία της εικόνας και η αναγνώριση θα πρέπει να γίνονται σε πραγματικό χρόνο, κάτι αρκετά δύσκολο ακόμα και σήμερα.

1.2.2. Έλεγχος πρόσβασης

Σε περιπτώσεις όπου θέλουμε να περιορίσουμε την πρόσβαση, όπως για παράδειγμα στο χώρο ενός γραφείου ή σε κάποιο υπολογιστή, η ομάδα των ανθρώπων η οποία πρέπει να αναγνωρίζεται είναι σχετικά μικρή. Το γεγονός αυτό, σε συνδυασμό και με το γεγονός ότι οι εικόνες συλλαμβάνονται σε πολύ ελεγχόμενο χώρο και συνθήκες, οδηγεί σε πολύ αξιόπιστα αποτελέσματα και χωρίς

κάποιον ιδιαίτερο χειρισμό από το χρήστη. Δε χρειάζεται να αγγίξει κάτι με το δάχτυλό του την παλάμη του ή να παρουσιάσει το μάτι του. Όταν βέβαια συνδυάζεται και με μία δεύτερη μέθοδο από τις παραπάνω, τότε η μέθοδος της αναγνώρισης προσώπου είναι ακόμα πιο αξιόπιστη. Η εφαρμογή αυτή έχει προσελκύσει πολύ το ενδιαφέρον με αποτέλεσμα να έχουν αναπτυχθεί πολλές αντίστοιχες εμπορικές εφαρμογές.



Εικόνα 5: Παράδειγμα εφαρμογής ενιαίου συστήματος ελέγχου πρόσβασης, παρακολούθησης επισκεπτών και τήρησης ωραρίου, βασισμένης στην τεχνολογία ανίχνευσης και αναγνώρισης προσώπων

1.2.3. Έξυπνες κάρτες

Οι έξυπνες κάρτες έχουν ενσωματωμένο ένα μικροεπεξεργαστή ή ένα τσιπ μνήμης το οποίο παρέχει την επεξεργαστική ισχύ για να εξυπηρετηθούν πολλές εφαρμογές. Οι κάρτες μνήμης απλά αποθηκεύουν δεδομένα. Οι έξυπνες κάρτες μπορούν να προσθέτουν, να διαγράφουν και να χειρίζονται γενικότερα τις πληροφορίες της μνήμης. Η έξυπνη κάρτα μπορεί επίσης να έχει ενσωματωμένα χαρακτηριστικά ασφαλείας καθώς και ενσωματωμένη κεραία ώστε ο αναγνώστης

των καρτών να μπορεί να διαβάζει την κάρτα από απόσταση. Όσον αφορά την αναγνώριση προσώπων, μπορεί στην κάρτα κατά την έκδοσή της να αποθηκευτούν τα βιομετρικά δεδομένα που αντιστοιχούν στο πρόσωπο του ιδιοκτήτη. Κατά την ταυτοποίηση συγκρίνονται τα δεδομένα που υπάρχουν αποθηκευμένα στην κάρτα με τα πραγματικά χαρακτηριστικά του ατόμου. Σε περίπτωση ταύτισης μπορούν τα δεδομένα να ενημερωθούν και να παραμένουν με αυτό τον τρόπο πάντα πρόσφατα και ενημερωμένα γεγονός που διασφαλίζει και μεγαλύτερη αξιοπιστία.

1.2.4. Συστήματα επιβολής του νόμου.

Με ένα σύστημα αναγνώρισης και ανάκτησης ενός προσώπου, οι αρχές μπορούν εύκολα να ξεχωρίσουν κάποιο ύποπτο άτομο. Δίνεται η δυνατότητα στις υπηρεσίες επιβολής του νόμου να αναζητήσουν και να αναγνωρίσουν υπόπτους γρήγορα, ακόμα και με ελλιπείς πληροφορίες, ακόμα και από ένα σκίτσο από τις αναμνήσεις ενός μάρτυρα. Παρομοίως, κάποιος ύποπτος μπορεί γρήγορα να ταυτοποιηθεί από video από συστήματα παρακολούθησης σε τράπεζες ή καταστήματα. Επίσης, η αναζήτηση εικόνων προσώπων με βάση κάποια χαρακτηριστικά μπορεί εύκολα να μειώσει τον αριθμό των εικόνων που μπορούν να χρησιμοποιηθούν για αναγνώριση από μάρτυρες. Ανασταλτικός παράγοντας για αυτή την εφαρμογή αποτελεί το μεγάλο μέγεθος μίας αντίστοιχης βάσης προσώπων και η δυσκολία αναζήτησης εικόνων σε αυτή.

1.2.5. Αλληλεπίδραση ανθρώπου - υπολογιστή

Για να επιτευχθεί πιο αποδοτική και πιο φιλική προς τον χρήστη αλληλεπίδραση με τον υπολογιστή, κάποια μέρη του ανθρωπίνου σώματος μπορούν να θεωρηθούν από τον υπολογιστή ως μονάδες εισόδου. Αυτό έχει κινητοποιήσει την έρευνα στην ανίχνευση την ανάλυση και την αναγνώριση της κίνησης και της στάσης του ανθρωπίνου σώματος, κυρίως των χεριών και του προσώπου. Μία σχετική εφαρμογή είναι και η αναγνώριση ανθρωπίνων συναισθημάτων ή ψυχολογικών καταστάσεων, όπως η χαρά, ο θυμός, η αηδία και

η έκπληξη, εντοπίζοντας και αναγνωρίζοντας τις αντίστοιχες εκφράσεις του προσώπου.

1.2.6. Ευρετηριασμός βάσεων δεδομένων πολυμέσων

Εικόνες οι οποίες περιέχουν ανθρώπινα πρόσωπα απαντώνται πολύ συχνά σε υλικό πολυμέσων όπως είναι οι ειδήσεις, τα αθλητικά, οι ταινίες, προσωπικές βιντεοσκοπήσεις και άλλα. Ο ευρετηριασμός αυτού του υλικού με ανίχνευση, παρακολούθηση και αναγνώριση των προσώπων είναι πολύ σημαντικός για την εύκολη περιήγηση, κατάτμηση, σύνοψη και την εύκολη επιλογή και αναζήτηση σε αυτό. Σε συνεργασία με συστήματα αναγνώριση φωνής, κατανόηση φυσικής γλώσσας, και αναγνώρισης εικόνας και κειμένου, η διαχείριση των προσώπων αποτελεί ένα πανίσχυρο εργαλείο για τον αυτόματο ευρετηριασμό, αναζήτηση, ανάκτηση και πρόσβαση στο υλικό πολυμέσων το οποίο αυξάνει συνεχώς σε όγκο.

Η δυσκολία της άμεσης εφαρμογής της διαδικασίας αναγνώρισης προσώπων σε εφαρμογές πολυμέσων είναι ότι συνήθως το σύνολο των ταυτοτήτων των προσώπων δεν είναι γνωστό. Έχουν αναπτυχθεί μέθοδοι, όπου από το ίδιο υλικό μπορεί να εξαχθεί και αυτή η πληροφορία της ταυτότητας του προσώπου που έχει αναγνωριστεί. Αυτό μπορεί να γίνει είτε χρησιμοποιώντας ηχητικά δεδομένα όπου μπορεί να αναφέρεται η ταυτότητα του ατόμου που απεικονίζεται είτε το όνομα εμφανίζεται παράλληλα στην εικόνα και μπορεί να εκμαιευτεί χρησιμοποιώντας τεχνικές οπτικής αναγνώρισης χαρακτήρων (OCR).

2. Επιλογές για τη λύση του προβλήματος

Έχουν αναπτυχθεί τα τελευταία χρόνια πολλές μέθοδοι για την επίλυση του προβλήματος της αναγνώρισης των προσώπων. Άλλες βασίζονται στην αντιμετώπιση του προσώπου σαν ένα ενιαίο σύνολο, μία εικόνα, η οποία πρέπει να ταυτοποιηθεί και ονομάζονται ολιστικές μέθοδοι, ενώ άλλες εστιάζονται σε συγκεκριμένα χαρακτηριστικά του προσώπου και ονομάζονται αναλυτικές μέθοδοι. Υπάρχει και μία τρίτη κατηγορία μεθόδων, οι υβριδικές, οι οποίες προσπαθούν να συνδυάσουν δεδομένα από χαρακτηριστικά του προσώπου, αλλά και από τη συνολική εικόνα του. Η κάθε μέθοδος εμφανίζεται να έχει τα προτερήματα αλλά και τα μειονεκτήματά της απέναντι σε διαφορετικές απόψεις του προβλήματος.

2.1. Ολιστικές Μέθοδοι

Στις ολιστικές μεθόδους κατατάσσονται οι περισσότεροι αλγόριθμοι που βασίζονται στην ανάλυση κυρίων συνιστωσών (PCA). Σαν δεδομένα εισόδου σ' αυτές τις μεθόδους έχουμε τη συνολική εικόνα του προσώπου το οποίο θέλουμε να αναγνωρίσουμε. Οι μέθοδοι αυτές βασίζονται στο επιχείρημα ότι το πρόβλημα της αναγνώρισης προσώπων μπορεί να αναχθεί σε μία χαμηλού επιπέδου διαδικασία επεξεργασίας εικόνας δύο διαστάσεων. Γεγονός είναι ότι από τα πρώιμα στάδια της ανάπτυξής του ο άνθρωπος έχει την ικανότητα να αναγνωρίζει πρόσωπα με καταπληκτική ταχύτητα. Επίσης σε πειράματα που έχουν γίνει σε πιθήκους υποστηρίζεται ότι έχουν απομονωθεί νευρώνες οι οποίοι ανταποκρίνονται επιλεκτικά σε πρόσωπα.

2.1.1. Μέθοδος Eigenface

Οι Turk και Pentland (Turk and Pentland 1991) παρουσιάζουν μία μέθοδο αναγνώρισης προσώπων, όπου οι εικόνες των προσώπων προβάλλονται στις κυρίες συνιστώσες των εικόνων που χρησιμοποιήθηκαν για την εκπαίδευση. Τα ιδιοδιανύσματα (eigenvectors) που προκύπτουν από αυτή τη μέθοδο ονομάζονται eigenfaces. Τα eigenfaces που προκύπτουν συγκρίνονται με γνωστά άτομα και

ταξινομούνται ανάλογα με την ομοιότητά τους με αυτά.

Η μέθοδος αναγνώρισης που υλοποιούμε και αναπτύσσουμε στην εργασία αυτή είναι βασισμένη στη μέθοδο αυτή.

Συνοπτικά, φαίνεται ότι η μέθοδος eigenfaces είναι ένας γρήγορος, απλός και πρακτικός αλγόριθμος. Ωστόσο, περιοριστικό παράγοντα αποτελεί το γεγονός ότι για βέλτιστη απόδοση απαιτείται υψηλός βαθμός συσχέτισης μεταξύ της έντασης των εικονοστοιχείων των εικόνων της εκπαίδευσης και των εικόνων προς αναγνώριση. Ο περιορισμός αυτός μπορεί να παρακαμφθεί με εκτεταμένη χρήση προεπεξεργασίας για την κανονικοποίηση των εικόνων, οι οποίες προορίζονται για την εκπαίδευση και την αναγνώριση.

Παραλλαγές βασισμένες πάνω στη μέθοδο eigenfaces: (Turk 2005; Pissarenko 2002; Leonardis and Bischof 2000; Tahir et al.)

2.1.2. Μέθοδος Fisherface

Η μέθοδος Fisherface προσπαθεί να βελτιώσει τη μέθοδο eigenface. Ο αλγόριθμος Fisherface μειώνει τον ιδιοχώρο (eigenspace), χρησιμοποιώντας τη γραμμική διακριτική ανάλυση Fisher (Fisher's Linear discriminant – FLD). Με τη μέθοδο αυτή ο ιδιοχώρος επιλέγεται έτσι ώστε ο λόγος της μεταξύ κλάσεων διασποράς προς τη διασπορά εντός της κλάσης να είναι μέγιστος.

Ας θεωρήσουμε το πρόβλημα αναγνώρισης όπου έχουμε ένα σύνολο c διακριτών κλάσεων. Ο μεταξύ των κλάσεων πίνακας διασποράς ορίζεται ως εξής:

$$S_b = \sum_{i=1}^c Pr(\Omega_i) (\mu_i - \mu) (\mu_i - \mu)^T \quad (1)$$

Όπου $Pr(\Omega_i)$ είναι η εκ των προτέρων πιθανότητα, μ_i είναι το μέσο δείγμα της κλάσης Ω_i και μ είναι το μέσο δείγμα όλων των κλάσεων.

Ο πίνακας διασποράς εντός της κλάσης ορίζεται ως εξής:

$$S_w = \sum_{i=1}^c (Pr(\Omega_i) \times \frac{1}{N_i} \sum_{y_k \in \Omega_i} (y_k - \mu_i) (y_k - \mu_i)^T) \quad (2)$$

Όπου N_i είναι ο αριθμός των δειγμάτων της κλάσης Ω_i . Αν το S_w δεν είναι μοναδικό, τότε ο πίνακας προβολής W_{fd} μπορεί να επιλεγεί ως:

$$W_{fd} = \arg \max_w \frac{|W^T S_b W|}{|W^T S_w W|} \quad (3)$$

Και στη συνέχεια, το W_{fd} μπορεί να υπολογιστεί λύνοντας το γενικευμένο πρόβλημα ιδιοτιμών:

$$S_b W = S_w W \lambda \quad (4)$$

Στην περίπτωση της εφαρμογής αναγνώρισης προσώπων, το S_w είναι πάντα μοναδικό. Επειδή η τάξη όμως του πίνακα $S_w \in \mathbb{R}^{n \times n}$, είναι το πολύ $N - c$, όπου N είναι ο αριθμός των εικόνων στα δεδομένα εκπαίδευσης και τυπικά το N είναι πολύ μικρότερο του n , που είναι ο αριθμός των pixels της κάθε εικόνας. Για να προσπεραστεί αυτή η δυσκολία χρησιμοποιείται συνήθως αρχικά η μέθοδος της ανάλυσης κυρίων συνιστωσών, ώστε να μειωθούν οι διαστάσεις των εικόνων από n σε $N - c$ ή και λιγότερο, και στη συνέχεια ξαναυπολογίζεται το S_w το οποίο πλέον θα είναι μοναδικό και η μέθοδος γραμμικής διακριτικής ανάλυσης του Fisher (Fisher's Linear Discriminant analysis ή FLD) μπορεί να χρησιμοποιηθεί για να υπολογίσει το W_{fd} που ονομάζεται και Fisherface.

Για τη μέθοδο ανάλυσης κυρίων συνιστωσών ο πίνακας προβολής W_{pca} επιλέγεται έτσι ώστε να μεγιστοποιείται η ορίζουσα του συνολικού πίνακα διασποράς των προβαλλόμενων δειγμάτων. Ο πίνακας διασποράς ορίζεται ως:

$$S = \sum_{k=1}^N (x_k - \mu)(x_k - \mu)^T \quad (5)$$

Όπου N είναι ο συνολικός αριθμός των δειγμάτων και μ είναι το μέσο δείγμα. Ο πίνακας W_{pca} υπολογίζεται λύνοντας το πρόβλημα ιδιοτιμών:

$$S W = W A \quad (6)$$

Από τον παραπάνω ορισμό του αλγορίθμου PCA, συμπεραίνουμε ότι ο πίνακας προβολής W_{pca} μεγιστοποιεί τις αποστάσεις μεταξύ όλων των δειγμάτων στο προβαλλόμενο διάστημα. Οι αποστάσεις όμως μεταξύ των δειγμάτων της εκπαίδευσης μπορεί να μην οφείλονται μόνο στις διαφορές μεταξύ των κλάσεων, αλλά και σε διαφορές μεταξύ των δειγμάτων της ίδιας κλάσης. Έτσι, σε ορισμένες περιπτώσεις, όπου τα δείγματα της ίδιας κλάσης μπορεί να διαφέρουν αρκετά μεταξύ τους, η μέθοδος PCA ίσως να μην αποδίδει τόσο καλά όσο αυτή η μέθοδος Fisherface. Χαρακτηριστικό παράδειγμα μιας τέτοιας περίπτωσης έχουμε όταν η φωτεινότητα των εικόνων του ίδιου προσώπου ποικίλει σε μεγάλο βαθμό (PIE face database).

Μέθοδοι βασισμένες σε παραλλαγές της μεθόδου Fisherface (Shan et al. 2002; Yina et al. 2005; Belhumeur et al. 1997; Gross et al. 2004)

2.1.3. Μέθοδος Μηχανών διανυσμάτων υποστήριξης.

Οι μηχανές διανυσμάτων υποστήριξης αποτελούν ένα σύνολο σχεσιακών μεθόδων εκπαίδευσης με επίβλεψη για σκοπούς ταξινόμησης. Έχοντας την ικανότητα να βλέπουν τα δεδομένα εισόδου σαν δύο σύνολα διανυσμάτων σε ένα n -διάστατο χώρο, οι μηχανές διανυσμάτων υποστήριξης μπορούν να κατασκευάσουν μία διαχωριστική υπερεπιφάνεια σ' αυτό το χώρο, η οποία μεγιστοποιεί το περιθώριο μεταξύ των δύο συνόλων των δεδομένων.

Οι μηχανές διανυσμάτων υποστήριξης όπως φαίνεται και από τον ορισμό τους αποτελούν μέθοδο δυαδικής ταξινόμησης. Για να μπορέσουν λοιπόν τα λύσουν προβλήματα q -κλάσεων, όπως είναι το πρόβλημα της αναγνώρισης προσώπων, υπάρχουν δύο πιθανές στρατηγικές. Η πρώτη προσέγγιση ενός προς όλα, απαιτεί την εκπαίδευση q μηχανών οι οποίες θα ξεχωρίζουν την κάθε μία κλάση από όλες τις υπόλοιπες. Στη δεύτερη προσέγγιση οι κλάσεις χωρίζονται ανά ζευγάρια και εκπαιδεύουμε $q(q-1)/2$ μηχανές οι οποίες ξεχωρίζουν τις δύο κλάσεις μεταξύ τους. Στη συνέχεια οι μηχανές μπορούν να οργανωθούν σε δέντρα όπου με

σταδιακό αποκλεισμό των κλάσεων καταλήγουμε σε αναγνώριση του προσώπου. (Heisele, Ho, and Poggio 2001; Zhu and Ji 2005)

2.1.4. Ανάλυση ανεξάρτητων συνιστωσών

Η ανάλυση ανεξάρτητων συνιστωσών είναι μία στατιστική και υπολογιστική τεχνική για την αποκάλυψη κρυμμένων παραγόντων που βρίσκονται πίσω από σύνολα τυχαίων μεταβλητών, μετρήσεων ή σημάτων. Η τεχνική αυτή ορίζει ένα γενικευμένο μοντέλο των παρατηρούμενων δεδομένων, τα οποία συνήθως δίνονται με τη μορφή μίας βάση δεδομένων από δείγματα. Στο μοντέλο αυτό τα δεδομένα θεωρούνται γραμμικά συνδυαστικά παράγωγα κάποιων αγνώστων μεταβλητών, καθώς και ο τρόπος συνδυασμού των μεταβλητών αυτών είναι άγνωστος. Οι μεταβλητές αυτές ονομάζονται ανεξάρτητες συνιστώσες των παρατηρούμενων δεδομένων και μπορούν να βρεθούν με τη μέθοδο αυτή.

Η ανάλυση ανεξαρτήτων συνιστωσών μοιάζει μόνο επιφανειακά με την ανάλυση κυρίων συνιστωσών. Αποτελεί μία πολύ δυνατή τεχνική με αρκετά προτερήματα που όμως η πολυπλοκότητά της σε σύγκριση με τη μέθοδο ανάλυσης κυρίων συνιστωσών και το μικρό σχετικά όφελος ως προς την αποτελεσματικότητα καθιστούν πρακτικά τη μέθοδο της ανάλυσης κυρίων συνιστωσών προτιμότερη. (Hynäriinen and E. Oja 2000)

2.1.5. Γραμμική διακριτική ανάλυση

Ο μετασχηματισμός Fourier και η γραμμική διακριτική ανάλυση είναι δύο συχνά χρησιμοποιούμενες τεχνικές επεξεργασίας και αναγνώρισης εικόνων. Η γραμμική διακριτική ανάλυση επιχειρεί να ανακαλύψει το γραμμικό συνδυασμό χαρακτηριστικών τα οποία διαχωρίζουν καλύτερα δύο ή περισσότερες κλάσεις προσώπων. Η λειτουργία του αλγορίθμου αντιστοιχεί όπως και της ανάλυσης κυρίων συνιστωσών σε μείωση των διαστάσεων του προβλήματος, με τη διαφορά ότι επιχειρεί ρητά να μοντελοποιήσει τη διαφορά μεταξύ των κατηγοριών των δεδομένων ενώ ο αλγόριθμος PCA δε λαμβάνει υπόψη καμία διαφορά στις κατηγορίες. (Fazl-Ersi, Zelek, and Tsotsos 2007; Xudong 2009)

2.1.6. Πιθανοθεωρητικά Νευρωνικά δίκτυα βασισμένα σε αποφάσεις

Στα πιθανοθεωρητικά νευρωνικά δίκτυα η διακριτική συνάρτηση των υποδικτύων βασίζεται όπως δείχνει και το όνομά τους σε πιθανοθεωρητικούς περιορισμούς. Τα δίκτυα είναι έτσι σχεδιασμένα ώστε να μοντελοποιούν τη συνάρτηση φυσικής λογαριθμικής ομοιότητας. Η ιδιαιτερότητα των δικτύων αυτών είναι ότι οι κανόνες εκπαίδευσής τους βασίζονται σε αποφάσεις και αποτελούν υβρίδιο όπου η εκπαίδευση τοπικά είναι χωρίς εποπτεία (unsupervised), ενώ συνολικά γίνεται με εποπτεία (supervised). Αυτό γίνεται με τον εξής τρόπο: Για κάθε πρότυπο εκπαίδευσης, το δίκτυο αφήνεται να ταξινομήσει το πρότυπο και στη συνέχεια ο εκπαιδευτής αποφασίζει αν η ταξινόμηση που έγινε ήταν σωστή ή λάθος. Στη συνέχεια τα βάρη του δικτύου προσαρμόζονται ανάλογα με την απόφαση αυτή που πάρθηκε από τον εκπαιδευτή. Η καλύτερη πρακτική για τη χρήση των δικτύων αυτών στην αναγνώριση προσώπων είναι η τεχνική διάκρισης μίας κλάσης από όλες τις υπόλοιπες. Μ' αυτό τον τρόπο έχουμε ένα δίκτυο για κάθε πρόσωπο τα οποία συνδέονται με ένα δίκτυο maxnet, το οποίο αναδεικνύει τη νικήτρια κλάση με βάση τις εξόδους των υποδικτύων. Η μέθοδος αυτή πλεονεκτεί γιατί απαιτείται λιγότερος χρόνος εκπαίδευσης των δικτύων και επίσης κατά την προσθήκη ενός νέου προσώπου αρκεί να εκπαιδευτεί μόνο ένα νέο δίκτυο και απλά να προσαρμοστεί ελάχιστα η εκπαίδευση των υπολοίπων.

Τα πιθανοθεωρητικά νευρωνικά δίκτυα μπορούν να χρησιμοποιηθούν με επιτυχία και στο στάδιο της ανίχνευσης προσώπων.

2.2. Μέθοδοι βασισμένες σε χαρακτηριστικά ή Αναλυτικές μέθοδοι

Οι μέθοδοι αυτές βασίζονται μόνο σε ορισμένα τμήματα του προσώπου και όχι σε ολόκληρη την εικόνα του προσώπου όπως γίνεται με τις προηγούμενες μεθόδους. Το επιχείρημα για την τακτική αυτή είναι το γεγονός ότι στις περισσότερες εικόνες προσώπων κάποια χαρακτηριστικά όπως τα μαλλιά ή το περίγραμμα του προσώπου μπορεί να διαφέρουν αρκετά από εικόνα σε εικόνα για το ίδιο πρόσωπο, ενώ κάποια χαρακτηριστικά όπως τα μάτια, η μύτη ή το στόμα και

κυρίως η κεντρική περιοχή του προσώπου είναι αυτή που παραμένει σχεδόν σταθερή αλλά και παρέχει το μεγαλύτερο μέρος της πληροφορίας για την αναγνώριση. Η επικέντρωση σ' αυτό το σημείο του προσώπου επιτυγχάνεται συνήθως με τη χρήση κάποιων φίλτρων τα οποία αυτοματοποιούν τη διαδικασία ανίχνευσης των περιοχών που μας ενδιαφέρουν.

2.2.1. Καθαρά γεωμετρικές μέθοδοι 2Δ ή 3Δ

Οι καθαρά γεωμετρικές μέθοδοι βασίζονται σε γεωμετρικά δεδομένα των χαρακτηριστικών του προσώπου, κυρίως τις αποστάσεις μεταξύ κάποιων σημείων αναφοράς. Οι μέθοδοι αυτές ενδείκνυνται να χρησιμοποιηθούν σε περιπτώσεις όπου ο αριθμός των δεδομένων εισόδου είναι πολύ μικρός, π.χ. στην περίπτωση που έχουμε μόνο μία εικόνα για κάθε πρόσωπο που θέλουμε να αναγνωρίσουμε. Για την ακρίβεια των δεδομένων, αποκλείεται η περίπτωση όπου τα χαρακτηριστικά αναγνωρίζονται αυτόματα, οπότε απαιτείται παρέμβαση για τις μετρήσεις ακριβείας που απαιτούνται ώστε να εξαχθούν τελικά τα δεδομένα εισόδου από τις εικόνες. Για την αναγνώριση ενός προσώπου, τα αντίστοιχα γεωμετρικά δεδομένα εξάγονται και συγκρίνονται με αυτά της βάσης μας χρησιμοποιώντας ειδικές στοχαστικές συναρτήσεις και αρκετά αποδοτικούς υπολογισμούς.

Με την εξέλιξη της τεχνολογίας και του οπτικού εξοπλισμού, η μέθοδος αυτή επεκτάθηκε και στην τρισδιάστατη έκδοσή της, όπου πλέον τα δεδομένα δεν αναφέρονται σε αποστάσεις όπως αυτές μετρώνται σε μία δισδιάστατη προβολή του προσώπου, αλλά σε πραγματικά τρισδιάστατα δεδομένα απεικόνισης του προσώπου στο χώρο. Μία υβριδική μέθοδος που συνδυάζει και τη μέθοδο ανάλυσης κυρίων συνιστωσών χρησιμοποιεί σαν δεδομένα εισόδου την τρισδιάστατη απεικόνιση του προσώπου όπως αυτή λαμβάνεται από ένα Laser scanner. Τα ιδιοδιανύσματα που προκύπτουν από την ανάλυση αυτή ονομάζονται *eigenheads* και ο γραμμικός συνδυασμός τους χρησιμοποιείται για την αναγνώριση.

Υπάρχει πληθώρα βιβλιογραφίας βασισμένη σε αυτή τη μέθοδο. Παραθέτουμε κάποιες χαρακτηριστικές μεθόδους αυτής της κατηγορίας: (Kakadiaris et al. 2006; Mian, Bennamoun, and Owens 2007; Smith and Kanade 2004; Akarun, Glöckberk, and Salah 2005; Huang, Heisele, and Blanz 2003)

2.2.2. Αντιστοίχιση δυναμικών συνδέσμων (Dynamic Link Matching)

Στην περίπτωση της αναγνώρισης προσώπων με τη μέθοδο της αντιστοίχισης δυναμικών συνδέσμων, έχουμε ένα νευρωνικό δίκτυο πολλαπλών στρωμάτων, όπου ένα στρώμα είναι το στρώμα του μοντέλου, και ένα άλλο το στρώμα της εικόνας. Κάθε νευρώνας του δικτύου χρησιμοποιεί ανάλυση κυματομορφών (wavelet analysis) για να εντοπίσει συγκεκριμένα χαρακτηριστικά του μοντέλου στην εικόνα. Το ενδιαφέρον χαρακτηριστικό της μεθόδου των δυναμικών συνδέσμων είναι ότι είναι πολύ ανεκτική σε μετακινήσεις και παραμορφώσεις που οφείλονται στην περιστροφή του προσώπου. Παρόλα αυτά η μέθοδος αυτή έχει και πολλά μειονεκτήματα, όπως η πολύ χαμηλή ταχύτητα αναγνώρισης που την καθιστούν πρακτικά μη εφαρμόσιμη.

(von der Malsburg, Wiskott, and Fellous 1997; Fazl-Ersi, Zelek, and Tsotsos 2007)

2.2.3. Συνελικτικά νευρωνικά δίκτυα

Τα συνελικτικά δίκτυα είναι μία ειδική μορφή νευρωνικών δικτύων πολλαπλών στρωμάτων. Όπως τα περισσότερα δίκτυα πολλαπλών στρωμάτων, έτσι κι αυτά εκπαιδεύονται με τον αλγόριθμο back propagation ή κάποια παραλλαγή του. Για την αναγνώριση προσώπων χρησιμοποιούνται σε εναλλαγή με ενδιάμεσα επιμεριστικά στρώματα τα οποία σταδιακά μειώνουν την πολυπλοκότητα και τις διαστάσεις των δεδομένων εισόδου διατηρώντας ταυτόχρονα την ιδιομορφία τους. (Lawrence et al. 1997) Τα συνελικτικά νευρωνικά δίκτυα είναι σχεδιασμένα να μπορούν να αναγνωρίζουν οπτικά πρότυπα απευθείας από εικονοστοιχεία εικόνων με ελάχιστη προεπεξεργασία και είναι αρκετά ανεκτικά σε παραμορφώσεις, μετακινήσεις και θόρυβο στα δεδομένα εισόδου. Το μοντέλο των συνελικτικών δικτύων βασίζεται στις ίδιες αρχές που διέπουν το οπτικό και το ακουστικό νεύρο

στα θηλαστικά.

2.3. Υβριδικές μέθοδοι

Και οι δύο προηγούμενες κατηγορίες αλγορίθμων έχουν μειονεκτήματα και προτερήματα. Σε ορισμένες περιπτώσεις οι ολιστικές μέθοδοι υπερτερούν, ενώ σε κάποιες άλλες περιπτώσεις υπερτερούν οι αναλυτικές μέθοδοι.

Παράγοντας Μεταβολής	Τοπικά Χαρακτηριστικά		Ολιστικές Μέθοδοι	
Μικρές μεταβολές		Καθόλου ευαισθησία		Μέτρια ευαισθησία
Μεγάλες μεταβολές		Μέτρια ευαισθησία		Μεγάλη ευαισθησία
Φωτισμός		Μεγάλη ευαισθησία		Μέτρια ευαισθησία
Έκφραση προσώπου		Καθόλου ευαισθησία		Μέτρια ευαισθησία
Πόζα		Μέτρια ευαισθησία		Μεγάλη ευαισθησία
Θόρυβος		Μεγάλη ευαισθησία		Μέτρια ευαισθησία
Τμηματική απόκρυψη		Καθόλου ευαισθησία		Μεγάλη ευαισθησία

Πίνακας 1: Σύγκριση της ευαισθησίας των ολιστικών μεθόδων και των μεθόδων βασισμένων σε χαρακτηριστικά σε μεταβολές στις εικόνες.

Στον Πίνακα 1 επιχειρείται μία γενικότερη σύγκριση των ολιστικών μεθόδων και των μεθόδων βασισμένων σε τοπικά χαρακτηριστικά ως προς την ευαισθησία τους σε μεταβολές των εικόνων και πως αυτές επηρεάζουν την αποτελεσματικότητά τους. Τα δεδομένα είναι βασισμένα σε πειραματικά αποτελέσματα (Martinez 2002; Tan et al. 2005; Heisele et al. 2001; Costen, Cootes, and Taylor 2002; Peter W. Hallinan 1999)

Οι υβριδικές μέθοδοι χρησιμοποιούν ταυτόχρονα κάποια τοπικά χαρακτηριστικά του προσώπου, αλλά και όλη την περιοχή του προσώπου για τη διαδικασία της αναγνώρισης, προσπαθώντας να συνδυάσουν τα πλεονεκτήματα και των δύο μεθόδων και να εξαλείψουν τα μειονεκτήματα της κάθε μεθόδου μεμονωμένης.

2.3.1. Modular eigenfaces, eigenmodules.

Χρησιμοποιώντας τη μέθοδο eigenface για την αναγνώριση επιμέρους χαρακτηριστικών του προσώπου, προκύπτουν ιδιοδιανύσματα τα οποία

ονομάζονται eigenmodules (eigeneyes, eigennooses, eigenmouths κτλ ανάλογα με το χαρακτηριστικό του προσώπου το οποίο έχουν εκπαιδευτεί να αναγνωρίζουν). Έχει υπολογιστεί ότι το πλήθος των χαρακτηριστικών που απαιτούνται για την αναγνώριση αυξάνει λογαριθμικά με το πλήθος των προσώπων που πρέπει να αναγνωριστούν. Για το λόγο αυτό, σε συστήματα όπου ο αριθμός των προσώπων είναι μεγάλος, η χρήση των eigenmodules γίνεται σε συνδυασμό με την ολιστική μέθοδο eigenface, ώστε να μειωθεί ο αριθμός των δεδομένων εκπαίδευσης που απαιτούνται για την εξαγωγή των eigenmodules με τη μέθοδο της ανάλυσης κυρίων συνιστωσών από τις εικόνες των προσώπων που έχουμε για την εκπαίδευση. (Leonardis and Bischof 2000; Pavlou 2005; Gross et al. 2004) Με τη μέθοδο αυτή μπορούν να προσπεραστούν κάποια από τα μειονεκτήματά της μεθόδου eigenface. Ένα σύστημα που χρησιμοποιεί αποκλειστικά τη μέθοδο eigenface μπορεί εύκολα να ξεγελαστεί από την αλλαγή σε κάποιο χαρακτηριστικό του προσώπου, όπως γυαλιά, καπέλο, μούσι, ενώ σε συνδυασμό με τα υπόλοιπα επιμέρους χαρακτηριστικά ενός τέτοιου υβριδικού συστήματος μπορούμε να έχουμε σωστή ταυτοποίηση.

2.3.2. Υβριδική ανάλυση τοπικών χαρακτηριστικών (Hybrid LFA)

Οι Penev και Atick πρότειναν το 1996 ένα υβριδικό μοντέλο για ένα σύστημα αναγνώρισης προσώπων το οποίο συνδυάζει τη μέθοδο της ανάλυσης κυρίων συνιστωσών με τη μέθοδο ανάλυσης τοπικών χαρακτηριστικών (Penev and Atick 1996). Η μέθοδος ανάλυσης τοπικών χαρακτηριστικών προσπαθεί να παρακάμψει το μειονέκτημα της μεθόδου PCA να αγνοεί κάποια τοπικά χαρακτηριστικά του προσώπου και να λαμβάνει υπ' όψιν μόνο τη συνολική εικόνα του προσώπου. Η μέθοδος της ανάλυσης τοπικών χαρακτηριστικών εστιάζει σε συγκεκριμένα χαρακτηριστικά του προσώπου, όπως είναι τα μάτια, τα φρύδια, η μύτη, το στόμα κτλ. Οι νευρώνες του δικτύου που χρησιμοποιείται έχουν πεδία εισόδου πρόσθιας τροφοδότησης, τα οποία συνδυάζονται ιδανικά με τις κατηγορίες των αντικειμένων και αποσυσχετίζουν τις εισόδους όσο το δυνατόν περισσότερο. Οι υπολειπόμενες συσχετίσεις καταστέλλονται με πλευρική αναστολή. Αυτό οδηγεί σε εξόδους οι

οποίες είναι αραιά κατανεμημένες, δηλαδή για κάθε συγκεκριμένη είσοδο, μόνο ένα τμήμα της εξόδου είναι ενεργό κάθε φορά. Η μέθοδος ανάλυσης τοπικών χαρακτηριστικών επιτυγχάνει πολύ καλή ανακατασκευή του προσώπου από τα χαρακτηριστικά που έχουν εξαχθεί κατά την εκπαίδευση. Ο συνδυασμός των δύο μεθόδων με ισορροπημένο τρόπο καταφέρνει να αντιλαμβάνεται κάποια γενικότερα χαρακτηριστικά όπως το μέγεθος, το σχήμα του προσώπου και οι συνθήκες φωτισμού, τα οποία γίνονται καλύτερα αντιληπτά με την ολιστική μέθοδο PCA καθώς και τα υπόλοιπα χαρακτηριστικά του προσώπου που γίνονται καλύτερα αντιληπτά με την τοπική μέθοδο LFA. (Ng and Cheung 2004)

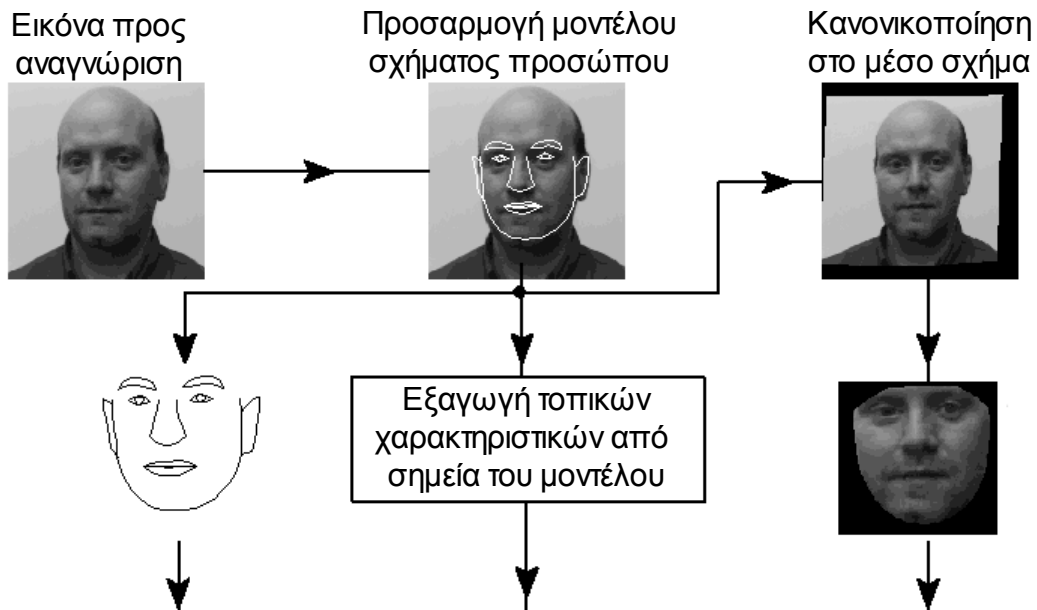
2.3.3. Αναγνώριση προσώπων με ευέλικτα εμφανισιακά μοντέλα

Το 1997 ο Ανδρέας Λανίτης (Lanitis, Taylor, and Cootes 1997) πρότεινε ένα σύστημα αναγνώρισης προσώπων το οποίο βασίζεται σε ευέλικτα εμφανισιακά μοντέλα. Το σύστημα υπολογίζει το μοντέλο του περιγράμματος του προσώπου, εντοπίζοντας έναν αριθμό από σημεία (152) πάνω στην εικόνα. Στη συνέχεια, η περιοχή του προσώπου διαχωρίζεται από το υπόβαθρο και χρησιμοποιώντας ένα μικρότερο αριθμό από σημεία – κλειδιά του σχήματος του προσώπου, το τμήμα της εικόνας του προσώπου κανονικοποιείται ώστε να μην περιέχει πλέον καμία πληροφορία για το σχήμα του προσώπου. Η καθαρή πληροφορία που απομένει είναι οι διαβαθμίσεις του γκρι στην περιοχή του προσώπου.

Στη συνέχεια, για την αναγνώριση του προσώπου συνδυάζονται τριών ειδών πληροφορίες:

- Το κανονικοποιημένο τμήμα της εικόνας του προσώπου, από το οποίο μέσω της κανονικοποίησης έχει αφαιρεθεί η πληροφορία του σχήματος του προσώπου.
- Από το μοντέλο του σχήματος του προσώπου μπορούμε να εξαγάγουμε ένα σύνολο τοπικών χαρακτηριστικών όπως είναι τα μάτια, το στόμα και τα αυτιά.
- Τέλος το ίδιο το μοντέλο του σχήματος του προσώπου παρέχει πληροφορίες σχετικά με το σχήμα και τη μορφή του προσώπου.

Τελικά τα παραπάνω δεδομένα αναλύονται με τη μέθοδο της διακριτής ανάλυσης και είτε συνδυάζονται ή χρησιμοποιούνται μεμονωμένα για την εξαγωγή χρήσιμων πληροφοριών που αφορούν την ταυτότητα του προσώπου ή κάποιο χαρακτηριστικό του (φύλο, έκφραση προσώπου).



Εικόνα 6: Διάγραμμα λειτουργίας του συστήματος αναγνώρισης προσώπων με τη μέθοδο των ευέλικτων μοντέλων

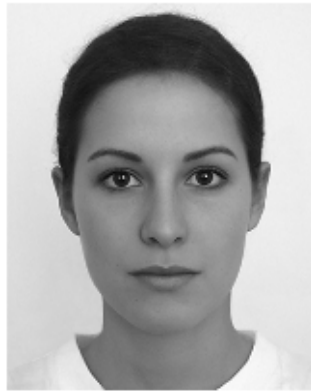
Η μέθοδος φαίνεται να έχει αρκετά ικανοποιητικά αποτελέσματα και είναι αρκετά ανεκτική σε μεταβολές του φωτισμού και της πόζας του προσώπου.

3. Η ανάλυση κυρίων συνιστωσών ως μέθοδος αναγνώρισης προσώπων

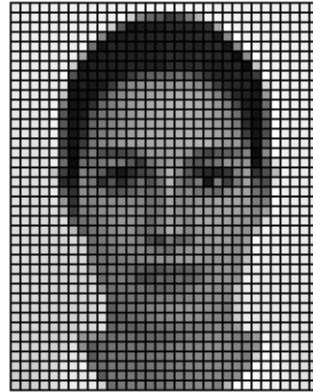
3.1. Μέθοδος ανάλυσης κυρίων συνιστωσών.

Η μέθοδος ανάλυσης κυρίων συνιστωσών (PCA) είναι μία στατιστική τεχνική που ανήκει σε μία γενικότερη κατηγορία με το όνομα ανάλυση παραγόντων (factor analysis). Ο σκοπός της μεθόδου PCA είναι ο προσδιορισμός της δομής της εξάρτησης πίσω από μια στοχαστική παρατήρηση πολλαπλών μεταβλητών. Όταν ο συσχετισμός μεταξύ των παρατηρούμενων μεταβλητών δεν είναι μηδενικός, τότε, ο αριθμός των διαστάσεων n , του χώρου των δεδομένων μας (ο αριθμός των παρατηρούμενων μεταβλητών) δεν αντιπροσωπεύει τον αριθμό των ανεξάρτητων μεταβλητών m , οι οποίες χρειάζονται πραγματικά για να περιγράψουν τα δεδομένα. Όσο πιο ισχυρή είναι η συσχέτιση μεταξύ των παρατηρούμενων μεταβλητών, τόσο μικρότερος είναι ο αριθμός των ανεξάρτητων μεταβλητών που μπορεί να τις περιγράψει επαρκώς.

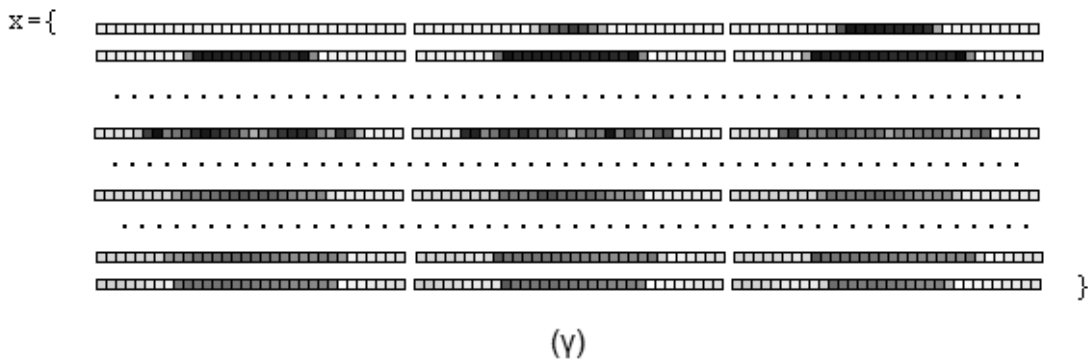
Το πρόσωπο, το οποίο είναι μία εικόνα, μπορεί να εκφραστεί και σαν διάνυσμα. Αν η εικόνα έχει πλάτος w και ύψος h , τότε ο αριθμός των συνιστωσών αυτού του διανύσματος θα είναι $w * h$. Κάθε pixel της εικόνας κωδικοποιείται σε μία συνιστώσα του διανύσματος. Η κατασκευή του διανύσματος αυτού από την εικόνα είναι πολύ απλή. Αρκεί να παραθέσουμε κάθε γραμμή από εικονοστοιχεία της εικόνας τη μία δίπλα στην άλλη, όπως φαίνεται στο σχήμα (2).



(α)



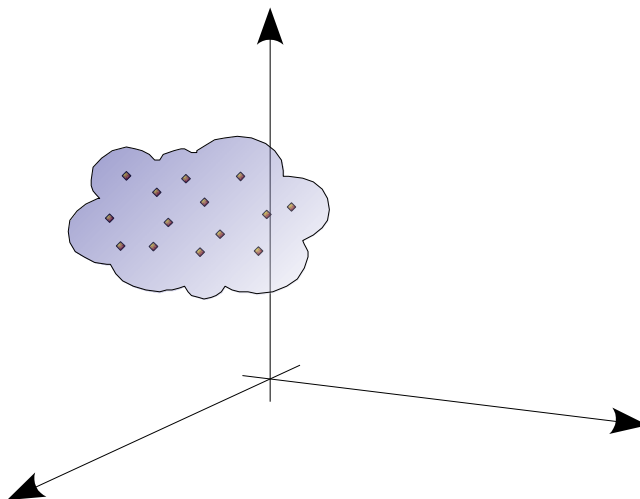
(β)



Σχήμα 2: Αναπαράσταση της εικόνας ενός προσώπου με τη μορφή διανύσματος: (α) Η αρχική εικόνα του προσώπου, (β) Η ψηφιοποιημένη μορφή της εικόνας. (γ) Το τελικό διάνυσμα x όπου κάθε γραμμή εικονοστοιχείων παρατίθεται σε σειρά. Χρησιμοποιείται σχετικά χαμηλή ανάλυση ώστε να είναι εμφανής η αναδιάταξη των εικονοστοιχείων.

Επειδή όλα τα πρόσωπα μοιάζουν και έχουν κάποια κοινά στοιχεία (δύο μάτια, μία μύτη, ένα στόμα), από το συνολικό χώρο που αντιπροσωπεύει το σύνολο των πιθανών διανυσμάτων τα οποία προκύπτουν από μία εικόνα αυτής της μορφής (διαστάσεων $w \times h$), μόνο ένα μικρό τμήμα, ένας περιορισμένος υποχώρος, περιλαμβάνει όλες τις εικόνες οι οποίες απεικονίζουν πρόσωπα (Σχ. 3).

Ακόμα, ούτε όλα τα pixels της εικόνας αντιστοιχούν στο πρόσωπο, και κάθε pixel εξαρτάται και από τα γειτονικά του. Αυτό καθιστά πολύ δύσκολο να καθορίσουμε τις διαστάσεις του υποχώρου των προσώπων, αλλά το σίγουρο είναι ότι είναι πολύ μικρότερο από το συνολικό χώρο.



Σχήμα 3: Το συνολικό διάστημα των πιθανών εικόνων και το υποδιάστημα του χώρου των εικόνων που απεικονίζουν πρόσωπα.

Ο σκοπός της μεθόδου της ανάλυσης κυρίων συνιστωσών είναι να μειώσει τις διαστάσεις του συνολικού χώρου των δεδομένων έτσι ώστε η νέα βάση να περιγράφει καλύτερα τα πρότυπα των δεδομένων εκπαίδευσης. Τα πρότυπα αναπαρίστανται στο νέο σύστημα συντεταγμένων σαν γραμμικός συνδυασμός των κυρίων συνιστωσών. Οι κύριες συνιστώσες που εξάγονται από τη μέθοδο αυτή είναι ανεξάρτητες μεταξύ τους, οπότε μεγιστοποιούνται οι διαφορές των αρχικών προτύπων που μας ενδιαφέρουν.

3.1.1. Ιστορικό και θεωρία της ανάλυσης κυρίων συνιστωσών.

Η ανάλυση κυρίων συνιστωσών είναι μία κλασική στατιστική μέθοδος η οποία είναι ουσιαστικά ισοδύναμη με τον μετασχηματισμό Karhunen–Loève που χρησιμοποιείται στη μονοδιάστατη και δισδιάστατη επεξεργασία σήματος. Η φιλοσοφία πίσω από τη μέθοδο PCA είναι αρκετά παλιά. Ξεκίνησε από τον Pearson το 1901 (Pearson 1901) σαν μία μεθοδολογία για την προσαρμογή επιπέδων με την έννοια των ελαχίστων τετραγώνων (η οποία ονομάζεται γραμμική παλινδρόμηση (linear regression)). Στη συνέχεια ο Hotelling το 1933 (Harold

Hotelling 1933) πρότεινε την τεχνική αυτή για την ανάλυση της δομής συσχέτισης μεταξύ πολλών τυχαίων μεταβλητών.

Οι n λοιπόν παρατηρούμενες μεταβλητές μπορούν να αναπαρασταθούν σαν συναρτήσεις των m λανθανουσών μεταβλητών οι οποίες ονομάζονται παράγοντες (factors), όπου $m < n$ και πολύ συχνά $m \ll n$. Όσο πιο απλή είναι η μαθηματική μορφή των συναρτήσεων, τόσο πιο οικονομική είναι η περιγραφή της δομής εξάρτησης των μεταβλητών. Τυπικά η μέθοδος PCA συσχετίζεται με γραμμικούς μετασχηματισμούς, οι οποίοι είναι η πιο απλή μαθηματική μορφή αναπαράστασης. Οι μεταβλητές αυτές καλούνται και παράγοντες (factors) ή χαρακτηριστικά (features) του πολυμεταβλητού τυχαίου σήματος, και το διάνυσμα το οποίο συγκροτούν είναι μέλος του χώρου συντελεστών.

Ας αναλογιστούμε ένα τυχαίο διάνυσμα $\mathbf{x} = [x_1, \dots, x_n]^T$ με μέσο $E\{\mathbf{x}\} = 0$ και πίνακα συνδιασποράς $R_x = E\{\mathbf{x}\mathbf{x}^T\} \in R^{n \times n}$. Στη μέθοδο PCA το διάνυσμα συντελεστής \mathbf{y} είναι ένας ορθογώνιος, γραμμικός μετασχηματισμός των δεδομένων:

$$\mathbf{y} = \mathbf{W}\mathbf{x} \quad (7)$$

όπου οι στήλες του \mathbf{W} σχηματίζουν μία ορθοκανονική βάση ενός υποχώρου L , έτσι ώστε $\mathbf{W}\mathbf{W}^T = \mathbf{I}$ και $L = \text{έκταση}(\mathbf{W})$. Η προβολή του \mathbf{x} στο L είναι η αναπαράσταση του \mathbf{x} από το \mathbf{y} :

$$\hat{\mathbf{x}} = \mathbf{W}^T \mathbf{y} = \mathbf{W}^T \mathbf{W}\mathbf{x} \quad (8)$$

Η μέθοδος PCA επιδιώκει την ελαχιστοποίηση του μέσου τετραγώνου του σφάλματος ανασχηματισμού.

$$J_e = E\{\|\mathbf{x} - \hat{\mathbf{x}}\|^2\} \quad (9)$$

$$= E\{\text{tr}[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T]\} \quad (10)$$

$$J_e = E \{ \|\mathbf{x} - \hat{\mathbf{x}}\|^2 \} \quad (9)$$

$$= \text{tr}(\mathbf{R}_x) - \text{tr}(\mathbf{W}\mathbf{R}_x\mathbf{W}^T) \quad (11)$$

από το οποίο και σύμφωνα με το γεγονός ότι $\text{tr}(\mathbf{A}) = \text{tr}(\mathbf{A}^T)$ και

$$\text{tr}(E\{\mathbf{W}^T \mathbf{W} \mathbf{x} \mathbf{x}^T \mathbf{W}^T \mathbf{W}\}) = \text{tr}(\mathbf{W}\mathbf{W}^T \mathbf{W}\mathbf{R}_x\mathbf{W}^T) \quad (12)$$

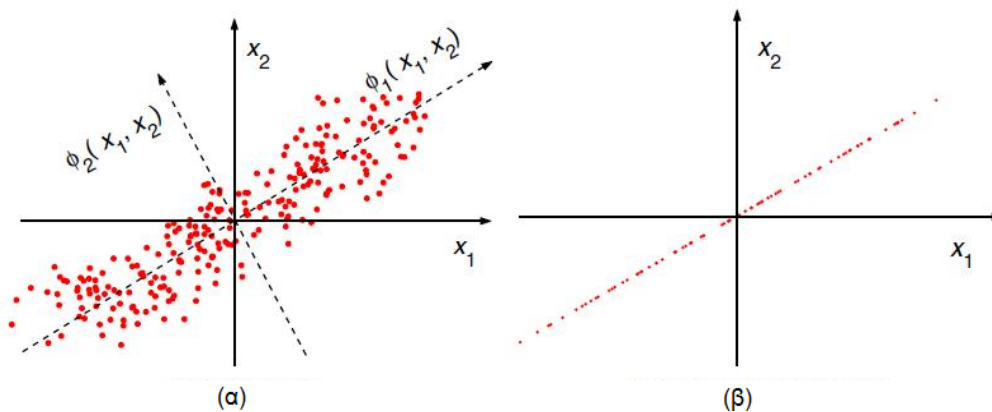
$$= \text{tr}(\mathbf{W}\mathbf{R}_x\mathbf{W}^T) \quad (13)$$

Παρατηρούμε ότι ο τελευταίος όρος (ας τον ονομάσουμε J_v) ισούται με τη διακύμανση του y , που επίσης ισούται με τη διακύμανση της προβολής $\hat{\mathbf{x}}$:

$$J_v = \text{tr}(\mathbf{W}\mathbf{R}_x\mathbf{W}^T) = E\{\text{tr}(\mathbf{y}\mathbf{y}^T)\} = \sum_{i=1}^m y_i^2 \quad (14)$$

$$= \text{tr}(\mathbf{W}^T \mathbf{W}\mathbf{R}_x\mathbf{W}^T \mathbf{W}) = E\{\text{tr}(\hat{\mathbf{x}}\hat{\mathbf{x}}^T)\} = \sum_{i=1}^n \hat{x}_i^2 \quad (15)$$

Κατά συνέπεια, η μεγιστοποίηση της διακύμανσης της προβολής είναι ισοδύναμη με την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος του ανασχηματισμού. Η μέθοδος PCA μπορεί ισοδύναμα να θεωρηθεί είτε σαν τεχνική μεγιστοποίησης της διακύμανσης, είτε σαν ένας αλγόριθμος ελαχίστων μέσων τετραγώνων.



Σχήμα 4: Μέθοδος PCA.

α) Με συμπαγείς γραμμές εμφανίζεται η αρχική βάση του συστήματος συντεταγμένων. Με διακεκομμένες φαίνεται η βάση του μετασχηματισμού Karhunen-Loève.

Τα σημεία έχουν επιλεγεί σε κανονικά διαστήματα σε ευθεία γραμμή περιστραμμένη κατά 30ο και στη συνέχεια έχουμε εφαρμόσει τυχαία μετατόπιση στις 2 διαστάσεις.

β) Η προβολή (ανακατασκευή σε μία διάσταση) των δεδομένων χρησιμοποιώντας μόνο την πρώτη κύρια συνιστώσα,

3.1.2. Ο κανόνας του Hebb

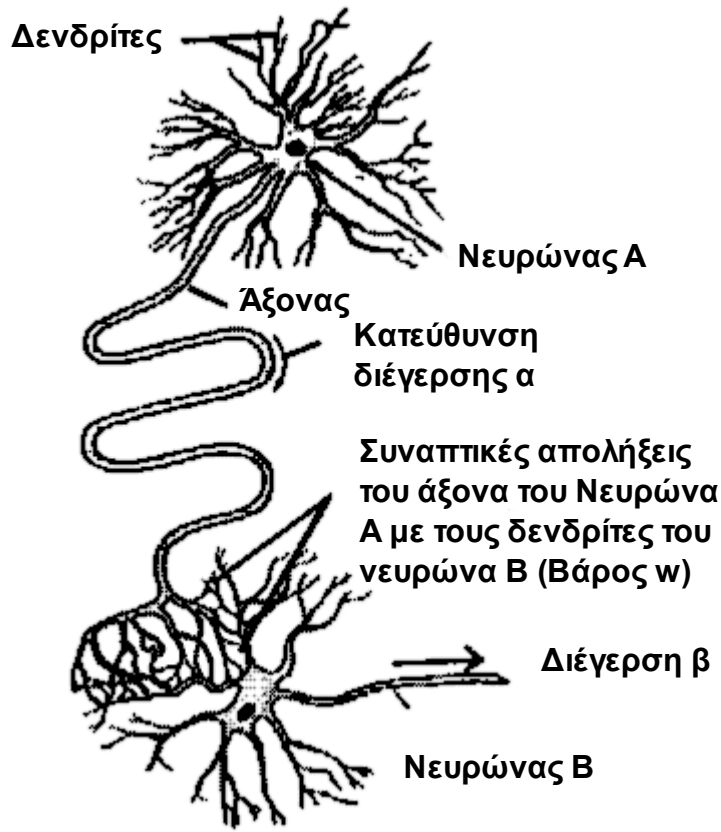
Απαραχγή για μία σειρά θεωριών μάθησης για τα νευρωνικά μοντέλα αποτέλεσε και το σημαντικό έργο του Donald Hebb που δημοσιεύτηκε στα τέλη της δεκαετίας του '40. Ο Hebb, στο 4ο κεφάλαιο του βιβλίου του The Organization of behavior (Hebb 1949) διατύπωσε την ακόλουθη υπόθεση σχετικά με την αυτοοργάνωση ενός συνόλου νευρώνων:

“Αν ο άξονας του κυττάρου A βρίσκεται αρκετά κοντά έτσι ώστε να διεγείρει τη κύτταρο B και επίμονα και επανειλημμένα λαμβάνει μέρος στην διέγερση του B τότε λαμβάνει χώρα κάποια διαδικασία ανάπτυξης ή κάποια αλλαγή στο μεταβολισμό στο ένα από τα δύο ή και στα δύο κύτταρα έτσι ώστε η αποτελεσματικότητα του κυττάρου A στο να διεγείρει το κύτταρο B αυξάνεται.”

Η θεωρία αυτή περιγράφει ένα βασικό μηχανισμό συναπτικής πλαστικότητας όπου η αποτελεσματικότητα της σύναψης αυξάνει από επίμονη και επαναλαμβανόμενη διέγερση του μετασυναπτικού κυττάρου από το προσυναπτικό. Η υπόθεση αυτή

υποστηρίζεται σοβαρά από πειραματικά δεδομένα και έχει προταθεί για την ανάλυση φαινομένων πλαστικότητας σε διάφορα μέρη του εγκεφάλου.

Σύμφωνα με τον Hebb, ο πιο πιθανός τρόπος αύξησης της αποτελεσματικότητας είναι η αύξηση της επιφάνειας του συναπτικού κόμβου που συνδέει τον άξονα του κυττάρου Α με τον λαμβάνοντα δενδρίτη του κυττάρου Β. Έτσι ο άξονας του πρώτου κυττάρου είτε αναπτύσσει νέους συναπτικούς κόμβους είτε αυξάνει την επιφάνεια των ήδη υπαρχόντων κόμβων. Αυτή η διαδικασία δημιουργεί δομικές αλλαγές στο νευρωνικό δίκτυο και αποτελεί στην ουσία μία διαδικασία μάθησης. Αν και ο παραπάνω κανόνας είναι αρκετά ασαφής για να χρησιμοποιηθεί πρακτικά έχουν αναπτυχθεί μαθηματικές θεωρίες βασισμένες στη Χεμπιανή φιλοσοφία που έχουν αρκετά σαφή μορφή για να χρησιμοποιηθούν στην πράξη. Παρόλο που η Χεμπιανή θεωρία ήταν βασισμένη στη μετάδοση των σημάτων μεταξύ των νευρικών κυττάρων μέσω ηλεκτρικών διεγέρσεων και στη συνέχεια ανακαλύφθηκε ότι η κυρίαρχη μορφή της μετάδοσης των σημάτων στις συνάψεις είναι χημική, τα σύγχρονα τεχνητά νευρωνικά δίκτυα εξακολουθούν να βασίζονται σ' αυτή τη θεωρία.



Εικόνα 7: Σύμφωνα με τον κανόνα του Hebb το συναπτικό βάρος μεταξύ του νευρώνα Α και του νευρώνα Β αυξάνεται ανάλογα με το γινόμενο των ενεργοποιήσεων των δύο νευρώνων

Μια απλή μαθηματική διατύπωση του κανόνα του Hebb είναι ότι η μεταβολή του συναπτικού βάρους w είναι ανάλογη του γινομένου της ενεργοποίησης b του νευρώνα Β και της ενεργοποίησης a του νευρώνα Α.

$$w(k+1) = w(k) + \beta(k)(b(k)a(k)) \quad (16)$$

Όπου η παράμετρος $\beta(k)$ είναι ένας μικρός, ενδεχόμενα μεταβλητός θετικός αριθμός, το βήμα εκπαίδευσης (learning rate). Αν υποθέσουμε ότι οι ενεργοποιήσεις των νευρώνων μπορούν να πάρουν είτε θετικές είτε αρνητικές τιμές, τότε σύμφωνα με την εξίσωση (16) το συναπτικό βάρος αυξάνεται όταν οι δύο νευρώνες είναι σε συμφωνία (και οι δύο θετικοί ή και οι δύο αρνητικοί), ενώ

αντίθετα το συναπτικό βάρος μειώνεται όταν οι δύο νευρώνες βρίσκονται σε αντίθεση (ο ένας θετικός και ο άλλος αρνητικός).

Ένα σημαντικό χαρακτηριστικό του Χεμπιανού μοντέλου μάθησης είναι η εντοπικότητα (locality). Σύμφωνα με την ιδιότητα αυτή, όλη η πληροφορία που χρειάζεται ο συναπτικός κόμβος για να αναπτυχθεί ή να συρρικνωθεί βρίσκεται τοπικά στον κόμβο αυτό.

3.2. Βασικές αρχές της αυτοοργάνωσης

Στα αυτοοργανούμενα συστήματα, τα οποία εκπαιδεύονται χωρίς επίβλεψη, η διαδικασία της μάθησης αποτελείται από επαναλαμβανόμενες μεταβολές των συναπτικών βαρών του νευρωνικού δικτύου σε ανταπόκριση στα πρότυπα ενεργοποίησης που χρησιμοποιούνται κατά την εκπαίδευση και σύμφωνα με κάποιους προκαθορισμένους κανόνες, μέχρι να αναπτυχθεί η τελική διαμόρφωση του δικτύου. Το κύριο ερώτημα είναι πώς μπορεί μέσω της διαδικασίας της αυτοοργάνωσης να προκύψει μία χρήσιμη διαμόρφωση. Η απάντηση βρίσκεται στην εξής παρατήρηση (Turing 1952):

Η καθολική τάξη μπορεί να προκύψει μέσα από τοπικές αλληλεπιδράσεις.

Η παρατήρηση αυτή έχει θεμελιώδη σημασία. Εφαρμόζεται στον εγκέφαλο αλλά και στα τεχνητά νευρωνικά δίκτυα. Πιο συγκεκριμένα, πολλές τυχαίες αλληλεπιδράσεις μεταξύ γειτονικών νευρώνων ενός δικτύου μπορούν να συνενωθούν σε καταστάσεις καθολικής τάξης και τελικά να οδηγήσουν σε συνεπή συμπεριφορά, που είναι και η ουσία της αυτοοργάνωσης.

Η αυτοοργάνωση του δικτύου πραγματοποιείται σε δύο διαφορετικά επίπεδα, τα οποία αλληλεπιδρούν με τη μορφή ενός βρόχου ανατροφοδότησης. Τα δύο επίπεδα είναι:

-Δραστηριότητα. Μερικά πρότυπα δραστηριότητας παράγονται από συγκεκριμένα δίκτυα σε απόκριση στα σήματα εισόδου.

-Συνδεδετικότητα. Η ισχύς των συνδέσεων (συναπτικών βαρών) του δικτύου μεταβάλλονται σε απόκριση στα νευρωνικά σήματα στα πρότυπα δραστηριότητας,

λόγω της πλαστικότητας των συνάψεων του δικτύου.

Η ανάδραση μεταξύ των αλλαγών στα συναπτικά βάρη και οι αλλαγές στα πρότυπα δραστηριότητας του δικτύου πρέπει να είναι θετική, έτσι ώστε να επιτευχθεί αυτοοργάνωση στο δίκτυο. Από αυτό προκύπτει και η πρώτη βασική αρχή της αυτοοργάνωσης (von der Malsburg, 1990):

ΑΡΧΗ 1. Οι μεταβολές στα συναπτικά βάρη τείνουν να αυτοενισχύονται.

Η διαδικασία της αυτοενίσχυσης περιορίζεται από την απαίτηση ότι οι αλλαγές που γίνονται στα συναπτικά βάρη πρέπει να είναι βασισμένες σε τοπικά διαθέσιμα σήματα, τα οποία πιο συγκεκριμένα είναι το προσυναπτικό σήμα και το μετασυναπτικό σήμα.. Οι απαιτήσεις της τοπικότητας και της αυτοενίσχυσης αρκούν για να καθορίσουν το μηχανισμό με τον οποίο μία ισχυρή σύναψη οδηγούν στη σύμπτωση του προσυναπτικού και μετασυναπτικού σήματος και κατά συνέπεια, η σύναψη ενισχύεται από αυτή τη σύμπτωση. Ο μηχανισμός αυτός στην ουσία είναι μία επαναδιατύπωση του κανόνα του Hebb για την εκπαίδευση του δικτύου.

Για να μπορέσει το σύστημα να σταθεροποιηθεί, θα πρέπει να υπάρχει συναγωνισμός για κάποιους περιορισμένους πόρους (π.χ. αριθμός των συνδέσεων, ενεργειακοί πόροι κτλ.). Πιο συγκεκριμένα, μία αύξηση στην ισχύ κάποιων συνδέσεων του δικτύου πρέπει να αντισταθμίζεται από μία αποδυνάμωση κάποιων άλλων. Έτσι μόνο οι επιτυχημένες συνάψεις θα μπορέσουν να αναπτυχθούν ενώ οι λιγότερο επιτυχημένες τείνουν να εξασθενήσουν και μπορεί τελικά να εξαφανιστούν. Αυτή η παρατήρηση οδηγεί στη δεύτερη αρχή της αυτοοργάνωσης:

ΑΡΧΗ 2. Ο περιορισμός των πόρων οδηγεί στον ανταγωνισμό μεταξύ των συνάψεων και κατά συνέπεια στην επιλογή των πλέον δυναμικά αναπτυσσομένων συνάψεων εις βάρος των υπολοίπων.

Η αρχή αυτή είναι εφικτή χάρη στην πλαστικότητα των συνάψεων.

Τέλος παρατηρούμε ότι μία και μόνο σύναψη από μόνη της δε μπορεί να αποφέρει επιθυμητά αποτελέσματα. Κατά συνέπεια απαιτείται συνεργασία μεταξύ μίας ομάδας συνάψεων, οι οποίες καταλήγουν στον ίδιο νευρώνα, και να μεταφέρουν σήματα τα οποία ταυτίζονται αρκετά ώστε να ενεργοποιήσουν το νευρώνα αυτό. Κατά συνέπεια:

ΑΡΧΗ 3. Οι μεταβολές στα συναπτικά βάρη έχουν την τάση να συνεργάζονται.

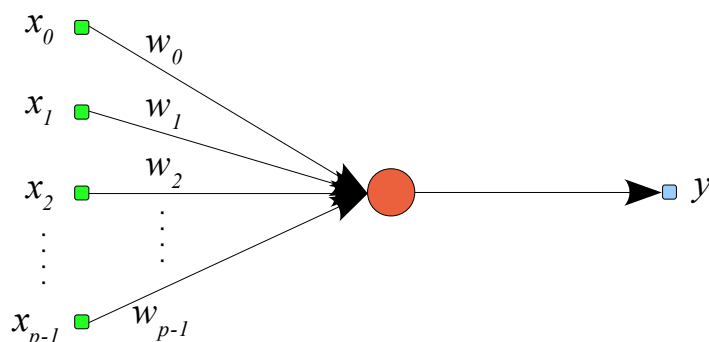
Η παρουσία μίας ισχυρής σύναψης μπορεί να ενισχύσει της υπόλοιπες συνάψεις, παρά τη συνολική ανταγωνιστικότητα του δικτύου. Αυτή η μορφή της συνεργασίας μπορεί να προκύψει λόγω της πλαστικότητας των συνάψεων ή λόγω ταυτόχρονων ερεθισμάτων των προσυναπτικών νευρώνων που προκαλούνται από την παρουσία των κατάλληλων συνθηκών στο εξωτερικό περιβάλλον.

Και οι τρεις αρχές αναφέρονται στο ίδιο το δίκτυο. Για να μπορέσει όμως να επιτευχθεί κάποια χρήσιμη λειτουργία επεξεργασίας δεδομένων μέσω της αυτοοργάνωσης, θα πρέπει από το περιβάλλον να υπάρχει πληθώρα στα πρότυπα ενεργοποίησης τα οποία τροφοδοτούν το δίκτυο.

3.3. Γραμμικός νευρώνας ως υλοποίηση εξαγωγής της πρώτης ιδιοτιμής

Υπάρχει μία στενή αντιστοιχία μεταξύ της συμπεριφοράς ενός αυτοοργανούμενου νευρωνικού δικτύου και της στατιστικής μεθόδου της ανάλυσης κυρίων συνιστωσών.

Ας αναλογιστούμε το μοντέλο ενός απλού γραμμικού νευρώνα, όπως εμφανίζεται στην παρακάτω εικόνα:



Σχήμα 5: Μοντέλο ενός απλού γραμμικού νευρώνα με p εισόδους

Το μοντέλο είναι γραμμικό γιατί με την έννοια ότι η έξοδος του νευρωνικού δικτύου είναι ο γραμμικός συνδυασμός των εισόδων του. Ο νευρώνας δέχεται ένα σύνολο από p σήματα εισόδου x_0, x_1, \dots, x_{p-1} μέσω p συνάψεων με βάρη w_0, w_1, \dots, w_{p-1} αντίστοιχα. Η έξοδος του μοντέλου y επομένως ορίζεται ως:

$$y = \sum_{i=0}^{p-1} w_i x_i \quad (17)$$

Έχοντας ένα μόνο νευρώνα, δε χρειάζονται διπλοί δείκτες για τον ορισμό των συναπτικών βαρών του δικτύου.

Τα βάρη κατά την εκπαίδευση του νευρώνα προσαρμόζονται σύμφωνα με τον κανόνα του Hebb. Το συναπτικό βάρος w_i μεταβάλλεται με το χρόνο, ενισχυόμενο όταν το σήμα εισόδου x_i και το σήμα εξόδου y ταυτίζονται. Πιο αναλυτικά:

$$w_i(k+1) = w_i(k) + \beta(k) y(k) x_i(k), \quad i=0, 1, \dots, p-1 \quad (18)$$

Όπου k είναι η διακριτή χρονική στιγμή και η παράμετρος $\beta(k)$ είναι το βήμα εκπαίδευσης τη χρονική στιγμή k , το οποίο θα μπορούσε να είναι και σταθερό ή

να μεταβάλλεται με το χρόνο. Γενικά είναι ένας μικρός θετικός αριθμός, ο οποίος παίζει καθοριστικό ρόλο στην ταχύτητα της εκπαίδευσης του δικτύου. Ωστόσο, ο κανόνας αυτός στη βασική του αυτή μορφή οδηγεί σε ολοένα αυξανόμενα συναπτικά βάρη w_i γεγονός το οποίο σε πρακτικό επίπεδο δεν μπορεί να γίνει αποδεκτό. Μπορούμε να υπερβούμε αυτό το πρόβλημα, ενσωματώνοντας κάποια μορφή κορεσμού ή κανονικοποίησης στον κανόνα εκπαίδευσης για την προσαρμογή των βαρών. Η χρήση της κανονικοποίησης έχει σαν αποτέλεσμα την εμφάνιση ανταγωνισμού μεταξύ των συνάψεων του νευρώνα για περιορισμένους πόρους, που είναι ουσιώδες για τη σταθερότητα του δικτύου. Από μαθηματικής απόψεως, η νέα, βολική μορφή της κανονικοποίησης έχει ως εξής: (Oja, 1982)

$$w_i(k+1) = \frac{w_i(k) + \beta(k)y(k)x_i(k)}{\left(\sum_{i=0}^{p-1} [w_i(k) + \beta(k)y(k)x_i(k)]^2\right)^{1/2}} \quad (19)$$

όπου το άθροισμα στον παρονομαστή περιλαμβάνει το σύνολο των συνάψεων που σχετίζονται με το νευρώνα. Κάνοντας την παραδοχή ότι ο βαθμός εκμάθησης β είναι πολύ μικρός μπορούμε να αναπτύξουμε την προηγούμενη εξίσωση ως:

$$w_i(k+1) = w_i(k) + \beta(k)y(k)[x_i(k) - y(k)w_i(k)] + O\beta(k)^2 \quad (20)$$

Για πολύ μικρό β , μπορούμε να παραλείψουμε τον όρο $O\beta(k)^2$ και η εξίσωση γίνεται:

$$w_i(k+1) = w_i(k) + \beta(k)y(k)[x_i(k) - y(k)w_i(k)] \quad (21)$$

Ο όρος $\beta(k)y(k)x_i(k)$ στο δεξί μέρος της εξίσωσης αναπαριστά τη μεταβολή σύμφωνα με τον κανόνα του Hebb του συναπτικού βάρους w_i και έτσι δρα ως

θετική ανατροφοδότηση και αντιπροσωπεύει την πρώτη αρχή της αυτοοργάνωσης. Ο αρνητικός όρος $-\beta(k)y(k)w_i(k)$ είναι υπεύθυνος για τη σταθερότητα του δικτύου σύμφωνα με τη δεύτερη αρχή της αυτοοργάνωσης. Μετατρέπει την είσοδο $x_i(k)$ σε μορφή που εξαρτάται από το σχετικό συναπτικό βάρος $w_i(k)$ και την έξοδο $y(k)$, όπως φαίνεται από:

$$x_i'(k) = x_i(k) - y(k)w_i(k) \quad (22)$$

το οποίο αντιπροσωπεύει την ενεργή είσοδο της i σύναψης. Επομένως ο κανόνας εκπαίδευσης μπορεί να ξαναγραφεί και ως:

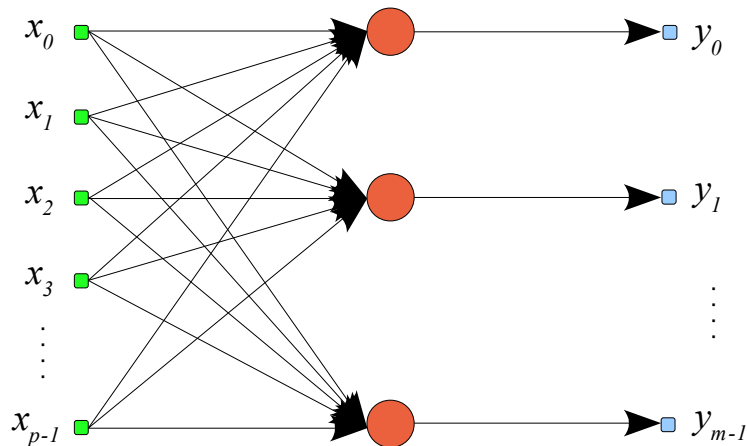
$$w_i(k+1) = w_i(k) + \beta(k)y(k)x_i'(k) \quad (23)$$

Αποδεικνύεται ότι ένα τέτοιο γραμμικό νευρωνικό μοντέλο με τον παραπάνω κανόνα εκπαίδευσης έχει την τάση να εξάγει την πρώτη κύρια συνιστώσα από τα διανύσματα εισόδου.

3.4. Αυτοοργανούμενη Ανάλυση Κυρίων Συνιστωσών

Θα δείξουμε στη συνέχεια ότι μπορούμε να γενικεύσουμε τον προηγούμενο αλγόριθμο εκπαίδευσης και να εκπαιδεύσουμε ένα δίκτυο πρόσθιας τροφοδότησης με ένα μόνο στρώμα γραμμικών νευρώνων. Ο σκοπός είναι να παράγουμε ένα δίκτυο, το οποίο μπορεί να εκτελεί ανάλυση κυρίων συνιστωσών με διανύσματα εισόδου οποιουδήποτε μεγέθους.

Ας αναλογιστούμε το παρακάτω δίκτυο, όπως φαίνεται στο σχήμα (6)



Σχήμα 6: Δίκτυο πρόσθιας τροφοδότησης με ένα μόνο στρώμα νευρώνων

Κάνουμε τις εξής παραδοχές όσον αφορά τη δομή του δικτύου:

1. Ο κάθε νευρώνας του στρώματος εξόδου του δικτύου είναι γραμμικός.
2. Το δίκτυο έχει p εισόδους και m εξόδους, τα οποία και τα δύο καθορίζονται. Επιπλέον το δίκτυο έχει πάντα λιγότερες εξόδους από ότι εισόδους ($m < p$).

Τα συστατικά του δικτύου τα οποία υπόκεινται στην εκπαίδευση είναι το σύνολο των βαρών των συνάψεων που συνδέουν τους i κόμβους του στρώματος εισόδου με τους υπολογιστικούς κόμβους j στο στρώμα εξόδου, όπου $i=0,1, \dots, p-1$ και $j=0,1, \dots, m-1$

Η έξοδος $y_j(k)$ του νευρώνα j τη χρονική στιγμή k παράγεται από το στρώμα εισόδου με σύνολο τιμών $\{x_i(k), i=0,1, \dots, p-1\}$ και δίνεται από τον παρακάτω τύπο:

$$y_j(k) = \sum_{i=0}^{p-1} w_{ji}(k) x_i(k), \quad j=0,1, \dots, m-1 \quad (24)$$

Τα συναπτικά βάρη $w_{ji}(k)$ προσαρμόζονται σύμφωνα με την γενικευμένη μορφή του κανόνα του Hebb όπως φαίνεται παρακάτω:

$$i=0,1, \dots, p-1 \quad (25)$$

$$\Delta w_{ji}(k) = \beta(k) [y_j(k)x_i(k) - y_j(k) \sum_{n=0}^j w_{ni}(k)y_n(k)], \quad j=0,1,\dots,m-1$$

Όπου $\Delta w_{ji}(k)$ είναι η μεταβολή που εφαρμόζεται στο συναπτικό βάρος $w_{ji}(k)$ τη χρονική στιγμή k , και $\beta(k)$ είναι πάλι το βήμα εκπαίδευσης. Η γενικευμένη μορφή του αλγορίθμου του Hebb (generalized Hebbian algorithm ή GHA) για ένα στρώμα m νευρώνων συμπεριλαμβάνει τον αλγόριθμο της εξίσωσης (18) για ένα νευρώνα σαν ειδική περίπτωση, όπου $j=0$.

Για να καταλάβουμε καλύτερα τη συμπεριφορά της γενικευμένης μορφής του αλγορίθμου του Hebb, ξαναγράφουμε την εξίσωση (25) στη μορφή:

$$\Delta w_{ji}(k) = \beta(k) y_j(k) [x_i'(k) - w_{ji}(k) y_j(k)], \quad \begin{matrix} i=0,1,\dots,p-1 \\ j=0,1,\dots,m-1 \end{matrix} \quad (26)$$

όπου $x_i'(k)$ είναι μία τροποποιημένη μορφή του i στοιχείου του διανύσματος εισόδου, συναρτήσει του δείκτη j όπως φαίνεται παρακάτω:

$$x_i'(k) = x_i(k) - \sum_{n=0}^{j-1} w_{ni}(k) y_n(k) \quad (27)$$

Για ένα συγκεκριμένο νευρώνα j , ο αλγόριθμος που περιγράφεται στην εξίσωση (26) έχει ακριβώς την ίδια μαθηματική μορφή με την εξίσωση (18), εκτός από το γεγονός ότι το σήμα εισόδου $x_i(k)$ αντικαθίσταται με την τροποποιημένη μορφή $x_i'(k)$ στην εξίσωση (26). Μπορούμε να προχωρήσουμε ένα βήμα παραπέρα και να ξαναγράψουμε την εξίσωση (27) σε μια μορφή που συμφωνεί με τις αξιώσεις του κανόνα του Hebb όπως φαίνεται παρακάτω:

$$\Delta w_{ji}(k) = \beta(k) y_j(k) x_i''(k) \quad (28)$$

όπου

$$x_i''(k) = x_i'(k) - w_{ji}(k) y_j(k) \quad (29)$$

Επομένως, παρατηρώντας ότι

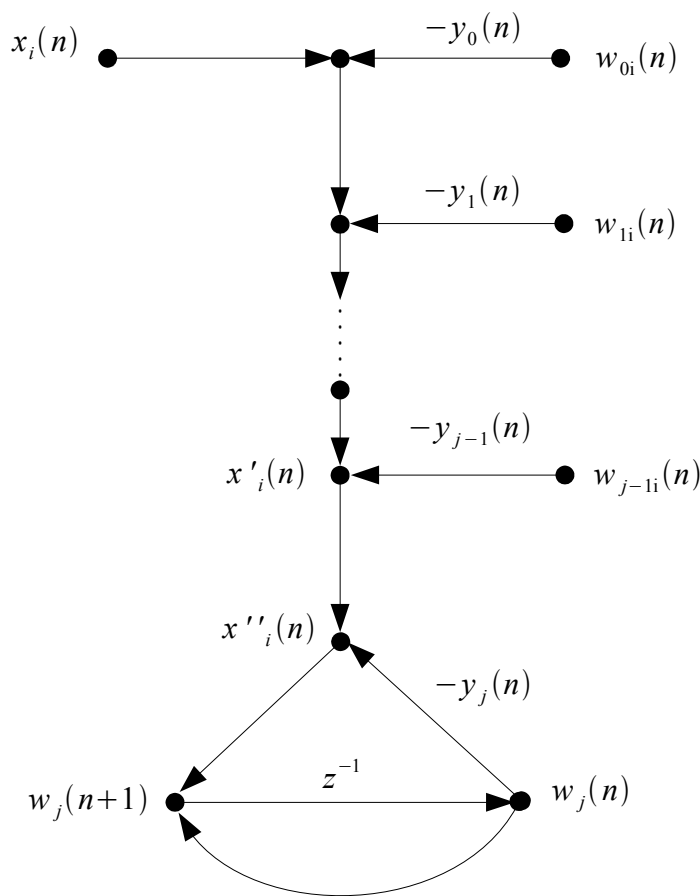
$$w_{ji}(k+1) = w_{ji} + \Delta w_{ji}(k) \quad (30)$$

και

$$w_{ji}(k) = z^{-1}[w_{ji}(k+1)] \quad (31)$$

όπου z^{-1} είναι ο τελεστής μοναδιαίας καθυστέρησης, μπορούμε να κατασκευάσουμε το διάγραμμα σήματος-ροής της γενικευμένης μορφής του αλγόριθμου του Hebb όπως περιγράφεται εδώ (Σχήμα 7).

Με την έννοια του τελεστή μοναδιαίας καθυστέρησης ορίζουμε τον παράγοντα που όταν εφαρμοστεί σε ένα δείγμα της ακολουθίας, τότε μας δίνει το αμέσως προηγούμενο δείγμα και συμβολίζεται ως z^{-1}



Σχήμα 7: Διάγραμμα ροής σήματος του γενικευμένου αλγόριθμου του Hebb

Για να καταλάβουμε καλύτερα πως ακριβώς λειτουργεί ο αλγόριθμος του Hebb ξαναγράφουμε τη μορφή του αλγορίθμου όπως φαίνεται παρακάτω:

$$\Delta w_j(k) = \beta(k) y_j(k) x'(k) - \beta(k) y_j^2(k) w_j(k) \quad (32)$$

όπου

$$x'(k) = x(k) - \sum_{n=0}^{j-1} w_n(k) y_n(k) \quad (33)$$

Το διάνυσμα $x'(k)$ παριστάνει μία τροποποιημένη μορφή του διανύσματος εισόδου. Σύμφωνα με την εξίσωση (32) μπορούμε να κάνουμε τις παρακάτω παρατηρήσεις:

1. Για τον πρώτο νευρώνα του δικτύου πρόσθιας τροφοδότησης έχουμε:

$$j=0: \quad x'(k) = x(k) \quad (34)$$

Σ' αυτή την περίπτωση, ο γενικευμένος αλγόριθμος του Hebb μειώνεται σε αυτόν της εξίσωσης (18) για ένα νευρώνα. Όπως έχουμε ήδη αναφέρει, γνωρίζουμε ότι ο νευρώνας αυτός θα ανακαλύψει την πρώτη κύρια συνιστώσα (δηλ. τη μεγαλύτερη ιδιοτιμή και το σχετικό ιδιοδιάνυσμα) του διανύσματος εισόδου $x(n)$.

2. Για το δεύτερο νευρώνα του δικτύου έχουμε:

$$j=1: \quad x'(k) = x(k) - w_0(k) y_0(k) \quad (35)$$

Δεδομένου ότι ο πρώτος νευρώνας έχει ήδη συγκλίνει στην πρώτη κύρια συνιστώσα, ο δεύτερος νευρώνας βλέπει ένα διάνυσμα εισόδου $x'(k)$ από το οποίο το πρώτο ιδιοδιάνυσμα έχει αφαιρεθεί από τον πίνακα συσχέτισης R. Κατά συνέπεια, ο δεύτερος νευρώνας εξάγει την πρώτη κύρια συνιστώσα του διανύσματος $x'(k)$, η οποία ισοδυναμεί με τη δεύτερη κύρια συνιστώσα (δηλ.

τη δεύτερη μεγαλύτερη ιδιοτιμή και το αντίστοιχο ιδιοδιάνυσμα) του αρχικού διανύσματος $x(k)$.

3. Για τον 3ο νευρώνα έχουμε:

$$j=2: \quad x'(k) = x(k) - w_0(k)y_0(k) - w_1(k)y_1(k) \quad (36)$$

Ας υποθέσουμε ότι οι δύο πρώτοι νευρώνες έχουν ήδη συγκλίνει στην πρώτη και δεύτερη κύρια συνιστώσα, όπως περιγράφηκε στα παραπάνω βήματα. Ο τρίτος νευρώνας τώρα βλέπει το διάνυσμα εισόδου $x'(k)$ από το οποίο τα δύο πρώτα ιδιοδιανύσματα έχουν ήδη αφαιρεθεί. Κατά συνέπεια εξάγει την πρώτη κύρια συνιστώσα του διανύσματος εισόδου $x'(k)$ που ορίζεται στην εξίσωση (36), η οποία ισοδυναμεί με την τρίτη κύρια συνιστώσα του αρχικού διανύσματος εισόδου $x(k)$.

4. Επαναλαμβάνοντας τα παραπάνω και για τους υπόλοιπους νευρώνες του δικτύου πρόσθιας τροφοδότησης είναι πλέον προφανές ότι η κάθε έξοδος του δικτύου, που εκπαιδεύτηκε σύμφωνα με τη γενικευμένη μορφή του αλγορίθμου του Hebb, αντιπροσωπεύει ένα συγκεκριμένο ιδιοδιάνυσμα του πίνακα συσχετίσεων του διανύσματος εισόδου και οι μεμονωμένες έξοδοι αντιστοιχούν στις ιδιοτιμές ταξινομημένες κατά φθίνουσα σειρά.

Η μέθοδος αυτή υπολογισμού των ιδιοδιανυσμάτων είναι παρόμοια με την τεχνική γνωστή ως τεχνική αποπληθωρισμού του Hotelling (Kreyzig, 1988) και ακολουθεί παρόμοια διαδικασία με τον ορθογωνισμό Gram – Schmidt (Strang, 1980).

Η περιγραφή νευρώνα - νευρώνα όπως παρουσιάστηκε παραπάνω έχει κυρίως ως σκοπό την απλοποίηση της επεξήγησης της διαδικασίας. Στην πράξη όλοι οι νευρώνες κατά την εκπαίδευση σύμφωνα με τη γενικευμένη μορφή του αλγορίθμου του Hebb συγκλίνουν όλοι μαζί ταυτόχρονα και ο συνολικός χρόνος εκπαίδευσης είναι μικρότερος από ότι αν εκπαιδεύονταν σειριακά, ένας κάθε φορά. Παρόλα αυτά, ο δεύτερος νευρώνας είναι απίθανο να συγκλίνει σωστά, προτού ο πρώτος νευρώνας είναι πολύ κοντά στο πρώτο ιδιοδιάνυσμα.

Περίληψη του γενικευμένου αλγορίθμου του Hebb (GHA)

Οι υπολογισμοί που εμπλέκονται στην υλοποίηση του αλγορίθμου του Hebb είναι απλοί και μπορούν να συνοψιστούν στα παρακάτω:

1. Αρχικοποίηση των συναπτικών βαρών του δικτύου w_{ji} με μικρές τυχαίες τιμές κατά τη χρονική στιγμή $k=1$. Ανάθεση μίας μικρής θετικής τιμής στο βήμα εκπαίδευσης β .

2. Για $k=1$, $j=0,1,\dots,m-1$ και για $i=0,1,\dots,p-1$, υπολογισμός των

$$y_j(k) = \sum_{i=0}^{p-1} w_{ji}(k) x_i(k) \quad (37)$$

$$\Delta w_{ji}(k) = \beta(k) [y_j(k) x_i(k) - y_j(k) \sum_{n=0}^j w_{nj}(k) y_n(k)] \quad (38)$$

όπου $x_i(k)$ είναι το i στοιχείο του p επί 1 διανύσματος $x(k)$ και m είναι ο επιθυμητός αριθμός των κυρίων συνιστωσών.

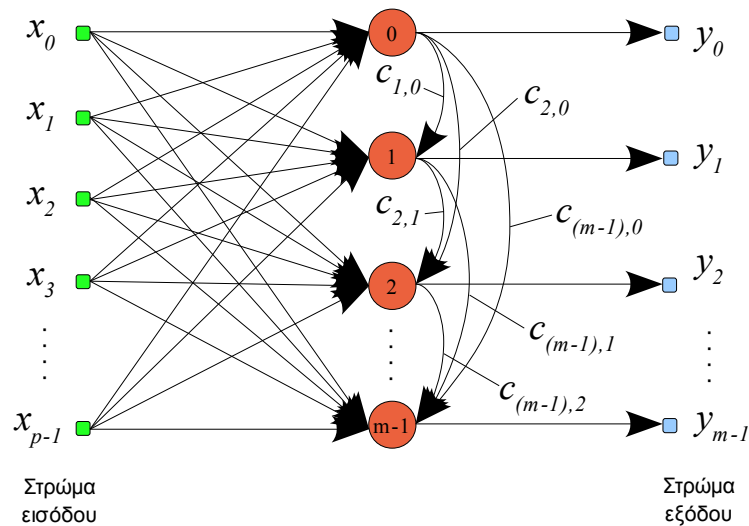
3. Αύξηση του k κατά 1, τυχόν προσαρμογή του βήματος εκπαίδευσης, και επαναλαμβάνουμε το βήμα 2 έως ότου τα συναπτικά βάρη w_{ji} σταθεροποιηθούν. Για μεγάλο k , τα συναπτικά βάρη w_{ji} του νευρώνα j συγκλίνουν προς το i στοιχείο του ιδιοδιανύσματος που σχετίζεται με την j ιδιοτιμή του πίνακα συσχετίσεων του διανύσματος εισόδου $x(k)$.

3.5. Προσαρμοστική Ανάλυση Κυρίων Συνιστωσών με Πλευρική Αναστολή

Ο γενικευμένος αλγόριθμος του Hebb που περιγράψαμε στο προηγούμενο κεφάλαιο βασίζεται αποκλειστικά σε συνδέσεις πρόσθιας τροφοδότησης για την εύρεση διαδοχικών ιδιοδιανυσμάτων. Ο Földiák (Földiák and Fdilir 1989) επέκτεινε τη δομή του δικτύου για τέτοιες εφαρμογές περιλαμβάνοντας και συνδέσεις ανατροφοδότησης αντίθετες προς τον κανόνα του Hebb. Το κίνητρο για την

τροποποίηση αυτή προήλθε από προηγούμενη εργασία των Barlon και Földiák (Barlow and Foldiak 1989) πάνω στην προσαρμοστικότητα και διακριτική ικανότητα του οπτικού κέντρου του εγκεφάλου, όπου υποστηριζόταν ότι αν οι νεωρώνες αλληλεπιδρούν σύμφωνα με κανόνες αντίθετους προς τον Hebb, τότε οι έξοδοι των νευρώνων καθορίζουν ένα σύστημα συντεταγμένων στο οποίο δεν υπάρχουν συσχετίσεις, ακόμα και όταν τα δεδομένα εισόδου έχουν ισχυρές συσχετίσεις. Οι Kung και Διαμαντάρας (Diamantaras and Kung 1996), βασιζόμενοι σε προηγούμενη εργασία του Oja (Erkki Oja and Karhunen 1982) και του Földiák (Földiák and Fdilir 1989) ανέπτυξαν ένα νέο αλγόριθμο, ονομαζόμενο αλγόριθμο APEX για τον αναδρομικό υπολογισμό των κυρίων συνιστωσών. Το ακρωνύμιο APEX σημαίνει Adaptive Principal components EXtraction (δηλ. προσαρμοστική εξαγωγή κυρίων συνιστωσών). Το ενδιαφέρον χαρακτηριστικό του αλγορίθμου APEX είναι ότι αν του δώσουμε τη j κύρια συνιστώσα, ο αλγόριθμος μπορεί με επαναληπτικό τρόπο να υπολογίσει τη $(j+1)$ κύρια συνιστώσα. Στο κεφάλαιο αυτό θα αναπτύξουμε τον αλγόριθμο APEX σαν ένα αλγόριθμο για τον προσαρμοστικό υπολογισμό των κυρίων συνιστωσών.

Το παρακάτω σχήμα (8) δείχνει το μοντέλο του δικτύου από το οποίο προκύπτει ο αλγόριθμος.



Σχήμα 8: Η δομή του νευρωνικού δικτύου APEX. Κάθε νευρώνας τροφοδοτείται πλαγίως και από όλους τους προηγούμενους νευρώνες (αυτούς που βρίσκονται από πάνω του).

Το διάνυσμα εισόδου x έχει πάλι διάσταση p , και αποτελείται από τα επιμέρους στοιχεία x_0, x_1, \dots, x_{p-1} . Και εδώ ο κάθε νευρώνας θεωρείται γραμμικός. Όπως φαίνεται και στο σχήμα υπάρχουν δύο είδη συναπτικών συνδέσεων στο δίκτυο:

- Συνδέσεις πρόσθιας τροφοδότησης από τους κόμβους εισόδου προς κάθε έναν από τους νευρώνες $0, 1, \dots, j$, όπου $j < p$

Ιδιαίτερο ενδιαφέρον παρουσιάζουν οι συνδέσεις του νευρώνα j . Οι συνδέσεις αυτές αναπαρίστανται με το διάνυσμα βαρών

$$w_j = [w_0(k), w_1(k), \dots, w_{p-1}(k)]^T \quad (39)$$

Οι συνδέσεις αυτές λειτουργούν διεγερτικά, σύμφωνα με τον κανόνα του Hebb και παρέχουν αυτοενίσχυση.

- Πλευρικές συνδέσεις από τις εξόδους των νευρώνων $0, 1, \dots, j-1$ προς το νευρώνα j , δίνοντας ανάδραση στο δίκτυο. Οι συνδέσεις αυτές αναπαρίστανται με το διάνυσμα βαρών

$$c_j(k) = [c_{j0}(k), c_{j1}(k), \dots, c_{j,j-1}(k)]^T \quad (40)$$

Οι πλευρικές αυτές συνδέσεις λειτουργούν κατασταλτικά, αντίθετα με τον κανόνα του Hebb.

Οι έξοδοι των νευρώνων $y_j(k)$ δίνονται από τον τύπο:

$$y_j(k) = w_j^T(k)x(k) + c_j^T(k)y_{j-1}(k) \quad (41)$$

όπου ο παράγοντας $w_j^T(k)x(k)$ οφείλεται στη συνεισφορά των συνδέσεων πρόσθιας τροφοδότησης, ενώ ο παράγοντας $c_j^T(k)y_{j-1}(k)$ οφείλεται στη συνεισφορά των πλευρικών συνδέσεων. Το διάνυσμα του σήματος ανατροφοδότησης $y_{j-1}(k)$ καθορίζεται από τις εξόδους των νευρώνων

$0, 1, \dots, j-1$:

$$y_{j-1}(k) = [y_0(k), y_1(k), \dots, y_j(k)]^T \quad (42)$$

Από τον τύπο (42) παρατηρούμε ότι οι συνάψεις που συνδέουν τις εισόδους με τις εξόδους συνεισφέρουν με θετικό πρόσημο στην έξοδο, ενώ οι πλάγιες συνδέσεις με αρνητικό. Οι πρώτες συνεργάζονται για την αύξηση της σχετικότητας μεταξύ των σημάτων εισόδου και εξόδου (σύμφωνα με τον κανόνα του Hebb ή Χεμπιανές), ενώ οι δεύτερες εργάζονται για την αποσυσχέτιση ή αλλιώς ορθογωνοποίηση των εξόδων (αντι-Χεμπιανές).

Ένα ιδιαίτερο χαρακτηριστικό του αλγορίθμου APEX είναι το γεγονός ότι τόσο οι Χεμπιανές όσο και οι αντι-Χεμπιανές συνάψεις εκπαιδεύονται με τον ίδιο αλγόριθμο και συγκεκριμένα με τον αλγόριθμο του Oja:

$$w_{ij}(k+1) = w_{ij}(k) + \beta(k)[y_i(k)x_j(k) - y_i(k)^2 w_{ij}(k)] \quad (43)$$

$$c_{ij}(k+1) = c_{ij}(k) + \beta(k)[y_i(k)y_j(k) - y_i(k)^2 c_{ij}(k)] \quad (44)$$

$$i=1, \dots, m, \quad j=1, \dots, i-1$$

Μια απόρροια της χρήσης του κανόνα του Oja είναι η εντοπικότητα του αλγορίθμου APEX. Πράγματι, ο κανόνας εκπαίδευσης της σύναψης w_{ij} η οποία συνδέει την είσοδο j με την έξοδο i χρησιμοποιεί μόνο τις τιμές x_j , y_i και w_{ij} . Σε διανυσματική μορφή έχουμε τέλος τους παρακάτω κανόνες εκπαίδευσης:

$$w_i(k+1) = w_i(k) + \beta(k)[y_i(k)x(k) - y_i(k)^2 w_i(k)] \quad (45)$$

$$c_i(k+1) = c_i(k) + \beta(k)[y_i(k)y_{(i-1)}(k) - y_i(k)^2 c_i(k)] \quad (46)$$

Παρατηρούμε ότι για τον πρώτο νευρώνα ($i=1$) το διάνυσμα c_1 δεν ορίζεται (είναι κενό). Συνεπώς ο κανόνας εκπαίδευσης για το νευρώνα αυτόν περιέχει μόνο την εξίσωση (45).

Η τοπολογία του δικτύου APEX διασφαλίζει ότι τα βάρη των πρόσθιων συνάψεων (W_i) συγκλίνουν κατά την εκπαίδευση στην πρώτη κύρια συνιστώσα και τα βάρη των πλευρικών συνάψεων τείνουν να μηδενιστούν:

$$c_{ij}[k] \rightarrow 0; i=1, \dots, m; j=1, \dots, i-1 \quad k \rightarrow \infty \quad (47)$$

Ο ρόλος του κάτω τριγωνικού πίνακα των πλευρικών βαρών είναι να διασφαλίσει την ορθογωνιοποίηση του πίνακα των βαρών σύμφωνα με τη διαδικασία Gram-Schmidt.[1]

3.6. Η βέλτιστη παράμετρος εκπαίδευσης β

Η στενή σχέση μεταξύ της ανάλυσης κυρίων συνιστωσών και βελτιστοποίησης ελαχίστων τετραγώνων οδηγεί φυσιολογικά και σε μία επίσης στενή σχέση μεταξύ του αλγορίθμου APEX και του Αναδρομικού Αλγόριθμου Ελαχίστων Τετραγώνων γνωστού και ως Recursive Least Squares (RLS). Ο κανόνας RLS ελαχιστοποιεί τη συνάρτηση κόστους

$$J_{LS}(\mathbf{w}) = \sum_{k=0}^N \gamma^{N-k} [\alpha(k) - \mathbf{b}(k)^T \mathbf{w}] \quad (48)$$

Όπου $\mathbf{b}(k)$ είναι ένα τυχαίο διανυσματικό σήμα ενώ $\alpha(k)$ είναι ένα τυχαίο βαθμωτό σήμα. Ο συντελεστής γ ($0 < \gamma \leq 1$) καλείται συντελεστής λήθης (forgetting factor) και μας δίνει μία εκθετικά φθίνουσα βαρύτητα στα δείγματα που συμμετέχουν στο άθροισμα. Στην πράξη, με τον συντελεστή γ ορίζουμε ένα παράθυρο μεγέθους περίπου $M = 1/1 - \gamma$ δειγμάτων τα οποία έχουν κάποια ουσιαστική επιρροή στον υπολογισμό του παραπάνω αθροίσματος.

Λύνοντας το παραπάνω πρόβλημα (Διαμαντάρας 2007) καταλήγουμε στον εξής αλγόριθμο:

$$w_{m,i}(k+1) = W_{m,i}(k) + \varepsilon(k) [x_i(k) - y_m(k) W_{m,i}(k)] \quad (49)$$

και

$$\varepsilon(k) = \left[\sum_{i=1}^k \gamma^{(k-i)} y_m(i)^2 \right]^{-1} y_m(k) \quad (50)$$

Θέτοντας

$$\beta(k) = \frac{1}{\sum_{i=1}^k \gamma^{k-i} y_m(i)^2} \quad (51)$$

και χρησιμοποιώντας διανυσματικό συμβολισμό παίρνουμε:

$$\mathbf{w}_m(k+1) = \mathbf{w}_m(k) + \beta(k) [y_m(k) \mathbf{x}(k) - y_m(k)^{2\text{bold}} \mathbf{w}_m(k)] \quad (52)$$

η οποία όπως παρατηρούμε είναι ισοδύναμη με τον κανόνα μάθησης APEX για τον m νευρώνα. Αντίστοιχα για τις αντι-χεμπιανές συνάψεις c_{ij} προκύπτει η παρακάτω εξίσωση:

$$c_m(k+1) = c_m(k) + \beta(k) [y(k) \mathbf{y}^{(m-1)}(k) - y(k)^2 c_m(k)] \quad (53)$$

Πλέον έχουμε μία συγκεκριμένη τιμή για την παράμετρο μάθησης β η οποία μπορεί να γραφεί με αναδρομική μορφή ως εξής:

$$\beta(k) = b(k-1) / [\gamma + y(k)^2 \beta(k-1)] \quad (54)$$

3.7. Άλλες μέθοδοι ανάλυσης κυρίων συνιστωσών

3.7.1. Ο κανόνας του Földiák

Το 1989 ο Földiák (Földiák and Fdilir 1989) πρότεινε ένα νευρωνικό δίκτυο που μοιάζει πολύ με το APEX με τη διαφορά ότι στον κανόνα Földiák υπάρχουν αμφίπλευρες πλάγιες συνδέσεις στους νευρώνες εξόδου, όποτε ο κάθε νευρώνας εξόδου συνδέεται με όλους τους υπόλοιπους.

Η σχέση εισόδου - εξόδου του δικτύου είναι η εξής:

$$y_i = \mathbf{w}_i^T \mathbf{x} - \sum_{j \neq i} c_{ij} y_j \quad (55)$$

Για τον υπολογισμό του διανύσματος εξόδου απαιτείται αντιστροφή πίνακα:

$$\mathbf{y} = (\mathbf{I} + \mathbf{C})^{-1} \mathbf{W} \mathbf{x} = \mathbf{F} \mathbf{x} \quad (56)$$

Αυτή η αντιστροφή του πίνακα αναιρεί την εντοπικότητα του αλγορίθμου.

Ο κανόνας εκπαίδευσης του δικτύου ακολουθεί επίσης το Χεμπιανό μοντέλο τόσο για τις πρόσθιες, όσο και για τις πλευρικές συνδέσεις:

$$w_{ij}(k+1) = w_{ij}(k) + \beta(k) [y_i(k) x_j(k) - y_i(k) w_{ij}^w(k)] \quad (57)$$

$$c_{ij}(k+1) = c_{ij}(k) + \beta(k) y_i(k) y_j(k) \quad (58)$$

$$i \neq j$$

Το δίκτυο του Földiák με m νευρώνες εξόδου μαθαίνει τον m -διάστατο κύριο υποχώρο του διανύσματος εισόδου. Το αποτέλεσμα του κανόνα του Földiák δεν είναι οι κύριες συνιστώσες, αλλά κάποιος άγνωστος γραμμικός συνδυασμός αυτών.

3.7.2. Ο κανόνας του υποχώρου (subspace rule)

Ο κανόνας αυτός αποτελεί μία επέκταση του μονοδιάστατου κανόνα του Oja σε πολλές διαστάσεις. Οι εξισώσεις που περιγράφουν τον κανόνα αυτό είναι οι παρακάτω:

$$\mathbf{y} = \mathbf{W} \mathbf{x} \quad (59)$$

$$\mathbf{W}(k+1) = \mathbf{W}(k) + \beta(k) [\mathbf{y}(k) \mathbf{x}(k)^T - \mathbf{y}(k)^T \mathbf{W}(k)] \quad (60)$$

Οι ομοιότητες είναι εμφανείς, αρκεί να συγκρίνουμε την παραπάνω εξίσωση (60) με την εξίσωση του κανόνα του Oja (21).

Όπως και ο κανόνας του Földiák, έτσι και ο κανόνας του υποχώρου δεν εξαγει ακριβώς τις κύριες συνιστώσες αλλά κάποιο άγνωστο γραμμικό συνδυασμό τους,

$$W = T[e_1, \dots, e_m]^T \quad (61)$$

3.8. Αναγνώριση - ταξινόμηση νέων προτύπων.

Τα νευρωνικά δίκτυα τα οποία χρησιμοποιούνται για την εξαγωγή των κυρίων συνιστωσών δεν μπορούν να χρησιμοποιηθούν άμεσα για την αναγνώριση - ταξινόμηση νέων προτύπων. Τροφοδοτώντας το δίκτυο με ένα άγνωστο πρότυπο, οι νευρώνες εξόδου μας παρέχουν τις κατάλληλες τιμές για τη βέλτιστη ανακατασκευή του προτύπου με βάση τις ιδιοτιμές που έχουν εξαχθεί κατά τη φάση της εκπαίδευσης του δικτύου.

Τι ακριβώς εννοούμε όμως με τον όρο ταξινόμηση προτύπου; Στην περίπτωση που έχουμε μόνο ένα νευρωνικό δίκτυο, εφόσον έχει εκπαιδευτεί με μια συγκεκριμένη κλάση προτύπων, η μόνη πληροφορία την οποία θα μπορούσε να μας δώσει είναι αν το άγνωστο αυτό πρότυπο με το οποίο το τροφοδοτήσαμε ανήκει στην κλάση αυτή ή όχι. Για να υλοποιηθεί ένα τέτοιο σύστημα θα μπορούσαμε στους νευρώνες εξόδου του δικτύου να προσαρμόσουμε ένα δεύτερο δίκτυο, ένα perceptron πολλαπλών επιπέδων ή ένα δίκτυο backpropagation, το οποίο θα εκπαιδευόταν κατάλληλα, ώστε στην έξοδο του να έχουμε δυο διακριτές τιμές - καταστάσεις που θα αντιστοιχούν στην ταξινόμηση ή όχι του προτύπου στην κλάση των προτύπων εκπαίδευσης. Αυτή η προσέγγιση έχει το μειονέκτημα ότι πλέον το δίκτυο PCA χρησιμοποιείται απλά σαν ένας μειωτής των διαστάσεων των δεδομένων εισόδου με σκοπό τη μείωση της πολυπλοκότητας της κύριας εργασίας της ταξινόμησης, η οποία θα γίνεται από το δευτερεύων δίκτυο.

Στην περίπτωση που έχουμε πολλαπλά νευρωνικά δίκτυα, εκπαιδευμένα το καθένα με ξεχωριστή κλάση, η ταξινόμηση έχει την έννοια εξαγωγής συμπεράσματος αν το άγνωστο διάνυσμα εισόδου ανήκει σε μία από τις κλάσεις των δικτύων και σε ποια, ή αν τελικά δεν ανήκει σε καμία από τις γνωστές κλάσεις. Πάλι θα μπορούσε να χρησιμοποιηθεί ένα δίκτυο MLP ή backpropagation με εισόδους τις εξόδους των δικτύων εξαγωγής κυρίων συνιστωσών. Σ' αυτή την περίπτωση όμως η πολυπλοκότητα θα αύξανε κατά πολύ και θα ήταν σχεδόν

εξαιρετικά δύσκολη η προσθήκη ενός δικτύου εκπαιδευμένου με μία άλλη κλάση ή η διαγραφή ενός υπάρχοντος δικτύου, γιατί θα έπρεπε το δεύτερο δίκτυο να εκπαιδευτεί από την αρχή.

Για να λυθούν τα παραπάνω προβλήματα μπορούμε να αξιοποιήσουμε μία χαρακτηριστική ιδιότητα των δικτύων εξαγωγής κυρίων συνιστωσών, την ιδιότητα να ελαχιστοποιούν το μέσο τετραγωνικό σφάλμα της ανακατασκευής του προτύπου. Το σφάλμα ανακατασκευής είναι ένα ενδεικτικό μέτρο της απόστασης του διανύσματος από τον υποχώρο της κλάσης του κάθε δικτύου. Κατά συνέπεια ο υποχώρος ο οποίος θα μας δώσει το μικρότερο σφάλμα υποδεικνύει την κλάση στην οποία το διάνυσμα που θέλουμε να αναγνωρίσουμε είναι πλησιέστερα. Χρησιμοποιώντας επιπλέον και κάποιο κατώφλι ανώτερου σφάλματος, μπορούμε να απορρίψουμε πρότυπα τα οποία δεν ανήκουν σε καμία από τις γνωστές κλάσεις.

Με τη μέθοδο αυτή η προσθήκη μιας νέας κλάσης είναι πολύ πιο απλή, αφού η μόνη απαίτηση είναι η προσθήκη ενός νέου δικτύου στο σύστημα και η εκπαίδευση μόνο του δικτύου αυτού με τα διανύσματα-πρότυπα της νέας κλάσης. Επίσης η ταξινόμηση καθεαυτή είναι μια διαδικασία με ελάχιστες υπολογιστικές απαιτήσεις, αφού το μόνο που απαιτείται είναι η χρήση του αντίστροφου πίνακα των βαρών των συνδέσμων του δικτύου και ο υπολογισμός του μέσου τετραγωνικού σφάλματος ανακατασκευής.

3.9. Άλλες εφαρμογές της ανάλυσης κυρίων συνιστωσών.

Η μέθοδος ανάλυσης κυρίων συνιστωσών παράγει τον βέλτιστο γραμμικό μετασχηματισμό απεικόνισης δεδομένων από τις n διαστάσεις στις m διαστάσεις, όπου $m < n$. Η ανακατασκευή που επιτυγχάνεται από την ανάλυση κυρίων συνιστωσών είναι βέλτιστη υπό την έννοια του ελαχίστου τετραγωνικού σφάλματος. Έτσι τα αρχικά n -διάστατα δεδομένα μπορούν να ανακατασκευαστούν από τη m -διάστατη απεικόνισή τους με το ελάχιστο μέσο τετραγωνικό σφάλμα (mean square error).

3.9.1. Συμπίεση εικόνας

Το γεγονός αυτό οδηγεί σε εφαρμογές της μεθόδου σε συμπίεση δεδομένων. Η ανακατασκευή των εικόνων η οποία επιτυγχάνεται με τη μέθοδο της ανάλυσης κυρίων συνιστωσών είναι βέλτιστη υπό την έννοια του ελαχίστου τετραγώνου.



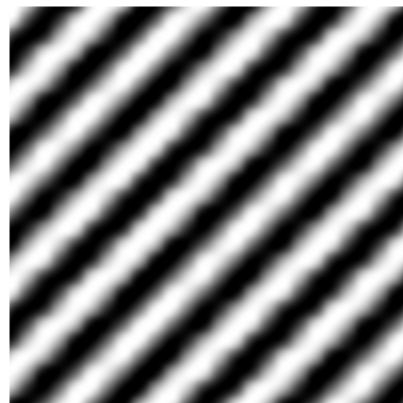
Μοτίβο Α



Μοτίβο Β



Μοτίβο Α



Μοτίβο Γ

Εικόνα 8: Ενώ το μέσο τετραγωνικό σφάλμα μεταξύ του Μοτίβου Α και του μοτίβου Β ισούται με το 100% της εικόνας, το μέσο τετραγωνικό σφάλμα μεταξύ του μοτίβου Α και Γ φτάνει μόλις το 50%

Η έννοια του ελαχίστου τετραγώνου μπορεί να γίνει πολύ εύκολα κατανοητή αν

παρατηρήσουμε την εικόνα 6. Το μοτίβο B, αν και δε φαίνεται με την πρώτη ματιά, είναι το αρνητικό του μοτίβου A και σαν αρνητικό, συγκρινόμενο με το A, το μέσο τετραγωνικό σφάλμα που προκύπτει είναι το 100% της εικόνας. Το μοτίβο Γ που αποτελεί μία περιστραμμένη εκδοχή του μοτίβου A, συγκρινόμενο με το μοτίβο A, το μέσο τετραγωνικό σφάλμα είναι μόλις το 50% της εικόνας. Το παραπάνω μοτίβο είναι ειδικά επιλεγμένο για να δείξουμε ότι το μέσο τετραγωνικό σφάλμα παρόλο που γενικά μπορεί να χρησιμοποιηθεί σαν ένα πολύ καλό μέτρο σύγκρισης δύο εικόνων, μπορεί σε ορισμένες περιπτώσεις να μην είναι και το καλύτερο.

Η διαδικασία συμπίεσης εικόνας έχει ως εξής:

- Αρχικά κωδικοποιούμε την εικόνα όπως έχουμε ήδη περιγράψει, αντιστοιχώντας τη φωτεινότητα του κάθε pixel σε έναν αριθμό από το 0 (μαύρο) ως το 255 (λευκό) και στη συνέχεια την τεμαχίζουμε σε διαδοχικά τμήματα μεγέθους 8x8 pixels.
- Στη συνέχεια μετατρέπουμε το κάθε τμήμα B_k σε ένα διάνυσμα $x(k)$ τοποθετώντας τη μία στήλη του B_k κάτω από την άλλη.
- Θεωρώντας τη φωτεινότητα του κάθε pixel ως μία τυχαία μεταβλητή εφαρμόζουμε τη μέθοδο της ανάλυσης κυρίων συνιστωσών στα διανύσματα $x(k)$ που προκύπτουν με την παραπάνω μέθοδο. Αφού τα διανύσματα είναι 64 διαστάσεων, υπάρχουν συνολικά 64 κύριες συνιστώσες που αντιστοιχούν στα 64 κύρια ιδιοδιανύσματα $e_{1,\dots}, e_{64}$.
- Προβάλουμε τα διανύσματα $\text{bold}\{x\}(k)$ σε μικρότερο χώρο m-διαστάσεων (όπου $m < 64$) χρησιμοποιώντας τα πρώτα m ιδιοδιανύσματα.

$$y(k) = [y_1(k), \dots, y_m(k)]^T = [e_{1,\dots}, e_m]^T x(k) \quad (62)$$

- Ανακατασκευάζουμε τα διανύσματα χρησιμοποιώντας την προβολή τους στο χώρο των m-διαστάσεων

$$\hat{x}(k) = [e_{1,\dots}, e_m] y(k) \quad (63)$$

- Από τα διανύσματα ανακατασκευάζουμε το κάθε block και τελικά

ανασυνθέτουμε την εικόνα.

Η όλη ιδέα βασίζεται πάνω στο γεγονός ότι η απώλεια των μικρότερων κυρίων συνιστωσών αντιστοιχεί και στο μικρότερο δυνατό σφάλμα ανακατασκευής των από τις m στις n διαστάσεις. Η συμπίεση έχει γενικά ικανοποιητικά αποτελέσματα διότι η πληροφορία βρίσκεται συνήθως συσσωρευμένη μέσα σε λίγες κύριες συνιστώσες. Στις εικόνες συνήθως τα γειτονικά pixels έχουν παρόμοιες τιμές και δεν έχουμε απότομες αλλαγές τις φωτεινότητες, οπότε στις περιοχές αυτές αντιστοιχούν οι πρώτες (μεγαλύτερες) κύριες συνιστώσες. Εκεί όπου υπάρχουν περισσότερο έντονες αλλαγές της φωτεινότητας είναι οι ακμές και μια και είναι λιγότερες στην εικόνα, αντιστοιχούν και στις υπόλοιπες (μικρότερες) κύριες συνιστώσες.

Κατά τη συμπίεση, εξαλείφοντας τις μικρότερες κύριες συνιστώσες, χάνεται τμήμα της πληροφορίας της εικόνας που αφορά τις ακμές, με αποτέλεσμα η εικόνα να είναι μεν αντιληπτή, αλλά να είναι θολή με έντονη την έλλειψη των λεπτομερειών στις ακμές. Αυτό γίνεται εμφανές στο παράδειγμα 1 της εικόνας 9, όπου έχουμε χρησιμοποιήσει μόνο $m=4$ κύριες συνιστώσες (συμπίεση 4/64). Η εικόνα γίνεται όλο και πιο ευκρινής όσο αυξάνουμε τον αριθμό των κυρίων συνιστωσών που χρησιμοποιούμε για την ανακατασκευή. Στα υπόλοιπα παραδείγματα που φαίνονται στην εικόνα έχουμε χρησιμοποιήσει 8 και 12 κύριες συνιστώσες. Παρατηρούμε ότι ακόμα και με μόλις 12 κύριες συνιστώσες (συμπίεση 12/64 ή 18,75%), η διαφορά της ανακατασκευασμένης εικόνας από την αρχική είναι ελάχιστα αντιληπτή.



1.



2.



3.



4.

Εικόνα 9: Παράδειγμα συμπίεσης εικόνας με τη μέθοδο της ανάλυσης κυρίων συνιστωσών. Η εικόνα τεμαχίζεται σε τμήματα των 8x8 pixels, μετατρέπονται σε διανύσματα μεγέθους 64 και εφαρμόζεται PCA. Τα τμήματα αυτά ανακατασκευάζονται με καλύτερη ποιότητα, όσο αυξάνεται το πλήθος των κυρίων συνιστωσών που χρησιμοποιούμε. (1) 4 PCs, (2) 8 PCs, (3) 12 PCs, (4) Η αρχική ασυμπίεστη εικόνα

3.9.2. Ανίχνευση επιθέσεων σε δίκτυα υπολογιστών

Τα δεδομένα τα οποία συλλέγονται για ανάλυση με σκοπό την ανίχνευση κάποιας δικτυακής επίθεσης ή εισβολής είναι συνήθως πολυδιάστατα με αποτέλεσμα να είναι δύσκολο να αναλυθούν και να οπτικοποιηθούν. Η ανάλυση κυρίων συνιστωσών χρησιμοποιείται για τη μείωση των διαστάσεων των δεδομένων που συλλέγονται ώστε να απλοποιηθεί η διαδικασία της ανάλυσης της δικτυακής κίνησης. Χρησιμοποιώντας την ανάλυση κυρίων συνιστωσών σε δεδομένα

γνωστών επιθέσεων, όπως Denial-of-Service (DoS) και διερευνητικές επιθέσεις (Network Probe attacks).

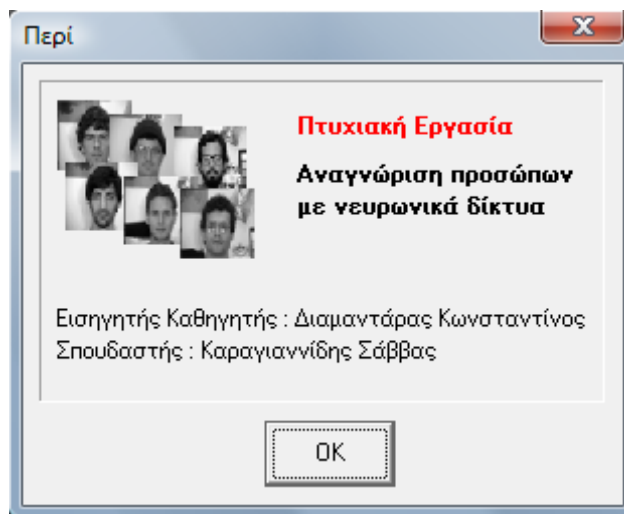
Τα δεδομένα της κίνησης του δικτύου μεταφράζονται σε χαρακτηριστικά διανύσματα, οι διαστάσεις των οποίων είναι τα ακόλουθα στοιχεία της επικεφαλίδας (header) των πακέτων:

- SIPx : Τα τέσσερα ψηφία ($x=[1-4]$) της διεύθυνσης IP της πηγής (Source IP).
- SPort: Ο αριθμός της πόρτας της πηγής (source port)
- DIPx : Τα τέσσερα ψηφία ($x=[1-4]$) της διεύθυνσης IP του προορισμού (Destination IP).
- SPort: Ο αριθμός της πόρτας προορισμού (destination port)
- Prot: Ο τύπος του πρωτοκόλλου επικοινωνίας (TCP, UDP, ICMP)
- PLen: Το μέγεθος του πακέτου σε bytes (Packet length).

Χρησιμοποιώντας σαν δεδομένα εισόδου τα δεδομένα από επιθέσεις της DARPA (DARPA 1998 intrusion detection data set), εκπαιδεύεται το σύστημα ώστε να αναγνωρίζει αποτελεσματικά προκαθορισμένους τύπους επιθέσεων. (Labib and Vemuri 2006)

4. Υλοποίηση του συστήματος αναγνώρισης προσώπων

Για την υλοποίηση του συστήματος αναγνώρισης των προσώπων αναπτύξαμε μια ευέλικτη εφαρμογή η οποία μας επιτρέπει να πειραματιστούμε με δεδομένα και διάφορες δομές δικτύων ώστε να μπορέσουμε να καταμετρήσουμε και την αποτελεσματικότητα ενός τέτοιου συστήματος.



Εικόνα 10: Η εφαρμογή "Αναγνώριση προσώπων με νευρωνικά δίκτυα" αναπτύχθηκε για τις ανάγκες υλοποίησης των αλγορίθμων που παρουσιάζονται στην παρούσα πτυχιακή εργασία

4.1. Τεχνικά στοιχεία υλοποίησης της εφαρμογής

Για την ανάπτυξη της εφαρμογής χρησιμοποιήσαμε τη γλωσσά προγραμματισμού Delphi. Ο λόγος που επιλέξαμε τη γλώσσα αυτή είναι ότι είναι μια γλωσσά γρήγορη, εύκολη στον προγραμματισμό και την αποσφαλμάτωση και προσφέρει ένα ολοκληρωμένο γραφικό περιβάλλον ανάπτυξης. Επίσης με την τεχνολογία των components που διαθέτει και την πληθώρα του ήδη υπάρχοντα κώδικα ο οποίος μπορεί να ξαναχρησιμοποιηθεί, μας δίνεται η δυνατότητα επέκτασης των δυνατοτήτων της εφαρμογής. Καθώς τα δεδομένα τα οποία διαχειρίζεται η

εφαρμογή είναι εικόνες, είναι σημαντική η δυνατότητα που δίνεται για τη διαχείριση πολλών ειδών αρχείων εικόνων, καθώς και η δυνατότητα λήψης εικόνας από διάφορες εξωτερικές πηγές.

Τέλος ο αντικειμενοστραφής προσανατολισμός της γλώσσας προγραμματισμού μας βολεύει στην περιγραφή των εννοιών που είδαμε μέχρι τώρα με τη μορφή αντικειμένων καθώς και να σχεδιάσουμε ένα interface για την εύκολη διαχείριση των ιδιοτήτων τους.

4.2. Περιγραφή των εννοιών/αντικειμένων της εφαρμογής

4.2.1. Εικόνες

Το βασικότερο στοιχείο με το οποίο ασχολείται η εφαρμογή από τα διανύσματα εκπαίδευσης μέχρι τα διανύσματα εξόδου και τις ιδιοτιμές αποτελούν εικόνες. Οι εικόνες αποτελούνται από ένα δισδιάστατο πίνακα εικονοστοιχείων (pixels). Η εικόνα μπορεί να είναι είτε έγχρωμη είτε να αποτελείται μόνο από αποχρώσεις του γκρι (grayscale). Στις περισσότερες των περιπτώσεων οι εικόνες με τις οποίες θα ασχοληθούμε είναι grayscale. Στην περίπτωση αυτή κάθε εικονοστοιχείο αντιστοιχεί σε θέση μνήμης ενός byte που αντιστοιχεί στη φωτεινότητα του εικονοστοιχείου. Στην περίπτωση που η εικόνα είναι έγχρωμη, το κάθε εικονοστοιχείο καταλαμβάνει 24 bits ή αλλιώς 3 bytes μνήμης (1 byte για κάθε βασικό χρώμα, κόκκινο πράσινο μπλε). Οι έγχρωμες εικόνες αντιμετωπίζονται γενικά στην εφαρμογή μας σαν 3 διαφορετικές grayscale εικόνες κάθε μια από τις οποίες αποτελείται μόνο από την ένταση του καθενός από τα βασικά χρώματα (RGB, κόκκινο πράσινο και μπλε).

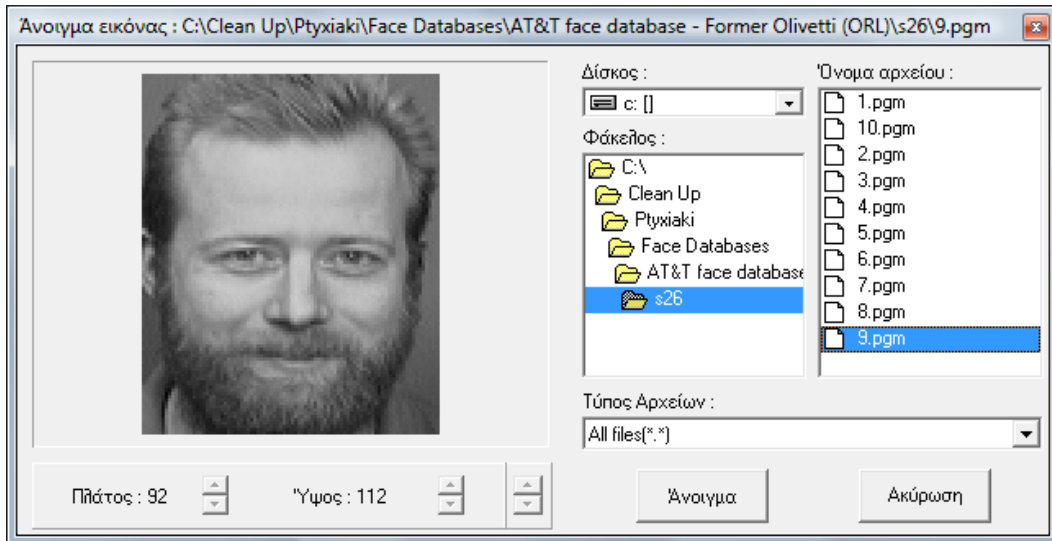
Οι εικόνες μπορούν να φορτωθούν απ' ευθείας από το δίσκο, εφόσον είναι αποθηκευμένες σε κάποια από τα υποστηριζόμενα format εικόνας. Η λίστα των format εικόνων που υποστηρίζονται είναι αρκετά ευρεία ώστε να καλύπτουν τις απαιτήσεις των διαφόρων διανομών από βάσεις δεδομένων εικόνων προσώπων που υπάρχουν σήμερα.

Τα αρχεία εικόνας που υποστηρίζονται είναι τα παρακάτω:

Τύπος Αρχείων	Κατάληξη Αρχείων
Windows bitmaps	(* .bmp; * .rle; * .dib)
Windows icons	(* .ico)
Windows metafiles	(* .wmf; * .emf)
JPG images	(* .jfif; * .jpg; * .jpe; * .jpeg)
Truevision images	(* .win; * .vst; * .vda; * .tga; * .icb)
Tagged image file format images	(* .tiff; * .tif; * .fax)
Encapsulated Postscript images	(* .eps)
ZSoft Paintbrush images	(* .pcx; * .pcc; * .scr)
Alias/Wavefront images	(* .rpf; * .rla)
SGL images	(* .sgi; * .rgba; * .rgb; * .bw)
Photoshop images	(* .psd; * .pdd)
Portable map images	(* .ppm; * .pgm; * .pbm)
Autodesk images	(* .cel; * .pic)
Kodak Photo-CD images	(* .pcd)
Dr. Halo images	(* .cut)
Paintshop Pro images	(* .psp)
Portable network graphic images	(* .png)

Πίνακας 2: Λίστα των τύπων αρχείων εικόνας που μπορούν να διαβαστούν από την εφαρμογή

Ακόμα και αν ένα format δεν υποστηρίζεται άμεσα, δίνεται η δυνατότητα εισαγωγής της εικόνας από κάποια άλλη εφαρμογή μέσω του clipboard, απλά αντιγράφοντας την εικόνα και επιλέγοντας από το βασικό μενού “Αρχείο -> Εικόνα -> Επικόλληση ως νέα”.



Εικόνα 11: Ο διάλογος ανοίγματος μίας εικόνας απ'ευθείας από το δίσκο

Από τη στιγμή που κάποια εικόνα έχει εισαχθεί στην εφαρμογή, παρέχονται και κάποιες πολύ βασικές λειτουργίες επεξεργασίας της, όπως η προσθήκη θορύβου, αυτόματη εξισορρόπηση της φωτεινότητας της εικόνας καθώς και η παράγωγη του αρνητικού της εικόνας (Μενού “Εικόνα -> Εξίσωση ιστογράμματος, Αρνητικό, Προσθήκη θορύβου...”).

4.2.2. Λίστες εικόνων

Καθώς τα δεδομένα μας σχεδόν πάντα είναι σύνολα από εικόνες, δημιουργήθηκε η ανάγκη να οργανωθούν οι εικόνες σε λίστες για την εύκολη μαζική διαχείριση και επεξεργασία τους.

Οι λίστες έχουν εμπλουτιστεί με μεταβλητές οι οποίες χαρακτηρίζουν την κάθε εικόνα της λίστας. Οι μεταβλητές αυτές είναι απαραίτητες για την επιλογή ενός υποσυνόλου των εικόνων της λίστας.

Π.χ. Έστω ότι έχουμε μια λίστα εικόνων η οποία περιέχει μια βάση προσώπων. Σ' αυτή τη λίστα μπορούμε να ορίσουμε μεταβλητές της μορφής person, glasses, beard, gender κτλ, οι οποίες θα χαρακτηρίζουν τις εικόνες παίρνοντας τις κατάλληλες τιμές. Έτσι η μεταβλητή person για όλες τις εικόνες του ίδιου ατόμου μπορεί να έχει μια μοναδική τιμή όπως το όνομα του ατόμου ή κάποιον αριθμό ο οποίος το χαρακτηρίζει. Αντίστοιχα, η μεταβλητή gender παίρνει την τιμή male ή

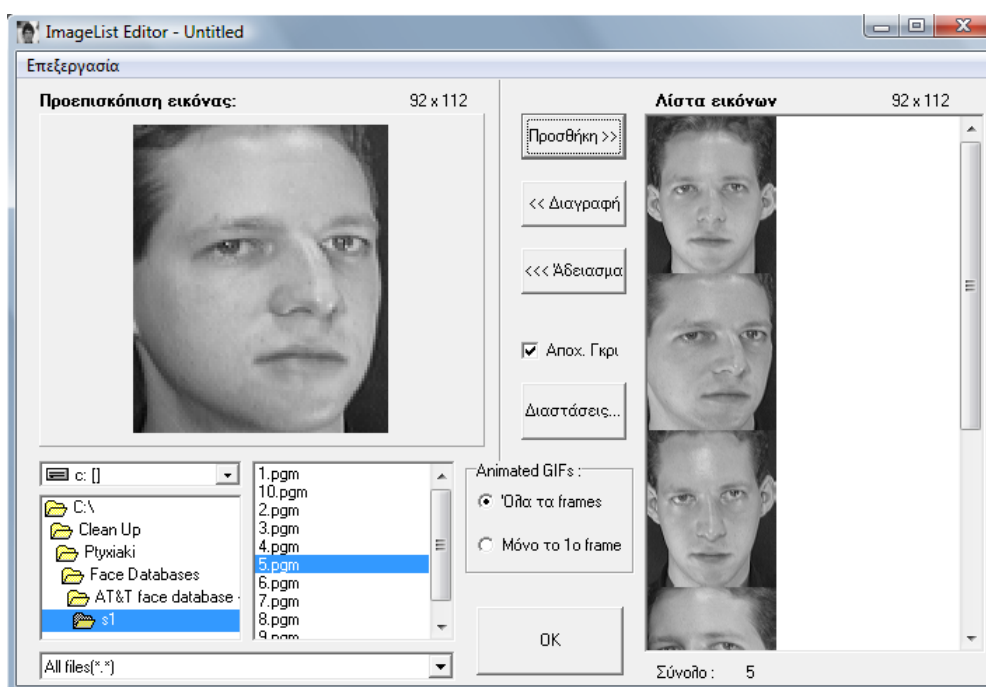
female ανάλογα με το φύλο του ατόμου που απεικονίζεται στην εικόνα.

Αυτή η δυνατότητα μας δίνει την ευελιξία να χρησιμοποιούμε την ίδια λίστα εικόνων για διαφορετικούς σκοπούς. Π.χ. Αν θέλουμε να δημιουργήσουμε ένα δίκτυο το οποίο αναγνωρίζει κάποιο συγκεκριμένο άτομο, αρκεί να το εκπαιδεύσουμε με εκείνες τις εικόνες της λίστας όπου η συνθήκη person = x αληθεύει, όπου x είναι το χαρακτηριστικό όνομα ή αριθμός του ατόμου. Αντίστοιχα, αν θέλουμε να εκπαιδεύσουμε ένα δίκτυο το οποίο αναγνωρίζει το φύλο του ατόμου που απεικονίζεται σε μια εικόνα, αρκεί να εκπαιδεύσουμε 2 δίκτυα, το ένα με τις εικόνες όπου gender = male και το άλλο με τις εικόνες όπου gender = female.

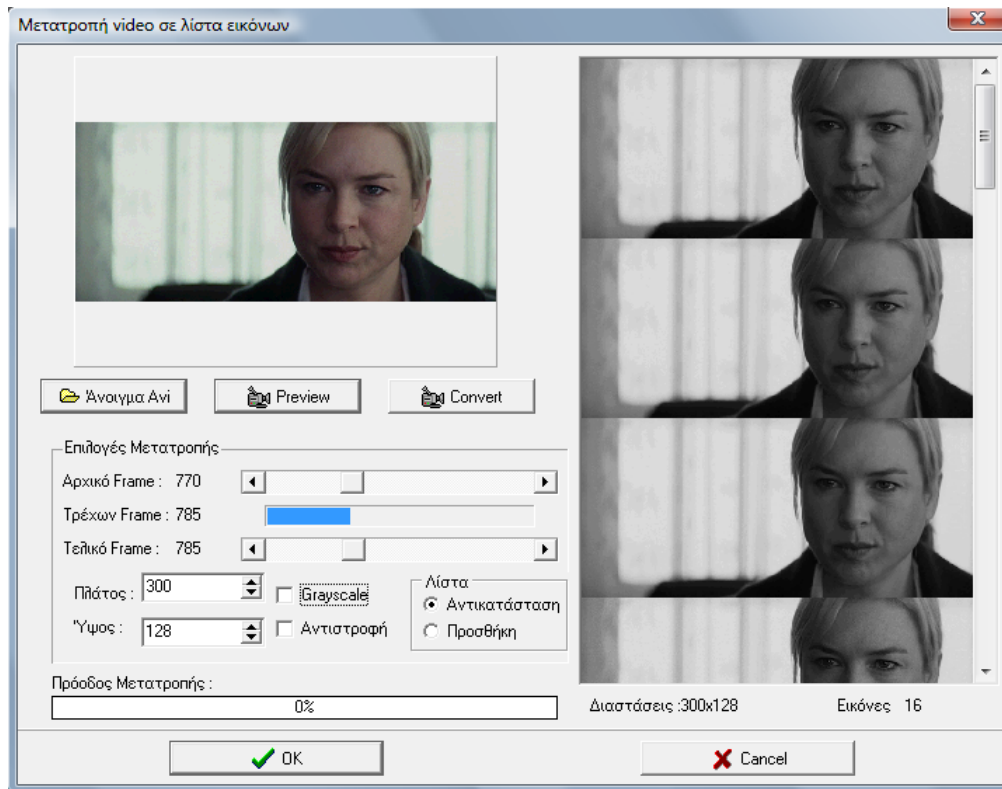
Η δημιουργία της λίστας των εικόνων γίνεται κυρίως με τρεις τρόπους:

α) Από επιμέρους εικόνες, δημιουργούμε τη λίστα και στη συνέχεια προσθέτουμε στη λίστα εικόνες που έχουμε αποθηκευμένες στο δίσκο. (Εικόνα 12)

β) Εισάγοντας εικόνες – frames από κάποιο αρχείο video. Η μορφή του αρχείου video θα πρέπει να είναι avi και ο codec που χρησιμοποιήθηκε για την κωδικοποίηση πρέπει να είναι συμβατός με το video for windows και να είναι εγκατεστημένος στο σύστημα. (Εικόνα 13)



Εικόνα 12: Δημιουργία λίστας εικόνων από μεμονωμένα αρχεία



Εικόνα 13: Ο διάλογος μετατροπής μίας ακολουθίας βίντεο σε λίστα εικόνων για την εισαγωγή του στην εφαρμογή

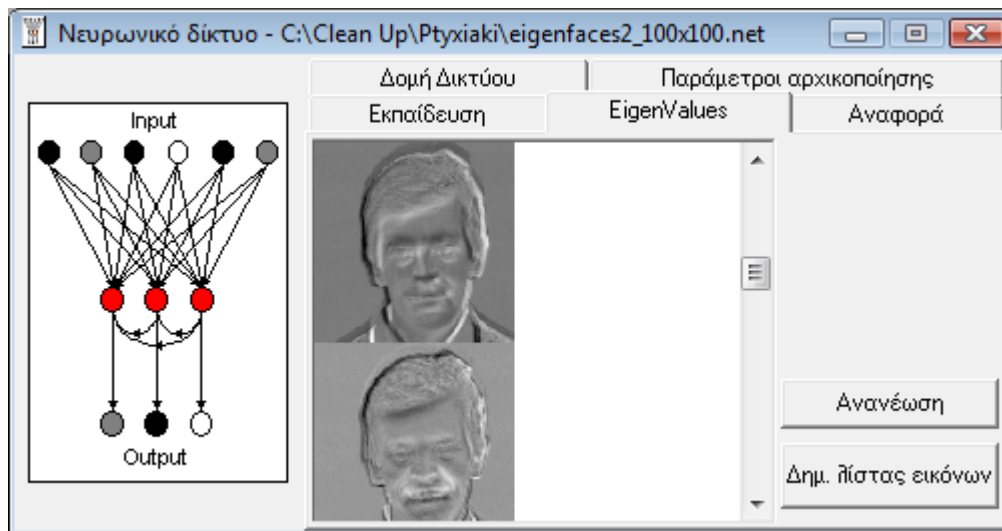
γ) Εισάγοντας εικόνες απ' ευθείας από κάποια πηγή video συνδεδεμένη στον υπολογιστή (webcam, κάρτα TV tuner, κτλ)

Οι επιλογές αυτές είναι διαθέσιμες από το μενού “Αρχείο -> Λίστα εικόνων -> Νέα..., Εισαγωγή από αρχείο βίντεο..., Εγγραφή...”.

4.2.3. Νευρωνικό δίκτυο

Αυτό είναι το αντικείμενο που αντιπροσωπεύει το ίδιο το νευρωνικό δίκτυο APEX. Κατά τη δημιουργία ενός νευρωνικού δικτύου απαιτείται η καταχώρηση των βασικότερων στοιχείων του δικτύου, όπως οι διαστάσεις των νευρώνων εισόδου και ο αριθμός των νευρώνων εξόδου. Λόγω της μορφής των δεδομένων εισόδου αλλά και τον τρόπο με τον οποίο εκπαιδεύεται το δίκτυο, προτιμούμε να δίνουμε τον αριθμό των νευρώνων εισόδου σε 2 διαστάσεις. Αυτό δεν επηρεάζει καθόλου τον τρόπο εκπαίδευσης του δικτύου. Ο αριθμός των νευρώνων εισόδου είναι

κανονικά το γινόμενο των 2 διαστάσεων. Αυτό που επηρεάζεται είναι ο τρόπος με τον οποίο λαμβάνονται τα δεδομένα εισόδου από την εικόνα, σε περίπτωση που οι διαστάσεις του δικτύου δεν ταυτίζονται με τις διαστάσεις της εικόνας - προτύπου εκπαίδευσης ή της εικόνας προς ταξινόμηση. Στην περίπτωση όπου οι διαστάσεις των εικόνων δεν αντιστοιχούν στις διαστάσεις του δικτύου, η εικόνα τεμαχίζεται σε μικρότερα τμήματα στις διαστάσεις του δικτύου ξεκινώντας από την επάνω δεξιά γωνιά. η μέθοδος αυτή έχει σαν εφαρμογή τη συμπίεση των εικόνων. Ένα τυπικό παράδειγμα αυτής της περίπτωσης είναι να έχουμε σαν πρότυπο εκπαίδευσης μία η περισσότερες εικόνες μεγάλων διαστάσεων (1024 x 768 ή και μεγαλύτερες) και ένα σχετικά μικρό δίκτυο με αρκετά μικρό αριθμό εξόδων (8 x 8 συνήθως με 8 νευρώνες εξόδου). Ο αριθμός των νευρώνων εξόδου (αριθμός κυρίων συνιστωσών που εξάγονται) επηρεάζει την ποιότητα της συμπίεσης. Όσο πιο κοντά ο αριθμός των νευρώνων εξόδου στον αριθμό των νευρώνων εισόδου, τόσο μικρότερη συμπίεση και απώλεια πληροφορίας έχουμε. Τμήματα της εικόνας που περισσεύουν στην κάτω και δεξιά πλευρά της εικόνας αγνοούνται. Εναλλακτικά, μπορεί η εικόνα να προσαρμοστεί στις διαστάσεις του δικτύου.



Εικόνα 14: Ένα νευρωνικό δίκτυο το οποίο έχει ήδη εκπαιδευτεί. Βλέπουμε τα eigenfaces τα οποία προκύπτουν από τα βάρη του κάθε νευρώνα εξόδου. Τα eigenfaces είναι ορατά στο παράθυρο του δικτύου ακόμα και κατά τη διάρκεια της εκπαίδευσης.

Άλλα στοιχεία που απαιτούνται είναι η μέγιστη και ελάχιστη τιμή των αρχικών

τυχαίων βαρών των νευρωνικών συνδέσμων, το β_0 και το γ . Γενικά το β_0 προτιμούμε να είναι μία τιμή κοντά στο 0 αλλά όχι μηδενική, και το γ μία τιμή κοντά στο 1.

4.2.4. Εκπαίδευση του νευρωνικού δικτύου

Η εκπαίδευση του νευρωνικού δικτύου γίνεται σύμφωνα με τους κανόνες των εξισώσεων (45) και (46). Οι εικόνες από τη λίστα εικόνων προς εκπαίδευση δίνονται τμηματικά ή ολόκληρες στους νευρώνες εισόδου του δικτύου. Σε κάθε επανάληψη λαμβάνονται διαδοχικά οι εικόνες ή τμήματα αυτών ανάλογα με τις διαστάσεις του δικτύου και των εικόνων και προσαρμόζονται τα βάρη του δικτύου. Για τον τερματισμό της εκπαίδευσης του δικτύου, βασιζόμαστε στο γεγονός ότι ο πίνακας των βαρών του δικτύου τείνει να γίνει ορθογώνιος, όπως έχουμε ήδη αναφέρει. Κατά συνέπεια στο τέλος της τέλει εκπαίδευσης του δικτύου θα πρέπει να ισχύει:

$$WW^T = I \quad (64)$$

Στην πραγματικότητα, τα βάρη δε φτάνουν σ' αυτή την ιδανική κατάσταση, αλλά η νόρμα της διαφοράς $I - WW^T$ δίνει ένα ικανοποιητικό μέτρο του βαθμού εκπαίδευσης του δικτύου. Όταν η παραπάνω διαφορά πλησιάζει ικανοποιητικά το 0, τότε θεωρούμε ότι το δίκτυο έχει εκπαιδευτεί ικανοποιητικά. Το μέτρο αυτό το αποκαλούμε σφάλμα βαρών.

Ένα αντίστοιχο μέτρο μπορούμε να εξάγουμε από τον πίνακα των βαρών C . Όπως γνωρίζουμε τα βάρη C κατά την εκπαίδευση έχουν την τάση να μηδενιστούν. Έτσι το άθροισμα

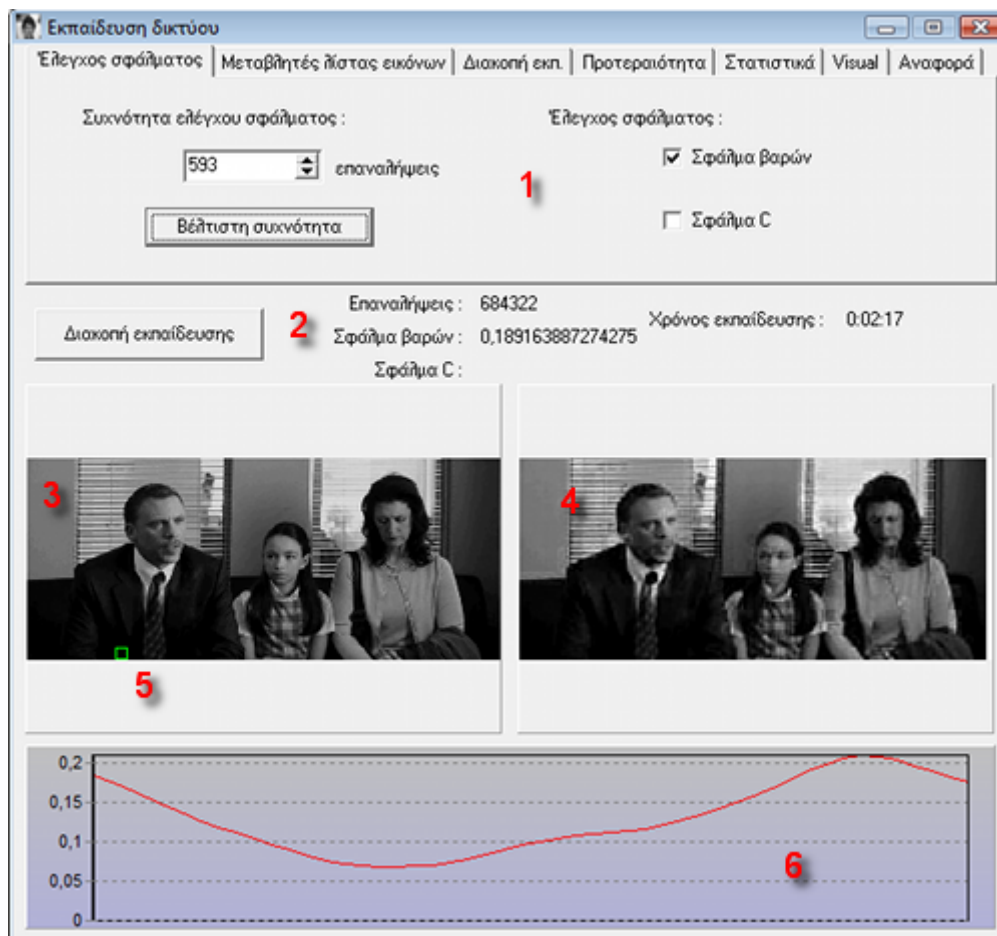
$$\sum_{i=1}^m \sum_{j=1}^{i-1} c_{ij}(k)^2 \rightarrow 0, \quad \text{όταν το } k \rightarrow \infty \quad (65)$$

Όπου m είναι ο αριθμός των νευρώνων εξόδου του δικτύου μας.

Μπορούμε να κανονικοποιήσουμε το παραπάνω άθροισμα διαιρώντας με τον αριθμό των βαρών C που είναι $n(n-1)/2$, οπότε η εξίσωση (65) γίνεται:

$$C_{error} = \frac{\sum_{i=1}^n \sum_{j=1}^{i-1} c_{ij}(k)^2}{n(n-1)/2} \quad (66)$$

Το παραπάνω άθροισμα είναι ιδανικότερο για τον έλεγχο της εκπαίδευσης του δικτύου γιατί απαιτεί λιγότερες πράξεις σε σχέση με τον πίνακα των πρόσθιων βαρών w .



Εικόνα 15: Τυπικό παράδειγμα εκπαίδευσης νευρωνικού δικτύου από λίστα εικόνων:

1) Περιοχή ρυθμίσεων των παραμέτρων της εκπαίδευσης, 2) Περιοχή προβολής της τρέχουσας κατάστασης της εκπαίδευσης (σφάλμα βαρών, χρόνος εκπαίδευσης κτλ) 3) Τρέχουσα εικόνα – δεδομένο εισόδου εκπαίδευσης 4) Ανακατασκευασμένη εικόνα εισόδου με βάση την τρέχουσα εκπαίδευση του δικτύου, 5) Τρέχων τμήμα της εικόνας το οποίο αντιστοιχεί στην είσοδο του δικτύου. 6) Γραφική παράσταση της μεταβολής του σφάλματος των βαρών κατά τη διάρκεια της εκπαίδευσης του δικτύου.

4.2.5. Σύστημα αναγνώρισης - ταξινόμησης προτύπων.

Το σύστημα αναγνώρισης αποτελείται από ένα σύνολο αντικειμένων νευρωνικών δικτύων τα οποία μπορούν να μοιράζονται μια κοινή λίστα εικόνων για την εκπαίδευση τους. Παρέχει την ευκολία μαζικών αλλαγών σε όλα τα δίκτυα που ανήκουν στο σύνολο, καθώς και τη δυνατότητα αυτόματης σειριακής εκπαίδευσης όλων των δικτύων. Η βασική όμως λειτουργία του συστήματος αναγνώρισης - ταξινόμησης είναι όπως υποδηλώνει και το όνομα του η ταξινόμηση προτύπων. Η ταξινόμηση όπως περιγράψαμε σε προηγούμενο κεφάλαιο γίνεται με βάση το ελάχιστο σφάλμα ανακατασκευής του προτύπου από όλα τα νευρωνικά δίκτυα του συνόλου. Το ελάχιστο σφάλμα αντιπροσωπεύει την υποκλάση η οποία είναι πλησιέστερα στο πρότυπο που θέλουμε να αναγνωρίσουμε. Με την παραπάνω μέθοδο προς το παρόν δεν υπάρχει η δυνατότητα να βγάλει το συμπέρασμα ότι το άγνωστο πρότυπο δεν ανήκει σε καμία από τις γνωστές κλάσεις. Θα μπορούσε όμως πολύ εύκολα να υλοποιηθεί ορίζοντας κάποιο κατώφλι (threshold) σφάλματος, το οποίο όταν ξεπερνιέται κατά την ανακατασκευή του προτύπου, τότε το πρότυπο δεν ανήκει σε κάποια από τις γνωστές κλάσεις.

4.2.6. Γεννήτρια αναφορών / εξαγωγή δεδομένων

Η γεννήτρια αναφορών χρησιμοποιείται για την εξαγωγή ωμών δεδομένων από την εφαρμογή για πιθανή χρήση ή επεξεργασία των δεδομένων σε κάποια άλλη εφαρμογή. Τα δεδομένα τα οποία μπορούν να εξαχθούν από την εφαρμογή είναι το σύνολο των βαρών των συνδέσεων των νευρώνων ενός νευρωνικού δικτύου.

4.3. Ροή λειτουργίας της εφαρμογής

Η τυπική λειτουργία της εφαρμογής μπορεί να περιγραφεί σε 5 τυπικά βήματα:

4.3.1. Εισαγωγή προτύπων εκπαίδευσης.

Κατά τη φάση της εισαγωγής των προτύπων εκπαίδευσης δημιουργούμε μια η περισσότερες λίστες εικόνων στις οποίες εισάγουμε τις εικόνες οι οποίες θα χρησιμοποιηθούν σαν πρότυπα για την εκπαίδευση του/των δικτύων μας. Για την εισαγωγή των δεδομένων επιλέγουμε είτε εικόνες είτε μια πηγή video. Σε κάθε

περίπτωση καλό είναι η ανάλυση των δεδομένων εισόδου να είναι όσο το δυνατόν χαμηλή, χωρίς όμως να χάνεται μεγάλο μέρος της πληροφορίας των εικόνων, στοχεύοντας στην εξισορρόπηση μεταξύ πλήθους δεδομένων και υπολογιστικού κόστους.

Αφού έχουμε δημιουργήσει τη λίστα πρέπει να δημιουργήσουμε μεταβλητές οι οποίες θα χαρακτηρίζουν τα διακριτά πρότυπα εισόδου. Αυτό θα μας βοηθήσει αργότερα στην επιλογή των εικόνων που θα πρέπει να χρησιμοποιηθούν

4.3.2. Δημιουργία δικτύου/δικτύων.

Αφού έχουμε οργανώσει τα δεδομένα εκπαίδευσης σε λίστες, ταξινομημένες κατάλληλα με τη βοήθεια μεταβλητών, μπορούμε να προχωρήσουμε στη δημιουργία του/των δικτύων.

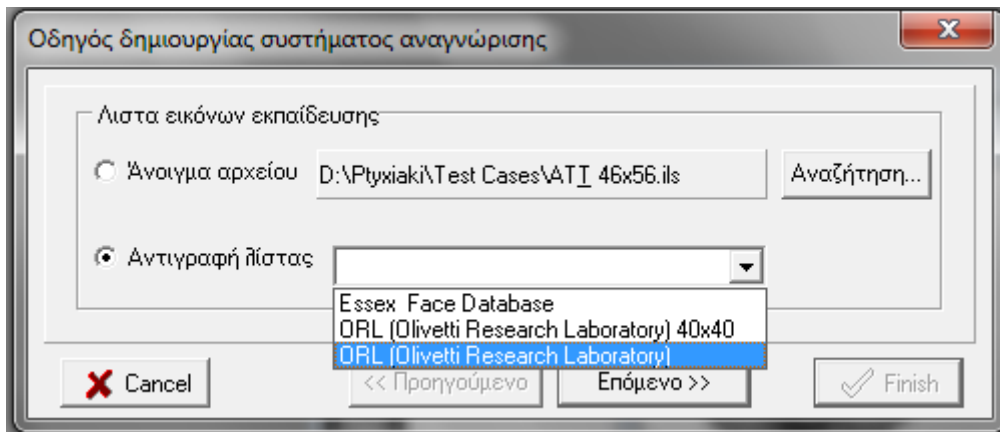
Το κατάλληλο μέγεθος του δικτύου για την εφαρμογή της αναγνώρισης είναι οι ακριβείς διαστάσεις των δεδομένων εκπαίδευσης. Διαφορετικές διαστάσεις θα είχαν νόημα μόνο για εφαρμογές συμπίεσης των δεδομένων.

4.3.3. Οδηγός (wizard) δημιουργίας συστήματος αναγνώρισης.

Ένα πολύ δυνατό χαρακτηριστικό της εφαρμογής είναι η ύπαρξη ενός οδηγού (wizard), ο οποίος μπορεί να αυτοματοποιήσει τη δημιουργία ενός συστήματος αναγνώρισης από μία λίστα εικόνων.

Ο οδηγός ενεργοποιείται από το menu “Αρχείο → Σύστημα Αναγνώρισης → Οδηγός Δημιουργίας”.

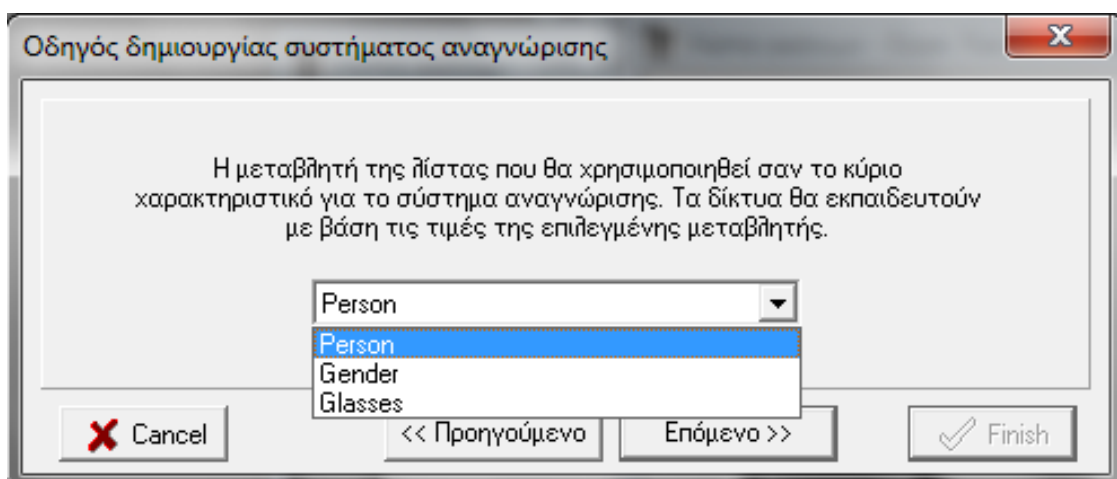
1ο βήμα: Στο πρώτο βήμα ζητείται η λίστα των εικόνων η οποία θα χρησιμοποιηθεί για την εκπαίδευση ή σαν βάση για τη δημιουργία του συστήματος αναγνώρισης. Εδώ υπάρχουν οι δυνατότητες να χρησιμοποιηθεί κάποια λίστα εικόνων η οποία είναι αποθηκευμένη στο σκληρό δίσκο ή να δημιουργηθεί ένα αντίγραφο από μία λίστα την οποία έχουμε ήδη ανοιχτή για επεξεργασία στην επιφάνεια εργασίας της εφαρμογής.



Εικόνα 16: Το πρώτο βήμα του οδηγού δημιουργίας του συστήματος αναγνώρισης, όπου επιλέγουμε τη λίστα εικόνων

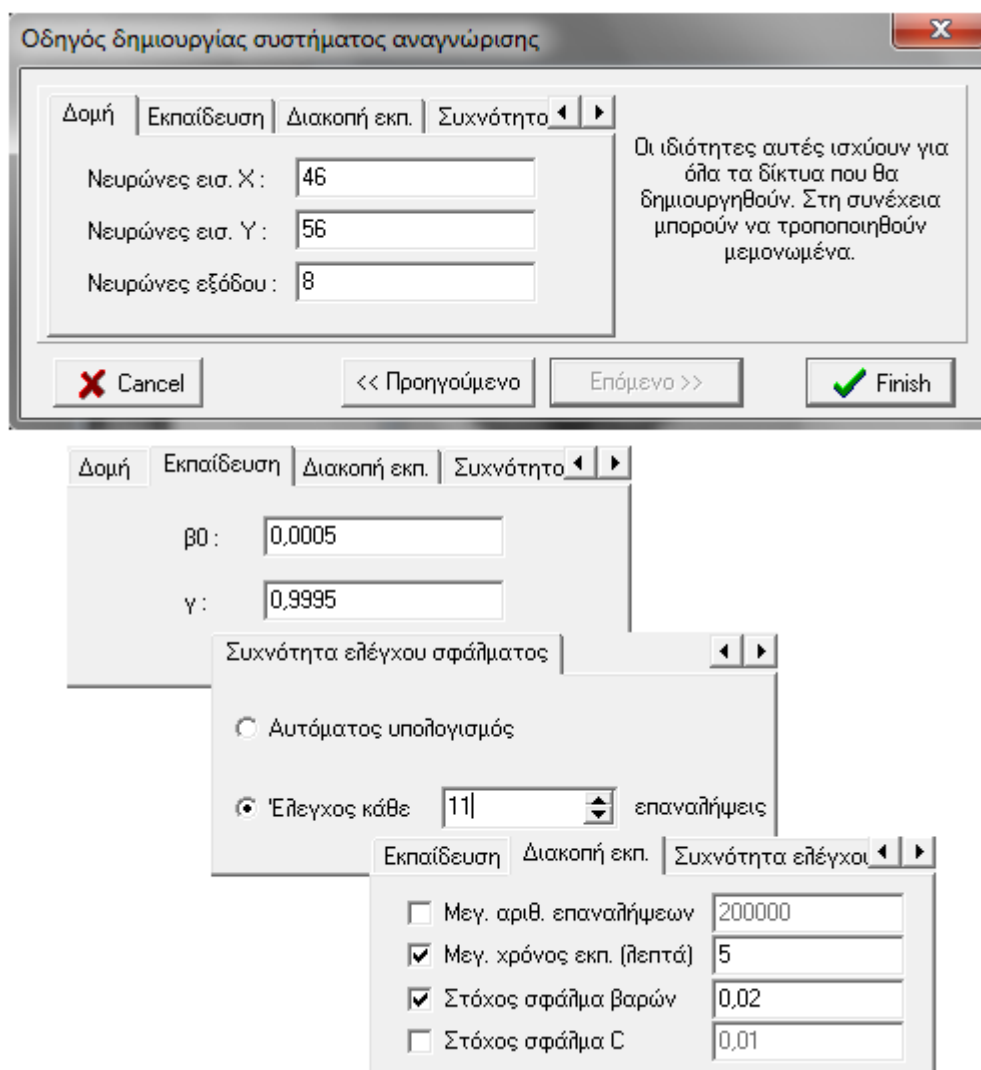
2ο βήμα: Από τη στιγμή που έχουμε επιλέξει τη λίστα των εικόνων, στο δεύτερο βήμα μας δίνεται η ευκαιρία να κάνουμε κάποιες τελευταίες τροποποιήσεις στη λίστα των εικόνων πριν προχωρήσουμε στο επόμενο βήμα.

3ο βήμα: Το 3ο βήμα είναι το πλέον καθοριστικό για το είδος της λειτουργίας του συστήματος αναγνώρισης. Εδώ μας δίνεται η δυνατότητα να επιλέξουμε από τη λίστα όλων των μεταβλητών που έχουν οριστεί στη λίστα εικόνων, τη μεταβλητή η οποία αντιστοιχεί στο χαρακτηριστικό των εικόνων που καθορίζει τις κλάσεις. Το σύνολο των τιμών της μεταβλητής αυτής καθορίζει τον αριθμό των δικτύων από τα οποία θα αποτελείται το σύστημα αναγνώρισης το οποίο θα δημιουργηθεί.



Εικόνα 17: Επιλογή της μεταβλητής η οποία θα καθορίσει τη δομή του συστήματος αναγνώρισης

4ο βήμα: Στο 4ο βήμα καθορίζουμε τις βασικότερες ρυθμίσεις που αφορούν τη δομή (διαστάσεις των νευρώνων εισόδου, πλήθος νευρώνων εξόδου) και την εκπαίδευση (β_0 , γ , συνθήκες διακοπής της εκπαίδευσης, συχνότητα ελέγχου σφάλματος) των επιμέρους δικτύων. Οι τιμές που θα δοθούν στο βήμα αυτό είναι καθοριστικές για την αποτελεσματικότητα και την ταχύτητα της εκπαίδευσης των δικτύων.

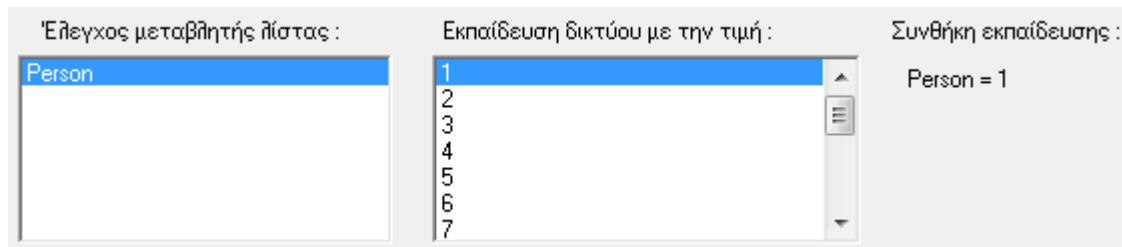


Εικόνα 18: Το σύνολο των παραμέτρων που καθορίζονται στο τελευταίο βήμα του οδηγού δημιουργίας του συστήματος αναγνώρισης. Όλα τα δίκτυα τα οποία θα δημιουργηθούν αυτόματα από τον οδηγό θα κληρονομήσουν τις ιδιότητες αυτές. Μετά τη δημιουργία των δικτύων από τον οδηγό μπορούμε να αλλάξουμε τις παραμέτρους αυτές σε κάθε δίκτυο μεμονωμένα.

4.3.4. Εκπαίδευση των δικτύων

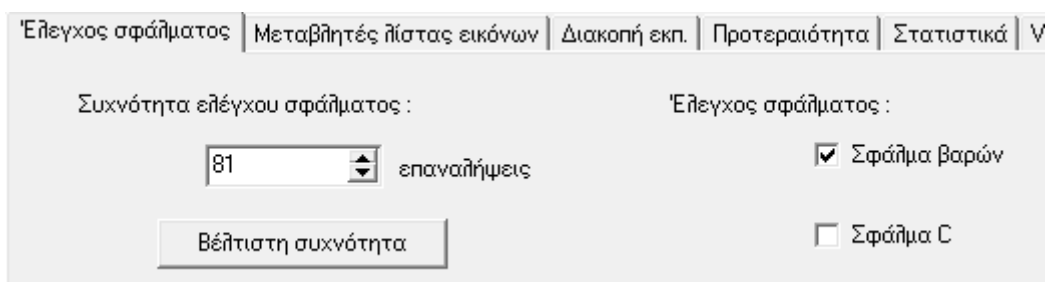
Εφόσον έχουμε δημιουργήσει το σύνολο των δικτύων μπορούμε να προχωρήσουμε στην εκπαίδευσή τους με τη λίστα εικόνων που έχουμε επιλέξει ή το υποσύνολο της λίστας με βάση κάποιο χαρακτηριστικό (το υποσύνολο των εικόνων, όπου μία μεταβλητή έχει μία συγκεκριμένη τιμή) (Εικόνα 15).

Η παράμετρος που ορίζει το χαρακτηριστικό το οποίο θα χρησιμοποιηθεί για την εκπαίδευση του δικτύου, εφόσον το δίκτυο έχει δημιουργηθεί αυτόματα από το προηγούμενο βήμα (οδηγός δημιουργίας συστήματος αναγνώρισης) είναι αυτόματα επιλεγμένη, με διαφορετική τιμή για κάθε δίκτυο.



Εικόνα 19: Εκπαίδευση νευρωνικού δικτύου APEX: Επιλογή της συνθήκης επιλογής του συνόλου των εικόνων εκπαίδευσης από τη συνολική λίστα εικόνων

Μία παράμετρος η οποία μπορεί να επιταχύνει πολύ τη διαδικασία της εκπαίδευσης είναι η Συχνότητα ελέγχου σφάλματος. Η παράμετρος αυτή καθορίζει κάθε πόσες επαναλήψεις του αλγορίθμου εκπαίδευσης θα υπολογίζεται το σφάλμα το οποίο και επιλέγουμε να ελέγχουμε για τον καθορισμό της κατάστασης της εκπαίδευσης του δικτύου. Λόγω του ότι η μεταβολή του σφάλματος σε κάθε βήμα της εκπαίδευσης είναι πολύ μικρό, μπορούμε να παραλείψουμε τον έλεγχο σε έναν αρκετά μεγάλο αριθμό επαναλήψεων, χωρίς να αλλοιώνεται το τελικό αποτέλεσμα.



Εικόνα 20: Επιλογή του τύπου σφάλματος που ελέγχουμε καθώς και της συχνότητας του ελέγχου

Οι παράμετροι που αφορούν στις συνθήκες διακοπής της εκπαίδευσης είναι εξίσου σημαντικές, ειδικά στην περίπτωση όπου θα αφήσουμε ένα σύνολο δικτύων να εκπαιδευτούν αυτόματα. Η δυνατότητα αυτή παρέχεται από το σύστημα αναγνώρισης (Κουμπί “Αυτόματη εκπαίδευση...”), εκπαιδεύοντας το ένα δίκτυο μετά το άλλο σε σειρά.



Εικόνα 21: Παράμετροι επιλογής των συνθηκών διακοπής της εκπαίδευσης του δικτύου

Οι δυνατότητες που παρέχονται για την επιλογή των συνθηκών διακοπής της εκπαίδευσης είναι τριών ειδών:

α) Χρονικός περιορισμός: Περιορίζουμε το χρόνο εκπαίδευσης του δικτύου. Η εκπαίδευση του δικτύου διακόπτεται όταν έχει παρέλθει ο συγκεκριμένος χρόνος εκπαίδευσης του δικτύου. Ο χρόνος εκπαίδευσης θα πρέπει να είναι συνεχόμενος και δεν συσσωρεύεται. Αν για παράδειγμα έχουμε ορίσει μέγιστο χρόνο εκπαίδευσης τα 5 λεπτά, αλλά στο 4ο λεπτό διακόψουμε την εκπαίδευση και στη συνέχεια επιλέξουμε να συνεχιστεί η εκπαίδευση, η εκπαίδευση δεν θα διακοπεί σε 1 λεπτό αλλά μετά από 5 λεπτά. Αν έχουμε επιλέξει χρονικό περιορισμό, κατά την εκπαίδευση προβάλλεται και ο χρόνος που απομένει.

β) Αριθμός επαναλήψεων: Περιορίζουμε τον αριθμό των συνολικών επαναλήψεων του αλγορίθμου της εκπαίδευσης. Ο περιορισμός αυτός είναι παρόμοιος με το χρονικό περιορισμό, αν και ο χρονικός περιορισμός είναι πιο συγκεκριμένος. Ο αριθμός των επαναλήψεων που απαιτούνται για την ικανοποιητική εκπαίδευση του δικτύου μπορεί να κυμαίνεται πολύ, από μερικές 100άδες έως πολλές χιλιάδες, ανάλογα με το $\beta\theta$ και το γ που έχουμε επιλέξει, τον αριθμό των δεδομένων εισόδου της εκπαίδευσης, αλλά και τη δομή του δικτύου που εκπαιδεύεται. Γενικά καλό είναι να αποφεύγεται αυτός ο περιορισμός ή να επιλέγουμε κάποια αρκετά μεγάλη τιμή ώστε να διασφαλίσουμε την ικανοποιητική εκπαίδευση του δικτύου.

γ) Σφάλμα βαρών w ή Σφάλμα C : Η ιδανική συνθήκη τερματισμού της εκπαίδευσης του δικτύου. Ορίζοντας ένα αρκετά μικρό σφάλμα – στόχο, εξασφαλίζουμε ότι η εκπαίδευση θα είναι ικανοποιητική. Το δίκτυο εκπαιδεύεται μέχρι το επιλεγμένο σφάλμα να καταφέρει να πέσει κάτω από την τιμή threshold που ορίζουμε.

Μπορούμε να επιλέξουμε περισσότερες από μία από τις παραπάνω συνθήκες. Η εκπαίδευση διακόπτεται όταν πληρείται έστω και μία από τις συνθήκες που θα επιλέξουμε.

4.3.5. Αναγνώριση άγνωστων προτύπων

Η αναγνώριση ενός άγνωστου προτύπου μπορεί να έχει είτε τη μορφή μίας μόνο εικόνας, είτε μίας λίστας εικόνων. Στην περίπτωση της μίας εικόνας η εικόνα ανακατασκευάζεται από κάθε δίκτυο του συστήματος αναγνώρισης και υπολογίζεται το μέσο τετραγωνικό σφάλμα της ανακατασκευής από το κάθε δίκτυο. Το δίκτυο το οποίο θα καταφέρει να ανακατασκευάσει την εικόνα με το μικρότερο σφάλμα είναι και το δίκτυο νικητής το οποίο θεωρούμε ότι αναγνωρίζει επιτυχώς την εικόνα.

Στο τέλος της διαδικασίας της αναγνώρισης προβάλλεται το αποτέλεσμα της αναγνώρισης, μία λίστα με τις ανακατασκευασμένες εικόνες από το κάθε δίκτυο καθώς και το μέσο τετραγωνικό σφάλμα της ανακατασκευής (Εικόνα 22).

Στην περίπτωση που έχουμε μία λίστα εικόνων για αναγνώριση, εφόσον η λίστα περιέχει τις σωστές αντίστοιχες πληροφορίες που αφορούν την αναγνώριση, μπορεί να δημιουργηθεί αυτόματα ένας πίνακας αποτελεσματικότητας του συστήματος αναγνώρισης. Στον πίνακα φαίνονται οι πραγματικές ταυτότητες των εικόνων και οι ταυτότητες οι οποίες αποδόθηκαν από το σύστημα αναγνώρισης. Κάτω από τον πίνακα αυτό υπάρχουν τα ποσοστά επιτυχίας και αποτυχίας του συστήματος (Εικόνα 23).

	ρότυπη εικό	Σφάλμα Αναγνώρισης	Χαρακτηριστικό δικτύου
Δίκτυο 8			1537,2470588235 Person=8
Δίκτυο 9			1773,85882352947 Person=9
Δίκτυο 10			1225,40000000001 Person=10
Δίκτυο 11			296,949019607842 Person=11
Δίκτυο 12			1454,25098039215 Person=12
Δίκτυο 13			1759,07058823537 Person=13

Εικόνα 22: Προβολή του αποτελέσματος αναγνώρισης μίας μεμονωμένης εικόνας. Το δίκτυο με το μικρότερο σφάλμα ανακατασκευής (το δίκτυο Νο 11 στη συγκεκριμένη περίπτωση) καθορίζει το αποτέλεσμα της αναγνώρισης

Αριστερά υπάρχει η λίστα με τις εικόνες - πρότυπα προς αναγνώριση. Κάνοντας διπλό κλικ σε κάθε εικόνα, παρουσιάζεται το αποτέλεσμα της αναγνώρισης της συγκεκριμένης μεμονωμένης εικόνας όπως αναφέρθηκε παραπάνω.

ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΓΝΩΡΙΣΗΣ

Προγματική ιδιότητα

Αναγνώριση ως	10	9	8
12	0	0	0
11	0	0	0
10	1	0	0
9	0	1	0
8	0	0	1
7	0	0	0
6	0	0	0
5	0	0	0

Correct	False
100 %	0 %

ΑΝΑΓΝΩΡΙΣΗ ΕΠΙΛΕΓΜΕΝΗΣ ΕΙΚΟΝΑΣ

Προγματική ιδιότητα	Αναγνώριση ως
Person=11	Person=11

Προβολή όλων

Εικόνα 23: Προβολή του αποτελέσματος αναγνώρισης μίας λίστας εικόνων

5. Πειραματικά δεδομένα

5.1. Βάσεις Προσώπων

Για τη μέτρηση της αποτελεσματικότητας ενός συστήματος αναγνώρισης προσώπων αλλά και τη σύγκρισή του με άλλα αντίστοιχα συστήματα απαιτείται η ύπαρξη ενός πρότυπου συνόλου δεδομένων, με βάση το οποίο μπορούν να εξαχθούν συγκρίσιμα αποτελέσματα ακόμα και όταν οι μέθοδοι που εφαρμόζονται είναι τελείως διαφορετικές.

Σήμερα υπάρχει μία πληθώρα από τέτοιες βάσεις προσώπων οι οποίες χρησιμοποιούνται ευρέως. Η επιλογή της κατάλληλης βάσης δεδομένων προσώπων θα πρέπει να γίνεται είτε σύμφωνα με το στόχο της μεθόδου αναγνώρισης, είτε σύμφωνα με την ιδιότητα του συστήματος που πρέπει να δοκιμαστεί (π.χ. πως ο αλγόριθμος συμπεριφέρεται όταν στις εικόνες υπάρχουν έντονες αλλαγές στο φωτισμό ή εικόνες με αλλαγές στις εκφράσεις του προσώπου). Σε περιπτώσεις όπου η εκπαίδευση ενός αλγορίθμου απαιτεί περισσότερες από μία εικόνες ανά κλάση, κάποιες βάσεις δεδομένων οι οποίες περιέχουν μία εικόνα ανά κλάση δεν μπορούν να θεωρηθούν κατάλληλες. Επίσης, στην περίπτωση του προβλήματος της ανίχνευσης προσώπου, όταν οι εικόνες είναι σε ελεγχόμενο περιβάλλον με καθορισμένο φωτισμό και ομοιογενές φόντο, σίγουρα δεν είναι το κατάλληλο σύνολο δεδομένων για τον έλεγχο της αποτελεσματικότητας του αλγορίθμου.

Παραθέτουμε κάποιες από τις πιο γνωστές βάσεις δεδομένων προσώπων και τα κυριότερα χαρακτηριστικά τους:

α) Βάση προσώπων FERET

(<http://face.nist.gov/colorferet/>).

Η βάση προσώπων FERET αποτελείται από ένα σύνολο 14.126 έγχρωμων εικόνων, οι οποίες καταγράφηκαν σε ημιελεγχόμενο περιβάλλον. Λόγω του μεγέθους της βάσης, εικόνες από διαφορετικές περιόδους καταγραφής (15 διακριτοί περίοδοι συνολικά, από τον Αύγουστο του 1993 έως τον Ιούλιο του 1996)

παρουσιάζουν κάποιες διαφορές μεταξύ τους, λόγω του ότι ο εξοπλισμός έπρεπε να αποθηκεύεται και να ξανατοποθετείται σε κάθε περίοδο. Οι εικόνες είναι από συνολικά 1199 διαφορετικά άτομα και κάποια σετ εικόνων είναι διπλά με την έννοια ότι το ίδιο σετ εικόνων ενός προσώπου καταγράφηκε ξανά, μετά από κάποιο χρονικό διάστημα που μπορεί να φτάνει και τα 2 χρόνια. Μία έκδοση που είχε δημοσιοποιηθεί αρχικά περιλάμβανε μόνο grayscale εικόνες.

β) Βάση προσώπων Yale A

(<http://cvc.yale.edu/projects/yalefaces/yalefaces.html>).

Υπάρχουν δύο εκδόσεις της βάσης προσώπων Yale, η A και η B. Η A είναι η προγενέστερη η οποία περιέχει 165 grayscale εικόνες σε μορφή GIF από 15 διαφορετικά άτομα. Για κάθε άτομο υπάρχουν 15 εικόνες μία για κάθε έκφραση ή διαμόρφωση (χαρά, λύπη, έκπληξη, διαφορετικές συνθήκες φωτισμού, με/χωρίς γυαλιά).



Εικόνα 24: Δείγμα εικόνων από τη βάση προσώπων Yale A

γ) Βάση προσώπων Yale B

(<http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html>).

Η δεύτερη έκδοση της βάσης προσώπων Yale (B) έχει μόνο 10 διαφορετικά άτομα. Για κάθε όμως άτομο έχει 576 grayscale εικόνες (9 διαφορετικές πόζες x 64 διαφορετικές συνθήκες φωτισμού) φτάνοντας συνολικά τις 5760 εικόνες. Όλες οι εικόνες είναι φωτισμένες από μία μοναδική πηγή φωτός.



Εικόνα 25: Δείγμα εικόνων από τη βάση προσώπων Yale B

δ) Βάση προσώπων PIE

(http://www.ri.cmu.edu/research_project_detail.html?project_id=418&menu_id=261)

Η βάση προσώπων PIE αποτελείται από 41368 έγχρωμες εικόνες από 68 διαφορετικά άτομα κάτω από 43 διαφορετικές συνθήκες φωτισμού, με 13 διαφορετικές πόζες και 4 διαφορετικές εκφράσεις. Το φόντο των εικόνων δεν είναι ομοιογενές.



Εικόνα 26: Δείγματα εικόνων από τη βάση προσώπων PIE

ε) Βάση προσώπων AT&T (πρώην Olivetti Research Labs)

(<http://www.cl.cam.ac.uk/Research/DTG/attarchive/facedatabase.html>)

Αποτελείται από συνολικά 400 grayscale εικόνες 40 διαφορετικών ατόμων. Οι εικόνες έχουν καταγραφεί σε διαφορετικές χρονικές στιγμές, με μεταβλητό φωτισμό και διαφορετικές εκφράσεις προσώπου (ανοιχτά – κλειστά μάτια, χαμογελώντας ή όχι, με γυαλιά – χωρίς γυαλιά). Όλες οι εικόνες έχουν σκούρο ομοιογενές φόντο και όλα τα άτομα έχουν την ίδια πόζα, χωρίς κλίση του προσώπου, με ελάχιστες

πλάγιες μετατοπίσεις.



Εικόνα 27: Δείγματα εικόνων της βάσης προσώπων AT&T (πρώην Olivetti Research Labs)

στ) Βάση προσώπων AR

(http://cobweb.ecn.purdue.edu/~aleix/aleix_face_DB.html)

Αποτελείται από πάνω από 4000 έγχρωμες εικόνες προσώπων 126 διαφορετικών ατόμων (70 άνδρες και 56 γυναίκες). Σε όλες τις εικόνες τα πρόσωπα κοιτάνε μπροστά αλλά συνδιάζονται οι διαφορετικές οι εκφράσεις του προσώπου, οι συνθήκες φωτισμού και επιπλέον υπάρχουν και συνθήκες απόκρυψης τμήματος του προσώπου από ένα ζευγάρι μαύρα γυαλιά και ένα κασκόλ.



Εικόνα 28: Δείγματα εικόνων της Βάσης εικόνων AR

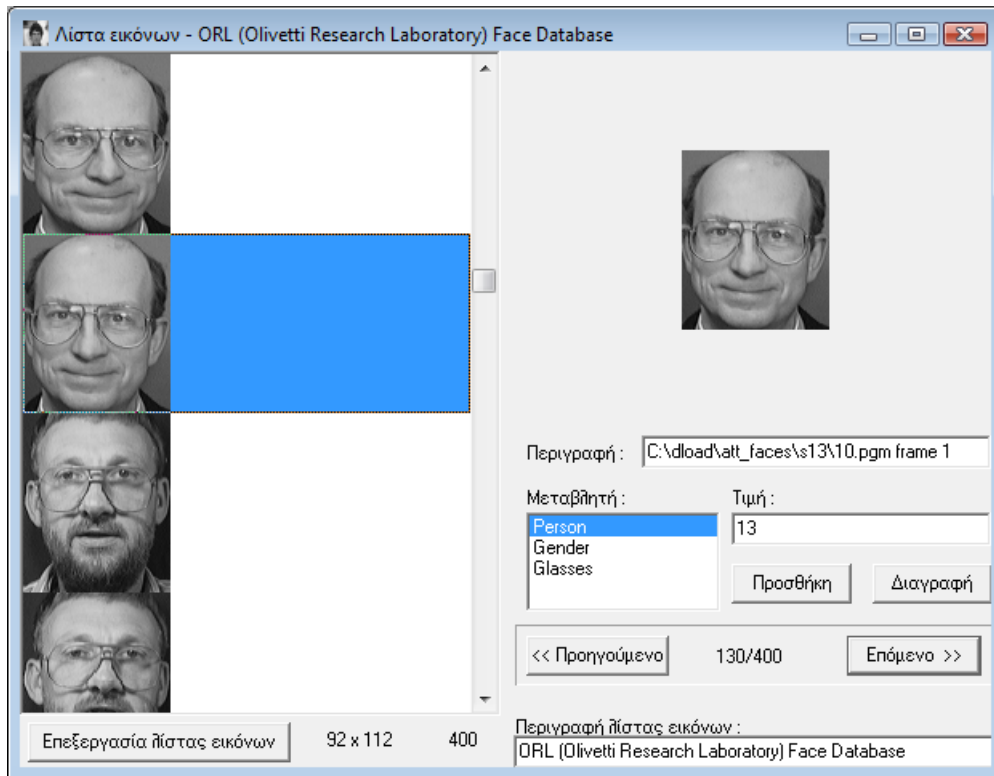
Υπάρχουν πολλές ακόμα βάσεις προσώπων, οι οποίες επικεντρώνονται σε κάποιο ιδιαίτερο χαρακτηριστικό των προσώπων, στις διαφορές των προσώπων ανάλογα με τη φυλή (πρόσωπα Κινέζων, Γιαπωνέζων), ακόμα και συλλογή εικόνων προσώπων από πραγματικά δεδομένα και όχι εργαστηριακά. Τα τελευταία χρόνια έχουν αναπτυχθεί και πολλές βάσεις προσώπων νέου τύπου, όπου πλέον τα δεδομένα είναι υψηλής ανάλυσης, τριών διαστάσεων, ή ακόμα και βίντεο προσώπων, ακόμα και με ψηφιακό ήχο. Μία ενημερωμένη λίστα από τις πιο σύγχρονες βάσεις προσώπων υπάρχει διαθέσιμη στην ιστοσελίδα:

Face Recognition Homepage: <http://www.face-rec.org/databases/>

Για την αξιολόγηση του αλγορίθμου που υλοποιήσαμε χρησιμοποιήσαμε δεδομένα από 4 διαφορετικές βάσεις προσώπων (AT&T, Essex Faces95, Essex Grimaces, Yale A).

5.2. Υλοποίηση συστήματος αναγνώρισης της βάσης προσώπων AT&T (πρώην Olivetti Research Labs).

Για τη δημιουργία της βάσης των προσώπων δημιουργούμε μία λίστα εικόνων στην οποία εισάγουμε το σύνολο των εικόνων. Η βάση της AT&T αποτελείται από δέκα (10) εικόνες προσώπων σαράντα (40) διαφορετικών ατόμων δηλαδή συνολικά 400 εικόνες προσώπων. Η συγκεκριμένη βάση είναι ιδανική για χρήση στην εφαρμογή μας και την έχουμε επιλέξει γιατί όλες οι εικόνες είναι ήδη επεξεργασμένες, έτσι ώστε τα πρόσωπα να είναι περίπου στην ίδια θέση χωρίς μεγάλες κλίσεις του προσώπου και με ομοιόμορφο φωτισμό. Αφού έχουμε εισάγει όλες τις εικόνες σε μία λίστα εικόνων προχωράμε στη δημιουργία μιας βοηθητικής μεταβλητής «person», η οποία θα μας βοηθήσει κατά την εκπαίδευση αλλά και κατά την αξιολόγηση της αποτελεσματικότητας του συστήματος αναγνώρισης. Καλό είναι στο σημείο αυτό να αποθηκεύσουμε τη λίστα σε αρχείο. Στη συνέχεια, για μεγαλύτερη ευκολία εφόσον έχουμε ετοιμάσει τη λίστα με την πλήρη βάση των προσώπων, χρησιμοποιούμε το wizard για την αυτόματη δημιουργία του συστήματος αναγνώρισης με βάση τη λίστα που έχουμε δημιουργήσει.



Εικόνα 29: Η βάση προσώπων όπως έχει περαστεί στην εφαρμογή. Είναι εμφανείς οι τρεις μεταβλητές που ορίσαμε για τον χαρακτηρισμό των εικόνων: Person, Gender και Glasses

Ξεκινώντας το wizard μας ζητείται η λίστα των εικόνων που θα χρησιμοποιηθεί για την εκπαίδευση των δικτύων που θα δημιουργηθούν. Επιλέγουμε τη λίστα που δημιουργήσαμε στο προηγούμενο βήμα και επιλέγουμε «επόμενο». Η λίστα φορτώνεται και μας δίνεται η δυνατότητα αν θέλουμε να την επεξεργαστούμε περαιτέρω. Συνεχίζουμε δίνοντας τα απαραίτητα δεδομένα για τη δημιουργία των δικτύων (διαστάσεις νευρώνων εισόδου και πλήθος νευρώνων εξόδου) καθώς και για την αρχικοποίηση των βαρών και τις παραμέτρους τις εκπαίδευσης τους. Επίσης επιλέγουμε και τη μεταβλητή της λίστας, με βάση την οποία θα δημιουργηθούν τα διαφορετικά δίκτυα. Για κάθε τιμή της μεταβλητής δημιουργείται και ένα δίκτυο, το οποίο στις συνθήκες εκπαίδευσης του θα έχει τη μεταβλητή αυτή να ισούται με την τρέχουσα τιμή.

Θα πρέπει να λάβουμε υπόψιν ότι το πλήθος των εικόνων-δεδομένων εισόδου που προκύπτει για κάθε δίκτυο θα πρέπει πάντα να είναι μεγαλύτερο του πλήθους των νευρώνων εξόδου, ειδικά η εκπαίδευση των δικτύων θα είναι αδύνατη. Εφόσον έχουνε πλέον δημιουργηθεί τα δίκτυα τα οποία θα αποτελούν το σύστημα αναγνώρισης των 40 προσώπων της βάσης, μπορούμε να προχωρήσουμε στην εκπαίδευση τους. Ο wizard έχει ήδη ρυθμίσει τα απαραίτητα φίλτρα στο κάθε δίκτυο έτσι ώστε το καθένα από αυτά να εκπαιδευτεί με τις αντίστοιχες εικόνες των προσώπων και μόνο αυτές (10 για κάθε πρόσωπο). πριν ξεκινήσουμε την εκπαίδευση των δικτύων καλόν είναι να δούμε τις βασικές ρυθμίσεις της εκπαίδευσης. Οι πλέον σημαντικές ρυθμίσεις είναι το γ και το β_0 , άλλα και στην περίπτωση που θέλουμε να αυτοματοποιήσουμε την εκπαίδευση, να ορίσουμε τη συνθήκη τερματισμού της διαδικασίας εκπαίδευσης του δικτύου. Εδώ προτιμούμε να ορίσουμε ένα αρκετά χαμηλό σφάλμα και όχι κάποιο χρονικό όριο, ώστε να διασφαλίσουμε ότι τα δίκτυά μας έχουν εκπαιδευτεί σωστά. Αν η εκπαίδευση των δικτύων δεν είναι πλήρης και ομοιόμορφη, τότε η αποτελεσματικότητα του συστήματος αναγνώρισης θα είναι αρκετά μειωμένη με αρκετά λάθη. ένα δίκτυο το οποίο δεν έχει εκπαιδευτεί ικανοποιητικά, σίγουρα θα δίνει μεγαλύτερο σφάλμα κατά την ανακατασκευή, με αποτέλεσμα να αυξάνονται οι πιθανότητες να έχουμε εσφαλμένη αναγνώριση, ειδικά στα πρότυπα, τα οποία ανήκουν στην κλάση την οποία εκπαιδεύτηκε να αναγνωρίζει.

Στη συνέχεια, αφήνουμε τα δίκτυα να εκπαιδευτούν. Αυτό μπορεί να γίνει είτε σειριακά, είτε παράλληλα. Η φάση της εκπαίδευσης είναι η πιο χρονοβόρα διαδικασία και μπορεί να διαρκέσει αρκετές ώρες. ο χρόνος εκπαίδευσης εξαρτάται από το μέγεθος των εικόνων, το μέγεθος των δικτύων άλλα και του σφάλματος-στόχου που έχουμε ορίσει.

όταν έχει ολοκληρωθεί και η εκπαίδευση όλων των δικτύων μπορούμε πλέον να δοκιμάσουμε το σύστημα αναγνώρισης, σε πρώτη φάση με τις εικόνες των προτύπων εκπαίδευσης. Με τον έλεγχο αυτό αναμένουμε πολύ υψηλά ποσοστά επιτυχίας που αγγίζουν το 100%. Αυτό αποτελεί απλά μία ένδειξη ότι τα δίκτυα μας έχουν εκπαιδευτεί σωστά. Χαμηλό ποσοστό επιτυχίας του συστήματος σε αυτή τη φάση θα σήμαινε με σιγουριά ότι κάποιο ή κάποια από τα δίκτυα του συστήματος

αναγνώρισης δεν έχουν εκπαιδευτεί πλήρως ή δεν έχουν εκπαιδευτεί σωστά.

5.3. Αναγνώριση αλλοιωμένων προτύπων

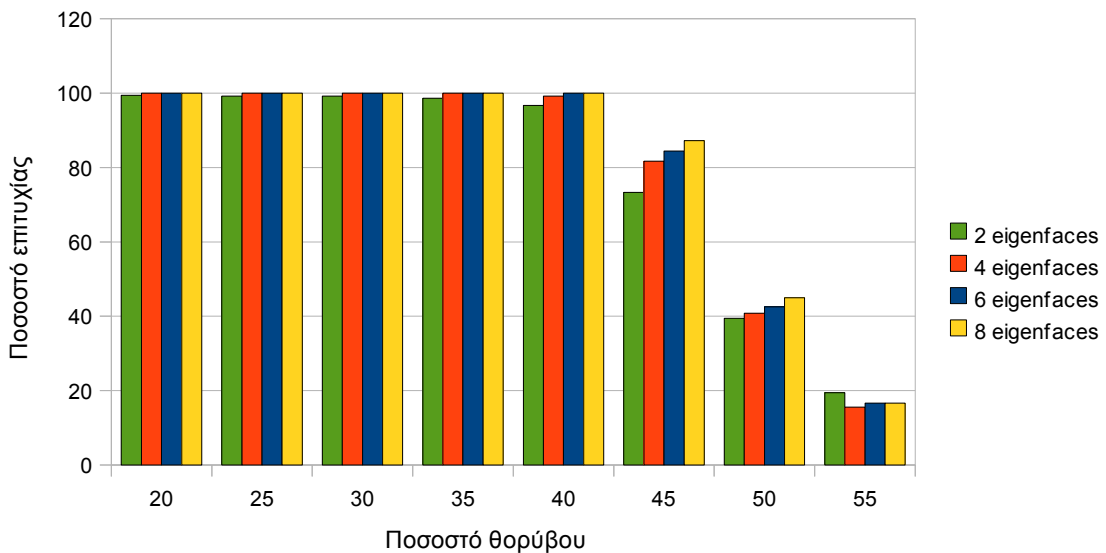
Η αποτελεσματικότητα του συστήματος φαίνεται μόνο χρησιμοποιώντας άγνωστες εικόνες, διαφορετικές από αυτές που χρησιμοποιήθηκαν κατά την εκπαίδευση. Για το σκοπό αυτό μπορούμε να κατασκευάσουμε κάποια αλλοιωμένα πρότυπα παραμορφώνοντας με κάποιο τρόπο τα αρχικά. Για το σκοπό αυτό χρησιμοποιούμε έναν αλγόριθμο ο οποίος προσθέτει κάποιες τυχαίες τιμές - θόρυβο στις εικόνες. Στη συνέχεια αφήνουμε το σύστημα να αναγνωρίσει αυτόματα τη λίστα με τις παραμορφωμένες εικόνες. Αυτό θα δείξει κυρίως το πόσο ανεκτικά είναι τα δίκτυα σε αλλοιώσεις. Αν και γενικά δεν είναι καλό τεστ, μια που γενικά τα δίκτυα PCA είναι αρκετά ευαίσθητα σε αλλαγές που αφορούν τη φωτεινότητα, τον προσανατολισμό του κεφαλιού άλλα και το φόντο της εικόνας, παρόλα αυτά το ποσοστό επιτυχίας του συστήματος φαίνεται να παραμένει αρκετά υψηλό, ακόμα και όταν το επίπεδο του θορύβου έχει αυξηθεί αρκετά.

Για να διαπιστώσουμε την ανεκτικότητα του συστήματος αναγνώρισης στον τυχαίο θόρυβο που αλλοιώνει τις εικόνες, δοκιμάσαμε την αποτελεσματικότητα του συστήματος με διαφορετικά ποσοστά θορύβου στις εικόνες, οι οποίες χρησιμοποιήθηκαν για την εκπαίδευση των δικτύων.

Για το πείραμα χρησιμοποιήθηκε η βάση εικόνων προσώπων της AT&T. Η ανάλυση των εικόνων μειώθηκε στο μισό (46 x 56 από 92 x 112). Η λίστα εικόνων που χρησιμοποιήθηκε για την αναγνώριση είναι η ίδια που χρησιμοποιήθηκε και για την εκπαίδευση των δικτύων, με μόνη διαφορά την προσθήκη του τυχαίου θορύβου. Επαναλάβαμε το πείραμα σε διάφορα ποσοστά θορύβου και με 4 διαφορετικές διαμορφώσεις του συστήματος αναγνώρισης, με 2,4,6 και 8 eigenfaces.

Θόρυβος	2 eigenfaces	4 eigenfaces	6 eigenfaces	8 eigenfaces
20,00%	99,44%	100,00%	100,00%	100,00%
25,00%	99,17%	100,00%	100,00%	100,00%
30,00%	99,17%	100,00%	100,00%	100,00%
35,00%	98,61%	100,00%	100,00%	100,00%
40,00%	96,67%	99,17%	100,00%	100,00%
45,00%	73,33%	81,67%	84,44%	87,22%
50,00%	39,44%	40,83%	42,50%	45,00%
55,00%	19,44%	15,56%	16,67%	16,67%

Πίνακας 3: Πίνακας αποτελεσματικότητας του αλγορίθμου ως προς το ποσοστό τυχαίου θορύβου και τον αριθμό των eigenfaces.



Σχήμα 9: Γραφική παράσταση του ποσοστού επιτυχίας του συστήματος αναγνώρισης σε σχέση με το ποσοστό τυχαίου θορύβου που εισάγουμε στα δεδομένα προς αναγνώριση

Από τα πειραματικά αποτελέσματα που φαίνονται στο παραπάνω γράφημα του σχήματος 9 συμπεραίνουμε ότι το σύστημα αναγνώρισης είναι αρκετά ανεκτικό σε παραμορφώσεις αυτής της μορφής, αφού το ποσοστό επιτυχίας παραμένει σε πολύ υψηλά ποσοστά, πολύ κοντά στο 100%, ακόμα και όταν το ποσοστό του θορύβου φτάνει το 40%, ακόμα και με 2 μόνο εξόδους σε κάθε δίκτυο. Το ποσοστό επιτυχίας είναι κάπως καλύτερο για ποσοστά θορύβου από 40% - 50%, όσο

περισσότερα eigenfaces χρησιμοποιούμε. Όταν όμως το ποσοστό θορύβου φτάνει το 50% πλέον το ποσοστό επιτυχίας είναι πολύ μικρό και περισσότερο τυχαίο.

5.4. Εκπαίδευση των δικτύων με τμήμα μόνο της βάσης AT&T

Η πραγματική αξιολόγηση της αποτελεσματικότητας του συστήματος μπορεί να γίνει μόνο με τελείως άγνωστες εικόνες. Για να προσομοιώσουμε μια τέτοια κατάσταση, από τη λίστα των εικόνων προσώπων αφαιρούμε κατά τη φάση της εκπαίδευσης μια εικόνα από κάθε άτομο. Μ' αυτό τον τρόπο έχουμε για την εκπαίδευση 9 εικόνες για κάθε άτομο και μια εικόνα άγνωστη, η οποία θα μας δείξει την αποτελεσματικότητα του συστήματος αναγνώρισης. Έχοντας μειώσει τον αριθμό των εικόνων ανά άτομο μειώνεται περαιτέρω και ο μέγιστος αριθμός των νευρώνων εξόδου των δικτύων που μπορούμε να έχουμε και κατά συνεπεία των κυρίων συνιστωσών που μπορούν να εξαχθούν από τις εικόνες εκπαίδευσης. Εκπαιδεύουμε για άλλη μια φορά τα δίκτυα και χρησιμοποιούμε είτε την πλήρη λίστα των προσώπων, είτε μόνο τη λίστα των εικόνων που παραλείψαμε από την εκπαίδευση.

Αναλυτικότερα, για το πείραμα αυτό χρησιμοποιήσαμε και πάλι τη βάση προσώπων της AT&T. Η ανάλυση των εικόνων μειώθηκε και εδώ στο μισό (46 x 56 από 92 x 112). Από το σύνολο των δέκα εικόνων του κάθε προσώπου, απομονώσαμε ένα τυχαίο δείγμα, διαχωρίζοντας το σύνολο της βάσης σε εννέα εικόνες ανά πρόσωπο ως το σύνολο εκπαίδευσης του κάθε δικτύου (συνολικά 360 εικόνες) και μία εικόνα προσώπου από κάθε άτομο ως το σύνολο των αγνώστων προτύπων που θα αναγνωριστεί (συνολικά 40 εικόνες), ώστε να αξιολογηθεί η αποτελεσματικότητα του συστήματος. Η εικόνα αυτή επιλέχθηκε τυχαία, χρησιμοποιώντας το σύστημα τυχαίου αποκλεισμού εικόνων της εφαρμογής. Το πείραμα επαναλήφθηκε 2 φορές επιλέγοντας να αποκλείσουμε διαφορετικές εικόνες κάθε φορά και να τις χρησιμοποιήσουμε για τη δοκιμή της αποτελεσματικότητας. Σε κάθε επανάληψη δημιουργήσαμε και πάλι 4 συστήματα αναγνώρισης με 2,4,6 και 8 νευρώνες εξόδου.

Τα αποτελέσματα των δοκιμών φαίνονται στον πίνακα (4).

ΔΟΚΙΜΗ	2 eigenfaces	4 eigenfaces	6 eigenfaces	8 eigenfaces
1	92,5% (37/40)	97,5% (39/40)	97,5% (39/40)	97,5% (39/40)
2	95,0% (38/40)	100% (39/40)	100% (40/40)	100% (40/40)

Πίνακας 4: Πίνακας αποτελεσματικότητας αναγνώρισης άγνωστου προτύπου (βάση AT&T)

5.5. Υλοποίηση συστήματος αναγνώρισης της βάσης προσώπων Yale A

Επαναλάβαμε το παραπάνω πείραμα χρησιμοποιώντας σαν δεδομένα εισόδου τη βάση προσώπων Yale A. Η βάση Yale περιλαμβάνει 11 εικόνες προσώπων 15 διαφορετικών ατόμων. Οι εικόνες αυτές είναι της μορφής

- 1) κεντρικός φωτισμός
- 2) με γυαλιά
- 3) χαρούμενη έκφραση
- 4) φωτισμός από αριστερά
- 5) χωρίς γυαλιά
- 6) κανονική έκφραση
- 7) φωτισμός από δεξιά
- 8) λυπημένη έκφραση
- 9) κοιμισμένος/η (κλειστά μάτια)
- 10) λυπημένη έκφραση
- 11) κλείνοντας το ένα μάτι

Γενικά οι εικόνες δεν έχουν μεγάλες διαφορές μεταξύ τους, με εξαίρεση της εικόνας 4 και 7, όπου ο ιδιαίτερος φωτισμός προκαλεί έντονη αλλαγή στο φόντο, λόγω της σκιάς του ατόμου.



Εικόνα 30: Έντονη αλλαγή στις εικόνες από μεταβολή του φωτισμού στις εικόνες της βάσης προσώπων Yale A

Για να αποφύγουμε τα προβλήματα τα οποία θα προκαλούσε αυτή η έντονη αλλαγή στις εικόνες κατά την εκπαίδευση, αποκλείσαμε τις συγκεκριμένες εικόνες από την εκπαίδευση των δικτύων και από τις υπόλοιπες αποκλείσαμε 1 εικόνα από κάθε πρόσωπο και εκπαιδεύσαμε τα δίκτυα με τις υπόλοιπες 8.

Έχοντας μόνο 8 εικόνες για εκπαίδευση, χρησιμοποιήσαμε δίκτυα με 60x44 εισόδους και 6 εξόδους.

Στην περίπτωση 1 που αποκλείσουμε τελείως τις 2 εικόνες ανά άτομο με τις αλλαγές στο φωτισμό, η αποτελεσματικότητα του συστήματος είναι 100%. Όλες οι εικόνες της εκπαίδευσης, αλλά και οι 15 άγνωστες αναγνωρίζονται με επιτυχία.

Στην περίπτωση 2 που περιλάβουμε και τις 2 εικόνες ανά άτομο με το διαφορετικό φωτισμό, το ποσοστό επιτυχίας πέφτει στο 89,7%. (17 εικόνες δεν αναγνωρίζονται σωστά, και οι 17 είναι από τις 30 τις οποίες αποκλείσαμε αρχικά).

ΔΟΚΙΜΗ	6 eigenfaces
1	100% (135/135)
2	89,7% (148/165)

Πίνακας 5: Πίνακας αξιοπιστίας του συστήματος κατά τις δοκιμές με τη βάση προσώπων Yale A

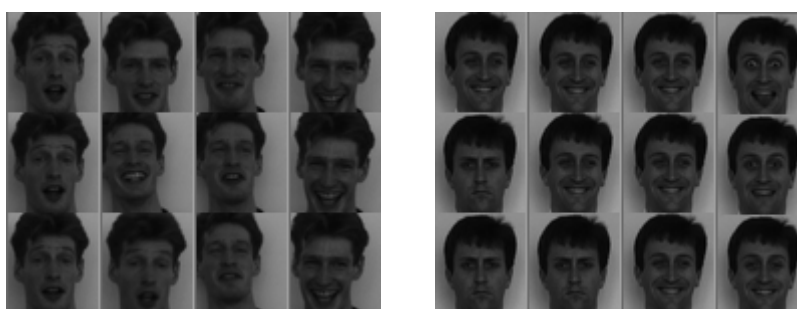
5.6. Υλοποίηση συστήματος αναγνώρισης της βάσης προσώπων Essex (Grimaces, Faces95)

Στο τελευταίο πείραμα χρησιμοποιήσαμε σαν δεδομένα τις βάσεις προσώπων του πανεπιστημίου του Essex Grimaces και Faces95.

5.6.1. Grimaces

Η Βάση προσώπων Grimaces είναι ιδανική καθώς όλα τα πρόσωπα είναι ομοιόμορφα χωρίς καμία άλλη αλλαγή εκτός από την έκφραση του προσώπου. Περιέχονται συνολικά 18 άτομα με 20 φωτογραφίες από το κάθε ένα.

Για το πείραμα χρησιμοποιήσαμε δίκτυα μεγέθους 45x50 εισόδους και 2,6 και 8 εξόδους. Από τις 20 εικόνες ανά άτομο, αποκλείσαμε 2 εικόνες από κάθε άτομο και εκπαιδεύσαμε τα δίκτυα με τις υπόλοιπες 18.



Εικόνα 31: Δείγματα εικόνων της βάσης προσώπων Grimaces του πανεπιστημίου του Essex

Στη συνέχεια επαναλάβουμε το πείραμα διαχωρίζοντας τις εικόνες του κάθε ατόμου της βάσης προσώπων σε 10 εικόνες εκπαίδευσης και 10 άγνωστες.

Τα αποτελέσματα σε κάθε περίπτωση ήταν θεαματικά, χωρίς κανένα σφάλμα κατά την αναγνώριση.

ΔΟΚΙΜΗ	2 eigenfaces	4 eigenfaces	6 eigenfaces	8 eigenfaces
1	100% (36/36)	100% (36/36)	100% (36/36)	100% (36/36)
2	100% (180/180)	100% (180/180)	100% (180/180)	100% (180/180)

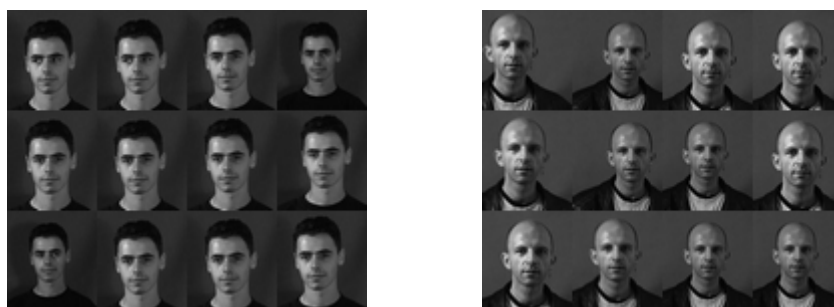
Πίνακας 6: Πίνακας αποτελεσματικότητας του συστήματος με τη βάση προσώπων Grimaces

5.6.2. Faces95

Σε δεύτερη φάση του πειράματος χρησιμοποιήσαμε τη βάση Faces95 πάλι του πανεπιστημίου του Essex.

Η βάση Faces95 περιέχει 20 εικόνες από κάθε ένα από συνολικά 72 διαφορετικά άτομα, 1440 εικόνες συνολικά. Τα χαρακτηριστικά της συγκεκριμένης βάσης προσώπων είναι τα εξής:

- Το φόντο σε όλες τις εικόνες είναι ομοιόμορφο και αποτελείται μόνο από μία κουρτίνα. Μικρές αλλαγές στο φόντο προκαλούνται από αλλαγές στις σκιές καθώς το άτομο μετακινείται μπροστά ή πίσω.
- Διαστάσεις του προσώπου. Οι διαστάσεις του προσώπου μεταβάλλονται αισθητά καθώς το άτομο πλησιάζει ή απομακρύνεται από το φωτογραφικό φακό.
- Στροφή του κεφαλιού, αλλαγή πόζας. Πολύ μικρές μεταβολές.
- Θέση του προσώπου στην εικόνα. Υπάρχει μετατόπιση του προσώπου ως προς τη θέση του μέσα στις εικόνες.
- Μεταβολές στο φωτισμό. Υπάρχει αλλαγή στο φωτισμό του προσώπου όταν το άτομο μετακινείται προς τα εμπρός
- Έκφραση προσώπου. Πολύ μικρές αλλαγές



Εικόνα 32: Δείγματα εικόνων της βάσης προσώπων faces95. Είναι εμφανής η μετακίνηση του ατόμου που προκαλεί αλλαγές στη θέση, στο μέγεθος και το φωτισμό του προσώπου στις εικόνες.

Επαναλάβουμε το πείραμα αποκλείοντας αρχικά 2 τυχαίες εικόνες από κάθε άτομο και εκπαιδεύοντας τα δίκτυα με τις υπόλοιπες 18 και στη συνέχεια αποκλείοντας 4 τυχαίες εικόνες και εκπαιδεύοντας τα δίκτυα με τις υπόλοιπες 16. Με καθένα από τα παραπάνω σετ εκπαιδεύσαμε ένα σύστημα με δίκτυα 4 εξόδων και ένα

σύστημα με 8 εξόδους. Τα αποτελέσματα φαίνονται στον πίνακα (7)

ΔΟΚΙΜΗ	4 eigenfaces	8 eigenfaces
2 άγνωστα πρότυπα	88,89% (128/144)	92,36% (133/144)
4 άγνωστα πρότυπα	90,63% (261/288)	92,01% (265/288)

Πίνακας 7: Πίνακας αξιοπιστίας του συστήματος κατά τις δοκιμές με τη βάση προσώπων Faces95 του πανεπιστημίου του Essex

6. Συμπεράσματα

Από το πείραμα με τη βάση προσώπων της AT&T παρατηρούμε ότι κατά την πρώτη δοκιμή υπάρχει μία εικόνα η οποία δεν αναγνωρίζεται σωστά ούτε και στην περίπτωση όπου έχουμε 8 eigenfaces. Εστιάζοντας στη λάθος αυτή αναγνώριση διαπιστώνουμε τα μειονεκτήματα της μεθόδου αναγνώρισης με eigenfaces. Η εικόνα (33) δείχνει την ανακατασκευή του άγνωστου προτύπου και το σφάλμα ανακατασκευής που οδηγεί σε λάθος αναγνώριση. Το άγνωστο πρότυπο - εικόνα ανήκει στο άτομο 28, αλλά το δίκτυο το οποίο έχει εκπαιδευτεί να αναγνωρίζει εικόνες του ατόμου 37 πετυχαίνει καλύτερη ανακατασκευή.

Δίκτυο 28		828,729411764708	Person=28
Δίκτυο 37		728,149019607841	Person=37

Εικόνα 33: Δείγμα εσφαλμένης αναγνώρισης του προτύπου κατά την 1η δοκιμή. Στην εικόνα φαίνονται στις στήλες από αριστερά προς τα δεξιά τα IDs των δικτύων, το πρότυπο προς αναγνώριση, η ανακατασκευή του άγνωστου προτύπου, το σφάλμα ανακατασκευής και το χαρακτηριστικό του δικτύου.

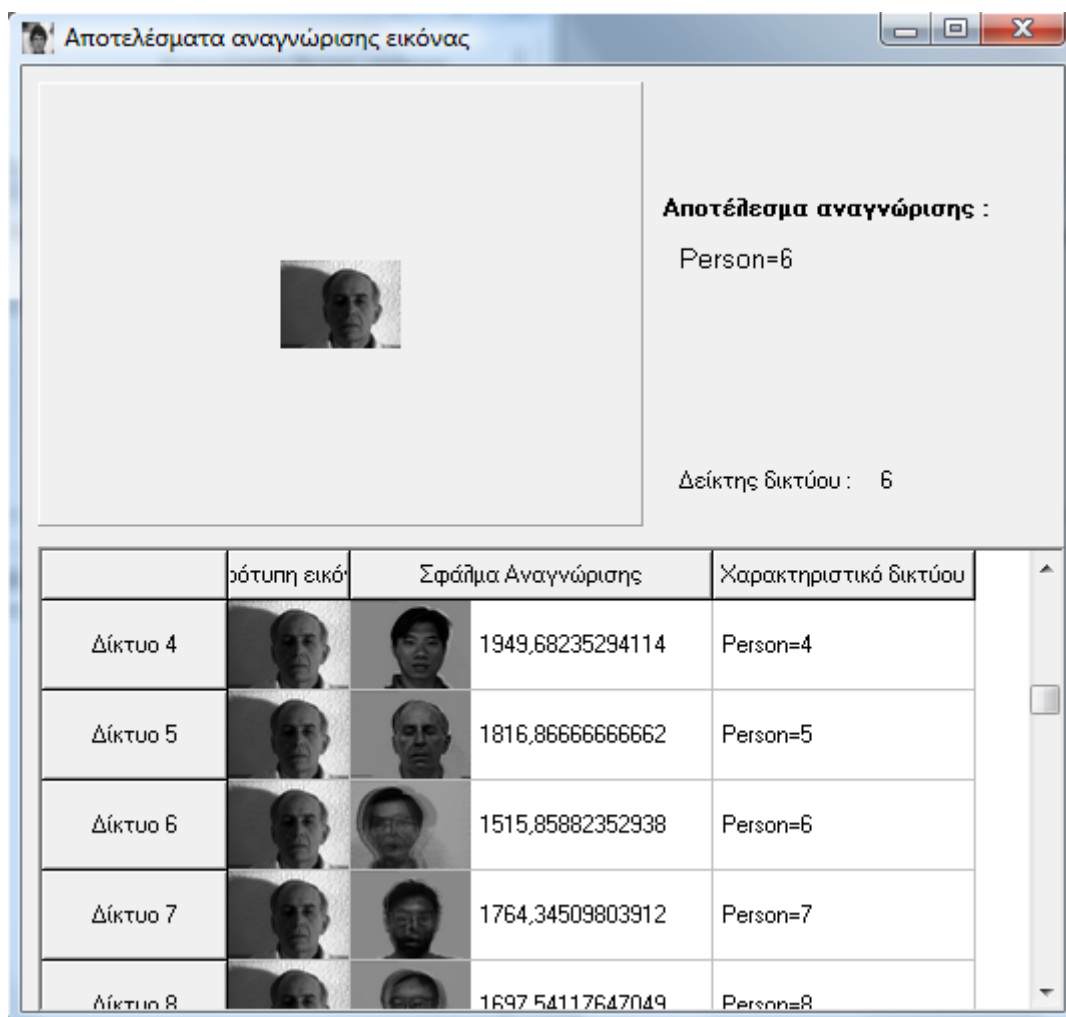
Μπορούμε να παρατηρήσουμε αρχικά ότι υπάρχει αρκετά μεγάλη ομοιότητα μεταξύ των δύο ατόμων τα οποία έχουν και τα δύο φαλάκρα, γυαλιά και μούσι. Αν όμως παρατηρήσουμε την ανακατασκευή του προτύπου από το δίκτυο 28 θα δούμε ότι το δίκτυο προσπαθεί να πετύχει την ανακατασκευή της κλίσης του προσώπου προς τα δεξιά, αλλά δεν φαίνεται να το καταφέρνει. Αν παρατηρήσουμε και το σύνολο των εικόνων που χρησιμοποιήθηκαν για την εκπαίδευση του συγκεκριμένου δικτύου (Εικόνα 34), θα διαπιστώσουμε ότι καμία από τις εικόνες δεν έχει παρόμοια κλίση.



Εικόνα 34: Το σύνολο των εικόνων που χρησιμοποιήθηκαν για την εκπαίδευση του δικτύου

Το γεγονός αυτό, αν και μεμονωμένο, όπως φαίνεται από τη δεύτερη δοκιμή όπου επιλέχθηκαν διαφορετικές τυχαίες πάλι εικόνες και τα αποτελέσματα δεν παρουσιάζουν κανένα σφάλμα κατά την αναγνώριση, αναγνωρίζοντας επιτυχώς και τα 40 άγνωστα πρότυπα, δεν παύει να αποτελεί ένδειξη της αδυναμίας της μεθόδου στην περίπτωση που υπάρχει έντονη κλίση του προσώπου και αυτό δεν αποτελεί μέρος της εκπαίδευσης.

Στη δεύτερη περίπτωση της χρήσης της βάσης προσώπων Yale παρουσιάζεται ένα αντίστοιχο φαινόμενο, αυτή τη φορά με έντονες αλλαγές στο φόντο της εικόνας. Παρόλο που οι συγκεκριμένες εικόνες έχουν δημιουργηθεί για να δοκιμάσουν την αποτελεσματικότητα του αλγορίθμου σε συνθήκες διαφορετικού φωτισμού, το κυρίως πρόβλημα στην περίπτωσή μας είναι η έντονη αλλαγή στο φόντο που την προκαλεί ο διαφορετικός φωτισμός. Η μέθοδος PCA σαν ολιστική μέθοδος δεν διαχωρίζει το φόντο της εικόνας από την εικόνα του προσώπου και έτσι αυτές οι έντονες αλλαγές (σκιά του ατόμου από την αλλαγή του φωτισμού) επιδρούν αρνητικά στην αποτελεσματικότητα του αλγορίθμου.



Εικόνα 35: Επίδραση της σκιάς στο φόντο και ο τρόπος που επηρεάζει την αναγνώριση. Η σκιά που πέφτει αριστερά έχει την τάση να μειώνει το σφάλμα ανακατασκευής στα δίκτυα, όπου το πρόσωπο στις εικόνες της εκπαίδευσης βρισκόταν αριστερά. Αντίστοιχο φαινόμενο παρατηρείται και με την σκιά προς την αντίθετη κατεύθυνση.

Στην τρίτη περίπτωση όπου χρησιμοποιήσαμε τη βάση προσώπων Faces95 αντιμετωπίσαμε δυσκολίες λόγω την έντονης μεταβολής της θέσης του προσώπου μέσα στην εικόνα. Τα αποτελέσματα εξαρτώνται αρκετά από τον τυχαίο διαχωρισμό των δύο συνόλων (εκπαίδευσης και ελέγχου). Παρόλα αυτά το αποτέλεσμα της αναγνώρισης δεν είναι καθόλου άσχημο αν λάβουμε υπόψη και το γεγονός ότι η βάση προσώπων Faces95 περιέχει και έναν αρκετά μεγάλο αριθμό από διαφορετικά άτομα (72). Και εδώ φαίνεται ότι η αύξηση του αριθμού των

eigenfaces που χρησιμοποιήθηκαν βελτιώνει μεν την αξιοπιστία του συστήματος αλλά πιθανότατα η διαφορά αυτή να μην είναι ικανή να δικαιολογήσει τις επιπλέον υπολογιστικές απαιτήσεις και τον επιπλέον χρόνο που απαιτείται για την εκπαίδευση.

Τα πειραματικά δεδομένα αποδεικνύουν ότι η μέθοδος PCA μπορεί να είναι μία εύκολη και εύχρηστη μέθοδος με εφαρμογή στην αναγνώριση προσώπων, που όμως, όπως οι περισσότερες ολιστικές μέθοδοι είναι αρκετά ευαίσθητη σε έντονες αλλαγές στις εικόνες, όπως η αλλαγή στην πόζα του προσώπου και οι έντονες αλλαγές στο φόντο των εικόνων.

Τα προβλήματα αυτά μπορούν να μειωθούν δραματικά αν η μέθοδος συνδυαστεί με κάποια μέθοδο προεπεξεργασίας των εικόνων όπως η ανίχνευση του προσώπου. Στα πειράματα τα οποία διεξήχθησαν, τα δεδομένα χρησιμοποιήθηκαν χωρίς καμία προεπεξεργασία γι' αυτό και σε ορισμένες περιπτώσεις η αποτελεσματικότητα του αλγορίθμου παρουσιάζεται μειωμένη. Στις περιπτώσεις όπου τα δεδομένα είχαν άμεσα την αναμενόμενη μορφή, τα αποτελέσματα ήταν θεαματικά.

Η μέθοδος PCA μπορεί να χρησιμοποιηθεί και σαν τμήμα ενός πιο σύνθετου ή υβριδικού συστήματος αναγνώρισης, χάρη στην πολύ αποτελεσματική ικανότητα που έχει να μειώνει τις διαστάσεις και την πολυπλοκότητα των δεδομένων εισόδου.

7. Βιβλιογραφία

- Acosta, E., L. Torres, A. Albiol, and E. J. Delp. 2002. An automatic face detection and recognition system for video indexing applications. In *IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS SPEECH AND SIGNAL PROCESSING*, 4:3644–3647.
- Akarun, L., B. Gökberk, and A. A. Salah. 2005. 3D face recognition for biometric applications. In *Proc. of the 13th European Signal Processing Conference (EUSIPCO)*.
- Barlow, H. B, and P. Foldiak. 1989. Adaptation and decorrelation in the cortex. *The computing neuron*: 54–72.
- Belhumeur, P. N, J. P Hespanha, D. J Kriegman, and others. 1997. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on pattern analysis and machine intelligence* 19, no. 7: 711–720.
- Costen, N. P., T. F. Cootes, and C. J. Taylor. 2002. Compensating for ensemble-specific effects when building facial models. *Image and Vision Computing* 20, no. 9-10 (August 1): 673-682. doi:10.1016/S0262-8856(02)00057-4. <http://www.sciencedirect.com/science/article/B6V09-461XPPM-2/2/c2b7adb77993e4cbcf8fc76d5e086e9e>.
- Diamantaras, K. I., and S. Y. Kung. 1996. *Principal Component Neural Networks: Theory and Applications*. Wiley-Interscience, February.
- Fazl-Ersi, E., J. S Zelek, and J. K Tsotsos. 2007. Robust face recognition through local graph matching. *Journal of Multimedia* 2, no. 5: 31.
- Földiák, Peter, and Peter Fdilir. 1989. Adaptive Network for Optimal Linear Feature Extraction. *Proceedings of the IEEE/INNS International Joint Conference on Neural Networks* 1 (June 18): pp. 401-405.
- Gross, R., S. Baker, I. Matthews, and T. Kanade. 2004. Face recognition across pose and illumination. *Handbook of Face Recognition*: 193–216.
- Harold Hotelling. 1933. *Analysis of a complex of statistical variables into principal components*. (Baltimore): Warwick & York.

http://openlibrary.org/b/OL22285550M/Analysis_of_a_complex_of_statistica_l_variables_into_principal_components.

- Hebb, D.O. 1949. *The Organization of Behavior: A Neuropsychological Theory*. New edition. Psychology Press.
- Heisele, B., P. Ho, and T. Poggio. 2001. Face recognition with support vector machines: Global versus component-based approach. In *Proc. 8th International Conference on Computer Vision*, 2:688–694.
- Heisele, B., T. Serre, M. Pontil, and T. Poggio. 2001. Component-based face detection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Vol. 1.
- Huang, J., B. Heisele, and V. Blanz. 2003. Component-based face recognition with 3D morphable models. *Lecture notes in computer science*: 27–34.
- Hyvärinen, A., and E. Oja. 2000. Independent component analysis: algorithms and applications. *Neural Networks* 13, no. 4-5: 411–430.
- Kakadiaris, I., G. Passalis, G. Toderici, N. Murtuza, and T. Theoharis. 2006. 3D face recognition. In *Proc. British Machine Vision Conference*, 200–208.
- Labib, K., and V. R Vemuri. 2006. An application of principal component analysis to the detection and visualization of computer network attacks. In *Annales des télécommunications*, 61:218.
- Lanitis, A., C. J Taylor, and T. F Cootes. 1997. Automatic interpretation and coding of face images using flexible models. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19, no. 7: 743–756.
- Lawrence, S., C. L Giles, A. C Tsoi, and A. D Back. 1997. Face recognition: A convolutional neural-network approach. *IEEE Transactions on Neural Networks* 8, no. 1: 98–113.
- Leonardis, A., and H. Bischof. 2000. Robust recognition using eigenimages. *Computer Vision and Image Understanding* 78, no. 1: 99–118.
- von der Malsburg, N. K.C, L. Wiskott, and J. M Fellous. 1997. Face recognition by elastic bunch graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19, no. 7: 775–779.
- Martinez, A. M. 2002. Recognizing imprecisely localized, partially occluded, and

- expression variant faces from a single sample per class. *IEEE Transactions on Pattern analysis and machine intelligence*: 748–763.
- Mian, Ajmal S., Mohammed Bennamoun, and Robyn Owens. 2007. Keypoint Detection and Local Feature Matching for Textured 3D Face Recognition. *International Journal of Computer Vision* 79, no. 1 (9): 1-12.
doi:10.1007/s11263-007-0085-5.
<http://www.springerlink.com/index/10.1007/s11263-007-0085-5>.
- Ng, J., and H. Cheung. 2004. Dynamic local feature analysis for face recognition. *Lecture Notes in Computer Science*: 234–240.
- Oja, Erkki, and Juha Karhunen. 1982. New methods for stochastic approximation of truncated Karhunen-Loève transform. *Proceedings, 6th Int. Conf. on Pattern Recognition* (October 19): 550-553.
- Pavlou, M. 2005. Face kernel extraction from local features. Citeseer.
- Pearson, K. 1901. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine* 2, no. 6: 572, 559.
- Penev, P. S, and J. J Atick. 1996. Local feature analysis: A general statistical theory for object representation. *Network: computation in neural systems* 7, no. 3: 477–500.
- Peter W. Hallinan. 1999. *Two- and Three-Dimensional Patterns of the Face*. AK Peters, Ltd., April 21. http://openlibrary.org/b/OL9846946M/Two-_and_Three-Dimensional_Patterns_of_the_Face.
- Pissarenko, D. 2002. *Eigenface-based facial recognition*. Citeseer.
- Shan, S., B. Cao, W. Gao, and D. Zhao. 2002. EXTENDED FISHERFACE FOR FACE RECOGNITION FROM A SINGLE EXAMPLE IMAGE PER PERSON. In *IEEE International Symposium on Circuits and Systems*.
- Smith, M. A, and T. Kanade. 2004. *Multimodal video characterization and summarization*. Kluwer Academic Pub.
- Tahir, N. M, A. Hussain, S. A Samad, and H. Hussain. Reconstruction of Human Posture by PCA.
- Tan, X., S. Chen, Z. H Zhou, and F. Zhang. 2005. Recognizing partially occluded, expression variant faces from single training image per person with SOM

- and soft k-NN ensemble. *IEEE Transactions on Neural Networks* 16, no. 4: 875–886.
- Turing, A. M. 1952. The chemical basis of morphogenesis. *Bulletin of Mathematical Biology* 52, no. 1: 153–197.
- Turk, M. 2005. Eigenfaces and beyond. *Face Processing: Advanced Modeling and Methods*: 55–86.
- Turk, M., and A. Pentland. 1991. Eigenfaces for recognition. *Journal of Cognitive Neuroscience* 3, no. 1: 71-86.
- Wang, H., and S. F Chang. 1997. A highly efficient system for automatic face region detection in MPEG video. *IEEE Transactions on Circuits and Systems for Video Technology* 7, no. 4: 615–628.
- Xudong, Jiang. 2009. Asymmetric Principal Component and Discriminant Analyses for Pattern Classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31 (May).
- Yina, B., X. Baia, Q. Shia, and Y. Suna. 2005. Enhanced Fisherface for Face Recognition*. *Journal of Information & Computational Science* 2, no. 3: 591–595.
- Zhu, Z., and Q. Ji. 2005. Robust real-time eye detection and tracking under variable lighting conditions and various face orientations. *Computer Vision and Image Understanding* 98, no. 1: 124–154.
- Διαμαντάρας, Κωνσταντίνος. 2007. *Τεχνητά Νευρωνικά Δίκτυα*.