



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



ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Ανάλυση και Ανάπτυξη Αναγνώρισης Προσώπου μέσω Raspberry Pi

Του φοιτητή
Ίντριτ Ντοκούζι
Αρ. Μητρώου: 134161

Επιβλέπων καθηγητής
Περικλής Χατζημίσιος

Θεσσαλονίκη
Σεπτέμβριος 2018

Περίληψη

Αυτή η πτυχιακή παρουσιάζει μια συνολική εικόνα του προβλήματος της αναγνώρισης προσώπου. Παρουσιάζεται μια έρευνα διαθέσιμων αλγορίθμων ανίχνευσης προσώπου καθώς και η εφαρμογή και οι δοκιμές διαφορετικών μεθόδων εξαγωγής χαρακτηριστικών. Επίσης γίνεται η ανάλυση και την ανάπτυξη αναγνώρισης προσώπου μέσω μιας μονής πλακέτας υπολογιστή, Raspberry Pi. Αυτή η πτυχιακή έγινε στα πλαίσια του Διαδικτύου Των Πράγματος (Internet of Things – IoT). Στόχος της εργασίας αυτή είναι να δημιουργήσουμε μια έξυπνη μηχανή που θα μπορούσε να χρησιμοποιηθεί παντού , είτε σε μια γωνία του σπιτιού, είτε γραφείο , είτε σε μια επιχείρηση κλπ. Στόχος της είναι να φέρει ασφάλεια και άμεση ενημέρωση στον χρήστη. Η λειτουργία της είναι απλή, αρκεί να είναι συνδεδεμένο στο δίκτυο του σπιτιού, να έχει μια μικρή οθόνη 5” πάνω στο raspberry και τροφοδότηση 5V . Ας πάρουμε ως παράδειγμα ένα γραφείο μιας επιχείρησης που δεν θέλουμε να μπαίνει κρυφά κανείς εννοώ λείπουμε. Αυτό που κάνουμε είναι να το τοποθετήσουμε σε ένα μη ορατό σημείο που να βλέπει την πόρτα του γραφείου. Τότε κάθε φορά που θα μπαίνει κάποιος στο γραφείο θα γίνει αναγνώριση του προσώπου του. Παρακάτω θα δείτε όλα τα κεφάλαια ξεκινώντας από το μηδέν μέχρι το τελικό επιθυμητό αποτέλεσμα.

Ευχαριστίες

Με το τέλος της διπλωματικής αυτής εργασίας θέλω να ευχαριστήσω θερμά τον επιβλέπων καθηγητή μου κ. Περικλής Χατζημίσιος του τμήματος πληροφορικής, για την ευκαιρία που μου έδωσε να ολοκληρώσω την πρακτική μου στο εξωτερικό για να μάθω νέα πράγματα και να μπορώ να ανοίξω νέες πόρτες στη ζωή μου μετά το πτυχίο. Επίσης και το θέμα της πτυχιακής μου ήταν πάνω σε ότι ασχολήθηκα κατά την διάρκεια της πρακτικής μου στο Λονδίνο το οποίο είναι πολύ ενδιαφέρον και έχει αρκετές προοπτικές στην αναζήτηση εργασίας.

Τέλος, ευχαριστώ την οικογένεια μου για την υποστήριξη τους όχι μόνο για την ολοκλήρωση της πρακτικής μου και πτυχιακής μου αλλά και για την συμπαράσταση που μου επέδειξαν όλα αυτά τα χρόνια των σπουδών μου.

ΠΕΡΙΕΧΟΜΕΝΑ

Περίληψη.....	1
Ευχαριστίες.....	1
ΚΕΦΑΛΑΙΟ 1 - Εισαγωγή.....	1
1.1 Κίνητρο.....	1
1.2 Εισαγωγή.....	1
ΚΕΦΑΛΑΙΟ 2 - Εισαγωγή στο Διαδίκτυο των Πραγμάτων (IoT)	1
2.1 Τι είναι το IoT.....	1
2.2 Τι σημαίνει για την επιχείρηση.....	1
ΚΕΦΑΛΑΙΟ 3 - Μηχανή Όρασης	1
3.1 Τι είναι η μηχανή όρασης	1
3.2 Ιστορική αναδρομή Μηχανής Όρασης	2
ΚΕΦΑΛΑΙΟ 4 - Ιστορική αναδρομή και συστήματα αναγνώρισης προσώπου	2
4.1 Λειτουργίες Προσώπου Αναγνώρισης.....	2
4.1.1 Επιβεβαίωση.....	2
4.1.2 Ταυτοποίηση.....	2
4.1.3 Λίστα παρακολούθησης	2
4.2 Αναγνώριση προσώπου Vendor Test 2002	3
4.3 Συμπεράσματα.....	3
ΚΕΦΑΛΑΙΟ 5 – Εργαλεία Ανάπτυξης.....	3
5.1 OpenCV	3
5.2 Raspberry Pi	3
5.3 Λειτουργικό Σύστημα (Raspbian).....	3
5.4 Ρύθμιση Raspberry Pi 3.....	3
5.4.1 Προετοιμασία της microSD κάρτας	3
5.4.2 Εκκίνηση του Raspberry Pi.....	3
5.4.2.1 Μέσω μιας επιτραπέζιας οθόνης	3
5.4.2.2 Μέσω Putty.....	3
5.4.2.3 Μέσω VNC viewer (Virtual Network Computing).....	3
ΚΕΦΑΛΑΙΟ 6 – Υλοποίηση Συστήματος	4
6.1 Δοκιμάζοντας την εγκατάσταση του OpenCV με Python	4
6.2 Ανίχνευση Προσώπου	5
6.2.1 Ανίχνευση προσώπου σε εικόνα	5
6.2.2 Ανίχνευση προσώπου σε πραγματικό χρόνο	5

6.3 Αναγνώριση Προσώπου.....	5
6.3.1 Αναγνώριση προσώπου με την χρήση γεωμετρικών χαρακτηριστικών	5
6.3.2 Αναγνώριση προσώπου με τη χρήση προτύπων ταιριάσματος.....	5
6.4 Προεπεξεργασία	5
6.5 Εξαγωγή Χαρακτηριστικών	5
6.6 Αντιστοίχιση Χαρακτηριστικών.....	5
6.7 Δεδομένα Προσώπου	6
6.7.1 IMM Face Database	6
6.7.2 IMM Frontal Face Database.....	6
6.8 Σετ δεδομένων που χρησιμοποιήθηκαν σε αυτήν την εργασία	6
6.8.1 Σετ 1	6
6.8.2 Σετ 2	6
6.9 Ανίχνευση Προσώπου: Έρευνα.....	6
6.10 Αντιπροσωπευτική εργασία ανίχνευσης προσώπου	6
6.11 Περιγραφή των επιλεγόμενων μεθόδων ανίχνευσης προσώπου.....	6
6.11.1 Γενικές πτυχές των αλγορίθμων ανίχνευσης προσώπων.....	6
6.11.2 EigenFaces.....	6
6.11.3 Fisherfaces	6
6.11.4 Νευρωνικά Δίκτυα	6
6.12 Το Πρακτικό Μέρος της Πτυχιακής.....	6
6.12.1 Επεξήγηση.....	6
6.12.2 Κώδικας.....	7
ΚΕΦΑΛΑΙΟ 7 - Συμπεράσματα και Μελλοντική Εργασία.....	7
7.1 Συμπεράσματα.....	7
7.2 Μελλοντική εργασία.....	7
(α) Ανίχνευση προσώπου	7
(β) Ανακατασκευή 3D Προσώπου.....	7
Βιβλιογραφία	8

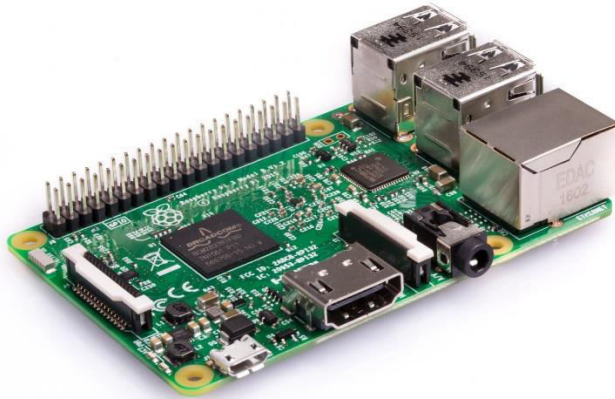
ΚΕΦΑΛΑΙΟ 1 - Εισαγωγή

1.1 Κίνητρο

Με το πέρασμα του χρόνου ο κόσμος έχει δει μια περίεργη εξέλιξη του υπολογιστή. Δέκα με δεκαπέντε χρόνια πριν υπήρχε ένας πόλεμος μεταξύ INTEL και AMD προσπαθώντας να γίνουν οι πρώτοι που πωλούσαν έναν επεξεργαστή των 1 GHz στον εγχώριο χρήστη. Χρόνια αργότερα η διαρροή θερμότητας τέθηκε ως ένα σοβαρό ζήτημα και οι εγχώριες επεξεργαστές άρχισαν να χτυπούν 3GHz και πάνω. Οι επεξεργαστές με δυο, τέσσερις και περισσότερους πυρήνες επεξεργασίας άρχισαν να αποτελούν προτεραιότητα αφού έγινε εκθετικά πιο προφανές ότι θα μπορούσαν να επιτυγχάνουν μεγαλύτερα οφέλη όχι εκτελώντας πράγματα πιο γρήγορα αλλά η πραγματοποίηση πολλών πραγμάτων ταυτόχρονα. Ακόμα και οι επεξεργαστές πολλαπλών πυρήνων με χαμηλότερες συχνότητες από τις προηγούμενες γενιές θα μπορούσαν να αναβαθμιστούν και να βελτιωθούν οι επιδόσεις τους από τους προγραμματιστές που προφέρουν καλή χρήση κωδικοποίησης πολλαπλών νημάτων. Από το δεύτερο μισό της πρώτης δεκαετίας του 2000,πραγματοποιήθηκε μια άλλη περίεργη διαδικασία. Με το λανσάρισμα του πρώτου iPhone της Apple το 2007 και τις μυριάδες των τηλεφώνων Android που ακολουθήσαν, οι φορητοί επεξεργαστές είδαν μια τεράστια ώθηση στη ζήτηση της αγοράς, επιδόσεις και στη χαμηλή κατανάλωση ενέργειας με επεξεργαστές ARM (Acorn RISC Machine) που έγινε επιθυμητή από τις πλήρες μονάδες x86. Τότε σε μικρό χρονικό διάστημα έλαβαν μέρος διάφορα κινητά που είχαν CPU με πολλούς πυρήνες και μερικά GHz. Ενώ αυτοί οι επεξεργαστές κινητών τηλεφώνων υψηλής απόδοσης εξακολουθούν να πωλούνται σε σχετικά υψηλές τιμές, η ταχεία εξέλιξη στον τομέα της πληροφορικής δίνει εναλλακτικές λύσεις χαμηλού κόστους που εξακολουθούν να είναι ικανές να εκτελούν πολύπλοκα καθήκοντα. Αυτό οδήγησε σε τεράστια αύξηση δημοτικότητας για τους single-board υπολογιστές. Αντί της αρθρωτής αρχιτεκτονικής του επιτραπέζιου υπολογιστή, αυτές είναι μη επεκτάσιμες ρυθμίσεις. Εξαιτίας αυτού είναι πολύ μικρού μεγέθους, έχουν πολύ χαμηλές απαιτήσεις ισχύος και ανάλογα με τις προδιαγραφές τους μπορούν επίσης να είναι σε πολύ χαμηλές τιμές. Με την ενσωμάτωση αυτών των νέων επεξεργαστών και την αύξηση της δημοτικότητας των

λειτουργικών συστημάτων που δημιουργήθηκαν για ARM αρχιτεκτονικές, οι single board υπολογιστές έχουν γίνει δεκτά πιο πολύ από ποτέ και το κοινό τους έχει αυξηθεί δραματικά.

Εδώ παίρνει μέρος ένα από τα πιο δημοφιλή συστήματα το οποίο είναι αναμφισβήτητο το Raspberry Pi. Είναι μακράν το πιο φθινό μηχανήμα που μπορείς να βρεις στην αγορά όπου για το βασικό μοντέλο η τιμή του ξεκινάει στα 25 δολάρια, μπορεί να εκτελέσει μια πλήρη, ειδικά καταρτισμένη διανομή Linux. Έχει επίσης μεγάλη επιτυχία. Στην σελίδα του Raspberry Pi Foundation ανακοίνωσε τον Νοέμβρη του 2013 στο επίσημο blog του ότι είχε πουλήσει 2 εκατομμύρια Raspberry Pi. Το γεγονός ότι είναι τόσο ευρέως διαδεδομένο αυξάνει σημαντικά την διαθέσιμη υποστήριξη, για παράδειγμα, λύσεις για προβλήματα είναι εύκολο να βρεθούν και επίσης υπάρχουν πολλές αποσυμπιεσμένες βιβλιοθήκες. Οι λύσεις αυτές υπάρχουν τόσο στην επίσημη ιστοσελίδα του Raspberry pi τόσο όσο και σε διάφορα Forums από διάφορους χρήστες. Υπάρχει μια ποικιλία διανομών Linux που είναι ειδικά συμπιεσμένες για και την υποστήριξη raspberry pi. Παρόλο που οι προδιαγραφές της προσδίδουν μια ιδιαίτερη αξία στη χρήση των μέσων ενημέρωσης, οι λάτρεις το χρησιμοποίησαν πιο πολύ για διαδίκτυο των πράγματων, όπως και για προσωπικούς διακομιστές, ελεγκτές ρομπότ και βεβαίως ως κανονικούς προσωπικούς υπολογιστές. Μονή Πλακέτα (Single board) υπολογιστές είναι, ουσιαστικά, μικροί, φθινοί, πλήρως εξοπλισμένοι υπολογιστές που παρέχουν πολύ καλή ευελιξία, ενώ είναι πολύ προσιτές.



Σχήμα 1 “Raspberry Pi 3”

Τα τελευταία χρόνια υπάρχει μια τεράστια εξέλιξη στα περιφερειακά ενός raspberry pi, δεν είναι μόνο η υπολογιστική ισχύς και η προσβασιμότητα αλλά και τα web-cams. Πρόκειται για μικρές ψηφιακές φωτογραφικές μηχανές που συνήθως χρησιμοποιούνται συχνά για την λήψη και την αναπαραγωγή ακίνητων εικόνων και βίντεο σε υπολογιστές. Συνδέονται συνήθως μέσω USB και είναι ικανά να καταγράφουν βίντεο υψηλής ανάλυσης μέχρι 1080p και στιγμιότυπα πολλών megapixel.

Ο συνδυασμός αυτών των εξελίξεων άνοιξε πρωτοφανή παράθυρα ευκαιριών. Ο συνδυασμός ενός single board υπολογιστή, μιας υψηλής ανάλυσης κάμερα ή μια απλή κάμερα και τις προηγμένες λύσεις λογισμικού που είναι διαθέσιμες σήμερα, είναι θεωρητικά δυνατή να αναπτύξει κανείς ένα μικρό, φορητό, φθινό και ευέλικτου καταγραφής βίντεου και επεξεργασίας βίντεου. Εφαρμογές όπως ανίχνευση κίνησης και η χαρτογράφηση επιπέδου δραστηριότητας μπορούν να γίνουν με χαμηλότερο κόστος και με μεγαλύτερη ευελιξία από ποτέ, δεδομένου ότι οι single board υπολογιστές μοιράζονται την αρχιτεκτονική με κοινούς υπολογιστές [1].

1.2 Εισαγωγή

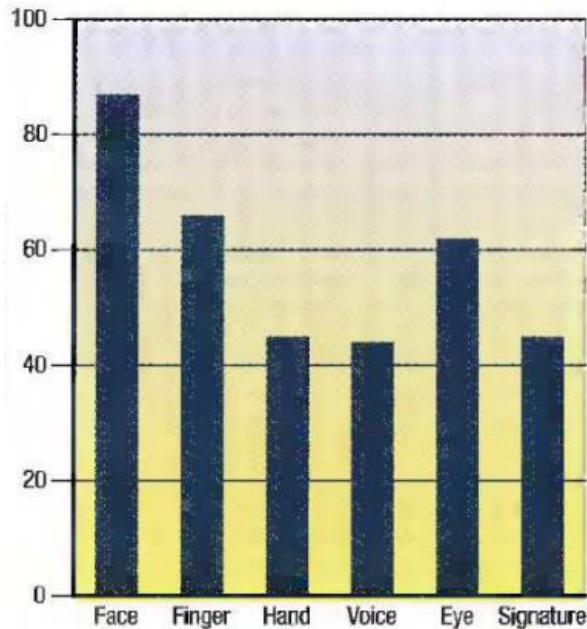
Η αναγνώριση προσώπου είναι μια διεργασία τόσο συνηθισμένη στους ανθρώπους, που το άτομο ούτε καν παρατηρεί το πόσες φορές εκτελείται καθημερινά στη ζωή του. Η αναγνώριση προσώπου ξεκίνησε από την δεκαετία του 1960, πρόσφατα έφτασε στην

προσοχή της επιστημονικής κοινότητας. Την τελευταία δεκαετία έχουν πραγματοποιηθεί πολλές τεχνικές ανάλυσης και προσομοιώσεις προσώπου. Ωστόσο η αξιοπιστία των συστημάτων αναγνώρισης προσώπου συνεχίζει να αποτελεί μεγάλη πρόκληση για την επιστημονική κοινότητα. Η πλαστογράφηση των δελτίων ταυτότητας ή η διείσδυση φυσικών και εικονικών περιέχων 'χακαροντας' αλφαριθμητικούς κωδικούς πρόσβασης, είναι ένα συμβάν που βλέπουμε συχνά στα μέσα μαζικής ενημέρωσης. Αυτά τα είδους προβλήματα έχουν προκαλέσει μια πραγματική αναγκαιότητα για αξιόπιστους, φιλικούς προς τον χρήστη και ευρέως αποδεκτούς μηχανισμούς ελέγχου για την αναγνώριση και επαλήθευση του ατόμου.

Σήμερα μας είναι γνωστή η Βιομετρία, το οποίο έχει αρχίσει να μπαίνει στη ζωή μας τα τελευταία χρόνια. Αρκετές χώρες προσπαθούν να το ενσωματώσουν στα διαβατήρια των ανθρώπων. Η βιομετρία βασίζεται στην εξακρίβωση της ταυτότητας στις εγγενείς πτυχές ενός συγκεκριμένου ανθρωπίνου όντος, εμφανίζεται ως βιώσιμη εναλλακτική λύση σε πιο παραδοσιακές προσεγγίσεις (όπως είναι οι κωδικοί PIN ή κωδικοί πρόσβασης). Μεταξύ των παλαιότερων βιομετρικών τεχνικών είναι η αναγνώριση δακτυλικών αποτυπωμάτων. Αυτή η τεχνική χρησιμοποιήθηκε στην Κίνα ήδη από το 700 μ.Χ. για την επίσημη πιστοποίηση των συμβάσεων. Αργότερα, στα μέσα 19^{ου} αιώνα, χρησιμοποιήθηκε για την αναγνώριση των προσώπων στην Ευρώπη. Μια σύγχρονη Βιομετρική τεχνική είναι η αναγνώριση ίριδας. Αυτή η τεχνική χρησιμοποιείτε τώρα αντί για ταυτοποίηση διαβατηρίου για τακτικά δρομολόγια σε ορισμένα αεροδρόμια στο Ηνωμένο Βασίλειο, Καναδά, και Ολλανδία. Αυτές οι τεχνικές είναι ακατάλληλες λόγω το ότι πρέπει να γίνει πρώτα αναγνώριση και αυθεντικοποίηση του κάθε ατόμου ξεχωριστά. Η αναγνώριση προσώπου από την άλλη πλευρά μπορεί να είναι μη παρεμβατική τεχνική. Αυτός είναι ένας από τους λόγους για τους οποίους η τεχνική αυτή έχει προσελκύσει αυξημένο ενδιαφέρον από την επιστημονική κοινότητα κατά την τελευταία δεκαετία.

Η αναγνώριση προσώπου έχει πολλά πλεονεκτήματα σε σχέση με άλλες βιομετρικές τεχνικές. Είναι φυσικό, μη ενοχλητικό και εύκολο στη χρήση. Σε μια μελέτη που εξετάζει τη συμβατότητα έξι βιομετρικών τεχνικών (πρόσωπο, δάκτυλο, χέρι, φωνή, μάτια, υπογραφή) με αναγνωρίσιμη μηχανή ταξιδιωτικών εγγραφών, τα χαρακτηριστικά του

προσώπου σημείωσαν το υψηλότερο ποσοστό συμβατότητας, Σχήμα 2. Σ 'αυτήν την μελέτη εξετάστηκαν παράμετροι όπως η εγγραφή, η ανανέωση, οι απατήσεις της μηχανής και η δημόσια αντίληψη. Ωστόσο, τα χαρακτηριστικά του προσώπου δεν πρέπει να θεωρούνται ως τα πιο αξιόπιστα βιομετρικά στοιχεία.



Σχήμα 2 “Σύγκριση 6 βιομετρικών στοιχείων”

Ωστόσο, η αυτοματοποιημένη αναγνώριση προσώπου μπορεί να χρησιμοποιηθεί σε πολλές περιοχές εκτός από τις εφαρμογές που προσανατολίζονται στην ασφάλεια όπως η ψυχαγωγία μέσω υπολογιστή και η προσαρμοσμένη αλληλεπίδραση μεταξύ υπολογιστή και ανθρώπου. Οι προσαρμοσμένες εφαρμογές αλληλεπίδρασης μεταξύ υπολογιστή και ανθρώπου θα υλοποιηθούν στο μέλλον όπως στα αυτοκίνητα, βοηθήματα για ανθρώπους με ειδικές ανάγκες, κτίρια κλπ. Το ενδιαφέρον για την αυτοματοποιημένη αναγνώριση προσώπου και ο αριθμός των αιτήσεων πιθανότητα να αυξηθούν ακόμα περισσότερο στο μέλλον . Αυτό θα μπορούσε να επιτευχθεί με την αυξημένη διείσδυση τεχνολογιών, όπως ψηφιακές φωτογραφικές μηχανές και το διαδίκτυο, και να οδηγήσει σε μεγαλύτερη ζήτηση για διαφορετικά συστήματα ασφάλειας.

Αν και οι άνθρωποι είναι αυτοί που δημιουργούν τους πάντες όπως και η αναγνώριση προσώπου ακόμη δεν έχει γίνει κατανοητό πως λειτουργεί αυτή η αναγνώριση. Για πολλά χρόνια οι ψυχοφυσικοί και νεύρο επιστήμονες ερευνούν εάν η αναγνώριση προσώπου γίνεται ολιστικά ή με τοπική ανάλυση χαρακτηριστικών δηλαδή γίνεται αναγνώριση προσώπου εξετάζοντας το πρόσωπο ως σύνολο ή εξετάζοντας τα τοπικά χαρακτηριστικά του προσώπου ανεξάρτητα. Είναι ωστόσο σαφές ότι οι άνθρωποι είναι μόνο ικανοί να κρατούν μια εικόνα προσώπου στο μυαλό σε μια δεδομένη στιγμή. Ένα κλασικό παράδειγμα ψευδαίσθησης που ονομάζεται 'Η γυναίκα μου και η πεθερά', η οποία εισήχθη στην ψυχολογική βιβλιογραφία από τον Edwin G. Boring. Τι βλέπεις? Μια μάγισσα ή μια νεαρή κοπέλα?

ΚΕΦΑΛΑΙΟ 2 - Εισαγωγή στο Διαδίκτυο των Πραγμάτων (IoT)

2.1 Τι είναι το IoT

Το διαδίκτυο των πράγματων μπορεί να μην είναι ένας νέος όρος στο κλάδο της πληροφορικής αλλά σίγουρα αποτελεί ένα καυτό θέμα. Ο Κέβιν Αστον στη δεκαετία του 2000 είχε θέσει το τι πρόκειται να γίνει το Διαδίκτυο των πράγματων στο εργαστήριο AutoID του MIT. Ο Ashton ήταν από τους πρώτους που είχε μια ιδέα απλή αλλά ισχυρή, αυτή η ιδέα έψαχνε τρόπους με τους οποίους η Proctor & Gamble θα μπορούσε να βελτιώσει τις δραστηριότητες της συνδέοντας τις πληροφορίες RFID με το διαδίκτυο. Αν όλα τα αντικείμενα στην καθημερινή ζωή ήταν εξοπλισμένα με αναγνωριστικά και ασύρματη συνδεσιμότητα, αυτά τα αντικείμενα θα μπορούσαν να επικοινωνούν μεταξύ τους και να διαχειρίζονται από τους υπολογιστές[2].

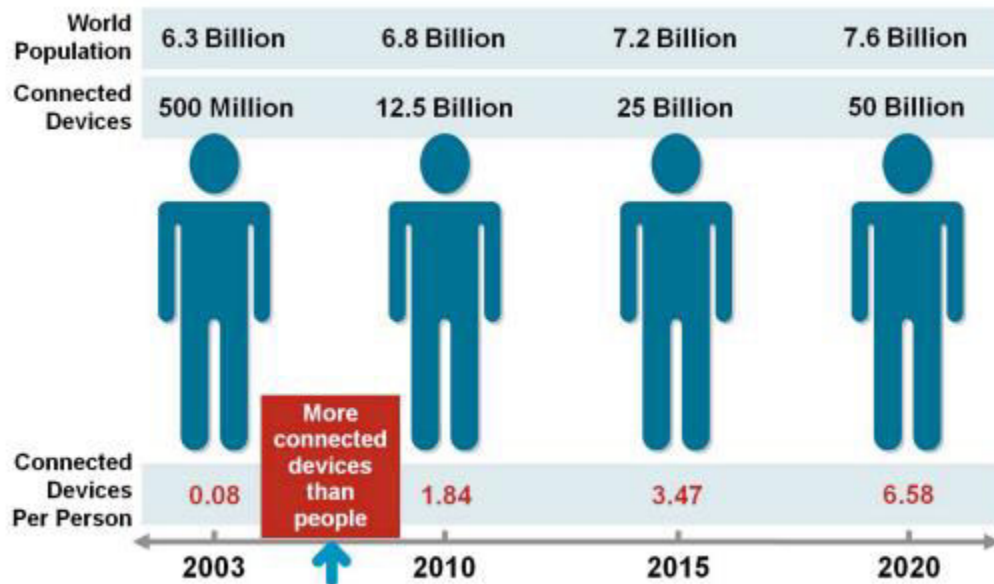
Το 1999 σε άρθρο του RFID, Ο Ashton έγραψε:

“Αν είχαμε υπολογιστές που γνώριζαν τα πάντα για πράγματα που ήταν γνωστά-χρησιμοποιώντας δεδομένα που έχουν συγκεντρώσει χωρίς τη βοήθεια του ανθρώπου-θα μπορούσαμε να εντοπίσουμε και να μετρήσουμε τα πάντα, και να μειώσουμε σημαντικά τα απόβλητα, τις απώλειες και το κόστος. Θα γνωρίζαμε τότε χρειάζεται αντικατάσταση, επισκευή ή ανάκληση. Πρέπει να ενδυναμώσουμε τους υπολογιστές με τα δικά τους μέσα συγκέντρωσης”

Πως θα μπορούσαμε να συνδέσουμε τα πάντα στον πλανήτη? Τι είδους ασύρματες επικοινωνίες θα μπορούσαν να ενσωματωθούν σε συσκευές? Ποιες αλλαγές πρέπει να γίνουν στη υπάρχουσα υποδομή του διαδικτύου για την υποστήριξη δισεκατομμυρίων νέων συσκευών επικοινωνίας? Τι θα μπορούσε να τροφοδοτήσει αυτές τις συσκευές? Τι πρέπει να αναπτυχθεί για να είναι οι λύσεις αποδοτικές?

Το θετικό είναι ότι τα περισσότερα από τα παραπάνω έχουν επιλυθεί. Το μέγεθος και το κόστος των ασυρμάτων έχει μειωθεί σημαντικά. Το IPv6 (Internet Protocol Version 6) μας επιτρέπει να εκχωρήσουμε μια διεύθυνση επικοινωνίας σε δισεκατομμύρια συσκευές. Επίσης πολλές ηλεκτρονικές εταιρείες κατασκευάζουν Wi-Fi και κινητά ασύρματης επικοινωνίας σε ένα ευρύ φάσμα συσκευών. Η κάλυψη δεδομένων κινητής

τηλεφωνίας έχει δείξει μεγάλο πρόοδο σε πολλά δίκτυα που προσφέρουν απίστευτα γρήγορες ταχύτητες. Αν και δεν είναι τέλεια, η αρχιτεκτονική και η τεχνολογία της μπαταρίας έχει βελτιωθεί σημαντικά και η ηλιακή επαναφόρτιση της μπαταρίας έχει ενσωματωθεί σε πολλές συσκευές. Σε μερικά χρόνια θα υπάρχουν δισεκατομμύρια αντικείμενα που θα συνδέονται στο δίκτυο. Για παράδειγμα, η ομάδα του διαδικτύου των πραγμάτων της Cisco προβλέπει ότι μέχρι το 2020 θα υπάρχουν πάνω από 50 δισεκατομμύρια συνδεδεμένες συσκευές [2].



Σχήμα 3 “Συνδεδεμένες συσκευές ανά άτομο με την πάροδο του χρόνου”

Το διαδίκτυο των πραγμάτων ουσιαστικά περιγράφει ένα σύστημα όπου αντικείμενα που υπάρχουν στο φυσικό κόσμο με αισθητήρες που είναι ενσωματωμένα ή εξωτερικά να μπορούν να συνδέονται ασύρματα στο διαδίκτυο παντού. Αυτοί οι αισθητήρες μπορούν να χρησιμοποιήσουν διάφορους τύπους τοπικών συνδέσεων όπως RFID (Radio-Frequency Identification), NFC(Near-Field Communication), Wi-Fi, Bluetooth και Zigbee. Οι αισθητήρες βεβαίως μπορούν να έχουν ευρυζωνική σύνδεση όπως GSM, GPRS, 3G και LTE(4G) .

Το διαδίκτυο των πράγματων θα:

Συνδέσει τόσο τα μη ζωντανά (άψυχα) όσο και τα ζωντανά πράγματα. Οι πρώτες δοκιμές και αναπτύξεις του διαδικτύου στο Διαδίκτυο των Πράγματων ξεκίνησε με την σύνδεση βιομηχανικού εξοπλισμού. Σήμερα, το IoT οραματίζεται να συνδέσει τα πάντα από τον βιομηχανικό εξοπλισμό με τα καθημερινά αντικείμενα. Τα είδη αντικειμένων κυμαίνονται από αεριοστρόβιλους μέχρι και αυτοκίνητα μέτριας χρησιμότητας. Μπορεί επίσης να περιλαμβάνει και ζωντανούς οργανισμούς όπως φυτά, ζώα και ανθρώπους. Ένα παράδειγμα, το πρόγραμμα παρακολούθησης αγελάδων στο Essex χρησιμοποιεί δεδομένα που συλλέγονται από τις ετικέτες προσδιορισμού θέσης που αυτά χρησιμεύουν να ελέγχουν τις αγελάδες από τυχόν αρρώστιες και παρακολούθηση συμπεριφοράς μέσα στο αγέλη. Nike+ Fuel Band και Fitbit είναι συσκευές που χρησιμοποιούνται από τους ανθρώπους. Αυτά είναι παραδείγματα για το πώς οι άνθρωποι συνδέονται στο Διαδίκτυο των Πράγματων. Η Cisco έχει επεκτείνει τον ορισμό του Διαδικτύου των Πράγματων, όπου περιλαμβάνονται οι άνθρωποι, χώροι, και αντικείμενα. Ουσιαστικά οτιδήποτε συνδέεται με έναν αισθητήρα και το διαδίκτυο μπορεί να συμμετάσχει στα νέα συνδεδεμένα οικοσυστήματα.

Χρήση αισθητήρων για συλλογή δεδομένων. Τα φυσικά αντικείμενα που είναι συνδεδεμένα θα διαθέτουν έναν ή περισσότερους αισθητήρες. Κάθε αισθητήρας θα παρακολουθεί μια συγκεκριμένη κατάσταση όπως τοποθεσία, κραδασμοί, κίνηση και θερμοκρασία. Αυτοί οι αισθητήρες θα συνδεθούν μεταξύ τους και στα συστήματα που καταλαβαίνουν ή παρουσιάζουν πληροφορίες από τα δεδομένα αισθητήρων. Αυτοί οι αισθητήρες θα παρέχουν νέες πληροφορίες στα συστήματα μια εταιρίας και στους ανθρώπους.

Τα δεδομένα του διαδικτύου των πράγματων διαφέρουν από το παραδοσιακό υπολογισμό. Τα δεδομένα μπορεί να είναι μικρού μεγέθους και συχνής μετάδοσης. Ο αριθμός των συσκευών ή κόμβων που είναι συνδεδεμένα στο διαδίκτυο είναι πολύ μεγαλύτερος στο IoT από ότι στους παραδοσιακούς υπολογιστές. Οι επικοινωνίες Machine-to-Machine και οι πληροφορίες που προέρχονται από τις συσκευές και το δίκτυο θα επιτρέψουν στις επιχειρήσεις να αυτοματοποιήσουν ορισμένες βασικές λειτουργίες χωρίς να εξαρτώνται από τις κεντρικές εφαρμογές και υπηρεσίες που βασίζονται στο cloud. Αυτά τα χαρακτηριστικά παρουσιάζουν ευκαιρίες για συλλογή ευρέως φάσματος δεδομένων αλλά και προκλήσεις όσον αφορά τον σχεδιασμό της κατάλληλης δικτύωσης και ασφάλειας [2].

2.2 Τι σημαίνει για την επιχείρηση

Το διαδίκτυο επηρεάζει κάθε επιχείρηση. Το κινητό και το διαδίκτυο των πράγματων θα αλλάξουν τους τύπους συσκευών που συνδέονται στα συστήματα μιας εταιρίας. Αυτές οι νέες συνδεδεμένες συσκευές θα παράγουν νέους τύπους δεδομένων. Το Διαδίκτυο των Πράγματων θα βοηθήσει στην αποκόμιση κέρδους από τις επιχειρήσεις, θα αξιοποιήσει τις πληροφορίες από ένα ευρύ φάσμα εξοπλισμού, θα βελτιώσει τις λειτουργίες και θα αυξήσει την ικανοποίηση του πελάτη. Το IoT θα βελτιώσει την δημόσια ασφάλεια, την μεταφορά και την υγειονομική περίθαλψη με καλύτερη πληροφόρηση και γρηγορότερη επικοινωνία αυτής της πληροφορίας. Τέλος, καταλήγουμε ότι το διαδίκτυο των πράγματων φέρνει τουλάχιστον τρία πλεονεκτήματα για τις επιχειρήσεις, αυτές είναι: Επικοινωνία, έλεγχος και εξοικονόμηση χρήματων [2].

ΚΕΦΑΛΑΙΟ 3 - Μηχανή Όρασης

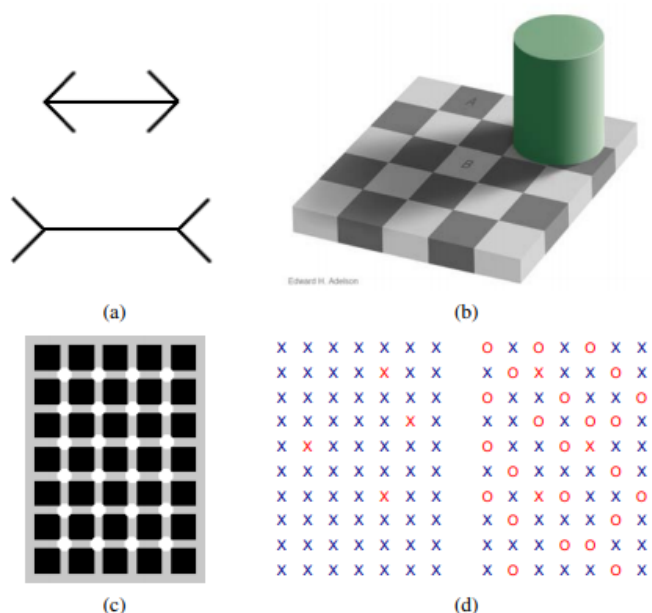
3.1 Τι είναι η μηχανή όρασης

Ως άνθρωποι, αντιλαμβανόμαστε την τρισδιάστατη δομή του κόσμου γύρω μας με προφανή ευκολία. Σκέψου όμως ποσό ζωντανή είναι η τρισδιάστατη αντίληψη όταν βλέπεις ένα βάζο λουλουδιών δίπλα σας στο τραπέζι. Μπορείς να περιγράψεις το σχήμα και την διαφάνεια κάθε πεταλιού μέσα από τις λεπτές μορφές και σκίασης που παίζουν στην επιφάνεια του και αδιατάρακτα μπορείς να ομαδοποιήσεις κάθε λουλούδι από τη πίσω σκηνή της φόντο (Σχήμα 4) [3].



Σχήμα 4 “Εύκολη ανάλυση της φωτό με το ανθρώπινο μάτι”

Κοιτάζοντας ένα πορτραίτο φωτογραφίας όπου είναι μια ομάδα ατόμων εύκολα μπορούμε να μετρήσουμε και να ονομάσουμε όλους τους ανθρώπους που βρίσκονται στην εικόνα και ακόμη να μαντέψουμε τα συναισθήματα τους από την εμφάνιση του προσώπου τους. Οι αντιληπτές ψυχολόγοι έχουν περάσει δεκαετίες προσπαθώντας να καταλάβουν πως λειτουργεί το οπτικό σύστημα και παρόλο που μπορούν να επινοήσουν οπτικές ψευδαισθήσεις για να αποδείξουν μερικές από τις θεωρίες τους, μια ολοκληρωμένη λύση σε αυτό το παζλ παραμένει ασαφής [4].

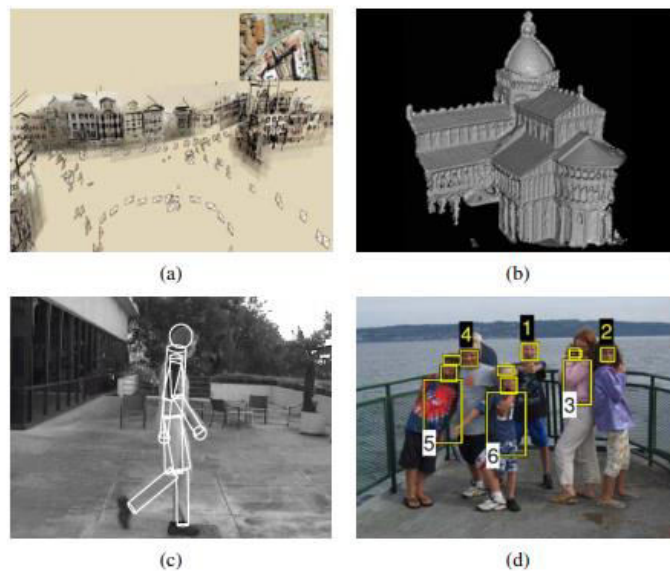


Σχήμα 5

Η παραπάνω εικόνα (a) είναι μια ψευδαίσθηση του Muller-Lyer όπου το μήκος των δυο γραμμών εμφανίζεται ως “διαφορετικό” πράγμα που δεν ισχύει. Στην (b) Το ‘λευκό’ τετράγωνο B στη σκιά και το ‘μαύρο’ τετράγωνο στο φως έχουν στην πραγματικότητα την ίδια τιμή απόλυτης έντασης. Η αντίληψη οφείλεται στην σταθερότητα της φωτεινότητας, στην προσπάθεια του οπτικού συστήματος να αποκλείει τον φωτισμό όταν ερμηνεύει τα χρώματα. Στην (c) Καθώς κουνάμε τα μάτια πάνω στο σχήμα, εμφανίζονται γρίζες κηλίδες πάνω στις διασταυρώσεις. Και (d) μετρήστε τα κόκκινα X στο αριστερό ήμισυ του σχήματος. Τώρα μέτρα τα X στο δεξιό μισό του σχήματος. Είναι πολύ πιο δύσκολο.

Ερευνητές στην μηχανική όραση έχουν ανάπτυξη παράλληλα και μαθηματικές τεχνικές για την ανάκτηση του τρισδιάστατου σχήματος και της εμφάνισης αντικειμένων σε εικόνες. Έχουμε τώρα αξιόπιστες τεχνικές για τον ακριβή υπολογισμό ενός μερικού τρισδιάστατου μοντέλου ενός περιβάλλοντος από χιλιάδες μερικώς επικαλυπτόμενες φωτογραφίες (Σχήμα 6a). Έχοντας ένα μεγάλο σύνολο προβολών ενός συγκεκριμένου αντικειμένου ή πρόσοψης, μπορούμε να δημιουργήσουμε ακριβή μοντέλα 3D επιφάνειας με την χρήση στερεοφωνικών προσαρμογών (Σχήμα 6b). Μπορούμε να παρακολουθήσουμε ένα άτομο που κινείται σε ένα πολύπλοκο υπόβαθρο (Σχήμα 6c). Μπορούμε ακόμη και με μέτρια επιτυχία να προσπαθήσουμε να βρούμε και να ονομάσουμε όλους τους ανθρώπους σε μια φωτογραφία χρησιμοποιώντας έναν

συνδυασμό ανιχνεύσεις και αναγνώρισης προσώπου, ρουχισμού και μαλλιών (Σχήμα 6d). Ωστόσο το όνειρο για πραγματοποίηση μιας μηχανής όρασης που να ερμηνεύσει μια εικόνα όπως ένα παιδί δυο ετών (για παράδειγμα να μετράει όλα τα ζώα σε μια εικόνα) παραμένει αόριστη. Γιατί η όραση είναι τόσο δύσκολη ? Εν μέρη, είναι επειδή η όραση είναι ένα αντίστροφο πρόβλημα, στο οποίο επιδιώκουμε να ανακτήσουμε κάποια άγνωστα δοσμένα δεδομένα για να καθορίσουμε πλήρως τη λύση. Επόμενος πρέπει να καταφύγουμε σε μοντέλα βασισμένα στη φυσική και πιθανολογούμενα μοντέλα για να αποσαφηνίσουμε τις δυνητικές λύσεις [4].

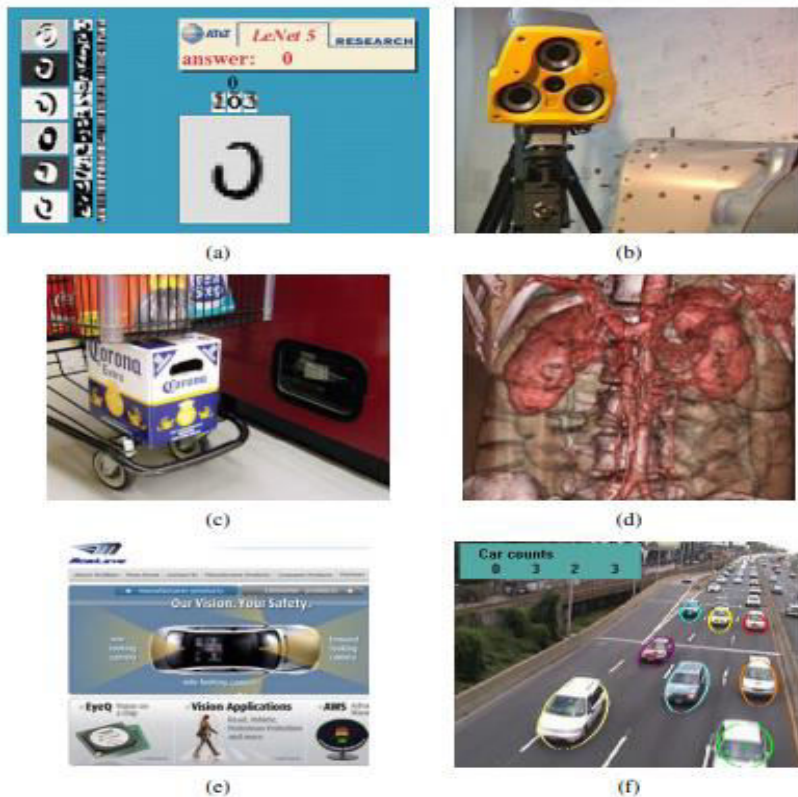


Σχήμα 6 “Μηχανική όραση με διαφορετικούς μηχανισμούς”

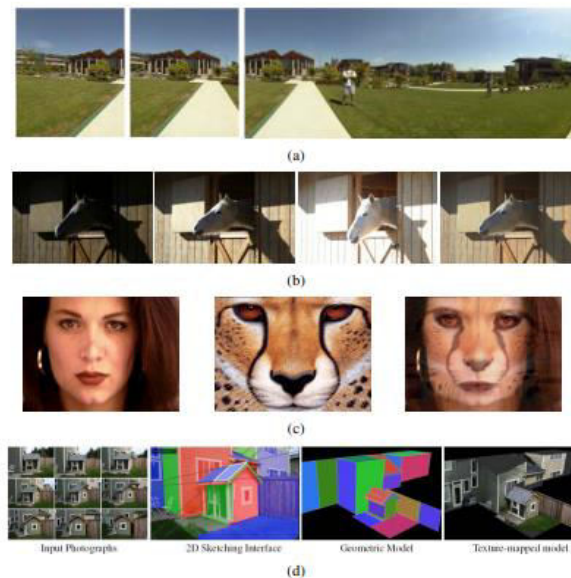
Τα μοντέλα που χρησιμοποιούμε στην μηχανική όραση αναπτύσσονται συνήθως με φυσική (ραδιομετρία, οπτική και σχεδιασμός αισθητήρων) και με γραφικά υπολογιστών. Και τα δυο πεδία αποτελούν πρότυπο για το πώς κινούνται και ζωντανεύουν τα αντικείμενα, πως το φως αντανακλά τις επιφάνειες τους, διασκορπίζεται από την ατμόσφαιρα, διαθλάται μέσω φακών της κάμερας (ή των ανθρωπίνων ματιών) και τελικά προβάλλεται πάνω σε ένα επίπεδο η καμπυλωτή εικόνα. Ενώ τα γραφικά του υπολογιστή δεν είναι ακόμα τέλεια (Δεν έχει πέτυχει ακόμα η δημιουργία ταινιών κινουμένων σχεδίων με ανθρωπίνους χαρακτήρες ώστε να γίνει αναγνώριση για

διαχωρισμό πραγματικού ανθρώπου από τα ρομπότ), σε περιορισμένους τομείς, όπως η δημιουργία μια σκηνής από καθημερινά αντικείμενα ή η ζωντανέψη των εξαφανισμένων πλασμάτων όπως δεινόσαυροι, η ψευδαίσθηση της πραγματικότητας είναι τέλεια [5].

Στην μηχανική όραση, προσπαθούμε να κάνουμε το αντίστροφο, δηλαδή περιγράφουμε τον κόσμο που βλέπουμε σε μια ή περισσότερες εικόνες και να ανασυνδέσουμε τις ιδιότητες του, όπως σχήμα, φωτισμό και οι διανομές χρωμάτων. Είναι καταπληκτικό που οι άνθρωποι και τα ζώα το κάνουν αυτό χωρίς κόπο ενώ οι αλγόριθμοι μηχανικής όρασης είναι τόσο επιρρεπής σε λάθη. Συνήθως άτομο που δεν έχουν ασχοληθεί με τον τομέα αυτό δυσκολεύονται να καταλάβουν το πρόβλημα.



Σχήμα 7 “ Μηχανική Όραση: Αλγόριθμοι και Εφαρμογές”



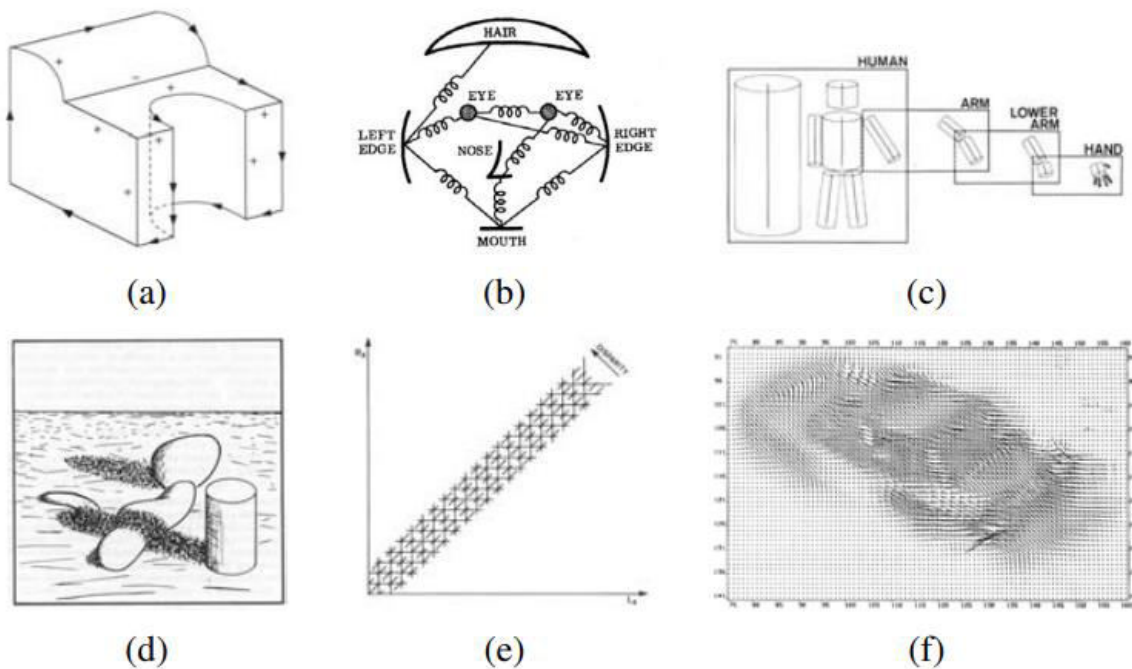
Σχήμα 8 “ Μηχανική Όραση: Αλγόριθμοι και Εφαρμογές”

Το ενδιαφέρον είναι ότι αυτές οι εφαρμογές σε πολλούς φοιτητές είναι οικεία. Είναι τουλάχιστον τεχνολογίες που οι φοιτητές αμέσως μπορούν να εκτιμήσουν και να τα χρησιμοποιήσουν για δικά τους μέσα ενημέρωσης. Δεδομένου ότι η μηχανική όραση είναι ένα δύσκολο θέμα, έχοντας ένα ευρύ φάσμα μαθηματικών και η εγγενώς δύσκολη φύση των προβλημάτων που επιλύονται, η διασκέδαση και τα σχετικά προβλήματα που μπορείς να δουλέψεις είναι ιδιαίτερα ενθαρρυντικά και εμπνευσμένα[3].

3.2 Ιστορική αναδρομή Μηχανής Όρασης

Όταν η μηχανική όραση ξεκίνησε κατά την δεκαετία του 1970, θεωρήθηκε ως μια ανακάλυψη για να μιμηθεί την ανθρώπινη νοημοσύνη και να προμηθεύει τα ρομπότ με έξυπνη συμπεριφορά. Εκείνη την εποχή, οι πρώτοι επιστήμονες της τεχνητής νοημοσύνης πιστεύαν ότι η τοποθέτηση την όρασης θα έκανε την επίλυση των δύσκολων προβλημάτων πιο εύκολο. Το διακεκριμένο μηχανική όραση του υπολογιστή από το ήδη υπάρχον πεδίο της ψηφιακής επεξεργασίας εικόνας ήταν επιθυμητό να ανακτήσει την τρισδιάστατη δομή του κόσμου από τις εικόνες και να τις χρησιμοποιήσει

ως ένα βήμα προς πλήρη κατανόηση της σκηνής. Winston (1975), Hanson και Riseman (1978) παρέχουν δυο ωραίες συλλογές των κλασικών εγγράφων από αυτή την πρώιμη περίοδο. Οι πρώιμες προσπάθειες κατανόησης σκηνής περιλάμβαναν την αφαίρεση των ακρών και τότε η συναγωγή δομής ενός 3D αντικειμένου από την τυπολογική δομή των 2D γραμμών. Διάφοροι αλγόριθμοι γραμμικής επισήμανσης αναπτύχθηκαν εκείνη την εποχή, οι οποίοι είχαν δημιουργηθεί από τους (Huffman 1971, Clowes 1971, Waltz 1975, Rosenfeld, Hummel, and Zucker 1976, Kanade 1980, Nalwa 1993) όπου έχουν δώσει μια καλή ανασκόπηση περ του θέματος αυτού (Σχήμα 9 a) [3].



Σχήμα 9 “Μηχανική Όραση: Αλγόριθμοι και Εφαρμογές”

Επίσης μελετήθηκε η τρισδιάστατη μοντελοποίηση των μη πολυεδρικών αντικειμένων. Μια δημοφιλής προσέγγιση χρησιμοποίησε γενικευμένους κυλίνδρους δηλαδή στερεά εκ περιστροφής και σάρωση κλειστών καμπύλων (Σχήμα 9c). Fischler and Elschlager(1973) την ονόμασαν ελαστική διάταξη διανυσματικών δομών (Σχήμα 9b). Αυτή είναι σήμερα μια από τις προτιμώμενες προσεγγίσεις που χρησιμοποιούνται στην αναγνώριση αντικειμένων.

Μια ποιοτική προσέγγιση για την κατανόηση των εντάσεων και των παραλλαγών σκίασης και η εξήγηση τους από τις επιδράσεις των φαινομένων σχηματισμού εικόνας, όπως ο προσανατολισμός της επιφάνειας και οι σκιές, υπερασπίστηκε ο Barrow και Tenenbaum (1981) στο έγγραφο τους για τις εγγενείς εικόνες (Σχήμα 9d) μαζί με τις σχετικές ιδέες 2 ½ D σκίτσα του Marr(1982). Αυτή η προσέγγιση ξαναβλέπει μια αναβίωση στο έργο του Tappan, Freeman και Adelson(2005). Σε εκείνη την εποχή είχαν αναπτυχθεί περισσότερες προσεγγίσεις στην μηχανική όραση συμπεριλαμβανομένου του πρώτου από πολλούς αλγόριθμους αλληλεπίδρασης στερεοφωνικών χαρακτηριστικών με βάση τα χαρακτηριστικά (Σχήμα 9e) (Dev 1974, Marr and Poggio 1976, Moravec 1977, Marr and Poggio 1979, Mayhew and Frisby 1981, Baker 1982, Barnard and Fischler 1982, Ohta and Kanade 1985, Grimson 1985, Pollard, Mayhew, and Frisby 1985, Prazdny 1985) και αλγόριθμοι οπτικής ροής με βάση την ένταση (Σχήμα 9f) (Horn and Schunck 1981; Huang 1981; Lucas and Kanade 1981; Nagel 1986). Η Πρώρη εργασία στην ανάκτηση της δομής 3D και της κίνησης κάμερας άρχισε περίπου σ' αυτήν την εποχή (Ullman 1979; Longuet-Higgins 1981) [4].

Μεγάλο μέρος της φιλοσοφίας για το πώς η όραση δουλεύει εκείνη τη στιγμή συνοψίζεται στο βιβλίο του David Marr(1982). Συγκεκριμένα, ο Marr προσέθεσε την ιδέα του για τρία επίπεδα περιγραφής ενός οπτικού συστήματος επεξεργασίας πληροφοριών. Αυτά τα τρία επίπεδα είναι:

- 1) Υπολογιστική θεωρία: Ποιος είναι ο στόχος του υπολογισμού και ποιοι είναι οι περιορισμοί που είναι γνωστοί ή μπορούν να αντιμετωπιστούν στο πρόβλημα;
- 2) Αναπαραστάσεις και Αλγόριθμοι: Πως παρουσιάζονται οι πληροφορίες εισόδου, εξόδου και ενδιάμεσου και ποιοι είναι οι αλγόριθμοι που χρησιμοποιούνται για τον υπολογισμό του επιθυμητού αποτελέσματος.
- 3) Υλοποίηση Hardware: Πως αντιστοιχίζονται οι Αναπαραστάσεις και Αλγόριθμοι σε πραγματικό Hardware για παράδειγμα σε ένα σύστημα βιολογικής όρασης ή ένα εξιδεικευμένο κομμάτι πυριτίου.

Αντιστρόφως πως μπορούν να χρησιμοποιηθούν περιορισμοί hardware για να καθοδηγήσουν την επιλογή της Αναπαραστάσης και του αλγορίθμου; Με την χρήση τσιπ γραφικών (GPU) και πολλών αρχιτεκτονικές πολλών πυρήνων για

την όραση μέσω υπολογιστή, αυτή η είδους ερώτηση γίνεται πάλι αρκετά σχετική.

Καταλήγουμε στο ότι μια προσεκτική ανάλυση των προδιαγραφών ενός προβλήματος και των γνωστών περιορισμών από την δημιουργία εικόνων και των προτέρων πρέπει να γίνει παντρειά με αποτελεσματικούς και ισχυρούς αλγορίθμους για να σχεδιαστεί ένας ολοκληρωμένος αλγόριθμος μηχανικής όρασης. Έτσι, βλέπουμε ότι οι οδηγίες του Marrs μας οδηγούν σε πλαισίωση και επίλυση προβλημάτων στο τομέα της πληροφορικής όπως ήταν 25 χρόνια πριν.

Στη δεκαετία του 1980, η προσοχή τέθηκε στις πιο εξελιγμένες μαθηματικές τεχνικές για την εκτέλεση ποσοτικής ανάλυσης εικόνας και σκηνής.

ΚΕΦΑΛΑΙΟ 4 - Ιστορική αναδρομή και συστήματα αναγνώρισης προσώπου

Ο πιο διαισθητικός τρόπος για να καταλάβετε την αναγνώριση προσώπου είναι αν γίνει εξέταση στα κυρία χαρακτηριστικά του προσώπου και να γίνει η σύγκριση με τα ίδια χαρακτηριστικά σε άλλα πρόσωπα. Μερικές από τις πρώτες μελέτες σχετικά με την αναγνώριση προσώπου έγιναν από τον Δαρβίνο και Γαλτον. Το έργο του Δαρβίνου περιλαμβάνει ανάλυση των διάφορων εκφράσεων προσώπου λογού διαφορετικών συναισθηματικών καταστάσεων, όπου ο Γαλτον μελέτησε το προφίλ προσώπου. Ωστόσο, οι πρώτες πραγματικές προσπάθειες για την ανάπτυξη ημιαυτομάτων συστημάτων αναγνώρισης προσώπου ξεκίνησαν τέλη της δεκαετίας του 1960 και στις αρχές του 1970^{οι} οποίοι ήταν βασισμένοι σε γεωμετρικές πληροφορίες. Εδώ έχουν τοποθετηθεί 'καρφίτσες' σε φωτογραφίες που εντοπίζουν τα κυρία χαρακτηριστικά ενός προσώπου όπως τα μάτια, μύτη, και τις γωνίες του στόματος. Οι σχετικές αποστάσεις και γωνίες υπολογίστηκαν από αυτές τις 'καρφίτσες' σε ένα κοινό σημείο αναφοράς και συγκρίθηκαν με τα δεδομένα αναφοράς. Στο Goldstein 1971 δημιουργήθηκε ένα σύστημα 21 υποκειμενικών δεικτών, όπως χρώμα μαλλιών και πάχος χείλους. Αυτοί οι δείκτες αποδείχτηκαν πολύ δύσκολο να αυτοματοποιηθούν λόγω της υποκειμενικής φύσης πολλών από τις μετρήσεις που έγιναν εντελώς με το χέρι [6].

Μια πιο συνεπής προσέγγιση για την αναγνώριση προσώπου έγινε από τον Fischler (1973) και αργότερα από τους Yuille (1992). Αυτή η προσέγγιση μετρά τα χαρακτηριστικά του προσώπου χρησιμοποιώντας πρότυπα μεμονωμένων χαρακτηριστικών του προσώπου και τα χαρτογραφεί σε ένα παγκόσμιο πρότυπο.

Συνοπτικά, οι πιο αναπτυγμένες τεχνικές κατά την διάρκεια των πρώτων σταδίων αναγνώρισης προσώπου επικεντρωθήκαν στην αυτόματη ανίχνευση των επιμέρους χαρακτηριστικών του προσώπου. Τα μεγαλύτερα πλεονεκτήματα αυτών των γεωμετρικών μεθόδων με βάση τα χαρακτηριστικά είναι η ευαισθησία στον φωτισμό και η διαισθητική κατανόηση των εξαγομένων χαρακτηριστικών. Ωστόσο, ακόμη και

σήμερα, οι τεχνικές ανίχνευσης και μέτρησης προσώπου δεν είναι αρκετά αξιόπιστες για την αναγνώριση των γεωμετρικών χαρακτηριστικών βασισμένο στο πρόσωπο αναγνώρισης και οι γεωμετρικές ιδιότητες είναι ανεπαρκείς για την αναγνώριση προσώπου. Λόγω του μειονεκτήματος της αναγνώρισης βάση γεωμετρικών χαρακτηριστικών, η τεχνική έχει εγκαταλειφθεί σταδιακά και έχει καταβληθεί προσπάθεια στην έρευνα ολιστικών τεχνικών βασισμένο σε χρώματα, η οποία έχει δώσει καλύτερα αποτελέσματα. Οι ολιστικές τεχνικές με βάση το χρώμα συγκεντρώνουν ένα σύνολο από διαφορετικές όψεις για να αποκτήσουν μια αντιστοιχία μεταξύ των εντάσεων των εικονοστοιχείων, ένας πλησιέστερος ταξινομητής μπορεί να χρησιμοποιηθεί για να ταξινομήσει νέα πρόσωπα όταν η νέα εικόνα ευθυγραμμίζεται πρώτα με το σύνολο των ήδη ευθυγραμμισμένων εικόνων. Με την εμφάνιση της τεχνικής EigenFaces, μια προσέγγιση στατιστικής μάθησης, αυτή η χονδροειδής μέθοδος ενισχύθηκε σημαντικά. Αντί της άμεσης σύγκρισης των εντάσεων των εικονοστοιχείων των διαφορετικών εικόνων του προσώπου, η διάσταση των εντάσεων εισόδου μειώθηκε αρχικά με ανάλυση κυρίας συνιστώσας (PCA) στην τεχνική EigenFace. Το EigenFace αποτελεί βασική συνιστώσα πολλών από τα συστήματα αναγνώρισης προσώπου που βασίζονται στην εικόνα που χρησιμοποιούνται σήμερα. Μια από τις τρέχουσες τεχνικές είναι το Fisherfaces. Αυτή η τεχνική χρησιμοποιείται ευρέως και είναι προτεινόμενη. Συνδυάζει τις ιδιομορφίες με την ανάλυση γραμμικής διακρίσεως Fisher (FLDA) για να αποκτήσει έναν καλύτερο διαχωρισμό των επιμέρους προσώπων. Σε Fisherfaces, η διάσταση των φορέων έντασης εισόδου μειώνεται με PCA και στη συνέχεια με FLDA όπου εφαρμόζεται για να αποκτήσει μια βέλτιστη προβολή για διαχωρισμού των προσώπων από διαφορετικά άτομα.

Μετά την ανάπτυξη τεχνικής FisherFace έχουν προταθεί πολλές σχετικές τεχνικές. Αυτές οι νέες τεχνικές αποσκοπούν στην παροχή μιας ακόμα καλύτερης προβολής για τον διαχωρισμό των προσώπων από διαφορετικά πρόσωπα. Προσπαθούν να ενισχύουν την ανθεκτικότητα στην αντιγραφή με διαφορά στον φωτισμό και στην εμφάνιση της εικόνας [6].

4.1 Λειτουργίες Προσώπου Αναγνώρισης

Τα τρία κυρία συστήματα αναγνώρισης προσώπου είναι:

- Επαλήθευση (Έλεγχος ταυτότητας) – Είμαι εγώ που ισχυρίζομαι ότι είμαι?
- Ταυτοποίηση (Αναγνώριση) – Ποιος είμαι?
- Λίστα παρακολούθησης – Ψάχνεις για έμενα?

Για την εξέταση των τριών καθηκόντων που περιγράφονται παραπάνω πρέπει να εφαρμοστούν διαφορετικά συστήματα.

Ποιο σχήμα πρέπει να χρησιμοποιηθεί εξαρτάται από την φύση της εφαρμογής [7].

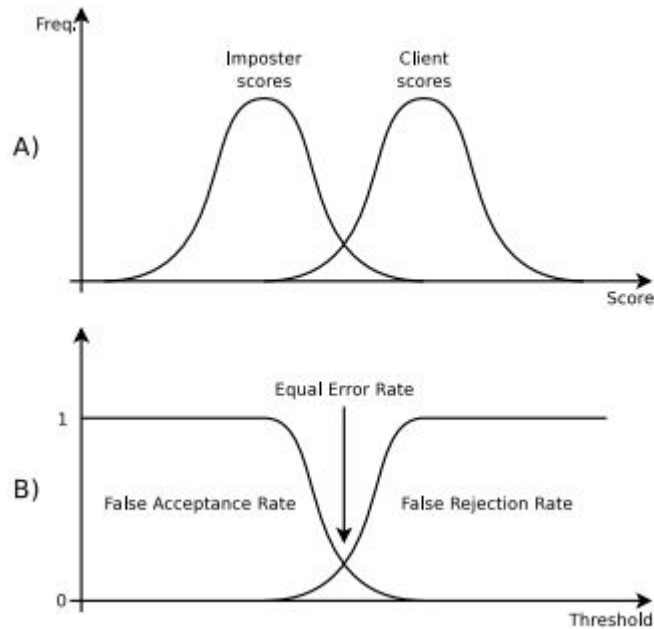
4.1.1 Επιβεβαίωση

Η εργασία επιβεβαίωσης απευθύνεται σε εφαρμογές που απαιτούν αλληλεπίδραση χρήστη με τη μορφή αξίωσης ταυτότητας, δηλαδή εφαρμογών πρόσβασης. Η δοκιμή επαλήθευσης διεξάγεται διαιρώντας τα πρόσωπα σε δυο ομάδες:

-Πελάτες, άνθρωποι που προσπαθούν να αποκτήσουν πρόσβαση χρησιμοποιώντας τη δική τους ταυτότητα.

-Απατεώνες, οι άνθρωποι που προσπαθούν να αποκτήσουν πρόσβαση χρησιμοποιώντας μια ψεύτικη ταυτότητα.

Το ποσοστό των απατεώνων που αποκτούν πρόσβαση αναφέρεται ως False Acceptance Rate (FAR) και το ποσοστό απόρριψης του πελάτη αναφέρεται ως ποσοστό ψευδούς απόρριψης (False Rejection Rate - FRR) για ένα συγκεκριμένο όριο. Μια απεικόνιση αυτού φαίνεται στο Σχήμα 10 [8].



Σχήμα 10 “Διάγραμμα ποσοστού πρόσβασης και μη πρόσβασης απατεώνων σε ένα σύστημα”

4.1.2 Ταυτοποίηση

Η διεργασία ταυτοποίησης απευθύνεται κυρίως σε εφαρμογές που δεν απαιτούν αλληλεπίδραση χρήστη, δηλαδή εφαρμογές επιτήρησης. Η δοκιμασία αναγνώρισης λειτουργεί από την υπόθεση ότι όλα τα πρόσωπα στη δοκιμασία είναι γνωστά πρόσωπα. Το ποσοστό των σωστών αναγνωρίσιμων αναφέρεται στη συνέχεια ως το ποσοστό ορθής ταυτοποίησης ή το ποσοστό των ψευδών ταυτοτήτων αναφέρεται ως ποσοστό ψευδούς αναγνώρισης [9].

4.1.3 Λίστα παρακολούθησης

Η διεργασία λίστας παρακολούθησης είναι μια γενίκευση της διεργασίας ταυτοποίησης που περιλαμβάνει άγνωστους ανθρώπους. Η δοκιμή της λίστας παρακολούθησης είναι σαν τη δοκιμή αναγνώρισης που αναφέρθηκε στο ποσοστό ορθής ταυτοποίησης ή στο

ποσοστό ψευδούς αναγνώρισης, αλλά μπορεί να έχει FAR (False Acceptance Rate) και FRR(False Rejection Rate) που σχετίζονται με αυτήν για να περιγράψει την ευαισθησία της λίστας παρακολούθησης δηλαδή ποσό συχνά είναι ένα άγνωστο ταξινομημένο ως άτομο στη λίστα παρακολούθησης (FAR) [9].

4.2 Αναγνώριση προσώπου Vendor Test 2002

Το 2002, η δοκιμή αναγνώρισης προσώπου Vendor Test 2002 εξέτασε μερικά από τα καλύτερα εμπορικά συστήματα αναγνώρισης προσώπου για την απόδοση τους στις τρεις κυρίες διεργασίες αναγνώρισης προσώπου που περιγράφονται στην ενότητα 4.1. Αυτή η δοκιμή χρησιμοποίησε 121589 εικόνες προσώπου από μια ομάδα 37437 διαφορετικών ανθρώπων. Τα διάφορα συστήματα που συμμετέχουν στη δοκιμή παρατίθενται στο σχήμα 11. Η αξιολόγηση πραγματοποιήθηκε σε εύλογα ελεγχόμενες συνθήκες εσωτερικού φωτισμού.

Company	Web site
AcSys Biometrics Corp	http://www.acsysbiometricscorp.com
C-VIS GmbH	http://www.c-vis.com
Cognitec Systems GmbH	http://www.cognitec-systems.com
Dream Mirh Co., Ltd	http://www.dreammirh.com
Eyematic Interfaces Inc.	http://www.eyematic.com
Iconquest	http://www.iconquesttech.com
Identix	http://www.identix.com
Imagis Technologies Inc.	http://www.imagistechnologies.com
Viisage Technology	http://www.viisage.com
VisionSphere Technologies Inc.	http://www.visionspheretech.com

Σχήμα 11 “Συμμετέχοντες στην Αναγνώριση Προσώπου του Vendor Test 2002”

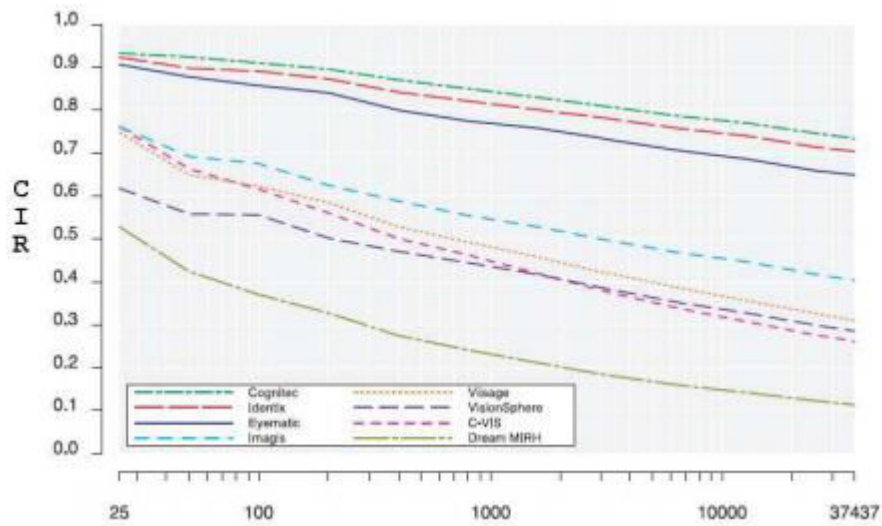
Τα συστήματα που παρέχουν τα καλύτερα αποτελέσματα στο vendor test παρουσιάζουν τα χαρακτηριστικά που παρατίθενται στο σχήμα 12.

Tasks	CIR	FRR	FAR
Identification	73%		
Verification		10%	1%
Watch list	56% to 77% ²		1%

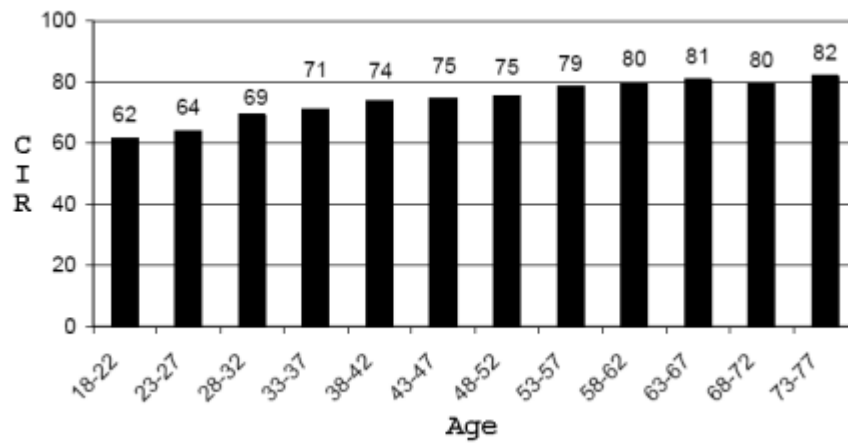
Σχήμα 12 “ Ποια από τα συστήματα είχε μεγαλύτερη επιτυχία στην αναγνώριση προσώπου στο Vendor Test 2002”

Τα επιλεγμένα συμπεράσματα από το Face Recognition Vendor test 2002 είναι:

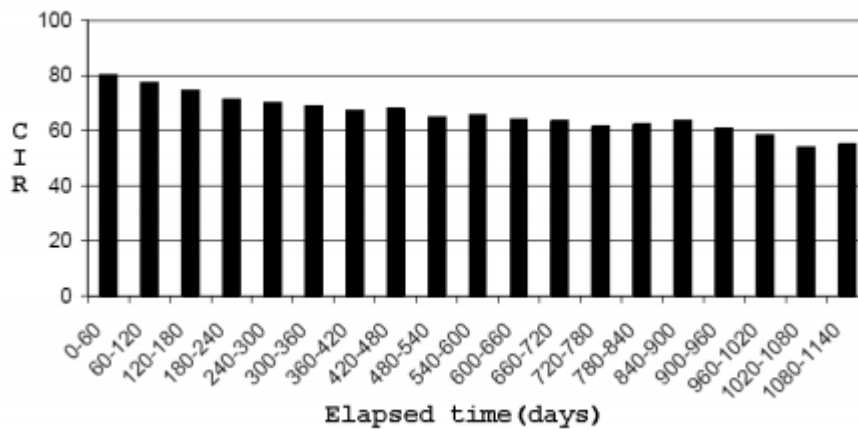
- Διεργασία αναγνώρισης αποφέρει καλύτερα αποτελέσματα για μικρότερες βάσεις δεδομένων. Η διεργασία αναγνώρισης έδωσε υψηλότερη βαθμολογία στη μικρότερη βάση δεδομένων που χρησιμοποιήθηκε. Η απόδοση αναγνώρισης παρουσίασε γραμμική μείωση σε σχέση με τον λογάριθμο του μεγέθους της βάσης δεδομένων. Για κάθε διπλασιασμό του μεγέθους των επιδόσεων της βάσης δεδομένων μειώθηκε κατά 2% έως 3% (Σχήμα 13).
- Τα συστήματα αναγνώρισης προσώπου έδειξαν την τάση να εντοπίζονται ευκολότερα οι ηλικιωμένοι από τους νεότερους. Τα τρία συστήματα με τις καλύτερες επιδόσεις εμφάνισαν μια μέση αύξηση των επιδόσεων κατά περίπου 5% για κάθε δέκα χρόνια γίνεται αύξηση ηλικίας του δοκιμαστικού πληθυσμού. Βλέπε σχήμα 14.
- Όσο περισσότερο χρόνο διαρκεί από την εκπαίδευση του συστήματος έως την παρουσίαση μιας νέας 'ενημερωμένης' εικόνας ενός ατόμου, τόσο η απόδοση της αναγνώρισης μειώνεται. Για τα τρία συστήματα με τις καλύτερες επιδόσεις, σημειώθηκε μέση μείωση περίπου 5%. Βλέπε σχήμα 15



Σχήμα 13



Σχήμα 14



Σχήμα 15

4.3 Συμπεράσματα

Το ενδιαφέρον της έρευνας είναι ότι τα αποτελέσματα από το Test Vendor 2002 δείχνουν υψηλότερη απόδοση αναγνώρισης των ηλικιωμένων σε σχέση με τους νεότερους. Επιπλέον, τα αποτελέσματα δείχνουν ότι γίνεται ακόμα πιο δύσκολο να εντοπιστούν οι άνθρωποι με την πάροδο του χρόνου καθώς το ανθρώπινο πρόσωπο μεταβάλλεται συνεχώς με την πάροδο του χρόνου. Τα αποτελέσματα της αναγνώρισης προσώπου στο Test Vendor 2002, που αναφέρεται στο σχήμα 12, είναι δύσκολο να ερμηνευθεί και συγκριθεί με άλλες δοκιμές, καθώς η αλλαγή στο πρωτόκολλο δοκιμής ή τα δεδομένα δοκιμών θα αποδώσει διαφορετικά αποτελέσματα. Ωστόσο, αυτά τα αποτελέσματα παρέχουν μια ένδειξη της απόδοσης των συστημάτων αναγνώρισης προσώπου.

ΚΕΦΑΛΑΙΟ 5 – Εργαλεία Ανάπτυξης

5.1 OpenCV

Το OpenCV είναι ένα σύνολο βιβλιοθηκών, γραμμένο σε C++, που υλοποιούν αλγόριθμους μηχανικής οράσεως και μηχανικής μάθησης. Περιλαμβάνει, μεταξύ πολλών άλλων, ανίχνευση και αναγνώριση προσώπου, διόρθωση εικόνας, ανάλυση κίνησης, παρακολούθηση αντικειμένων και ανίχνευση αντικειμένων. Το OpenCV ξεκίνησε ως ερευνητικό πρόγραμμα της Intel. Αυτή τη στιγμή τη διατηρεί ο OpenCV.org, ένας μη κερδοσκοπικός οργανισμός. Λόγω του εξαιρετικά ευρέως πεδίου εφαρμογής της, το OpenCV αναπτύσσεται σε όλο τον κόσμο και χρησιμοποιείται από διάφορους οργανισμούς, καθιστώντας το de facto βιβλιοθήκη λογισμικού για την ανίχνευση αντικειμένων. Υπάρχουν επίσης αρκετές επεκτάσεις της λειτουργικότητας του (που έχουν OpenCV ως απαίτηση), όπως το BazAR και το LibPaBOD.

Το OpenCV υλοποιεί περισσότερους από δυο χιλιάδες αλγόριθμους και έχει C,C++ και Python διεπαφές. Είναι ανοικτού κώδικα και κυκλοφορεί υπό την άδεια BSD, που σημαίνει ότι μπορεί να μεταγλωττιστεί σε νέες πλατφόρμες και οι λεπτομέρειες των υλοποιήσεων μπορούν να αναλυθούν. εξαιτίας αυτού, το OpenCV έχει ήδη συνταχθεί για χρήση στο Raspberry Pi. Δεδομένου ότι εξακολουθεί να είναι ένα μεγάλο πακέτο λογισμικού, εάν η διαδικασία σύνταξης εκτελείται στο ίδιο το Raspberry pi, ο χρόνος σύνταξης μπορεί να διαρκέσει αρκετές ώρες. Αυτό δεν αποτελεί πρόβλημα, ωστόσο, καθώς η διασταύρωση είναι επίσης μια επιλογή με αρκετά μαθήματα διαθέσιμα σε όλο το διαδίκτυο [9].

5.2 Raspberry Pi

Το Raspberry Pi είναι μια σειρά χαμηλού κόστους, μικρού μεγέθους, single-board υπολογιστή που αναπτύχθηκε από το Raspberry Pi Foundation στο Ηνωμένο Βασίλειο. Η πρόθεση πίσω από την δημιουργία του Raspberry Pi ήταν να προωθήσει τη διδασκαλία βασικών δεξιοτήτων πληροφορικής στα σχολεία και το έχει πέτυχει αυτό με επιτυχία. Το Raspberry Pi έχει επεκτείνει τα αποτυπώματα του πολύ πέρα από τον επιδιωκόμενο σκοπό του διεισδύοντας στην αγορά των ενσωματωμένων συστημάτων και έρευνας. Τα μοντέλα Raspberry Pi είναι –A, A+, B και B+ τα οποία βασίζονται στο σύστημα SoC (System on a chip), Broadcom BCM2835, το οποίο περιλαμβάνει CPU ARM11 700MHz. Η RPi δεύτερης γενιάς χρησιμοποιεί τετραπύρηνο ARM Cortex-A7, το πρώτο πολύ πύρηνο Raspberry Pi. Τα Raspberry Pi A και B χρησιμοποιούν κάρτες SD για εκκίνηση και μόνιμη αποθήκευση, εννοώ τα μοντέλα A+, B+ και Pi 2 χρησιμοποιούν κάρτες microSD. Τα μοντέλα A και A+ έχουν 256 MB μνήμη RAM, τα B και B+ έχουν 512MB μνήμης RAM και το Pi 2 διαθέτει 1 GB μνήμης RAM [14].

Μέχρι τώρα υπάρχουν έξι μεγάλα μοντέλα Raspberry Pi, αυτά είναι:

- Μοντέλο A
- Μοντέλο A+
- Μοντέλο B
- Μοντέλο B+
- Raspberry Pi2
- Raspberry Pi3

Το Raspberry Pi Foundation παρέχει λογισμικό Debian και Linux ARM για λήψη. Η Python είναι η κυρία πλατφόρμα προγραμματισμού και γλώσσες όπως C, C++, Java, Perl, Ruby που μπορούν να χρησιμοποιηθούν για να προγραμματισμό στο Raspberry Pi. Που στην δική μου περίπτωση έχω χρησιμοποιήσει Python.

Για την μηχανή όρασης μπορούμε να χρησιμοποιήσουμε όλα τα μοντέλα Raspberry Pi. Εδώ χρησιμοποιήθηκε Raspberry Pi 3 [9].

Raspberry Pi 3 Specifications

SoC: Broadcom BCM2837

CPU: 4× ARM Cortex-A53, 1.2GHz

GPU: Broadcom VideoCore IV

RAM: 1GB LPDDR2 (900 MHz)

Networking: 10/100 Ethernet, 2.4GHz 802.11n wireless

Bluetooth: Bluetooth 4.1 Classic, Bluetooth Low Energy

Storage: microSD

GPIO: 40-pin header, populated

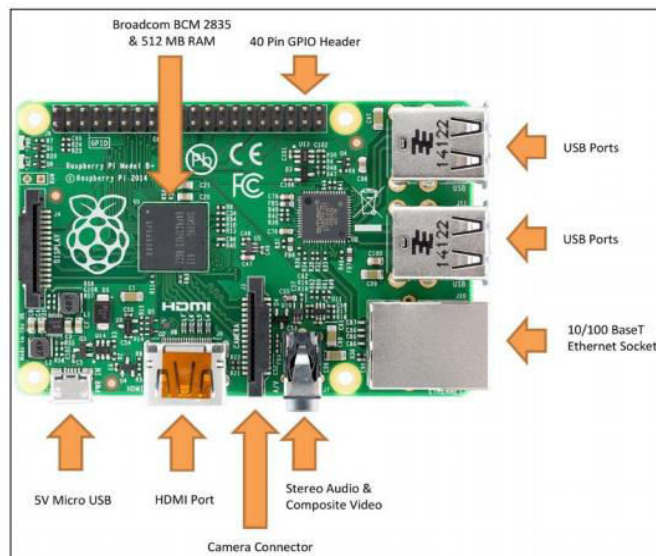
Ports: HDMI, 3.5mm analogue audio-video jack, 4× USB 2.0, Ethernet, Camera

5.3 Λειτουργικό Σύστημα (Raspbian)

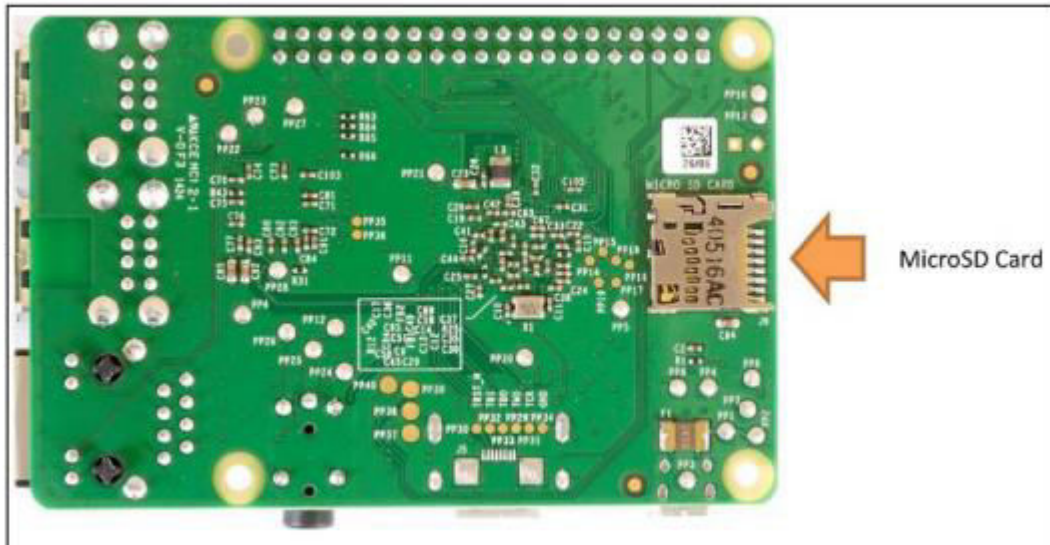
Το Raspbian είναι μια ανεπίσημη παραλλαγή του Debian Wheezy armhf που έχει συνταχθεί για δύσκολο κώδικα που θα τρέχει σε Raspberry Pi υπολογιστή. Πρόκειται για ένα δωρεάν λογισμικό βασισμένο στο Debian που είναι βελτιστοποιημένο για λογισμικό του Raspberry Pi [9].

5.4 Ρύθμιση Raspberry Pi 3

Η μπροστινή όψη του Raspberry Pi



Σχήμα 16 “ Η Μπροστινή όψη του Raspberry Pi”



Σχήμα 17 “Η πίσω όψη του Raspberry Pi”

Για να γίνει η εγκατάσταση λογισμικού στο Raspberry Pi θα χρειαστούμε:

- Raspberry Pi 3
- Τροφοδότηση (είτε με καλώδιο micro Usb, είτε φορτιστή οπου θα μεταφέρει τροφοδότηση 5^A)
- Πληκτρολόγιο
- Κάρτα μνήμης τουλάχιστον 4GB
- Ποντική
- Επιτραπέζια οθόνη με HDMI είσοδο
- Έναν υπολογιστή οπου θα γίνει η εγγραφή του λογισμικού στη κάρτα μνήμης.

5.4.1 Προετοιμασία της microSD κάρτας

Ακόμα και για κάποιον που δεν έχει ασχοληθεί με Raspberry Pi του είναι πιο εύκολο από ποτέ η προετοιμασία την κάρτας μνήμης. Υπάρχουν πολλά βοηθήματα παντού στο διαδίκτυο αλλά εν συντομία θα δώσω ένα παράδειγμα εδώ.

Πρώτα από όλα θα πρέπει να βρούμε την εφαρμογή οπού θα κάνει εγκατάσταση του λογισμικού στην κάρτα μνήμης. Ένα τέτοιου είδους αρχείου είναι το Win32 Disk Imager το οποίο είναι διαθέσιμο προς λήψη στο διαδίκτυο. Αφού γίνει λήψη του αρχείου τότε το κάνουμε εγκατάσταση στον υπολογιστή μας.

Το επόμενο βήμα είναι να βρούμε το λογισμικό που θέλουμε να εγκαταστήσουμε στην κάρτα μνήμης, εμείς θέλουμε το Raspbian. Αυτό είναι διαθέσιμο προς λήψη στην ιστοσελίδα του Raspberry Pi. Το κατεβάζουμε και το κάνουμε αποσυμπίεση το οποίο θα είναι σε μορφή img.

Αμέσως μετά τρέχουμε το πρόγραμμα Win32Imager.exe και γράφουμε το img αρχείο στην κάρτα μνήμης.

Αφού ολοκληρωθεί το βγάζουμε και το τοποθετούμε στην θύρα κάρτας μνήμης του Raspberry Pi.

5.4.2 Εκκίνηση του Raspberry Pi

Αφού ολοκληρώσαμε την εγκατάσταση λογισμικού στην κάρτα μνήμης τότε θα χρειαστούμε μια επιτραπέζια οθόνη με HDMI είσοδο. Αφού το συνδέσουμε με το raspberry Pi την οθόνη τότε θα έχουμε γραφικό περιβάλλον Raspbian.

Για να έχουμε πλήρη έλεγχο πρέπει να είναι συνδεδεμένα στο Raspberry pi πληκτρολόγιο και ποντική επίσης.

Στη συνέχεια θα ζητάει όνομα χρήστη και κωδικό. Αυτά τα δίνει στις οδηγίες του raspberry pi στο βιβλιάριο που παρέχεται μέσα στο κουτί. Για λόγους ασφάλειας ο κωδικός πρέπει να αλλάχτεί συχνά, να μην αφήσετε το default δηλαδή. Αφού γίνει είσοδος στο Desktop του Raspbian τότε πρέπει να τρέξουμε το τερματικό για να κάνουμε ενημέρωση και αναβάθμιση.

Αφού γίνουν αυτά μέσω του τερματικού δίνουμε την εντολή raspi-config για να κάνουμε κάποιες απαραίτητες ρυθμίσεις. Η κυριότερη ρύθμιση είναι να γίνει επέκταση του συστήματος αρχείων (Expand Filesystem).

5.4.2.1 Μέσω μιας επιτραπέζιας οθόνης

Η εκκίνηση του Raspberry Pi μέσω επιτραπέζιας οθόνης είναι η πιο εύκολη λύση. Η διαδικασία εκκίνησης περιγράφεται παραπάνω. Το Raspberry pi εδώ χρησιμοποιείτε σαν επιτραπέζιος υπολογιστής. Μπορείς να κάνεις τα πάντα όπως σε ένα κανονικό υπολογιστή. Τι γίνεται όμως σε περίπτωση που κάποιος δεν έχει επιτραπέζια οθόνη αλλά μόνο έναν φορητό υπολογιστή. Αυτό δείτε το στο επόμενο υποκεφάλαιο.

5.4.2.2 Μέσω Putty

Putty είναι μια εφαρμογή που αποκτάς πρόσβαση στο raspberry pi εισάγοντας το IP δικτύου. Προσοχή ! Πρέπει να είναι στο ίδιο δίκτυο αν η σύνδεση είναι ασύρματη ή σε περίπτωση που δεν έχει συνδεθεί στο Wi-Fi λόγω το ότι δεν μπορείς να το ρυθμίσεις χωρίς οθόνη τότε η λύση είναι μέσω καλωδίου Ethernet μπορείς να αποκτήσεις πρόσβαση. Ένα κόλπο να ενεργοποιήσεις το ssh για να αποκτήσεις πρόσβαση στο raspberry Pi χωρίς οθόνη είναι να δημιουργήσεις ένα κενό αρχείο μέσα στα αρχεία της κάρτας μνήμης με ονομασία ssh. Αυτό το κάνεις εισάγοντας την κάρτα σε έναν οποιοδήποτε υπολογιστή. Αφού γίνουν αυτά μέσω του ip αποκτάς πρόσβαση στο raspberry Pi. Ο χώρος που θα εμφανιστεί είναι ένα τερματικό όπου μπορείς να δουλέψεις χωρίς να χρειάζεται να έχεις ορατό το desktop. Είναι ένας τρόπος που μπορείς να δουλέψεις τις εργασίες σου όπου και να είσαι, αρκεί να έχεις μαζί σου φορητό υπολογιστή και το raspberry pi μαζί με Ethernet καλώδιο.

5.4.2.3 Μέσω VNC viewer (Virtual Network Computing)

Σαυτό το υποκεφάλαιο θα εξηγήσω έναν ακόμη εύκολο τρόπο για να έχετε το desktop του raspberry pi στην οθόνη του φορητού υπολογιστή. Πρώτα από όλα πρέπει να τρέχει το Putty και να γίνει ο έλεγχος του raspberry pi από κει. Αφού έχει γίνει αυτό πρέπει να γίνει εγκατάσταση του VNC server στο Raspberry Pi. Μόλις γίνει εγκατάσταση πρέπει να το τρέξουμε με την εντολή `tightvncserver` και θα μας εμφάνιση τον αριθμό θύρας σύνδεσης το οποίο είναι 1 συνήθως.

Αφήνοντας το Putty ανοιχτό τρέχουμε το VNC στο φορητό υπολογιστή οπού επίσης το έχουμε κάνει εγκατάσταση. Όταν τρέξει τότε πρέπει να δημιουργήσουμε ένα δικό μας κωδικό για λογούς ασφάλειας ώστε να γίνει επιβεβαίωση κάθε φορά που θέλουμε να συνδεθούμε. Αφού έχουν γίνει αυτά πάμε στην τελική ευθεία για να κάνουμε την σύνδεση, βάζουμε ip:1 και συνδεόμαστε στο raspberry pi και τότε έχουμε στην οθόνη του υπολογιστή μας Raspbian περιβάλλον. Το διαφορετικό εδώ σε σχέση με το Putty είναι ότι εδώ βλέπουμε την κάμερα σε ζωντανή λήψη ενώ στο Putty μόνο τρέχουμε την κάμερα αλλά δεν μπορούμε να δούμε ζωντανά την λήψη. Γίνεται απλά αποθήκευση. Πρέπει να συνδεθούμε σε μια οθόνη για να δούμε το βίντεο.

ΚΕΦΑΛΑΙΟ 6 – Υλοποίηση Συστήματος

6.1 Δοκιμάζοντας την εγκατάσταση του OpenCV με Python

Τώρα έχουμε μια μηχανή που τρέχει Raspbian OS. Πρέπει να βεβαιωθούμε ότι η ενσύρματη σύνδεση στο διαδίκτυο είναι ενεργή με λογική ταχύτητα για αυτήν την δραστηριότητα. Ας προετοιμάσουμε το Pi για μηχανική όραση:

- 1) Σύνδεση του Pi μέσω ασύρματης σύνδεσης ή με ένα καλώδιο Ethernet μέσω του δρομολογητή.
- 2) Τρέξετε την παρακάτω εντολή για να γίνει επανεκκίνηση του δικτύου:

-sudo service networking restart

- 3) Βεβαιωθείτε ότι το Raspberry Pi είναι συνδεδεμένο στο διαδίκτυο πληκτρολογώντας αυτή την εντολή:

-ping -c4 www.google.com

- 4) Με τη σειρά εκτελείστε τις παρακάτω εντολές:

-Το Advanced Package Tool (apt) είναι ένα εργαλείο που χρησιμοποιείτε για να κάνει εγκατάσταση και διαγραφή του λογισμικού Debian και τις παραλλαγές του. Για την ενημέρωση λογισμικού χρησιμοποιούμε την παρακάτω εντολή:

-sudo apt-get update

-Αυτή η εντολή συγχρονίζει τη λίστα των πακέτων από την πηγή. Όλα τα πακέτα ανανεώνονται. Αυτή η εντολή πρέπει να πραγματοποιηθεί πριν την εντολή αναβάθμισης.

-sudo apt-get upgrade

-Η παρακάτω εντολή θα εγκαταστήσει τις πιο πρόσφατες εκδόσεις του ήδη εγκατεστημένου λογισμικού. Τα παρωχημένα πακέτα/βοηθητικά προγράμματα δεν καταργούνται αυτόματα. Αν το λογισμικό είναι ήδη ενημερωμένο τότε μένει όπως είναι.

-sudo rpi-update

-Αυτή η εντολή χρησιμοποιείται για την αναβάθμιση firmware. Ο πυρήνας και το firmware εγκαθίστανται ως πακέτα του Debian, και ως εκ τούτου, θα λάβουμε επίσης τις ενημερώσεις. Αυτά τα πακέτα ενημερώνονται σπάνια μετά από εκτεταμένες δοκιμές.

- 5) Τώρα θα πρέπει να εγκαταστήσουμε μερικά απαραίτητα πακέτα για το OpenCV. Στο παρακάτω πίνακα υπάρχει μια λίστα πακέτων που πρέπει να εγκαταστήσουμε. Πρώτα πρέπει να ελέγξουμε αν το Pi είναι συνδεδεμένο στο διαδίκτυο και για να γίνει η εγκατάσταση του πρέπει να πληκτρολογήσουμε `sudo apt-get install <Όνομα πακέτου>`, όπου το όνομα πακέτου είναι ένα από αυτά που βρίσκονται στο παρακάτω πίνακα:

libopencv-dev	libpng3	libdc1394-22-dev
build-essential	libpnglite-dev	libdc1394-22
libavformat-dev	zlib1g-dbg	libdc1394-utils
x264	zlib1g	libv4l-0
v4l-utils	zlib1g-dev	libv4l-dev
ffmpeg	pngtools	libpython2.6
libc2.3	libtiff4-dev	python-dev
libcvaux2.3	libtiff4	python2.6-dev
libhighgui2.3	libtiffxx0c2	libgtk2.0-dev
python-opencv	libtiff-tools	libpngwriter0-dev
opencv-doc	libjpeg8	libpngwriter0c2
libcv-dev	libjpeg8-dev	libswscale-dev
libcvaux-dev	libjpeg8-dbg	libjpeg-dev
libhighgui-dev	libavcodec-dev	libwebp-dev
python-numpy	libavcodec53	libpng-dev
python-scipy	libavformat53	libtiff5-dev
python-matplotlib	libstreamer0.10-0-dbg	libjasper-dev
python-pandas	libstreamer0.10-0	libopenexr-dev
python-nose	libstreamer0.10-dev	libgdal-dev
v4l-utils	libxine1-ffmpeg	python-tk
libgtkglext1-dev	libxine-dev	python3-dev
libpng12-0	libxine1-bin	python3-tk
libpng12-dev	libunicap2	python3-numpy
libpng+-dev	libunicap2-dev	libeigen3-dev

Σχήμα 18 “ Πακέτα προς εγκατάσταση”

Για παράδειγμα, αν θέλουμε να εγκαταστήσουμε το x264 το οποίο είναι ένα πακέτο που μας επιτρέπει να διαβάζουμε διάφορες μορφές αρχείων βίντεο από το δίσκο και να μπορούμε να δουλεύουμε απευθείας σε ροές βίντεο. Το μόνο που πρέπει να κάνουμε είναι να πληκτρολογήσουμε `sudo apt-get install x264`. Αυτό θα εγκαταστήσει το απαραίτητο πακέτο. Με τον ίδιο ακριβώς τρόπο γίνεται και η εγκατάσταση των υπολοίπων πακέτων [9]. Εάν το πακέτο είναι ήδη εγκατεστημένο στο Pi, τότε θα μας δώσει αυτό το μήνυμα:

```
pi@pi02 - $ sudo apt-get install x264
Reading package lists... Done
Building dependency tree
Reading state information... Done
x264 is already the newest version.
0 upgraded, 0 newly installed, 0 to remove and 0 not
upgraded.
```

Σε αυτή την περίπτωση δεν υπάρχει λόγος ανησυχίας. Το πακέτο που ήθελες έχει ήδη εγκατασταθεί και είναι ήδη ενημερωμένο. Απλά προχωρήστε στην εγκατάσταση υπολοίπων πακέτων στη λίστα μια-μια.

6) Τα παραπάνω πακέτα του πίνακα πρέπει να εγκατασταθούν οπωσδήποτε πριν γίνει η εγκατάσταση του OpenCV. Θα εξηγήσω εν συντομία τα βήματα για την εγκατάσταση του.

Αφού έχουν ολοκληρωθεί τα 4 βήματα του κεφαλαίου 8 πρέπει να γίνει εγκατάσταση ορισμένων εργαλείων ανάπτυξης [9].

-Το Cmake μας βοηθά να διαμορφώσουμε την διαδικασία δημιουργίας του OpenCV. Η εντολή για την εγκατάσταση του Cmake είναι:

-sudo apt-get install build-essential cmake pkg-config

-Στη συνέχεια, πρέπει να εγκαταστήσουμε ορισμένα I/O πακέτα εικόνας που θα μας επιτρέπουν να φορτώσουμε διάφορες μορφές αρχείων εικόνας από το δίσκο. Παράδειγμα τέτοιων μορφών αρχείων περιλαμβάνουν JPEG,PNG,TIFF κλπ. Η εντολή για αυτό είναι:

-sudo apt-get install libjpeg-dev libtiff5-dev libjasper-dev libpng12-dev

-Όπως χρειαζόμαστε πακέτα για για την φόρτωση εικόνων έτσι χρειάζεται να εγκαταστήσουμε πακέτα για την φόρτωση βίντεο και να μπορούμε να δουλεύουμε με ροές βίντεων. Η εντολές για αυτό είναι:

-sudo apt-get install libavcodec-dev libavformat-dev libswscale-dev libv4l-dev

-sudo apt-get install libxvidcore-dev libx264-dev

-Η βιβλιοθήκη OpenCV συνοδεύεται από μια υπομονάδα με όνομα 'highgui', το οποίο χρησιμοποιείται για την εμφάνιση εικόνων στην οθόνη μας και την κατασκευή βασικών GUI (Γραφικό περιβάλλον χρήστη). Προκειμένου να καταρτίσουμε την υπομονάδα

'highgui' πρέπει να εγκαταστήσουμε την βιβλιοθήκη ανάπτυξης GTK. Το GTK είναι η βιβλιοθήκη που δημιουργεί το Γραφικό περιβάλλον του χρήστη. Χρησιμοποιώντας αυτήν την βιβλιοθήκη, μπορούμε να δημιουργήσουμε προγράμματα ανοιχτού κώδικα, ελεύθερα ή εμπορικά. Η εντολή για την εγκατάσταση της βιβλιοθήκης GTK είναι:

-sudo apt-get install libgtk2.0-dev

-Πολλές λειτουργίες εντός του OpenCV μπορούν να βελτιστοποιηθούν περαιτέρω με την εγκατάσταση κάποιων επιπλέον εξαρτήσεων:

-sudo apt-get install libatlas-base-dev gfortran

-Και τέλος ας εγκαταστήσουμε και τις δυο εκδόσεις Python2.7 και Python3 ώστε να μπορέσουμε να μεταγλωττίσουμε το OpenCV με Python:

-sudo apt-get install python2.7-dev python3-dev

**Εάν παραβλέψουμε το παραπάνω βήμα, ενδέχεται να παρατηρήσουμε σφάλμα στο Python.h, το αρχείο κεφαλίδας δεν βρέθηκε κατά την εκτέλεση του make για να μεταγλωττίσουμε το OpenCV. **

-Τώρα που έχουμε εγκαταστήσει τα απαραίτητα πακέτα, πρέπει να κάνουμε λήψη του OpenCV 3.1.0. Η εντολή για αυτό είναι:

-cd ~

-wget -O opencv.zip https://github.com/Itseez/opencv/archive/3.1.0.zip

-unzip opencv.zip

-Όμως για να είναι ολοκληρωμένο το OpenCV πρέπει να εγκαταστήσουμε και το Opencv_Contrib γιατί υπάρχουν κάποια κρυμμένα αρχεία που με αυτό θα είναι εμφανείς και θα μπορούμε να δουλεύουμε. Η εντολή για αυτό είναι:

```
-wget -O opencv_contrib.zip https://github.com/Itseez/opencv_contrib/archive/3.1.0.zip  
-unzip opencv_contrib.zip
```

Σημείωση* Βεβαιωθείτε ότι το OpenCV και OpenCV_Contrib είναι ίδιας έκδοσης. Αν οι αριθμοί έκδοσης δεν ταιριάζουν τότε πιθανότητα να εμφανιστούμε στη στιγμή μεταγλώττισης είτε στη στιγμή εκτέλεσης. *

-Πριν μεταγλωττίσουμε το OpenCV πρέπει να εγκαταστήσουμε το pip, ένας διαχειριστής πακέτων της Python.

```
-wget https://bootstrap.pypa.io/get-pip.py  
-sudo python get-pip.py
```

-Η εγκατάσταση των πακέτων virtualenv και virtualenvwrapper δεν είναι αναγκαστικό να γίνει η εγκατάσταση τους για να εγκαταστήσουμε το OpenCV, αλλά το συνιστώ απολυτά να τα εγκαταστήσουμε καθώς σε άλλα μαθήματα PyImageSearch στο μέλλον θα επηρεάσουν τα εικονικά περιβάλλοντα Python. Οπότε είναι καλό να γίνει εγκατάσταση και των δυο πακέτων. Παρακάτω θα δούμε ότι είναι σημαντικά αυτά τα δυο πακέτα.

Πρώτα από όλα είναι σημαντικό να καταλάβουμε ότι ένα εικονικό περιβάλλον είναι ένα ειδικό εργαλείο που χρησιμοποιείτε για να διατηρεί τις εξαρτήσεις που απαιτούνται από διαφορετικά έργα σε ξεχωριστά μέρη δημιουργώντας απομονωμένα, ανεξάρτητα περιβάλλοντα Python για καθένα από αυτά [9].

```
-sudo pip install virtualenv virtualenvwrapper  
-sudo rm -rf ~/.cache/pip
```

-Τώρα που εγκαταστήσαμε το `virtualenv` και `virtualenvwrapper` πρέπει να ενημερώσουμε το `~/.profile` αρχείο προσθέτοντας αυτές τις γραμμές. Μέσα στο αρχείο μπαίνουμε με την εντολή `nano ~/.profile`.

`# virtualenv and virtualenvwrapper`

`-export WORKON_HOME=$HOME/.virtualenvs`

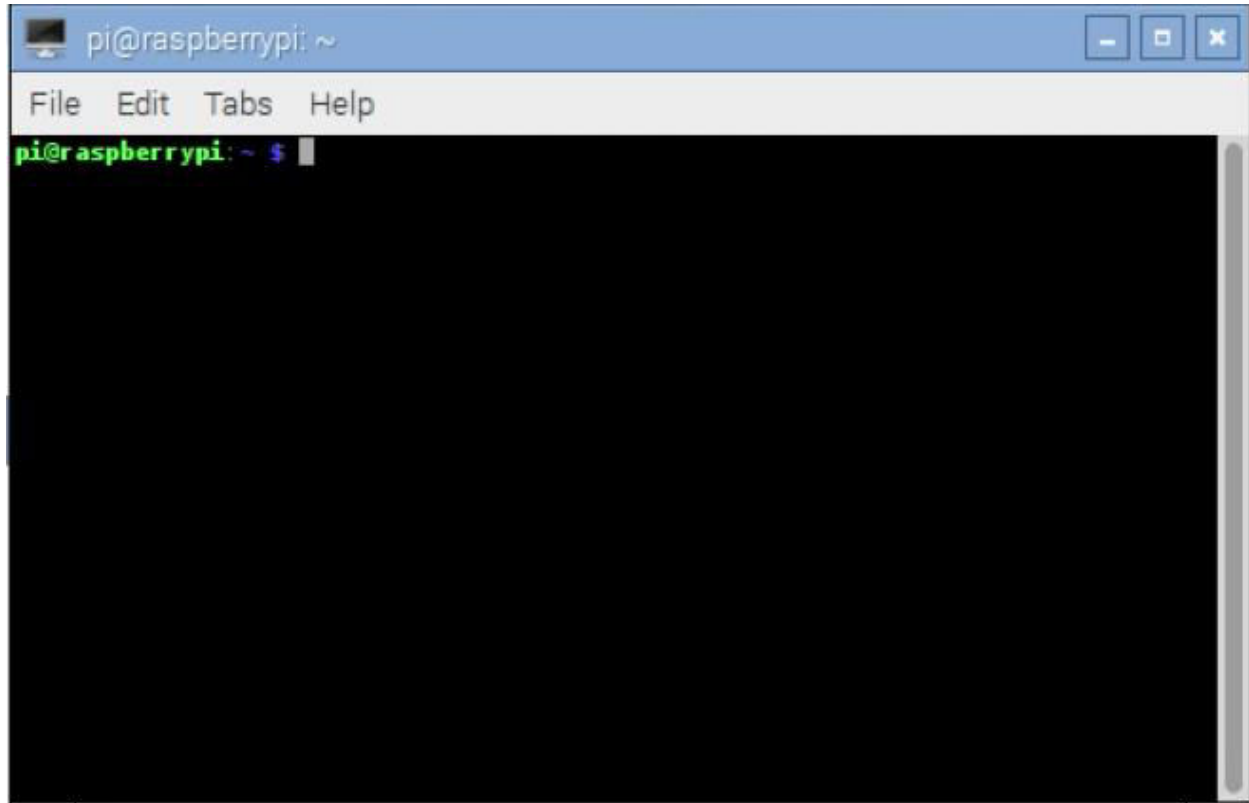
`-source /usr/local/bin/virtualenvwrapper.sh`

-Στη συνέχεια ας δημιουργήσουμε Python περιβάλλον που θα χρησιμοποιήσουμε για την ανάπτυξη μηχανικής όρασης με την παρακάτω εντολή:

`-mkvirtualenv cv -p python2`

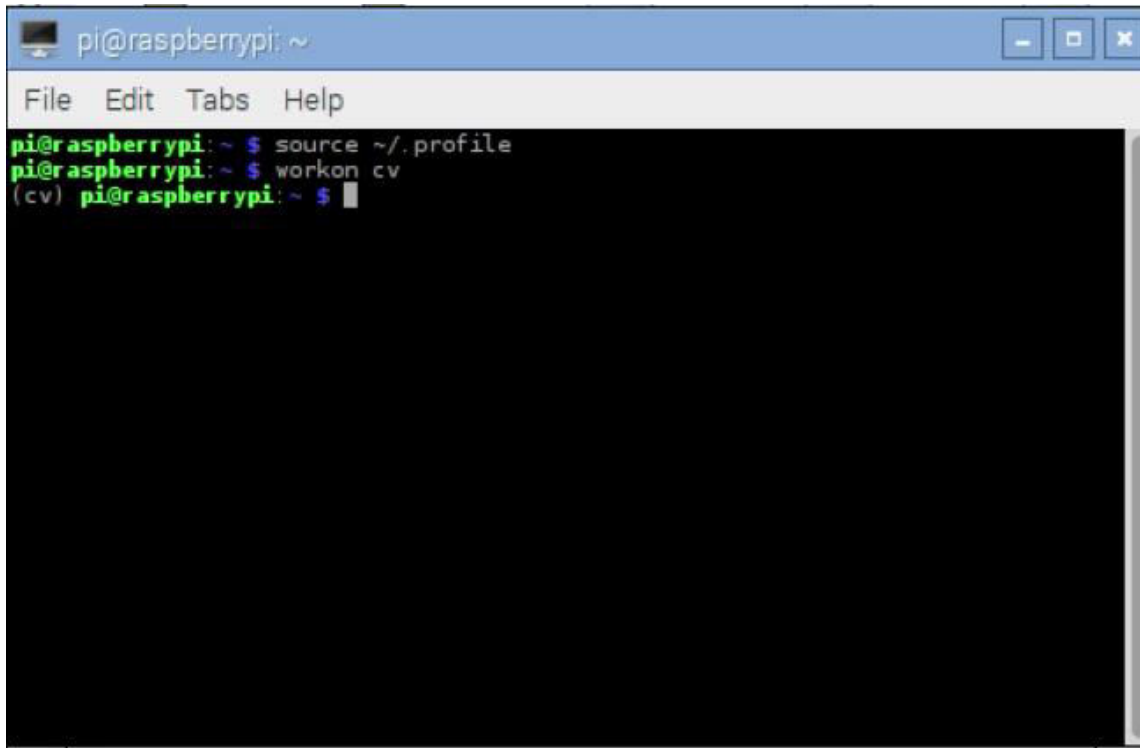
Η παραπάνω εντολή είναι για να δουλέψουμε σε `python2.7`, αν θέλουμε να δουλεύουμε σε Python 3 τότε είναι η ίδια εντολή αντί για `python2.7` στο τέλος βάζουμε `python3`.

-Αν δεν ξέρεις εάν είσαι σε `python` περιβάλλον πρέπει στο τερματικό να έχεις μπροστά (`cv`). Αν είναι όπως η παρακάτω εικόνα τότε δεν βρίσκεσαι σε Python περιβάλλον.



Σχήμα 19 “Το τερματικό του Raspberry Pi”

Για να είσαι σε Python περιβάλλον πρέπει να δώσεις την εντολή [source ~/.profile](#) και μετά δίνεις την εντολή [workon cv](#). Όπως στην παρακάτω εικόνα



```
pi@raspberrypi: ~  
File Edit Tabs Help  
pi@raspberrypi: ~ $ source ~/.profile  
pi@raspberrypi: ~ $ workon cv  
(cv) pi@raspberrypi: ~ $
```

Σχήμα 20 “Εικονικό περιβάλλον Open CV”

Αφού ολοκληρωθεί και αυτό το βήμα τότε είμαστε έτοιμοι να μεταγλωττίσουμε το OpenCV. Αφού έχουμε σιγουρευτεί ότι βρισκόμαστε σε cv εικονικό περιβάλλον τότε πάμε να χτίσουμε το CMake Που μας βοηθά να δουλεύουμε πιο ευκολά τις εργασίες μας. Οι εντολές αυτές είναι:

```
cd ~/opencv-3.1.0/  
$ mkdir build  
$ cd build  
$ cmake -D CMAKE_BUILD_TYPE=RELEASE \  
-D CMAKE_INSTALL_PREFIX=/usr/local \  
-D INSTALL_PYTHON_EXAMPLES=ON \  
-D OPENCV_EXTRA_MODULES_PATH=~/opencv_contrib-3.1.0/modules \  
-D BUILD_EXAMPLES=ON..
```

Πριν προχωρήσουμε στο τελικό βήμα ολοκληρώσεις, πρέπει να είμαστε σίγουροι για την έξοδο της CMake. Τι πρέπει να παρατηρήσουμε για να δούμε αν η έξοδος της CMake είναι σωστή? Τότε υπάρχουν δυο περιπτώσεις. Εάν κάνεις μεταγλώττιση OpenCV για Python2.7 τότε σιγουρέψου ότι η ενότητα του Python2.7 περιέχει έγκυρες

διαδρομές στο Interpreter, Libraries, Numpy και Packages path. Βλέπε παρακάτω εικόνα

```
-- Use Intel VA-API/OpenCL: NO
-- Use Eigen: NO
-- Use Cuda: NO
-- Use OpenCL: YES
-- Use custom HAL: NO
--
-- OpenCL:
-- Version: dynamic
-- Include path: /home/pi/opencv-3.1.0/3rdparty/include/opencv/1.2
-- Use AMDFFT: NO
-- Use AMDBLAS: NO
--
-- Python 2:
-- Interpreter: /home/pi/.virtualenvs/cv/bin/python2.7 (ver 2.7.9)
-- Libraries: /usr/lib/arm-linux-gnueabi/libpython2.7.so (ver 2.7.9)
-- numpy: /home/pi/.virtualenvs/cv/local/lib/python2.7/site-packages/numpy/core/include (ver 1.18.4)
-- packages path: lib/python2.7/site-packages
--
-- Python 3:
-- Interpreter: /usr/bin/python3.4 (ver 3.4.2)
-- Libraries: /usr/lib/arm-linux-gnueabi/libpython3.4m.so (ver 3.4.2)
-- numpy: /usr/lib/python3/dist-packages/numpy/core/include (ver 1.8.2)
-- packages path: lib/python3.4/site-packages
--
-- Python (for build): /home/pi/.virtualenvs/cv/bin/python2.7
--
-- Java:
-- ant: NO
-- JNI: NO
-- Java wrappers: NO
-- Java tests: NO
```

Σχήμα 21 “Έλεγχος εγκύρων διαδρομών στην εγκατάσταση του Python 2.7”

Το ίδιο ισχύει εάν κάνεις μεταγλώττιση με Python3:

```
-- Xine: NO
-- gPhoto2: NO
--
-- Parallel framework: pthreads
--
-- Other third-party libraries:
-- Use IPP: NO
-- Use VA: NO
-- Use Intel VA-API/OpenCL: NO
-- Use Eigen: NO
-- Use Cuda: NO
-- Use OpenCL: YES
-- Use custom HAL: NO
--
-- OpenCL:
-- Version: dynamic
-- Include path: /home/pi/opencv-3.1.0/3rdparty/include/opencv/1.2
-- Use AMFFFT: NO
-- Use AMDLLAS: NO
--
-- Python 2:
-- Interpreter: /usr/bin/python2.7 (ver 2.7.9)
-- Libraries: /usr/lib/arm-linux-gnueabi/libpython2.7.so (ver 2.7.9)
-- numpy: /usr/lib/python2.7/dist-packages/numpy/core/include (ver 1.8.2)
-- packages path: lib/python2.7/dist-packages
--
-- Python 3:
-- Interpreter: /home/pi/.virtualenvs/cv/bin/python3.4 (ver 3.4.2)
-- Libraries: /usr/lib/arm-linux-gnueabi/libpython3.4m.so (ver 3.4.2)
-- numpy: /home/pi/.virtualenvs/cv/lib/python3.4/site-packages/numpy/core/include (ver 1.10.4)
-- packages path: lib/python3.4/site-packages
```

Σχήμα 22

Και τέλος είμαστε έτοιμοι να μεταγλωττίσουμε το OpenCV με την παρακάτω εντολή:

Make -j4

Το -j4 είναι μια εντολή που ελέγχει τον αριθμό των πυρήνων για την αξιοποίηση τους την ώρα που γίνεται μεταγλώττιση του OpenCV. Στη δικιά μου περίπτωση χρησιμοποίησα Raspberry Pi 3 το οποίο είναι τετραπύρρηνο για αυτό και ο αριθμός 4 για να μπορεί να μεταγλωττιστεί πιο γρήγορα.

Σε κάποιες περιπτώσεις όμως βγάζει λάθος όταν χρησιμοποιούμε πολλαπλές πυρήνες, τότε η λύση είναι να κάνουμε την μεταγλώττιση ξανά χρησιμοποιώντας μόνο έναν πυρήνα [10].

Αφού το OpenCV έχει τελειώσει τότε πρέπει να μοιάζει με την παρακάτω εικόνα.

```
[ 98%] Built target tutorial_npr_demo
Scanning dependencies of target tutorial_planar_tracking
[ 98%] Building CXX object samples/cpp/MakeFiles/tutorial_planar_tracking.dir/tutorial_code/features2D/AKAZE_tracking/plana
r_tracking.cpp.o
Linking CXX executable ../../bin/cpp-tutorial-objectDetection
Linking CXX executable ../../bin/cpp-tutorial-objectDetection2
Linking CXX executable ../../bin/cpp-tutorial-non_linear_svm
[ 98%] Built target tutorial_objectDetection
[ 98%] Built target tutorial_objectDetection2
Scanning dependencies of target tutorial_pointPolygonTest_demo
Scanning dependencies of target tutorial_video-input-psnr-ssim
[ 98%] [ 98%] Building CXX object samples/cpp/MakeFiles/tutorial_pointPolygonTest_demo.dir/tutorial_code/ShapeDescriptors/p
ointPolygonTest_demo.cpp.o
Built target tutorial_non_linear_svm
[ 98%] Building CXX object samples/cpp/MakeFiles/tutorial_video-input-psnr-ssim.dir/tutorial_code/HighGUI/video-input-psnr-
ssim/video-input-psnr-ssim.cpp.o
Scanning dependencies of target tutorial_video-write
[100%] Building CXX object samples/cpp/MakeFiles/tutorial_video-write.dir/tutorial_code/HighGUI/video-write/video-write.cpp
.o
Linking CXX executable ../../bin/cpp-tutorial-pointPolygonTest_demo
Linking CXX executable ../../bin/cpp-tutorial-video-write
[100%] Built target tutorial_pointPolygonTest_demo
[100%] Built target tutorial_video-write
Linking CXX executable ../../bin/cpp-tutorial-video-input-psnr-ssim
Linking CXX executable ../../bin/cpp-tutorial-planar_tracking
[100%] Built target tutorial_video-input-psnr-ssim
[100%] Built target tutorial_planar_tracking

real    72m12.945s
user    228m31.480s
sys     6m28.110s
(cv) pi@raspberrypi: ~/opencv-3.1.0/build $
```

Σχήμα 23 “Ολοκλήρωση εγκατάστασης OpenCV”

Σχεδόν φτάνουμε στο τέλος, μερικά βήματα ακόμα και θα είμαστε έτοιμοι να χρησιμοποιήσουμε το Raspberry Pi 3 με OpenCV 3.

Αφού ολοκληρώσαμε το προηγούμενο βήμα χωρίς κανένα λάθος τότε το OpenCV πρέπει να είναι εγκατεστημένο στο: [/usr/local/lib/python2.7/site-packages](#).

Για να το ελέγξουμε χρησιμοποιούμε την [ls](#) εντολή όπως παρακάτω.

```
ls -l /usr/local/lib/python2.7/site-packages
```

και ως απάντηση πρέπει να πάρουμε:

```
total 1852
```

```
-rw-r--r-- 1 root staff 1895772 Mar 20 20:00 cv2.so
```

Το τελικό βήμα είναι να κάνουμε συμβολικό σύνδεσμο(Συμβολικός σύνδεσμος είναι ένα όρος για οποιοδήποτε αρχείο που περιέχει μια αναφορά σε άλλο αρχείο ή κατάλογο με την μορφή απολυτής ή σχετικής διαδρομής και επηρεάζει την ανάλυση ονομάτων της διαδρομής) του OpenCV στο εικονικό περιβάλλον του Python2.7.

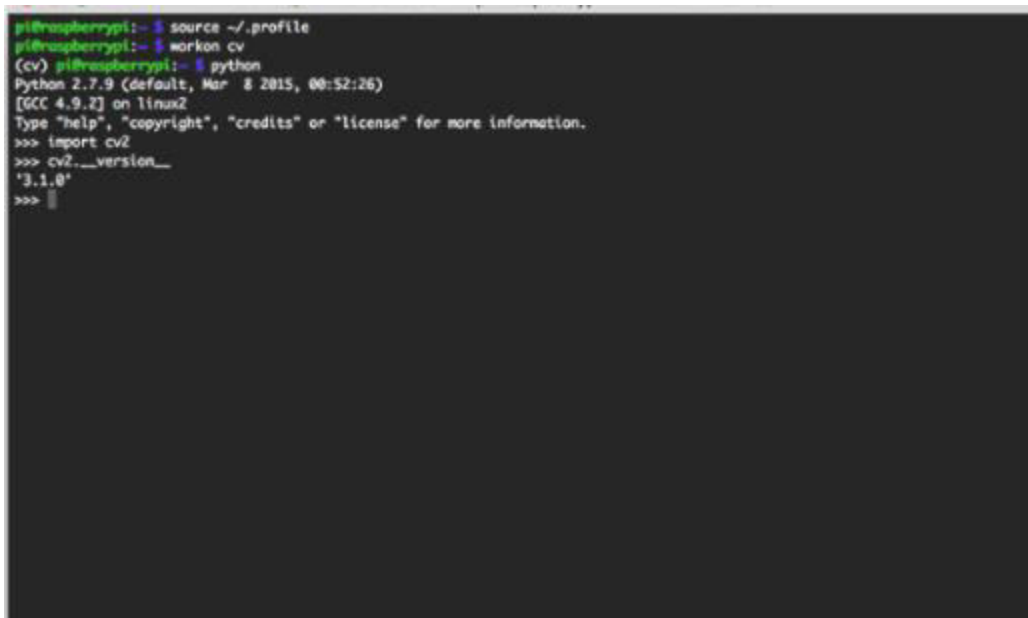
```
cd ~/.virtualenvs/cv/lib/python2.7/site-packages/
```

In -s /usr/local/lib/python2.7/site-packages/cv2.so cv2.so

Το ίδιο κάνουμε αν χρησιμοποιούμε το Python 3.

Εννιώ ολοκληρώσαμε όλα τα βήματα ήρθε η ώρα του ελέγχου του OpenCV. Ανοίγουμε καινούργιο τερματικό και δίνουμε τις παρακάτω εντολές:

```
$ source ~/.profile  
$ workon cv  
$ python  
>>> import cv2  
>>> cv2.__version__  
'3.1.0'  
>>>
```



```
pi@raspberrypi:~$ source ~/.profile  
pi@raspberrypi:~$ workon cv  
(cv) pi@raspberrypi:~$ python  
Python 2.7.9 (default, Mar 8 2015, 00:52:26)  
[GCC 4.9.2] on linux2  
Type "help", "copyright", "credits" or "license" for more information.  
>>> import cv2  
>>> cv2.__version__  
'3.1.0'  
>>> |
```

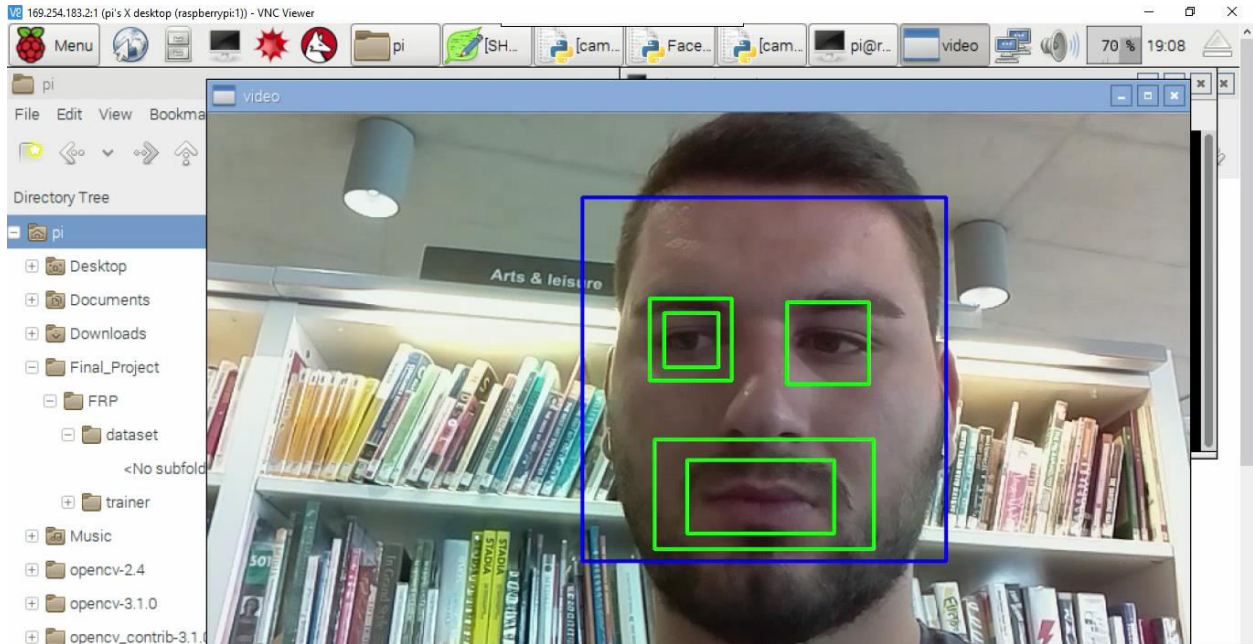
Σχήμα 24 “Έλεγχος έκδοσης OpenCV”

6.2 Ανίχνευση Προσώπου

Το πρόβλημα της αναγνώρισης προσώπου είναι η ανίχνευση προσώπου. Αυτό το γεγονός φαίνεται αρκετά περίεργο στους νέους ερευνητές στον τομέα αυτό. Ωστόσο, πριν η αναγνώριση προσώπου γίνει δυνατή, πρέπει να είσαι σε θέση να βρεις τα αξιοπιστία σημεία σε ένα πρόσωπο και τα ορόσημα του. Αυτό είναι ουσιαστικά ένα πρόβλημα τμηματοποίησης και σε πρακτικά συστήματα, το μεγαλύτερο μέρος της προσπάθειας πηγαίνει στην επίλυση αυτού του καθήκοντος. Στην πραγματικότητα, η πραγματική αναγνώριση που βασίζεται στα χαρακτηριστικά που εξάγονται από αυτά τα ορόσημα προσώπου είναι μόνο ένα τελευταίο βήμα [10]. Υπάρχουν δυο τύποι προβλημάτων ανίχνευσης προσώπου:

- 1) Ανίχνευση προσώπου σε εικόνες
- 2) Ανίχνευση προσώπου σε πραγματικό χρόνο

6.2.1 Ανίχνευση προσώπου σε εικόνα



Εικόνα 25 “Ανίχνευση Προσώπου”

Η παραπάνω εικόνα είναι μια επιτυχημένη ανίχνευση προσώπου σε μια εικόνα με μετωπική όψη ενός ανθρωπίνου προσώπου [11].

Τα περισσότερα συστήματα ανίχνευσης προσώπου προσπαθούν να εξάγουν ένα σύνολο του προσώπου, εξαλείφοντας έτσι το μεγαλύτερο μέρος της πίσω όψης της εικόνας και άλλων περιοχών του κεφαλιού ενός ατόμου, όπως τα μαλλιά που δεν είναι απαραίτητα για την εργασία αναγνώρισης προσώπου. Με τις στατικές εικόνες, αυτό γίνεται συχνά τρέχοντας ένα “παράθυρο” απέναντι στην εικόνα. Το σύστημα ανίχνευσης προσώπου τότε κρίνει εάν υπάρχει πρόσωπο μέσα στο παράθυρο. Δυστυχώς, με στατικές εικόνες υπάρχει ένας πολύ μεγάλος χώρος αναζήτησης πιθανόν θέσεων ενός προσώπου σε μια εικόνα. Τα πρόσωπα μπορεί να είναι μεγάλα ή μικρά και να τοποθετούνται οπουδήποτε από το επάνω αριστερό έως το κάτω δεξιά μέρος της εικόνας [12].

Τα περισσότερα συστήματα ανίχνευσης προσώπου χρησιμοποιούν μια προσέγγιση μάθησης που βασίζεται σε ένα παράδειγμα που αποφασίζει εάν υπάρχει ή όχι πρόσωπο στο παράθυρο αυτή τη στιγμή. Εάν ένα νευρωνικό δίκτυο ή κάποιος άλλος ταξινομητής εκπαιδεύεται χρησιμοποιώντας επίβλεψη μάθησης με παραδείγματα 'προσώπου' και 'μη πρόσωπο', επιτρέποντας έτσι να ταξινομήσει μια εικόνα ως 'πρόσωπο' ή 'μη πρόσωπο'. Δυστυχώς, εννοώ είναι εύκολο να βρούμε παραδείγματα για το πώς να βρεθεί ένα πρόσωπο σε μια εικόνα, πως όμως θα βρούμε ένα αντιπροσωπευτικό παράδειγμα εικόνων που να αντιπροσωπεύουν τα μη πρόσωπα? Επομένως, συστήματα ανίχνευσης χρησιμοποιούν παραδείγματα βασισμένα στη μάθηση χρειάζεται χιλιάδες εικόνες προσώπου και μη προσώπου για να έχει μια αποτελεσματική εκπαίδευση. Οι Rowley, Baluja Και Kanade 1996 χρησιμοποίησαν 1025 εικόνες προσώπου και 8000 εικόνες χωρίς πρόσωπο για την εκπαίδευση του συστήματος [12].

Υπάρχει όμως μια τεχνική για το πώς να προσδιορίσουμε εάν υπάρχει ένα πρόσωπο μέσα στο παράθυρο του συστήματος ανίχνευσης προσώπου, χρησιμοποιώντας την αντιστοίχιση προτύπων. Η διαφορά μεταξύ ενός σταθερού μοτίβου προσώπου και παραθύρου είναι ο υπολογισμός και η οριοθέτηση. Εάν ένα παράθυρο περιέχει ένα πρότυπο που είναι κοντά στο πρόσωπο τότε το παράθυρο κρίνεται ότι περιέχει ένα πρόσωπο. Μια εφαρμογή της αντιστοίχισης προτύπων που ονομάζεται Πρότυπα Συσχετισμού χρησιμοποιεί μια ολόκληρη τράπεζα προτύπων σταθερού μεγέθους για την ανίχνευση των χαρακτηριστικών του προσώπου σε μια εικόνα. Χρησιμοποιώντας διάφορα πρότυπα διαφορετικών μεγεθών, εντοπίζονται πρόσωπα διαφορετικών κλίμακων. Η άλλη εφαρμογή της προσαρμογής του προτύπου είναι να χρησιμοποιεί ένα παραμορφώσιμο πρότυπο. Αντί να χρησιμοποιήσουμε διάφορα πρότυπα σταθερού μεγέθους, χρησιμοποιούμε ένα παραμορφώσιμο πρότυπο και εκεί με την αλλαγή του μεγέθους του προτύπου ελπίζει να ανιχνευτεί ένα πρόσωπο σε μια εικόνα.

Ένα σχέδιο ανίχνευσης προσώπου που σχετίζεται με την αντιστοίχιση προτύπων είναι οι μεταβλητές εικόνες. Ο γεγονός ότι η τοπική κανονική δομή της κατανομής φωτεινότητας ενός προσώπου παραμένει σε μεγάλο βαθμό αμετάβλητη κάτω από

διαφορετικές συνθήκες φωτισμού χρησιμοποιείται για να κατασκευάσει ένα χωρικό πρότυπο του προσώπου που αντιστοιχεί στενά στα χαρακτηριστικά του προσώπου. Με απλά λόγια η μέση ένταση άσπρου- μαύρου ή αλλιώς γκρι κλίμακας στα ανθρώπινα πρόσωπα χρησιμοποιείται ως βάση για την ανίχνευση προσώπου. Για παράδειγμα, σχεδόν τα πάντα σε μια περιοχή ματιών ατόμων είναι πιο σκούρα από το μέτωπο του ή τη μύτη του. Επομένως μια εικόνα ταιριάζει με το πρότυπο εάν ικανοποιεί τις σχέσεις 'πιο σκούρο από' και 'λαμπρότερο από' .

6.2.2 Ανίχνευση προσώπου σε πραγματικό χρόνο

Η ανίχνευση προσώπου σε πραγματικό χρόνο περιλαμβάνει την ανίχνευση ενός προσώπου από μια σειρά πλαισίων από μια συσκευή λήψης βίντεο. Ενώ οι απατήσεις υλικού για ένα τέτοιο σύστημα είναι πολύ πιο αυστηρές, από την μηχανική όραση, η ανίχνευση προσώπου σε πραγματικό χρόνο είναι μια πολύ απλούστερη διαδικασία από την ανίχνευση προσώπου σε μια στατική εικόνα. Αυτό συμβαίνει επειδή σε αντίθεση με το μεγαλύτερο μέρος του περιβάλλοντος μας, οι άνθρωποι κινούνται συνεχώς. Επομένως, η ανίχνευση προσώπου σε πραγματικό χρόνο έχει γίνει ένα σχετικά απλό πρόβλημα και είναι δυνατόν ακόμη και σε αδόμητα και ανεξέλεγκτα περιβάλλοντα χρησιμοποιώντας αυτές τις πολύ απλές τεχνικές επεξεργασίας εικόνας και κανόνες συλλογισμού [12].

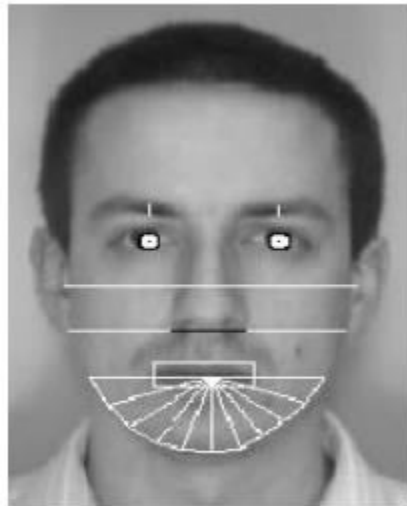
6.3 Αναγνώριση Προσώπου

Τις τελευταίες δεκαετίες έχουν προταθεί πολλές τεχνικές για την αναγνώριση προσώπου. Πολλές από τις τεχνικές που προτάθηκαν κατά τα αρχικά στάδια της μηχανικής όρασης δεν μπορούν να θεωρηθούν επιτυχείς, αλλά σχεδόν όλες οι πρόσφατες προσεγγίσεις στο πρόβλημα αναγνώρισης προσώπων ήταν αξιόπιστες. Σύμφωνα με έρευνα του Brunelli και Roggio(1993), όλες οι προσεγγίσεις στην αναγνώριση ανθρώπινου προσώπου μπορούν να χωριστούν σε δυο κατηγορίες:

- A) Γεωμετρικά χαρακτηριστικά
- B) Αντιστοίχιση προτύπων

6.3.1 Αναγνώριση προσώπου με την χρήση γεωμετρικών χαρακτηριστικών

Αυτή η τεχνική περιλαμβάνει τον υπολογισμό ενός συνόλου γεωμετρικών χαρακτηριστικών όπως το πλάτος και το μήκος της μύτης, η θέση του στόματος και το σχήμα του πηγουνιού κλπ. από την εικόνα του προσώπου που θέλουμε να αναγνωρίσουμε. Αυτό το σύνολο χαρακτηριστικών συνδυάζεται με τα χαρακτηριστικά γνωστών ατόμων. Μια κατάλληλη μετρική όπως η Ευκλείδεια απόσταση μπορεί να χρησιμοποιηθεί για να βρεθεί ο πλησιέστερος συνδυασμός. Η πιο πρωτοποριακή εργασία στην αναγνώριση προσώπου έγινε χρησιμοποιώντας γεωμετρικά χαρακτηριστικά, επίσης ο Crow έκανε σχετικά πρόσφατες εργασίες πάνω στον τομέα αυτό [13].



Σχήμα 26 “ Γεωμετρικά χαρακτηριστικά για την αναγνώριση προσώπου”

Το πλεονέκτημα της χρήσης γεωμετρικών χαρακτηριστικών ως βάση για την αναγνώριση προσώπου είναι ότι η αναγνώριση είναι δυνατή ακόμη και πολύ χαμηλές

αναλύσεις και με θορυβώδεις εικόνες. Αν και το πρόσωπο δεν μπορεί να προβληθεί λεπτομερώς, η συνολική γεωμετρική του διαμόρφωση μπορεί να εξαχθεί για αναγνώριση προσώπου.

Το κύριο μειονέκτημα της τεχνικής είναι ότι η αυτοματοποιημένη εξαγωγή των γεωμετρικών χαρακτηριστικών του προσώπου είναι πολύ δύσκολη. Η αυτοματοποιημένη αναγνώριση με βάση την εξαγωγή χαρακτηριστικών είναι επίσης πολύ ευαίσθητη στην κλιμάκωση και την περιστροφή ενός προσώπου στο επίπεδο εικόνας. Αυτό είναι εμφανές όταν εξετάζουμε τα αποτελέσματα της Kanade οπού ανέφερε το ποσοστό αναγνώρισης μεταξύ 45-75% με βάση δεδομένων μόνο 20 ατόμων. Ωστόσο, αν αυτά τα χαρακτηριστικά εξάγονται με το χέρι μπορούν να ληφθούν αρκετά ικανοποιητικά αποτελέσματα σύμφωνα με το Goldstain (1971), Kaya και Kobayashi (1972).

6.3.2 Αναγνώριση προσώπου με τη χρήση προτύπων ταιριάσματος

Αυτό είναι παρόμοιο με την τεχνική αντιστοίχισης προτύπων που χρησιμοποιείται στην ανίχνευση προσώπου, εκτός από το σημείο που προσπαθούμε να μην ταξινομήσουμε μια εικόνα ως 'πρόσωπο' ή 'μη πρόσωπο' αλλά προσπαθούμε να αναγνωρίσουμε ένα πρόσωπο [14].



Σχήμα 27 “ Όλο το πρόσωπο , μύτη, στόμα και μάτια που χρησιμοποιούνται για την αναγνώριση προσώπου”

Η βάση στρατηγικής αντιστοίχισης προτύπου είναι να εξάγει ολόκληρες τις περιοχές του προσώπου και να τις συγκρίνει με τις αποθηκευμένες εικόνες γνωστών ατόμων. Για άλλη μια φορά, η Ευκλείδεια απόσταση μπορεί να χρησιμοποιηθεί για να βρει τον πλησιέστερο ταίριασμα. Ωστόσο, υπάρχουν πολύ πιο εξελιγμένες μέθοδοι αντιστοίχισης προτύπων για αναγνώριση προσώπου. Αυτά περιλαμβάνουν εκτεταμένη Προ επεξεργασία και μετασχηματισμό των τιμών εξαγομένης έντασης γκριζου επιπέδου. Για παράδειγμα, οι Turk και Pentland χρησιμοποίησαν την βασική ανάλυση συστατικών, γνωστό και ως προσέγγιση EigenFaces, για την Προ επεξεργασία γκριζου χρησιμοποίησαν διαγράμματα που κωδικοποιήθηκαν χρησιμοποιώντας φίλτρα Gabor για την Προ επεξεργασία των εξαγομένων περιοχών [12].

Μια έρευνα σχετικά με τα γεωμετρικά χαρακτηριστικά έναντι του προτύπου για την αναγνώριση προσώπου από τους Brunelli και Poggio (1993) κατάληξε στο συμπέρασμα ότι αν και μια στρατηγική που βασίζεται σε χαρακτηριστικά μπορεί να προσφέρει υψηλότερη ταχύτητα αναγνώρισης και μικρότερες απαιτήσεις μνήμης.

6.4 Προεπεξεργασία

Ο σκοπός του βήματος προ επεξεργασία προσώπου είναι η ομαλοποίηση της ανίχνευσης προσώπου, έτσι ώστε να μπορεί να επιτευχθεί μια ισχυρή εξαγωγή χαρακτηριστικών. Ανάλογα με την εφαρμογή, η προ επεξεργασία προσώπου περιλαμβάνει: Παράταξη κατά σειρά (Μετάφραση, περιστροφή, κλιμάκωση) και κανονικοποίηση /συσχέτιση φωτός.

6.5 Εξαγωγή Χαρακτηριστικών

Σκοπός της εξαγωγής χαρακτηριστικών είναι η εξαγωγή ενός συμπαγούς συνόλου διαπροσωπικών διακριτικών γεωμετρικών ή φωτομετρικών χαρακτηριστικών του προσώπου.

6.6 Αντιστοίχιση Χαρακτηριστικών

Η αντιστοίχιση προτύπων είναι η πραγματική διαδικασία αναγνώρισης. Τα πρότυπα χαρακτηριστικών που λαμβάνεται από την εξαγωγή χαρακτηριστικών αντιστοιχεί σε κλάσεις (άτομα) εικόνων προσώπου που έχουν ήδη εγγραφεί σε μια βάση δεδομένων. Οι αλγόριθμοι αντιστοίχισης ποικίλλουν από τα αρκετά προηγμένα συστήματα όπως τα Νευρωνικά δίκτυα.

6.7 Δεδομένα Προσώπου

Προκειμένου να κατασκευάσουμε /να εκπαιδεύσουμε και να δοκιμάσουμε αξιόπιστα τους αλγόριθμους αναγνώρισης προσώπων απαιτούνται σημαντικές βάσεις δεδομένων εικόνων προσώπου. Πολλές βάσεις δεδομένων που χρησιμοποιούνται για μη εμπορικούς σκοπούς είναι διαθέσιμες στο διαδίκτυο, είτε δωρεάν είτε έναντι μικρού ποσού χρήματων. Αυτές οι βάσεις δεδομένων καταγράφονται υπό διάφορες συνθήκες και με διάφορες εφαρμογές. Οι ακόλουθες ενότητες περιγράφουν εν συντομία μερικές από τις διαθέσιμες βάσεις δεδομένων που είναι ευρέως γνωστές και χρησιμοποιούνται [14].

6.7.1 IMM Face Database

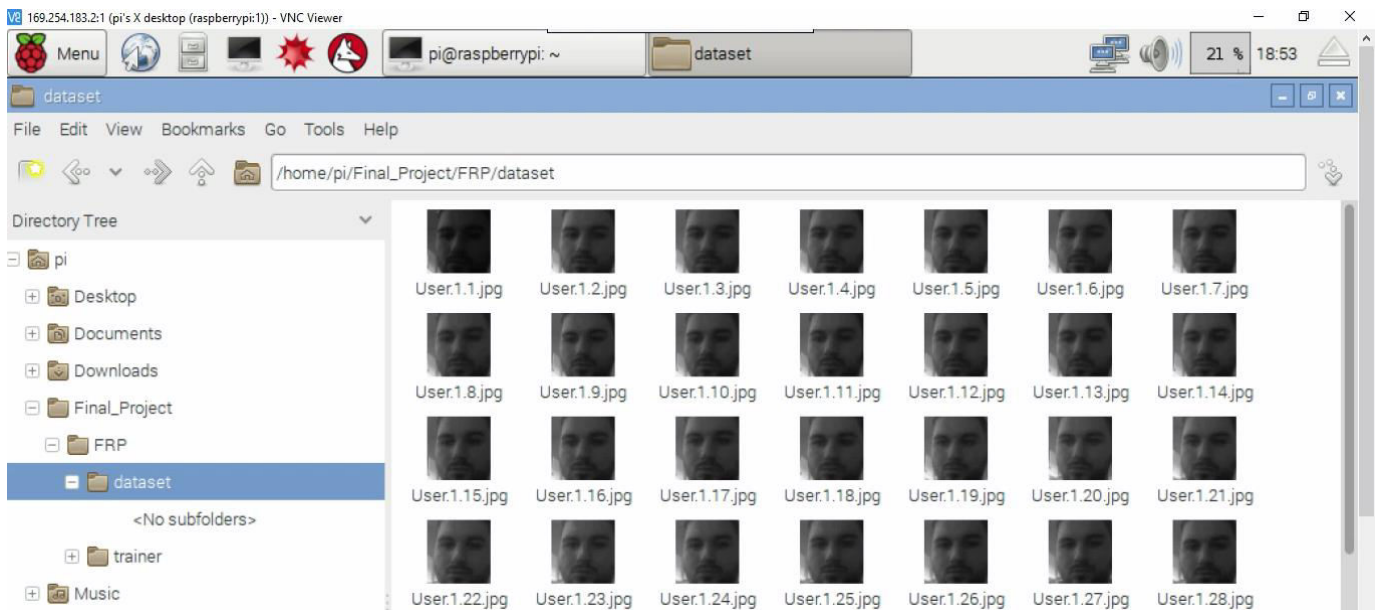
Η βάση δεδομένων προσώπου IMM καταγράφηκε το 2001 στο τμήμα πληροφορικής και μαθηματικής μοντελοποίησης του Τεχνικού Πανεπιστημίου της Δανίας. Η βάση δεδομένων περιέχει εικόνες από 40 άτομα. Από αυτούς 33 είναι αρσενικού γένους και 7 θηλυκού γένους. Καταγράφηκε κατά την διάρκεια μια συνεδρίας και αποτελείται από 7 εικόνες άνα άτομο με αποτέλεσμα να υπάρχουν συνολικά 240 εικόνες. Οι 7 εικόνες του κάθε ατόμου καταγράφηκαν υπό διάφορες εκφράσεις του προσώπου, σημεία θέασης κάμερας και φωτισμού. Οι περισσότερες από τις εικόνες έχουν καταγραφεί σε χρώμα, ενώ οι υπόλοιπες έχουν καταγραφεί σε γκρι κλίμακα, όλες με ανάλυση 640x480 pixel. Τα ορόσημα βρίσκονται σε 58 διαφορετικά σημεία του σχήματος [12].

6.7.2 IMM Frontal Face Database

Η βάση δεδομένων προσώπου IMM Frontal Face πραγματοποιήθηκε το 2006 στο τμήμα πληροφορικής και μαθηματικής μοντελοποίησης του τεχνικού πανεπιστημίου της Δανίας. Η βάση δεδομένων περιέχει εικόνες από 12 άτομα. Όλα αρσενικού γένους. Η βάση δεδομένων καταγράφηκε κατά την διάρκεια μιας συνεδρίας και αποτελείται από 10 εικόνες ο καθένας με αποτέλεσμα να υπάρχουν συνολικά 120 εικόνες. Οι 10 εικόνες του κάθε ατόμου καταγράφηκαν υπό διάφορες εκφράσεις του προσώπου. Όλες οι εικόνες καταγράφονται σε χρώμα με ανάλυση 2560x1920 pixel. Τα ορόσημα βρίσκονται σε 73 διαφορετικά σημεία του σχήματος [12].

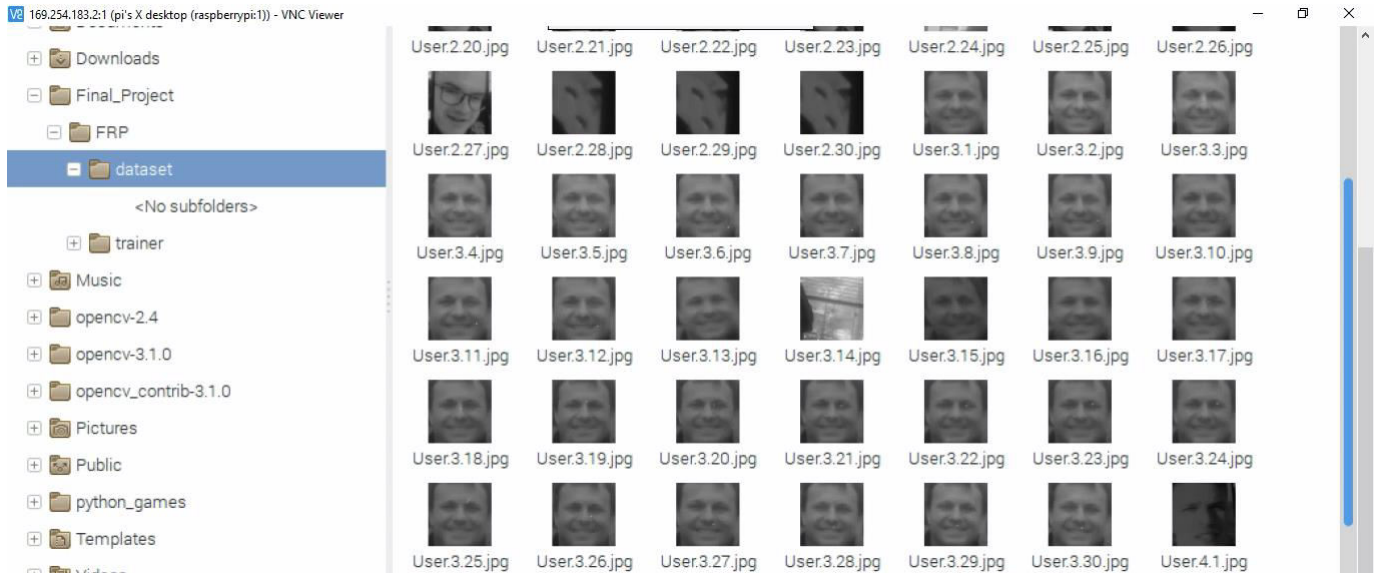
6.8 Σετ δεδομένων που χρησιμοποιήθηκαν σε αυτήν την εργασία

6.8.1 ΣΕΤ 1



Εικόνα 28 “Σετ εικόνων που χρησιμοποιήθηκαν για την αναγνώριση προσώπου”

6.8.2 ΣΕΤ 2



6.9 Ανίχνευση Προσώπου: Έρευνα

Αυτό το κεφάλαιο ασχολείται με το πρόβλημα της ανίχνευσης προσώπου. Δεδομένου ότι κυρίως αντικείμενο αυτής της εργασίας είναι η αναγνώριση προσώπου, αυτό το κεφάλαιο θα χρησιμεύει ως εισαγωγή σε ήδη αναπτυγμένους αλγορίθμους για την ανίχνευση προσώπου. Όπως περιεγράφηκε προηγουμένως στο κεφάλαιο 9, η ανίχνευση προσώπου είναι το απαραίτητο πρώτο βήμα σε ένα σύστημα αναγνώρισης προσώπου. Ο σκοπός ανίχνευσης προσώπου είναι να εντοπίσει και να εξαγάγει την περιοχή προσώπου από το φόντο της εικόνας. Ωστόσο, δεδομένου ότι το ανθρώπινο πρόσωπο είναι ένα εξαιρετικά δυναμικό αντικείμενο που παρουσιάζει μεγάλο βαθμό μεταβλητότητας στην εμφάνιση, η αυτόματη ανίχνευση προσώπου παραμένει ένα δύσκολο έργο. Το πρόβλημα περιπλέκεται περαιτέρω από τις συνεχείς αλλαγές του προσώπου με την πάροδο του χρόνου των ακολουθών παραμέτρων:

- Η τρισδιάστατη θέση του προσώπου
- Αφαιρούμενα χαρακτηριστικά, όπως γυαλιά και γένια
- Έκφραση Προσώπου
- Μερική απόφραξη του προσώπου, π.χ τρίχες, κασκόλ και γυαλιά ηλίου

-Προσανατολισμός του προσώπου

-Συνθήκες φωτισμού

Παρακάτω θα δείτε ακριβώς την περιγραφή αυτών των δυο ορών ανίχνευση προσώπου και εντοπισμός προσώπου.

Ανίχνευση προσώπου: Η διαδικασία ανίχνευσης όλων των προσώπων (εάν υπάρχουν) σε μια δεδομένη εικόνα.

Εντοπισμός Προσώπου: Η διαδικασία εντοπισμού ενός προσώπου σε μια δεδομένη εικόνα, δηλαδή υποτίθεται ότι η εικόνα περιέχει μια και μόνο όψη [15].

Έχουν αναπτυχθεί περισσότερες από 150 μέθοδοι ανίχνευσης προσώπου, αν και εδώ εξετάζεται μόνο ένα μικρό υποσύνολο. Οι μέθοδοι ανίχνευσης προσώπου χωρίζονται σε τέσσερις κατηγορίες:

-Μέθοδοι που βασίζονται στη γνώση: Οι μέθοδοι που βασίζονται στη γνώση χρησιμοποιούν ένα σύνολο κανόνων που περιγράφουν τι πρέπει να συλλαμβάνει. Οι κανόνες κατασκευάζονται από την διαισθητική ανθρώπινη γνώση των συστατικών το προσώπου και μπορούν να είναι απλές σε σχέσεις μεταξύ χαρακτηριστικών προσώπου.

-Χαρακτηριστικά αμετάβλητων προσεγγίσεων: Σκοπός των αντικειμενικών αμετάβλητων προσεγγίσεων είναι η αναζήτηση διαρθρωτικών χαρακτηριστικών, τα οποία είναι αμετάβλητα στις μεταβολές των συνθηκών θέσης και φωτισμού.

-Μέθοδοι αντιστοίχισης προτύπων: Οι μέθοδοι αντίστοιχης προτύπων δημιουργούν ένα ή περισσότερα πρότυπα (μοντέλα) για την περιγραφή χαρακτηριστικών προσώπου. Η συσχέτιση μεταξύ μιας εικόνας εισόδου και κατασκευασμένου μοντέλου επιτρέπει στη μέθοδο να διακρίνει την περίπτωση προσώπου ή του μη προσώπου.

-Μέθοδοι βασισμένες σε εμφάνιση: Οι μέθοδοι βασισμένες σε εμφάνιση χρησιμοποιούν τη στατιστική ανάλυση και την μηχανική μάθηση για να εξαγάγουν τα σχετικά χαρακτηριστικά ενός προσώπου ώστε να είναι σε θέση να διακρίνουν μεταξύ των

εικόνων προσώπου και των μη προσώπων. Τα χαρακτηριστικά αποτελούνται τόσο από τις γεωμετρικές πληροφορίες όσο και από τις φωτομετρικές πληροφορίες [13].

6.10 Αντιπροσωπευτική εργασία ανίχνευσης προσώπου

Οι αντιπροσωπευτικές μέθοδοι των τεσσάρων κατηγοριών που περιεγράφηκαν παραπάνω συνοψίζονται σχήμα 29. Μόνο οι μέθοδοι που βασίζονται στην εμφάνιση περιγράφονται περαιτέρω σε αυτήν την πτυχιακή, καθώς τα ανωτέρα αποτελέσματα φαίνεται να έχουν αναφερθεί χρησιμοποιώντας αυτές τις μεθόδους σε συγκρίσει με τις άλλες τρεις κατηγορίες [16].

Approach	Representative Work
Knowledge-based	Multiresolution rule-based method [57]
Feature invariant - Facial Features - Texture - Skin Color - Multiple Features	Grouping of edges [36] Space Gray-Level Dependence matrix of face pattern [14] Mixture of Gaussian [58] Integration of skin color, size and shape [34]
Template matching - Predefined face templates - Deformable Templates	Shape templates [13] Active Shape Models [35]
Appearance-based method - Eigenfaces & Fisherfaces - Neural Network ¹ - Deformable Models	Eigenvector decomposition and clustering [54] Ensemble of neural networks and arbitration schemes [47] Active Appearance Models [10]

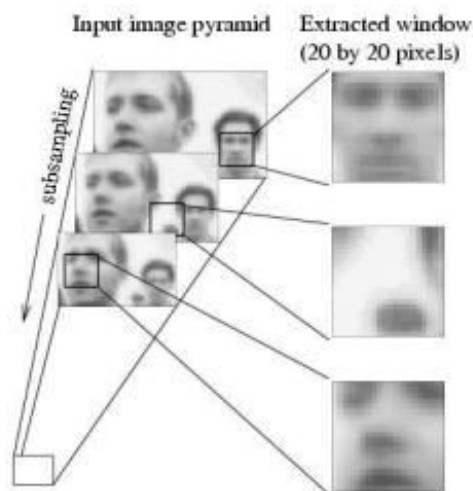
Σχήμα 29 “ Ομαδοποίηση μεθόδων για ανίχνευση προσώπου σε μια μονή εικόνα”

6.11 Περιγραφή των επιλεγόμενων μεθόδων ανίχνευσης προσώπου

Σε αυτήν την ενότητα περιγράφονται οι μέθοδοι EigenFaces, Fisherfaces, Νευρωνικά δίκτυα και Ενεργά μοντέλα εμφάνισης, με ιδιαίτερη έμφαση στα μοντέλα ενεργών εμφάνισης. Τα μοντέλα ενεργών εμφάνισης παρουσιάζουν σαφή πλεονεκτήματα για τους σκοπούς αναγνώρισης προσώπου, τα οποία θα περιγράψουν αργότερα στην πτυχιακή.

6.11.1 Γενικές πτυχές των αλγορίθμων ανίχνευσης προσώπων

Οι περισσότεροι αλγόριθμοι ανίχνευσης προσώπου λειτουργούν με τη συστηματική ανάλυση των υπό περιφερειών μιας εικόνας. Ένα παράδειγμα για τον τρόπο απόκρυψης αυτών των υπό περιφερειών θα μπορούσε να είναι η λήψη εντός δευτερεύοντος εικονοστοιχείου 20x20 pixels. Στην επάνω αριστερή γωνία της αρχικής εικόνας και η συνέχιση της λήψης δευτερευόντων εικόνων σε ένα προκαθορισμένο πλαίσιο. Όλα αυτά τα δευτερεύοντα σήματα αξιολογούνται έπειτα χρησιμοποιώντας αλγόριθμο ανίχνευσης προσώπου. Η υπό δειγματοληψία της εικόνας με τρόπο πυραμίδας που επιτρέπει την σύλληψη διαφορετικών μεγεθών. Αυτό απεικονίζεται στην σχήμα 30[15].

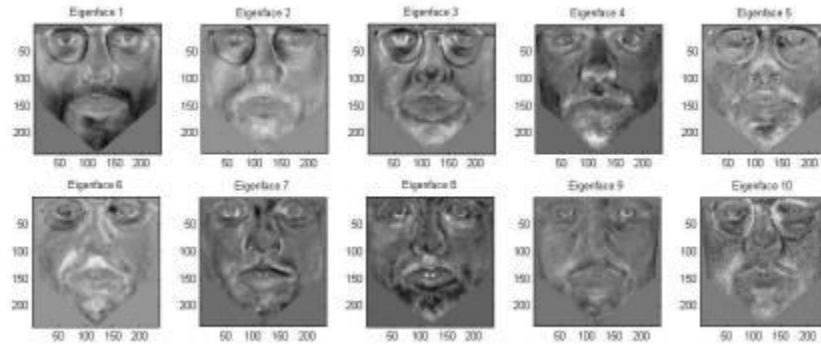


Σχήμα 30 “Απεικόνιση της υποδειγματοληψίας μιας εικόνας υπό μορφή πυραμίδας”

6.11.2 EigenFaces

Η μέθοδος EigenFace χρησιμοποιεί PCA για να δημιουργήσει ένα σύνολο εικόνων EigenFace. Παραδείγματα εικόνων EigenFace εμφανίζονται στο σχήμα 31. Αυτές οι EigenFaces, μπορούν να συνδυαστούν γραμμικά για να αναδημιουργήσουν τις εικόνες του αρχικού σετ εκπαίδευσης. Όταν εισάγετε μια νέα εικόνα, ένα σφάλμα (ξ) μπορεί να υπολογιστεί από την καλύτερη ανακατασκευή εικόνας χρησιμοποιώντας τα EigenFaces

στη νέα εικόνα. Εάν οι EigenFaces κατασκευάζονται από μια βάση δεδομένων μεγάλου προσώπου, το μέγεθος σφάλματος (ξ) μπορεί να χρησιμοποιηθεί για να προσδιοριστεί εάν μια νεοεισαχθείσα εικόνα περιέχει πρόσωπο ή όχι [16].



Σχήμα 31 “Παράδειγμα από 10 EigenFaces”

6.11.3 Fisherfaces

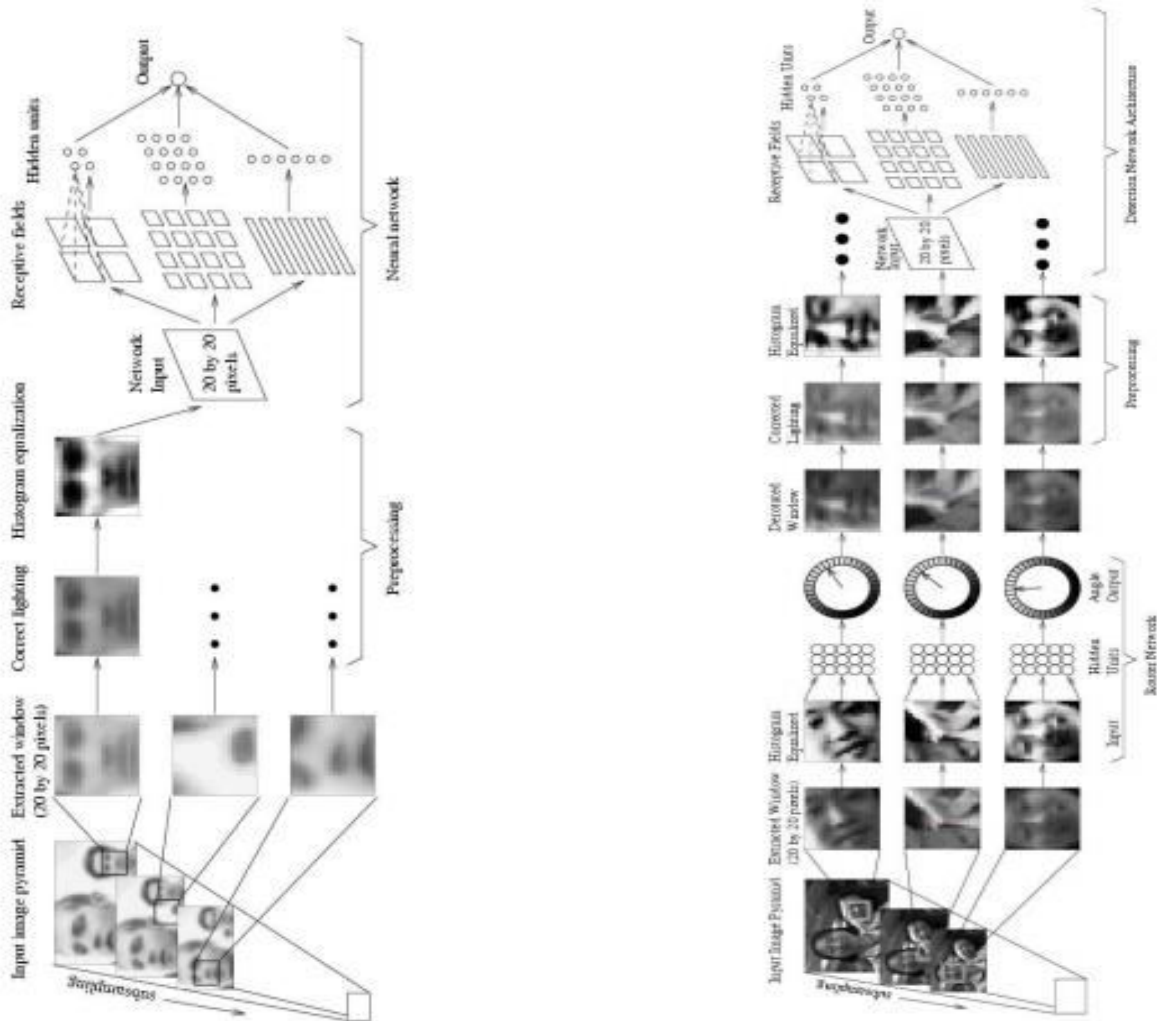
Όπως οι EigenFaces, οι Fisherfaces κατασκευάζουν ένα υποσύστημα στο οποίο ο αλγόριθμος μπορεί να κάνει διάκριση μεταξύ των εικόνων του προσώπου και του μη προσώπου [16].

6.11.4 Νευρωνικά Δίκτυα

Σε μια προσέγγιση νευρωνικού δικτύου τα χαρακτηριστικά από μια εικόνα εξάγονται και τροφοδοτούνται σε ένα νευρωνικό δίκτυο. Ένα τεράστιο μειονέκτημα των νευρωνικών δικτύων είναι ότι μπορούν να συντονιστούν εκτενώς, όσον αφορά τη λήψη αποφάσεων σχετικά με τις μεθόδους μάθησης και τον αριθμό των στρωμάτων, κόμβων κλπ. Ένα από τα σημαντικότερα έργα ανίχνευσης προσώπου νευρωνικών δικτύων έχει γίνει από του Rowley. Χρησιμοποίησε ένα νευρωνικό δίκτυο για να ταξινομήσει τις εικόνες σε [-1,1], όπου -1 και 1 υποδηλώνουν μια εικόνα με πρόσωπο και μια εικόνα χωρίς πρόσωπο, αντίστοιχα. Κάθε παράθυρο 20x20 pixel χωριστικό σε τέσσερα 10x10 pixel, 16 5x5 pixel και 6 υπό παράθυρα των 20x5 pixel [17].

Ένας κρυμμένος κόμβος στο νευρωνικό δίκτυο τροφοδοτήθηκε σε κάθε ένα από αυτά τα δευτερεύοντα παράθυρα, αποδίδοντας συνολικά 26 κρυμμένους κόμβους. Ένα διάγραμμα του σχεδιασμού νευρωνικού δικτύου από του Rowley φαίνεται στο σχήμα

32. Το νευρωνικό δίκτυο μπορεί να βελτιωθεί με την προσθήκη ενός επιπλέον νευρωνικού δικτύου για τον προσδιορισμό της περιστροφής ενός παραθύρου εικόνας. Αυτό θα επιτρέψει στο σύστημα να καταγράψει πρόσωπα που δεν είναι κάθετα ευθυγραμμισμένα στην εικόνα εισόδου, δείτε σχήμα 32 [18].



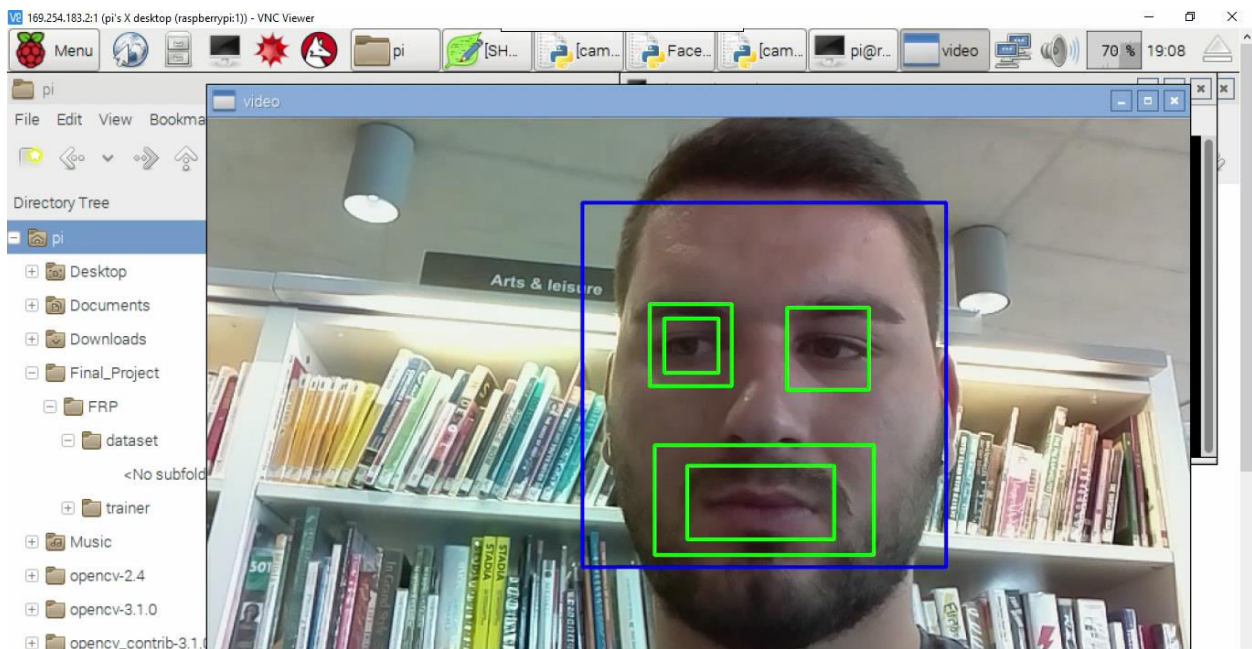
Σχήμα 32 “Εξελιγμένο διάγραμμα νευρωνικού δικτύου”

6.12 Το Πρακτικό Μέρος της Πτυχιακής

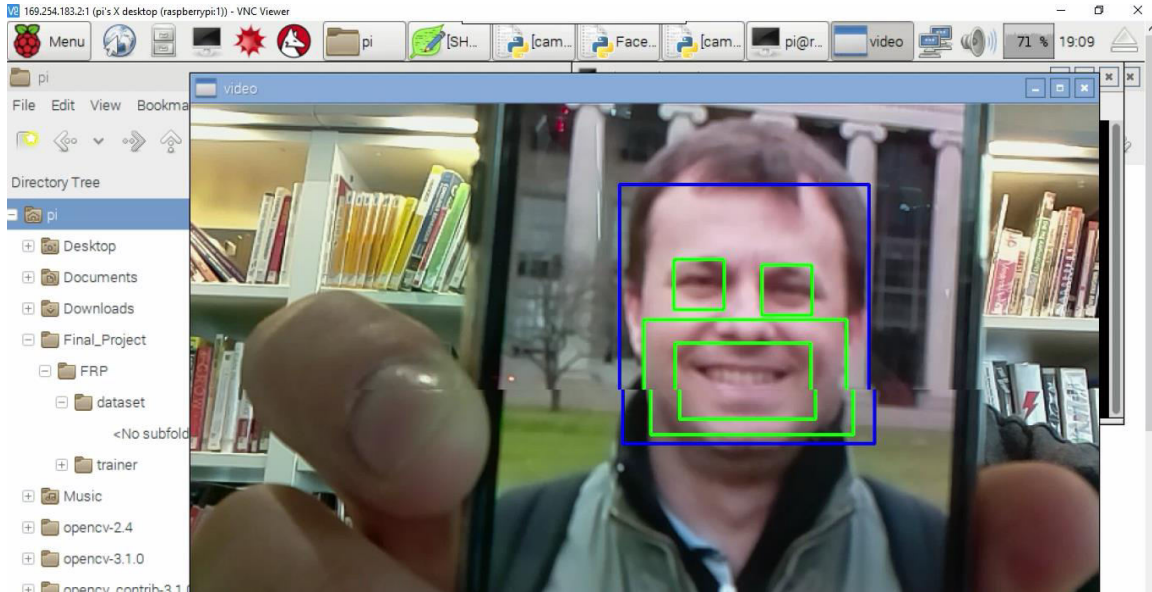
6.12.1 Επεξήγηση

Στα παραπάνω κεφάλαια ήταν μια επεξήγηση στο θεωρητικό κομμάτι της πτυχιακής μου για να καταλάβουμε περίπου την σημασία της αναγνώρισης προσώπου. Έφτασε η ώρα λοιπόν να δούμε πως όλη αυτή η θεωρία που είπαμε παραπάνω φαίνεται σε πρακτικό κομμάτι. Πρώτα από όλα θα δούμε δυο εικόνες οι οποίες η μια είναι ανίχνευση προσώπου όχι μόνο σαν σύνολο αλλά και τα βασικά μέρη του προσώπου όπως μύτη, μάτια, στόμα. Επίσης και ανίχνευση συναισθημάτων όπως το γέλιο.

Η πρώτη εικόνα είναι ανίχνευση προσώπου σε ζωντανή λήψη βίντεου.

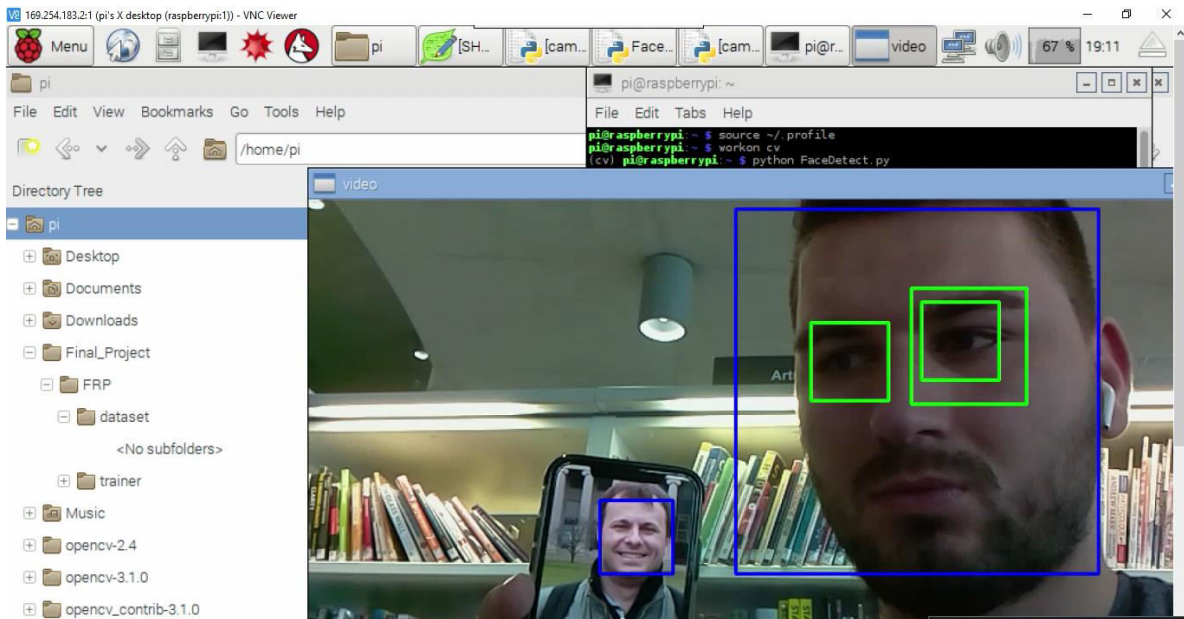


Σχήμα 33 “Ανίχνευση προσώπου σε ζωντανή λήψη βίντεου”



Σχήμα 34 “Ανίχνευση προσώπου και γέλιου σε εικόνα”

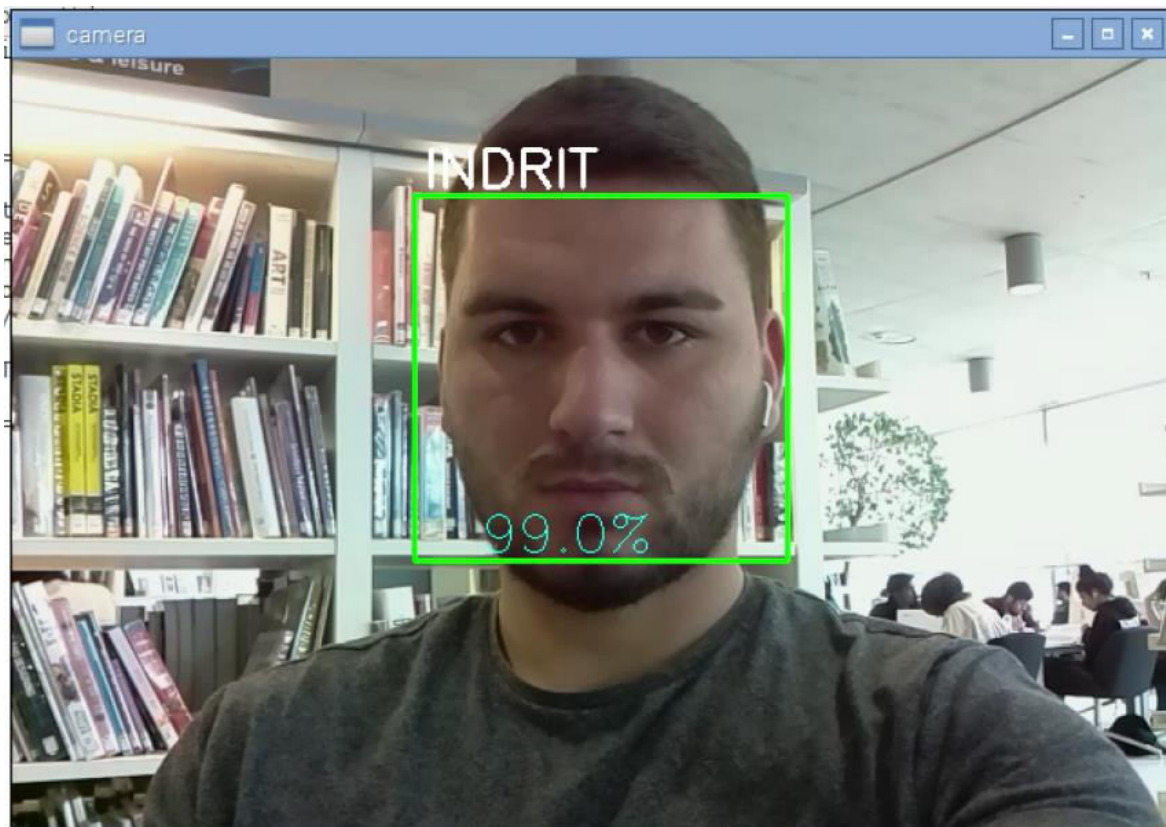
Αφού είδαμε ξεχωριστά την ανίχνευση προσώπου σε μια εικόνα και σε ζωντανή λήψη βίντεου, πάμε να τα δούμε ταυτόχρονα την ανίχνευση προσώπου.



Σχήμα 35 “Ανίχνευση προσώπου και στις δυο παραπάνω περιπτώσεις ταυτόχρονα”

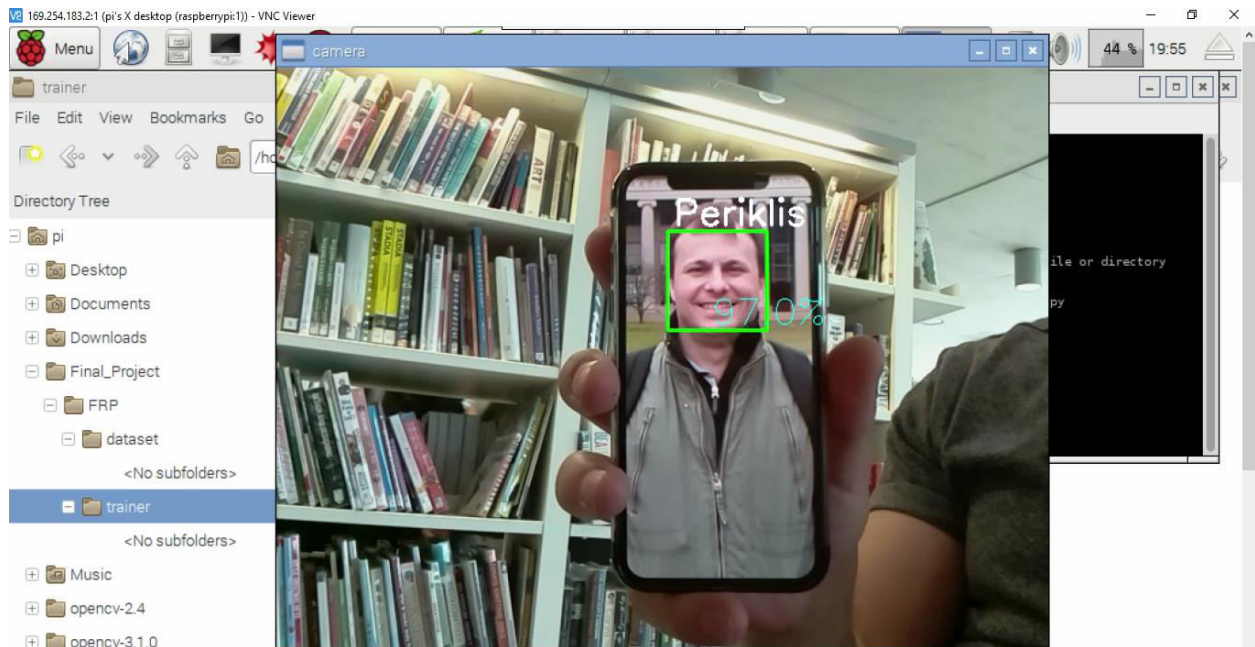
Αυτά ήταν για το κομμάτι της ανίχνευσης προσώπου. Παρακάτω θα δούμε την αναγνώριση προσώπου μέσω μιας εικόνας, ζωντανής λήψης βίντεου ταυτόχρονα.

Η παρακάτω εικόνα αναγνωρίζει το πρόσωπο αναγράφοντας πάνω από το πλαίσιο το όνομα του και κάτω από το πλαίσιο το ποσοστό. Τι σημαίνει το ποσοστό; Το ποσοστό δείχνει ποσό τις 100 ταιριάζει αυτό που βλέπει η κάμερα με αυτά που έχουμε εισάγει στο dataset της εφαρμογής μας. Τα dataset που χρησιμοποιήθηκαν στην εφαρμογή μπορείτε να τα δείτε στο κεφάλαιο 6. Μετά τα dataset έχει γίνει εκπαίδευση των εικόνων αυτών και αφού γίνει η εκπαίδευση αυτή τότε δημιουργείται ένα αρχείο trainer.yml οπού έχοντας το αρχείο αυτό τρέχουμε το κυρίως πρόγραμμα μας και γίνεται η αναγνώριση.

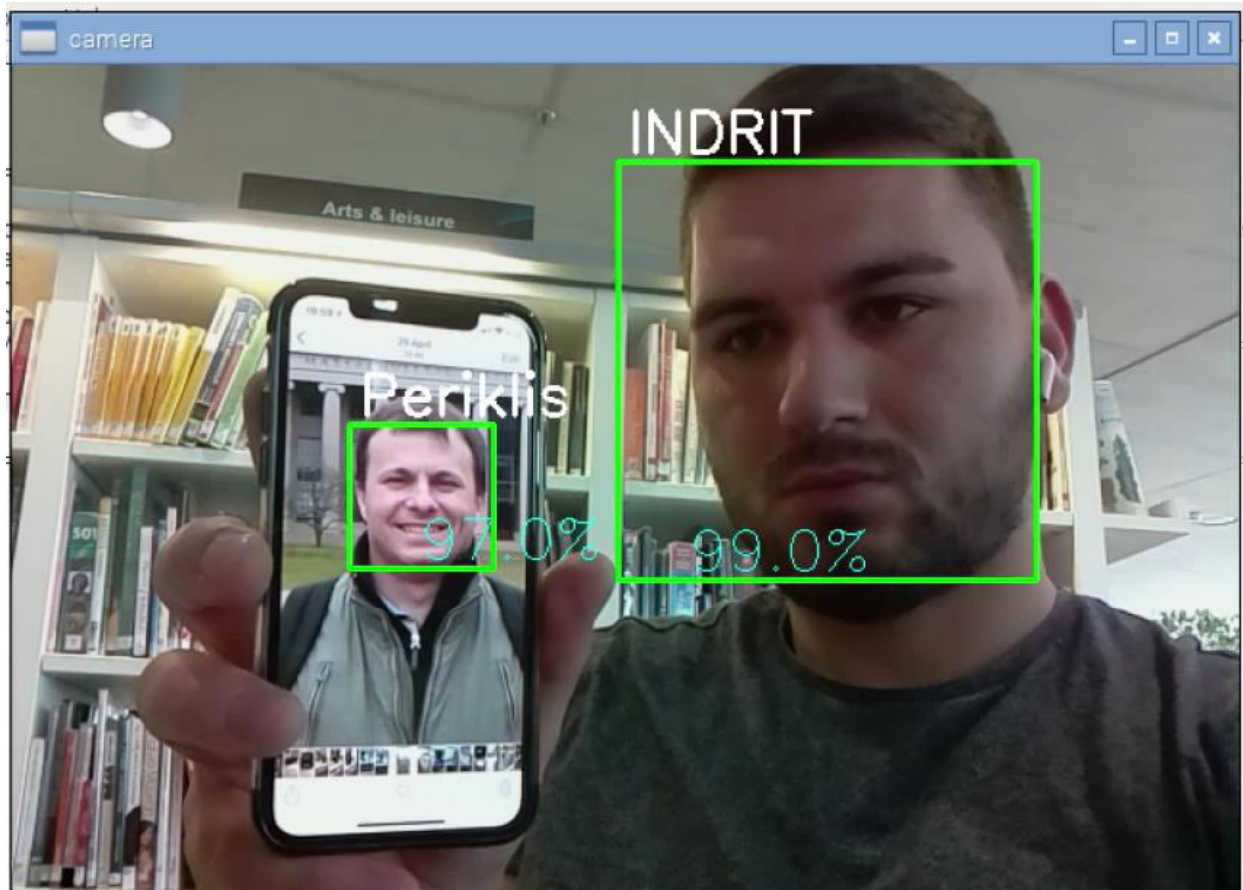


Σχήμα 36 “Αναγνώριση Προσώπου ”

Η δεύτερη εικόνα είναι η αναγνώριση του προσώπου του επιβλέπων καθηγητή μου Περικλής μέσω μιας εικόνας του.



Σχήμα 38 “Αναγνώριση προσώπου σε εικόνα ”



Σχήμα 39 “Αναγνώριση προσώπου και των δυο μας ταυτόχρονα”

6.12.2 Κώδικας

Ανίχνευση Προσώπου:

Πριν προχωρήσουμε στην αναγνώριση προσώπου πάμε να δούμε πρώτα τον κώδικα που χρειάστηκε για να γίνει ανίχνευση προσώπου μέσω OpenCV με γλώσσα Python.

```

import numpy as np
import cv2

faceCascade = cv2.CascadeClassifier('/home/pi/Final_Project/haarcascade_frontalface_default.xml')
eyeCascade = cv2.CascadeClassifier('/home/pi/Final_Project/haarcascade_eye.xml')
smileCascade = cv2.CascadeClassifier('/home/pi/Final_Project/haarcascade_smile.xml')

cap = cv2.VideoCapture(0)
cap.set(3,820) # set Width
cap.set(4,480) # set Height

while True:
    ret, img = cap.read()
    img = cv2.flip(img, -1)
    gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
    faces = faceCascade.detectMultiScale(
        gray,
        scaleFactor=1.3,
        minNeighbors=5,
        minSize=(30, 30)
    )

    for (x,y,w,h) in faces:
        cv2.rectangle(img, (x, y), (x+w, y+h), (255, 0, 0), 2)
        roi_gray = gray[y:y+h, x:x+w]
        roi_color = img[y:y+h, x:x+w]

        eyes = eyeCascade.detectMultiScale(
            roi_gray,
            scaleFactor= 1.5,
            minNeighbors=5,
            minSize=(5, 5),
            )

        for (ex, ey, ew, eh) in eyes:
            cv2.rectangle(roi_color, (ex, ey), (ex + ew, ey + eh), (0, 255, 0), 2)

        smile = smileCascade.detectMultiScale(
            roi_gray,
            scaleFactor= 1.5,
            minNeighbors=15,
            minSize=(25, 25),
            )

        for (xx, yy, ww, hh) in smile:
            cv2.rectangle(roi_color, (xx, yy), (xx + ww, yy + hh), (0, 255, 0), 2)

    cv2.imshow('video', img)

    k = cv2.waitKey(30) & 0xff
    if k == 27: # press 'ESC' to quit
        break

cap.release()
cv2.destroyAllWindows()

```

Αναγνώριση προσώπου:

Face Dataset code:

Ο Κώδικας για να γίνει η συλλογή εικόνας προσώπου και η αποθήκευση τους σε ένα αρχείο ώστε ο επόμενος κώδικας να μπορεί να τα βρίσκει και να τα εκπαιδεύσει.

```
import cv2
import os

cam = cv2.VideoCapture(0)
cam.set(3, 640) # set video width
cam.set(4, 480) # set video height

face_detector = cv2.CascadeClassifier('/home/pi/Final_Project/FRP/haarcascade_frontalface_default.xml')

# For each person, enter one numeric face id
face_id = input('\n enter user id end press <return> ==>  ')

print("\n [INFO] Initializing face capture. Look the camera and wait ...")
# Initialize individual sampling face count
count = 0

while(True):

    ret, img = cam.read()
    img = cv2.flip(img, -1) # flip video image vertically
    gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
    faces = face_detector.detectMultiScale(gray, 1.3, 5)

    for (x,y,w,h) in faces:

        cv2.rectangle(img, (x,y), (x+w,y+h), (255,0,0), 2)
        count += 1

        # Save the captured image into the datasets folder
        cv2.imwrite("dataset/User." + str(face_id) + '.' + str(count) + ".jpg", gray[y:y+h,x:x+w])

        cv2.imshow('image', img)

    k = cv2.waitKey(100) & 0xFF # Press 'ESC' for exiting video
    if k == 27:
        break
    elif count >= 30: # Take 30 face sample and stop video
        break

# Do a bit of cleanup
print("\n [INFO] Exiting Program and cleanup stuff")
cam.release()
cv2.destroyAllWindows()
```

Face Training code:

Αφού ολοκληρώθηκε η συλλογή εικόνας προσώπου τώρα πρέπει να γίνει εκπαίδευση του προσώπου.

```

import cv2
import numpy as np
from PIL import Image
import os

# Path for face image database
path = 'dataset'

recognizer = cv2.face.createLBPHFaceRecognizer()
recognizer.save("trainer.yml")
detector = cv2.CascadeClassifier("/home/pi/Final_Project/FRF/haarcascade_frontalface_default.xml");

# function to get the images and label data
def getImagesAndLabels(path):

    imagePaths = [os.path.join(path, f) for f in os.listdir(path)]
    faceSamples=[]
    ids = []

    for imagePath in imagePaths:

        PIL_img = Image.open(imagePath).convert('L') # convert it to grayscale
        img_numpy = np.array(PIL_img, 'uint8')

        id = int(os.path.splitext(imagePath)[-1].split(".")[1])
        faces = detector.detectMultiScale(img_numpy)

        for (x,y,w,h) in faces:
            faceSamples.append(img_numpy[y:y+h,x:x+w])
            ids.append(id)

    return faceSamples,ids

print ("\n [INFO] Training faces. It will take a few seconds. Wait ...")
faces,ids = getImagesAndLabels(path)
recognizer.train(faces, np.array(ids))

# Save the model into trainer/trainer.yml
recognizer.save("trainer/trainer.yml") # recognizer.save() worked on Mac, but not on Pi

# Print the numer of faces trained and end program
print ("\n [INFO] {0} faces trained. Exiting Program".format(len(np.unique(ids))))

```

Το αρχείο που δημιουργείτε μετά την εκτέλεση του κώδικα αυτού είναι το παρακάτω το οποίο είναι ένας πίνακας Matrix του OpenCV, η οποία έχει μετατρέψει τα δεδομένα του dataset σε πληροφορίες κατάλληλες για μηχανική μάθηση.



trainer.yml

Face_Recognition_Code:

Τέλος, αφού έχουμε ολοκληρώσει τα προηγούμενα βήματα με επιτυχία πρέπει να δημιουργήσουμε το κυρίως πρόγραμμα ώστε να λειτουργήσει η αναγνώριση προσώπου

Για να λειτουργήσει ο κώδικας χρειαζόμαστε τα αρχεία haarcascade τα οποία είναι απαραίτητα για την αναγνώριση προσώπου. Αυτά τα αρχεία υπάρχουν έτοιμα μετά την εγκατάσταση του OpenCV. Συγκεκριμένα στο αρχείο DATA του OpenCV. Στην δικιά μου περίπτωση χρησιμοποίησα τα παρακάτω xml αρχεία:



haarcascade_frontalface_default.xml



haarcascade_eye.xml



haarcascade_smile.xml

Υπάρχουν πολλά περισσότερα αρχεία για πολλά αλλά είδους εργασιών.

```
import cv2
import rumpy as rp
import time
import os

recognizer = cv2.face.createLBPHFaceRecognizer()
recognizer.load('trainer/trainer.yml')
cascadePath = "haarcascade_frontalface_default.xml"
faceCascade = cv2.CascadeClassifier(cascadePath);

font = cv2.FONT_HERSHEY_SIMPLEX

#iniciate id counter
id = 0

# names related to ids: example ==> indrit: id=1, etc
names = ['None', 'INDRIT', 'Gerry', 'Periklis', 'Kostas' ]

# Initialize and start realtime video capture
cam = cv2.VideoCapture(0)
cam.set(3, 640) # set video width
cam.set(4, 480) # set video height

# Define min window size to be recognized as a face
minW = 0.1*cam.get(3)
minH = 0.1*cam.get(4)

while True:

    ret, img =cam.read()
    img = cv2.flip(img, -1) # Flip vertically
```



```

gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)

faces = faceCascade.detectMultiScale(
    gray,
    scaleFactor = 1.2,
    minNeighbors = 5,
    minSize = (int(minW), int(minH)),
)

for (x, y, w, h) in faces:

    cv2.rectangle(img, (x, y), (x+w, y+h), (0, 255, 0), 2)

    id = recognizer.predict(gray[y:y+h, x:x+w])
    confidence = recognizer.predict(gray[y:y+h, x:x+w])

    # id=None
    # confidence=None
    # print('x=', x, 'y=', y, 'w=', w, 'h=', h)

    # Check if confidence is less than 100 ==> "0" is perfect match
    if (confidence < 100):
        id = names[id]
        confidence = " {0}%".format(round(100 - confidence))
    else:
        id = "unknown"
        confidence = " {0}%".format(round(100 - confidence))

    cv2.putText(img, str(id), (x+5, y-5), font, 1, (255, 255, 255), 2)
    cv2.putText(img, str(confidence), (x+5, y+h-5), font, 1, (255, 255, 0), 1)

cv2.imshow('camera', img)

k = cv2.waitKey(10) & 0xff # Press 'ESC' for exiting video
if k == 27:
    break

# Do a bit of cleanup
print("\n [INFO] Exiting Program and cleanup stuff")
cam.release()
cv2.destroyAllWindows()

```


ΚΕΦΑΛΑΙΟ 7 - Συμπεράσματα και Μελλοντική Εργασία

7.1 Συμπεράσματα

Η αύξηση κατανάλωσης αγαθών όπως τα αυτοκίνητα και οι οικιακές συσκευές προβλέπει μια νέα αγορά για τεχνολογίες όπως η αναγνώριση προσώπου. Επιπλέον, λόγω των διάφορων γεγονότων που συμβαίνουν στον κόσμο προκάλεσαν δημόσια ζήτηση για ιδιωτική ασφάλεια και ασφάλεια σε δημοσίου χώρους. Απαιτήσεις που θα μπορούσαν να επιτευχθούν εν μέρει με την ανάπτυξη αυτών των τεχνολογιών. Στόχοι της πτυχειακής είναι να συζητήσουμε και να συνοψίζουμε την διαδικασία αναγνώρισης προσώπου, να εξετάζουμε να τις διαθέσιμες τεχνικές αναγνώρισης του προσώπου, να σχεδιάζουμε και να αναπτύξουμε όσο το δυνατόν έναν αρκετά ικανοποιητικό αλγόριθμο αναγνώρισης προσώπου με την βοήθεια του Raspberry Pi.

7.2 Μελλοντική εργασία

(α) Ανίχνευση προσώπου

Σύμφωνα με την εφαρμογή, η σωστή ανίχνευση προσώπου είναι πολύ σημαντική σε μια διαδικασία αναγνώρισης προσώπου. Για το μέλλον θα μπορούσε η απόδοση αναγνώρισης προσώπου να βελτιωθεί σημαντικά ελαχιστοποιώντας την ψευδή ανίχνευση των προσώπων. Όπως είπαμε προηγούμενος μερικοί αλγόριθμοι δυσκολεύονται να αναγνωρίζουν το πρόσωπο σε μια εικόνα δεδομένου ότι δεν έχει εκπαιδευτεί να χειρίζεται την μεταβολή των εικόνων μη προσώπου με αποτέλεσμα να έχουμε πιθανή εσφαλμένη ταξινόμηση. Επίσης θα μπορούσαμε στο μέλλον να κάνουμε τους αλγόριθμους να μην είναι τόσο ευαίσθητοι σε συνθήκες φωτισμού, οπού αυτό θα ήταν μια εξέλιξη με μεγάλο ενδιαφέρον.

(β) Ανακατασκευή 3D Προσώπου

Η αναδόμηση προσώπου σε 3D από 2D εικόνες είναι μια πάρα πολύ ενδιαφέρουσα εξέλιξη. Ο τελικός στόχος είναι η εξαγωγή περισσότερων διακριτικών πληροφοριών 3D

από τις οποίες θα μπορούσε να επιτευχθεί καλύτερη αναγνώριση προσώπου σε σύγκριση με των 2D εικόνων. Επίσης ενδιαφέρον θέμα θα αποτελούσε η αναγνώριση μη μετωπικών προσώπων μέσω 3D εικόνων. Το ενδιαφέρον εδώ είναι το όριο στην αναγνώριση προσώπου, αφού η ανακατασκευή ενός τρισδιάστατου σχήματος χωρίς την χρήση τρισδιάστατων σαρωτών θα ήταν χρήσιμη σε πολλές εφαρμογές.

Βιβλιογραφία

- [1] David Emanuel Ribeiro Gaspar, Raspberry Pi: a Smart Video Monitoring Platform, Embedded Systems and Computer Engineering, November 2014
- [2] Lopez Research, An Introduction to the Internet of Things (IoT), November 2013
- [3] Carl-Martin Ivask, Raspberry Pi based System for Visual Object Detection and Tracking Bachelor's Thesis, Faculty of Information Technology Department of Computer Control, 2015
- [4] K. Revathi¹ , Jaya Bharathi.M² , Saranya .U, Face recognition using image processing for visuallychallenged, , July 2015
- [5] Jens Fagertun, Face Recognition, Technical University of Denmark Informatics and Mathematical Modelling, 2005
- [6] Igor Petrov, Raspberry Pi based System for Visual Detection of Fluid Level, Tallinn University of Technology Faculty of Information Technology Department of Computer Control, 2014
- [7] Otieno, Ezra Odondi, Computer Vision through the Raspberry-PI: Counting Objects, The University of Nairobi Department of Electrical and Information Engineering,2016
- [8] Trygve Christian Eftestøhl, Development of face recognition system for use on the NAO robot, FACULTY OF SCIENCE AND TECHNOLOGY, 2016
- [9] Lalendra Sumitha Balasuriya, Frontal View Human Face Detection and Recognition, Department of Statistics and Computer Science University of Colombo, 2000
- [10] ISAC TÖRNBERG, Real time object tracking on Raspberry Pi 2, STOCKHOLM, 2016
- [11] Adam Eriksson och Jonas Holmberg, Real time image recognition of facial features for detecting a true breath,2015
- [12] Alejandro Ariza Bueno, Eye detection and face recognition in watch-list applications, Master in Science in Telecommunication Engineering & Management,2014
- [13] Richard Szeliski, Computer Vision: Algorithms and Applications, September 3, 20
- [14] Adrian Rosebrock, Raspberry Pi 3 +Raspbian Jessie + OpenCV, 2016
- [15] Ashwin Pajankar, Raspberry Pi Computer Vision Programming, May 2015
- [16] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. IEEE Trans. Pattern Analysis and Machine Intelligence, July 1997
- [17] M. Bichsel and A. Pentland. Human face recognition and the face image set's topology. Computer Vision, Graphics, and Image Processing, 1994.
- [18] H. Cevikalp, M. Neamtu, M. Wilkes, and A. Barkana. Discriminative common vectors for face recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005.

- [19] Alpert, S., Galun, M., Basri, R., and Brandt, A. , Image segmentation by probabilistic bottom-up aggregation and cue integration. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition ,Minneapolis, 2007
- [20] C. Darwin. The Expression of the Emotions in Man and Animals. London: John Murray, 1872
- [21] J. Fagertun and M. B. Stegmann. The IMM frontal face database. Technical Report 11, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Aug 2005.
- [22] M. A. Fischler and R. A. Elschlager. The representation and matching of pictorial structures. IEEE Transactions on Computers, 1973
- [23] I. Gauthier and M. Behrmann. Can face recognition really be dissociated from object recognition? Journal of Cognitive Neuroscience, 1999
- [24] D. D. Gomez, J. Fagertun, and B. K. Ersbøll. A face recognition algorithm based on multiple individual discriminative models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005.
- [25] A. Lanitis, C. J. Taylor, and T. F. Cootes. An automatic face identification system using flexible appearance models. Image and Vision Computing, 1995