



ΣΥΣΤΗΜΑ ΑΝΑΓΝΩΡΙΣΗΣ ΠΡΟΣΩΠΩΝ ΜΕ ΧΡΗΣΗ WEB CAMERA ΚΑΙ OPENCV

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Κωνσταντίνος Βλάχος

ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΣΕΡΡΩΝ
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ & ΕΠΙΚΟΙΝΩΝΙΩΝ

ΣΥΣΤΗΜΑ ΑΝΑΓΝΩΡΙΣΗΣ ΠΡΟΣΩΠΩΝ ΜΕ
ΧΡΗΣΗ WEB CAMERA ΚΑΙ OPENCV
ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Κωνσταντίνος Βλάχος, Α.Μ. 1049

ΣΕΡΡΕΣ, 2013

Υπεύθυνη Δήλωση: Βεβαιώνω ότι είμαι συγγραφέας αυτής της πτυχιακής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της, είναι πλήρως αναγνωρισμένη και αναφέρεται στην πτυχιακή εργασία. Επίσης έχω αναφέρει τις όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε αυτές αναφέρονται ακριβώς είτε παραφρασμένες. Επίσης βεβαιώνω ότι αυτή η πτυχιακή εργασία προετοιμάστηκε από εμένα προσωπικά ειδικά για τις απαιτήσεις του προγράμματος σπουδών του τμήματος Πληροφορικής & Επικοινωνιών του Τ.Ε.Ι. Σερρών.

ΣΥΣΤΗΜΑ ΑΝΑΓΝΩΡΙΣΗΣ ΠΡΟΣΩΠΩΝ ΜΕ ΧΡΗΣΗ WEB CAM-
ERA ΚΑΙ OPENCV

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Κωνσταντίνος Βλάχος, Α.Μ. 1049

Επιβλέπων: Δρ. Γεώργιος Παυλίδης
Επιστημονικός Συνεργάτης ΤΕΙ Σερρών

ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΣΕΡΡΩΝ
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ & ΕΠΙΚΟΙΝΩΝΙΩΝ

στη μητέρα μου Ικασία

Περίληψη

Οι εφαρμογές που απαιτούν την εξακρίβωση της ταυτότητας ενός ατόμου πολλαπλασιάζονται συνεχώς και μαζί με αυτές αυξάνεται η ανάγκη για τη χρήση μεθόδων πιστοποίησης που δεν θα απαιτούν από τους ανθρώπους να θυμούνται πολύπλοκους κωδικούς πρόσβασης ή να μεταφέρουν μαγνητικές κάρτες, ταυτότητες ή άλλου είδους έγγραφα που θα επαληθεύουν την ταυτότητα τους. Η χρήση βιομετρικών μεθόδων πιστοποίησης, δηλαδή μεθόδων που στηρίζονται στην αναγνώριση φυσικών χαρακτηριστικών ή χαρακτηριστικών της συμπεριφοράς του ανθρώπου είναι μια λύση σε αυτό το πρόβλημα. Μεταξύ των διαφόρων βιομετρικών μεθόδων που χρησιμοποιούνται, όπως η αναγνώριση ίριδας, των δακτυλικών αποτυπωμάτων, της γεωμετρίας της παλάμης, της στάσης του σώματος και του βαδίσματος, της φωνής, ακόμη και του τρόπου πληκτρολόγησης ή υπογραφής, η αναγνώριση προσώπου είναι σίγουρα μία από τις φιλικότερες προς το χρήστη, γιατί ουσιαστικά μιμείται τον τρόπο που χρησιμοποιούν οι άνθρωποι για να αναγνωρίζουν ο ένας τον άλλο.

Abstract

In recent years, more and more applications require the identification of a person along with the need for using methods of certification that will not require remembering complicated passwords or carrying magnetic cards, identification cards or other identification documents. The use of biometric methods of identification based on natural features or characteristics of human behavior is one solution to this problem. Among the various biometric methods already being used, such as the recognition of iris, fingerprint, hand geometry, posture and gait, voice, or even the way of typing or signing, face recognition is definitely one of the most user-friendly as it mimics the natural way people recognize each other.

Πρόλογος

Ένα σύστημα αναγνώρισης προσώπου (facial recognition system) ασχολείται με την αυτόματη αναγνώριση ή ταυτοποίηση ενός προσώπου από μια ψηφιακή εικόνα ή από ένα βίντεο. Ένας από τους τρόπους για να πραγματοποιηθεί η αναγνώριση, είναι συγκρίνοντας επιλεγμένα χαρακτηριστικά του προσώπου από την εικόνα με τα χαρακτηριστικά από άλλα πρόσωπα σε μια βάση δεδομένων. Χρησιμοποιείται συνήθως σε συστήματα ασφαλείας και μπορεί να συγκριθεί με άλλα βιομετρικά συστήματα, όπως τα συστήματα αναγνώρισης δακτυλικών αποτυπωμάτων ή ίριδας ματιού.

Το πρόβλημα της αναγνώρισης προσώπου είναι βασικά ένα πρόβλημα ταξινόμησης όπου κάθε πρόσωπο στη βάση δεδομένων αναπαριστά μια κατηγορία στην οποία τα εισερχόμενα δεδομένα πρόκειται να αντιστοιχιστούν. Το πρόβλημα της ταξινόμησης επιδεινώνεται με το μέγεθος των δεδομένων. Κατά τα τελευταία δέκα χρόνια περίπου, η αναγνώριση προσώπου έχει γίνει ένα δημοφιλές θέμα έρευνας και μελέτης για τους επιστήμονες και επίσης μια από τις πιο πετυχημένες εφαρμογές ανάλυσης εικόνας. Λόγω της φύσης του προβλήματος, δεν είναι μόνο ερευνητές της επιστήμης των υπολογιστών που ενδιαφέρονται για το θέμα αυτό, αλλά νευροεπιστήμονες και ψυχολόγοι επίσης.

Κίνητρο για τη διεξαγωγή αυτής της πτυχιακής εργασίας υπήρξε ο συνεχώς αυξανόμενος ρόλος που διαδραματίζει η τεχνολογία σε όλους τους τομείς της κοινωνίας μας. Ενώ μέχρι πρόσφατα η αναγνώριση προσώπων ήταν δύσκολη υπόθεση, στις μέρες μας ολοένα και περισσότερες τεχνικές αναπτύσσονται με αξιόλογα αποτελέσματα.

Σκοπός αυτής της πτυχιακής εργασίας είναι η διερεύνηση του προβλήματος της υλοποίησης συστημάτων εντοπισμού και αναγνώρισης προσώπων με βάση κάμερες και ειδικότερα κάμερες διαδικτύου χαμηλού κόστους και η υλοποίηση συστήματος αναγνώρισης προσώπων με την εκμετάλλευση των δυνατοτήτων που παρέχει η βιβλιοθήκη προγραμματισμού OpenCV και ειδικότερα η μεθοδολογία της ανάλυσης κυρίων συνιστωσών (Principal Components Analysis – PCA) με χρήση των ιδιοεικόνων προσώπων (eigenfaces).

Η εργασία διαρθρώνεται ως εξής:

- ✓ Στο Πρώτο Κεφάλαιο παρουσιάζονται οι έννοιες, τεχνικές και τα αντικείμενα της βιομετρικής και της αναγνώρισης προσώπων
- ✓ Στο Δεύτερο Κεφάλαιο περιγράφονται οι τεχνικές αναγνώρισης προσώπων και η χρησιμοποίηση των eigenfaces
- ✓ Στο τρίτο κεφάλαιο περιγράφεται η εφαρμογή που υλοποιήθηκε με στόχο την αναγνώριση προσώπων με χρήση απλού εξοπλισμού

χαμηλού κόστους. Αρχικά περιγράφονται οι απαιτήσεις υλοποίησης, η χρήση της γλώσσας προγραμματισμού C++ και η βιβλιοθήκη προγραμματισμού OpenCV. Στη συνέχεια αναλύεται η εφαρμογή που αναπτύχθηκε και τα πειραματικά αποτελέσματα που ελήφθησαν. Αναφέρονται τα συμπεράσματα από τα πειραματικά αποτελέσματα και συζητείται η μελλοντική εξέλιξη των τεχνικών αναγνώρισης προσώπων

- ✓ Στο παράρτημα Α παρέχονται οι οδηγίες εγκατάστασης και χρήσης, καθώς και ο πηγαίος κώδικας της πιλοτικής εφαρμογής που αναπτύχθηκε.

Η πτυχιακή αυτή οφείλεται στη βοήθεια κάποιων ανθρώπων που θέλω να ευχαριστήσω θερμά εδώ. Θέλω να εκφράσω τις ευχαριστίες μου στον κ. Γεώργιο Παυλίδη για την εμπιστοσύνη που μου έδειξε, την καθοδήγηση του και τις άψογες συνθήκες συνεργασίας μαζί του, αλλά και την ευκαιρία που μου έδωσε να ασχοληθώ με ένα τόσο ενδιαφέρον αντικείμενο μελέτης. Δεν θα μπορούσα να παραλείψω να ευχαριστήσω τους γονείς μου Γιώργο και Ικασία, την αδερφή μου Μόνικα και την Έφη για την ηθική και υλική υποστήριξη τους όλα αυτά τα χρόνια.

*Κωνσταντίνος Βλάχος
Σέρρες, 2013*

ΕΙΚΟΝΕΣ

Σχήμα 1. Λειτουργικό τμήμα ενός προσαρμοζόμενου συστήματος αναγνώρισης προτύπου	20
Σχήμα 2. Διάγραμμα ροής ενός συστήματος αναγνώρισης προσώπου ...	22
Σχήμα 3. Εικόνες ιδιοδιανυσμάτων (eigenfaces)	25
Σχήμα 4. Παράδειγμα χρήσης της ΑΚΣ για τη μείωση διάστασης σε 2Δ δεδομένα	25
Σχήμα 5. Προσέγγιση του τετραγωνικού κύματος με σειρές Fourier	27
Σχήμα 6. Ανακατασκευή εικόνας με υπολογισμό και άθροισμα των βαρών	28
Σχήμα 7. Μη αποδοτική και αποδοτική ταξινόμηση	29
Σχήμα 8. Εικόνες εκπαίδευσης.....	30
Σχήμα 9. Έξι eigenfaces για την διαδικασία της εκπαίδευσης διαλεγμένα από τη βάση δεδομένων MIT-CBCL, με τυχαία σειρά.....	32
Σχήμα 10. Πέντε (5) εικόνες στο σύνολο εκπαίδευσης	35
Σχήμα 11. Υποχώρος προσώπου	35
Σχήμα 12. Γραφικό περιβάλλον του Dev-C++	39
Σχήμα 13. Μπλοκ διάγραμμα πρώτου δομικού στοιχείου του συστήματος	44
Σχήμα 14. Μπλοκ διάγραμμα δεύτερου δομικού στοιχείου του συστήματος	45
Σχήμα 15. Λειτουργικό διάγραμμα συστήματος αναγνώρισης προσώπου	45
Σχήμα 16. Δείγμα εικόνας προσώπου από την κάμερα	46
Σχήμα 17. Δείγμα εικόνας προσώπου μετά από τη μετατροπή σε εικόνα γκριζών τόνων	47
Σχήμα 18. Δείγμα εξομαλυμένης και κανονικοποιημένης εικόνας προσώπου	47
Σχήμα 19. Δείγμα φωτογραφίας κατά το σενάριο πειραματισμών 1	48
Σχήμα 20. Δείγμα φωτογραφίας κατά το σενάριο πειραματισμών 2	48
Σχήμα 21. Δείγμα φωτογραφίας κατά το σενάριο πειραματισμών 3	48
Σχήμα 22. Δείγμα φωτογραφίας κατά το σενάριο πειραματισμών 4	49
Σχήμα 23. Δείγμα από τη βάση δεδομένων όπου εμφανίζονται τα πρώτα 64 eigenfaces	49
Σχήμα 24. Αναγνώριση τυπικού προσώπου	50
Σχήμα 25. Αναγνώριση προσώπου με γυαλιά ηλίου	51
Σχήμα 26. Αναγνώριση προσώπου με καπέλο	51
Σχήμα 27. Αναγνώριση προσώπου με καπέλο και γυαλιά ηλίου.....	51
Σχήμα 28. Αναγνώριση προσώπου με σκούφο	51
Σχήμα 29. Αναγνώριση προσώπου με σκούφο και γυαλιά ηλίου	52
Σχήμα 30. Αναγνώριση προσώπου από φωτογραφία	52
Σχήμα 31. Διάγραμμα αποδοτικότητας.....	52
Σχήμα 32. Ιστοσελίδα λήψης της βιβλιοθήκης προγραμματισμού OpenCV	57

Σχήμα 33. Βήματα εγκατάστασης της OpenCV	58
Σχήμα 34. Ιστοσελίδα λήψης της Dev-C++	58
Σχήμα 35. Βήματα εγκατάστασης του Dev-C++	59
Σχήμα 36. Μενού «Εργαλεία» και επιλογές του μεταγλωττιστή	59
Σχήμα 37. Επιλογή του συμβόλου + για προσθήκη νέου μεταγλωττιστή με όνομα OpenCV	60
Σχήμα 38. Προσθήκη των συγκεκριμένων εντολών	60
Σχήμα 39. Προσθήκη φακέλου όπου βρίσκονται τα dll (dynamic library files)	60
Σχήμα 40. Προσθήκη φακέλου επικεφαλίδων για συγγραφή κώδικα C ..	61
Σχήμα 41. Προσθήκη φακέλου επικεφαλίδων για συγγραφή κώδικα C++	61
Σχήμα 42. Εκτέλεση του αρχείου	61
Σχήμα 43. Γραφικό περιβάλλον της εφαρμογής.....	62
Σχήμα 44. Επιλογή εγγραφής νέου φοιτητή	62
Σχήμα 45. Προετοιμασία για συλλογή εικόνων προσώπου	62
Σχήμα 46. Αναγνώριση προσώπου (εάν υπάρχει στη βάση δεδομένων) .	63
Σχήμα 47. Εισαγωγή προσωπικών στοιχείων φοιτητή	63
Σχήμα 48. Επιλογή αποθήκευσης προσωπικών δεδομένων	63
Σχήμα 49. Αποθήκευση προσώπου στη βάση δεδομένων	64
Σχήμα 50. Ολοκλήρωση αποθήκευσης προσώπων	64
Σχήμα 51. Επισκόπηση λεπτομερειών της εκπαίδευσης	65
Σχήμα 52. Σύνδεση και αποσύνδεση χρήστη	65
Σχήμα 53. Σύνδεση φοιτητή	65
Σχήμα 54. Άγνωστο πρόσωπο	66
Σχήμα 55. Αποσύνδεση φοιτητή	66
Σχήμα 56. Στοιχεία συγγραφέα της εφαρμογής και έξοδος από την εφαρμογή	66

ΠΙΝΑΚΕΣ

Πίνακας I. Αναγνώριση χρηστών ανάλογα με τα σενάρια των πειραμάτων	50
Πίνακας II. Ποσοστό βεβαιότητας στην αναγνώριση των χρηστών των πειραμάτων.....	50

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1. ΒΙΟΜΕΤΡΙΚΗ	11
Τι είναι η βιομετρική.....	11
Το πρόσωπο και η αναγνώριση του	13
ΚΕΦΑΛΑΙΟ 2. ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΤΗΣ ΑΝΑΓΝΩΡΙΣΗΣ ΠΡΟΤΥΠΩΝ	19
Εισαγωγή.....	19
Κλάσεις προτύπων και πρότυπα	20
Εκπαίδευση και Εκμάθηση	20
Περίγραμμα ενός συστήματος αναγνώρισης προτύπων	20
Αναγνώριση προσώπου.....	21
ΑΚΣ-Ανάλυση Κύριων Συνιστωσών (PCA–Principal Component Analysis)	24
Αναγνώριση προσώπου με χρήση eigenfaces	26
Προσέγγιση με βάση τη Θεωρία της Πληροφορίας	29
Ο αλγόριθμος για την εύρεση eigenfaces	30
Εύρεση των βαρών.....	32
Η διαδικασία της αναγνώρισης.....	32
Μετρήσεις αποστάσεων (distance measures)	33
Επιλογή κατωφλίου (threshold)	34
Υποχώρος προσώπου (face space).....	35
ΚΕΦΑΛΑΙΟ 3. ΣΥΣΤΗΜΑ ΑΝΑΓΝΩΡΙΣΗΣ ΠΡΟΣΩΠΩΝ ΜΕ ΧΡΗΣΗ EIGENFACES.....	37
Απαιτήσεις υλοποίησης συστήματος αναγνώρισης προσώπων	37
Πρόγραμμα εργασιών υλοποίησης	37
Περί της γλώσσας προγραμματισμού C++	38
Ανάπτυξη σε περιβάλλον Dev-C++.....	39
Open Source Computer Vision Library - OpenCV	39
Ανίχνευση προσώπων με την OpenCV	40
Η συνάρτηση ανίχνευσης.....	41
Οι Μέθοδοι ανίχνευσης	43
Υλοποίηση του συστήματος αναγνώρισης προσώπων.....	43
Πειράματα και συμπεράσματα.....	46
Σκέψεις για βελτιώσεις και μελλοντικές εφαρμογές	53
ΒΙΒΛΙΟΓΡΑΦΙΑ	55
ΠΑΡΑΡΤΗΜΑ Α	57
Οδηγίες εγκατάστασης υποδομής υλοποίησης.....	57
Οδηγίες χρήσης της εφαρμογής από το CD	61

Κεφάλαιο 1. Βιομετρική

Τι είναι η βιομετρική

Στο σημερινό περιβάλλον της αυξημένης ανάγκης για οργάνωση και ασφάλεια, οι μέθοδοι της ταυτοποίησης και αναγνώρισης έχουν εξελιχθεί σε ένα αναγκαίο χαρακτηριστικό της σημερινής τεχνολογίας και βρίσκουν εφαρμογή στις εισόδους κτιρίων, στον έλεγχο πρόσβασης σε υπολογιστές, στα ΑΤΜ ή και σε υπηρεσίες του ταχυδρομείου. Μεγάλο ρόλο παίζουν στην πάταξη του εγκλήματος. Η ανάγκη για αξιόπιστη ταυτοποίηση έχει ως αποτέλεσμα την αύξηση του ενδιαφέροντος για τη χρήση βιομετρικών στοιχείων και συστημάτων.

Η βιομετρική ταυτοποίηση είναι η τεχνική του αυτόματου εντοπισμού και επαλήθευσης ενός ατόμου μέσω ενός φυσικού χαρακτηριστικού γνώρισματος ή προσωπικού χαρακτηριστικού. Ο όρος «αυτόματου εντοπισμού» σημαίνει ότι το βιομετρικό σύστημα αναγνώρισης πρέπει να προσδιορίσει ή να επαληθεύσει ένα φυσικό χαρακτηριστικό γνώρισμα ή προσωπικό χαρακτηριστικό γρήγορα με ελάχιστη ή καμία παρέμβαση από το χρήστη. Τα συστήματα βιομετρικής τεχνολογίας αναπτύχθηκαν για χρήση σε υψηλού επιπέδου συστήματα ασφαλείας και επιβολής του νόμου.

Το βασικό στοιχείο της βιομετρικής τεχνολογίας είναι η ικανότητά της να εντοπίζει έναν άνθρωπο και να επιβάλει την ασφάλεια. Βιομετρικό στοιχείο είναι ένα μοναδικό και μετρήσιμο χαρακτηριστικό του ανθρώπου που μπορεί να χρησιμοποιηθεί για να αναγνωρίζει αυτόματα ένα άτομο ή να επαληθεύει την ταυτότητα του. Τα βιομετρικά χαρακτηριστικά και γνώσματα χωρίζονται σε συμπεριφορικά και φυσιολογικά. Συμπεριφορικό είναι το βιομετρικό στοιχείο που είναι σχετικό με τη συμπεριφορά ενός ατόμου και περιλαμβάνει [1]:

- Φωνή
- Υπογραφή
- Τρόπο δακτυλογράφησης
- Στάση σώματος και βάδισμα

Τα φυσιολογικά βιομετρικά συστήματα χρησιμοποιούν φυσιολογικά βιομετρικά στοιχεία. Φυσιολογικό είναι το βιομετρικό στοιχείο που είναι σχετικό με τη μορφή του σώματος και περιλαμβάνει:

- Σχήμα και δομή δακτυλικού αποτυπώματος
- Σχήμα και δομή παλάμης - χεριού

- Σχήμα και χαρακτηριστικά προσώπου
- Δομή της ίριδας του ματιού
- Μορφή του αμφιβληστροειδή χιτώνα του ματιού

Ένα σύστημα αναγνώρισης προσώπου επιτρέπει την αναγνώριση ατόμου με την εμφάνισή του σε μια κάμερα ασφαλείας. Οι άνθρωποι αναγνωρίζουν συχνά ο ένας τον άλλον από το πρόσωπο. Η αυτόματη αναγνώριση προσώπου βασίζεται στην ίδια ιδέα. Η αναγνώριση προσώπου είναι η πιο επιτυχημένη διαδικασία για την επιτήρηση των ανθρώπων. Η τεχνολογία των συστημάτων αναγνώρισης προσώπου χρησιμοποιείται για την βελτίωση της ασφάλειας και είναι ένας από τους σημαντικούς τομείς στη βιομηχανία των βιομετρικών συστημάτων.

Το ενδιαφέρον για τα συστήματα αναγνώρισης προσώπου ενισχύεται από τη διαθεσιμότητα και το χαμηλό κόστος των συστημάτων παρακολούθησης και τον ολοένα αυξανόμενο αριθμό των καμερών που εγκαθίστανται στους χώρους εργασίας και στους δημόσιους χώρους. Αν και τα συστήματα αναγνώρισης προσώπου βρίσκονται ακόμα σε φάση έρευνας και ανάπτυξης, πολλά συστήματα είναι εμπορικά διαθέσιμα και αρκετοί ερευνητικοί οργανισμοί εργάζονται για την ανάπτυξη πιο αξιόπιστων συστημάτων.

Υπάρχουν πολλοί λόγοι για την επιλογή της αναγνώρισης προσώπου ως μέθοδο αναγνώρισης και ταυτοποίησης ατόμων. Οι πιο σημαντικά αναφέρονται παρακάτω:

1. Δεν απαιτείται φυσική αλληλεπίδραση εκ μέρους του χρήστη - ελεγχόμενου
2. Τα ποσοστά επαλήθευσης είναι υψηλά και η αναγνώριση ακριβής
3. Δεν χρειάζεται κάποιος ειδικός για να ερμηνεύσει το αποτέλεσμα της σύγκρισης
4. Δεν απαιτείται επέκταση ή αλλαγή τυχούσας υπάρχουσας υλικοτεχνικής υποδομής
5. Είναι τα ελάχιστα βιομετρικά συστήματα που επιτρέπουν την παθητική αναγνώριση ατόμου μέσα σε πλήθος (π.χ.: τον εντοπισμό τρομοκράτη σε ένα αεροδρόμιο)

Τα βασικά βήματα επεξεργασίας ενός συστήματος αναγνώρισης προσώπου συνοψίζονται στα ακόλουθα:

1. Αναγνώριση ενός δείγματος προσώπου από ένα σύνολο
2. Χρήση της Ανάλυσης Κύριων Συνιστωσών - PCA (χρησιμοποιώντας τη μέθοδο των eigenfaces)

3. Χρήση μια απλής προσέγγισης για την αναγνώριση προσώπου και σύγκρισή της με τη μέθοδο των eigenfaces
4. Ανάδειξη και αιτιολογία της καλύτερης προσέγγισης

Το πρόσωπο και η αναγνώριση του

Στην επιστήμη της ψυχολογίας ένα πρόσωπο είναι μια πλούσια πηγή πληροφοριών για την ανθρώπινη συμπεριφορά. Οι κινήσεις προσώπου δείχνουν συγκίνηση, ρυθμίζουν την κοινωνική συμπεριφορά, αποκαλύπτουν τη χρήση του εγκεφάλου και σηματοδοτούν την ανάπτυξη αισθημάτων στα βρέφη. Γεωμετρικά, το πρόσωπο είναι τρισδιάστατη μεταβλητή επιφάνεια με μεγάλο πλήθος μη-πολυγωνικών περιοχών, ενώ τα χαρακτηριστικά του προσώπου σχετίζονται με την εν λόγω γεωμετρία του προσώπου. Υπολογιστικά, τα δεδομένα εικόνας μπορούν να χαρακτηριστούν ως διανύσματα. Για παράδειγμα, μια $p \times q$ δισδιάστατη εικόνα αντιστοιχίζεται σε ένα διάνυσμα $x \in R^{p \times q}$. Έστω $X = (x_1; x_2; \dots; x_i; \dots; x_N)$ ένας πίνακας δεδομένων $n \times N$, όπου x_i είναι ένα διάνυσμα προσώπου διάστασης n που έχει εξαχθεί από μια εικόνα προσώπου $p \times q$, όπου $n = p \times q$ ο συνολικός αριθμός pixel στο πρόσωπο. Η σημασία του ανθρώπινου προσώπου δεν είναι μόνο ανατομική και λειτουργική, αλλά έχει επιπλέον ένα γενικότερο ατομικό και διαπροσωπικό ρόλο καθώς παρουσιάζει ένα έντονο σημασιολογικό περιεχόμενο. Αποτελεί για κάθε άνθρωπο ένα εντελώς ιδιαίτερο γνώρισμα, του οποίου η σπουδαιότητα έχει κοινωνικές, ψυχολογικές, καλλιτεχνικές και φιλοσοφικές προεκτάσεις [2].

Η αναγνώριση προσώπου αποτελεί μια πρόταση ταυτοποίησης ανθρώπων χωρίς τη χρήση εγγράφων, ούτε συστημάτων επαφής ή αλληλεπίδρασης όπως είναι η αναγνώριση δακτυλικών αποτυπωμάτων, της ίριδας του ματιού, κ.ο.κ. Ο όρος αναγνώριση προσώπων αναφέρεται σε δυο διαφορετικά σενάρια εφαρμογής [3],[4]:

- την ταυτοποίηση (identification) και
- την πιστοποίηση (verification ή authentication)

Και στις δυο περιπτώσεις, πριν από την αναγνώριση λαμβάνει χώρα το στάδιο της καταγραφής (enrollment) των εικόνων των οποίων η ταυτότητα ή/και η έκφραση είναι γνωστή. Κατά την καταγραφή υπολογίζονται τα χαρακτηριστικά (features) των προσώπων και αποθηκεύονται σε μια βάση δεδομένων που αποτελεί τη συλλογή (gallery) του συστήματος. Μετά την καταγραφή το σύστημα μπορεί να δεχτεί νέα άγνωστα πρόσωπα δοκιμής (probe faces) και να αναγνωρίσει την ταυτότητα ή την έκφραση τους.

Στην περίπτωση της ταυτοποίησης το σύστημα καλείται να βρει την ταυτότητα του ατόμου της συλλογής στο οποίο ανήκει η εικόνα δοκιμής.

Το σύστημα απαντά στο ερώτημα «ποιος είναι» ο εικονιζόμενος. Εδώ υπάρχουν δύο ακόμα υποπεριπτώσεις, η ταυτοποίηση ανοιχτού και κλειστού συνόλου. Με τον όρο κλειστό σύνολο εννοούμε ότι η εικόνα δοκιμής ανήκει σε κάποιο από τα άτομα της συλλογής, ενώ με τον όρο ανοιχτό εννοούμε ότι η εικόνα δοκιμής μπορεί να ανήκει και σε άτομο που δεν έχει προηγουμένως καταγραφεί στη συλλογή. Στην πρώτη περίπτωση, το σύστημα συγκρίνει την εικόνα δοκιμής μία προς μία με όλες τις εικόνες συλλογής για την εκτίμηση ενός μέτρου ομοιότητας (similarity score), σύμφωνα με το οποίο αποφασίζει σε ποιο άτομο ανήκει η εικόνα δοκιμής. Στη δεύτερη περίπτωση το καλύτερο μέτρο ομοιότητας συγκρίνεται σε δεύτερη φάση με ένα κατώφλι (threshold) για να ελεγχθεί μήπως η εικόνα δοκιμής δεν ανήκει σε κανένα από τα άτομα της συλλογής. Σε κάθε δοκιμή λοιπόν κατασκευάζεται ένας μονοδιάστατος πίνακας από τα μέτρα ομοιότητας της εικόνας δοκιμής με κάθε εικόνα συλλογής, στη συνέχεια τα άτομα της συλλογής κατατάσσονται με φθίνουσα σειρά των μέτρων ομοιότητας και τέλος επιστρέφεται η ταυτότητα του ατόμου με το καλύτερο μέτρο.

Στην *πιστοποίηση*, το σύστημα αναγνώρισης προσώπων καλείται να εξακριβώσει αν η ταυτότητα που επικαλείται κάποιος χρήστης είναι αληθής ή όχι, καλείται δηλαδή να απαντήσει στο ερώτημα «είναι πράγματι αυτός που ισχυρίζεται ότι είναι» κι έτσι να διακρίνει αν πρόκειται για εξουσιοδοτημένο χρήστη (legitimate user) ή για πλαστοπροσωπία (impostor). Για να γίνει αυτό, η εικόνα δοκιμής συγκρίνεται με τις εικόνες συλλογής του ατόμου του οποίου η ταυτότητα επικαλείται και στη συνέχεια το μέτρο ομοιότητας που προκύπτει συγκρίνεται με ένα κατώφλι για την αποδοχή ή όχι του αιτήματος. Ενώ λοιπόν η ταυτοποίηση είναι μια διαδικασία σύγκρισης ένα-προς-πολλά (1:N), η πιστοποίηση είναι μια διαδικασία σύγκρισης ένα προς ένα (1:1) κι έχει δυαδικό αποτέλεσμα: Ναι, Όχι.

Η απόδοση ενός συστήματος πιστοποίησης εξαρτάται από την επιλογή του κατωφλίου, γιατί αυτό καθορίζει τη συναλλαγή (trade-off) ανάμεσα στην ασφάλεια και την άνεση του χρήστη. Όσο υψηλότερο είναι το κατώφλι τόσο πιο αυστηρό γίνεται το σύστημα και η ασφάλεια αυξάνεται. Ταυτόχρονα όμως αυξάνεται και η πιθανότητα να απορριφθούν εξουσιοδοτημένοι χρήστες με αποτέλεσμα το σύστημα να γίνεται λιγότερο φιλικό προς το χρήστη. Το αντίστροφο συμβαίνει όταν μειώνεται το κατώφλι. Η συμπεριφορά του συστήματος, όπως και σε άλλες περιπτώσεις στη θεωρία ανίχνευσης σημάτων, παριστάνεται με ειδικές καμπύλες απόδοσης με τη γενική ονομασία λειτουργικές χαρακτηριστικές του δέκτη (Receiver Operating Characteristic - ROC), οι οποίες είναι γραφικές απεικονίσεις που δείχνουν την απόδοση ενός δυαδικού ταξινομητή καθώς μεταβάλλεται το κατώφλι αναγνώρισης. Το ROC αποτελεί αναπαράσταση του ρυθμού των ορθά αναγνωρισμένων

(TPR – true positive rate), για διαφορετικές τιμές του κατωφλίου. Ο ρυθμός ορθά αναγνωρισμένων (TPR) είναι η ευαισθησία, ενώ ο ρυθμός εσφαλμένα αναγνωρισμένων (FPR) είναι το συμπλήρωμα ως προς 1 της ιδιαιτερότητας (specificity)[5],[6].

Οι δυνατότητες της γεωμετρίας του προσώπου για την αναγνώριση προσώπων και εκφράσεων ουσιαστικά άρχισαν να ερευνώνται κατά τα δέκα τελευταία χρόνια. Το υψηλό κόστος των αισθητήρων σε συνδυασμό με τη χαμηλή ποιότητα των δεδομένων που παρήγαγαν και το μεγάλο χρόνο απόκρισης οδήγησαν τη πλειοψηφία των ερευνητών στη χρήση έγχρωμων εικόνων ή εικόνων γκριζών τόνων και την ανάπτυξη φωτομετρικών (photometry) μεθόδων αναγνώρισης. Στον τομέα της αναγνώρισης προσώπων προτάθηκε ένας τεράστιος αριθμός τεχνικών, πολλές από τις οποίες έχουν ήδη χρησιμοποιηθεί για την ανάπτυξη εμπορικών συστημάτων. Με την πρόοδο των αισθητήρων βάθους τα τελευταία χρόνια και τη διαθεσιμότητα δεδομένων στην επιστημονική κοινότητα για την διεξαγωγή πειραμάτων, η έρευνα γύρω από την 2Δ ή 3Δ αναγνώριση προσώπων και εκφράσεων δέχτηκε σημαντική ώθηση. Όλα τα συστήματα αναγνώρισης προτύπων ακολουθούν μια τυπική αρχιτεκτονική που περιλαμβάνει τέσσερα βασικά στάδια επεξεργασίας [5],[6]:

- λήψη μετρήσεων για κάποιες από τις ιδιότητες του αντικειμένου
- προ-επεξεργασία των μετρήσεων για τη μείωση του θορύβου ή την κανονικοποίηση των μετρήσεων
- εξαγωγή χαρακτηριστικών, με τα οποία γίνεται η διάκριση των προτύπων και
- ταξινόμηση, που περιλαμβάνει τη σύγκριση των χαρακτηριστικών του αντικειμένου είτε με κάποια κατώφλια είτε με αντίστοιχα προ-αποθηκευμένα χαρακτηριστικά γνωστών αντικειμένων, ώστε να αποφασιστεί σε ποια κλάση κατατάσσεται το αντικείμενο

Τα συστήματα 3Δ αναγνώρισης προσώπων και εκφράσεων ακολουθούν την ίδια τυπική αρχιτεκτονική με την επιπλέον ιδιαιτερότητα ότι τα δυο πρώτα στάδια, της καταγραφής του βάθους και της προ-επεξεργασίας γίνονται μαζί, ενώ τα στάδια της εξαγωγής χαρακτηριστικών και της ταξινόμησης χρησιμοποιούν παρόμοιες τεχνικές. Η καταγραφή της επιφάνειας του προσώπου γίνεται με μεθόδους που μπορούν να χωριστούν σε δυο μεγάλες κατηγορίες, τις *ενεργητικές* και τις *παθητικές*.

Οι *ενεργητικές μέθοδοι* στηρίζονται στο φωτισμό της επιφάνειας με ένα γνωστό μοτίβο φωτός (light pattern) και καλούνται γενικά τεχνικές δομημένου φωτισμού (structured light techniques). Το μοτίβο φωτός προβάλλεται μέσω ενός συστήματος προβολής πάνω στο αντικείμενο και

μια κάμερα συλλαμβάνει την ανάκλασή του πάνω στην επιφάνεια του αντικειμένου. Στη συνέχεια, για κάθε σημείο της επιφάνειας υπολογίζεται η αντίστοιχη θέση του πάνω στην εικόνα και μέσω αυτής της αντιστοιχίας και απλών τριγωνομετρικών υπολογισμών (αρχή του τριγωνισμού) εκτιμάται η απόσταση του σημείου από το επίπεδο της κάμερας. Το μοτίβο φωτός μπορεί να είναι ένα απλό σημείο ή μια γραμμή. Επειδή όμως η χρήση τόσο απλών προτύπων καταλήγει σε χρονοβόρες διαδικασίες αντιστοίχισης, τελευταία χρησιμοποιούνται πιο πολύπλοκα πρότυπα που επιτρέπουν παράλληλη επεξεργασία, όπως πλέγματα ή χρωματικά κωδικοποιημένες λωρίδες. Επίσης για να μην ενοχλείται ο χρήστης από την προβολή φωτός, γίνονται προσπάθειες για την αντικατάσταση των ορατών προτύπων από υπέρυθρες ακτίνες.

Οι *παθητικοί μέθοδοι* χρησιμοποιούν στερεοσκοπικές μεθόδους για τον υπολογισμό του βάθους, στηρίζονται δηλαδή στην επεξεργασία δύο ή περισσότερων εικόνων του αντικειμένου από διαφορετικές όψεις. Το κύριο πρόβλημα σ' αυτές τις μεθόδους είναι το πρόβλημα της αντιστοίχισης, η εύρεση δηλαδή των εικονοστοιχείων (pixels) σε όλες τις εικόνες που αντιστοιχούν στο ίδιο 3D σημείο της επιφάνειας. Για την επίλυση του, συνήθως χρησιμοποιείται η μέθοδος της οπτικής ροής (optical flow) σε συνδυασμό με την εφαρμογή επιπολικής γεωμετρίας που περιορίζει το εύρος της αναζήτησης. Ένα μειονέκτημα της συγκεκριμένης μεθόδου είναι ότι δεν μπορεί να εφαρμοστεί για αντικείμενα που δε διαθέτουν υφή (texture). Αυτό βέβαια μπορεί να αντιμετωπιστεί εφαρμόζοντας ενεργητικές τεχνικές προβολής κατάλληλου πρότυπου υφής. Σε γενικές γραμμές, το ανθρώπινο πρόσωπο περιέχει αρκετή πληροφορία για την επίλυση του προβλήματος της αντιστοίχισης χωρίς να χρειάζεται η προβολή προτύπου υφής.

Η επιλογή της μεθόδου εξαρτάται κυρίως από την εφαρμογή για την οποία προορίζεται. Γενικά, είναι θέμα συμβιβασμού ανάμεσα στην ταχύτητα, το κόστος και την ποιότητα των παραγόμενων δεδομένων. Οι παθητικοί μέθοδοι είναι ταχύτερες και φθηνότερες, αλλά παράγουν δεδομένα με περισσότερο θόρυβο. Από την άλλη πλευρά, οι ενεργητικοί μέθοδοι είναι πιο ακριβείς, αλλά είναι ακριβότεροι και απαιτούν την ακινητοποίηση του χρήστη για το διάστημα της καταγραφής.

Οι δυο τεχνικές στηρίζονται σε φωτομετρικές μεθόδους (2D) για την εκτίμηση του βάθους. Αυτό σημαίνει ότι ενώ το βάθος εν γένει δεν εξαρτάται από τις συνθήκες φωτισμού, ωστόσο η καταγραφή (sensing) του μπορεί να επηρεάζεται από αυτές.

Από πλευράς υλοποίησης, οι δύο κύριες προσεγγίσεις για την αναγνώριση προσώπου είναι:

- η *γεωμετρική προσέγγιση* που χρησιμοποιεί τη χωροταξική διάταξη των χαρακτηριστικών του προσώπου. Αυτό σημαίνει ότι

τα βασικά γεωμετρικά χαρακτηριστικά του προσώπου όπως τα μάτια, η μύτη και το στόμα αναγνωρίζονται πρώτα και στη συνέχεια τα πρόσωπα ταξινομούνται με βάση διαφορών γεωμετρικών αποστάσεων και γωνιών μεταξύ των χαρακτηριστικών

- η *εικονογραφική προσέγγιση* που χρησιμοποιεί πρότυπα των χαρακτηριστικά του προσώπου. Η μέθοδος αυτή χρησιμοποιεί τα πρότυπα των κύριων χαρακτηριστικών του προσώπου και ολόκληρο το πρόσωπο για να πραγματοποιήσει την αναγνώριση του προσώπου

Εκτός από αυτές τις δύο βασικές τεχνικές υπάρχουν και άλλες πιο πρόσφατες βάσει προτύπου προσεγγίσεις, που σχηματίζουν πρότυπα από την κλίση της εικόνας και η ανάλυση κύριων συνιστωσών.

Λαμβάνοντας υπόψη ότι μια εικόνα προέρχεται από ένα πλαίσιο από ακολουθία βίντεο, η αυτόματη αναγνώριση προσώπου είναι μια ιδιαίτερα πολύπλοκη διαδικασία που περιλαμβάνει τον εντοπισμό και τη αναγνώριση των προσώπων σε ένα μη ελεγχόμενο χώρο χρησιμοποιώντας εξομάλυνση και μετά αναγνώριση. Το ανθρώπινο πρόσωπο είναι ένα πολύ δύσκολο πρότυπο για την ανίχνευση και την αναγνώριση του, γιατί ενώ η ανατομία του είναι άκαμπτη (όλα τα πρόσωπα έχουν την ίδια δομή), την ίδια στιγμή υπάρχουν πολλοί περιβαλλοντικοί και προσωπικοί παράγοντες που επηρεάζουν την εμφάνιση του προσώπου. Το κύριο πρόβλημα της αναγνώρισης του προσώπου είναι η μεγάλη μεταβλητότητα των καταγεγραμμένων εικόνων, οι συνθήκες φωτισμού, οι εκφράσεις του προσώπου, η χρήση των καλλυντικών, διαφορετικό στυλ στα μαλλιά, παρουσία γυαλιών, γενειάδα, κλπ. Εικόνες από το ίδιο άτομο που ελήφθησαν σε διαφορετικές χρονικές στιγμές, μερικές φορές μπορεί να εμφανίζουν περισσότερη μεταβλητότητα λόγω των προαναφερθέντων παραγόντων (προσωπικότητα), από τις εικόνες διαφορετικών ατόμων λόγω φύλου, φυλής, ηλικίας και ατομικών χαρακτηριστικών. Ένας τρόπος αντιμετώπισης των προσωπικών χαρακτηριστικών είναι η προσθήκη στο σύστημα εκπαίδευσης επιπρόσθετων εικόνων που να καλύπτουν αυτή τη διακύμανση.

Τα υπολογιστικά μοντέλα αναγνώρισης προσώπου παρουσιάζουν ιδιαίτερο ενδιαφέρον, γιατί μπορούν να συμβάλουν σε θεωρητικές και πρακτικές διαπιστώσεις. Συστήματα που αναγνωρίζουν πρόσωπα θα μπορούσαν να χρησιμοποιηθούν σε μια ποικιλία προβλημάτων, όπως η αναγνώριση εγκληματιών, συστημάτων ασφαλείας, εικόνες και επεξεργασίας φιλμ, και αλληλεπίδραση μεταξύ ανθρώπου και υπολογιστή. Για παράδειγμα, η δυνατότητα διαμόρφωσης ενός συγκεκριμένου μοντέλου-προσώπου μέσω της εξαγωγής του από ένα μεγάλο πλήθος αποθηκευμένων μοντέλων προσώπου θα καθιστούσε δυνατή την βελτίωση αναγνώρισης εγκληματιών. Ακόμη και η

δυνατότητα να ανιχνεύει μόνο τα πρόσωπα, χωρίς να είναι ικανό να τα αναγνωρίσει, μπορεί να είναι σημαντική. Η ανίχνευση προσώπων στις εικόνες είναι ένα σημαντικό πρόβλημα.

Κεφάλαιο 2. Βασικές Έννοιες της Αναγνώρισης Προτύπων

Εισαγωγή

Η ανάγκη για βελτιωμένα συστήματα πληροφοριών έχει ενταθεί, από τη στιγμή που οι πληροφορίες είναι ένα σημαντικό στοιχείο στη λήψη αποφάσεων. Ένα από τα σημαντικότερα προβλήματα στον σχεδιασμό των σύγχρονων πληροφοριακών συστημάτων είναι η αυτόματη αναγνώριση προτύπων (patterns). Η αναγνώριση θεωρείται ως ένα από τα βασικά χαρακτηριστικά της ανθρώπινης φύσης, καθώς και άλλων ζωντανών οργανισμών. Ένα πρότυπο είναι η περιγραφή ενός αντικειμένου. Ο άνθρωπος θεωρείται ένα πολύ εξελιγμένο πληροφοριακό σύστημα, λόγω της έμφυτης ικανότητας του στην αναγνώριση προτύπων. Τα πρότυπα μπορούν να χωριστούν σε δύο μεγάλες κατηγορίες [7]:

- **Σαφή πρότυπα (*Recognition of concrete items*)**. Περιλαμβάνει οπτική και ακουστική αναγνώριση προτύπων. Αυτή η διαδικασία αναγνώρισης περιλαμβάνει την αναγνώριση και ταξινόμηση των προτύπων στο χώρο (spatial) και στο χρόνο (temporal). Μερικά παραδείγματα χωρικών προτύπων είναι οι τυπωμένοι χαρακτήρες, τα δακτυλικά αποτυπώματα, τα φυσικά αντικείμενα, οι χάρτες καιρού και οι εικόνες. Τα χρονικά πρότυπα περιλαμβάνουν ακουστικές κυματομορφές, χρονοσειρές, ηλεκτροεγκεφαλογραφήματα και σήματα ραντάρ.
- **Αφηρημένα πρότυπα (*Recognition of abstract items*)**. Από την άλλη πλευρά, ένα παλιό επιχείρημα ή μια λύση σε ένα πρόβλημα μπορεί να αναγνωρίζεται. Αυτή η διαδικασία περιλαμβάνει την αναγνώριση αφηρημένων προτύπων και μπορεί να ονομαστεί εννοιολογική αναγνώριση. Η αναγνώριση των συγκεκριμένων προτύπων από τον άνθρωπο μπορεί να θεωρηθεί ως ένα ψυχοφυσιολογικό πρόβλημα, το οποίο περιλαμβάνει τη σχέση μεταξύ ενός ατόμου και ενός φυσικού ερεθίσματος. Η ανθρώπινη αναγνώριση στην πραγματικότητα είναι ο υπολογισμός των σχετικών πιθανοτήτων ότι τα δεδομένα εισόδου μπορούν να συσχετιστούν με ένα σύνολο από στατιστικά γνωστών ατόμων που εξαρτώνται από την εμπειρία του παρελθόντος μας και τα οποία θα σχηματίσουν τις ενδείξεις και τις πληροφορίες για την αναγνώριση. Οπότε, το πρόβλημα της αναγνώρισης προτύπων είναι η διαδικασία διαχώρισης των δεδομένων εισόδου μεταξύ των πληθυσμών με χρήση της αναζήτησης χαρακτηριστικών ή

αμετάβλητων χαρακτηριστικών μεταξύ των μελών ενός πληθυσμού.

Κλάσεις προτύπων και πρότυπα

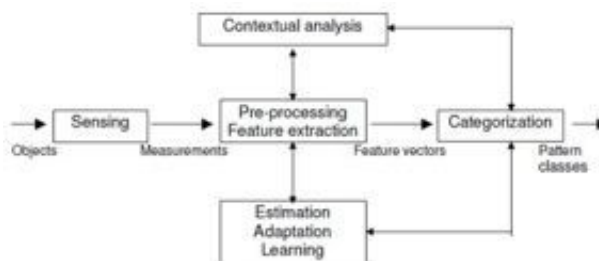
Αναγνώριση Προτύπων (Pattern Recognition) είναι ο επιστημονικός κλάδος που ασχολείται με την εύρεση μεθόδων για την ταξινόμηση (classification) των προτύπων σε κατηγορίες. Τις κατηγορίες αυτές θα τις ονομάζουμε κλάσεις ή τάξεις (classes) [8].

- Μια **κλάση προτύπων** είναι μια κατηγορία που καθορίζεται από όμοια χαρακτηριστικά ή ιδιότητες. Τα χαρακτηριστικά μιας κλάσης προτύπων είναι οι κοινές/όμοιες ιδιότητες από όλα τα πρότυπα που ανήκουν σε αυτή την κλάση. Τέτοια χαρακτηριστικά συχνά αναφέρονται ως *intraset features*. Τα χαρακτηριστικά γνωρίσματα που αντιπροσωπεύουν τις διαφορές ανάμεσα στις κλάσεις προτύπων αναφέρονται ως *interest features*.
- Ένα **πρότυπο** είναι η περιγραφή οποιουδήποτε μέλους της κατηγορίας που αντιπροσωπεύει η κλάση προτύπων. Για λόγους ευκολίας, τα πρότυπα συνήθως αναπαρίστανται από ένα διάνυσμα.

Εκπαίδευση και Εκμάθηση

Η εκμάθηση και η εκπαίδευση λαμβάνουν χώρα μόνο κατά τη διάρκεια της φάσης του σχεδιασμού (ή ενημέρωσης) ενός συστήματος αναγνώρισης προτύπων. Μόλις αποκτηθούν αποδεκτά αποτελέσματα από την εκπαίδευση των προτύπων, το σύστημα είναι έτοιμο να εκτελέσει αναγνώριση σε δείγματα που ελήφθησαν από το περιβάλλον στο οποίο αναμένεται να λειτουργήσει. Η ποιότητα στην απόδοση της αναγνώρισης καθορίζεται σε μεγάλο βαθμό από το πόσο κοντά τα πρότυπα εκπαίδευσης βρίσκονται σε σχέση με τα αρχικά στοιχεία με τα οποία το σύστημα βρίσκεται σε κανονική λειτουργία.

Περίγραμμα ενός συστήματος αναγνώρισης προτύπων



Σχήμα 1. Λειτουργικό τμήμα ενός προσαρμοζόμενου συστήματος αναγνώρισης προτύπων

Στο Σχήμα 1 παρουσιάζεται ένα διάγραμμα λειτουργικού τμήματος ενός προσαρμοζόμενου συστήματος αναγνώρισης προτύπου. Παρά το γεγονός ότι η διάκριση μεταξύ της βέλτιστης απόφασης και της εξαγωγής προεπεξεργασμένων χαρακτηριστικών δεν είναι απαραίτητη, η έννοια της λειτουργικής ανάλυσης βοηθάει στην κατανόηση της αναγνώρισης προτύπων.

Η σωστή αναγνώριση εξαρτάται από το ποσό των διακριτών πληροφοριών που περιέχονται στις μετρήσεις και την αποτελεσματική αξιοποίησή τους. Σε ορισμένες εφαρμογές, οι σχετικές πληροφορίες (contextual information) είναι απαραίτητες για την επίτευξη μιας ακριβούς αναγνώρισης. Για παράδειγμα, στην αναγνώριση χειρόγραφων χαρακτήρων και στην ταξινόμηση των δακτυλικών αποτυπωμάτων, οι σχετικές πληροφορίες (contextual information) είναι απαραίτητες. Όταν θέλουμε να σχεδιάσουμε ένα σύστημα αναγνώρισης προτύπου το οποίο είναι ανθεκτικό σε στρεβλώσεις, ευέλικτο σε μεγάλες αποκλίσεις των προτύπων, και ικανό για διαμόρφωση (self-adjustment), βρισκόμαστε αντιμέτωποι με το πρόβλημα της προσαρμοστικότητας.

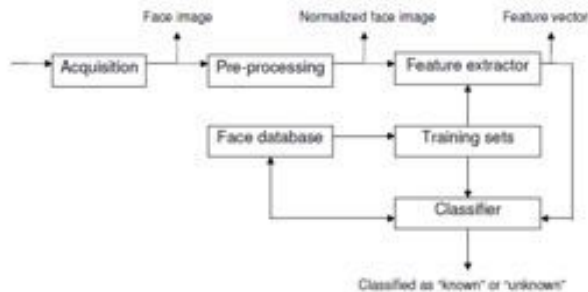
Αναγνώριση προσώπου

Η ανίχνευση προσώπου θεωρείται ως ένα πρόβλημα ταξινόμησης (classification) δύο κλάσεων, όπου μία περιοχή μιας εικόνας ταξινομείται ως «πρόσωπο» ή «μη-πρόσωπο». Με την ανίχνευση προσώπου εντοπίζουμε ανθρώπινα πρόσωπα σε μια εικόνα ανεξαρτήτως [4]:

- της θέσης τους
- του μεγέθους τους (κλίμακας)
- του προσανατολισμού τους (στροφή στο επίπεδο της εικόνας)
- της στάσης/πόζας τους (στροφή εκτός επιπέδου εικόνας)
- της έκφρασής τους (ύπαρξη μορφασμών)
- της ύπαρξης άλλων δομικών στοιχείων (μούσια, μουστάκια, γυαλιά)
- του περιεχομένου της εικόνας (ύπαρξη εμποδίων ή άλλων προσώπων)
- και του φωτισμού και των συνθηκών αποτύπωσης της εικόνας (ευαισθησία κάμερας, ανάλυση)

Όλα τα παραπάνω αποτελούν τις κύριες αιτίες της μεγάλης δυσκολίας που εμφανίζει το πρόβλημα της ανίχνευσης ενός προσώπου. Οι παραλλαγές αυτές στην εμφάνιση του προσώπου, στη στάση, στην έκφραση και τον φωτισμό κάνουν πολύπλοκο το πολύπτυχο (manifold) του χώρου των προσώπων και δυσδιάκριτα τα όρια προσώπων / μη-προσώπων. Για να

είναι εφικτή η διαχείριση αυτής της πολύπλοκης κατάστασης απαιτείται ένας μη γραμμικός ταξινομητής, καθώς επίσης και ένα μεγάλο σύνολο δειγμάτων εκπαίδευσης, τα οποία θα εμπεριέχουν αυτήν την ποικιλοπλοκότητα στην εμφάνιση των προσώπων. Ένα επίσης σημαντικό θέμα, είναι η ταχύτητα για την απόδοση του συστήματος σε πραγματικό χρόνο. Έχει γίνει σημαντική ερευνητική προσπάθεια για την δημιουργία σύνθετων και γρήγορων ταξινομητών και από τη δεκαετία του '90 έχει γίνει σημαντική πρόοδος στο θέμα αυτό.



Σχήμα 2. Διάγραμμα ροής ενός συστήματος αναγνώρισης προσώπου

Υπάρχουν έξι κύρια στάδια σε ένα σύστημα αναγνώρισης προσώπου, όπως φαίνεται στο Σχήμα 2 [9],[10]:

- 1) **Είσοδος εικόνας με πρόσωπο (the acquisition module).** Αυτό είναι το σημείο εισόδου της διαδικασίας της αναγνώρισης προσώπου. Είναι το στάδιο όπου ο χρήστης παρουσιάζει μια εικόνα με πρόσωπο στο σύστημα. Η εικόνα προσώπου μπορεί να είναι ένα αρχείο εικόνας που βρίσκεται σε σκληρό δίσκο, από κάμερα ή από σαρωτή.
- 2) **Προ-επεξεργασία (the pre-processing module).** Η προ-επεξεργασία με κανονικοποίηση της έντασης των εικονοστοιχείων (pixels) βοηθάει στην διόρθωση των αποκλίσεων των παραμέτρων αποτύπωσης των εικόνων μέσω των καμερών, όπως επίσης και των παραλλαγών στις συνθήκες φωτισμού. Μερικά βήματα που μπορούν να εφαρμοστούν στην διαδικασία της προ-επεξεργασίας είναι:
 - **Τροποποίηση μεγέθους εικόνας (image size normalization):** Τροποποίηση όλων των προτύπων προσώπων σε ένα συγκεκριμένο μέγεθος (π.χ. σε 128×128 pixels).
 - **Εξισορρόπηση Ιστογράμματος (histogram equalization):** αντισταθμίζει τις επιδράσεις φωτισμού εξαιτίας διαφορετικών συνθηκών φωτισμού και διαφορετικών καμπύλων απόκρισης της κάθε κάμερας λήψης.
 - **Φίλτρο μεσαίας τιμής (median filter):** Για την εξομάλυνση (smoothing) των ακμών και μείωση του θορύβου μιας εικόνας

που έχει αποκτηθεί με χρήση κάμερας, μπορεί να χρησιμοποιηθεί το φίλτρο ενδιάμεσης τιμής.

- **Υψηπερατό φίλτρο (*high-pass filtering*):** Η ανάδειξη και ο τονισμός των λεπτομερειών μιας εικόνας μπορεί να γίνει με τη χρήση υψηπερατών φίλτρων. Αυτό μπορεί να επιτευχθεί εύκολα αν από την αρχική εικόνα αφαιρέσουμε μια εξομαλυμένη έκδοση της (για παράδειγμα μετά από την εφαρμογή μιας μάσκας μέσης τιμής ή καλύτερα ενός φίλτρου Gauss). Το υψηπερατό φιλτράρισμα (*high pass filtering*) στο πεδίο του χώρου μπορεί να γίνει με την χρησιμοποίηση μασκών που αποτελούνται από ένα μείγμα θετικών και αρνητικών συντελεστών. Όπως είναι φανερό, τα υψηπερατά φίλτρα τονίζουν τις λεπτομέρειες της εικόνας, όπως είναι οι ακμές. Ένα υψηπερατό φίλτρο πρέπει να έχει παρόμοια απόκριση ανεξάρτητα από την διεύθυνση μεταβολής της φωτεινότητας. Για το σκοπό αυτό, οι συντελεστές κατανέμονται συμμετρικά ως προς το κέντρο της μάσκας. Επίσης, πρέπει να έχει θετικούς συντελεστές στο κέντρο της μάσκας και αρνητικούς στην περιφέρεια.
- **Αφαίρεση φόντου (*background removal*):** Περιορισμός των μη επιθυμητών δομών του υποβάθρου (θόρυβος) κοντά στα όρια του πλαισίου ενός προτύπου προσώπου.
- **Επιδιόρθωση διαβαθμισμένου φωτισμού (*illumination normalization*):** Ανεύρεση μιας γραμμικής συνάρτησης η οποία να ταιριάζει στην διαβάθμιση των τιμών έντασης του παραθύρου του προσώπου και της οποίας οι τιμές αφού αφαιρεθούν, επιδιορθώνουν ακραίες διαβαθμίσεις φωτισμού (σκιές που δημιουργούνται από ακραίες γωνίες φωτισμού).

3) **Εξαγωγή χαρακτηριστικών (*the feature extraction module*).** Μετά την εκτέλεση της διαδικασίας προ-επεξεργασίας (αν είναι απαραίτητη), η κανονικοποιημένη/ομαλοποιημένη (*normalized*) εικόνα προσώπου παρουσιάζεται στο στάδιο της εξαγωγής χαρακτηριστικών για να βρεθούν τα βασικά χαρακτηριστικά που πρόκειται να χρησιμοποιηθούν για την ταξινόμηση. Αυτό το στάδιο είναι υπεύθυνο για τη σύνθεση ενός διανύσματος χαρακτηριστικών γνωρισμάτων που είναι αρκετά σαφές ώστε να εκπροσωπήσει την εικόνα του προσώπου.

4) **Ταξινόμηση.** Με τη βοήθεια ενός ταξινομητή προτύπου, τα εξαγόμενα χαρακτηριστικά της εικόνας προσώπου συγκρίνονται με αυτά που είναι αποθηκευμένα σε μια βιβλιοθήκη με αποθηκευμένα πρόσωπα (ή βάση δεδομένων). Μετά τη σύγκριση, η εικόνα προσώπου θα ταξινομηθεί ως γνωστή (*known*) ή άγνωστη (*unknown*).

- 5) **Σύνολο εκπαίδευσης (training set).** Τα πρότυπα που χρησιμοποιούνται για την εκπαίδευση του ταξινομητή αποτελούν το σύνολο εκπαίδευσης.
- 6) **Βιβλιοθήκη προσώπων ή βάση δεδομένων προσώπων (face library or face database).** Από τη στιγμή που ταξινομούνται οι εικόνες προσώπου ως άγνωστες, μπορούν να προστεθούν σε μια βιβλιοθήκη (ή σε μια βάση δεδομένων) μαζί με το διάνυσμα των χαρακτηριστικών γνωρισμάτων (feature vector) για να χρησιμοποιηθούν σε συγκρίσεις. Το στάδιο ταξινόμησης κάνει άμεση χρήση της βιβλιοθήκης προσώπων.

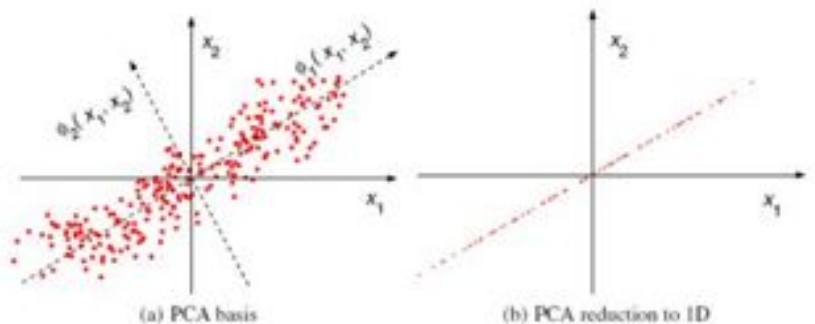
ΑΚΣ-Ανάλυση Κύριων Συνιστωσών (PCA–Principal Component Analysis)

Η ανάλυση των κύριων συνιστωσών (ΑΚΣ) είναι μία στατιστική μέθοδος η οποία μετασχηματίζει γραμμικά ένα σύνολο μεταβλητών σε ένα σύνολο νέων μη συσχετιζόμενων μεταβλητών. Από τις νέες μεταβλητές, που ονομάζονται κύριες συνιστώσες, μόνο ένα μέρος αυτών θα χρησιμοποιηθεί για την εξαγωγή συμπεράσματος. Η μελέτη δύο ή τριών μη συσχετιζόμενων μεταβλητών είναι ευκολότερη από την μελέτη του συνόλου των αρχικών μεταβλητών. Οι Kirby και Sirovich (1987) [11] χρησιμοποίησαν την ΑΚΣ για να αναπαραστήσουν πρόσωπα και οι Turk και Pentland [12] για την αναγνώριση προσώπων. Η ΑΚΣ είναι μια χρήσιμη στατιστική μέθοδος που βρίσκει εφαρμογή σε τομείς όπως η αναγνώριση προσώπου και η συμπίεση της εικόνας και είναι μια τεχνική για την εύρεση προτύπων σε δεδομένα πολλών διαστάσεων. Η ΑΚΣ, που συνήθως αναφέρεται ως η χρήση των *eigenfaces* (βλ. Σχήμα 3), είναι η μέθοδος που χρησιμοποίησαν οι Kirby και Sirovich το 1987. Με την ΑΚΣ, οι προς εκπαίδευση και οι προς εξέταση εικόνες θα πρέπει να έχουν το ίδιο μέγεθος, αλλά και να κανονικοποιηθούν, ώστε τα μάτια και το στόμα να ευθυγραμμιστούν. Στη συνέχεια η μέθοδος ΑΚΣ χρησιμοποιείται για να μειωθεί η διάσταση των δεδομένων με χρήση συμπίεσης δεδομένων. Η μείωση αυτή σε διαστάσεις αφαιρεί πληροφορίες που δεν είναι χρήσιμες και μετασχηματίζει τη δομή του προσώπου σε ένα ορθοκανονικό σύνολο αξόνων στις διευθύνσεις όπου τα δεδομένα παρουσιάζουν μέγιστη συνδιακύμανση (*eigenfaces*). Όλες οι εικόνες προσώπων του συνόλου εκπαίδευσης προβάλλονται πάνω στον υποχώρο προσώπου (*face space*) προκειμένου να βρεθεί ένα σύνολο βαρών που περιγράφει τη συμμετοχή κάθε διανύσματος στον υποχώρο προσώπου. Για να γίνει ταυτοποίηση μιας νέας εικόνας πρέπει να προβληθεί το διάνυσμα που την αναπαριστά στον υποχώρο προσώπου ώστε να υπολογιστεί ένα αντίστοιχο σύνολο βαρών. Συγκρίνοντας τα βάρη της εικόνας με το σύνολο βαρών των εικόνων εκπαίδευσης μπορεί να γίνει η ταυτοποίηση της.

Στο Σχήμα 4 παρουσιάζεται σχηματικό παράδειγμα που επιδεικνύει την εφαρμογή της ΑΚΣ σε διανύσματα δύο διαστάσεων με αποτέλεσμα την προβολή τους σε μία διάσταση με βέλτιστο τρόπο (με κριτήριο τη μέγιστη συνδιασπορά). Η μέθοδος αναγνωρίζει αρχικά τους άξονες που μπορούν να εκφράσουν τη μέγιστη συνδιασπορά και εν συνεχεία προβάλλει τα διανύσματα πάνω στον άξονα μέγιστης συνδιασποράς.



Σχήμα 3. Εικόνες ιδιοδιανυσμάτων (eigenfaces)



Σχήμα 4. Παράδειγμα χρήσης της ΑΚΣ για τη μείωση διάστασης σε 2Δ δεδομένα

Η μέθοδος ΑΚΣ χρησιμοποιεί ιδιοδιανύσματα (eigenvectors) και ιδιοτιμές (eigenvalues) για την αναπαράσταση των εικόνων προσώπου. Τα ιδιοδιανύσματα μπορούν να θεωρηθούν ως ένα σύνολο χαρακτηριστικών τα οποία χαρακτηρίζουν τη διακύμανση μεταξύ των εικόνων προσώπου.

Εκτός της ΑΚΣ υπάρχουν και άλλοι μέθοδοι με μεγαλύτερη αποδοτικότητα:

- **Ανάλυση Ανεξαρτήτων Συνιστωσών – (ICA-Independent Component Analysis):** Σε μια διεργασία, όπως η αναγνώριση προσώπου ή εκφράσεων, στην οποία σημαντική πληροφορία μπορεί να περιέχεται στις υψηλής τάξης σχέσεις μεταξύ των pixels, είναι λογική η προσδοκία ότι καλύτερες εικόνες βάσης μπορούν να προκύψουν από μεθόδους ευαίσθητες σε αυτές της υψηλής τάξης ροπές. Η Ανάλυση ανεξάρτητων συνιστωσών (Independent Component Analysis, ICA), που αποτελεί μια γενίκευση της μεθόδου PCA, είναι μια τέτοια μέθοδος η οποία προσπαθεί να υπολογίσει μια στατιστικά ανεξάρτητη διανυσματική βάση. Για εφαρμογές επεξεργασίας εικόνων του προσώπου έχουν αναπτυχθεί δύο αρχιτεκτονικές βασισμένες στην Ανάλυση Ανεξάρτητων Συνιστωσών. Η πρώτη αρχιτεκτονική (Architecture I) παράγει στατιστικά ανεξάρτητες εικόνες βάσης, ενώ η δεύτερη (Architecture II) παράγει στατιστικά ανεξάρτητους συντελεστές.
- **Γραμμική Διακρίνουσα Ανάλυση (LDA-Linear Discriminant Analysis):** Η γραμμική διακρίνουσα ανάλυση ή η λεγόμενη μέθοδος των Fisherfaces, είναι μία μέθοδος που στοχεύει στην υψηλή διαχωριστικότητα μεταξύ των διαφόρων μοτίβων που πρέπει να ταξινομηθούν. Η γραμμική διακρίνουσα του Fisher ομαδοποιεί τις εικόνες που ανήκουν στην ίδια κλάση ενώ ταυτόχρονα διαχωρίζει αυτές που ανήκουν σε διαφορετικές κλάσεις. Οι εικόνες προβάλλονται από το διδιάστατο χώρο σε ένα χώρο C διαστάσεων, όπου C ο αριθμός των κλάσεων των εικόνων (ή διαφορετικά ο αριθμός των διαφορετικών ατόμων στη βάση των προσώπων).

Αναγνώριση προσώπου με χρήση eigenfaces

Η μέθοδος των eigenfaces έχει βάση από τις πιο θεμελιώδεις ιδέες στα μαθηματικά και επεξεργασία σήματος - τις σειρές Fourier. Μια μικρή περίληψη στην δυνατότητα των σειρών Fourier θα μας αποκομίσει πολλές γνώσεις στο τι ακριβώς η μέθοδος eigenfaces (ή ΑΚΣ) κάνει.

Οι μαθηματικοί του δέκατου όγδοου αιώνα, μεταξύ των οποίων και οι Daniel Bernoulli και Leonard Euler, περιέγραψαν το πρόβλημα της ταλάντωσης τεντωμένης χορδής με μερικές διαφορικές εξισώσεις οι οποίες δεν είχαν λύσεις που να ήταν συνάρτηση «στοιχειωδών συναρτήσεων». Η πρότασή τους στην δυσκολία αυτή ήταν η εισαγωγή σειρών με άπειρους όρους συναρτήσεων ημιτόνου και συνημιτόνου που ικανοποιούσαν τις εξισώσεις. Στις αρχές του δέκατου ένατου αιώνα, ο Joseph Fourier, ενώ μελετούσε το πρόβλημα της ροής θερμότητας,

ανέπτυξε μια περιεκτική θεωρία για παρόμοιες σειρές. Τελικά, οι σειρές αυτές πήραν το όνομά του.

Έστω ότι η $f(x)$ είναι περιοδική με περίοδο 2π , ορίζεται στο διάστημα $c \leq x \leq c + 2\pi$ και ικανοποιεί τις συνθήκες του Dirichlet:

1. Η $f(x)$ είναι ορισμένη στο $c \leq x \leq c + 2\pi$, όπου $c = -\pi$, εκτός ίσως σε πεπερασμένο πλήθος σημείων
2. Η $f(x)$ είναι περιοδική εκτός του $c \leq x \leq c + 2\pi$ με περίοδο $2(c + 2\pi)$
3. Η $f(x)$ και $f'(x)$ είναι τμηματικά συνεχείς στο $c \leq x \leq c + 2\pi$

Τότε η σειρά Fourier που αντιστοιχεί στην $f(x)$ δίνεται από την

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(nx) + b_n \sin(nx)] \quad (1)$$

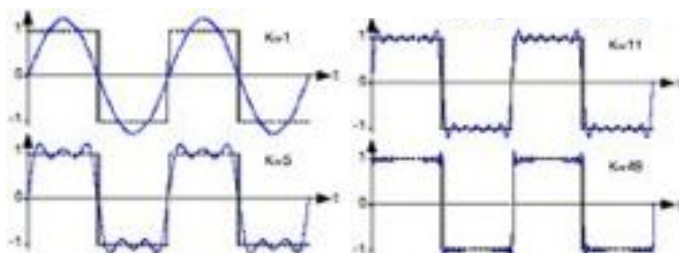
όπου οι συντελεστές Fourier a_0 , a_n και b_n είναι:

$$a_0 = \frac{1}{\pi} \int_c^{c+2\pi} f(x) dx$$

$$a_n = \frac{1}{\pi} \int_c^{c+2\pi} f(x) \cos(nx) dx$$

$$b_n = \frac{1}{\pi} \int_c^{c+2\pi} f(x) \sin(nx) dx$$

Στο Σχήμα 5 παρουσιάζεται ένα τυπικό παράδειγμα που δείχνει την εφαρμογή της (1) ή των σειρών Fourier για την προσέγγιση ενός τετραγωνικού κύματος.



Σχήμα 5. Προσέγγιση του τετραγωνικού κύματος με σειρές Fourier

Ένα τετράγωνο κύμα (που δίνεται σε μαύρο) μπορεί να προσεγγιστεί με χρήση μιας σειράς ημιτόνων και συνημιτόνων (το αποτέλεσμα αυτής της άθροισης εμφανίζεται με μπλε χρώμα). Σε οριακή περίπτωση θα

μπορούσαμε να ανακατασκευάσουμε το τετραγωνικό κύμα με χρήση μόνο απλών ημιτόνων και συνημιτόνων.

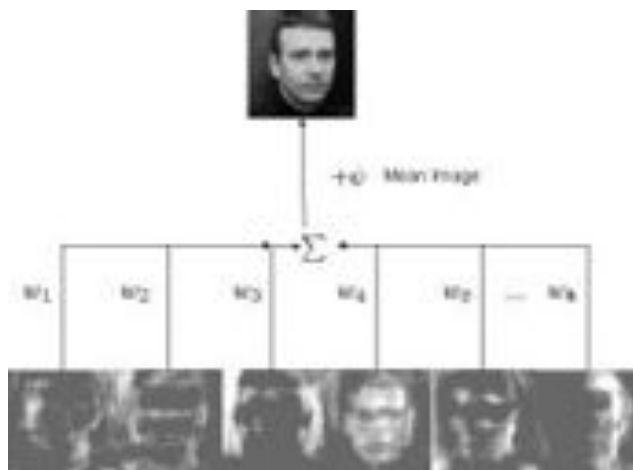
Η ιδέα πίσω από την μέθοδο των eigenfaces είναι παρόμοια. Ο στόχος είναι η εκπροσώπηση ενός προσώπου από ένα γραμμικό συνδυασμό από ένα σύνολο εικόνων (στη σειρά Fourier οι όροι ήταν απλά ημίτονα και συνημίτονα). Δηλαδή [13]:

$$\Phi_i = \sum_{j=1}^{\kappa} w_j u_j$$

όπου το Φ_i αντιπροσωπεύει το i -οστό πρόσωπο (αφού έχει αφαιρεθεί το μέσο διάνυσμα), το w_j τα βάρη και το u_j τα ιδιοδιανύσματα.

Με λίγα λόγια, θέλουμε να βρούμε ένα σύνολο από εικόνες (που ονομάζονται eigenfaces και είναι ιδιοδιανύσματα των δεδομένων εκπαίδευσης) οι οποίες αν προστεθούν κατάλληλα θα μας παράγουν μια εικόνα που μας ενδιαφέρει.

Ένα χαρακτηριστικό παράδειγμα είναι το εξής: Στο Σχήμα 6, ένα πρόσωπο που ήταν στη βάση των δεδομένων εκπαίδευσης (training database) ανακατασκευάστηκε με υπολογισμό και άθροισμα των βαρών όλων των προσώπων και στη συνέχεια προστέθηκε σε αυτά η μέση εικόνα (μέσο πρόσωπο). Σημειώστε ότι στο Σχήμα οι eigenfaces δεν είναι τοποθετημένες με σειρά σπουδαιότητας αλλά έχουν επιλεγεί τυχαία από μια ομάδα 70 εικόνων από τη βάση δεδομένων MIT-CBCL (η φωτεινότητα των eigenfaces έχει ρυθμιστεί ώστε να φαίνονται καθαρότερες, ως εκ τούτου η φωτεινότητα της ανακατασκευασμένης εικόνας φαίνεται διαφορετική από εκείνη των εικόνων εκπαίδευσης).



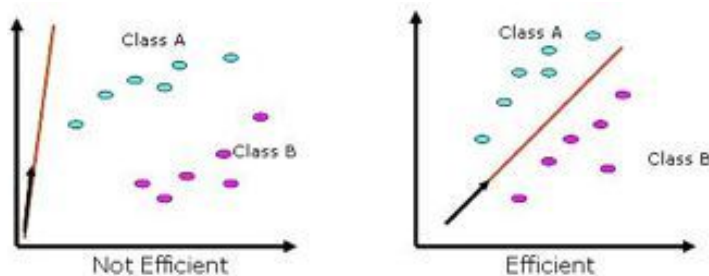
Σχήμα 6. Ανακατασκευή εικόνας με υπολογισμό και άθροισμα των βαρών

Προσέγγιση με βάση τη Θεωρία της Πληροφορίας

Η ιδέα της μεθόδου των eigenfaces θεωρεί την αναγνώριση προσώπου ως 2-Δ πρόβλημα αναγνώρισης και αυτό βασίζεται στην υπόθεση ότι κατά τη διαδικασία της αναγνώρισης τα πρόσωπα θα είναι σε όρθια θέση (upright) και εμπρόσθια (frontal). Επομένως, δεν είναι απαραίτητες 3-Δ λεπτομερείς πληροφορίες σχετικά με το πρόσωπο. Αυτό μειώνει την πολυπλοκότητα σε μεγάλο βαθμό.

Πριν παρουσιαστεί η μέθοδος της αναγνώριση προσώπου με χρήση eigenfaces, γινόταν χρήση χαρακτηριστικών, όπως η απόσταση ανάμεσα στα μάτια, τα αυτιά κτλ. Αυτό δεν ήταν πολύ αποτελεσματικό. Η μέθοδος των eigenfaces χρησιμοποιεί μια προσέγγιση της Θεωρία της Πληροφορίας, όπου οι πιο σημαντικές πληροφορίες του προσώπου κωδικοποιούνται σε μια ομάδα (σύνολο) προσώπων με την οποία θα υπάρχει καλύτερη διάκριση (αναγνώριση) στα πρόσωπα. Μετατρέπει τις εικόνες προσώπου σε μια σύνολο από πρόσωπα εκπαίδευσης, που ουσιαστικά αποτελούν τις κύριες συνιστώσες των εικόνων προσώπου.

Οι κυριότερες συνιστώσες (ή ιδιοδιανύσματα) προσπαθούν να αναπαραστήσουν τα δεδομένα πιο αποδοτικά. Αυτό είναι ιδιαίτερα χρήσιμο για τη μείωση του υπολογιστικού φόρτου. Για να γίνει αυτό κατανοητό, ας υποθέσουμε ότι έχουμε 60 σημεία, από αυτά περίπου τα 40 μπορεί να είναι ασήμαντα και μόνο 20 αντιπροσωπεύουν τη διακύμανση των δεδομένων ως σημαντική. Οπότε μπορούμε να χρησιμοποιήσουμε μόνο τα 20. Αυτό φαίνεται στο Σχήμα 7.



Σχήμα 7. Μη αποδοτική και αποδοτική ταξινόμηση

Μια τέτοια προσέγγιση της θεωρίας της πληροφορίας θα κωδικοποιήσει όχι μόνο τα τοπικά χαρακτηριστικά, αλλά και τα γενικά. Όταν βρούμε τις κύριες συνιστώσες ή τα ιδιοδιανύσματα του συνόλου της εικόνας, κάθε ιδιοδιάνυσμα έχει κάποια συνεισφορά από κάθε πρόσωπο που χρησιμοποιείται στο σύνολο εκπαίδευσης. Οπότε, τα ιδιοδιανύσματα εμφανίζονται ως πρόσωπα. Αυτά μοιάζουν με εικόνες φάντασμα ή eigenfaces. Κάθε εικόνα στο σύνολο εκπαίδευσης μπορεί να αναπαρασταθεί ως ένας γραμμικός συνδυασμός των βαρών των προσώπων εκπαίδευσης.

Ο αριθμός των eigenfaces που θα έχουμε θα είναι ίσος με τον αριθμό των εικόνων του συνόλου εκπαίδευσης. Έστω αυτός ο αριθμός M . Όπως ανέφερα και πριν από μια παράγραφο, μερικά από τα eigenfaces είναι πιο σημαντικά στην κωδικοποίηση της διακύμανσης των εικόνων προσώπου, οπότε μπορούμε επίσης να υπολογίσουμε κατά προσέγγιση τα πρόσωπα χρησιμοποιώντας μόνο τα K πιο σημαντικά eigenfaces.

Οπότε καταλήγουμε στα εξής τρία σημεία:

1. Υπάρχουν M εικόνες στο σύνολο εκπαίδευσης
2. Υπάρχουν K σημαντικά eigenfaces με τα οποία γίνεται πιο αποτελεσματικότερη αναγνώριση προσώπου. Σημειώνεται ότι $K < M$.
3. Όλες οι εικόνες είναι πίνακες διαστάσεων $N \times N$, όπου μπορούν να αναπαρασταθούν ως διάνυσμα διάστασης $N^2 \times 1$. Η ίδια λογική εφαρμόζεται σε εικόνες που δεν έχουν το ίδιο μήκος και πλάτος. Για παράδειγμα, μια εικόνα μεγέθους 112×112 , μπορεί να αναπαρασταθεί ως ένα διάνυσμα διάστασης 12544 ή ως ένα σημείο στο χώρο διαστάσεων 12544.

Ο αλγόριθμος για την εύρεση eigenfaces

1. Απόκτηση M εικόνων εκπαίδευσης $I_1, I_2, \dots, I_n$. Είναι σημαντικό οι εικόνες να είναι κεντραρισμένες (βλ. Σχήμα 8).



Σχήμα 8. Εικόνες εκπαίδευσης

2. Αναπαράσταση κάθε εικόνας I_i ως ένα διάνυσμα Γ_i .

$$I_i = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{bmatrix}_{N \times N} \xrightarrow{\text{concatenation}} \begin{bmatrix} a_{11} \\ \vdots \\ a_{1N} \\ \vdots \\ a_{2N} \\ \vdots \\ a_{NN} \end{bmatrix}_{N^2 \times 1} = \Gamma_i$$

3. Εύρεση του μέσου διανύσματος προσώπου Ψ .

$$\Psi = \sum_{i=1}^M \Gamma_i$$

4. Αφαίρεση του μέσου προσώπου από το κάθε μέσο διάνυσμα προσώπου Γ_i για να πάρουμε ένα σύνολο διανυσμάτων Φ_i . Ο σκοπός της αφαίρεσης της μέσης εικόνας από το κάθε διάνυσμα εικόνας είναι η εξαγωγή μόνο των ιδιαίτερων χαρακτηριστικών κάθε προσώπου με την αφαίρεση των κοινών χαρακτηριστικών του.

$$\Phi_i = \Gamma_i - \Psi$$

5. Εύρεση του πίνακα συνδιασποράς C : $C = AA^T$, όπου $A = [\Phi_1, \Phi_2, \dots, \Phi_M]$

Σημειώστε ότι ο πίνακας C είναι διαστάσεων $N^2 \times N^2$ και ο A διαστάσεων $N^2 \times M$

6. Τώρα θα υπολογίσουμε τα ιδιοδιανύσματα u_i του πίνακα C . Όμως ο C είναι πίνακας διαστάσεων $N^2 \times N^2$ και θα επιστρέψει N^2 ιδιοδιανύσματα διαστάσεων N^2 το καθένα. Αυτός ο αριθμός για μια τυπική εικόνα είναι πολύ μεγάλος και ο υπολογισμός του θα οδηγούσε το σύστημα σε έλλειψη μνήμης. Οπότε:

7. Αντί για τον πίνακα AA^T , έστω ο πίνακας $A^T A$. Ο A έχει διαστάσεις $N^2 \times M$, οπότε ο πίνακας $A^T A$ θα έχει διαστάσεις $M \times M$. Με την εύρεση των ιδιοδιανυσμάτων αυτού του πίνακα, καταλήγουμε ότι επιστρέφει M ιδιοδιανύσματα με διάσταση $M \times 1$ το καθένα. Ας ονομάσουμε αυτά τα ιδιοδιανύσματα v_i .

Από τις ιδιότητες των πινάκων έχουμε ότι $u_i = Av_i$. Το v_i υπολογίστηκε νωρίτερα. Οπότε με τη χρήση του v_i , μπορούμε να υπολογίσουμε τα M μεγαλύτερα ιδιοδιανύσματα του πίνακα AA^T . Το $M \ll N^2$, επειδή το M είναι το σύνολο των εικόνων εκπαίδευσης.

8. Εύρεση των M καλύτερων ιδιοδιανυσμάτων του $C = AA^T$, με χρήση της σχέσης $u_i = Av_i$. Επίσης, $\|u_i\| = 1$



Σχήμα 9. Έξι eigenfaces για την διαδικασία της εκπαίδευσης διαλεγμένα από τη βάση δεδομένων MIT-CBCL, με τυχαία σειρά

9. Επιλογή των K καλύτερων ιδιοδιανυσμάτων. Η επιλογή των συγκεκριμένων ιδιοδιανυσμάτων γίνεται βάσει εμπειρίας.

Εύρεση των βαρών

Τα ιδιοδιανύσματα που υπολογίστηκαν στην προηγούμενη ενότητα, u_i όταν μετασχηματιστούν σε έναν πίνακα με την αντίστροφη διαδικασία του βήματος 2, εμφανίζονται ως πρόσωπο. Από το γεγονός ότι είναι ιδιοδιανύσματα και έχουν εμφάνιση σαν προσώπου προκύπτει η ονομασία τους 'eigenfaces' (ιδιοεικόνες). Μερικές φορές λέγονται εικόνες φάντασμα λόγω της περίεργης εμφάνισής τους.

Κάθε πρόσωπο στο σύνολο εκπαίδευσης (εκτός από το μέσο πρόσωπο), Φ_i μπορεί να αναπαρασταθεί ως ένας γραμμικός συνδυασμός των ιδιοδιανυσμάτων u_i : $\Phi_i = \sum_{j=1}^K w_j u_j$, όπου τα u_j είναι eigenfaces.

Αυτά τα βάρη υπολογίζονται ως: $w_j = u_j^T \Phi_i$

Κάθε κανονική εικόνα εκπαίδευσης (normalized training image) αναπαρίσταται ως ένα διάνυσμα.

$$\Omega_i = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_K \end{bmatrix}, \text{ όπου } i = 1, 2, \dots, M$$

Αυτό σημαίνει ότι πρέπει να υπολογιστεί ένα τέτοιο διάνυσμα που αντιστοιχεί σε κάθε εικόνα του σύνολο εκπαίδευσης και να αποθηκευτεί ως πρότυπο (template).

Η διαδικασία της αναγνώρισης

Έστω ότι έχουμε βρει τα eigenfaces για τις εικόνες εκπαίδευσης, τα βάρη τους μετά την επιλογή των καταλληλότερων ιδιοδιανυσμάτων και έχουμε αποθηκεύσει αυτά τα διανύσματα που αντιστοιχούν στην κάθε εικόνα εκπαίδευσης. Για να αναγνωριστεί ένα άγνωστο πρόσωπο (probe face) πρέπει:

1. Να εξομαλύνουμε (normalize) την εισερχόμενη προς εξέταση εικόνα Γ με $\Phi = \Gamma - \Psi$

2. Μετά την προβάλλουμε στο χώρο των ιδιοδιανυσμάτων (eigen-space) για να υπολογίσουμε τα βάρη $w_i = u_i^T \Phi$
3. Οπότε το Φ μπορεί να αντιστοιχηθεί ως

$$\Omega = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_K \end{bmatrix}$$

Από τη στιγμή που το διάνυσμα βάρους για την άγνωστη εικόνα έχει υπολογιστεί, επόμενο βήμα είναι η ταξινόμησή της. Για τη διαδικασία της ταξινόμησης μπορούμε να χρησιμοποιήσουμε υπολογισμό απόστασης ως εξής:

- Εύρεση του $e_r = \min \|\Omega - \Omega_i\|$. Δηλαδή παίρνουμε το διάνυσμα βάρους της άγνωστης εικόνας και υπολογίζουμε την απόστασή του με τα διανύσματα βάρους που σχετίζονται με την κάθε μία εικόνα εκπαίδευσης, ως απόλυτη διαφορά. Από τις απόλυτες διαφορές που προκύπτουν επιλέγεται η ελάχιστη τιμή.
- Εάν το $e_r < \theta$, όπου θ είναι το κατώφλι, τότε το πρόσωπο προς εξέταση αναγνωρίζεται ως το πρόσωπο που δίνει τη μικρότερη απόλυτη διαφορά.
- Εάν το $e_r > \theta$, τότε το πρόσωπο προς εξέταση θεωρείται άγνωστο.

Μετρήσεις αποστάσεων (distance measures)

Σε πολλές εφαρμογές είναι απαραίτητο να βρίσκουμε την απόσταση μεταξύ δύο εικονοστοιχείων ή μεταξύ δύο στοιχείων μιας εικόνας. Δυστυχώς δεν υπάρχει ένας και μοναδικός τρόπος για τον προσδιορισμό της απόστασης στις ψηφιακές εικόνες. Η απόσταση μπορεί να υπολογισθεί με πολλούς διαφορετικούς τρόπους. Για κάθε εικονοστοιχείο p , q και r κάθε μετρική απόσταση θα πρέπει να ικανοποιεί τις παρακάτω ιδιότητες:

1. $d(p, q) \geq 0$
2. $d(p, q) = 0$ αν και μόνο αν $p = q$
3. $d(p, q) = d(q, p)$
4. $d(p, r) \leq d(p, q) + d(q, r)$

Η πιο διαδεδομένη απόσταση είναι η Ευκλείδεια. Μια άλλη απόσταση που λαμβάνει υπόψη τη διασπορά των δεδομένων είναι η απόσταση Mahalanobis. Η απόσταση Mahalanobis παρουσιάζει γενικά καλύτερη απόδοση στο συγκεκριμένο πρόβλημα.

Ευκλείδεια απόσταση (Euclidean distance). Αυτή είναι η πιο συχνά χρησιμοποιούμενη μέθοδος μέτρησης απόστασης. Εκφράζει απλά τη

γεωμετρική απόσταση δύο εικονοστοιχείων σε μια εικόνα. Υπολογίζεται ως:

$$\|x - y\|_e = \sqrt{\sum_i (x_i - y_i)^2}$$

Η μέθοδος αυτή έχει σημαντικά πλεονεκτήματα. Όμως οι αποστάσεις μπορεί να επηρεάζονται σημαντικά από διαφορετικές κλιμακώσεις τμημάτων της εικόνας. Επίσης, αν για παράδειγμα μια από τις διαστάσεις μιας εικόνας εκφράζεται σε *cm* και μετατραπεί σε *mm*, τότε η σχετική Ευκλείδεια απόσταση των αντικειμένων διαφοροποιείται.

Απόσταση Mahalanobis (Mahalanobis distance). Ο τύπος που χρησιμοποιείται για τον υπολογισμό της απόστασης Mahalanobis είναι ο εξής:

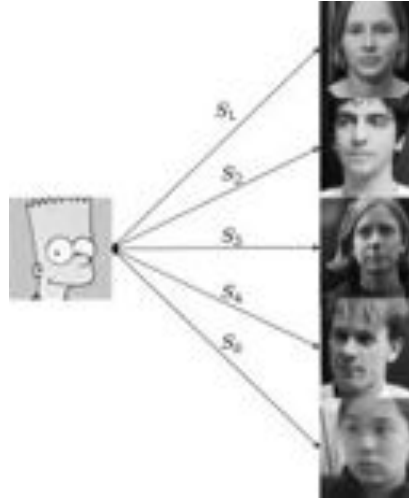
$$d(x, y) = \sqrt{(x - y)^T C^{-1} (x - y)}$$

Όπου x , y διανύσματα και C ο πίνακας συνδιασποράς μεταξύ των δεδομένων. Η μέθοδος αυτή χρησιμοποιείται ιδιαίτερα για την εύρεση της ομοιότητας μεταξύ των σημείων και θεωρείται ανώτερη από την Ευκλείδεια απόσταση, διότι δεν εξαρτάται μόνο από τη μέση τιμή αλλά και από το συσχετισμό ανάμεσα στις μεταβλητές καθώς και από την κατανομή των σημείων.

Επιλογή κατώφλιου (threshold)

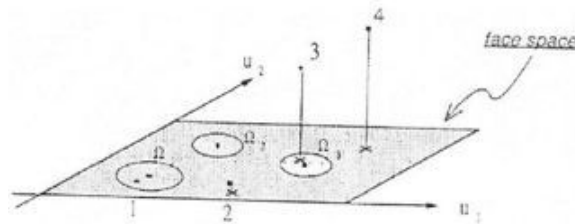
Έστω ότι έχουμε μόνο πέντε εικόνες στο σύνολο εκπαίδευσης και μια εικόνα προς αναγνώριση που δεν είναι στο σύνολο εκπαίδευσης (Σχήμα 10). Η βαθμολογία για κάθε μία από τις 5 εικόνες θα βρεθεί ως την προς αναγνώριση εικόνα. Ακόμη κι αν καμία εικόνα της προς αναγνώριση εικόνας δε βρίσκεται καταχωρημένη στη βάση δεδομένων, εξακολουθεί να μπορεί να χαρακτηριστεί ως η εικόνα της οποίας το σκορ (απόσταση) είναι το χαμηλότερο. Είναι σαφές ότι αυτή είναι μια ανωμαλία που πρέπει να εξετάσουμε. Για αυτό το σκοπό μελετούμε το κατώφλι με ευρεστικό τρόπο.

Έστω λοιπόν ότι η εικόνα προς αναγνώριση είναι μια εικόνα Simpson (μη-πρόσωπο) (Σχήμα 10). Επομένως, η εικόνα αυτή θα συγκριθεί με κάθε μία από τις εικόνες εκπαίδευσης. Ας υποθέσουμε ότι το S4 είναι το χαμηλότερο σκορ από όλες τις εικόνες εκπαίδευσης. Αυτό σημαίνει ότι η εικόνα προς αναγνώριση θα χαρακτηριστεί ως η εικόνα με το σκορ S4 ενώ προφανώς δεν είναι αυτή και φυσικά δεν ανήκει ούτε στη βάση δεδομένων. Για να επιλέξουμε το κατώφλι κάτω από το οποίο η εικόνα δε θα ταυτίζεται με κάποια της βάσης δεδομένων, διαλέγουμε ένα μεγάλο σύνολο τυχαίων εικόνων (με πρόσωπο και χωρίς) και υπολογίζουμε στη συνέχεια τις βαθμολογίες για εικόνες ανθρώπων στη βάση δεδομένων, αλλά και για αυτό το τυχαίο σύνολο και αναλόγως επιλέγουμε το κατώφλι θ .



Σχήμα 10. Πέντε (5) εικόνες στο σύνολο εκπαίδευσης

Υποχώρος προσώπου (face space)



Σχήμα 11. Υποχώρος προσώπου

Μια απλοποιημένη αναπαράσταση του υποχώρου προσώπου φαίνεται στο Σχήμα 11. Οι εικόνες ενός προσώπου και ιδίως τα πρόσωπα του συνόλου εκπαίδευσης θα πρέπει να βρίσκονται κοντά στον υποχώρο προσώπου, ο οποίος γενικά περιγράφει τις εικόνες που μοιάζουν με πρόσωπο. Η απόσταση προβολής e_r θα πρέπει να είναι κάτω από ένα συγκεκριμένο κατώφλι θ . Οι εικόνες γνωστών ατόμων πέφτουν κοντά σε μια κλάση προσώπου μέσα στον υποχώρο προσώπου. Υπάρχουν τέσσερις πιθανοί συνδυασμοί για την είσοδο μιας εικόνα προσώπου:

1. *Κοντά σε κλάση προσώπου και κοντά στον υποχώρο προσώπου:* Αυτό συμβαίνει όταν η εικόνα προς εκπαίδευση (probe) είναι ενός γνωστού ατόμου (*known* = η εικόνα του προσώπου είναι ήδη καταχωρημένη στη βάση δεδομένων).
2. *Κοντά στον υποχώρο προσώπου, αλλά μακριά από την κλάση του προσώπου:* Αυτό συμβαίνει όταν η εικόνα προς εκπαίδευση (probe image) είναι εικόνα προσώπου που δεν ανήκει σε κανέναν στη

βάση δεδομένων, δηλαδή είναι μακριά από οποιαδήποτε κλάση προσώπου (face class).

3. *Μακριά από τον υποχώρο προσώπου και κοντά στην κλάση προσώπου:* Αυτό συμβαίνει όταν η εικόνα προς εκπαίδευση δεν είναι πρόσωπο, ωστόσο εξακολουθεί να μοιάζει με μια συγκεκριμένη κλάση προσώπου που είναι αποθηκευμένη στη βάση δεδομένων.
4. *Μακριά από τον υποχώρο προσώπου και την κλάση προσώπου:* Όταν η εικόνα προς εκπαίδευση δεν είναι εικόνα προσώπου, δηλαδή είναι μακριά από τον υποχώρο προσώπου και δεν μοιάζει με καμία κλάση προσώπου που είναι αποθηκευμένη στη βάση δεδομένων.

Από τους τέσσερις συνδυασμούς, ο τρίτος παράγει τα περισσότερα ψευδώς θετικά αποτελέσματα. Για την αποφυγή τους συνιστάται η χρήση διαδικασίας ανίχνευσης προσώπου (face detection) σε τέτοια συστήματα, ώστε να αποκλείονται εικόνες που δεν περιέχουν πρόσωπο.

Κεφάλαιο 3. Σύστημα αναγνώρισης προσώπων με χρήση eigenfaces

Απαιτήσεις υλοποίησης συστήματος αναγνώρισης προσώπων

Στο πλαίσιο της παρούσας εργασίας τέθηκαν αρχικά κάποιες προδιαγραφές και περιορισμοί στην υλοποίηση του συστήματος αναγνώρισης προσώπων. Αρχικά, η εφαρμογή αναγνώρισης θα έπρεπε να είναι διαδικτυακή (web based), φορητή (portable) και ανεξάρτητη πλατφόρμας (platform independent). Η γλώσσα προγραμματισμού C/C++ είναι κατάλληλη για την υλοποίηση της εφαρμογής. Από πλευράς υλικού υπολογιστικού συστήματος τέθηκαν ιδιαίτερα χαμηλές προδιαγραφές που αντιστοιχούν σε σύστημα τύπου netbook ώστε να καταστούν εμφανείς οι υπολογιστικές απαιτήσεις του αλγορίθμου. Το λογισμικό ανάπτυξης του συστήματος θα μπορούσε να είναι οποιοσδήποτε μεταγλωττιστής (compiler) C/C++, όπως Dev C++ και η βιβλιοθήκη αλγορίθμων επεξεργασίας εικόνας και όρασης υπολογιστών OpenCV.

Πρόγραμμα εργασιών υλοποίησης

Η υλοποίηση του συστήματος χωρίστηκε στις εξής ενότητες εργασίας (φάσεις υλοποίησης) για αποδοτικότερη ολοκλήρωση:

- *Ανάλυση απαιτήσεων:* Σε αυτή τη φάση συγκεντρώθηκαν οι πλήρεις προδιαγραφές, για να επιτευχθεί καλύτερη κατανόηση του έργου.
- *Βιβλιογραφική έρευνα:* Σε αυτή τη φάση του έργου συγκεντρώθηκαν οι απαραίτητες πληροφορίες σχετικά με την ανίχνευση και αναγνώριση προσώπων και τον τρόπο λειτουργίας της. Η ανάλυση προγραμματίστηκε για ένα διάστημα 2 μηνών λόγω της ανάγκης για αναζήτηση σε ξενόγλωσση βιβλιογραφία.
- *Υλοποίηση και διαμόρφωση του συστήματος:* Η βασική φάση προγραμματισμού για την υλοποίηση του συστήματος.
- *Έλεγχος λειτουργίας και απόδοσης του συστήματος:* Η δοκιμή του συστήματος απαιτεί αναζήτηση και διόρθωση λαθών στον προγραμματισμό, δοκιμή σωστής εκτέλεσης και εκτίμηση της απόδοσης του συστήματος καθώς και ενδεχόμενες βελτιστοποιήσεις.
- *Παράδοση:* Στη φάση αυτή περιλαμβάνεται η συγγραφή της παρούσας εργασίας και της παρουσίασης.

Περί της γλώσσας προγραμματισμού C++

Η C++ είναι μια γενικού σκοπού γλώσσα προγραμματισμού Η/Υ. Θεωρείται μέσου επιπέδου γλώσσα, καθώς περιλαμβάνει ένα συνδυασμό χαρακτηριστικών από γλώσσες υψηλού και χαμηλού επιπέδου. Υποστηρίζει δομημένο, αντικειμενοστραφή και γενικό προγραμματισμό.

Η γλώσσα αναπτύχθηκε από τον Bjarne Stroustrup το 1979 στα εργαστήρια Bell της AT&T, ως βελτίωση της ήδη υπάρχουσας γλώσσας προγραμματισμού C, και αρχικά ονομάστηκε 'C with Classes', δηλαδή «C με Κλάσεις». Μετονομάστηκε σε C++ το 1983. Οι βελτιώσεις ξεκίνησαν με την προσθήκη κλάσεων, και ακολούθησαν, μεταξύ άλλων, εικονικές συναρτήσεις, υπερφόρτωση τελεστών, πολλαπλή κληρονομικότητα, πρότυπα κ.α.

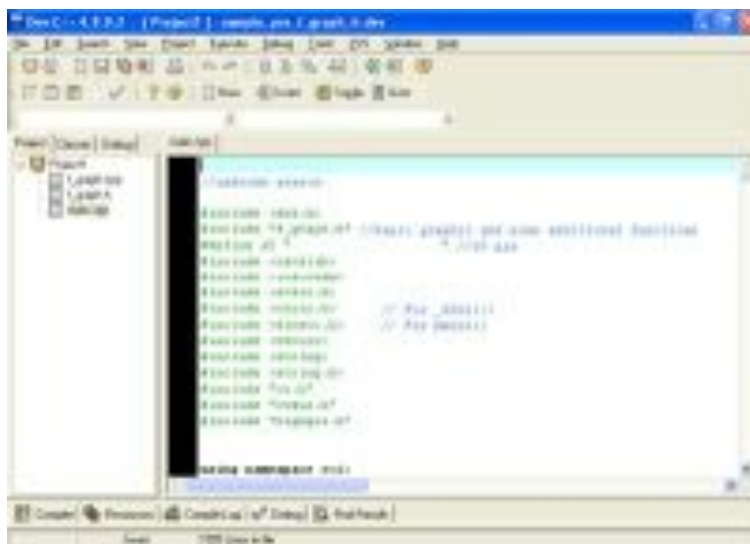
Η γλώσσα ορίστηκε παγκοσμίως, το 1998, με το πρότυπο ISO/IEC 14882:1998. Η τρέχουσα έκδοση αυτού του προτύπου είναι αυτή του 2003, η ISO/IEC 14882:2003. Μια καινούρια έκδοση είναι υπό ανάπτυξη, γνωστή ανεπίσημα με την ονομασία C++0x.

Στο βιβλίο του 'The Design and Evolution of C++' (1994), ο Bjarne Stroustrup περιγράφει κάποιους κανόνες που χρησιμοποιεί για το σχεδιασμό της C++:

- η C++ είναι σχεδιασμένη ως μια γενικής χρήσης γλώσσα με στατικούς τύπους, που είναι όσο *αποτελεσματική και φορητή, όσο η C*
- η C++ είναι σχεδιασμένη να υποστηρίζει άμεσα και σφαιρικά πολλά είδη προγραμματισμού (*δομημένος προγραμματισμός, αντικειμενοστραφής προγραμματισμός, γενικός προγραμματισμός*)
- η C++ είναι σχεδιασμένη να *δίνει επιλογές στον προγραμματιστή*, ακόμα κι αν του επιτρέπει να επιλέξει λανθασμένα
- η C++ είναι σχεδιασμένη να είναι όσο το δυνατόν *συμβατή με τη C*, ώστε να διευκολύνει τη μετάβαση από τη C
- η C++ *αποφεύγει χαρακτηριστικά που αναφέρονται σε συγκεκριμένες πλατφόρμες ή δεν είναι γενικής χρήσης*
- η C++ *δεν έχει κόστος για χαρακτηριστικά της γλώσσας που δεν χρησιμοποιούνται*
- η C++ είναι σχεδιασμένη να *λειτουργεί χωρίς κάποιο εξελιγμένο προγραμματιστικό περιβάλλον*

Ανάπτυξη σε περιβάλλον Dev-C++

Το Dev-C++ (BloodShed Software) [14] είναι ένα ιδιαίτερα χρήσιμο περιβάλλον προγραμματισμού που μπορεί να χρησιμοποιηθεί ακόμη και από νεοεισερχόμενους στον κόσμο του προγραμματισμού χρησιμοποιώντας τις γλώσσες C/C++. Το περιβάλλον αποτελείται από ένα IDE-Ολοκληρωμένο Περιβάλλον Ανάπτυξης, που αναπτύσσεται στη γλώσσα Delphi (Σχήμα 12). Από το περιβάλλον έχει κανείς πολλαπλές επιλογές χρήσης: Φίλτρο εκκαθάρισης κώδικα (code purifier), εφαρμογή σύνταξης, διορθωτής σφαλμάτων, δημιουργία εκτελέσιμων, κοκ. Το Dev-C++ παρέχει πλήρη διαχείριση έργου (project) με απόλυτα προσαρμοσμένες οθόνες για επεξεργασία έργων. Προσφέρει υποστήριξη για εκτύπωση και σύνδεση με βιβλιοθήκες ανοικτού κώδικα (open source libraries).



Σχήμα 12. Γραφικό περιβάλλον του Dev-C++

Open Source Computer Vision Library - OpenCV

Η βιβλιοθήκη 'Open Source Computer Vision Library' (OpenCV) [15] είναι μία ελεύθερα διαθέσιμη, ανοικτού κώδικα βιβλιοθήκη σε C++, που αφορούν την επεξεργασία εικόνας και την τεχνητή όραση. Η OpenCV παρέχει μια στέρεη υποδομή για την τεχνητή όραση που επιτρέπει στους ειδικούς να δουλέψουν σε ένα υψηλότερο επίπεδο από το στοιχειώδες. Με την τύπου BSD (Berkeley Software Distribution) άδεια χρήσης της, η OpenCV παρέχεται από την Intel για ελεύθερη εμπορική και ερευνητική χρήση. Η υποστήριξη της OpenCV στην τεχνητή όραση είναι εκτεταμένη. Υποστηρίζει αλγορίθμους για την ανάγνωση, απεικόνιση, επεξεργασία και αποθήκευση στατικών αλλά και κινούμενων ψηφιακών εικόνων. Παρέχονται ένα πλήθος από ρουτίνες για την επεξεργασία εικόνας και

υποστηρίζονται ρουτίνες υπολογιστικής γεωμετρίας, αναγνώρισης προτύπων και αλγόριθμοι ανίχνευσης αντικειμένων. Αξίζει δε να σημειωθεί ότι η βιβλιοθήκη είναι συνεχώς υπό ανάπτυξη με τη διαρκή προσθήκη νέων αλγορίθμων.

Ανίχνευση προσώπων με την OpenCV

Η OpenCV [16] παρέχει ένα χαμηλού και υψηλού επιπέδου περιβάλλον ανάπτυξης (Application Programming Interface-API) για την ανίχνευση προσώπων ή άλλων αντικειμένων, ανάλογα με την εκπαίδευση που έχει προηγηθεί στον ταξινομητή αντικειμένων που προσφέρει. Ο ταξινομητής αντικειμένων που προσφέρει είναι αυτός που έχει προταθεί από τους Viola & Jones [17],[18] και έχει βελτιωθεί από τον Lienhart [19],[20] κ.ά. Ο ταξινομητής για τα πρόσωπα έχει εκπαιδευτεί με εκατοντάδες θετικά δείγματα προσώπων μεγέθους 20x20 εικονοστοιχείων και αυθαίρετα αρνητικά δείγματα ίδιου μεγέθους. Η OpenCV υλοποιεί τον αλγόριθμο *Haartraining* για την εκπαίδευση ενός ταξινομητή με ένα δεδομένο σύνολο θετικών και αρνητικών δειγμάτων, που αφορούν ένα συγκεκριμένο αντικείμενο. Ο αλγόριθμος αυτός δημιουργεί ένα xml αρχείο που περιέχει τα δεδομένα που αφορούν τον εν λόγω ταξινομητή. Το αρχείο αυτό μπορεί να χρησιμοποιηθεί στη συνέχεια για την ανίχνευση του συγκεκριμένου αντικειμένου. Αφού εκπαιδευτεί ο ταξινομητής, μπορεί να εφαρμοστεί σε μία περιοχή ενδιαφέροντος μιας εικόνας εισόδου. Ο ταξινομητής αποκρίνεται με "1" αν η περιοχή περιέχει το αναζητούμενο αντικείμενο και με "0" σε αντίθετη περίπτωση. Για να αναζητηθεί το αντικείμενο σε ολόκληρη την εικόνα, το παράθυρο αναζήτησης μπορεί να μετακινηθεί σε διαφορετικές θέσεις και να ελεγχθεί η απόκριση του ταξινομητή. Ο ταξινομητής έχει σχεδιαστεί έτσι ώστε το παράθυρο αναζήτησης να μπορεί να τροποποιηθεί ως προς τις διαστάσεις του εύκολα, για να ανευρεθούν αντικείμενα διαφορετικών διαστάσεων, αντί να απαιτείται αντίστοιχη μεταβολή των διαστάσεων της ίδιας της εικόνας. Έτσι, για να βρεθεί ένα αντικείμενο άγνωστων διαστάσεων σε μία εικόνα, η διαδικασία αναζήτησης μπορεί να γίνει όσες φορές χρειάζεται σε διαφορετικές κλίμακες.

Ο τελικός ταξινομητής αποτελείται από επίπεδα απλούστερων ταξινομητών σε διάταξη καταρράκτη (cascade). Κάθε επίπεδο του καταρράκτη αποτελείται από ένα ισχυρό ταξινομητή (strong classifier) που κατασκευάζεται από απλούστερους αδύναμους ταξινομητές (weak classifiers) με την τεχνική της ενίσχυσης (boosting). Οι τεχνικές ενίσχυσης που υποστηρίζονται είναι η Discrete AdaBoost, η Real AdaBoost, η Gentle AdaBoost και η LogitBoost. Οι αδύναμοι ταξινομητές που χρησιμοποιούνται είναι δέντρα απόφασης (decision trees) με 1 έως 4 κόμβους. Οι αδύναμοι ταξινομητές παίρνουν ως είσοδο ορθογώνια χαρακτηριστικά (features) τύπου Haar, που καθορίζονται από τη θέση, τις διαστάσεις, τον προσανατολισμό και την μορφή τους και παίρνουν τιμές

μία σταθμισμένη διαφορά των εντάσεων των εικονοστοιχείων στις περιοχές που ορίζουν. Τα αθροίσματα των εντάσεων υπολογίζονται με τη χρήση της εικόνας ολοκλήρωμα (integral image) που έχει προϋπολογιστεί. Έτσι μια χαμηλού επιπέδου ρουτίνα, για παράδειγμα, παρέχει τη δυνατότητα στο χρήστη να ελέγξει μία συγκεκριμένη περιοχή της εικόνας για το αν περιέχει ή όχι πρόσωπο. Βοηθητικές ρουτίνες υπολογίζουν τις εικόνες ολοκληρώματα, και ρυθμίζουν την κλίμακα αναζήτησης του ταξινομητή για πρόσωπα διαφορετικών μεγεθών κλπ. Αντίθετα, μία υψηλού επιπέδου συνάρτηση, όπως η `cvHaarDetectObjects` κάνει όλα τα προηγούμενα, χρησιμοποιώντας τον ταξινομητή που είναι αποθηκευμένος σε xml αρχείο, παρέχοντας αυτόματα τα αποτελέσματα ανίχνευσης στο χρήστη, πράγμα που είναι τις περισσότερες φορές αρκετό.

Η συνάρτηση ανίχνευσης

Η συνάρτηση `cvHaarDetectObjects` βρίσκει ορθογώνιες περιοχές σε μια δεδομένη εικόνα που είναι πιθανό να περιέχουν αντικείμενα που ο ταξινομητής έχει εκπαιδευτεί να διακρίνει και επιστρέφει αυτές τις περιοχές σαν μία σειρά από ορθογώνια. Η συνάρτηση σαρώνει την εικόνα αρκετές φορές σε διαφορετικές κλίμακες, και κάθε φορά, για κάθε περιοχή, εφαρμόζει τον ταξινομητή για να ελέγξει την ύπαρξη του προς ανίχνευση αντικειμένου. Αφού προχωρήσει η διαδικασία και συλλεχθούν τα υποψήφια ορθογώνια (περιοχές που επαληθεύονται από τον ταξινομητή), ομαδοποιούνται σε ομάδες επικαλυπτόμενων ορθογωνίων και επιστρέφονται σαν μία σειρά από μέσα ορθογώνια πλαίσια, ένα για κάθε ομάδα. Οι προκαθορισμένες παράμετροι (`scale_factor=1.1`, `min_neighbors=3`, `flags=0`) είναι ρυθμισμένες για μεγαλύτερης ακρίβειας αν και χρονοβόρα ανίχνευση. Για ταχύτερη λειτουργία σε εικόνες βίντεο οι ρυθμίσεις είναι:

```
scale_factor=1.2,min_neighbors=2,flags=
CV_HAAR_DO_CANNY_PRUNING, min_size=<minimum possible face
size>
```

Η συνάρτηση `cvHaarDetectObjects` δέχεται τις εξής παραμέτρους:

```
CvSeq* cvHaarDetectObjects( const CvArr* image,
                             CvHaarClassifierCascade* cascade,
                             CvMemStorage* storage,
                             double scale_factor=1.1,
                             int min_neighbors=3,
                             int flags=0,
                             CvSize min_size=cvSize(0,0) );
```

`image`

Η εικόνα στην οποία ανιχνεύονται τα αντικείμενα

`cascade`

Η αναφορά στο χρησιμοποιούμενο ταξινομητή

`storage`

Μνήμη για την αποθήκευση της σειράς των ορθογωνίων που ανιχνεύτηκαν τα αντικείμενα.

`scale_factor`

Ο συντελεστής κατά τον οποίο μεγεθύνεται το παράθυρο ανίχνευσης μεταξύ των διαδοχικών σαρώσεων της εικόνας, π.χ., 1.1 σημαίνει αύξηση του παραθύρου κατά 10%. Στην εφαρμογή έχει χρησιμοποιηθεί η τιμή 1.1.

`min_neighbors`

Ο ελάχιστος αριθμός (μείον 1) των γειτονικών ορθογωνίων που αντιστοιχούν στο ίδιο αντικείμενο και συγχωνεύονται σε ένα. Όλες οι ομάδες με μικρότερο αριθμό ανιχνευμένων ορθογωνίων απορρίπτονται ως εσφαλμένες ανιχνεύσεις. Εάν η `min_neighbors` είναι 1, η συνάρτηση δεν κάνει καμία ομαδοποίηση και επιστρέφει όλα τα ανιχνευμένα υποψήφια ορθογώνια. Αυτή η λειτουργία είναι χρήσιμη στην περίπτωση που ο χρήστης θέλει να εφαρμόσει μια εξειδικευμένη διαδικασία ομαδοποίησης. Στην εφαρμογή έχει χρησιμοποιηθεί η τιμή 3.

`flags`

Παράμετρος που καθορίζει τον τρόπο λειτουργίας. Ένας διαθέσιμος τρόπος λειτουργίας είναι ο `CV_HAAR_DO_CANNY_PRUNING`. Σ' αυτή την περίπτωση, η συνάρτηση χρησιμοποιεί τη μέθοδο ανίχνευσης ακμών του Canny για να απορρίψει κάποιες περιοχές που περιέχουν πολύ λίγες ή πάρα πολλές ακμές σε σχέση με το αντικείμενο που ανιχνεύεται. Οι συγκεκριμένες τιμές έχουν ρυθμιστεί έτσι ώστε να επιταχύνεται η διαδικασία στην ανίχνευση προσώπων. Στην εφαρμογή έχουν χρησιμοποιηθεί οι μέθοδοι `FIND_BIGGEST_OBJECT` και `DO_ROUGH_SEARCH`. Με τη χρησιμοποίηση της μεθόδου `FIND_BIGGEST_OBJECT`, η συνάρτηση βρίσκει το μεγαλύτερο αντικείμενο (εάν υπάρχει) στην εικόνα. Οπότε η ακολουθία εξόδου θα περιέχει ένα (ή μηδέν) στοιχείο. Η μέθοδος `DO_ROUGH_SEARCH` χρησιμοποιείται μόνο όταν έχει οριστεί η μέθοδος `FIND_BIGGEST_OBJECT` και το `min_neighbors>0`. Εάν η σημαία έχει οριστεί, η συνάρτηση δε ψάχνει για αντικείμενα μικρότερου μεγέθους από τη στιγμή που έχει βρει το βασικό αντικείμενο στη τρέχουσα κλίμακα. Όταν η τιμή του `min_neighbors` είναι σταθερή εμφανίζεται μεγαλύτερο ορθογώνιο ανίχνευσης από τη συνηθισμένη λειτουργία. Όμως είναι πιο γρήγορη και με μια μεγαλύτερη τιμή του `min_neighbors` επιστρέφει καλύτερη ακρίβεια.

`min_size`

Ελάχιστο μέγεθος παραθύρου από το οποίο ξεκινά η ανίχνευση. Η προκαθορισμένη τιμή είναι το μέγεθος των δειγμάτων με τα οποία εκπαιδεύτηκε ο ταξινομητής (~20×20 για την ανίχνευση προσώπων). Στην εφαρμογή έχει χρησιμοποιηθεί η τιμή 20x20.

Οι Μέθοδοι ανίχνευσης

Η OpenCV παρέχει 4 μεθόδους ανίχνευσης προσώπων σε κατά μέτωπο (frontal) στάση, με τις οποίες έχει ήδη εκπαιδευτεί ο ταξινομητής και είναι αποθηκευμένες σε αντίστοιχα αρχεία xml, τα οποία μπορούν να φορτωθούν με τη χρήση της συνάρτησης cvLoad:

- Stump-based 24x24 Discrete Adaboost with cascade (haarcascade_frontalface_default.xml)
- Stump-based 20x20 Gentle Adaboost with cascade (haarcascade_frontalface_alt.xml)
- CART-based (2-nodes) 20x20 Gentle Adaboost with cascade (haarcascade_frontalface_alt2.xml)
- Stump-based 20x20 Gentle Adaboost with tree of stage classifiers (haarcascade_frontalface_alt_tree.xml)

Αυτές οι τέσσερις διαθέσιμες μέθοδοι ταξινόμησης θα αναφέρονται σε συντομία στη συνέχεια ως *default*, *alt*, *alt2* και *tree*.

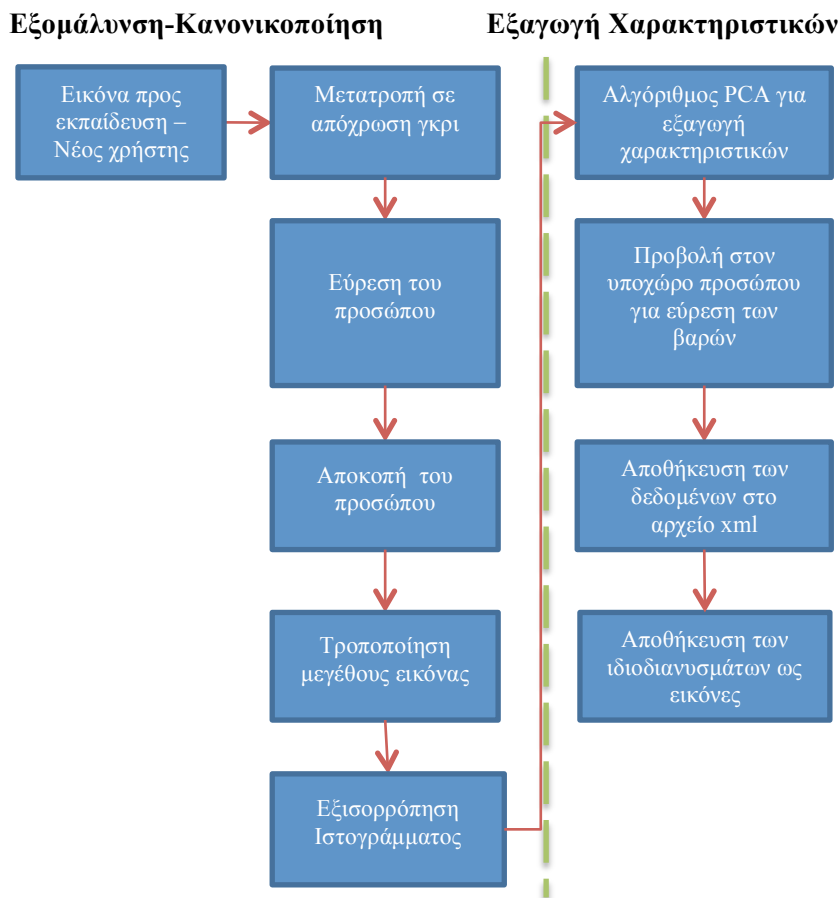
Υλοποίηση του συστήματος αναγνώρισης προσώπων

Το σύστημα που υλοποιήθηκε αποτελείται από δύο βασικά δομικά στοιχεία:

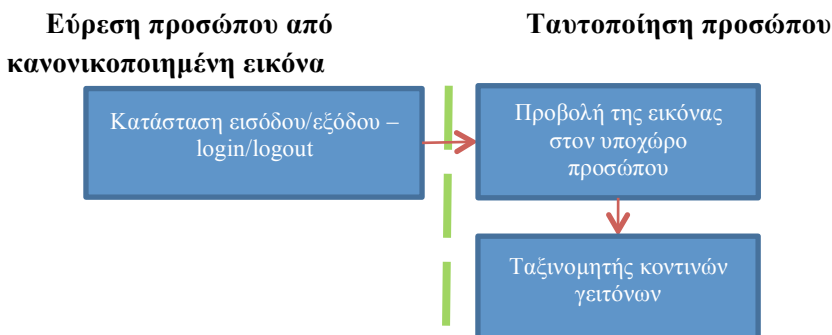
- *Αναγνώριση νέου προσώπου*: εύρεση του προσώπου στην εικόνα, εξομάλυνση, εξαγωγή χαρακτηριστικών του προσώπου και αποθήκευση στη βάση δεδομένων (βλ. Σχήμα 13)
- *Ταυτοποίηση προσώπου*: εντοπισμός προσώπου στην εικόνα και ταυτοποίηση του με εντοπισμό στη βάση δεδομένων (βλ. Σχήμα 14)

Στο Σχήμα 13 περιγράφεται συνοπτικά η λειτουργία της αποθήκευσης νέου προσώπου στη βάση δεδομένων του συστήματος. Κατά τη διαδικασία, αρχικά επιλέγεται η αποθήκευση στη βάση δεδομένων κάποιων τυπικών προσωπικών δεδομένων και στη συνέχεια ενεργοποιείται το παράθυρο της εικόνας της κάμερας. Πραγματοποιείται αυτόματος εντοπισμός του προσώπου στην εικόνα και ακολουθεί εξομάλυνσή της για την αποθήκευση των eigenfaces στη βάση δεδομένων, μέσω του αλγόριθμου PCA. Στο Σχήμα 14 περιγράφεται συνοπτικά η λειτουργία της ταυτοποίησης προσώπου. Με την επιλογή της εισόδου στο σύστημα ενεργοποιείται η κάμερα και το παράθυρο εικόνας της κάμερας και αυτομάτως γίνεται εύρεση του προσώπου. Η ταυτοποίηση του προσώπου επιτυγχάνεται με τη εύρεση των βαρών και σύγκριση της απόστασής τους από το κατώφλι που έχει δοθεί.

Το σύστημα αναγνώρισης προσώπου συνολικά αποτελείται γενικά από τέσσερα υποσυστήματα, όπως φαίνεται στο Σχήμα 15: την *ανίχνευση*, την *ευθυγράμμιση*, την *εξαγωγή χαρακτηριστικών* και την *σύγκριση*, όπου ο εντοπισμός και η κανονικοποίηση των προσώπων, μέσω της ανίχνευσης και της ευθυγράμμισης, προηγούνται της αναγνώρισης, μέσω της εξαγωγής των χαρακτηριστικών και της σύγκρισης με τη βάση των εγγεγραμμένων προσώπων.

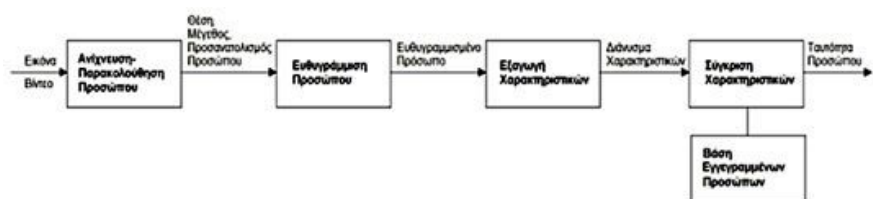


Σχήμα 13. Μπλοκ διάγραμμα πρώτου δομικού στοιχείου του συστήματος



Σχήμα 14. Μπλοκ διάγραμμα δεύτερου δομικού στοιχείου του συστήματος

Η *ανίχνευση προσώπου* (face detection) αναφέρεται στον έλεγχο αν υπάρχουν πρόσωπα σε μία εικόνα και στον προσδιορισμό της θέσης και των διαστάσεών τους. Ο *εντοπισμός προσώπου* (face localization) αναφέρεται ειδικότερα στον προσδιορισμό της θέσης και των διαστάσεων του προσώπου, στην περίπτωση που σε μία εικόνα υπάρχει σίγουρα ένα μόνο πρόσωπο. Στην περίπτωση του βίντεο τα πρόσωπα που ανιχνεύονται πρέπει να ανιχνεύονται διαρκώς σε κάθε πλαίσιο, με ένα σύστημα παρακολούθησης προσώπου (face tracking).



Σχήμα 15. Λειτουργικό διάγραμμα συστήματος αναγνώρισης προσώπου

Η *ευθυγράμμιση προσώπου* (face alignment/registration) στοχεύει στον εντοπισμό του προσώπου με μεγαλύτερη ακρίβεια, στον προσδιορισμό του μεγέθους και του προσανατολισμού του (πόζα), ενώ η ανίχνευση προσώπου πετυχαίνει απλά ένα χοντρικό προσδιορισμό της θέσης και του μεγέθους του προσώπου. Για την ευθυγράμμιση του προσώπου επιλέγονται χαρακτηριστικά σημεία αναφοράς ή ορόσημα του προσώπου (fiducial points/landmarks), από συγκεκριμένα σημεία ενδιαφέροντος (interest points) που εντοπίζονται στην εικόνα του προσώπου, όπως τα μάτια, τα φρύδια, η μύτη και το στόμα. Τα πρόσωπα κανονικοποιούνται σε σχέση με τις γεωμετρικές τους ιδιότητες, όπως το μέγεθος και ο προσανατολισμός, χρησιμοποιώντας γεωμετρικούς μετασχηματισμούς. Τέλος, μια περαιτέρω κανονικοποίηση μπορεί να εφαρμοστεί, σε σχέση με το φωτισμό και τις χρωματικές συνιστώσες της εικόνας του προσώπου, ώστε η εικόνα να περιέχει συγκεκριμένα επίπεδα διαβαθμίσεων του γκρι. Η *εξαγωγή χαρακτηριστικών* (feature extraction) έρχεται μετά τη γεωμετρική και φωτομετρική κανονικοποίηση για να προσφέρει την απαραίτητη πληροφορία, ώστε να μπορεί να γίνει διάκριση μεταξύ των προσώπων διαφορετικών ατόμων, ανεξάρτητα από τις γεωμετρικές και φωτομετρικές διαφοροποιήσεις που υπάρχουν στις εικόνες.

Κατά την *σύγκριση των προσώπων* (face matching), το εξαχθέν διάνυσμα χαρακτηριστικών του άγνωστου δοκιμαστικού προσώπου συγκρίνεται με τα εγγεγραμμένα πρόσωπα της βάσης και όταν βρεθεί κατάλληλο ταίρι με επαρκή αξιοπιστία, συνάγεται η ταυτότητα του άγνωστου προσώπου, διαφορετικά η ταυτότητά του παραμένει άγνωστη.

Τα αποτελέσματα της αναγνώρισης προσώπου εξαρτώνται ιδιαιτέρως από τα χαρακτηριστικά που αντιπροσωπεύουν το πρόσωπο και από τις μεθόδους ταξινόμησης που χρησιμοποιούνται για να τα διακρίνουν. Ωστόσο η διαδικασία εντοπισμού και κανονικοποίησης των προσώπων παίζει σημαντικό ρόλο στην εξαγωγή αποτελεσματικών χαρακτηριστικών για την αναγνώριση.

Πειράματα και συμπεράσματα

Για την καλύτερη κατανόηση της λειτουργίας αλλά και για την εκτίμηση της απόδοσης του συστήματος που αναπτύχθηκε πραγματοποιήθηκαν μια σειρά από πειράματα. Στις παραγράφους που ακολουθούν παρουσιάζονται παραδείγματα χρήσης του συστήματος καθώς και πειραματικά αποτελέσματα, ενώ τέλος, παρατίθενται σχόλια σχετικά με πλεονεκτήματα και μειονεκτήματα της μεθοδολογίας.

Συνολικά το σύστημα που υλοποιήθηκε εκτελεί τη διαδικασία αναγνώρισης σε δύο βήματα. Στο πρώτο βήμα γίνεται εντοπισμός προσώπου και στο δεύτερο αναγνώριση. Ο αλγόριθμος εντοπισμού προσώπου ακολουθεί τα εξής 5 στάδια:

1. Αποθήκευση εικόνας από την κάμερα



Σχήμα 16. Δείγμα εικόνας προσώπου από την κάμερα

2. Μετατροπή της εικόνας σε εικόνα γκριζων τόνων



Σχήμα 17. Δείγμα εικόνας προσώπου μετά από τη μετατροπή σε εικόνα γκριζών τόνων

3. Εντοπισμός προσώπου στην εικόνα γκριζών τόνων
4. Αποκοπή του καρέ για εμφάνιση μόνο της περιοχής του προσώπου
5. Εξομάλυνση και κανονικοποίηση της εικόνας προσώπου



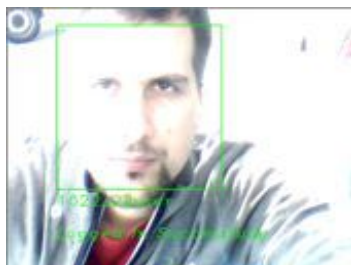
Σχήμα 18. Δείγμα εξομαλυνμένης και κανονικοποιημένης εικόνας προσώπου

Στη συνέχεια το σύστημα ακολουθεί δύο διαφορετικές διαδρομές ανάλογα με τη λειτουργία που εκτελείται. Μια διαδρομή αφορά στη διαδικασία εκμάθησης και δημιουργίας βάσης δεδομένων, ενώ η άλλη διαδρομή αφορά στη διαδικασία αναγνώρισης προσώπου σε σχέση με τα πρόσωπα που έχουν ήδη εκπαιδεύσει το σύστημα:

- Όταν το σύστημα είναι σε φάση εκμάθησης και δημιουργίας βάσης δεδομένων τότε απαιτείται η συγκέντρωση αρκετών φωτογραφιών υπό διαφορετική κλίση προσώπου και συνθηκών φωτισμού για κάθε πρόσωπο. Προφανώς σε κάθε περίπτωση αρχικά εφαρμόζονται τα πέντε στάδια του εντοπισμού όπως έχουν περιγραφεί παραπάνω. Στη συνέχεια εφαρμόζεται ο αλγόριθμος *eigenfaces* για τη δημιουργία εικόνων-«υπογραφής» για τα πρόσωπα που εισάγονται στη βάση δεδομένων και με τον τρόπο αυτό εκπαιδεύουν το σύστημα.
- Όταν το σύστημα καλείται να αναγνωρίσει ένα πρόσωπο που εμφανίζεται στην κάμερα τότε το πρόσωπο αυτό συγκρίνεται με τα πρόσωπα-υπογραφές της βάσης δεδομένων με βάση ένα μέτρο απόστασης και αναγνωρίζεται ως αυτό που αντιστοιχεί στη μικρότερη απόσταση.

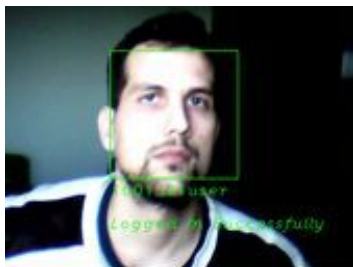
Για τον έλεγχο της αποδοτικότητας του συστήματος, δημιουργήθηκε μια βάση δεδομένων 50 χρηστών με 20 διαφορετικές πόζες ο καθένας. Η βάση αυτή αποτελείται συνολικά από 1000 φωτογραφίες. Από τα προκαταρκτικά πειράματα που πραγματοποιήθηκαν μετά από την ολοκλήρωση του συστήματος φάνηκε ότι η αποδοτικότητα του συστήματος εξαρτάται αποκλειστικά από τις συνθήκες φωτισμού. Έτσι, το τελικό σενάριο εκτίμησης της απόδοσης που σχεδιάστηκε αφορούσε την επιλογή 15 τυχαίων χρηστών (από τους 50 της βάσης δεδομένων) και την εξέταση της απόδοσης του συστήματος στην αναγνώρισή τους υπό την επίδραση 4 διαφορετικών συνθηκών φωτισμού:

- (1) εξωτερικός χώρος υπό το φως ημέρας και τον ήλιο



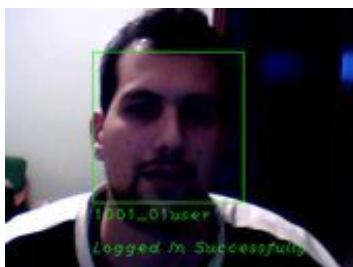
Σχήμα 19. Δείγμα φωτογραφίας κατά το σενάριο πειραματισμών 1

- (2) εσωτερικός χώρος υπό το φως ημέρας



Σχήμα 20. Δείγμα φωτογραφίας κατά το σενάριο πειραματισμών 2

- (3) εσωτερικός χώρος υπό το φως από λαμπτήρα πυρακτώσεως



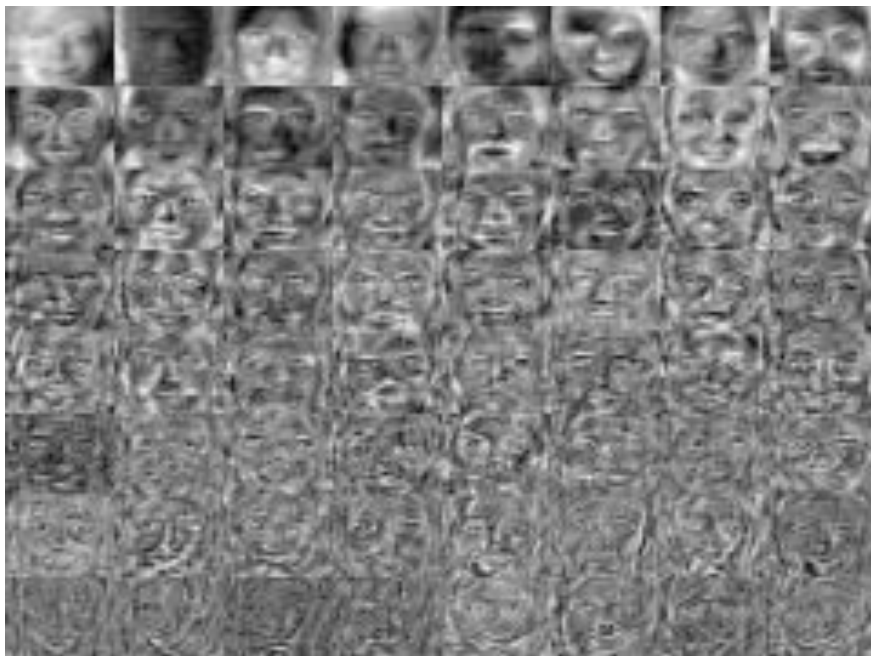
Σχήμα 21. Δείγμα φωτογραφίας κατά το σενάριο πειραματισμών 3

- (4) εσωτερικός χώρος υπό το φως δύο λαμπτήρων πυρακτώσεως



Σχήμα 22. Δείγμα φωτογραφίας κατά το σενάριο πειραματισμών 4

Στο Σχήμα 23 παρουσιάζεται μέρος της βάσης δεδομένων που δημιουργήθηκε με εστίαση στα πιο σημαντικά 64 eigenfaces.



Σχήμα 23. Δείγμα από τη βάση δεδομένων όπου εμφανίζονται τα πρώτα 64 eigenfaces

Στους πίνακες που ακολουθούν εμφανίζονται τα αποτελέσματα των πειραμάτων. Ο Πίνακας I παρουσιάζει συγκεντρωτικά και με αριθμητική κωδικοποίηση για τους χρήστες το χρήστη που αναγνωρίζεται από το σύστημα σε κάθε σενάριο πειραματισμών. Για παράδειγμα στη γραμμή με τίτλο «Σενάριο (1)» και στη στήλη «1» εμφανίζεται ο αριθμός 22. Αυτό σημαίνει ότι κατά το σενάριο (1) και ενώ παρουσιάστηκε στην κάμερα ο χρήστης 1, το σύστημα αναγνώρισε το χρήστη 22. Στην τελευταία στήλη παρουσιάζεται ποσοστιαία η απόδοση ορθής αναγνώρισης για το κάθε σενάριο. Παρατηρούμε ότι στις περιπτώσεις των σεναρίων (2) και (3) η επιτυχία είναι πλήρης (100%) ενώ στα σενάρια (1) και (4) κυρίως όταν

εμφανίζεται έντονος φωτισμός, η απόδοση ορθής αναγνώρισης ελαττώνεται.

Πίνακας Ι. Αναγνώριση χρηστών ανάλογα με τα σενάρια των πειραμάτων

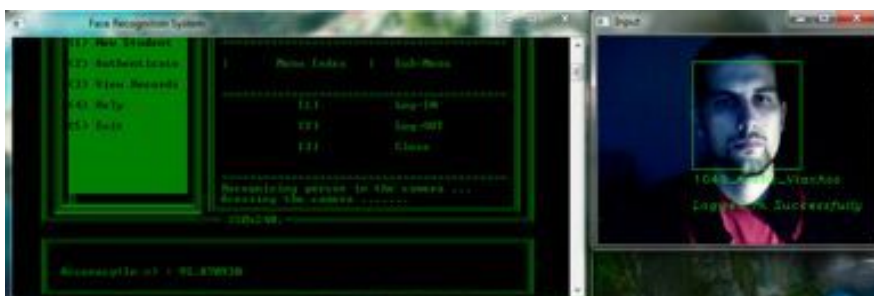
Χρήστης προς αναγνώριση	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Απόδοση (%)
Χρήστης που αναγνωρίζεται																
Σενάριο (1)	22	2	31	38	5	6	7	8	9	10	38	31	13	38	15	60,00
Σενάριο (2)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	100,00
Σενάριο (3)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	100,00
Σενάριο (4)	31	31	42	4	5	6	4	42	9	10	11	31	42	4	15	46,67

Ο Πίνακας ΙΙ παρουσιάζει αντίστοιχα τα ποσοστά βεβαιότητας στην αναγνώριση των χρηστών κατά τα σενάρια πειραμάτων που πραγματοποιήθηκαν. Φαίνεται και πάλι η υπεροχή του συστήματος στις περιπτώσεις των σεναρίων (2) και (3) όπου παρατηρείται υψηλότατο ποσοστό βεβαιότητας της αναγνώρισης. Αντίστοιχα τα αποτελέσματα δείχνουν ελάττωση του ποσοστού βεβαιότητας για τα σενάρια (1) και (4).

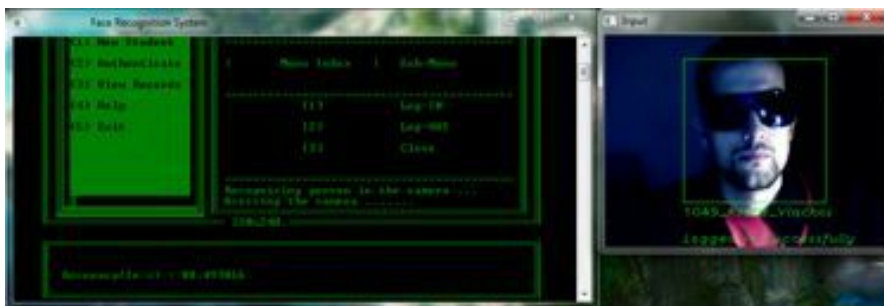
Πίνακας ΙΙ. Ποσοστό βεβαιότητας στην αναγνώριση των χρηστών των πειραμάτων

Σενάριο πειράματος	Ποσοστό βεβαιότητας αναγνώρισης (%)															Μέση τιμή (%)
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Σενάριο (1)	88,71	89,93	87,85	88,91	89,97	90,11	89,37	89,92	91,05	90,03	88,87	88,12	89,87	88,57	90,01	89,42
Σενάριο (2)	95,34	96,48	94,46	97,12	94,31	95,42	93,25	92,15	94,37	95,22	92,57	95,61	95,72	94,23	95,82	94,80
Σενάριο (3)	92,37	92,52	92,21	92,38	93,15	91,57	92,82	92,23	91,57	92,82	92,92	93,13	92,57	91,97	92,25	92,43
Σενάριο (4)	86,93	86,85	87,23	87,23	87,57	86,52	88,22	88,57	84,97	89,59	88,73	89,53	87,23	88,43	89,97	87,84

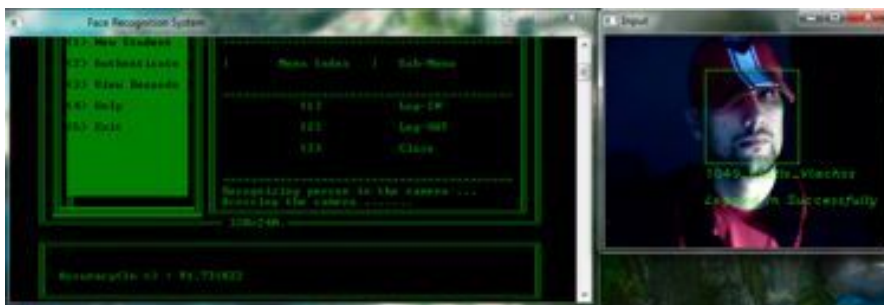
Παράλληλα με τα βασικά σενάρια πειραμάτων πραγματοποιήθηκε και σειρά επιπρόσθετων δοκιμών με προσπάθεια κάλυψης χαρακτηριστικών προσώπου για την εκτίμηση της αποδοτικότητας του συστήματος σε αντίστοιχες μη-ιδανικές συνθήκες φωτογράφισης. Στα Σχήματα που ακολουθούν (Σχήμα 24-Σχήμα 30) δίνονται μια σειρά από στιγμιότυπα από το λογισμικό που αναπτύχθηκε για τις διάφορες αυτές μη-ιδανικές περιπτώσεις αναγνώρισης προσώπου.



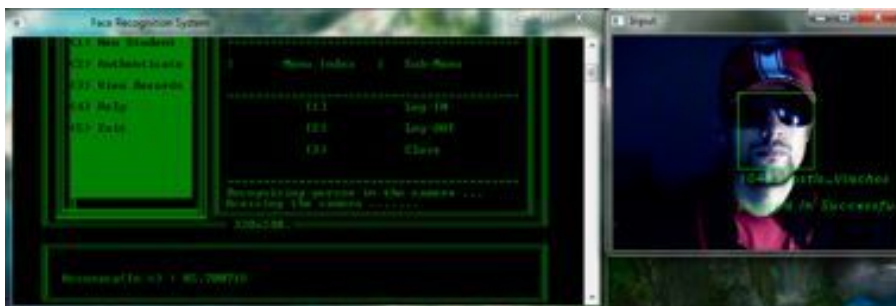
Σχήμα 24. Αναγνώριση τυπικού προσώπου



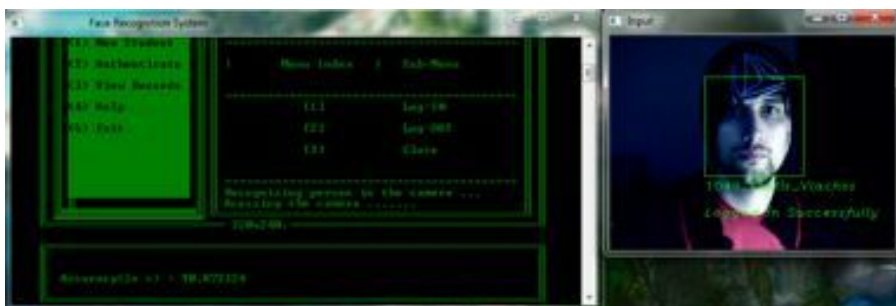
Σχήμα 25. Αναγνώριση προσώπου με γυαλιά ηλίου



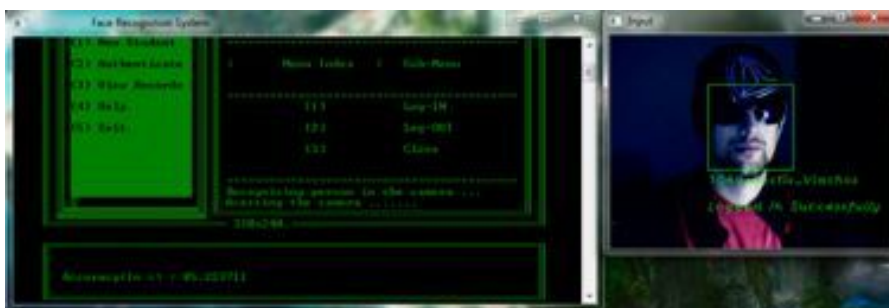
Σχήμα 26. Αναγνώριση προσώπου με καπέλο



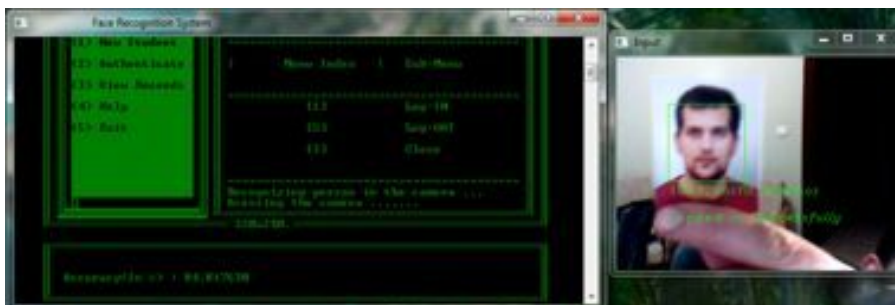
Σχήμα 27. Αναγνώριση προσώπου με καπέλο και γυαλιά ηλίου



Σχήμα 28. Αναγνώριση προσώπου με σκούφο

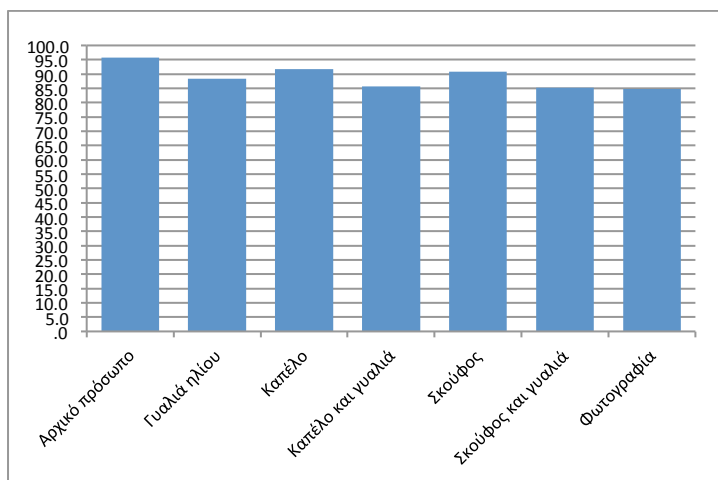


Σχήμα 29. Αναγνώριση προσώπου με σκούφο και γυαλιά ηλίου



Σχήμα 30. Αναγνώριση προσώπου από φωτογραφία

Σημειώνεται ότι στις περιπτώσεις που ο χρήστης φοράει καπέλο με γυαλιά ηλίου ή σκούφο με γυαλιά ηλίου, η αναγνώριση επιτυγχάνεται μόνο σε συγκεκριμένη απόσταση (κοντινή) από τη κάμερα. Σε μακρινή απόσταση ή ελαφρώς πλάγια δεν αναγνωρίζεται το πρόσωπο. Στο Σχήμα 31 φαίνεται ένα διάγραμμα αποδοτικότητας που επιτυγχάνεται για τις διάφορες περιπτώσεις εμφάνισης προσώπου.



Σχήμα 31. Διάγραμμα αποδοτικότητας

Από το Σχήμα 31 μπορούμε να διαπιστώσουμε πως τα καλύτερα ποσοστά αποδοτικότητας επιτυγχάνονται με καπέλο (91,53%) και με σκούφο (90,87%). Αυτό, βέβαια, ήταν εν μέρει αναμενόμενο καθώς η περιοχή του προσώπου που καλύπτεται είναι ψηλά στο κεφάλι, εκεί που τα χαρακτηριστικά του προσώπου είναι λιγότερα αξιόπιστα για την αναγνώριση. Όταν καλύπτονται χαρακτηριστικά όπως τα μάτια και η μύτη, η αποδοτικότητα πέφτει από επτά (7) έως δέκα (10) εκατοστιαίες μονάδες και βρίσκονται στο όριο της αναγνώρισης. Στην περίπτωση της *φωτογραφίας* έχουμε τη χαμηλότερη αποδοτικότητα (84,82%). Βέβαια στην περίπτωση αυτή εξακολουθεί να πραγματοποιείται αναγνώριση χρήστη και αυτό είναι ενδεχομένως ένα ζήτημα ασφαλείας για το σκοπό της συγκεκριμένης εφαρμογής.

Σκέψεις για βελτιώσεις και μελλοντικές εφαρμογές

Για τη βελτίωση της απόδοσης αναγνώρισης, υπάρχουν πολλά πράγματα που μπορούν να γίνουν, μερικά από τα οποία είναι αρκετά εύκολο να εφαρμοστούν. Παραδείγματος χάριν, θα μπορούσε να προστεθεί επεξεργασία χρώματος, ανίχνευση ακρών (edge detection) κλπ. Συνήθως μπορεί να βελτιωθεί η ακρίβεια αναγνώρισης προσώπου με τη χρησιμοποίηση περισσότερων εικόνων εισαγωγής, τουλάχιστο 50 ανά άτομο, με τη λήψη περισσότερων φωτογραφιών κάθε προσώπου, ιδιαίτερα από διαφορετικές γωνίες και φωτισμό. Υπάρχουν διάφορες απλές τεχνικές που θα μπορούσαν να χρησιμοποιηθούν για να ληφθούν περισσότερες εικόνες εκπαίδευσης, παράγοντας νέες εικόνες από τις ήδη υπάρχουσες:

- Δημιουργία αντιγράφων με καθρέφτισμα των εικόνων, έτσι ώστε να διπλασιαστούν οι εικόνες εκπαίδευσης
- Μετατόπιση ή περιστροφή των εικόνων για τη λήψη εναλλακτικών εικόνων για την εκπαίδευση
- Προσθήκη θορύβου εικόνας για βελτίωση της ανοχής στο θόρυβο

Είναι σημαντικό να υπάρχουν πολλές παραλλαγές των όρων για κάθε πρόσωπο, έτσι ώστε ο ταξινομητής να είναι σε θέση να αναγνωρίσει το πρόσωπο σε διαφορετικές συνθήκες έντασης και θέσης της πηγής φωτισμού.

Μια άλλη βελτίωση περιλαμβάνει τον τρόπο μέτρησης της απόστασης. Ο βασικός αλγόριθμος Eigenface χρησιμοποιεί την Ευκλείδεια απόσταση. Η εφαρμογή της απόστασης Mahalanobis αναφέρεται στη βιβλιογραφία ότι συνήθως δίνει καλύτερα αποτελέσματα. Αυτό εξηγείται από το γεγονός ότι όταν προβάλλεται μια εικόνα προσώπου επάνω στον υποχώρο του PCA παρατηρείται ένα ποσοστό «παραμόρφωσης», η οποία δεν είναι ίδια για κάθε άξονα του υποχώρου. Οι συντεταγμένες που αντιστοιχούν στις μεγαλύτερες ιδιοτιμές (eigenvalues) παραμορφώνονται πολύ περισσότερο

από τις συντεταγμένες που συνδέονται με μικρότερες ιδιοτιμές. Επειδή η Ευκλείδεια απόσταση αγνοεί αυτή την «ασυμμετρία», η χρησιμοποίησή της για τη μέτρηση της απόστασης δίνει την ίδια σημασία προς όλους του άξονες του υποχώρου. Επειδή η απόσταση Mahalanobis σχεδιάστηκε ώστε να λαμβάνει υπόψη αυτή την «ασυμμετρία» αναμένεται να δώσει καλύτερα αποτελέσματα αν εφαρμοστεί.

Νέες τεχνικές, όπως η τρισδιάστατη αναγνώριση προσώπου ή τα Μοντέλα Ενεργούς Εμφάνισης (Active Appearance Models) είναι δυνατό να δίνουν καλύτερα αποτελέσματα και σε κάθε περίπτωση προτείνεται η κατά αντιπαράσταση δοκιμή τους με την τεχνική των *eigenfaces* για την ανάδειξη της βέλτιστης λύσης ανάλογα με την εφαρμογή.

Με την εκπόνηση της πτυχιακής αυτής εργασίας έγινε φανερό στο συγγραφέα της ότι το πεδίο της Αναγνώρισης Προσώπων είναι ένα ευρύ πεδίο μιας άκρως ενδιαφέρουσας έρευνας που απαιτεί τη χρήση νέων, αλλά και δοκιμασμένων μεθόδων και τεχνικών από τα Γραφικά, την Επεξεργασία Εικόνας, την Τεχνητή Όραση και την Αναγνώριση Προτύπων.

Με το πέρας και την πιλοτική υλοποίηση του συστήματος ταυτοποίησης στο πλαίσιο της παρούσας εργασίας, κατέστη φανερό ότι μια πιθανή εφαρμογή των συστημάτων αναγνώρισης προσώπων (και συγκεκριμένα της υλοποίησης αυτής της πτυχιακής εργασίας) βρίσκεται στην τριτοβάθμια εκπαίδευση και στην ενσωμάτωση της στις Ηλεκτρονικές Γραμματείες των Σχολών, με στόχο την ταυτοποίηση προσώπων, στην πρόσβαση σε ηλεκτρονικές υπηρεσίες, την πρόσβαση σε εργαστηριακούς χώρους, την αυτόματη καταχώρηση παρουσιών, την ταυτοποίηση κατά την διεξαγωγή εξετάσεων, κοκ, αφαιρώντας σημαντικό γραφειοκρατικό και ελεγκτικό βάρος από το προσωπικό των Σχολών. Η ταχύτητα όσο και η ακρίβεια του συστήματος κρίνονται ικανοποιητικές για παρόμοιες εφαρμογές και θα μπορούσαν με μικρό κόστος να αποτελέσουν τμήμα της «ηλεκτρονικής διακυβέρνησης» των Σχολών τριτοβάθμιας εκπαίδευσης κατά το μέλλον.

Βιβλιογραφία

- [1] Wikipedia: Biometrics, <http://en.wikipedia.org/wiki/Biometrics>, 20 January 2013.
- [2] Joydip Dhar, “Intelligent Biometric System using PCA and R-LDA”, Indian Institute of IT and Management, Information Technology, 1 January 2009.
- [3] S. Z. Li and A. K. Jain (ed.). “Handbook of Face Recognition”, Springer 2005.
- [4] M.-H. Yang, D. Kriegman, and N. Ahuja, “Detecting Faces in Images: A Survey”. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002.
- [5] D. Blackburn, J. Bone, and P. Phillips, “Face Recognition Vendor Test 2000”, Evaluation Report, Tech. Rep., February 2001.
- [6] P. Phillips, P. Grother, R. Micheals, D. Blackburn, E. Tabassi, and J. Bone, “Face Recognition Vendor Test 2002”, Evaluation Report, Tech. Rep., March 2003.
- [7] John Peter Jesan, “The neural approach to pattern recognition”, Ubiquity archive, Volume 2004 Issue April, Pages 2-2, ACM New York, NY, USA.
- [8] National Programme on Technology Enhanced Learning: Remote Sensing, Information Extraction, http://nptel.iitm.ac.in/courses/Webcourse-contents/IIT-KANPUR/ModernSurveyingTech/lectureD_28/D_28_2.htm
- [9] H. A. Rowley, S. Baluja, and T. Kanade. “Neural Network-Based Face Detection”. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998.
- [10] K-K. Sung and T. Poggio. “Example-Based Learning for View-Based Human Face Detection”. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998.
- [11] Sirovich, L, Kirby, M., “A low-dimensional procedure for the characterization of human faces”, Journal of Optical Society American A, Vol. 4, No. 3, 1987, pp. 519-524.
- [12] M. Turk and A. Pentland, “Eigenfaces for recognition, Journal of Cognitive Neuroscience”, vol. 3, no. 1, 1991, pp. 71-86.
- [13] Shubhendu Trivedi, Face Recognition using Eigenfaces and Distance Classifiers: A Tutorial, <http://onionesquereality.wordpress.com/2009/02/11/face-recognition-using-eigenfaces-and-distance-classifiers-a-tutorial/>, February 11, 2009
- [14] Bloodshed Dev-C++, <http://www.bloodshed.net/devcpp.html>

- [15] OpenCV Library,
<http://www.intel.com/research/mrl/research/opencv>
- [16] G. Bradski, A. Kaehler and V. Pisarevsky, "Learning-Based Computer Vision with Intel's Open Source Computer Vision Library". In Intel Technology Journal, Vol. 09-Iss.02, May 2005.
- [17] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001.
- [18] P. Viola and M. Jones, "Robust Real-time Object Detection". In IEEE ICCV Workshop on Statistical and Computational Theories of Vision, 2001.
- [19] R. Lienhart and J. Maydt, "An Extended Set of Haar-like Features for Rapid Object Detection". IEEE ICIP 2002.
- [20] R. Lienhart, A. Kuranov, and V. Pisarevsky, "Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection". MRL Technical Report, Intel Labs, 2002.

Παράρτημα Α

Οδηγίες εγκατάστασης υποδομής υλοποίησης

Στο παρόν παράρτημα παρατίθενται οι οδηγίες εγκατάστασης όλης της απαιτούμενης υποδομής λογισμικών για την υλοποίηση του συστήματος αναγνώρισης προσώπων για περιβάλλον Microsoft Windows.

Για την ανάπτυξη και υλοποίηση της εφαρμογής «Simple Face Recognition System» είναι **αναγκαία** η λήψη της βιβλιοθήκης προγραμματισμού OpenCV (συγκεκριμένα η έκδοση 2.1 με όνομα αρχείου «OpenCV-2.1.0-win32-vs2008.exe») από την ιστοσελίδα ανάπτυξής της στη διεύθυνση <http://sourceforge.net/projects/opencvlibrary/files/opencv-win/2.1/> (συμπεριλαμβάνεται στο CD).



Σχήμα 32. Ιστοσελίδα λήψης της βιβλιοθήκης προγραμματισμού OpenCV

Μετά τη λήψη του αρχείου προχωράμε στην εγκατάστασή του, όπως εμφανίζεται στις παρακάτω εικόνες (Σχήμα 33).





Σχήμα 33. Βήματα εγκατάστασης της OpenCV

Ένα δεύτερο προαπαιτούμενο είναι η λήψη και εγκατάσταση προγράμματος για γλώσσα προγραμματισμού C++. Στην περίπτωση μας χρησιμοποιήθηκε η Dev-C++ (συγκεκριμένα η έκδοση 4.9.9.2 με όνομα αρχείου «devcpp-4.9.9.2_setup.exe») από την ιστοσελίδα <http://www.bloodshed.net/dev/devcpp.html> (συμπεριλαμβάνεται στο CD).



Σχήμα 34. Ιστοσελίδα λήψης της Dev-C++

Μετά τη λήψη του αρχείου είναι δυνατή η εγκατάστασή του, όπως εμφανίζεται στις παρακάτω εικόνες (Σχήμα 35).



Σχήμα 35. Βήματα εγκατάστασης του Dev-C++

Το τρίτο και τελευταίο βήμα είναι η σύνδεση της βιβλιοθήκης προγραμματισμού OpenCV με το πρόγραμμα Dev-C++, όπως φαίνεται στις παρακάτω εικόνες (Σχήμα 36-Σχήμα 41). Αρχικά θα πρέπει να προσθέσουμε ορισμένες κεφαλίδες (header files).



Σχήμα 36. Μενού «Εργαλεία» και επιλογές του μεταγλωττιστή



Σχήμα 37. Επιλογή του συμβόλου + για προσθήκη νέου μεταγλωττιστή με όνομα OpenCV



Σχήμα 38. Προσθήκη των συγκεκριμένων εντολών



Σχήμα 39. Προσθήκη φακέλου όπου βρίσκονται τα dll (dynamic library files)



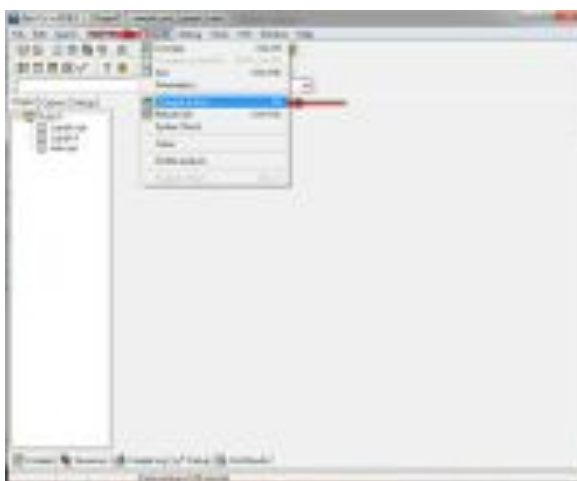
Σχήμα 40. Προσθήκη φακέλου επικεφαλίδων για συγγραφή κώδικα C



Σχήμα 41. Προσθήκη φακέλου επικεφαλίδων για συγγραφή κώδικα C++

Οδηγίες χρήσης της εφαρμογής από το CD

Αντιγράφουμε στην επιφάνεια εργασίας του Η/Υ από το CD, το φάκελο «Source Code». Ο φάκελος περιέχει δεκατρία (13) αρχεία. Τρέχουμε το «sample_pro_f_graph_h.exe».



Σχήμα 42. Εκτέλεση του αρχείου



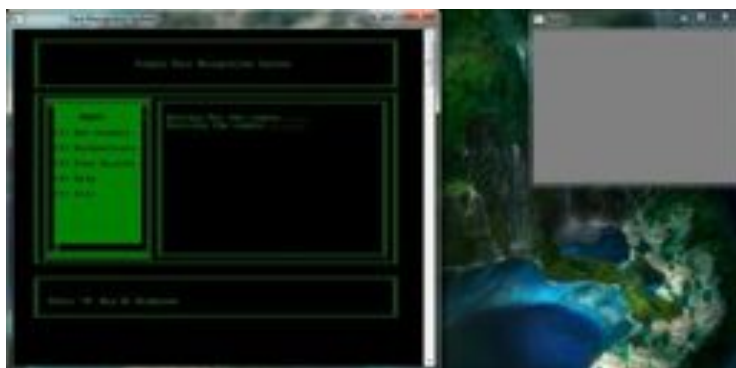
Σχήμα 43. Γραφικό περιβάλλον της εφαρμογής

Για την εισαγωγή δεδομένων στην εφαρμογή πρέπει να είναι επιλεγμένο το παράθυρο *DOS*. Για εγγραφή ενός νέου φοιτητή επιλέγουμε την επιλογή «<1> New Student» από το κύριο μενού.



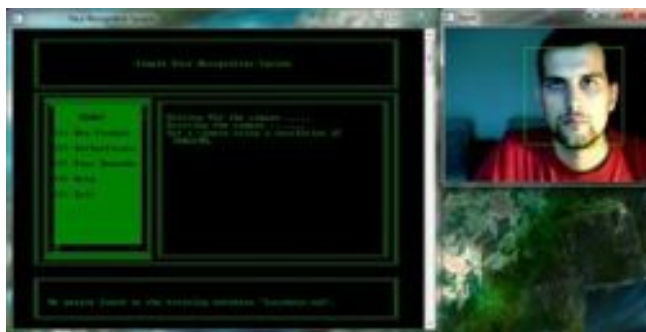
Σχήμα 44. Επιλογή εγγραφής νέου φοιτητή

Εμφανίζεται ένας διάλογος με τις τρεις (3) βασικές οδηγίες της εφαρμογής. Πληκτρολογούμε «1» για να συνεχίσουμε ή «0» για να ακυρώσουμε.

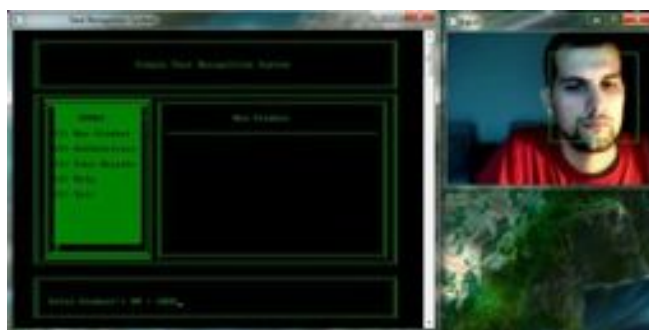


Σχήμα 45. Προετοιμασία για συλλογή εικόνων προσώπου

Πατάμε το πλήκτρο «N» όταν το άτομο είναι έτοιμο για εκπαίδευση.



Σχήμα 46. Αναγνώριση προσώπου (εάν υπάρχει στη βάση δεδομένων)



Σχήμα 47. Εισαγωγή προσωπικών στοιχείων φοιτητή

Εάν δε πατήσουμε το πλήκτρο «N», στην κάμερα εντοπίζεται το πρόσωπο του άγνωστου και εμφανίζει ότι δε βρέθηκε στη βάση δεδομένων (*No people found in the training database 'facedata.xml'*).

Εισάγουμε τα προσωπικά δεδομένα του φοιτητή (αριθμός μητρώου, ονοματεπώνυμο (χωρίς κενό), τμήμα, κωδικός πρόσβασης) και πατάμε το πλήκτρο «y», εάν θέλουμε να τα σώσουμε ή «n» να τα ακυρώσουμε.



Σχήμα 48. Επιλογή αποθήκευσης προσωπικών δεδομένων

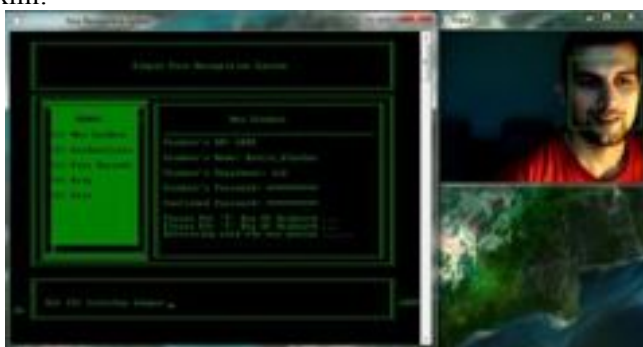
Από τη στιγμή που θα συναινέσουμε στην αποθήκευση των προσωπικών δεδομένων, η εφαρμογή αυτόματα θα αποθηκεύει το πρόσωπο που φαίνεται στην webcam (processed frontal face). Όσες περισσότερες εικόνες αποθηκεύσει, τόσο το καλύτερο για την αποδοτικότητα της

αναγνώρισης. Ο χρήστης μπορεί να κουνήσει το κεφάλι του και να φορέσει γυαλιά, κασκόλ κτλ. Οι εικόνες αποθηκεύονται ως αρχεία PGM στο φάκελο data και τα ονόματα τους φαίνονται στο αρχείο κειμένου train.txt.



Σχήμα 49. Αποθήκευση προσώπου στη βάση δεδομένων

Μετά την επιθυμητή αποθήκευση προσώπων, πατάμε το πλήκτρο «T» για να ξεκινήσει η εκπαίδευση. Η εφαρμογή θα “κολλήσει” για ένα διάστημα 5 με 30 δευτερολέπτων (εξαρτάται από τον αριθμό εικόνων και την υπολογιστική ισχύ) και θα συνεχίσει από τη στιγμή που θα έχει εκπαιδευτεί το έξτρα πρόσωπο. Το αρχείο της βάσης δεδομένων είναι το facedata.xml.



Σχήμα 50. Ολοκλήρωση αποθήκευσης προσώπων

Με την ολοκλήρωση της αποθήκευσης των προσώπων βγαίνει μήνυμα με τον αριθμό μητρώου, το ονοματεπώνυμο του φοιτητή και η αυτοπεποίθηση (confidence) της αναγνώρισης. Να σημειωθεί πως η αυτοπεποίθηση κυμαίνεται στο πεδίο [0.0, 1.0]. Είναι άμεσα συνδεδεμένη με τη Ευκλείδεια Απόσταση. Εάν βρίσκεται ανάμεσα στο πεδίο [0.0, 0.5] τότε οι εικόνες είναι διαφορετικές μεταξύ τους, ενώ στο πεδίο [0.5, 1.0] οι εικόνες είναι όμοιες μεταξύ τους.

Πατώντας το πλήκτρο «Esc» μας δίνεται η δυνατότητα του αρχικού μενού. Επιλέγοντας το «<3> View Records» βλέπουμε συγκεκριμένες λεπτομέρειες από την εκπαίδευση, όπως τον αριθμό ιδιοδιανυσμάτων, τις

εικόνες προς εκπαίδευση, την ανάλυση της κάμερας και το χρόνο που χρειάστηκε για να γίνει ο εντοπισμός προσώπου.



Σχήμα 51. Επισκόπηση λεπτομερειών της εκπαίδευσης

Επιλέγοντας το «<2> Authenticate» μας δίνεται η δυνατότητα της σύνδεσης και αποσύνδεσης του χρήστη με την αναγνώριση του προσώπου του.



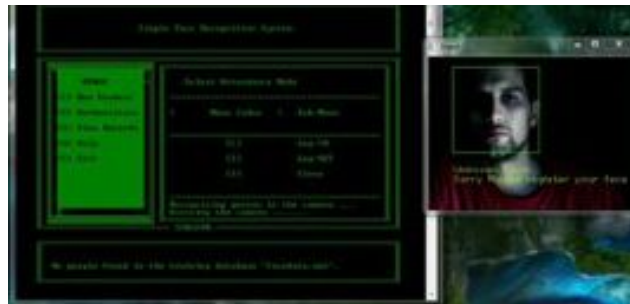
Σχήμα 52. Σύνδεση και αποσύνδεση χρήστη



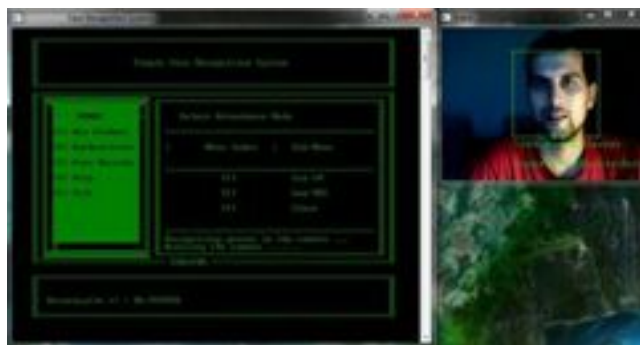
Σχήμα 53. Σύνδεση φοιτητή

Εμφανίζεται στο παράθυρο της webcam ο αριθμός μητρώου, το ονοματεπώνυμο του φοιτητή και το γεγονός ότι έχει συνδεθεί. Εάν κάποιος φοιτητής δεν υπάρχει στη βάση δεδομένων βγαίνει το μήνυμα «Unknown Face. Sorry please register your face». Να σημειωθεί πως

κατά τη σύνδεση και αποσύνδεση (αναγνώριση χρήστη) εμφανίζεται ο όρος Accuracy (ακρίβεια σε %). Θεωρείται πως με πάνω από 80% υπάρχει σωστή αναγνώριση του προσώπου.



Σχήμα 54. Άγνωστο πρόσωπο



Σχήμα 55. Αποσύνδεση φοιτητή

Η επιλογή «<4> Help» εμφανίζει τον συγγραφέα της εφαρμογής. Τέλος η επιλογή «<5> Exit» κλείνει την εφαρμογή.



Σχήμα 56. Στοιχεία συγγραφέα της εφαρμογής και έξοδος από την εφαρμογή

[Αυτή η σελίδα είναι εσκεμμένα λευκή]

[Αυτή η σελίδα είναι εσκεμμένα λευκή]

ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΣΕΡΡΩΝ
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ & ΕΠΙΚΟΙΝΩΝΙΩΝ

ΣΥΣΤΗΜΑ ΑΝΑΓΝΩΡΙΣΗΣ ΠΡΟΣΩΠΩΝ
ΜΕ ΧΡΗΣΗ WEB CAMERA ΚΑΙ OPENCV
ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Κωνσταντίνος Βλάχος, 2013

Οι εφαρμογές που απαιτούν την εξακρίβωση της ταυτότητας ενός ατόμου πολλαπλασιάζονται συνεχώς και μαζί με αυτές αυξάνεται η ανάγκη για τη χρήση μεθόδων πιστοποίησης που δεν θα απαιτούν από τους ανθρώπους να θυμούνται πολύπλοκους κωδικούς πρόσβασης ή να μεταφέρουν μαγνητικές κάρτες, ταυτότητες ή άλλου είδους έγγραφα που θα επαληθεύουν την ταυτότητα τους. Η χρήση βιομετρικών μεθόδων πιστοποίησης, δηλαδή μεθόδων που στηρίζονται στην αναγνώριση φυσικών χαρακτηριστικών ή χαρακτηριστικών της συμπεριφοράς του ανθρώπου είναι μια λύση σε αυτό το πρόβλημα. Μεταξύ των διαφόρων βιομετρικών μεθόδων που χρησιμοποιούνται, όπως η αναγνώριση ίριδας, των δακτυλικών αποτυπωμάτων, της γεωμετρίας της παλάμης, της στάσης του σώματος και του βαδίσματος, της φωνής, ακόμη και του τρόπου πληκτρολόγησης ή υπογραφής, η αναγνώριση προσώπου είναι σίγουρα μία από τις φιλικότερες προς το χρήστη, γιατί ουσιαστικά μιμείται τον τρόπο που χρησιμοποιούν οι άνθρωποι για να αναγνωρίζουν ο ένας τον άλλο.

TECHNOLOGICAL EDUCATIONAL INSTITUTION OF SERRES
SCHOOL OF TECHNOLOGICAL APPLICATIONS
DEPARTMENT OF INFORMATICS AND TELECOMMUNICATIONS

FACE RECOGNITION USING
A WEB CAMERA AND OPENCV
THESIS

Konstantinos Vlachos, 2013

In recent years, more and more applications require the identification of a person along with the need for using methods of certification that will not require remembering complicated passwords or carrying magnetic cards, identification cards or other identification documents. The use of biometric methods of identification based on natural features or characteristics of human behavior is one solution to this problem. Among the various biometric methods already being used, such as the recognition of iris, fingerprint, hand geometry, posture and gait, voice, or even the way of typing or signing, face recognition is definitely one of the most user friendly as it mimics the natural way people recognize each other.