



**Εξώφυλλο Αναφοράς Πτυχιακής Εργασίας  
Τεχνολογικό Εκπαιδευτικό Ίδρυμα Κρήτης**

**Σχολή Τεχνολογικών Εφαρμογών**

**Τμήμα Εφαρμοσμένης Πληροφορικής & Πολυμέσων**



**Πτυχιακή Εργασία**

**Τίτλος: Αυτόματη Οπτική Ταυτοποίηση Προσώπου**

**Καραταράκη Αικατερίνη ΑΜ: 1836**

**Καραταράκη Μαρία ΑΜ: 2127**

**Επιβλέπων Καθηγητής: Τριανταφυλλίδης Γιώργος**

**Επιτροπή Αξιολόγησης:**

**Ημερομηνία Παρουσίασης:**

## Ευχαριστίες

Θα θέλαμε να εκφράσουμε τις θερμές μας ευχαριστίες στον επιβλέποντα καθηγητή μας κ. Τριανταφυλλίδη Γιώργο για τη πολύτιμη βοήθεια και υποστήριξη που μας πρόσφερε κατά τη διάρκεια εκπόνησης της πτυχιακής μας εργασίας ,καθώς επίσης για το έναυσμα που μας έδωσε για να ασχοληθούμε με την αναγνώριση προσώπου (face recognition).

Η άριστη συνεργασία, η καθοδήγηση και η κατανόηση που μας επέδειξε συνέβαλαν στην ολοκλήρωση της παρούσας πτυχιακής. Επιπρόσθετα ,τις θερμότερες ευχαριστίες μας οφείλουμε να εκφράσουμε στα μέλη της οικογένεια μας για την ψυχολογική στήριξη που μας παρείχαν.

## Περιεχόμενα

Εξώφυλλο Αναφοράς Πτυχιακής Εργασίας.....	1
Ευχαριστίες .....	2
Abstract .....	8
Σύνοψη.....	9
1. Εισαγωγή .....	10
1.1 Περίληψη .....	11
1.2 Κίνητρο Για την Διεξαγωγή της Εργασίας .....	11
1.3 Σκοπός και Στόχοι Εργασίας .....	11
1.4 Δομή Εργασίας.....	11
2. Βιομετρική Αναγνώριση .....	12
2.1 Εισαγωγή στην Βιομετρική Αναγνώριση .....	12
2.1.1 Τι είναι η βιομετρική αναγνώριση;.....	12
2.1.2 Περιγραφή και δυνατότητες ενός βιομετρικού συστήματος.....	12
2.1.3 Το πρόβλημα της αναγνώρισης προσώπων .....	14
2.1.4 Εφαρμογές της αναγνώρισης προσώπων .....	16
2.1.5 Προβλήματα στατικής ταύτισης .....	18
2.1.5 Προβλήματα Δυναμικής Ταύτισης .....	20
2.1.6 Τα βασικά υποπροβλήματα στην αναγνώριση προσώπων .....	20
3. Σχέδιο δράσης για την εκπόνηση της εργασίας .....	22
Ταυτότητα και Ταυτοποίηση-Στάδια Αυτόματης Ταυτοποίησης.....	22
3.1 Η έννοια της ταυτότητας.....	22
3.2 Η έννοια της Ταυτοποίησης.....	23
3.2.1 Η χρησιμότητα της ταυτοποίησης.....	24
3.2.2 Οι κίνδυνοι της ταυτοποίησης .....	25
3.2.3 Διαδικασίες Ταυτοποίησης .....	26
3.2.4 Χαρακτηριστικά προς Ταυτοποίηση.....	26
3.3 Ιστορική Ανασκόπηση Βιομετρικής Ταυτοποίησης και Επιβεβαίωσης Ταυτότητας .....	27
3.4 Στάδια Αυτόματης Αναγνώρισης Προσώπων.....	28
3.4.1 Αναπαράσταση .....	30
3.4.2 Αναπαράσταση προσώπων - Εξαγωγή χαρακτηριστικών .....	30
3.4.3 Αναπαράσταση προσώπων με χρήση των τιμών της κλίμακας του γκρι.....	33
3.4.4 Γεωμετρικά χαρακτηριστικά από εικόνες εμπρόσθιας όψης και προφίλ.....	36
4. Κεφάλαιο -Κύριο Μέρος Πτυχιακής .....	46
4.1 Κατηγοριοποίηση Αλγορίθμων.....	46
4.1.1 Αλγόριθμος αναγνώρισης προσώπου .....	46

4.1.2 Ταξινόμηση αλγορίθμων.....	46
4.1.3 Αλγόριθμοι που βασίζονται σε μοντέλα (Model - based Algorithms).....	47
4.1.4 Αλγόριθμοι που βασίζονται στην εμφάνιση (Appearance – Based Algorithms).....	49
4.1.5 Αποστάσεις και συναρτήσεις ομοιότητας.....	50
4.1.6 Ανάλυση σε Κύριες Συνιστώσες (Principal Component Analysis, PCA) .....	51
4.1.7 Independent Component Analysis (ICA).....	58
4.1.8 Γραμμική Διακρίνουσα Ανάλυση (Linear Discriminant Analysis, LDA) .....	60
4.1.9 Αλγόριθμοι πυρήνα (Kernel Algorithms) .....	65
4.1.10 Ανάλυση σε Κύριες Συνιστώσες πυρήνα (Kernel PCA) .....	66
4.1.10 Γραμμική Διακρίνουσα Ανάλυση του πυρήνα (Kernel Fisher LDA).....	69
4.2 Υλοποίηση Αλγορίθμων .....	72
4.2.1 Εισαγωγή.....	72
4.2.2 Βάση ORL Face .....	72
4.2.3 Υλοποίηση και εφαρμογή αλγορίθμων.....	72
4.2.4 Συγκεντρωτικά Αποτελέσματα .....	85
5. Πειραματικό Μέρος.....	86
5.1 Η ιστορία του Matlab.....	86
5.1.1 Τι είναι μια γραφική διεπαφή χρήστη .....	87
5.2 Αναγνώριση Προσώπου Με τη Χρήση Matlab .....	87
5.2.1 Υλοποίηση του PCA.....	87
5.2.2 Πειραματικά Αποτελέσματα .....	96
Βιβλιογραφία .....	106

Εικόνα 1: Κοινά βιομετρικά χαρακτηριστικά.....	12
Εικόνα 2: Χαρακτηριστικό διάγραμμα ενός βιομετρικού συστήματος.....	13
Εικόνα 3: Ομοζυγωτικά δίδυμα.....	14
Εικόνα 4: Εκφράσεις προσώπων.....	14
Εικόνα 5: (α) Mugs hot φωτογραφία (front view) (β) Mug shot φωτογραφία(profile) (γ) (δ) Φωτογραφίες σε διπλώματα οδήγησης, διαβατήρια κλπ.....	19
Εικόνα 6: (α)(β)Φωτογραφίες με ανομοιόμορφο φόντο.....	19
Εικόνα 7: Η ταυτότητα και οι διαφορετικές υπό-ταυτότητες ανά εφαρμογή.....	23
Εικόνα 8: Χαρακτηριστικά προς ταυτοποίηση.....	27
Εικόνα 10: Ολοκληρωματικές προβολές.....	32
Εικόνα 12: Αναπαράσταση προσώπου ως άθροισμα ιδιοδιανυσμάτων.....	33
Εικόνα 13: Σημεία σε προφίλ και frontal views.....	36
Εικόνα 14: Εικόνες εκπαίδευσης αλγορίθμου.....	53
Εικόνα 15: Εικόνα προς εξέταση.....	56
Εικόνα 16: Face space.....	57
Εικόνα 17: Σημεία που διαχωρίζουν πλήρως σε δυο κλάσεις κατά την προβολή τους σε ευθεία.....	60
Εικόνα 18: Σημεία που συμπίπτουν χωρίς να είναι εφικτός ο ξεκάθαρος διαχωρισμός τους σε.....	60
Εικόνα 19: (α) Καλός διαχωρισμός κλάσεων. (β) Κακός διαχωρισμός κλάσεων.....	61
Εικόνα 20: Αρχή λειτουργίας του αλγόριθμου γραμμικής διακρίνουσας ανάλυσης (LDA).....	62
Εικόνα 21: (α)(β) Τα πέντε πρώτα διανύσματα βάσης του r-LDA.....	64
Εικόνα 22: Ένα παράδειγμα κατηγοριοποίησης προτύπων σε δύο κλάσεις [76]. Στα αριστερά: δείγματα στον δισδιάστατο χώρο εισόδου, όπου χρειάζεται ένα μη γραμμικό ελλειψοειδές όριο απόφασης (decision boundary) για το διαχωρισμό των κλάσεων A και B. Στα δεξιά: δείγματα που προβάλλονται στον τρισδιάστατο χώρο χαρακτηριστικών όπου ένα γραμμικό υπερεπίπεδο μπορεί να διαχωρίσει τις κλάσεις.....	66
Εικόνα 23: Παράδειγμα κλάσης της ORL.....	72
Εικόνα 24: Απόδοση ταυτοποίησης PCA με χρήση Mahalanobis.....	74
Εικόνα 25: Απόδοση ταυτοποίησης PCA με χρήση Ευκλείδειας.....	75
Εικόνα 26: Απόδοση ταυτοποίησης PCA με χρήση Cityblock.....	75
Εικόνα 27: Απόδοση ταυτοποίησης PCA με χρήση Cosine.....	76
Εικόνα 28: Απόδοση ταυτοποίησης LDA με χρήση Mahalanobis.....	77
Εικόνα 29: Απόδοση ταυτοποίησης LDA με χρήση Ευκλείδειας.....	77
Εικόνα 30: Απόδοση ταυτοποίησης LDA με χρήση Cityblock.....	78
Εικόνα 31: Απόδοση ταυτοποίησης LDA με χρήση Cosine.....	78
Εικόνα 32: Απόδοση ταυτοποίησης KPCA με χρήση Mahalanobis.....	79
Εικόνα 33: Απόδοση ταυτοποίησης KPCA με χρήση Ευκλείδειας.....	80
Εικόνα 34: Απόδοση ταυτοποίησης KPCA με χρήση Cityblock.....	81
Εικόνα 35: Απόδοση ταυτοποίησης KPCA με χρήση Cosine.....	81
Εικόνα 36: Απόδοση ταυτοποίησης KFA με χρήση Mahalanobis.....	82
Εικόνα 37: Απόδοση ταυτοποίησης KFA με χρήση Ευκλείδειας.....	83
Εικόνα 38: Απόδοση ταυτοποίησης KFA με χρήση CityBlock.....	84
Εικόνα 39: Απόδοση ταυτοποίησης KFA με χρήση Cosine.....	84
Εικόνα 40: Διάγραμμα Λειτουργίας του PCA.....	87
Εικόνα 41: Εκκίνηση διεπαφής.....	88
Εικόνα 42: Επιλογή Εικόνας μέσα από φάκελο.....	89
Εικόνα 43: Κουμπί Στιγμιότυπο Κάμερας.....	90
Εικόνα 44: Επιλογή εικόνας μέσω κάμερας.....	90
Εικόνα 45: Αποθήκευση Εικόνας μέσω Κάμερας.....	91

Εικόνα 46: Επιλογή Εικόνας.....	91
Εικόνα 47: Επιλογή εικόνας από φάκελο .....	92
Εικόνα 48: Παράθυρο επιλογής ή προσθήκης χρήστη .....	92
Εικόνα 49: Καταχώρηση γνωρισμάτων .....	93
Εικόνα 50: Προσθήκη εικόνας στη βάση .....	94
Εικόνα 51: (α)Αναγνώριση προσώπου (β)Στοιχεία ατόμου .....	94
Εικόνα 52: Προσθήκη περισσότερων από μια φωτογραφιών στη βάση .....	95
Εικόνα 53: Πληροφορίες Βάσης.....	95
Εικόνα 54: Ερώτηση επιβεβαίωσης διαγραφής .....	95
Εικόνα 55: Επιβεβαίωση διαγραφής.....	96
Εικόνα 56: Πρόσωπα Βάσης.....	96
Εικόνα 57: Εικόνα με πλάγια όψη .....	97
Εικόνα 58: Εικόνα Προσώπου με γυαλιά .....	98
Εικόνα 59: Λανθασμένη αναγνώριση σε ανομοιόμορφο φόντο .....	99
Εικόνα 60: Σωστή αναγνώριση σε ανομοιόμορφο φόντο.....	100
Εικόνα 61: Σωστή αναγνώριση με χρήση καπέλου .....	101
Εικόνα 62: Σωστή Ταυτοποίηση με αλλαγή φωτισμού .....	102
Εικόνα 63: Σωστή ταυτοποίηση με κλίση του κεφαλιού σε συνδυασμό με έκφραση προσώπου .....	103

Πίνακας 1: Συνήθειες εφαρμογές της αναγνώρισης προσώπων.....	18
Πίνακας 2: Δεδομένα που χρησιμοποιήθηκαν σε διάφορες μελέτες.....	30
Πίνακας 3: Χαρακτηριστικά που χρησιμοποιήθηκαν για τη διαδικασία σε μελέτες αναγνώρισης προσώπων. $D(x,y)$ = Ευκλείδεια απόσταση των χαρακτηριστικών $x$ και $y$ .....	31
Πίνακας 4: Αναγνώριση προσώπων σε διάφορες αναλύσεις και με χρήση 5 διαφορετικών μετρικών	34
Πίνακας 5: Αναγνώριση προσώπων σε διάφορες αναλύσεις και με χρήση 5 διαφορετικών μετρικών. Επιτυχής θεωρείται η αναγνώριση αν η προς έλεγχο φωτογραφία και κάποια από τις τρεις πλησιέστερες προς αυτήν αποθηκευμένες ανήκουν στο ίδιο πρόσωπο.....	35
Πίνακας 6: Αναγνώριση προσώπων σε διάφορες αναλύσεις και με χρήση 5 διαφορετικών μετρικών. Επιτυχής θεωρείται η αναγνώριση αν η προς έλεγχο φωτογραφία και κάποια από τις πέντε πλησιέστερες προς αυτήν αποθηκευμένες ανήκουν στο ίδιο πρόσωπο.....	35
Πίνακας 7: Χαρακτηριστικά για την περιγραφή του προσώπου (βλέπε Εικόνα 10). Η $d(i,j)$ είναι η Ευκλείδεια απόσταση ανάμεσα στα σημεία $i$ και $j$ .....	36
Πίνακας 8: Πίνακας αλγορίθμων.....	73
Πίνακας 9: Ποσοστά αναγνώρισης PCA.....	76
Πίνακας 10: Ποσοστά αναγνώρισης LDA.....	79
Πίνακας 11: Ποσοστά αναγνώρισης kernel PCA.....	82
Πίνακας 12: Ποσοστά αναγνώρισης KFA.....	85
Πίνακας 13: Συγκεντρωτικός πίνακας αποστάσεων.....	85

## Abstract

One of the rapidly developing technologies in the field of security is undoubtedly biometrics, which is an advanced method of verifying the identity of an individual based on certain physical characteristics that are unique to each individual, such as fingerprint, palm, iris of the eye etc. Biometrics is already used by state services in official documents, such as passports. Those subject to biometric control, in order to prove their identity, they should make an action, for example, to touch their finger to a "reader" fingerprint or look at an iris recognition device, which has high success in the degree of identification of the individual, it also creates a psychological "inconvenience".

For this reason biometrics is gaining ground and is incorporated in many applications with face recognition. A combination of systems, based primarily on cameras and computer programs, can identify an individual without having the subject to do any action. These systems scan through a camera the geometry of the face composed of a number of parameters such as the distance between the eye, the size of the mouth, the face shape and many others. These elements combined together create a unique model for each person. Recognizing persons using biometrics can be considered more functional and easy to use, but are much more complex and technologically difficult, because it is based on the analysis of a combination of characteristics of an individual. Today, however, "smart" programs have been developed that analyze facial images from cameras and can reliably identify an individual. Facial recognition can be used in many applications, with the main role of controlling physical access to various rooms in which only certain people are allowed to enter. Also, it is very likely in the future for security reasons that facial recognition systems will evolve to control who is or is not licensed to operate a computer or even who will be able to drive a car. Furthermore, facial recognition with modern technological methods can be used by security services to search for suspicious people in public places, for example in airports.

The aim of the thesis is to explore optical face recognition and will investigate the success rates of various algorithms, and propose an implementation. The method implemented in this thesis is the method of principal components analysis (PCA).

Algorithms that were evaluated for face recognition:

1. Principal Component Analysis (Principal Component Analysis, PCA).
2. Linear Discriminant analysis (Linear Discriminant Analysis, LDA).
3. Kernel Algorithms
4. Principal Component Analysis kernel (Kernel PCA)
5. Principal Discriminant Analysis of the core (Kernel Fisher LDA)



## Σύνοψη

Μία από τις τάχιστα αναπτυσσόμενες τεχνολογίες στον τομέα της ασφάλειας είναι αναμφισβήτητη η βιομετρία, η οποία αποτελεί μια προηγμένη μέθοδο πιστοποίησης της ταυτότητας ενός ατόμου με βάση κάποια φυσικά χαρακτηριστικά που είναι μοναδικά σε κάθε άνθρωπο, όπως το δακτυλικό αποτύπωμα, η παλάμη, η ίριδα του ματιού κ.ά. Ήδη η βιομετρία αξιοποιείται και σε επίσημα έγγραφα κρατικών υπηρεσιών, όπως στα διαβατήρια. Όσοι υπόκεινται στη διαδικασία έλεγχου των φυσικών χαρακτηριστικών τους, προκειμένου να αποδειχτεί η ταυτότητά τους, πρέπει να κάνουν μία ενέργεια, για παράδειγμα να ακουμπήσουν το δάχτυλό τους σε έναν «αναγνώστη» αποτυπωμάτων ή να κοιτάξουν σε μια συσκευή αναγνώρισης της ίριδας, κάτι που μπορεί να εξασφαλίζει σε μεγάλο βαθμό την ταυτοποίηση του ατόμου, δημιουργεί όμως και μια ψυχολογική «ενόχληση».

Για τον λόγο αυτό η βιομετρική μέθοδος που φαίνεται να κερδίζει έδαφος και να διεισδύει σε πολλές εφαρμογές είναι η αναγνώριση προσώπου. Ένας συνδυασμός συστημάτων, που βασίζεται κατά κύριο λόγο σε κάμερες και σε προγράμματα υπολογιστών, μπορεί να προβεί στην ταυτοποίηση ενός ατόμου χωρίς αυτό να χρειάζεται να κάνει κάποια ενέργεια. Τα συστήματα αυτά σκανάρουν μέσω μίας κάμερας τη γεωμετρία του προσώπου που αποτελείται από μια σειρά παραμέτρων όπως η απόσταση μεταξύ των ματιών, το μέγεθος του στόματος, το σχήμα του προσώπου και πολλά άλλα. Τα στοιχεία αυτά συνδυαζόμενα μεταξύ τους δημιουργούν ένα πρότυπο που για κάθε άνθρωπο είναι μοναδικό. Η μέθοδος αναγνώρισης του προσώπου μπορεί να θεωρείται πιο λειτουργική και εύχρηστη, είναι όμως αρκετά πιο πολύπλοκη και δύσκολη τεχνολογικά, γιατί βασίζεται στην ανάλυση ενός συνδυασμού χαρακτηριστικών του ανθρώπου. Σήμερα όμως έχουν αναπτυχθεί «έξυπνα» προγράμματα, τα οποία αναλύουν τις εικόνες του προσώπου από τις κάμερες και μπορούν να εξάγουν ασφαλή συμπεράσματα για την ταυτοποίηση ενός ατόμου. Η αναγνώριση προσώπου μπορεί να αξιοποιηθεί σε πολλές εφαρμογές, με κύρια τον έλεγχο φυσικής πρόσβασης σε διάφορους χώρους, στους οποίους μόνο συγκεκριμένα πρόσωπα έχουν άδεια να εισέρχονται. Επίσης, είναι πολύ πιθανό στο μέλλον για λόγους ασφαλείας να αναπτυχθούν συστήματα αναγνώρισης προσώπου που θα ελέγχουν το ποιος θα έχει ή όχι άδεια να χειρίζεται έναν υπολογιστή ή ακόμη και το ποιος θα μπορεί να οδηγήει ή όχι ένα αυτοκίνητο. Παράλληλα, η αναγνώριση προσώπου με σύγχρονες τεχνολογικές μεθόδους θα μπορεί να αξιοποιείται από υπηρεσίες ασφαλείας για την αναζήτηση ύποπτων ατόμων σε δημόσιους χώρους, για παράδειγμα στα αεροδρόμια.

Σκοπός της πτυχιακής εργασίας είναι να εξεταστούν τεχνικές αυτόματης οπτικής ταυτοποίησης προσώπων, θα διερευνηθούν τα ποσοστά επιτυχίας των αλγορίθμων και θα προταθεί μια υλοποίηση. Η μέθοδος που υλοποιήθηκε στην παρούσα πτυχιακή εργασία είναι η μέθοδος της Ανάλυσης Κυρίων Συνιστωσών (PCA).

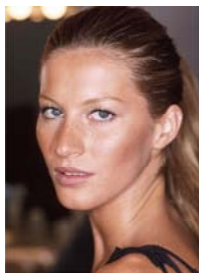
Αναλυτικά οι αλγόριθμοι που εξετάστηκαν για την ταυτοποίηση είναι:

1. Ανάλυση σε Κύριες Συνιστώσες (Principal Component Analysis, PCA).
2. Γραμμική Διακρίνουσα Ανάλυση (Linear Discriminant Analysis, LDA).
3. Αλγόριθμοι Πυρήνα
4. Ανάλυση σε Κύριες Συνιστώσες πυρήνα (Kernel PCA)
5. Γραμμική Διακρίνουσα Ανάλυση του πυρήνα (Kernel Fisher LDA)

## 1. Εισαγωγή

Η αναγνώριση προσώπου είναι μια διαδικασία που οι άνθρωποι πραγματοποιούν στην καθημερινή τους ζωή. Ο ανθρώπινος εγκέφαλος έχει την ικανότητα να αναγνωρίζει πρόσωπα, βάση μιας πολύπλοκης διαδικασίας. Στις μέρες μας, με την ύπαρξη ισχυρών υπολογιστικών συστημάτων επεξεργασίας, η αυτόματη αναγνώριση προσώπων με την χρήση ψηφιακών απεικονίσεων των αντίστοιχων προσώπων προσπαθεί να προσεγγίσει την ικανότητα του ανθρώπινου εγκέφαλου.

Η αναγνώριση προσώπου με τη χρήση υπολογιστικής όρασης (computer vision) είναι μια διαδικασία που συγκρίνει μια εικόνα προσώπου (εικόνα βάσης ελέγχου) με εικόνες προσώπων αποθηκευμένες σε μία βάση δεδομένων. Δύο διαφορετικά προβλήματα, με διαφορετικές προσεγγίσεις ως προς την επίλυση τους είναι το πρόβλημα της ταυτοποίησης και το πρόβλημα της εξακρίβωσης της ύπαρξης του ατόμου στη βάση. Η ταυτοποίηση πραγματοποιείται με την εύρεση εκείνης της εικόνας της βάσης που έχει τις περισσότερες «ομοιότητες» με την εικόνα ελέγχου σύμφωνα με κάποιο ποσοτικό κριτήριο. Το πρόβλημα ταυτοποίησης είναι «οριοθετημένο» μιας και το πρόσωπο ελέγχου θεωρείται ότι βρίσκεται στην βάση. Τα χαρακτηριστικά του προσώπου ελέγχου συγκρίνονται με αυτά της βάσης με την χρήση κατάλληλων συναρτήσεων ομοιότητας, ολοκληρώνοντας με αυτόν τον τρόπο τη διαδικασία της αναγνώρισης.



Who is she?



Η ανάλυση που προσώπου για την εξαγωγή χαρακτηριστικών του προσώπου είναι ιδιαίτερα σημαντική, για αυτό το σκοπό οι εικόνες των προσώπων πρέπει να είναι καλής ανάλυσης και τα χαρακτηριστικά ευδιάκριτα .

## 1.1 Περίληψη

Η αναγνώριση προσώπων έχει γνωρίσει τα τελευταία χρόνια ιδιαίτερη ανάπτυξη, γεγονός που οφείλεται κυρίως στο μεγάλο εύρος των εφαρμογών που χρησιμοποιούν συστήματα αναγνώρισης προσώπων και την διαθεσιμότητα κατάλληλης τεχνολογίας μετά από πολλά χρόνια έρευνας. Επιπλέον, το πρόβλημα της αναγνώρισης εικόνων του ανθρώπινου προσώπου από τον υπολογιστή εξακολουθεί να προσελκύει ερευνητές που δραστηριοποιούνται στις περιοχές της επεξεργασίας εικόνας, της αναγνώρισης προτύπων, των νευρωνικών δικτύων, της υπολογιστικής όρασης, των γραφικών του υπολογιστή κ.α.

Σκοπός της πτυχιακής εργασίας είναι να εξεταστούν τεχνικές αυτόματης οπτικής ταυτοποίησης προσώπων, θα διερευνηθούν τα ποσοστά επιτυχίας των αλγορίθμων και θα προταθεί μια υλοποίηση. Η μέθοδος που υλοποιήθηκε στην παρούσα πτυχιακή εργασία είναι η μέθοδος της Ανάλυσης Κυρίων Συνιστωσών (PCA).

Αναλυτικά οι αλγόριθμοι που εξετάστηκαν για την ταυτοποίηση είναι:

1. Ανάλυση σε Κύριες Συνιστώσες (Principal Component Analysis, PCA).
2. Γραμμική Διακρίνουσα Ανάλυση (Linear Discriminant Analysis, LDA).
3. Αλγόριθμοι Πυρήνα
4. Ανάλυση σε Κύριες Συνιστώσες πυρήνα (Kernel PCA)
5. Γραμμική Διακρίνουσα Ανάλυση του πυρήνα (Kernel Fisher LDA)

## 1.2 Κίνητρο Για την Διεξαγωγή της Εργασίας

Κίνητρο της παρούσας εργασίας είναι η σύγκριση διαφορετικών αλγορίθμων στην αναγνώριση προσώπων, εξέταση των παραγόντων που την επηρεάζουν και υλοποίηση ενός αλγορίθμου.

## 1.3 Σκοπός και Στόχοι Εργασίας

1. Ανάλυση λογισμικού Face Recognition χρησιμοποιώντας Eigenfaces.
2. Σύγκριση Τεχνικών
3. Εξέταση των παραγόντων που επηρεάζουν την έκβαση του αποτελέσματος.
4. Δυνατότητες Βελτίωσης του υπάρχοντος Λογισμικού.

## 1.4 Δομή Εργασίας

Στο πρώτο μέρος αναλύεται η έννοια της βιομετρικής αναγνώρισης, παρουσιάζονται εφαρμογές της αναγνώρισης προσώπων και τα προβλήματα που προκύπτουν.

Στο δεύτερο κεφάλαιο εξετάζεται η έννοια ταυτοποίηση, η χρησιμότητά της και οι κίνδυνοι που παραμονεύουν και στάδια για την αυτόματη αναγνώριση προσώπων.

Στο τρίτο κεφάλαιο κατηγοριοποιούμε τους αλγόριθμους ανάλογα με τα χαρακτηριστικά τους και στη συνέχεια τους υλοποιούμε με σκοπό την αποδοτικότητά τους και την αποτελεσματικότητά τους.

Στο τέταρτο μέρος γίνεται δημιουργία μιας διεπαφής με βάση του αλγορίθμου PCA και δοκιμές βάση ορισμένων συνθηκών.

Στο τέλος της εργασίας παραθέτουμε τα συμπεράσματά μας.

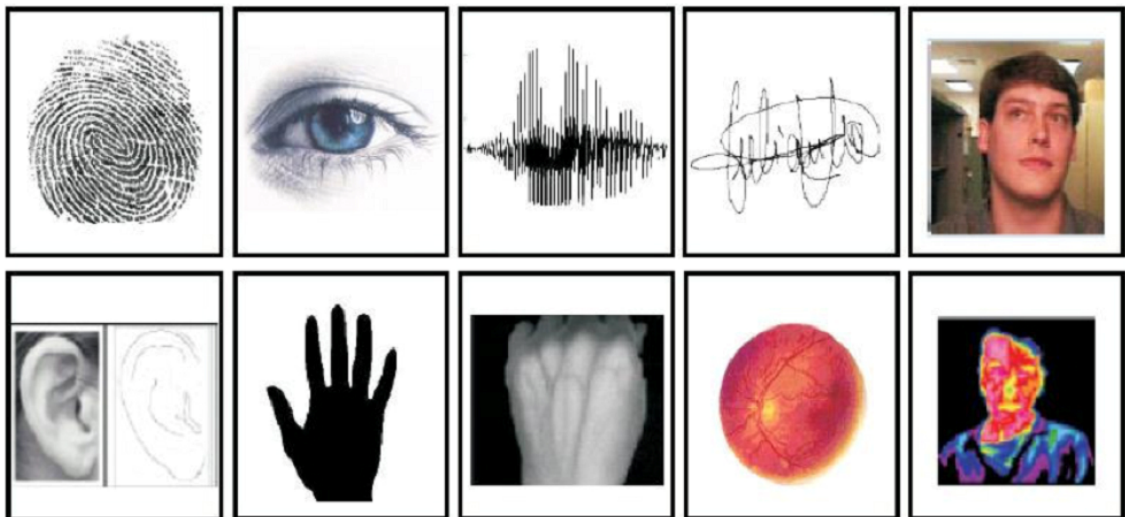
## 2. Βιομετρική Αναγνώριση

### 2.1 Εισαγωγή στην Βιομετρική Αναγνώριση

#### 2.1.1 Τι είναι η βιομετρική αναγνώριση;

Στη βιομετρική αναγνώριση αναπτύσσονται και μελετώνται μέθοδοι που χρησιμοποιούνται για την αναγνώριση ανθρώπων με βάση χαρακτηριστικά που εξάγονται από την ανθρώπινη φυσιολογία (physical traits) είτε από τη ανθρώπινη συμπεριφορά (behavioral traits). Στην επιστήμη υπολογιστών, ο όρος βιομετρική αναγνώριση αναφέρεται σε τεχνολογίες μέτρησης και ανάλυσης ανθρώπινων χαρακτηριστικών φυσιολογίας ή/ και συμπεριφοράς με σκοπό την αναγνώριση ατόμων. Παραδείγματα χαρακτηριστικών φυσιολογίας αποτελούν το δακτυλικό υποτύπωμα, ίριδα του ματιού, πρότυπα πρόσωπα και γεωμετρία χεριού ενώ χαρακτηριστικά συμπεριφοράς η υπογραφή, το περπάτημα, τα πρότυπα δακτυλογράφησης και η φωνή.

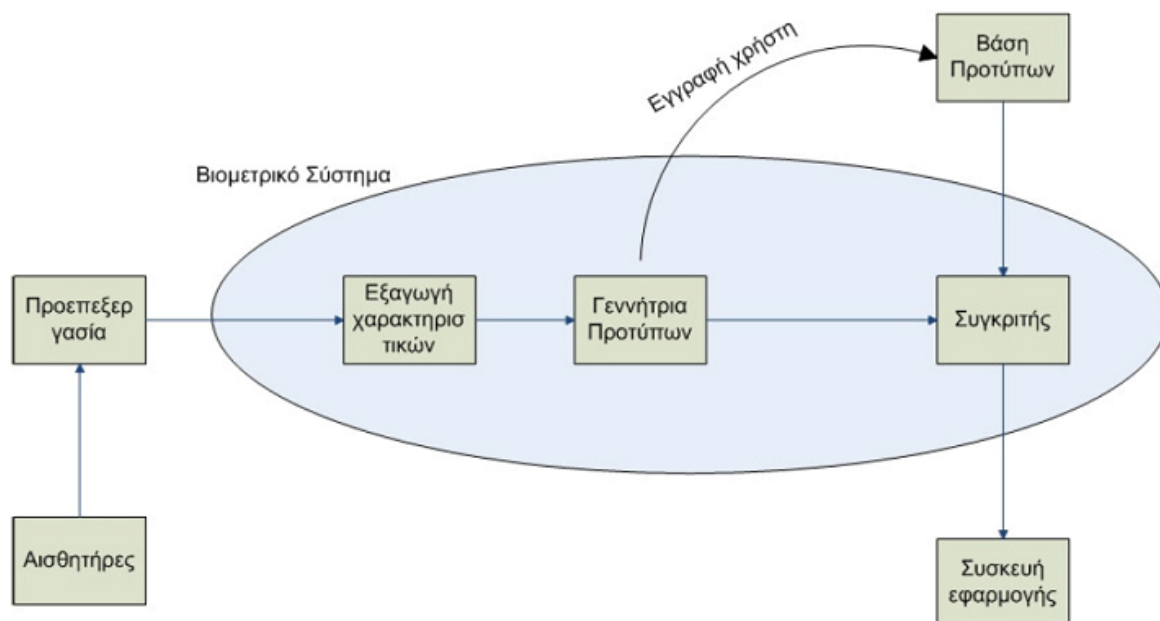
Στην Εικόνα 1 φαίνονται παραδείγματα κοινών βιομετρικών χαρακτηριστικών.



Εικόνα 1: Κοινά βιομετρικά χαρακτηριστικά

#### 2.1.2 Περιγραφή και δυνατότητες ενός βιομετρικού συστήματος

Στην Εικόνα 2 παρουσιάζεται το γενικό διάγραμμα ενός βιομετρικού συστήματος. Οι κύριες λειτουργίες που μπορεί να επιτελέσει το σύστημα είναι η Εγγραφή του χρήστη και η Σύγκριση. Κατά την Εγγραφή, τα βιομετρικά δεδομένα ενός χρήστη καταγράφονται στην βάση δεδομένων, ενώ κατά την διάρκεια της Σύγκρισης νέα δεδομένα συγκρίνονται με τα ήδη υπάρχοντα στη βάση. Το πρώτο υποσύστημα, οι αισθητήρες, αποτελεί το συνδετικό κρίκο ανάμεσα στον πραγματικό κόσμο και το βιομετρικό. Οι αισθητήρες είναι οι υπεύθυνοι για τη συλλογή των απαιτούμενων δεδομένων του βιολογικού χαρακτηριστικού (π.χ. κάμερα για την συλλογή εικόνων προσώπου, μικρόφωνο για συλλογή φωνής κτλ.) Η λειτουργία τους καθορίζεται κάθε φορά σύμφωνα με τα χαρακτηριστικά της εκάστοτε εφαρμογής.



**Εικόνα 2:** Χαρακτηριστικό διάγραμμα ενός βιομετρικού συστήματος

Στο στάδιο της επεξεργασίας γίνονται οι κατάλληλες ενέργειες για την απομάκρυνση του θορύβου και την ενίσχυση του σήματος που λαμβάνεται από τα αισθητήρια και την κανονικοποίηση των δεδομένων. Το τρίτο στάδιο αυτό της εξαγωγής των χαρακτηριστικών, είναι ίσως το πιο σημαντικό του όλου συστήματος. Σε αυτό το βήμα πρέπει να αποφασιστεί ποιά χαρακτηριστικά θα εξαχθούν και πως. Επιπλέον, αυτή η διαδικασία πρέπει να γίνει με συγκεκριμένη αποδοτικότητα και να μην απαιτείται μεγάλο χρονικό διάστημα για την ολοκλήρωσή της. Μετά από αυτό το στάδιο έχει ληφθεί ένα διάνυσμα χαρακτηριστικών (feature vector) με το οποίο χρησιμοποιείται για την δημιουργία ενός προτύπου. Ένα πρότυπο είναι μια σύνθεση από όλα τα χαρακτηριστικά που μπορούν να εξαχθούν από αρχικά δεδομένα. Η γεννήτρια προτύπων δηλαδή χρησιμοποιεί το παραπάνω διάνυσμα αλλάζοντας τα χαρακτηριστικά του (όπως είναι το μήκος του) για καλύτερη απόδοση του συστήματος. Το πρότυπο πρέπει να είναι όσο μικρότερο γίνεται για λόγους αύξησης της αποδοτικότητας αλλά δεν πρέπει να απορρίπτονται πολλές λεπτομέρειες από την πληροφορία για λόγους ελάττωσης της διάκρισης μεταξύ των διαφορετικών προτύπων. Τέλος το σύστημα εκτελεί είτε την διαδικασία της Εγγραφής είτε Σύγκρισης αποθηκεύοντας στη μια το πρότυπο σε κάποια βάση δεδομένων ή εξάγοντας στην άλλη ένα σκορ ομοιότητας μεταξύ των νέων και αποθηκευμένων δεδομένων. Το σκορ που εξάγεται χρησιμοποιείται από την εφαρμογή για τη λήψη κάποιας απόφασης.

Κλείνοντας την περιγραφή ενός βιομετρικού συστήματος πρέπει να αναφέρουμε τις δύο βασικές δυνατότητες στην αναγνώριση ατόμων, την Επαλήθευση (Verification) και την Ταυτοποίηση (Identification). Η Επαλήθευση απαντά στο ερώτημα : << Είναι αυτό το άτομο που ισχυρίζεται ότι είναι;>> Για παράδειγμα κάποιος ισχυρίζεται ότι είναι ένας από τους εγγεγραμμένους του συστήματος. Κατά τη διάρκεια της Επαλήθευσης εξάγονται νέα βιομετρικά δεδομένα από το συγκεκριμένο άτομο τα οποία συγκρίνονται με τα ήδη υπάρχοντα στη βάση δεδομένων και έτσι εξακριβώνεται η ταυτότητα του χρήστη. Αποτελεί μία ένα-προς-ένα σύγκριση (one-to-one Match) . Από την άλλη, στη Ταυτοποίηση εξάγονται βιομετρικά δεδομένα από το άτομο και συγκρίνονται με τα αντίστοιχα όλων των εγγεγραμμένων χρηστών. Αυτή είναι μια διαδικασία Σύγκρισης ένα-προς-πολλά (one-to-many Match). Είναι φανερό ότι η Ταυτοποίηση είναι μια πιο δύσκολη και χρονοβόρα

διαδικασία από την Επαλήθευση αφού για την εξακρίβωση της ταυτότητας απαιτούνται τόσες συγκρίσεις όσες και οι καταχωρήσεις στην βάση δεδομένων που για τα περισσότερα συστήματα ο αριθμός τους είναι πολύ μεγάλος.

### 2.1.3 Το πρόβλημα της αναγνώρισης προσώπων

Οι άνθρωποι εντοπίζουν και αναγνωρίζουν πρόσωπα σε μια σκηνή με χαρακτηριστική ευκολία. Η ικανότητα αυτή αναπτύσσεται σταδιακά κατά τη διάρκεια της παιδικής ηλικίας και μαζί με τη δυνατότητα εκτίμησης της έκφρασης ενός προσώπου παίζουν ένα πολύ σημαντικό ρόλο στη κοινωνική ζωή των ανθρώπων. Η ευκολία εντούτοις με την οποία εκτελείται η παραπάνω διαδικασία στον άνθρωπο δεν πρέπει να μας παραπλανεί. Τα πρόσωπα αντιπροσωπεύουν σύνθετες, πολυδιάστατες, γεμάτες πληροφορία διεγέρσεις, πράγμα που καθιστά την ανάπτυξη ενός υπολογιστικού συστήματος αναγνώρισης προσώπων εξαιρετικά δύσκολη.

Το πρόσωπο είναι ένα μοναδικό χαρακτηριστικό για τα ανθρώπινα όντα. Ακόμα και τα πρόσωπα δύο ομοζυγωτικών διδύμων διαφέρουν. Εντοπίζουμε, αναγνωρίζουμε και διακρίνουμε τα πρόσωπα με χαρακτηριστική ευκολία και απλή παρατήρηση. Η ευκολία αναγνώρισης των προσώπων από τον άνθρωπό ήταν ένας από τους λόγους για την επιλογή γραμματοσήμων το οποία περιλάμβαναν πρόσωπα: ήταν πολύ δυσκολότερο να γίνουν πλαστογραφίες γραμματοσήμων γιατί αυτές γίνονταν εύκολα αντιληπτές. Η μοναδικότητα των προσώπων είναι επίσης ο βασικός λόγος για την ευρεία χρήση του σε εφαρμογές όπου η ταυτοποίηση ανθρώπων είναι ιδιαίτερα σημαντική.



**Εικόνα 3:** Ομοζυγωτικά δίδυμα



**Εικόνα 4:** Εκφράσεις προσώπων

Στο σημείο αυτό πρέπει να γίνει διάκριση μεταξύ ταυτοποίησης προσώπων και ταυτοποίησης ανθρώπων με βάση τη συνολική τους συμπεριφορά. Στην ταυτοποίηση προσώπων χρησιμοποιείται αποκλειστικά το οπτικό σύστημα του ανθρώπου. Στην γενικότερη ταυτοποίηση ανθρώπων χρησιμοποιούνται και άλλα στοιχεία, όπως το βάδισμα, η στάση του σώματος, η φωνή ή το ντύσιμο. Παρόλο που το πρόσωπο και τα χαρακτηριστικά του χρησιμοποιούνται ευρύτατα για την αναγνώριση

ανθρώπων με κανένα τρόπο δεν αποτελούν τους μοναδικούς τρόπους αναγνώρισης. Στο τέλος του προηγούμενου αιώνα, ένα σύστημα βασισμένο σε μετρήσεις άλλων χαρακτηριστικών του σώματος, όπως το μήκος και εύρος του κεφαλιού, το μήκος του αριστερού πέλματος, το μήκος του μεσαίου δακτύλου του αριστερού χεριού και ποδιού, είχε χρησιμοποιηθεί στη Γαλλία για την ταυτοποίηση φυλακισμένων. Πρόσφατα αναπτύχθηκαν τεχνικές για ταυτοποίηση ανθρώπων βασισμένες στη δομή των γονιδίων.

Ένα σύστημα εντοπισμού και ταυτοποίησης προσώπων θα ήταν χρήσιμο σε μια πλειάδα εφαρμογών όπως αναγνώριση εγκληματιών, πρόσβαση σε συστήματα ασφαλείας όπως ταμεία τραπεζών (ATM) και γενικά σε κάθε εφαρμογή που απαιτεί πρόσβαση βάση κάποιου αποδεικτικού στοιχείου ταυτότητας. Τέτοια συστήματα στη παρούσα φάση χρησιμοποιούν στοιχεία όπως κάρτες, PIN και κωδικούς πρόσβασης. Πρόσβαση βασισμένη σε αναγνώριση προσώπων θα μπορούσε να είναι περισσότερο αξιόπιστη και πιο απλή. Η αλήθεια πάντως είναι ότι σε εφαρμογές που απαιτούν υψηλή αξιοπιστία όσον αφορά την ταυτοποίηση, η αναγνώριση προσώπου δεν επαρκεί –τουλάχιστον με τα υπάρχοντα δεδομένα. Αντίθετα σε εφαρμογές όπως τα συστήματα επικοινωνίας ανθρώπου μηχανής, ανάκλησης στοιχείων από βάσεις δεδομένων με βάση το περιεχόμενο, τηλεδιάσκεψη κλπ, η αναγνώριση προσώπου είναι πιο εύκολα υλοποιήσιμη και σε πολλές περιπτώσεις απαραίτητη.

Η ανάπτυξη ενός υπολογιστικού μοντέλου αναγνώρισης προσώπων είναι μια πρόκληση για την ανθρώπινη εφευρετικότητα και φιλοδοξία. Οι πιθανές εφαρμογές ενός τέτοιου μοντέλου, ενδεχομένως σε μελλοντικό χρόνο, χρησιμοποιούμενου για ασφάλεια κτιρίων και σταθμών εργασίας, αναγνώριση κακοποιών, επαλήθευση πιστωτικών καρτών, βελτίωσης της επικοινωνίας ανθρώπου - μηχανής κος, το καθιστούν ιδιαίτερα σημαντικό.

Το πρόβλημα χονδρικά τίθεται ως εξής: Δοθέντων φωτογραφιών ή ακολουθιών βίντεο από μια σκηνή, αναγνώρισε ένα ή περισσότερα άτομα χρησιμοποιώντας μια υπάρχουσα βάση δεδομένων από πρόσωπα. Συμπληρωματικές πληροφορίες, όπως ηλικία, φύλο, φυλή, μπορεί να είναι διαθέσιμες για κάθε αποθηκευμένο στη βάση πρόσωπο, ούτως ώστε να περιορίζεται η έρευνα. Διάφορα υποπροβλήματα πηγάζουν για την αυτοματοποίηση της παραπάνω διαδικασίας:

- (1) Απομόνωση και κατάτμηση του προσώπου μέσα από ένα –γενικά- θορυβώδες περιβάλλον.
- (2) Εξαγωγή ευσταθών χαρακτηριστικών από την περιοχή του προσώπου. Ευσταθή χαρακτηριστικά είναι εκείνα, τα οποία δεν επηρεάζονται από την κλίμακα, τον φωτισμό, την κλίση-στροφή του προσώπου, την αμφίεση και τη γήρανση.
- (3) Χρήση χαρακτηριστικών ή άλλων παρόμοιων τεχνικών με σκοπό τη μείωση της διάστασης του προβλήματος και τη δημιουργία προτύπων. Το πρόβλημα τελικά ανάγεται στην αναγνώριση προτύπων και τη χρήση κατάλληλων μετρικών για ταυτοποίηση-κατηγοριοποίηση.

Επιπλέον το πρόβλημα μπορεί να διαχωριστεί σε δύο κατηγορίες με βάση τη πηγή πληροφορίας, αν δηλαδή αναφέρεται σε σταθερές ή κινούμενες εικόνες. Η αναγνώριση προσώπου από σταθερές εικόνες έχει πλεονεκτήματα και μειονεκτήματα σε σχέση με την αναγνώριση από ακολουθίες βίντεο, ως εκ τούτου οι χρησιμοποιούμενες τεχνικές στις δύο αυτές κατηγορίες ποικίλουν.

Ο τρόπος με τον οποίο γίνεται η αναγνώριση προσώπων από τον άνθρωπο έχει απασχολήσει τους τομείς της Ψυχοφυσικής και της Νευροφυσιολογίας για περισσότερα από 20 χρόνια. Η μοναδικότητα των προσώπων, το ερώτημα αν η αναγνώριση γίνεται ολιστικά ή με χρήση επιμέρους χαρακτηριστικών, η ανάλυση των εκφράσεων του προσώπου για σκοπούς αναγνώρισης, η οργάνωση

της μνήμης για τα πρόσωπα και διάφορα άλλα θέματα απασχόλησαν τις παραπάνω επιστήμες. Οι μηχανικοί έδειξαν ενδιαφέρον για την έρευνα αυτή με σκοπό την υλοποίηση ενός αποδοτικού συστήματος αναγνώρισης. Μερικά από τα αποτελέσματα είναι πράγματι ενδιαφέροντα αν και τελικά η αναγνώριση προσώπων από μηχανές αναπτύχθηκε ανεξάρτητα από τις μελέτες της Ψυχοφυσικής και Νευροφυσιολογίας.

Σημαντική προσπάθεια στον τομέα της αναγνώρισης προσώπων από μηχανές είχε γίνει τα πρώτα χρόνια της δεκαετίας του '70. Κατά τη διάρκεια της δεκαετίας του '80 υπήρξε μια σχετική στασιμότητα. Από τις αρχές της δεκαετίας του '90 παρατηρείται μια σημαντική δραστηριότητα και ανανεωμένο ενδιαφέρον, το οποίο εκφράζεται με τη δημοσίευση σε τακτική βάση άρθρων σε περιοδικά όπως το *IEEE Transactionson PAMI*, και τη διοργάνωση ειδικών παρουσιάσεων –sessions– σε μεγάλα διεθνή συνέδρια (*IEEE ICIP, IEEE ICASSP, IEEE Multimedia*). Έναυσμα για την ανανέωση του ενδιαφέροντος και περαιτέρω δραστηριοποίηση έδωσε η εργασία [1]. Εκτοτε μια μεγάλη ποικιλία προσεγγίσεων έχει εμφανιστεί στη διεθνή βιβλιογραφία.

#### **2.1.4 Εφαρμογές της αναγνώρισης προσώπων**

Εμπορικές και νομικής φύσεως εφαρμογές της αναγνώρισης προσώπου καταγράφονται στον Πίνακα 1. Τα δεδομένα εισόδου των εφαρμογών αυτών ποικίλουν από στατικές φωτογραφίες με ελεγχόμενο φόντο έως ακολουθίες βίντεο με κανένα περιορισμό στη λήψη και την ανάλυση. Δεδομένης της ανομοιογένειας των πηγών εισόδου είναι φυσικό επακόλουθο η ύπαρξη πολλών μεθόδων που έχουν σχέση με πεδία όπως η επεξεργασία, ανάλυση και κατανόηση εικόνας και η αναγνώριση προτύπων, που προσπαθούν να αντιμετωπίσουν επιμέρους ζητήματα.

Χονδρικά μπορεί κάποιος να ταξινομήσει τις τεχνικές και τα επιμέρους προβλήματα σε δύο κατηγορίες: στατικής ταύτισης -δεν υπάρχει δυνατότητα παρακολούθησης κίνησης- και δυναμικής ταύτισης -στην οποία υπάρχει δυνατότητα παρακολούθησης της χρονικής μεταβολής των δεδομένων. Ακόμη και στο πλαίσιο των ανωτέρω γενικών κατηγοριών υπάρχουν σημαντικές διαφοροποιήσεις ανάλογα με τη συγκεκριμένη εφαρμογή. Οι διαφορές που αφορούν στην ποιότητα της εικόνας - χωρική ανάλυση, στην πολυπλοκότητα του φόντου η οποία επηρεάζει αποφασιστικά την ευκολία με την οποία μπορεί να εντοπιστεί το πρόσωπο μέσα στην εικόνα, στην ύπαρξη ενός καλά ορισμένου κριτηρίου ταύτισης καθώς και στο ποσοστό, στο είδος και στην υφή της παρέμβασης από τον άνθρωπο -όπως στις εφαρμογές 4 και 5 του Πίνακα 1. Σε μερικές εφαρμογές, όπως στον εντοπισμό των μεταβολών που επήλθαν με την γήρανση απαιτείται η εύρεση των μετασχηματισμών εκείνων που μετά την εφαρμογή τους οι νέες εικόνες θα είναι παρόμοιες με τις αναμενόμενες.

Με βάση τα προηγούμενα προκύπτουν τρία διαφορετικά προβλήματα που αφορούν στις εφαρμογές του Πίνακα 1. Συγκεκριμένα η ταύτιση, η ανίχνευση ομοιότητας και η εφαρμογή μετασχηματισμών. Οι εφαρμογές 1, 2, 3, 10 και 11 θέτουν κυρίως προβλήματα ταύτισης μιας εικόνας προσώπου με κάποια άλλη. Οι εφαρμογές 4-7 αναφέρονται στη δημιουργία μιας εικόνας προσώπου όμοιας με αυτήν που θυμούνται ή περιγράφουν αυτόπτες μάρτυρες. Τέλος οι εφαρμογές 8 και 9 εμπλέκουν την εφαρμογή μετασχηματισμών και τροποποιήσεων σε μια εικόνα με βάση στοιχεία που είναι χρήσιμα σε άλλες εφαρμογές. Κάθε μια από τις εφαρμογές που αναφέρονται στον Πίνακα 1 θέτει διαφορετικές απαιτήσεις και περιορισμούς στη διαδικασία της αναγνώρισης προσώπου. Η ταύτιση απαιτεί όπως η προς ταυτοποίηση εικόνα ανήκει σε ένα σύνολο από εικόνες οι οποίες έχουν επιλεγεί από το υπολογιστικό σύστημα. Η ανίχνευση ομοιότητας εκτός από την ταύτιση απαιτεί επιπλέον οι επιλεγείσες εικόνες να ομοιάζουν με αυτές που αυτόπτες μάρτυρες ανακαλούν στη μνήμη τους. Η τελευταία απαίτηση θέτει ένα σοβαρό περιορισμό στο κριτήριο ταύτισης το οποίο πρέπει να



είναι αντίστοιχο με αυτό που χρησιμοποιούν οι άνθρωποι. Οι εφαρμογές μετασχηματισμού από την άλλη πλευρά πρέπει να προσομοιάζουν τον τρόπο με τον οποίο λειτουργεί η γήρανση -ή τουλάχιστον αυτό που αναμένουν οι άνθρωποι να συμβεί ως συνέπεια της παρόδου του χρόνου.

A/A	ΕΦΑΡΜΟΓΕΣ	ΠΛΕΟΝΕΚΤΗΜΑΤΑ	ΜΕΙΟΝΕΚΤΗΜΑΤΑ
1(α)	Πιστωτικές κάρτες, ATM μηχανήματα, δίπλωμα οδήγησης, διαβατήριο, προσωπική Ταυτοποίηση.	Καλή ποιότητα εικόνων, ελεγχόμενες συνθήκες λήψης, εύκολος εντοπισμός του προσώπου	Μη υπάρχουσες βάσεις, μεγάλο μέγεθος βάσεων, σπάνια χρησιμοποιούμενη αναζήτηση
1(β)	Mug shots ταύτιση	Ελεγχόμενες συνθήκες λήψης, περισσότερες από μια φωτογραφίες για κάθε άτομο.	
2	Ασφάλεια τραπεζών / καταστημάτων	Υψηλή ζήτηση, γεωγραφικά και τοπικά περιορισμένη αναζήτηση.	Μη ελεγχόμενο φόντο, μικρός αριθμός φωτογραφιών
3	Επισκόπηση - παρακολούθηση πλήθους	Υψηλή ζήτηση, μικρό μέγεθος αρχείων, Εκμετάλλευση κίνησης	Μη ελεγχόμενο φόντο, χαμηλή ποιότητα εικόνων ,εκτέλεση σε πραγματικό χρόνο.
4	Ταυτοποίηση από ειδικούς	Υψηλή ζήτηση, δυνατότητα αναβάθμισης των φωτογραφιών	Χαμηλή ποιότητα εικόνων, απαραίτητη η άδεια από τις αστυνομικές αρχές.
5	Ανακατασκευή προσώπου με βάση μαρτυρίες	Όρια που τίθενται από τους αυτόπτες μάρτυρες	Αμφίβολη ομοιότητα
6	Ηλεκτρονικό Βιβλίο από mugshots	Όρια που τίθεται από την περιγραφή	Κούραση παρατηρητή
7	Ηλεκτρονική παράταξη υπόπτων	Όρια που τίθεται από την περιγραφή	Κούραση παρατηρητή
8	Ανακατασκευή προσώπου από υπολειπόμενα στοιχεία	Υψηλή ζήτηση	Απαιτείτε η χρήση δεδομένων από τη φυσιολογία
9	Υπολογισμός αλλαγών που επήλθαν με την Πάροδο του χρόνου (aging)	Υψηλή ζήτηση	Απαιτείται η ύπαρξη συγκεκριμένου παραδείγματος
10	Αναζήτηση με βάση το περιεχόμενο	Χρήση χρώματος	Ετερογενείς βάσεις πολυμέσων

11	Υποδιαίρεση ειδήσεων σε θεματικές ενότητες.	Χρήση χρώματος	Ανομοιογένεια συνθηκών studio και εξωτερικών λήψεων
----	---	----------------	---

**Πίνακας 1:** Συνήθεις εφαρμογές της αναγνώρισης προσώπων

Η τεχνολογία αναγνώρισης προσώπου, δεν χρησιμοποιείται μόνο για την αναγνώριση εγκληματιών από τις αστυνομικές αρχές. Πολλά προγράμματα διαχείρισης φωτογραφικών άλμπουμ, χρησιμοποιούν μεθόδους για να αναγνωρίσουν αυτόματα πρόσωπα και να τα ομαδοποιήσουν για λογαριασμό του χρήστη:

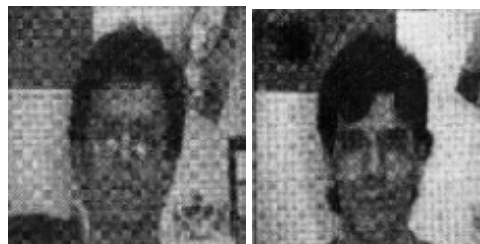
- Το iPhoto της Apple οργανώνει τις φωτογραφίες του χρήστη αυτόματα ανάλογα με τα πρόσωπα που εικονίζονται και είναι διαθέσιμα στο χρήστη ανάμεσα σε πολλαπλές εφαρμογές.
- Το Picassa της Google δίνει τη δυνατότητα στο χρήστη να επιλέξει την προβολή φωτογραφιών με συγκεκριμένο πρόσωπο και το ίδιο το πρόγραμμα διασύνδεει τις φωτογραφίες με το πρόσωπο και τις προβάλλει.
- Το Picture Motion Browser (PMB) της Sony, αναλύει τις φωτογραφίες και τις ομαδοποιεί ανάλογα με τα πρόσωπα που εμφανίζονται.
- Το Facebook χρησιμοποιεί πλέον σύστημα αναγνώρισης προσώπου, κάνοντας αυτόματα επισήμανση (tagging) στα πρόσωπα που υπάρχουν.
- Το Windows Live Photo Gallery στην τελευταία έκδοση χρησιμοποιεί επίσης τέτοιο σύστημα.
- Το fotobounce έδινε τη δυνατότητα διαχείρισης των φωτογραφιών του χρήστη ανάλογα με τους εικονιζόμενους ανθρώπους.
- Το [face.com](http://face.com) δίνει έτοιμο API για τους χρήστες που θέλουν να εισάγουν τεχνολογία αναγνώρισης προσώπου, μετρώντας 35,000,000,000+ πρόσωπα στη βάση δεδομένων του (κατά την 30η/11/2011). Τώρα το έχει αποκτήσει το [facebook.com](http://facebook.com).
- Ξεκλείδωμα εφαρμογών με αναγνώριση προσώπου στο Android. Σκοπός ύπαρξής της είναι το κλείδωμα των εφαρμογών που εμείς θα ορίσουμε για λόγους ασφαλείας και το ξεκλείδωμά τους με τη βοήθεια του προσώπου μας. Έτσι μπορούμε να κλειδώσουμε εφαρμογές όπως π.χ. SMS, Gmail, Gallery με ευαίσθητα προσωπικά δεδομένα αλλά και όποια άλλη επιλέξουμε.

### 2.1.5 Προβλήματα στατικής ταύτισης

Η πιο συνηθισμένη εφαρμογή της κατηγορίας αυτής είναι η ταύτιση «mugshot» φωτογραφιών. Τυπικά στις φωτογραφίες αυτού του είδους, ο φωτισμός είναι ελεγχόμενος και λαμβάνονται δύο φωτογραφίες μια εμπρόσθιας όψης –ανφάς- και μία πλάγιας όψης -προφίλ. Παρόλο που γενικά οι προδιαγραφές λήψης των φωτογραφιών μπορούν να τεθούν από τις αστυνομικές αρχές, δεν υπάρχει κάποιο συγκεκριμένο πρότυπο και σαν αποτέλεσμα υπάρχει μια ποικιλία ως προς τη μορφή τους ανάλογα με την περιοχή. Οι προδιαγραφές που μπορούν να τεθούν αφορούν στο φόντο, στις συνθήκες φωτισμού, στην χωρική ανάλυση και στην απόσταση του εικονιζόμενου από την κάμερα. Θέτοντας περιορισμούς όπως οι προηγούμενοι, τόσο η διαδικασία εντοπισμού, όσο και η διαδικασία ταύτισης απλοποιούνται σε μεγάλο βαθμό. Παραδείγματα mugshots φωτογραφιών δίνονται στην Εικόνα 5



**Εικόνα 5:** (α) Mugshot φωτογραφία (front view) (β) Mugshot φωτογραφία(profile) (γ) (δ) Φωτογραφίες σε διπλώματα οδήγησης, διαβατήρια κλπ.



(α)

(β)

**Εικόνα 6:** (α)(β)Φωτογραφίες με ανομοιόμορφο φόντο

Η εφαρμογή 2 είναι περισσότερο σύνθετη από την εφαρμογή 1 κυρίως εξαιτίας της μη ελεγχόμενης διαδικασίας λήψης των φωτογραφιών. Το φόντο δεν είναι υποχρεωτικά ομοιόμορφο και κατά συνέπεια η διαδικασία εντοπισμού και κατάτμησης καθίσταται δυσκολότερη. Επιπλέον η ποιότητα των εικόνων τείνει να είναι χαμηλή. Μια προσέγγιση φωτογραφιών της εφαρμογής αυτής δίνεται στις Εικόνες 6(α) και (β). Πρέπει να τονιστεί ότι ουσιαστικά η εφαρμογή 2 βρίσκεται ανάμεσα στη στατική και τη δυναμική ταύτιση δεδομένου ότι σε πολλές περιπτώσεις εκτός από φωτογραφίες από συνήθεις φωτογραφικές μηχανές λαμβάνονται και εικόνες από βιντεοκάμερα. Όπως και στην εφαρμογή 1, μεταβολές στα πρόσωπα οι οποίες προέρχονται από γήρανση, μεταμπίωση και απώλεια ή αλλαγή μαλλιών πρέπει να λαμβάνονται υπόψη στη διαδικασία ταύτισης και εξαγωγής χαρακτηριστικών. Στις εφαρμογές 1 και 2 το κριτήριο ταύτισης μπορεί να είναι ποσοτικό επιτρέποντας κατάταξη μερικών από τις επιλογές του συστήματος σύμφωνα με τη σειρά επιτυχίας.

Οι εφαρμογές 4-7 αφορούν την εύρεση ή δημιουργία ενός προσώπου όμοιου με αυτό που κάποιος αυτόπτης μάρτυρας είδε ή περιέγραψε. Στην εφαρμογή 4 ένας ειδικός καλείται να βεβαιώσει ότι το πρόσωπο στη δοθείσα εικόνα αντιστοιχεί στα αναζητούμενο άτομο. Είναι πιθανό το πρόσωπο της φωτογραφίας να είναι μεταμφιεσμένο ή να καλύπτεται μερικώς από άλλα πρόσωπα ή αντικείμενα. Τυπικά, της εφαρμογής αυτής προηγείται κάποια διαδικασία κατάταξης, η οποία δημιουργεί ένα σύνολο πιθανών προσώπων τα οποία έχουν σημαντικές ομοιότητες και ο ειδικός καλείται με προσεκτική έρευνα να βρει ποια από αυτές αντιστοιχεί στο αναζητούμενο άτομο. Στην εφαρμογή 5 ο αυτόπτης μάρτυρας καλείται να συνθέσει μια προσωπογραφία του αναζητούμενου ατόμου χρησιμοποιώντας μια βιβλιοθήκη από χαρακτηριστικά προσώπου όπως μάτια, μύτες, χείλη κλπ. Για παράδειγμα η βιβλιοθήκη μπορεί να περιέχει μύτες μακριές, κοντές, καμπυλωτές επίπεδες κλπ. Κάποια από τις οποίες θα ταιριάζει περισσότερο με αυτήν που ο μάρτυρας έχει στη μνήμη του. Η εφαρμογή 6 αναφέρεται στην ηλεκτρονική πλοήγηση σε συλλογή από φωτογραφίες. Στην εφαρμογή

7 ο μάρτυρας καλείται να αναγνωρίσει τον ύποπτο από ένα σύνολο φωτογραφιών μερικές από τις οποίες περιέχουν και λανθασμένες επιλογές. Τυπικά στις εφαρμογές 4-7 η ποιότητα των εικόνων είναι σχετικά χαμηλή και πέρα από την ταύτιση ζητείται και η εύρεση προσώπων που ομοιάζουν με αυτό που ο μάρτυρας έχει στη μνήμη του. Το κριτήριο ταύτισης είναι δύσκολο να καθοριστεί ποσοτικά, δεδομένου ότι πρέπει να βρεθούν μετρικές οι οποίες προσομοιάζουν την διαδικασία ταύτισης που πραγματοποιεί ο άνθρωπος. Αξίζει να σημειωθεί ότι όταν ο άνθρωπος καλείται να αναγνωρίσει κάποιο πρόσωπο από μια συλλογή από φωτογραφίες, τα λάθη αναγνώρισης αυξάνονται όσο αυξάνεται ο αριθμός των φωτογραφιών. Συμπερασματικά οι εφαρμογές 4-7 απαιτούν ισχυρή αλληλεπίδραση μεταξύ των αλγορίθμων που χρησιμοποιούνται και αποτελεσμάτων από τους τομείς της Ψυχοφυσικής και της Νευροφυσιολογίας.

Οι εφαρμογές 8 και 9 εμπλέκουν μετασχηματισμό της τρέχουσας εικόνας έτσι ώστε το εμφανιζόμενο πρόσωπο να μοιάζει με αυτό που αναμένεται μετά από την πάροδο κάποιου χρονικού διαστήματος ή με αυτό που θα έπρεπε να είναι. Οι εφαρμογές αυτές είναι ακόμη δυσκολότερες από τις εφαρμογές 4-6 δεδομένου ότι απαιτούν την ενσωμάτωση μηχανισμών εξομάλυνσης και πρόβλεψης στους αλγορίθμους.

### **2.1.5 Προβλήματα Δυναμικής Ταύτισης**

Η εφαρμογή 3 καθώς και περιπτώσεις της εφαρμογής 2 υπάγονται στη δυναμική ταύτιση, δεδομένου ότι είναι διαθέσιμη μια ακολουθία βίντεο και όχι μεμονωμένες φωτογραφίες. Εικόνες προερχόμενες από βιντεοκάμερες τείνουν να έχουν **χαμηλή ποιότητα**. Επιπλέον σε εφαρμογές επισκόπησης πλήθους το φόντο είναι εξαιρετικά ανομοιογενές, πράγμα που καθιστά τη διαδικασία εντοπισμού και κατάτμησης του προσώπου πολύ δύσκολη. Από την άλλη πλευρά, η ύπαρξη κίνησης αποτελεί ένα πολύ χρήσιμο στοιχείο για την κατάτμηση προσώπων από κινούμενα άτομα. Ακόμη, με αλγόριθμους ανακατασκευής 3-D αντικείμενων, με βάση τη κίνηση, μπορεί να επιτευχθεί μερική ανακατασκευή του προσώπου και να αντιμετωπιστούν θέματα μεταμφίεσης σχετικά καλύτερα από ότι στη στατική ταύτιση. Ένας από τους ισχυρότερους περιορισμούς στις εφαρμογές δυναμικής ταύτισης είναι η απαίτηση για εκτέλεση σε πραγματικό χρόνο.

Πρέπει να τονιστεί ότι οι ευρέως διαφοροποιούμενοι περιορισμοί που τίθενται στις επιμέρους εφαρμογές καθιστούν αναγκαία τη χρήση διαφορετικών μεθόδων βαθμολόγησης για την αξιολόγηση των υπάρχοντων αλγορίθμων και συστημάτων.

### **2.1.6 Τα βασικά υποπροβλήματα στην αναγνώριση προσώπων**

Πριν την ταυτοποίηση ενός προσώπου πρέπει πρώτα αυτό να εντοπιστεί. Η ανίχνευση και ο εντοπισμός προσώπων είναι το πρώτο στάδιο της αναγνώρισης προσώπων. Το επόμενο στάδιο είναι η ανάπτυξη μηχανισμών ταυτοποίησης του προσώπου με βάση γνωστές πληροφορίες. Ένα σχετιζόμενο πρόβλημα είναι η αποτελεσματική αναπαράσταση του προσώπου, π.χ. μέσω ενός δισδιάστατου πίνακα τιμών ή ενός συνόλου από χαρακτηριστικά. Προβλήματα αντίστοιχα με την αναγνώριση προσώπων είναι και αυτά της αναγνώρισης του φύλου ή της φυλής.

### *Αναπαράσταση Προσώπων*

Η αναπαράσταση των προσώπων είναι ιδιαίτερα σημαντική στην αναγνώριση. Όλα τα γνωστά πρόσωπα πρέπει να αναπαριστώνται σε κάποια μορφή. Τα νεοεντοπισμένα πρόσωπα πρέπει επίσης να αναπαριστώνται σε κατάλληλη μορφή ώστε η διαδικασία ταύτισης να είναι αφενός εφικτή και αφετέρου αποτελεσματική. Ένα πρόσωπο είναι δύσκολο να αναπαρασταθεί πλήρως από τα μεμονωμένα χαρακτηριστικά του, π.χ. από το αν έχει μακριά μύτη, σκούρα μαλλιά και οβάλ σχήμα. Αυτός είναι και ο λόγος για τον οποίο οι αστυνομικές αρχές δημιουργούν σκίτσα με βάση τις περιγραφές των αυτόπτων μαρτύρων που περιγράφουν το πρόσωπο με βάση επιμέρους χαρακτηριστικά. Η αναπαράσταση των προσώπων πρέπει να είναι συμπαγής, αλλά χωρίς την απώλεια σημαντικής πληροφορίας. Για παράδειγμα, αναπαράσταση σε ανάλυση 512x512, με βάθος χρώματος 6 bpp, μπορεί να είναι επαρκής συγκρινόμενη με 8 bpp. Ομοίως φωτογραφίες με ανάλυση μικρότερη από 512x512 μπορεί να είναι ικανοποιητικές τόσο για τον εντοπισμό όσο και για την ταυτοποίηση. Ο τρόπος αναπαράστασης των προσώπων επηρεάζεται αλλά και επηρεάζει το σχήμα ταύτισης το οποίο χρησιμοποιείται για την αναγνώριση.

### *Ανίχνευση και Εντοπισμός Προσώπων*

Σε πολλές περιπτώσεις οι συνθήκες κάτω από τις οποίες έχει ληφθεί η φωτογραφία είναι ελεγχόμενες, όπως π.χ. σε φωτογραφίες που πραγματοποιούνται από τις αστυνομικές αρχές. Ο εντοπισμός της ακριβούς θέσης του προσώπου στις φωτογραφίες αυτές είναι σχετικά εύκολος και σε αρκετές περιπτώσεις δεν είναι καν απαραίτητος. Στις περισσότερες περιπτώσεις εντούτοις, το εάν σε μια φωτογραφία υπάρχουν πρόσωπο, ή πρόσωπα, δεν είναι εκ των προτέρων γνωστό –πολύ περισσότερο δε, δεν είναι γνωστή η ακριβής θέση του. Διάφοροι παράγοντες, όπως η παρουσία μαλλιών που καλύπτουν μέρος του προσώπου, η ύπαρξη γενειάδας, το βάψιμο, η μερική επικάλυψη από άλλα αντικείμενα κοκ, επηρεάζουν τη διαδικασία εντοπισμού του προσώπου σε μια εικόνα γιατί καλύπτουν μερικά από τα βασικά χαρακτηριστικά του. Ένας άλλος ουσιαστικός παράγοντας είναι η κλίμακα και η γωνία κλίσης του προσώπου στις φωτογραφίες. Το πρόβλημα της κλίμακας και του προσανατολισμού αποτρέπει τον εντοπισμό μέσω ενός απλού προτύπου.

### 3. Σχέδιο δράσης για την εκπόνηση της εργασίας

## Ταυτότητα και Ταυτοποίηση-Στάδια Αυτόματης Ταυτοποίησης

### 3.1 Η έννοια της ταυτότητας

Η έννοια της ταυτότητας σχετίζεται με το χαρακτηρισμό και την αναπαράσταση ενός ατόμου ή μιας ομάδας και όσον αφορά στη δομή αυτού του χαρακτηρισμού. Για παράδειγμα η Ταυτότητα μπορεί να διαφοροποιηθεί σύμφωνα με τις διαφορετικές εκφάνσεις, όπως η προσωπική ταυτότητα (προσωπική), η βιολογική ταυτότητα (DNA – Deoxyribonucleic acid), η κοινωνική ταυτότητα (συμμετοχή) ή η νομική ταυτότητα και να τις συναθροίζει με βάση τη χρήση σε διαφορετικές καταστάσεις (όπως δραστηριότητες ελεύθερου χρόνου, συναλλαγές, εργασιακές ή κοινωνικές αλληλεπιδράσεις). Η έννοια της ταυτότητας μπορεί να εφαρμοσθεί σε ένα φυσικό, ηθικό ή αφηρημένο άτομο (όπως ένας οργανισμός ή ομάδα).

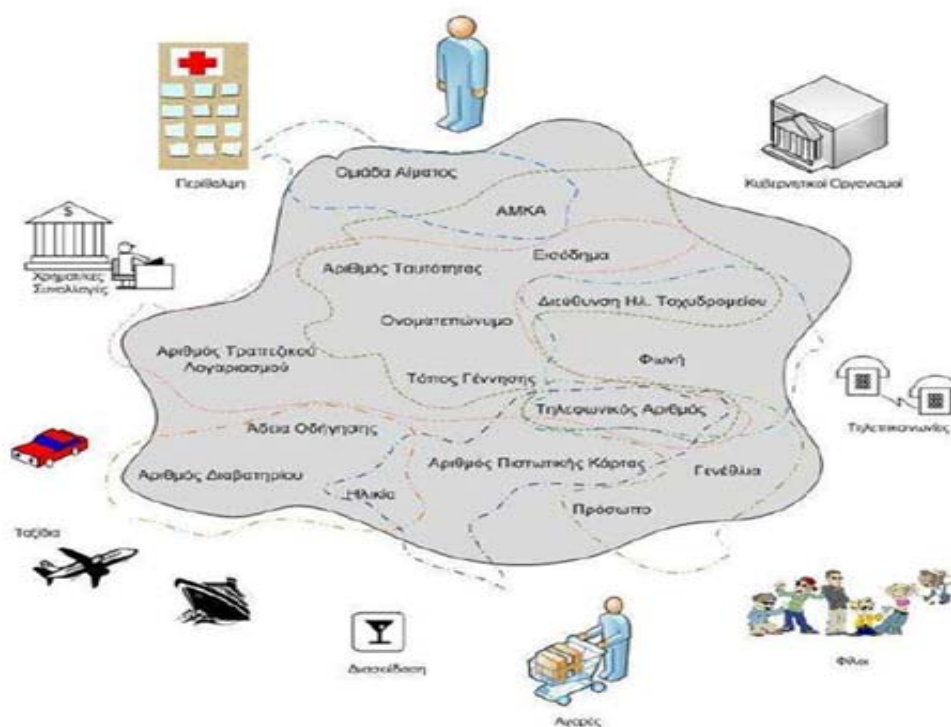
Μια δεύτερη κατηγοριοποίηση της ταυτότητας παρουσιάζεται στα Τρία Στρώματα Ταυτότητας [2]. Στο μοντέλο αυτό ο Durand διακρίνει τρεις κατηγορίες (στρώματα) ταυτοτήτων:

Η προσωπική ταυτότητα (η εσωτερική και παντοτινή ταυτότητα). Αυτή είναι η παντοτινή και προσωπική ταυτότητα που ανήκει και ελέγχεται αποκλειστικά από το άτομο.

Η συλλογική ταυτότητα (ανατεθειμένη ταυτότητα). Η ταυτότητα αυτή σχετίζεται με ένα συγκεκριμένο πλαίσιο (για παράδειγμα μια επιχειρηματική σχέση) και αναπαριστά ένα προσωρινά ανατεθειμένο και εκδομένο χαρακτηριστικό για το άτομο όπως ο εργασιακός του τίτλος, το τηλεφωνικό του νούμερο κλπ.

Η προωθητική ταυτότητα (αφηρημένη ή ενοποιημένη ταυτότητα). Αυτή η ταυτότητα αφορά στο αποτέλεσμα της ανάλυσης του προφίλ ενός ατόμου. Το άτομο δεν θεωρείται πραγματικά ως ένα πρόσωπο (το άτομο δεν έχει όνομα), αλλά ως το αποτέλεσμα φιλτραρίσματος που λαμβάνει χώρα επί ενός συνόλου χαρακτηριστικών. Ένα παράδειγμα θα μπορούσε να είναι ο πελάτης ο οποίος είναι νεαρής ηλικίας, έχει πανεπιστημιακή εκπαίδευση, είναι κάτοχος μοτοσυκλέτας παλαιότητας μεγαλύτερης των τριών ετών και παίζει τένις, με τον οποίο επικοινωνεί ο πωλητής.

Ενώ το μοντέλο αυτό μπορεί να φαίνεται πολύ απλουστευμένο για να ενσωματώνει όλη την πολυπλοκότητα της έννοιας της Ταυτότητας, εισάγει βασικές ιδιότητες στην Ταυτότητα και πιο συγκεκριμένα την ιδιότητα του προσωρινού, την ιδιότητα της εξάρτησης και την ιδιότητα του συγκεκριμένου.



Εικόνα 7: Η ταυτότητα και οι διαφορετικές υπό-ταυτότητες ανά εφαρμογή

### 3.2 Η έννοια της Ταυτοποίησης

Στην προηγούμενη παράγραφο παρουσιάστηκε η περιγραφική όψη της Ταυτότητας, δηλαδή μια εννοιολογικοποίηση που βασίζεται σε ένα σύνολο από ξεχωριστά χαρακτηριστικά τα οποία ανήκουν σε ένα αντικείμενο και τα οποία αποτελούν την ταυτότητά του. Στην τρέχουσα παράγραφο θα παρουσιαστεί η διαδικαστική όψη της Ταυτότητας, με άλλα λόγια οι διαδικασίες με τις οποίες η ταυτότητα ενός αντικειμένου συνδέεται και συγκεκριμένα πως η πληροφορία της ταυτότητας αποκαλύπτεται και χρησιμοποιείται. Έτσι, καλύπτει πλήθος εννοιών όπως η ανωνυμία, η μη συνδεσιμότητα και οι ταυτοποιητές.

Η Ταυτοποίηση είναι ως επί το πλείστον μερική (αποκάλυψη μόνο μιας μερικής ταυτότητας) και χρησιμοποιείται σε ένα συγκεκριμένο πλαίσιο και για ένα συγκεκριμένο σκοπό, όπως είναι η εκχώρηση πρόσβασης σε ένα πόρο ή μια λειτουργία. Θα πρέπει να σημειωθεί ότι και η σύνδεση με προφίλ (profiling) αντιπροσωπεύει μια έννοια η οποία μπορεί να σχετίζεται με την ταυτοποίηση. Ο όρος αυτός αφορά στη διαδικασία κατασκευής ή εφαρμογής ενός προφίλ ενός ατόμου ή μιας ομάδας. Ένα προφίλ αποτελείται από πρότυπα συσχετισμένων δεδομένων [3].

### 3.2.1 Η χρησιμότητα της ταυτοποίησης

Τα κύρια πλαίσια εντός των οποίων η ταυτοποίηση βρίσκει χρησιμότητα είναι τα ακόλουθα:

Έλεγχος πρόσβασης σε περιορισμένους πόρους ή περιοχές.  
Εκμετάλλευσης της πληροφορίας ταυτότητας.  
Παρακολούθηση για απόδοση ευθυνών.

#### *Έλεγχος πρόσβασης σε περιορισμένους πόρους ή περιοχές (πιστοποίηση)*

Ο έλεγχος αφορά τόσο στην πιστοποίηση όσο και στην διαχείριση πρόσβασης. Η πιστοποίηση σχετίζεται με την επιβεβαίωση της ταυτότητας ενός ατόμου και συγκεκριμένα με την διασφάλιση ότι το άτομο είναι όντως αυτό που ισχυρίζεται ότι είναι. Η διαχείριση του δικαιώματος πρόσβασης σχετίζεται με τον ρόλο του ατόμου εντός ενός συγκεκριμένου πλαισίου και τις κατηγορίες των λειτουργιών που έχουν ανατεθεί και οι οποίες επιτρέπονται στο άτομο (για παράδειγμα μπορεί να υφίσταται κάποιος περιορισμός σχετικά με τη χρήση ενός πόρου ή τις επιτρεπόμενες λειτουργίες σε μια συγκεκριμένη περιοχή).

Από την οπτική γωνία του χρήστη, η ταυτοποίηση μπορεί να εφαρμοσθεί σε ένα άλλο άτομο ή ένα οργανισμό. Στοιχεία ταυτότητας μπορούν επίσης να χρησιμοποιηθούν από ένα άτομο προκειμένου να διασφαλίσει ότι ένα άλλο άτομο ή οργανισμός είναι όντως αυτό που ισχυρίζεται ότι είναι. Έτσι, για παράδειγμα, μια διεύθυνση ηλεκτρονικού ταχυδρομείου παρέχει κάποιου βαθμού ταυτοποίηση σχετιζόμενη με τον αποστολέα, ενώ ο Ενιαίος Εντοπιστής Πόρων (Unified Resource Locator - URL) μπορεί να παρέχει κάποια μέσα για την πιστοποίηση ενός οργανισμού.

#### *Αναγνώριση με σκοπό την εκμετάλλευση πληροφορίας της ταυτότητας (γνώση)*

Ένας επίσης σημαντικός λόγος για την ταυτοποίηση ενός ατόμου είναι να επιτρέψει την πρόσβαση σε σχετική πληροφορία η οποία μπορεί να είναι απαραίτητη για το σύστημα, αλλά και να προσδώσει μια σειρά από πλεονεκτήματα στο άτομο, όπως για παράδειγμα επιτρέποντας πιο προσαρμοσμένη και αποτελεσματική αλληλεπίδραση του συστήματος και του ατόμου. Σε άλλες περιπτώσεις, η πρόσβαση σε αυτήν την πληροφορία αποτελεί υποχρεωτική συνθήκη για την παράδοση μιας συγκεκριμένης υπηρεσίας.

Αντιστρόφως, η πρόσβαση στην πληροφορία αυτή μπορεί να αποδειχθεί επιζήμια για το άτομο. Για παράδειγμα, εμπορικές εταιρείες άμεσης προώθησης μπορούν να εκμεταλλευτούν την πληροφορία αυτή προκειμένου να χειριστούν πιο αποτελεσματικά προς όφελός τους τον τελικό καταναλωτή στον οποίο στοχεύουν, με την πληροφορία αυτή να προσφέρει ένα άδικο πλεονέκτημα έναντι του ατόμου κατά τη διάρκεια μιας διαπραγμάτευσης.



### 3.2.2 Οι κίνδυνοι της ταυτοποίησης

Είναι σημαντικό να αναφερθεί ότι η Ταυτοποίηση, ανεξαρτήτως της χρήσης της, μπορεί να συνοδεύεται από σημαντικές αρνητικές επιπτώσεις για το άτομο είτε λόγω λανθασμένης ταυτοποίησης είτε λόγω μη επιθυμητής ταυτοποίησης.

#### *Λανθασμένη Ταυτοποίηση*

Οι περισσότερες περιπτώσεις λανθασμένης ταυτοποίησης αφορούν σε σφάλματα των συστημάτων αναγνώρισης ή σε επιτυχημένες προσπάθειες κλοπής η πλαστογράφησης ταυτότητας (identity theft – identity fraud). Η πρώτη αιτία των σφαλμάτων οφείλεται στις εγγενείς αδυναμίες των συστημάτων αναγνώρισης σε τεχνικό επίπεδο, σε σφάλματα κατά την καταχώρηση των δεδομένων ή σε απροσεξία σε επίπεδο ανθρώπου. Εστιάζοντας στις περιπτώσεις κλεμμένης ή πλαστής ταυτότητας, οι επιπτώσεις μιας τέτοιας ενέργειας μπορεί να έχει κυμαινόμενες επιπτώσεις αναλόγως με την εφαρμογή και το είδος της πληροφορίας η οποία χρησιμοποιείται για την ταυτοποίηση αλλά και για την πληροφορία η οποία συνδέεται με την ταυτοποίηση αυτή. Από τις πιο σοβαρές επιπτώσεις αυτών είναι η αποκάλυψη εμπιστευτικής πληροφορίας σε τρίτους ή η απώλεια μεγάλων χρηματικών ποσών (μέσω για παράδειγμα αν μια επιχείρηση εξαπάτησης πείσει ένα άτομο να αποκαλύψει της πληροφορίες της πιστωτικής του κάρτας). Παραδείγματα κλοπής ταυτότητας αποτελούν η κλοπή των κωδικών πρόσβασης και η κλοπή εγγράφων ταυτότητας μεταξύ των άλλων. Περιπτώσεις απάτης ταυτότητας αποτελούν η πανομοιότυπη αναπαραγωγή βιομετρικών δεδομένων και το μάντεμα των κωδικών πρόσβασης.

#### *Μη επιθυμητή ταυτοποίηση*

Η μη επιθυμητή ταυτοποίηση αφορά στις περιπτώσεις κατά τις οποίες πληροφορία σχετική με το άτομο αποκαλύπτεται σε τρίτους χωρίς τη συναίνεσή του ή / και τη γνώση του. Για τον λόγο αυτόν, αυτός ο κίνδυνος είναι άμεσα συνδεδεμένος με την ιδιωτικότητα του ατόμου. Παραδείγματα απαντώνται σε διάφορους τομείς της ανθρώπινης δραστηριότητας. Έτσι στις εμπορικές συναλλαγές και κυρίως στις ηλεκτρονικές, η αποκάλυψη πληροφορίας σε τρίτους για το άτομο χωρίς τη συναίνεσή του μπορεί να χρησιμοποιηθεί για την ενίσχυση της στοχευόμενης προώθησης ενός προϊόντος μέσω της χειραγώγησής του. Στον εργασιακό χώρο, πληροφορία για το άτομο σχετικά με τις πολιτικές του πεποιθήσεις τη συμμετοχή του σε μια οργάνωση, το ιατρικό του ιστορικό μεταξύ των άλλων μπορεί να επιφέρει σημαντικές επιπτώσεις στο άτομο, συμπεριλαμβανομένων της εργασιακής πίεσης, της διάκρισης, ακόμα και της απώλειας της εργασίας. Όσον αφορά στον κοινωνικό χώρο, τόσο η διαρροή των πολιτικών πεποιθήσεων του ατόμου όσο και των σεξουαλικών του προτιμήσεων μπορεί να έχει επίσης σημαντικό αρνητικό αντίκτυπο.

Είναι σημαντικό σε αυτό το σημείο να τονίσουμε ότι αυτού του είδους ο κίνδυνος ταυτοποίησης προκύπτει από την αποκάλυψη τόσο των στοιχείων τα οποία είναι άμεσα συνδεδεμένα με τις λειτουργίες του συστήματος το οποίο και την πραγματοποιεί όσο και στοιχείων από άλλα συστήματα και βάσεις δεδομένων με τα οποία επιτρέπει σύνδεση. Η δεύτερη περίπτωση είναι εφικτή

όταν το αναγνωριστικό για το άτομο είναι κοινό στα συστήματα τα οποία μοιράζονται πληροφορίες όπως ο Αριθμός Φορολογικού Μητρώου (ΑΦΜ), ο Αριθμός Δελτίου Ταυτότητας (ΑΔΤ), ο Αριθμός Μητρώου Κοινωνικής Ασφάλισης (ΑΜΚΑ). Η νομιμότητα και αποδοχή τέτοιων πράξεων εξαρτάται από την αναγκαιότητα αυτών καθώς και από το νομικό υπόβαθρο το οποίο μπορεί να τις στηρίζει και να τις καταδικάζει. Έτσι, για παράδειγμα η Τράπεζα της Ελλάδος στην προσπάθειά της για πρόληψη και καταστολή της νομιμοποίησης εσόδων από εγκληματικές δραστηριότητες και της χρηματοδότησης της τρομοκρατίας στα πλαίσια της 3<sup>ης</sup> Κοινοτικής Οδηγίας καθορίζει μια σειρά ενεργειών για τους Χρηματοπιστωτικούς Οργανισμούς και τα Πιστωτικά Ιδρύματα η οποία προϋποθέτει την συνεργασία των τελευταίων προκειμένου να εντοπίζουν, προλαμβάνουν, αποτρέπουν και αποφέρουν συναλλαγές τέτοιου είδους [4].

### 3.2.3 Διαδικασίες Ταυτοποίησης

Με βάση τα παραπάνω διακρίνουμε δύο προσεγγίσεις για την ταυτοποίηση ενός ατόμου, τη φανερή (explicit) και την υποκρυπτόμενη (implicit) ταυτοποίηση [5]. Όπως φανερώνουν και οι όροι, η πρώτη περίπτωση αφορά στην ταυτοποίηση του ατόμου για την οποία το άτομο έχει γνώση και ίσως ακόμη και να συμμετέχει σε αυτήν, ενώ η δεύτερη αφορά στην ταυτοποίηση του ατόμου δίχως το τελευταίο να έχει γνώση και να έχει δώσει τη συναίνεσή του. Η εν γνώσει ταυτοποίηση του ατόμου μπορεί να λάβει χώρα μέσω της παρουσίασης εκ μέρους του ατόμου ή ενός εξουσιοδοτημένου ατόμου του δελτίου ταυτότητας, της επαγγελματικής του κάρτας, της πιστωτικής του κάρτας ή των βιομετρικών του χαρακτηριστικών μεταξύ των άλλων. Επίσης, στην κατηγορία αυτή εμπίπτουν οι μηχανισμοί συλλογής των δεδομένων ταυτότητας, του ατόμου κατά τη φάση αρχικής καταχώρησης η οποία λαμβάνει χώρα με τη συναίνεση του ατόμου.

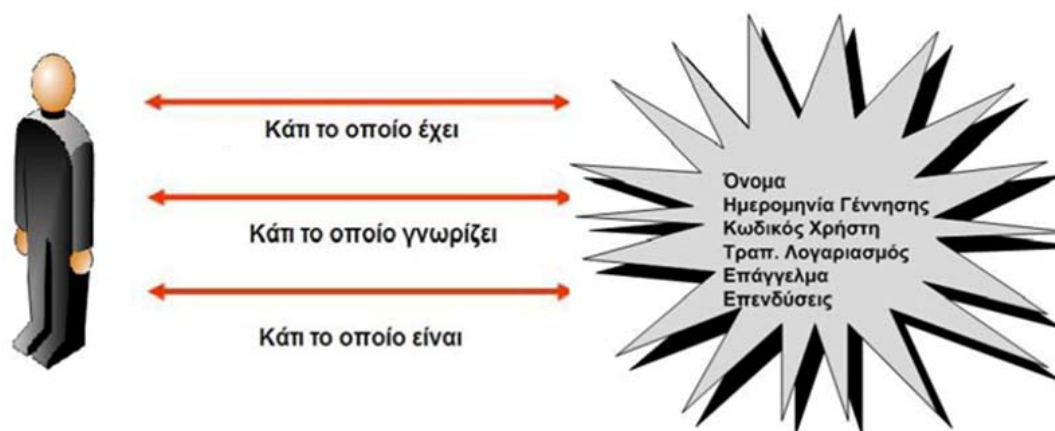
Στην περίπτωση της υποκρυπτόμενης ταυτοποίησης, η οποία είναι και άμεσα συνδεδεμένη με τους κινδύνους μη επιθυμητής ταυτοποίησης οι οποίοι παρουσιάστηκαν στην προηγούμενη παράγραφο, τόσο η διαδικασία ταυτοποίησης όσο και η λήψη πληροφορίας ταυτότητας λαμβάνουν χώρα χωρίς τη συγκατάθεση ή/και τη γνώση του ατόμου. Έτσι, για παράδειγμα με βάση αναγνωριστικά του ατόμου, όπως είναι οι κωδικού του, η φυσική του παρουσία σε ένα ή περισσότερους χώρους, είναι δυνατή η συλλογή των ιχνών της δράσης του ατόμου αυτού και η εξαγωγή συμπερασμάτων και πληροφοριών για την ταυτότητα και τη δραστηριότητα αυτού εν αγνοία του.

### 3.2.4 Χαρακτηριστικά προς Ταυτοποίηση

Η διαδικασία ταυτοποίησης ενός ατόμου μπορεί να βασίζεται σε πληθώρα χαρακτηριστικών του τα οποία γενικώς διακρίνονται σε:

- Κάτι που το άτομο γνωρίζει: αφορά σε πληροφορία την οποία το άτομο γνωρίζει, όπως είναι ο προσωπικός του κωδικός για την πρόσβασή του σε ένα σύστημα.
- Κάτι που το άτομο έχει: περιλαμβάνει περιπτώσεις κατά τις οποίες το άτομο είναι κάτοχος ενός αντικειμένου αναγνωριστικού, όπως είναι μια κάρτα, ψηφιακά πιστοποιητικά, κλπ.
- Κάτι που το άτομο είναι ή κάνει: σχετίζεται με γνωρίσματα με τα οποία το άτομο είναι άμεσα συνδεδεμένα, με χαρακτηριστικό παράδειγμα τα βιομετρικά του χαρακτηριστικά.

Όλα αυτά τα χαρακτηριστικά διαφέρουν ως προς την πρακτικότητά τους, το κόστος εφαρμογής τους, το πεδίο εφαρμογής τους, την αξιοπιστία τους, την αποτελεσματικότητά τους, την ασφάλεια τους αλλά και τον σεβασμό τους προς την ιδιωτικότητα του ατόμου.



Εικόνα 8: Χαρακτηριστικά προς ταυτοποίηση

### 3.3 Ιστορική Ανασκόπηση Βιομετρικής Ταυτοποίησης και Επιβεβαίωσης Ταυτότητας

Η χρήση των βιομετρικών συστημάτων δεν αποτελεί κάποια νέα μέθοδο ταυτοποίησης και επιβεβαίωσης της ταυτότητας των ατόμων. Έτσι, από τις απαρχές του πολιτισμού οι άνθρωποι βασίζονταν στο ανθρώπινο πρόσωπο προκειμένου να διακρίνουν γνωστά και μη άτομα. Με τα χρόνια η διαδικασία αυτή έγινε όλο και πιο πολύπλοκη καθώς ο πληθυσμός αυξανόταν και καθώς η εφεύρεση νέων μεθόδων μεταφοράς έφερνε νέα άτομα στις κάποτε μικρές κοινωνίες.

Στο πέρασμα των αιώνων πληθώρα ανθρώπινων χαρακτηριστικών χρησιμοποιήθηκαν ως μια πιο επίσημη μορφή ταυτοποίησης. Συγκεκριμένα, σε μια σπηλιά η οποία χρονολογείται ότι είναι παραπάνω από 31000 ετών, βρέθηκαν τοιχογραφίες οι οποίες εκτιμάται ότι δημιουργήθηκαν από προϊστορικούς ανθρώπους οι οποίοι έμεναν εκεί. Γύρω από τις τοιχογραφίες αυτές βρέθηκαν διάφορα αποτυπώματα ανθρώπινης παλάμης τα οποία θεωρείται ότι αποτελούσαν “υπογραφή” του δημιουργού τους [6]. Στοιχεία επίσης συνηγορούν ότι η χρήση των αποτυπωμάτων ως μέσο ταυτοποίησης χρονολογείται στα 500 π.Χ. όταν οι Βαβυλώνιοι αποτύπωναν τις συναλλαγές τους πάνω σε πλάκες από πυλό οι οποίες συμπεριελάμβαναν δακτυλικά αποτυπώματα [7]. Ο Ισπανός εξερευνητής και συγγραφέας Joao de Barros κατέγραψε μεταξύ των άλλων ότι η παρόμοια χρήση των δακτυλικών αποτυπωμάτων απαντάται και στην Κίνα του 14<sup>ου</sup> αιώνα οπότε και οι έμποροι τα χρησιμοποιούσαν κατά τις εμπορικές τους συναλλαγές. Μεταξύ των άλλων βασίζονταν σε αυτό το βιομετρικό χαρακτηριστικό καθώς και το αποτύπωμα του ποδιού τα οποία ελάμβαναν μέσω μελανιού για να ξεχωρίζουν τα παιδιά [8]. Σύμφωνα με τον ίδιο συγγραφέα, οι αρχαίοι Αιγύπτιοι έμποροι βασίζονταν σε φυσικούς περιγραφείς για τον διαχωρισμό των εμπιστων εμπόρων γνωστής φήμης και με προγενέστερες επιτυχείς συναλλαγές από νέους εμπόρους.

Στα μεταγενέστερα χρόνια η χρήση βιομετρικών χαρακτηριστικών παρατηρήθηκε πιο έντονα. Το 1980 ο Alphonse Bertillone, ένας Γάλλος αστυνόμος μελέτησε τη μηχανική του σώματος καθώς

και μετρήσεις προκειμένου να καταστεί δυνατή η σύλληψη εγκληματιών [9]. Η αστυνομία βασίστηκε στη μέθοδο αυτή, την επονομαζόμενη Bertillonage μέθοδο, μέχρι τη στιγμή που αναγνώρισε λανθασμένα κάποιους ανθρώπους. Η μέθοδος αυτή γρήγορα εγκαταλείφθηκε, ενώ ξαναεισήχθη από τον Richard Edward Henry, της Scotland Yard.

Στις αρχές του 20<sup>ου</sup> αιώνα ο Karl Pearson, ένας ερευνητής εφαρμοσμένων μαθηματικών μελέτησε τη βιομετρία στο University College of London. Κατά τη διάρκεια της έρευνάς του οδηγήθηκε σε σημαντικές ανακαλύψεις στο χώρο αυτό μέσω της μελέτης στατιστικής ιστορίας και συσχέτισης, την οποία και εφάρμοσε την εξέλιξη των ζώων. Το έργο του αυτό περιελάμβανε τη μέθοδο των στιγμών, το Pearson σύστημα των καμπύων, συσχέτιση και την  $\chi^2$  δοκιμή.

Αργότερα, κατά τις δεκαετίες 1960 και 1970, διαδικασίες βιομετρικής ταυτοποίησης μέσω της υπογραφής αναπτύχθηκαν, αλλά ο βιομετρικός χώρος έμεινε σταθερός μέχρι τη στιγμή που τα στρατιωτικά γραφεία και τα γραφεία ασφαλείας προχώρησαν στην έρευνα και στην ανάπτυξη της βιομετρικής τεχνολογίας πέραν των δακτυλικών αποτυπωμάτων.

Τα τελευταία χρόνια και με διαρκώς αυξανόμενη την προσπάθεια για βελτίωση των συστημάτων ταυτοποίησης και επιβεβαίωσης ταυτότητας τα οποία κυρίως βασίζοντας σε κωδικούς και κάρτες, συνοδευόμενα από σειρά προβλημάτων, παρατηρείται μια έντονη δραστηριότητα στο χώρο των βιομετρικών συστημάτων. Μία σημαντική εφαρμογή αυτών αποτέλεσε το Super Bowl της Tampa, Florida, το 2001 [10] κατά το οποίο έγινε χρήση συστήματος αναγνώρισης προσώπου με κάθε εικόνα προσώπου από τους 100000 θεατές να έχει καταγραφεί από κάμερες ασφαλείας και να έχει ελεγχθεί ηλεκτρονικά σε σχέση με εικόνες κακοποιών από την αστυνομία της Tampa. Αυτή η κίνηση συνδυαζόμενη από τη μη επιτυχή λειτουργία του συστήματος προκάλεσε πληθώρα αντιδράσεων από υποστηρικτές των κοινωνικών ελευθεριών οι οποίοι κατήγγειλαν τότε τις βιομετρικές τεχνολογίες ως εχθρούς της ελευθερίας του ατόμου.

Παρ' όλα αυτά και ειδικά μετά από το τρομοκρατικό κτύπημα της 11<sup>ης</sup> Σεπτεμβρίου 2001 στις ΗΠΑ, πολλές αρχές παγκοσμίως προχώρησαν στην εγκατάσταση βιομετρικών τεχνολογιών στα αεροδρόμια προκειμένου να εντοπίσουν υπόπτους για τρομοκρατία. Λίγα χρόνια αργότερα και μετά από πληθώρα βομβιστικών επιθέσεων στον υπόγειο σιδηρόδρομο, η Μεγάλη Βρετανία υιοθέτησε τεχνολογίες αναγνώρισης προσώπου καταλήγοντας σήμερα το Λονδίνο να έχει περισσότερες από 200000 κάμερες σε δημόσιους χώρους για λόγους ασφαλείας.

### **3.4 Στάδια Αυτόματης Αναγνώρισης Προσώπων**

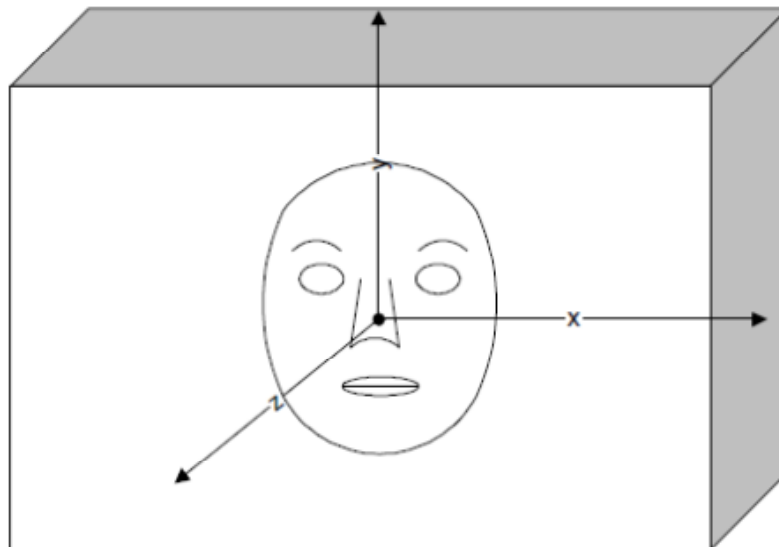
Η διαδικασία της αναγνώρισης μπορεί να διαιρεθεί σε τρία διαδοχικά στάδια. Σε μια δοσμένη εικόνα ερευνάται, αν αυτή απεικονίζει κάποιο πρόσωπο, το οποίο και εξάγεται από τον περιβάλλοντα χώρο. Στη συνέχεια έχουμε αναπαράσταση του προσώπου με την εξαγωγή διαφόρων χαρακτηριστικών ώστε να μειωθεί η διάσταση του προβλήματος. Τέλος ακολουθεί η διαδικασία της ταυτοποίησης του προσώπου με κάποιο υπάρχον στη βάση ή η προσθήκη του σε αυτήν.

Στο κεφάλαιο αυτό πραγματοποιείται μια ανασκόπηση των τεχνικών που έχουν εμφανιστεί στη βιβλιογραφία σχετικά με τα τρία επιμέρους προβλήματα της αναγνώρισης προσώπων. Σε κάποιες περιπτώσεις δίνονται και μερικές προσομοιώσεις, για βελτίωση της κατανόησης των αλγορίθμων αλλά και για να δοθεί η ευκαιρία να τονιστούν και να σχολιαστούν κάποια κρίσιμα ζητήματα.

Η έρευνα στην αναγνώριση προσώπων ξεκίνησε από τα μέσα του περασμένου αιώνα από τον Francis Galton. Πρέπει να τονιστεί ότι στη γενική περίπτωση κανένα από τα επιμέρους προβλήματα που σχετίζονται με την αναγνώριση προσώπων δεν έχει λυθεί και πρακτικά μόνο τα προβλήματα της ανίχνευσης και της ταυτοποίησης έχουν προσελκύσει ουσιαστικό ενδιαφέρον από τους ερευνητές. Στη περίπτωση της ταυτοποίησης η θέση του προσώπου είναι εκ των προτέρων γνωστή ή μπορεί να εκτιμηθεί με ευριστικό τρόπο. Σε πολλές περιπτώσεις ισχυρές παραδοχές τίθενται για να διευκολύνουν το πρόβλημα της ταυτοποίησης:

- (1). Το απεικονιζόμενο πρόσωπο είναι σε εμπρόσθια όψη ή σε πλάγια όψη (προφίλ). Με την παραδοχή αυτή είναι ευκολότερο να εντοπιστούν τα βασικά χαρακτηριστικά του.
- (2). Το πρόσωπο έχει μηδενική ή πολύ μικρή κλίση (περί τον άξονα  $z$  που περνά από το κέντρο της φωτογραφίας - βλ. Εικόνα 9).
- (3). Δεν υπάρχει επικάλυψη τμήματος του προσώπου.
- (4). Τα μαλλιά δεν επικαλύπτουν τμήμα του προσώπου, το πρόσωπο δεν φορά σκούρα γυαλιά και δεν υπάρχουν ουλές σε αυτό.

Πρέπει να αναφερθεί ότι οι περισσότερες από τις βάσεις δεδομένων που έχουν χρησιμοποιηθεί αποτελούνται από λευκούς άνδρες και περιλαμβάνουν σχετικά μικρό αριθμό φωτογραφιών.



**Εικόνα 9:** Σύστημα συντεταγμένων με βάση τη φωτογραφία

Αναφορά	Πλήθος Φωτογραφιών	Ανάλυση Χρώματος	Κβαντισμός	Σχόλια
Baron	150	512x480	6 bpp	Η ανάλυση μειώθηκε σε 128x120
Buhr	100	512x512	8 bpp	
Campell	50	256x256	6 bpp	
Craw et al	1000	128x128	-	Ο έλεγχος έγινε σε μόνο 20 πρόσωπα
Goldstein et al	255	-	-	Προφίλ τα οποία σχεδιάστηκαν από Ζωγράφους
Harmon et al	124	-	-	97 άνδρες +27 γυναίκες
Harmon et al	130	-	-	81 άνδρες +49 γυναίκες
Kaufman	120	240x356	1 bpp	-
Nixon	6	128x128	8 bpp	-
Petajan	-	244x248	-	Για τον εντοπισμό χειλιών μόνο
Riccia	506	-	-	Τα δεδομένα εισάγονται από τον χρήστη
Wong	6	256x256	8 bpp	-
Wu and Huang	18	-	-	-
Olivetti R.L.	400	112x92	8 bpp	40 πρόσωπα από 10 φωτογραφίες το καθένα

**Πίνακας 2:** Δεδομένα που χρησιμοποιήθηκαν σε διάφορες μελέτες

Στον Πίνακα 1 δίνονται κάποια στοιχεία για τις βάσεις εικόνων και τη μορφή των δεδομένων που χρησιμοποιήθηκαν, σε διάφορες εργασίες. Όπου λείπουν στοιχεία αυτά δεν αναφέρονται στη συγκεκριμένη εργασία. Σε πολλές περιπτώσεις τα δεδομένα εισόδου στο σύστημα δημιουργήθηκαν χειρωνακτικά (manually).

### 3.4.1 Αναπαράσταση

Δύο είδη αναπαραστάσεων χρησιμοποιούνται στη αναγνώριση και ταυτοποίηση προσώπων: Δισδιάστατοι πίνακες τιμών και διανύσματα χαρακτηριστικών (τα χαρακτηριστικά εδώ δεν αναφέρονται στα χαρακτηριστικά του προσώπου αλλά σε κάποιες τιμές οι οποίες περιγράφουν με αποτελεσματικό τρόπο τη δομή του).

### 3.4.2 Αναπαράσταση προσώπων - Εξαγωγή χαρακτηριστικών

Αποδοτική αναπαράσταση ενός προσώπου θεωρείται κάθε αναπαράσταση η οποία μειώνει τη διάσταση της εικόνας διατηρώντας όσο το δυνατό μεγαλύτερο όγκο πληροφορίας. Σε πολλές από τις αρχικές μελέτες η αναπαράσταση των προσώπων γινόταν με χρήση γεωμετρικών χαρακτηριστικών, όπως αποστάσεων, γωνιών και λόγων αποστάσεων μεταξύ βασικών σημείων του προσώπου. Τέτοια σημεία έπρεπε να μπορούν να εντοπίζονται εύκολα και με ακρίβεια. Μια τέτοια αντιμετώπιση δε λαμβάνει υπόψη πληροφορία από την υφή του προσώπου, γεγονός που οδήγησε στην επέκταση του

συνόλου των χρησιμοποιούμενων χαρακτηριστικών συμπεριλαμβάνοντας σε αυτά και πληροφορία υφής, είτε με άμεση μορφή (πίνακες φωτεινότητας διαφόρων περιοχών), είτε με χρήση στατιστικών μεγεθών (moment invariants), είτε με τη βοήθεια κάποιων μετασχηματισμών (Fourier descriptors).

Αξίζει να σημειωθεί ότι τελικά με τον όρο χαρακτηριστικά αναφερόμαστε σε κάθε αναπαράσταση του προσώπου σε μικρότερη διάσταση, είτε αυτή γίνεται με γεωμετρικά μεγέθη, είτε με πίνακες, είτε με στατιστικά μεγέθη. Με αυτό υπόψη μπορούμε να διακρίνουμε τέσσερις κατηγορίες [11]:

(α) Οπτικά χαρακτηριστικά.

Στη κατηγορία αυτή περιλαμβάνονται τα γεωμετρικά μεγέθη, οι πίνακες περιγραφής της υφής περιοχών, οι ακμές κ.ο.κ.. Η εξαγωγή τέτοιων χαρακτηριστικών βασίζεται σε παραδοσιακές τεχνικές της επεξεργασίας εικόνας όπως το χωρικό φιλτράρισμα, η ταύτιση με πρότυπο, τα ενεργά περιγράμματα –active contours- κ.ο.κ. Στον Πίνακα 3 παρουσιάζονται τέτοια χαρακτηριστικά, τα περισσότερα από τα οποία έχουν χρησιμοποιηθεί κατά κόρον σε παλαιότερες μελέτες. Τονίζεται πάντως ότι πολλά από αυτά χρησιμοποιήθηκαν σε μελέτες στις οποίες η ταύτιση πραγματοποιείται από ανθρώπους και όχι από υπολογιστές.

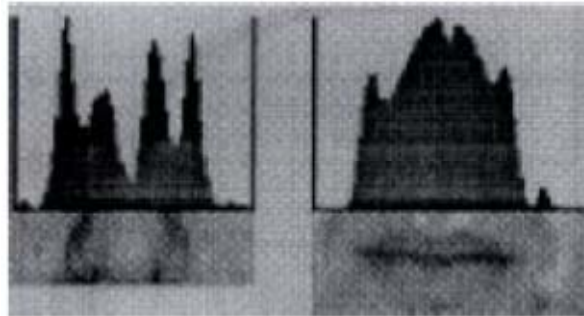
Κατηγορία	Χαρακτηριστικά
Συνολικά	Ύψος, περίγραμμα, εμβαδόν, κέντρο βάρους
Μάτια	Σχήμα, χρώμα, εμβαδόν, άνοιγμα, κέντρο, φωτεινότητα γύρω από την ίριδα, D(άνω βλέφαρο, κάτω βλέφαρο), D(εσωτερική γωνία ματιού, εξωτερική γωνία ματιού), D(κέντρο αριστερού ματιού, κέντρο δεξιού ματιού), D(εξωτερική γωνία αριστερού ματιού, εξωτερική γωνία δεξιού ματιού)
Βλέφαρα	Εμβαδόν, απόσταση από την καμπύλη του ματιού
Φρύδια	Πάχος, D(μάτι, φρύδι)
Χείλη	Πάχος, σχήμα, πρότυπο, πλάτος
Στόμα	Πρότυπο, πλάτος, μήκος, εμβαδόν, άνοιγμα
Μύτη	Πρότυπο, σχήμα, πλάτος, μήκος, εμβαδόν ρουθουνιών
Αυτιά	Μήκος, σχήμα, εμβαδόν
Μαλλιά	Χρώμα, υφή, περιοχή κάλυψης προσώπου, μήκος
Μάγουλα	Φωτεινότητα, υφή
Αποστάσεις	Πλάτος προσώπου στο ύψος των ματιών, D(πηγούνι, γραμμή ματιών), D(πηγούνι, κέντρο χειλιών), D(μάτια, γραμμή μαλλιών), D(μάτια, κέντρο μύτης), D(κορυφή πηγουνιού, κέντρο προσώπου), D(εσωτερική κόχη ματιού, κέντρο προσώπου), D(κέντρο ματιού, κέντρο φρυδιού), D(κέντρο προσώπου, κέντρο φρυδιών), D(εσωτερική κόχη ματιού, κόχη στόματος), D(κέντρο ματιού, κέντρο στόματος), D(κορυφή μύτης, κέντρο στόματος)
Εμβαδά	A(εσωτερικές κόχες ματιών, κέντρο στόματος), A(κέντρα φρυδιών, κέντρο στόματος), A(κέντρα ματιών, κέντρο προσώπου)
Λόγοι	D(κέντρο προσώπου, άνοιγμα στόματος)/ D(πηγούνι, άνοιγμα στόματος) D(πηγούνι, άνοιγμα στόματος)/ D(κορυφή μύτης, άνοιγμα στόματος)
Άλλα	Περίγραμμα πηγουνιού, περίγραμμα μετώπου, μέγεθος μετώπου

**Πίνακας 3:** Χαρακτηριστικά που χρησιμοποιήθηκαν για τη διαδικασία σε μελέτες αναγνώρισης προσώπων.  $D(x,y)$ = Ευκλείδεια απόσταση των χαρακτηριστικών  $x$  και  $y$

(β) Στατιστικά Χαρακτηριστικά

Τέτοια χαρακτηριστικά είναι το ιστόγραμμα ολόκληρης της εικόνας ή περιοχών αυτής και οι ολοκληρωματικές προβολές εικόνων ακμών στον κάθετο και οριζόντιο άξονα. Έστω για παράδειγμα η υποεικόνα  $I(x,y)$  ορισμένη στη περιοχή  $[x_1, x_2] [y_1, y_2]$ . Οι ολοκληρωματικές προβολές κατά τον κάθετο και οριζόντιο άξονα δίνονται από τις σχέσεις:

$$V(x) = \sum_{y=y_1}^{y_2} I(x, y), \quad H(y) = \sum_{x=x_1}^{x_2} I(x, y), \quad (1)$$



**Εικόνα 10:** Ολοκληρωματικές προβολές

Εξαγωγή στατιστικών χαρακτηριστικών από περιοχές του προσώπου γίνεται επίσης με χρήση κεντρικών ροπών (central moments) και moment invariants. Τα moment invariants είναι γραμμικοί συνδυασμοί κανονικοποιημένων κεντρικών ροπών και έχουν το σημαντικό χαρακτηριστικό της αναισθησίας ως προς τη θέση, προσανατολισμό και αλλαγή κλίμακας [12].

(γ) Χαρακτηριστικά συντελεστών μετασχηματισμού

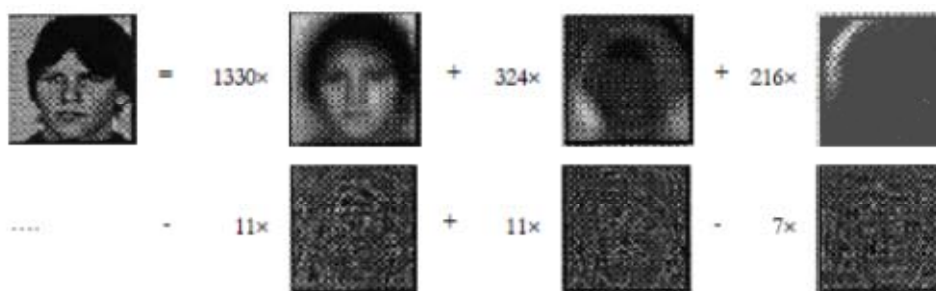
Η χρήση μετασχηματισμών όπως Fourier, Hadamard για περιγραφή ορίων περιοχών είναι αρκετά διαδεδομένη. Οι περιγραφείς Fourier και οι υπογραφές (signatures) είναι επίσης δύο τεχνικές που χρησιμοποιούνται ευρέως. Οι συντελεστές των μετασχηματισμών λαμβάνονται ως χαρακτηριστικά περιγραφής των επιμέρους περιοχών της εικόνας.

(δ) Αλγεβρικά Χαρακτηριστικά

Τα αλγεβρικά χαρακτηριστικά αναπαριστούν εσωτερικές ιδιότητες της εικόνας η οποία θεωρείται διδιάστατος πίνακας. Η τεχνική εξαγωγής τέτοιων χαρακτηριστικών βασίζεται στην αποσύνθεση πινάκων.

Οι πιο διαδεδομένες τεχνικές είναι ο μετασχηματισμός Karhunen Loeve -KL και η SVD -Singular Value Decomposition. Στο μετασχηματισμό KL κάθε εικόνα αναπτύσσεται με βάση τα ιδιοδιανύσματα του πίνακα μεταβλητότητας ο οποίος εκτιμάται πάνω σε ένα σύνολο από εικόνες προσώπων –βλέπε Εικόνα 11. Οι συντελεστές της ανάπτυξης αυτής αποτελούν τα αλγεβρικά χαρακτηριστικά με βάση τα οποία αναπαρίσταται η εικόνα.





Εικόνα 12: Αναπαράσταση προσώπου ως άθροισμα ιδιοδιανυσμάτων

### 3.4.3 Αναπαράσταση προσώπων με χρήση των τιμών της κλίμακας του γκρι

Η απλούστερη αναπαράσταση της φωτογραφίας ενός προσώπου είναι ο πίνακας με τις τιμές των στοιχείων της. Προφανώς μια τέτοια αναπαράσταση δεν είναι ιδιαίτερα συμπαγής αλλά είναι επιθυμητή όταν ο βασικός στόχος είναι η ευρωστία (δύναμη, ισχύς) του συστήματος. Σε συστήματα τα οποία χρησιμοποιούν άλλα χαρακτηριστικά για τη περιγραφή των προσώπων, αποθηκεύεται επίσης ο πίνακας με τις gray scale τιμές, έστω και αν δεν χρησιμοποιείται. Σε μεγάλες βάσεις δεδομένων με πολλές εικόνες, η αναπαράσταση με gray scale τιμές δεν είναι αποδοτική. Παρόλα αυτά, με δεδομένο ότι αναπαράσταση σε ανάλυση 32x32 και κβαντισμό 4 bpp είναι ικανοποιητική τόσο για την ανίχνευση όσο και την ταυτοποίηση, πολλές φορές οι πίνακες με τις gray scale τιμές υποδειγματοληπτούνται στη παραπάνω ανάλυση, επιτρέποντας με τον τρόπο αυτό αποδοτικότερη εφαρμογή των αλγορίθμων.

*Ο ρόλος της χωρικής ανάλυσης σε διαδικασίες ταύτισης με βάση τις gray scale τιμές*

Ο Πίνακας 2 δείχνει πειραματικά αποτελέσματα για τη διαδικασία αναγνώρισης με χρήση διαφόρων αναλύσεων και πέντε διαφορετικών μετρικών. Συγκεκριμένα έστω  $R1$  και  $R2$  δύο διαστάστοι πίνακες διάστασης  $k \times l$  που αντιστοιχούν στις gray scale τιμές δύο εικόνων. Ορίζουμε της παρακάτω μετρικές:

$$M1 = \max(SVD\left\{\left|\frac{R1}{\text{mean}(R1)} - \frac{R2}{\text{mean}(R2)}\right|\right\}) \Rightarrow \text{μέγιστη ιδιοτιμή του πίνακα των απόλυτων διαφορών μεταξύ } R1 \text{ και } R2$$

$$M2 = \sqrt{\sum \text{diag}\left\{\left|\frac{R1}{\text{mean}(R1)} - \frac{R2}{\text{mean}(R2)}\right|^2\right\}} \Rightarrow \text{η Frobenius νόρμα του πίνακα των απόλυτων διαφορών μεταξύ } R1 \text{ και } R2 \text{ οι οποίοι έχουν διαιρεθεί με τη μέση τιμή τους.}$$

$$M3 = \frac{1}{k \cdot l} \sum \left|\frac{R1}{\text{mean}(R1)} - \frac{R2}{\text{mean}(R2)}\right| \Rightarrow \text{άθροισμα των απόλυτων διαφορών -pixel προς pixel- των πινάκων } R1 \text{ και } R2$$

$M4 = \sqrt{\sum \text{diag}(|R1-R2|^T \cdot |R1-R2|)} \Rightarrow$  η Frobenius νόρμα του πίνακα των απόλυτων διαφορών μεταξύ R1 και R2

$M5 = \text{corr}(R1, R2) \Rightarrow$  η συσχέτιση των πινάκων R1 και R2

Η βάση δεδομένων που χρησιμοποιήθηκε είναι της Olivetti Research Laboratory (ORL) η οποία είναι μια από τις πλέον χρησιμοποιούμενες και συχνά χρησιμοποιείται ως βάση σύγκρισης διαφορετικών αλγορίθμων. Οι συγκρίσεις πραγματοποιήθηκαν με βάση τους πίνακες των gray scale τιμών και σε όλες τις περιπτώσεις ο κβαντισμός ήταν 8 bpp. Επιτυχής θεωρείται η αναγνώριση αν η προς έλεγχο φωτογραφία και η πλησιέστερη προς αυτήν αποθηκευμένη ανήκουν στο ίδιο πρόσωπο. Από τα αποτελέσματα προκύπτουν μερικά σημαντικά συμπεράσματα:

(1) Ακόμα και σε ανάλυση 28x23 η αναγνώριση είναι αποτελεσματική. Σημειώνεται πολύ μικρή μείωση στο βαθμό επιτυχίας του συστήματος -σε σχέση με τη μέγιστη ανάλυση.

(2) Η μέγιστη ανάλυση δεν δίνει υποχρεωτικά και τα καλύτερα αποτελέσματα σε όλες τις μετρικές. Αυτό είναι σύμφωνο και προς την ολοκληρωτική -και όχι διαφορική- συμπεριφορά του ματιού.

(3) Ο Μέσος Χρόνος Αναγνώρισης (MXA) είναι 1.51, 3.81 και 19.82 sec αντίστοιχα για τις τρεις διαφορετικές αναλύσεις και για τις πέντε μετρικές συνολικά και αναφέρεται σε βάση αποτελούμενη από 200 φωτογραφίες.

(4) Σχολιασμός της αποδοτικότητας των μετρικών γίνεται σε επόμενη παράγραφο.

Μετρική	M1	M2	M3	M4	M5
Αποτυχημένες Αναγνώρισεις	6	7	5	7	7
MXA (sec)	1.51				
Ανάλυση	28 x 23				
Μετρική	M1	M2	M3	M4	M5
Αποτυχημένες Αναγνώρισεις	4	5	3	5	7
MXA (sec)	3.81				
Ανάλυση	56 x 46				
Μετρική	M1	M2	M3	M4	M5
Αποτυχημένες Αναγνώρισεις	5	6	3	5	5
MXA (sec)	19.82				
Ανάλυση	112 x 92				

**Πίνακας 4:** Αναγνώριση προσώπων σε διάφορες αναλύσεις και με χρήση 5 διαφορετικών μετρικών

Μετρική	M1	M2	M3	M4	M5
Αποτυχημένες Αναγνώρισεις	5	6	3	4	7
Ανάλυση	28 x 23				
Μετρική	M1	M2	M3	M4	M5
Αποτυχημένες Αναγνώρισεις	3	3	2	5	5
Ανάλυση	56 x 46				
Μετρική	M1	M2	M3	M4	M5
Αποτυχημένες Αναγνώρισεις	3	3	1	4	5
Ανάλυση	112 x 92				

**Πίνακας 5:** Αναγνώριση προσώπων σε διάφορες αναλύσεις και με χρήση 5 διαφορετικών μετρικών. Επιτυχής θεωρείται η αναγνώριση αν η προς έλεγχο φωτογραφία και κάποια από τις τρεις πλησιέστερες προς αυτήν αποθηκευμένες ανήκουν στο ίδιο πρόσωπο.

Μετρική	M1	M2	M3	M4	M5
Αποτυχημένες Αναγνώρισεις	3	4	1	3	3
Ανάλυση	28 x 23				
Μετρική	M1	M2	M3	M4	M5
Αποτυχημένες Αναγνώρισεις	1	3	1	3	2
Ανάλυση	56 x 46				
Μετρική	M1	M2	M3	M4	M5
Αποτυχημένες Αναγνώρισεις	1	2	1	2	3
Ανάλυση	112 x 92				

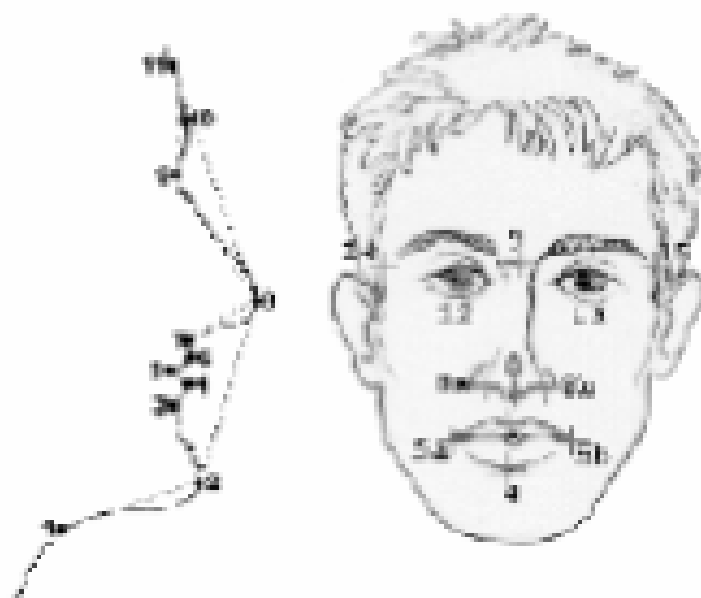
**Πίνακας 6:** Αναγνώριση προσώπων σε διάφορες αναλύσεις και με χρήση 5 διαφορετικών μετρικών. Επιτυχής θεωρείται η αναγνώριση αν η προς έλεγχο φωτογραφία και κάποια από τις πέντε πλησιέστερες προς αυτήν αποθηκευμένες ανήκουν στο ίδιο πρόσωπο.

Ο Πίνακας 4 δείχνει αποτελέσματα από το προηγούμενο πείραμα, μόνο που τώρα επιτυχημένη θεωρείται η αναγνώριση, αν η προς έλεγχο φωτογραφία και κάποια από τις τρεις πλησιέστερες προς αυτήν αποθηκευμένες ανήκουν στο ίδιο πρόσωπο. Αντίστοιχα, ο Πίνακας 5 δείχνει αποτελέσματα στα οποία επιτυχημένη θεωρείται η αναγνώριση, αν η προς έλεγχο φωτογραφία και κάποια από τις πέντε πλησιέστερες προς αυτήν αποθηκευμένες, ανήκουν στο ίδιο πρόσωπο. Τα αποτελέσματα των Πινάκων 4 και 5 δείχνουν την ευρωστία της αναγνώρισης για τις διάφορες μετρικές.

### 3.4.4 Γεωμετρικά χαρακτηριστικά από εικόνες εμπρόσθιας όψης και προφίλ

Η αναπαράσταση προσώπων με κάποιο διάνυσμα γεωμετρικών χαρακτηριστικών είναι η πιο διαδεδομένη πρακτική όταν χρησιμοποιούνται εικόνες προφίλ. Αντίθετα σε εικόνες εμπρόσθιας όψης η αναπαράσταση με τις τιμές της κλίμακας του γκρι είναι πιο διαδεδομένη. Πάντως η χρήση γεωμετρικών χαρακτηριστικών για την αναπαράσταση προσώπων έχει ανακτήσει νέο ενδιαφέρον με την ανάπτυξη του προτύπου MPEG-4.

Σε εικόνες προφίλ οι τιμές των χαρακτηριστικών παράγονται με βάση κάποια σημεία στη οριογραμμή του, όπως η εγκοπή ανάμεσα στα φρύδια και τη μύτη, η κορυφή της μύτης, η εγκοπή ανάμεσα στη μύτη και το άνω χείλος και η κορυφή του πηγουνιού. Τα χαρακτηριστικά είναι συνήθως αποστάσεις και γωνίες ανάμεσα στα χαρακτηριστικά αυτά σημεία. Ο Πίνακας 6 παρουσιάζει ένα σύνολο από συχνά χρησιμοποιούμενα γεωμετρικά χαρακτηριστικά και για τις δύο κατηγορίες εικόνων.



**Εικόνα 13:** Σημεία σε προφίλ και frontal views

Χαρακτηριστικά σε εικόνες profile	Περιγραφή	Χαρακτηριστικά σε εικόνες frontal view	Περιγραφή
Γωνία 1-2-3	Μέτρηση πηγουνιού	Λόγος $d(14,15) / d(4,9)$	Μέτρηση πλάτους προσώπου
Γωνία 7-8-9	Μέτρηση μύτης	Λόγος $d(12,13) / d(14,15)$	Απόσταση ματιών
Λόγος $d(8,10) / d(2,8)$ πηγούνι	Μακρύ — κοντό	Λόγος $d(5a, 5b) / d(14,15)$	Μέγεθος στόματος

**Πίνακας 7:** Χαρακτηριστικά για την περιγραφή του προσώπου (βλέπε Εικόνα 10). Η  $d(i,j)$  είναι η Ευκλείδεια απόσταση ανάμεσα στα σημεία  $i$  και  $j$ .

### 3.4.5 Υβριδική αναπαράσταση

Συχνά για την αναπαράσταση προσώπων χρησιμοποιείται συνδυασμός των gray scale τιμών και αλλά χαρακτηριστικών όπως αυτά των Πινάκων 3 και 7. Για παράδειγμα ο Campell [13] χρησιμοποίησε χαρακτηριστικά όπως η τιμή φωτεινότητας των μαλλιών και των παρειών σε συνδυασμό με ένα πίνακα 32x32 με τις gray scale τιμές της περιοχής των ματιών.

Στις πρώτες προσπάθειες για αναγνώριση προσώπων, οι οποίες κατά βάση στηρίζονταν σε εικόνες προφίλ από πορτρέτα, χρησιμοποιήθηκαν ευρέως τεχνικές τμηματικής προσέγγισης καμπυλών. Συγκεκριμένα ο Galton [14][15] δημιούργησε ένα σύστημα, στο οποίο πορτρέτα προφίλ περιγράφονται με χρήση ενός μαθηματικού τύπου. Κάθε τύπος αποτελείται από τέσσερις ομάδες εικόνων οι οποίες περιέχουν πέντε εικόνες η καθεμιά. Τα τμήματα των προφίλ κοντά στα χαρακτηριστικά σημεία, ή οι συνδέσεις των σημείων αυτών, περιγράφονται με τη χρήση απλών αριθμών. Οι αριθμοί αυτοί υπολογίζονται από κάποιο πίνακα στον οποίο διαφορετικά σχήματα καμπυλών περιγράφονται από συγκεκριμένους αριθμούς (δείκτες).

#### Εντοπισμός – Κατάτμηση προσώπου

Το πρώτο βήμα στη διαδικασία αναγνώρισης είναι ο εντοπισμός της επακριβούς θέσης του προσώπου στη φωτογραφία. Σε πολλές, παρόλα αυτά εργασίες, που πραγματεύονται την ταυτοποίηση προσώπων η θέση του προσώπου θεωρείται εκ των προτέρων γνωστή, ή μπορεί με ελάχιστη προσπάθεια να εντοπιστεί με βάση τη διαδικασία σύλληψης των φωτογραφιών. Τέτοια παραδείγματα εικόνων αποτελούν οι φωτογραφίες οι οποίες λαμβάνονται από τις αστυνομικές αρχές –mugshots.

Ο εντοπισμός προσώπων έχει προσεγγιστεί με δύο διαφορετικές τακτικές. Στη πρώτη προσέγγιση το πρόσωπο θεωρείται ως μια αυτόνομη μονάδα η οποία μοντελοποιείται με τεχνικές υπολογιστικής όρασης.

Στις εργασίες των Govindaraju et al. [11][12], για ανίχνευση προσώπων σε φωτογραφίες από εφημερίδες, το πρόσωπο μοντελοποιείται από δύο ευθείες, που αντιπροσωπεύουν τις πλευρές του προσώπου, και δύο τόξα, για το πηγούνι και το άνω μέρος του κεφαλιού. Για κάθε ένα από τα τέσσερα αυτά τμήματα ορίζονται τέσσερις παράμετροι: το μήκος του τμήματος, η χορδή που σχηματίζεται από τα άκρα του, η επιφάνεια ανάμεσα στη χορδή και το καμπύλο τμήμα, και το κέντρο βάρους της επιφάνειας αυτής. Το μέγεθος των προσώπων δεν είναι γνωστό αλλά εκτιμάται ευριστικά με πληροφορία που λαμβάνεται και από τη λεζάντα της φωτογραφίας. Για παράδειγμα αν η λεζάντα αναφέρει ότι υπάρχουν τρία πρόσωπα στην εικόνα, περιορισμοί μπορούν να τεθούν όσον αφορά το μέγιστο μέγεθος του προσώπου, δεδομένου του μεγέθους της φωτογραφίας. Η πληροφορία της λεζάντας αναπαρίσταται με ένα σημασιολογικό (semantic) δίκτυο στο πλαίσιο μιας συλλογιστικής διαδικασίας. Η ανίχνευση των γραμμών και των τόξων πραγματοποιείται με μια παραλλαγή του μετασχηματισμού Hough. Για την διερεύνηση του κατά πόσο υπάρχει πρόσωπο στην εικόνα το μοντέλο (οι δύο γραμμές και τα δύο τόξα και η χωροταξική συσχέτιση τους) συγκρίνεται με τις γραμμές και τα τόξα της προς εξέταση εικόνας. Το κέντρο βάρους το τεσσάρων τμημάτων λαμβάνεται ως το κέντρο του προσώπου. Οι περιορισμοί που τίθενται είναι: (α) το πρόσωπο κοιτάζει προς τη κάμερα (β) δεν υπάρχει ούτε κλίση –βλέπε Σχήμα A4.1, στροφή περί τον άξονα z- ούτε περιστροφή - στροφή περί τον άξονα y. Μειονεκτήματα της μεθόδου είναι η αδυναμία εντοπισμού προσώπων με μικρό μέγεθος, λόγω

αποτυχίας αποτελεσματικής ανίχνευσης των ακμών της εικόνας. Η εφαρμογή της τεχνικής πραγματοποιήθηκε σε 10 εικόνες και παρόλο που δεν απέτυχε σε καμία να βρει το πρόσωπο έδωσε αρκετά false alarms (εντόπισε πρόσωπα εκεί που δεν υπήρχαν).

Μια από τις πρώτες εργασίες οι οποίες ασχολήθηκαν με τη διερεύνηση αν σε κάποια φωτογραφία υπάρχει πρόσωπο είναι η [16]. Στην εργασία αυτή υπολογίζεται η εικόνα ακμών από την αρχική φωτογραφία και στη συνέχεια συγκρίνεται με ένα πρότυπο οβάλ αντικείμενο το οποίο μπορεί να μεταβάλλεται σε μέγεθος και θέση. Σε θέσεις στις οποίες ανιχνεύεται πιθανή ύπαρξη προσώπου αυτή επιβεβαιώνεται με ανίχνευση στις προσδοκώμενες θέσεις ακμών που είναι γνωστό ότι αντιστοιχούν σε χαρακτηριστικά του προσώπου όπως μάτια, στόμα κλπ.

Στην εργασία [17] οι αρχικές φωτογραφίες φιλτράρονται και στη συνέχεια δημιουργούνται οι εικόνες ακμών. Σε αυτές ανιχνεύονται οι ακμές, οι οποίες ενδεχομένως αντιστοιχούν στο περίγραμμα του προσώπου, και στη συνέχεια προβάλλονται ξανά στην αρχική εικόνα όπου και πραγματοποιείται λεπτομερέστερη επεξεργασία για τον επακριβή εντοπισμό τους. Διάφορα ευριστικά τεχνάσματα χρησιμοποιούνται για τη συνένωση των ακμών. Μετά τον πλήρη εντοπισμό του περιγράμματος του προσώπου πραγματοποιείται ο εντοπισμός των επιμέρους χαρακτηριστικών στις αναμενόμενες θέσεις με την εφαρμογή και πάλι ευριστικών τεχνικών.

Στην εργασία του Craw [18] η ανίχνευση του προσώπου πραγματοποιείται με την τεχνική ταύτισης με πρότυπο -template matching. Ο εντοπισμός του προσώπου με την τεχνική αυτή έχει αποδειχτεί πειραματικά πολύ αποτελεσματικός, ιδιαίτερα αν το πρόσωπο κοιτάζει κατευθείαν στη κάμερα και έχει μικρές γωνίες κλίσης και περιστροφής. Αν δεν υπάρχει κάποια πληροφόρηση σχετικά με το μέγεθος του προσώπου, η ανίχνευση πρέπει να πραγματοποιηθεί χρησιμοποιώντας πρότυπα σε διαφορετικές αναλύσεις, κάτι που επιβαρύνει σε μεγάλο βαθμό την υπολογιστική πολυπλοκότητα του αλγορίθμου. Η τεχνική ταύτισης με πρότυπο είναι επιρρεπής σε σφάλματα στην ανίχνευση προσώπων που φοράνε γυαλιά ή έχουν γενειάδα. Για την αποφυγή τέτοιων προβλημάτων συχνά χρειάζεται η χρήση τροποποιημένων προτύπων που να λαμβάνουν υπόψη τις παραμέτρους αυτές. Τέλος ένα πολύ σημαντικό θέμα είναι η απόφαση κατά πόσο σε μια φωτογραφία υπάρχει πρόσωπο ή όχι. Στη τεχνική ταύτισης με πρότυπο, η θέση του προσώπου επιλέγεται ως η θέση με το καλύτερο ταίριασμα με το πρότυπο. Το γεγονός αυτό δεν εξασφαλίζει και το ότι στη συγκεκριμένη θέση υπάρχει πρόσωπο. Η εφαρμογή κάποιου κατωφλίου είναι αυθαίρετη και δεν μπορεί να στηριχθεί σε τεκμηριωμένα στοιχεία. Τέλος στη τεχνική ταιριάσματος με πρότυπο σημαντικό στοιχείο είναι η επιλογή του μέτρου ταύτισης. Πειραματικά αποδεικνύεται ότι μια παραλλαγή της μετρικής  $H$  δίνει τα καλύτερα αποτελέσματα.

Σε μια νεότερη εργασία του, ο Craw [19] περιγράφει ένα σύστημα για την αναγνώριση και μέτρηση των χαρακτηριστικών του προσώπου. Η εργασία αυτή είχε σχεδιαστεί για δεικτοδότηση (indexing) βάσεων από mugshot φωτογραφίες που χρησιμοποιούσαν οι αστυνομικές αρχές. Ο στόχος ήταν ο εντοπισμός 40 χαρακτηριστικών σημείων σε μια φωτογραφία εμπρόσθιας όψης. Η επιλογή των σημείων αυτών έγινε με βάση την εργασία του Richard [20] η οποία και χρησιμοποιήθηκε για την αξιολόγηση των αποτελεσμάτων. Η ανίχνευση του προσώπου πραγματοποιείται με τη χρήση ενός πολυγωνικού προτύπου (Motion template) σε μια διαδικασία Apple to al. Η πολυγωνική μορφή του προτύπου επιτρέπει το μετασχηματισμό του όσον αφορά τη θέση, κλίμακα, κλίση και περιστροφή. Η θέση του προσώπου βρίσκεται με τεχνικές template indexing και διαδοχικούς μετασχηματισμούς του προτύπου. Η γεωμετρία του ανθρώπινου προσώπου μπορεί να χρησιμοποιηθεί για να τεθούν περιορισμοί στους πιθανούς μετασχηματισμούς. Μετά από ένα χονδρικό εντοπισμό του προσώπου εφαρμόζονται μετασχηματισμοί σε μεμονωμένα πολύγωνα για τον επακριβή εντοπισμό του περιγράμματος. Ο εντοπισμός των χαρακτηριστικών σημείων πραγματοποιείται με βάση μοντέλα των

χαρακτηριστικών τα οποία οι συγγραφείς ονομάζουν feature text. Σύμφωνα με τα αποτελέσματα της εργασίας το περίγραμμα του προσώπου ανιχνεύθηκε και στις 50 φωτογραφίες, σε 43 πλήρως, ενώ στις υπόλοιπες 7 υπήρξε απώλεια του τμήματος που αντιστοιχεί στο πηγούνι εξαιτίας της ύπαρξης μουστακιού και γενειάδας. Δεν εντοπίστηκαν κυρίως τα χαρακτηριστικά σημεία που αντιστοιχούν στα φρύδια, δεδομένου ότι για αυτά οι συγγραφείς δεν είχαν προβλέψει κάποιο συγκεκριμένο μοντέλο. Με δεδομένη τη χρήση των πολυγωνικών μετασχηματισμών και της template indexing βελτιστοποίησης το σύστημα είναι αρκετά πολύπλοκο υπολογιστικά.

Η εργασία [21] είναι από τις λίγες που ασχολήθηκαν με την κατάτμηση του προσώπου σε μη ομοιογενές φόντο. Μετά από επεξεργασία της αρχικής εικόνας για την δημιουργία της εικόνας ακμών, η τεχνική που προτείνεται ομαδοποιεί εκείνες τις ακμές οι οποίες ανήκουν σε μια έλλειψη με την οποία έχει μοντελοποιηθεί το πρόσωπο. Οι παράμετροι της έλλειψης είναι το κέντρο  $(x_0, y_0)$  και οι ημιάξονες  $a, b$ .

Όπως φαίνεται και από το σύνολο των παραμέτρων δεν λαμβάνεται πρόνοια για ανίχνευση προσώπων με κλίση -στροφή ως προς τον  $z$  άξονα- ενώ αντίθετα μπορούν ανιχνευθούν πρόσωπα με στροφή περί τον άξονα  $y$  εφόσον έχει γίνει αποτελεσματική ανίχνευση των ακμών.

Μια τεχνική κατάτμησης προσώπου παρόμοια με την ταύτιση με πρότυπο αλλά με καλύτερο θεωρητικό υπόβαθρο στηρίζεται στο μετασχηματισμό Karhunen-Loeve. Κάθε υποεικόνα προβάλλεται στο χώρο των ιδιοδιανυσμάτων του ζητούμενου χαρακτηριστικού και οι συντελεστές της προβολής συγκρίνονται με τις ιδιοτιμές του χαρακτηριστικού. Η διαφορά αυτή ορίζεται ως DFFS (Distance-From-Feature-Space) και προφανώς σε κάθε pixel της εικόνας αντιστοιχεί μια τέτοια απόσταση. Το pixel με τη μικρότερη απόσταση προσδιορίζει τη θέση του χαρακτηριστικού (εφόσον βέβαια η απόσταση αυτή είναι μικρότερη από κάποιο σκαλοπάτι, ώστε να είμαστε βέβαιοι, ότι το ζητούμενο χαρακτηριστικό υπάρχει στην υπό εξέταση εικόνα).

Συγκεκριμένα έστω ένα σύνολο από  $N$  πρότυπα χαρακτηριστικών προσώπου –μάτια, μύτη στόμα κοκ-  $\{I_1, I_2, \dots, I_N\}$  όπου  $I_i \in R^{n \times m}$ . Με λεξικογραφική διάταξη των προτύπων  $I_i$  σχηματίζουμε το σύνολο των διανυσμάτων  $\{v_1, v_2, \dots, v_N\}$  με  $v_i \in R^L$  και  $L = n \times m$ . Οι συναρτήσεις βάσης του μετασχηματισμού KL για το παραπάνω σύνολο των διανυσμάτων προκύπτουν από την επίλυση του προβλήματος ιδιοτιμών:

$$\Lambda \cdot \Phi^T \Sigma \Phi \quad (2)$$

όπου  $\Sigma$  είναι ο πίνακας συμμεταβλητότητας,  $\Phi$  είναι ο πίνακας των ιδιοδιανυσμάτων του  $\Sigma$ , και  $\Lambda$  είναι ο αντίστοιχος διαγώνιος πίνακας των ιδιοτιμών.

Ένα διάνυσμα περιγραφής  $y_i$  κάθε υποεικόνας  $x_i$  -απεικονιζόμενης μέσω του αντίστοιχου διανύσματος λεξικογραφικής διάταξης- προκύπτει από τη μερική προβολή της στους άξονες του μετασχηματισμού KL:

$$y_i = \Phi_k^T \cdot \hat{x}_i \quad (3)$$

όπου  $\hat{x}_i = x_i - \mu$ ,  $\mu = \sum_{j=1}^N v_j$  και  $\Phi_k$  είναι ο πίνακας με τα ιδιοδιανύσματα στήλες που αντιστοιχούν στις  $k$  μεγαλύτερες ιδιοτιμές του  $\Sigma$ .

Ορίζοντας το σφάλμα ανακατασκευής:

$$e_{KLT}(k, i) = (\hat{x}_i - \tilde{x}_i)^T \cdot (\hat{x}_i - \tilde{x}_i)$$

$$\text{όπου } \tilde{x}_i = \Phi_k \cdot y_i, \quad y_i = \Phi_k^T \cdot \hat{x}_i \quad (4)$$

η υποεικόνα με την ελάχιστη DFFS δίνεται από τη σχέση:

$$x_{opt} = \arg \min_i [e_{KLT}(k, i)] \quad (5)$$

Τεχνικές ανίχνευσης του προσώπου ως αυτόνομης μονάδας με χρήση νευρωνικών δικτύων έχουν επίσης αναπτυχθεί [22]. Η ανίχνευση προσώπων με χρήση νευρωνικών δικτύων μοιάζει ίσως περισσότερο από κάθε άλλη τεχνική με τη λειτουργία του ανθρώπινου εγκεφάλου. Δυστυχώς όμως παρουσιάζει όλα τα μειονεκτήματα των νευρωνικών δικτύων όπως η δυσκολία μάθησης και γενίκευσης. Επιπλέον υπάρχει σχετική δυσκολία υλοποίησης των αλγορίθμων που περιγράφονται στις διάφορες εργασίες.

Η δεύτερη προσέγγιση στην ανίχνευση προσώπου απαιτεί την ανίχνευση πρώτα κάποιων επιμέρους χαρακτηριστικών του προσώπου και στη συνέχεια με βάση τη χωροταξική τους διάταξη τον εντοπισμό όλου του προσώπου. Παρόλο που οποιοδήποτε από τα χαρακτηριστικά του προσώπου μπορεί να εντοπιστεί πρώτα τυπική επιλογή είναι τα μάτια.

Στην εργασία του Yulie [23] δίνεται μια τεχνική εντοπισμού των ματιών με χρήση deformable templates (παραμορφώσιμων προτύπων). Μια συνάρτηση ενέργειας με παραμέτρους τις ακμές, τις κορυφές και τις κοιλάδες της gray scale εικόνας χρησιμοποιείται για την ταύτιση. Το πρότυπο εφαρμόζεται στα επιμέρους τμήματα της εικόνας και οι παράμετροί του μεταβάλλονται για να επιτύχουν την ελαχιστοποίηση της συνάρτησης ενέργειας, δηλαδή τη βέλτιστη ταύτιση. Η ελαστικότητα του προτύπου επιτρέπει αλλαγές στην κλίμακα και στην οριοθέτηση των ματιών. Με τον τρόπο αυτό μπορούν να εντοπιστούν τα μάτια ανεξάρτητα από μεταβολές στη κλίμακα, κλίση και στις συνθήκες φωτεινότητας. Εκτός από τα μάτια, με την ίδια τεχνική εντοπίζονται και τα χείλη. Το πρότυπο για τα μάτια αποτελείται από ένα κύκλο ο οποίος αντιστοιχεί στη κόρη, δύο παραβολικά τμήματα, που αντιστοιχούν στις οριογραμμές ματιών και βλεφάρων, και δύο σημεία τα οποία αντιστοιχούν στα κέντρα των δύο λευκών περιοχών του ματιού. Η τεχνική αυτή είναι αρκετά αποτελεσματική αλλά αντιμετωπίζει δύο βασικά προβλήματα: οι συντελεστές της συνάρτησης ενέργειας υπολογίζονται ευριστικά, με αποτέλεσμα το πρότυπο να αντιμετωπίζει προβλήματα γενίκευσης και ο υπολογιστικός φόρτος είναι ιδιαίτερα υψηλός. Συγκεκριμένα η ανίχνευση των ματιών απαιτεί χρόνους της τάξης των πέντε λεπτών.

Ο Nixon στην εργασία του [24] προτείνει μια τεχνική στην οποία ανιχνεύει τα μάτια χρησιμοποιώντας το μετασχηματισμό Hough. Η ίριδα μοντελοποιείται ως κύκλος και η οριογραμμή ματιών και βλεφάρων προσεγγίζεται με τη χρήση μιας τροποποιημένης εκθετικής συνάρτησης. Στην εργασία του Craw [19] τα κέντρα των ματιών εντοπίζονται ως ένα ζεύγος από ελάχιστα της φωτεινότητας της εικόνας, με αναζήτηση τους κάτω από τα φρύδια. Είναι προφανές ότι της ανίχνευσης των ματιών προηγείται η ανίχνευση των φρυδιών.

Ως τελικό συμπέρασμα αναφέρεται ότι η ανίχνευση του προσώπου με χρήση πρώτα των επιμέρους χαρακτηριστικών μπορεί να εφαρμοστεί με επιτυχία μόνο σε περιπτώσεις στις οποίες το πρόσωπο στη φωτογραφία βρίσκεται σε αρκετά μεγάλη ανάλυση -ώστε να υπάρχει επαρκής πληροφορία για την μοντελοποίηση των επιμέρους χαρακτηριστικών- και τα χαρακτηριστικά δεν



έχουν επικαλύψεις. Από την άλλη πλευρά, στις περιπτώσεις όπου ο εξαιρετικά ακριβής εντοπισμός του προσώπου είναι πολύ σημαντικός, η χρήση των επιμέρους χαρακτηριστικών ίσως είναι αναπόφευκτη. Ένα επιπλέον στοιχείο είναι η εκ των προτέρων γνώση που έχουμε για τις θέσεις των χαρακτηριστικών στο πρόσωπο. Αν για κάποιο λόγο η ανίχνευση ενός χαρακτηριστικού αποτύχει, αυτό μπορεί να γίνει αντιληπτό μετά την ανίχνευση κάποιου άλλου, εφόσον η γεωμετρική διάταξη των δύο χαρακτηριστικών δεν συμφωνεί με την ανατομία του προσώπου.

### **Ταυτοποίηση Προσώπων**

Η ταυτοποίηση είναι το επόμενο και πιο σημαντικό στάδιο στην αναγνώριση προσώπου. Υπάρχουν δύο απόψεις σχετικές με την ταυτοποίηση: (α) Το πρόσωπο πρέπει να συσχετιστεί με κάποιο όνομα και (β) πολλές εμφανίσεις του ίδιου προσώπου πιθανόν με διαφορετικές αναλύσεις, συνθήκες φωτισμού και οριοθέτησης πρέπει να αναγνωρίζονται ως μία. Η πρώτη άποψη αναφέρεται περισσότερο στην λογική των ανθρώπων και για την υλοποίηση της σε περιβάλλον υπολογιστή χρειάζεται πρώτα η υιοθέτηση της δεύτερης άποψης. Ο πρώτος που ασχολήθηκε εκτεταμένα με τη διαδικασία ταυτοποίησης προσώπων ήταν ο Galton [14][15] στο τέλος του 19ου αιώνα. Με βάση τα δεδομένα εισόδου -διαθέσιμες φωτογραφίες- η ταυτοποίηση γίνεται με χρήση εικόνων προφίλ ή με φωτογραφίες εμπρόσθιας όψης.

Στα πρώτα χρόνια της ερευνητικής δραστηριότητας για την ταυτοποίηση προσώπων οι περισσότερες τεχνικές χρησιμοποιούσαν εικόνες προφίλ. Αντίθετα στην σύγχρονη βιβλιογραφία αναφέρονται τεχνικές που χρησιμοποιούν εικόνες εμπρόσθιας όψης. Παρόλο που η αφετηρία των δύο προσεγγίσεων διαφέρει, οι αρχές στις οποίες στηρίζονται είναι παρόμοιες. Μια τυπική ακολουθία βημάτων που ακολουθούνται στην ταυτοποίηση προσώπων και στις δύο προσεγγίσεις δίνεται στη συνέχεια:

- (1) Προσδιόρισε ένα σύνολο από ανεξάρτητα χαρακτηριστικά για την αναπαράσταση του προσώπου.
- (2) Αναπαράστησε όλα τα γνωστά πρόσωπα με βάση τα χαρακτηριστικά του βήματος 1 και αποθήκευσε τα σε μία βάση δεδομένων
- (3) Υπολόγισε τα χαρακτηριστικά του προς ταυτοποίηση προσώπου
- (4) Χρησιμοποίησε μια διαδικασία ταύτισης συνδυασμένη με κατάλληλη μετρική για να βρεις την βέλτιστη ταύτιση με τα γνωστά πρόσωπα.

Η διαφοροποίηση των τεχνικών που έχουν προταθεί για την ταυτοποίηση προσώπου έγκειται στην επιλογή των χαρακτηριστικών και της διαδικασίας ταύτισης. Στις περισσότερες περιπτώσεις ο αριθμός των γνωστών προσώπων στη βάση δεδομένων είναι σχετικά μικρός -λιγότερος από 100 πρόσωπα (βλέπε Πίνακα 4).

Ο αριθμός των χαρακτηριστικών ποικίλει από 10-40.

Η πρώτη εργασία στην ταυτοποίηση προσώπων από τον Galton [14][15] στηρίχθηκε σε εικόνες προφίλ. Ένα σύνολο από πέντε πρωτεύοντα σημεία που αντιστοιχούν στην εγκοπή ανάμεσα στη μύτη και τα φρύδια - σημείο 9 στην Εικόνα 12, στη κορυφή της μύτης, στην εγκοπή ανάμεσα στη μύτη και το άνω χείλος, στο τμήμα διαχωρισμού των χειλιών, και στη κορυφή του πηγουνιού ορίζονται στην οριογραμμή του προφίλ. Με βάση τα σημεία αυτά παράγονται τα χαρακτηριστικά τα

οποία χρησιμοποιούνται για την ταύτιση των προσώπων. Η διαδικασία ταύτισης που ακολούθησε ο Galton ήταν η έξης:

Έστω  $X$  τα σύνολα των χαρακτηριστικών που αντιστοιχούν στα αποθηκευμένα στη βάση γνωστά πρόσωπα. Αν  $Y$  είναι το σύνολο των χαρακτηριστικών του προς ταυτοποίηση προσώπου τότε αναζητείται το σύνολο των χαρακτηριστικών  $Z \in X$  το οποίο προσομοιάζει το σύνολο  $Y$  με ένα προκαθορισμένο βαθμό ομοιότητας. Η ιδιαιτερότητα στη διαδικασία ταύτισης έγκειται στο γεγονός ότι δύο αντίστοιχα χαρακτηριστικά  $y_i \in Y$  και  $z_i \in Z$  δεν μπορούν να διαφέρουν περισσότερο από κάποιο προκαθορισμένο όριο. Η τεχνική αυτή ταύτισης είναι συνηθισμένη σε προσεγγίσεις που χρησιμοποιούν γεωμετρικά χαρακτηριστικά τα οποία είναι λίγα στον αριθμό.

Ο Harmon [25] χρησιμοποίησε μια παρόμοια με τον Galton τεχνική για την ταυτοποίηση προσώπων από προφίλ. Χρησιμοποίησε εννέα θεμελιώδη σημεία διατηρώντας τα πέντε σημεία του Galton και άλλα τέσσερα που αντιστοιχούν στο μέτωπο, στις κορυφές των δύο χειλιών και το λαιμό - Εικόνα A4.1. Με βάση τα θεμελιώδη σημεία δημιουργείται το σύνολο των χαρακτηριστικών που περιλαμβάνει αποστάσεις και γωνίες μεταξύ των σημείων, καθώς και το εμβαδόν κάποιων τριγώνων που αυτά σχηματίζουν. Ο βαθμός ομοιότητας μεταξύ των συνόλων από χαρακτηριστικά  $Y$  και  $Z$  υπολογίζεται με μια παραλλαγή της Ευκλείδειας απόστασης:

$$d^2(Y, Z) = \sum_{i=1}^n \frac{y_i^2 - z_i^2}{\sigma_i^2} \quad (6)$$

όπου  $n$  είναι ο αριθμός των χαρακτηριστικών και  $\sigma^2$  είναι η διασπορά τους. Τα προφίλ από 256 πρόσωπα εισήχθησαν στον υπολογιστή, στη συνέχεια δε τα θεμελιώδη σημεία εντοπίζονται αυτόματα. Για παράδειγμα για τον εντοπισμό των σημείων που αντιστοιχούν στις κορυφές των χειλιών και την εγκοπή του στόματος ακολουθείται η παρακάτω διαδικασία:

- (1) Ξεκίνα από το σημείο εγκοπής της μύτης και του άνω χείλους και προχώρα προς τα κάτω
- (2) Υπολόγισε τη καμπυλότητα κάθε σημείου που ανήκει στην οριογραμμή του προφίλ και
- (3) Επέλεξε τα τρία σημεία που αντιστοιχούν στις μέγιστες καμπυλότητες που έχουν τα κοίλα προς τα μέσα, προς τα έξω και πάλι προς τα μέσα. Τα σημεία αυτά αντιστοιχούν στα ζητούμενα θεμελιώδη σημεία.

Σε μια νεότερη εργασία του [26] προσέθεσε περισσότερα πρόσωπα, μερικά από τα οποία ανήκαν σε γυναίκες. Χρησιμοποιώντας τα ίδια θεμελιώδη σημεία κατασκεύασε ένα 11-Δ διάνυσμα χαρακτηριστικών, αποτελούμενο από την προεξοχή της μύτης, το εμβαδόν και την διακύμανση του προφίλ καθώς και αρκετές αποστάσεις όπως οι  $D_{12}$ ,  $D_{14}$ ,  $D_{15}$ ,  $D_{16}$ ,  $D_{18}$  και  $D_{79}$  όπου  $D_{ij}$  είναι η απόσταση μεταξύ των θεμελιωδών σημείων  $i$  και  $j$ . Για τη διαδικασία ταύτισης χρησιμοποιήθηκε μια τεχνική διαμέρισης συνόλων η οποία στηρίζεται στο γεγονός ότι, αν έστω και ένα χαρακτηριστικό από δύο συγκρινόμενα προφίλ διαφέρει σε μεγάλο βαθμό, τότε τα προφίλ ανήκουν σε διαφορετικά πρόσωπα. Μετά τη διαδικασία διαμέρισης, η ταύτιση επιτυγχάνεται με βάση την μικρότερη Ευκλείδεια απόσταση ανάμεσα στα διανύσματα χαρακτηριστικών που ανήκουν στην ίδια διαμέριση (υποσύνολο). Οι συγγραφείς της εργασίας κατέληξαν στο συμπέρασμα ότι ένα 10-Δ διάνυσμα χαρακτηριστικών είναι επαρκές για εύρωστη και αποτελεσματική ταυτοποίηση. Στη εργασία [27] τα θεμελιώδη σημεία αυξήθηκαν σε έντεκα και προστέθηκαν κάποια επιπλέον χαρακτηριστικά. Επίσης για πρώτη φορά η σύλληψη των εικόνων προφίλ έγινε αυτόματα με χρήση κάμερας και με ομογενές και με μεγάλη αντίθεση (contrast) φόντο.

Οι Kaufman και Breeding [28] σχεδίασαν επίσης ένα σύστημα ταυτοποίησης προσώπων με χρήση προφίλ. Τα χαρακτηριστικά τα οποία χρησιμοποίησαν ήταν συντελεστές υπολογισμένοι από την πολική μορφή της συνάρτησης αυτοσυσχέτισης. Χρησιμοποίησαν επίσης συναρτήσεις moment invariants.

Διάφορες άλλες τεχνικές έχουν επίσης χρησιμοποιηθεί για ταυτοποίηση με βάση τα προφίλ. Μια σημαντική παράμετρος είναι ο τρόπος εισαγωγής των προφίλ στον υπολογιστή. Πολλές προσεγγίσεις χρησιμοποίησαν για το σκοπό αυτό συστήματα τηλεμετρίας - range finders.

Οι σύγχρονες τεχνικές ταυτοποίησης χρησιμοποιούν κυρίως εικόνες εμπρόσθιας όψης, οι οποίες μεταξύ άλλων εισάγονται ευκολότερα στον υπολογιστή. Μια από τις πρώτες τεχνικές της κατηγορίας αυτής προτάθηκε από τον Goldstein [29]. Τα πρόσωπα σχεδιάζονταν από ζωγράφους, και οι τιμές των χαρακτηριστικών υπολογίζονταν από διάφορους κριτές. Αρχικά χρησιμοποιήθηκαν 34 χαρακτηριστικά τα οποία όμως περιορίστηκαν τελικά στα 22, δεδομένου ότι δεν ήταν όλα μεταξύ τους ανεξάρτητα (π.χ. το πάχος του άνω και κάτω χείλους). Κάθε χαρακτηριστικό μπορούσε να πάρει μια τιμή στη κλίμακα 1-5 (χαμηλό – ενδιάμεσες τιμές - υψηλό). Το σύνολο μάθησης αποτελείτο από 255 πρόσωπα. Τα χαρακτηριστικά περιελάμβαναν διάφορες μετρήσεις χαρακτηριστικών προσώπου όπως την απόχρωση και υφή των μαλλιών, τα φρύδια, τα μάτια, τη μύτη, το στόμα, τα αυτιά, τα μάγουλα και το μέτωπο. Μετά την εισαγωγή των δεδομένων το σύστημα υπολόγιζε την Ευκλείδεια απόσταση ανάμεσα στο προς ταυτοποίηση πρόσωπο και όλα τα πρόσωπα της βάσης δεδομένων. Το αποθηκευμένο πρόσωπο με τη μικρότερη απόσταση εθεωρείτο το ζητούμενο. Από τα συμπεράσματα των συγγραφέων προέκυψε ότι μόνο 6-7 χαρακτηριστικά είχαν ουσιαστική συμμετοχή στη διαδικασία ταύτισης. Οι Kaya και Kobayashi [30] πρότειναν ένα αντίστοιχο σύστημα, αλλά χρησιμοποίησαν τις αποστάσεις ανάμεσα στα χαρακτηριστικά του προσώπου για την κατασκευή του διανύσματος χαρακτηριστικών.

Ο Buhr [31] χρησιμοποίησε 33 πρωτεύοντα και 12 δευτερεύοντα χαρακτηριστικά για τη διαδικασία της ταυτοποίησης. Ανάμεσα στα πρωτεύοντα χαρακτηριστικά ήταν 21 αποστάσεις -π.χ. αποστάσεις ανάμεσα στα κέντρα των ματιών και του στόματος- 4 διαφορές τεταγμένων, 4 εμβαδά τριγώνων, 2 λόγοι αποστάσεων και δύο ειδικά χαρακτηριστικά που αφορούσαν το εμβαδόν της περιοχής των ματιών. Το σχήμα ταύτισης αποτελείτο από ένα γραμμικό δέντρο αποφάσεων. Ο Wong [32] χρησιμοποίησε μια παρόμοια τεχνική με σχήμα ταύτισης ένα συνδυασμό Ευκλείδεια απόστασης και της διαφοράς του αθροίσματος των χαρακτηριστικών. Διάφορες άλλες τεχνικές χρησιμοποιούν ταύτιση με βάση τα ελάχιστα τετράγωνα ή την ομαδοποίηση (clustering).

Ο Baron [33] χρησιμοποίησε τις gray scale τιμές της εικόνας ως χαρακτηριστικά. Συγκεκριμένα μετά από την ανίχνευση των ματιών, με χρήση ενός 20x23 πίνακα ταύτισης και εφαρμογή της συσχέτισης, οι εικόνες κανονικοποιούνται, διατηρώντας την απόσταση ανάμεσα στα μάτια σταθερή. Στη συνέχεια, μετά από υποδειγματοληψία, η ανάλυση των εικόνων μειώνεται σε 15x16 pixels. Για τη διαδικασία ταύτισης, η κανονικοποιημένη προς ταυτοποίηση εικόνα συγκρίνεται με κάθε εικόνα της βάσης δεδομένων και οι εικόνες για τις οποίες η τιμή συσχέτισης είναι υψηλή περνάνε στο δεύτερο στάδιο ελέγχου. Στο στάδιο αυτό ελέγχεται η συσχέτιση των επιμέρους χαρακτηριστικών του προσώπου τα οποία είναι αποθηκευμένα σε ξεχωριστούς πίνακες ταύτισης. Ταύτιση θεωρείται ότι επιτυγχάνεται όταν τα τρία τέταρτα των χαρακτηριστικών των δύο προσώπων παρουσιάζουν υψηλό βαθμό συσχέτισης.

Υπάρχουν προσεγγίσεις της ταυτοποίησης προσώπων οι οποίες χρησιμοποιούν εξειδικευμένο λογισμικό.

Συγκεκριμένα οι Preston [34] και Taylor [35] χρησιμοποίησαν ένα οπτικό σύστημα αναγνώρισης και ένα δίκτυο φωτο-πολλαπλασιαστών αντίστοιχα για το σκοπό αυτό. Βέβαια, ο αριθμός των προσώπων που χρησιμοποιήθηκαν ήταν εξαιρετικά περιορισμένος -6 και 10 πρόσωπα αντίστοιχα.

Συμπερασματικά για τη διαδικασία ταύτισης πρέπει να έχουμε υπόψη μας μερικά βασικά σημεία:

### **Επιλογή των χαρακτηριστικών**

Συστήματα τα οποία χρησιμοποιούν χαρακτηριστικά και όχι τις gray scale τιμές της εικόνας θα πρέπει να επιλέξουν ένα κατάλληλο και αποτελεσματικό σύνολο από χαρακτηριστικά για την ταυτοποίηση. Έχουν προταθεί πολλά χαρακτηριστικά στις διάφορες ερευνητικές εργασίες. Στους Πίνακες 3 και 7 παρουσιάζονται αρκετά από αυτά τα οποία διαφέρουν ανάλογα με το αν χρησιμοποιούνται εικόνες εμπρόσθιας όψης ή προφίλ.

### **Σχήματα ταύτισης**

Πολλές τεχνικές έχουν χρησιμοποιηθεί για τη διαδικασία ταύτισης. Στην πλειοψηφία τους είναι κλασσικές τεχνικές ή παραλλαγές τους. Συνοπτικά οι πιο δημοφιλείς φαίνεται ότι είναι οι:

(1) *Ευκλείδια απόσταση*. Αποτελεί την πρώτη επιλογή στη διαδικασία ταύτισης. Ο στόχος είναι να ευρεθεί το πρόσωπο εκείνο στη βάση δεδομένων του οποίου το διάνυσμα χαρακτηριστικών παρουσιάζει την ελάχιστη Ευκλείδια απόσταση με το αντίστοιχο διάνυσμα του προς ταυτοποίηση προσώπου. Η χρήση της Ευκλείδειας απόστασης είναι αποτελεσματική ανάλογα με τα χαρακτηριστικά που χρησιμοποιούνται.

(2) *Ομαδοποίηση*. Τεχνικές ομαδοποίησης χρησιμοποιούνται σχεδόν αποκλειστικά όταν τα πρόσωπα αναπαριστώνται από διανύσματα χαρακτηριστικών. Διακρίνονται για την ευστάθεια τους αλλά όχι για τη τέλεια διαχωριστική τους ικανότητα. Τυπικά παραδείγματα αποτελούν οι αλγόριθμοι k-means και τα νευρωνικά δίκτυα SOM.

(3) *Διαμερισμός συνόλων*. Εφαρμόστηκε σε πολλές περιπτώσεις σε τεχνικές που χρησιμοποιούν διανύσματα γεωμετρικών χαρακτηριστικών όπως αποστάσεις, γωνίες κλπ. Η διαδικασία στηρίζεται στην σταδιακή απόρριψη διανυσμάτων τα οποία διαφέρουν σημαντικά έστω και σε ένα επιμέρους χαρακτηριστικό. Η τελική επιλογή πραγματοποιείται από ένα πολύ μικρό υποσύνολο των προσώπων της βάσης δεδομένων και συχνά απαιτεί τη χρήση μια άλλης μετρικής όπως η Ευκλείδια απόσταση.

(4) *Συσχέτιση*. Δημοφιλής διαδικασία ταύτισης για τεχνικές που χρησιμοποιούν τις gray scale τιμές της εικόνας ως χαρακτηριστικά. Πάντως πειραματικά αποδεικνύεται ότι δεν είναι η βέλτιστη επιλογή για την ταύτιση.

(5) *Συνδυασμός των ανωτέρω*. Ο συνδυασμός των ανωτέρω τεχνικών ταύτισης τις περισσότερες φορές είναι επιβεβλημένος από την ίδια τη μορφή τους. Δεν αποτελεί έκπληξη επομένως το γεγονός ότι πολλοί ερευνητές ακολούθησαν την τακτική αυτή. Άλλες προσεγγίσεις όπως τα γραμμικά δέντρα απόφασης και τα ελάχιστα τετράγωνα έχουν επίσης χρησιμοποιηθεί αλλά δεν είναι τόσο διαδεδομένες.

Στους Πίνακες 4 - 6 πραγματοποιήσαμε μια σύγκριση πέντε διαφορετικών μετρικών οι οποίες χρησιμοποιούνται για ταύτιση πινάκων. Από τα αποτελέσματα της σύγκρισης προκύπτει ότι οι

μετρικές  $M1$  και  $M3$  όχι μόνο επιτυγχάνουν υψηλότερο βαθμό επιτυχούς ταυτοποίησης αλλά είναι λιγότερο επιρρεπείς στη χωρική ανάλυση των εικόνων. Επίσης παρουσιάζουν σημαντική ευρωστία, δεδομένου ότι ακόμα και οι αποτυχημένες ταυτοποιήσεις δεν είναι ιδιαίτερα απομακρυσμένες από τα σωστά πρότυπα που βρίσκονται αποθηκευμένα στη βάση δεδομένων. Σημειώνεται ότι οι εικόνες που χρησιμοποιήθηκαν στο πείραμα δεν είχαν κανονικοποιηθεί ως προς την απόσταση των ματιών, μια τακτική η οποία είναι σχεδόν επιβεβλημένη στις τεχνικές που χρησιμοποιούν για την ταυτοποίηση τους πίνακες των gray scale τιμών.

## 4. Κεφάλαιο -Κύριο Μέρος Πτυχιακής

### 4.1 Κατηγοριοποίηση Αλγορίθμων

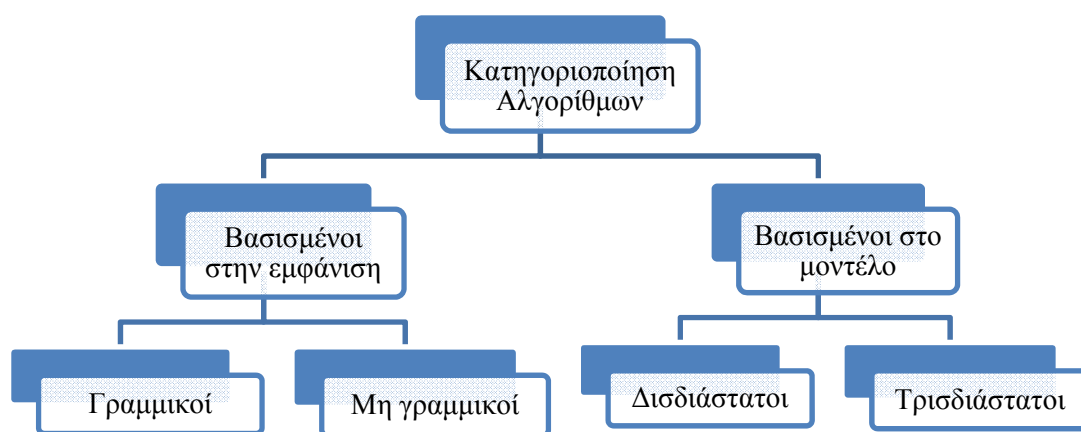
#### 4.1.1 Αλγόριθμος αναγνώρισης προσώπου

Σύμφωνα με τον καθολικό ορισμό του αλγορίθμου, Αλγόριθμος Αναγνώρισης Προσώπου ορίζεται μια πεπερασμένη σειρά ενεργειών, αυστηρά καθορισμένων και εκτελέσιμων σε πεπερασμένο χρόνο, που στοχεύει στην επίλυση του προβλήματος της αναγνώρισης του προσώπου. Το πρόβλημα αναγνώρισης προσώπου μπορούμε να το εκφράσουμε ορίζοντας αρχικά τα στοιχεία εισόδου τα οποία είναι ένα σύνολο  $A$  από εικόνες, όπως επίσης και μια επιπλέον εικόνα  $Test$  και θέλουμε να αναγνωρίσουμε. Έξοδος του προβλήματος είναι μια εικόνα από το αρχικό σύνολο  $A$  ή μια εικόνα της οποίας το χρώμα είναι μαύρο. Η λύση του προβλήματος ορίζεται ως την εύρεση του κατάλληλου προσώπου από το σύνολο  $A$ , το οποίο πρόσωπο έχει συντελεστή ομοιότητας μεγαλύτερο ή ίσο με  $\lambda$  όταν συσχετιστεί με το πρόσωπο  $Test$  και η εμφάνισή του στην έξοδο. Ειδικά αν η συσχέτιση είναι μικρότερη του  $\lambda$  τότε στην έξοδο θα πρέπει να εμφανίσουμε μια μαύρη εικόνα, που θεωρούμε πως υπάρχει.

Για την επίλυση λοιπόν του παραπάνω προβλήματος μπορούμε να αναπτύξουμε πολλούς αλγόριθμους, κύριο χαρακτηριστικό όλων θα είναι οι είσοδοι και οι έξοδοι του προβλήματος, ενώ μπορεί ο καθένας να υπολογίζει με διαφορετικό τρόπο τη συσχέτιση της εικόνας  $Test$  με το σύνολο των εικόνων  $A$ . Άρα ίσως και το αποτέλεσμα, δηλαδή η απόφαση που παίρνει ο κάθε αλγόριθμος να διαφέρει αφού θα έχει υπολογιστεί με διαφορετικό τρόπο και κριτήρια συσχέτισης. Επίσης ένα άλλο σημείο που μπορεί να διαφέρουν μεταξύ τους οι αλγόριθμοι είναι ο χρόνος εκτέλεσης κάτι που μπορούμε να το ανάγουμε και αυτό στους διαφορετικούς τρόπους υπολογισμού της συσχέτισης των εικόνων  $A$  με την  $Test$  εικόνα.

#### 4.1.2 Ταξινόμηση αλγορίθμων

Η αναγνώριση προσώπου είναι μια διαδικασία αναγνώρισης μοτίβου (pattern recognition). Το πρόσωπο σαν ένα τρισδιάστατο αντικείμενο χαρτογραφείται μέσω της προβολής μιας εικόνας που το αποτυπώνει στις δύο διαστάσεις. Ένα γενικό μοντέλο αναγνώρισης απαρτίζεται από δύο στάδια. Το πρώτο αφορά την εξαγωγή πληροφορίας από την εικόνα και τη μοντελοποίηση της με βάση συγκεκριμένες διαδικασίες. Το δεύτερο στάδιο είναι αυτό της ταξινόμησης και ταυτοποίησης κάνοντας χρήση μίας βάσης προσώπων για τα οποία έχει ακολουθηθεί η ίδια διαδικασία εξαγωγής πληροφορίας. Οι αλγόριθμοι αναγνώρισης προσώπου μπορούν να κατηγοριοποιηθούν βάσει δύο κύριων προσεγγίσεων στους αλγόριθμους που βασίζονται στην εμφάνιση (appearance based) και στους αλγόριθμους που βασίζονται στη δημιουργία μοντέλων (model based) ικανών να περιγράψουν το ανθρώπινο πρόσωπο (Σχήμα 1) [36].



**Σχήμα 1:** Κατηγοριοποίηση Αλγορίθμων

Αρκετοί αλγόριθμοι και των δυο κατηγοριών, καθώς και μια σειρά παραλλαγών τους εμφανίστηκαν τις τελευταίες δεκαετίες.

#### 4.1.3 Αλγόριθμοι που βασίζονται σε μοντέλα (Model - based Algorithms)

Τα συστήματα αναγνώρισης προσώπων που βασίζονται σε μοντέλα, χρησιμοποιούν τεχνικές για να κατασκευάζουν ένα μοντέλο που βασίζεται στη γεωμετρία του ανθρώπινου προσώπου τέτοιο ώστε να είναι ικανό να περιγράψει τις διαφορές ανάμεσα στα άτομα. Είναι αναγκαία η προαπαιτούμενη γνώση της δομής και των χαρακτηριστικών του προσώπου βάσει της οποίας σχεδιάζεται το μοντέλο. Για παράδειγμα η γνώση ότι το πρόσωπο έχει δυο μάτια, μια μύτη και ένα στόμα τα οποία είναι τοποθετημένα από πάνω προς τα κάτω με συγκεκριμένη σειρά αποτελεί μια καλή βάση για τη δημιουργία ενός μοντέλου.

Πρωτεργάτες της αυτόματης αναγνώρισης προσώπου θεωρούνται οι [37], [38], [39]. Στις εργασίες αυτές, το 1964 και 1965 χρησιμοποιήθηκε ο υπολογιστής για να αναγνωριστούν ανθρώπινα πρόσωπα. Η αναλυτική δημοσίευση των αποτελεσμάτων της έρευνας αυτής, η οποία είχε χρηματοδοτηθεί από υπηρεσία πληροφοριών, δεν πραγματοποιήθηκε ποτέ. Από τις λίγες δημοσιεύσεις που προέκυψαν, η μέθοδος που χρησιμοποιήθηκε βασιζόταν σε μοντελοποίηση του ανθρώπινου προσώπου όπου οι εικόνες ελέγχου συγκρίνονταν με μια μεγάλη βάση εικόνων και κατέληγαν στην εξαγωγή ενός συνόλου παρόμοιων εικόνων. Η επιτυχία της αναγνώρισης ήταν άρρηκτα συνδεδεμένη με τον λόγο του αριθμού των εικόνων του συνόλου που προέκυπτε ως προς το σύνολο των εικόνων της βάσης. Το σύστημα το οποίο αναπτύχθηκε ονομάστηκε «άνθρωπος - μηχανή» μιας και ο χρήστης αρχικά υπολόγιζε τις συντεταγμένες διάφορων χαρακτηριστικών από τις φωτογραφίες και στη συνέχεια τα αποτελέσματα αυτά τροφοδοτούσαν τη μηχανή για την επίτευξη της αναγνώρισης. Τα χαρακτηριστικά από τα οποία γινόταν η εξαγωγή των συντεταγμένων (μάτια, μύτη, ρινική οδός κλπ.) χρησιμοποιούνταν για τον υπολογισμό 20 αποστάσεων, όπως η απόσταση της μύτης από το στόμα, η απόσταση ανάμεσα στις άκρες των ματιών κλπ. Ο χρήστης μπορούσε να επεξεργάζεται 40 περίπου εικόνες την ώρα και κατά την δημιουργία της βάσης, κάθε πρόσωπο της

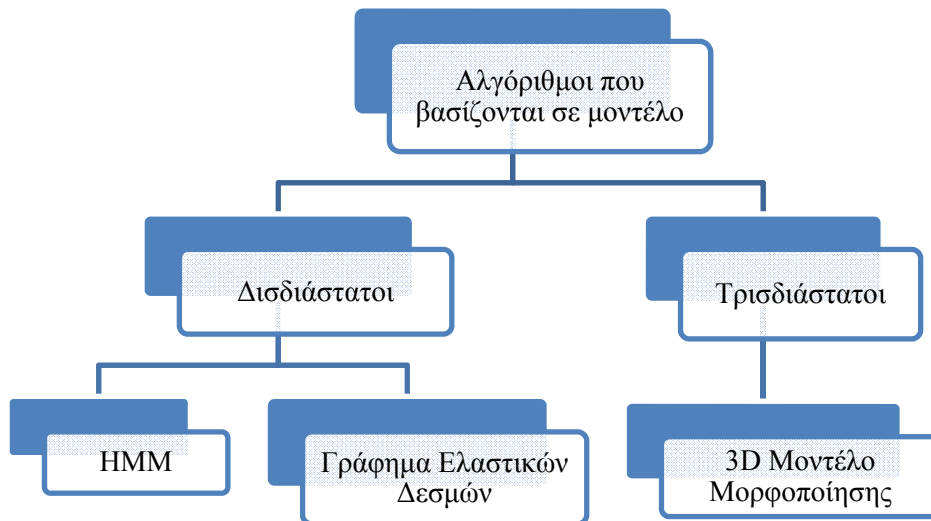
φωτογραφίας συνδεόταν με τη λίστα των αποστάσεων. Στη φάση της αναγνώρισης το σύνολο των αποστάσεων των εικόνων ελέγχου συγκρίνονταν με τις αποστάσεις των εικόνων της βάσης. Το πρωτεύον αυτό σύστημα αναγνώρισης ήταν υπεραπλουστευμένο και είχε πολλά προβλήματα λόγω των διαφορετικών γωνιών λήψης και της διαφορετικής απόστασης κάθε προσώπου από τον φακό (διαφορετική κλίμακα) των εικόνων. Ο κύριος συγγραφέας των [37], [38], [39], στην προσπάθειά του να εξαλείψει τα προβλήματα αυτά προσπάθησε να κανονικοποιήσει τις εικόνες υπολογίζοντας την κλίμακα και τη γωνία λήψης χρησιμοποιώντας επτά εικόνες από κάθε πρόσωπο.

Στην εργασία [40], δημιουργήθηκε ένα από τα πρώτα συστήματα αναγνώρισης μέσω μιας διαδικασίας αυτόματης εύρεσης των χαρακτηριστικών ενός προσώπου χρησιμοποιώντας ένα μοντέλο. Βρίσκοντας τις συντεταγμένες των άκρων των ματιών, της μύτης κλπ. μιας κατά πρόσωπο απεικόνιση των εικόνων, το σύστημα αυτό μπορούσε να αναγνωρίσει το πρόσωπο συγκρίνοντας με Ευκλείδεια απόσταση τις συντεταγμένες των χαρακτηριστικών των εικόνων ελέγχου με τις εκείνες της βάσης γνωστών προσώπων [41].

Πιο πρόσφατα συστήματα που βασίζονται σε μοντέλα είναι αυτό των γραφημάτων Ελαστικών Δεσμών [42], που δημιουργήθηκε το 1997, το σύστημα που βασίζεται σε κρυμμένα πρότυπα Markov (HMM) [12] το οποίο και βελτιώθηκε το 1998 [43] καθώς και περισσότερο σύνθετα συστήματα τρισδιάστατης μορφοποίησης. Οι αλγόριθμοι που βασίζονται σε μοντέλα συνήθως περιλαμβάνουν τα εξής τρία βήματα:

- Κατασκευή του μοντέλου
- Εφαρμογή του μοντέλου στην εικόνα προσώπου
- Δημιουργία διανύσματος που βασίζεται στις παραμέτρους εφαρμογής του μοντέλου στο πρόσωπο και σύγκριση με τα αντίστοιχα διανύσματα των γνωστών προσώπων.

Οι κυριότεροι αλγόριθμοι που βασίζονται σε μοντέλα φαίνονται στο Σχήμα 2.



**Σχήμα 2:** Αλγόριθμοι που βασίζονται σε μοντέλα



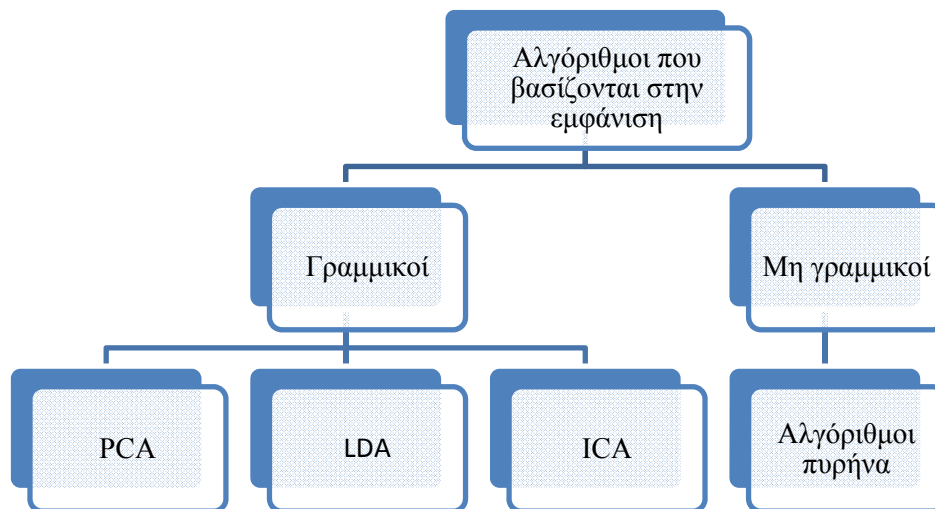
#### 4.1.4 Αλγόριθμοι που βασίζονται στην εμφάνιση (Appearance – Based Algorithms)

Πολλές προσεγγίσεις για την αναγνώριση προσώπων βασίζονται απευθείας στις εικόνες χωρίς την χρήση γεωμετρικών μοντέλων. Οι περισσότερες από αυτές τις τεχνικές εξαρτώνται από την αναπαράσταση των εικόνων ενώ δημιουργούν διανυσματικούς «χώρους προσώπων». Κάθε εικόνα  $m \times n$  μπορεί να θεωρηθεί ως ένα διάνυσμα στο  $R^{mn}$ . Στον χώρο αυτό κάθε εικόνα-πρόσωπο αντιπροσωπεύεται από ένα σημείο. Αρκετές από τις μεθόδους που βασίζονται στην εμφάνιση χρησιμοποιούν στατιστικές μεθόδους για τη ανάλυση της κατανομής των διανυσμάτων των εικόνων στον χώρο προσώπων και παράγουν μια αναπαράσταση (διάνυσμα χαρακτηριστικών) ικανή να περιγράψει το κάθε πρόσωπο. Για κάθε εικόνα ελέγχου δημιουργείται το διάνυσμα χαρακτηριστικών και στη συνέχεια συγκρίνεται με αυτά των «εικόνων εκπαίδευσης» που δημιούργησαν τον διανυσματικό χώρο. Η χρήση της διανυσματικής αναπαράστασης επιτρέπει την χρήση τεχνικών εκμάθησης για την εκπαίδευση των συστημάτων αναγνώρισης. Το πρόβλημα της αναγνώρισης σε αυτήν την περίπτωση μπορεί να αντιμετωπιστεί σαν πρόβλημα ταξινόμησης-αντιστοίχισης στον χώρο και ταυτόχρονα σαν πρόβλημα εκπαίδευσης μηχανών.

Στην [44] χρησιμοποιήθηκε η ανάλυση των κύριων συνιστωσών για να επιλύσουν το πρόβλημα της αναγνώρισης προσώπου. Αυτό αποτέλεσε σημαντικό επίτευγμα μιας και απέδειξε ότι πολύ μικρά μεγέθη πληροφορίας ήταν αρκετά να κωδικοποιήσουν ικανοποιητικά μια εικόνα προσώπου. Το 1991, [45] με την χρήση των ιδιοδιανυσμάτων, αναπτύχθηκε ένας τρόπος προσδιορισμού της ύπαρξης ή όχι ενός προσώπου σε μία εικόνα. Η τεχνική αυτή οδήγησε στη συνέχεια την δημιουργία συστημάτων αυτόματης αναγνώρισης προσώπου, σε πραγματικό χρόνο. Η προσέγγιση αυτή είχε αρκετούς περιορισμούς αλλά παρόλα αυτά αύξησε κατακόρυφα το ενδιαφέρον της επιστημονικής κοινότητας για την δημιουργία πλήρως αυτοματοποιημένων συστημάτων αναγνώρισης. Το ευρύ κοινό ήρθε σε επαφή με την εφαρμογή του συστήματος αυτού όταν τον Ιανουάριο του 2001, σε έναν αγώνα ποδοσφαίρου, εικόνες προσώπου που «τραβήχτηκαν» από τις κάμερες του συστήματος επιτήρησης ταυτοποιήθηκαν με μια βάση εικόνων της αστυνομίας. Η επίδειξη αυτή δημιούργησε ταυτόχρονα πολλές συζητήσεις για τη διασφάλιση των προσωπικών δεδομένων.

Οι συγγραφείς της [46] δημοσίευσαν το 1997, μια συγκριτική έρευνα ανάμεσα στην χρήση των κύριων συνιστωσών (PCA) και της γραμμικής διακρίνουσας ανάλυσης (LDA) για την αναγνώριση προσώπων. Ακολούθησαν, παρόμοιες έρευνες , [47], [48], και άλλες. Οι [49], [50], δημοσίευσαν έρευνα γύρω από τη χρήση της ανάλυσης των Ανεξάρτητων Συνιστωσών (ICA) όπου οι Ανεξάρτητες Συνιστώσες βρίσκονται μέσω της μεγιστοποίησης της εντροπίας σε ένα νευρωνικό δίκτυο  $N$  εισόδων -  $N$  εξόδων με σιγμοειδή συνάρτηση μεταφοράς, βασισμένο στον αλγόριθμο 16 [51]. Στην έρευνα [52] το πρόβλημα της εύρεσης των ανεξάρτητων συνιστωσών αντιμετωπίστηκε με την χρήση σταθερού αριθμού επαναλήψεων για την εκπαίδευση ενός νευρωνικού δικτύου  $N$  εισόδων -  $N$  εξόδων δημιουργώντας τον αλγόριθμο που είναι γνωστός και ως FASTICA

Οι κυριότεροι αλγόριθμοι αναγνώρισης προσώπου που βασίζονται στην εμφάνιση φαίνονται στο Σχήμα 3



Σχήμα 3: Αλγόριθμοι αναγνώρισης προσώπου που βασίζονται στην εμφάνιση

#### 4.1.5 Αποστάσεις και συναρτήσεις ομοιότητας

Η έννοια της απόστασης είναι θεμελιώδης στην πολυμεταβλητή ανάλυση. Σκοπός της απόστασης είναι να μετρήσει «πόσο απέχουν» δύο παρατηρήσεις. Στη διαδικασία αναγνώρισης που εφαρμόζεται στα διάφορα συστήματα αναγνώρισης προσώπου η έννοια της απόστασης λαμβάνει την σημασία της ομοιότητας των προσώπων. Οι αλγόριθμοι αναγνώρισης προσώπου παράγουν διανύσματα χαρακτηριστικών τόσο για τις εικόνες ελέγχου όσο και για τις εικόνες που σχηματίζουν τη βάση εκπαίδευσης. Έστω  $P_i = (p_1, \dots, p_n)$  το διάνυσμα χαρακτηριστικών μιας εικόνας ελέγχου και  $Q_i = (q_1, \dots, q_n)$ ,  $i = 1 \dots m$ , τα διανύσματα χαρακτηριστικών των  $m$  εικόνων της βάσης. Οι αποστάσεις για τα διανύσματα  $P$  και  $Q$  ορίζονται ως εξής:

- **Απόσταση Manhattan ή Cityblock ή L1 norm.**

Για  $k=1$  στην απόσταση Minkowski ορίζεται η απόσταση Cityblock (οικοδομικών τετραγώνων) που αποτελεί το άθροισμα των απόλυτων διαφορών των διανυσμάτων  $p$  και  $q$ . Η απόσταση αυτή είναι γνωστή και ως απόσταση Manhattan η νόρμα L1 και ορίζεται από την σχέση:

$$d(P, Q) = \sum_{i=1}^n |p_i - q_i| \quad (7)$$

- **Ευκλείδεια απόσταση**

Για  $k=2$  η απόσταση Minkowski ορίζει την Ευκλείδεια απόσταση των διανυσμάτων  $P$  και  $Q$ ,

$$d(P, Q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (8)$$

Χρησιμοποιώντας την νόρμα (norm) της διαφοράς των δύο διανυσμάτων  $P$  και  $Q$  η ευκλείδεια απόσταση δίνεται από την σχέση:

$$d(P, Q) = \|P - Q\| = \sqrt{(P - Q)^T(P - Q)} = \sqrt{\|P\|^2 + \|Q\|^2 - 2PQ} \quad (9)$$

- **Ομοιότητα συνημίτονου (Cosine)**

Η ομοιότητα συνημίτονου είναι ένα μέτρο που υπολογίζει την ομοιότητα ανάμεσα σε δύο διανύσματα υπολογίζοντας το συνημίτονο της γωνίας  $\varphi$  που σχηματίζουν. Ορίζεται από την σχέση:

$$d(P, Q) = \cos(\varphi) = 1 - \frac{P^T Q}{\sqrt{P^T P} \sqrt{Q^T Q}} \quad (10)$$

- **Απόσταση Mahalanobis**

Η απόσταση Mahalanobis παρουσιάστηκε από τον P.C Mahalanobis το 1936 ως μέτρο απόστασης που λαμβάνει υπόψη την ασύμμετρη κατανομή των διαστάσεων των συγκρινόμενων διανυσμάτων  $P$  και  $Q$ . Βασίζεται στη συσχέτιση μεταξύ μεταβλητών βάση των οποίων αναγνωρίζονται και αναλύονται διαφορετικά μοτίβα κατανομής των τιμών. Διαφέρει από την Ευκλείδεια απόσταση μιας και λαμβάνει

υπόψη τις συσχετίσεις (correlations). Ορίζεται από την σχέση:

$$dmah(P, Q) = \sqrt{(p_i - q_i)^T C^{-1} (p_i - q_i)} \quad (11)$$

όπου  $C$  η συνδιακύμανση των  $P$  και  $Q$ .

#### 4.1.6 Ανάλυση σε Κύριες Συνιστώσες (Principal Component Analysis, PCA)

Ένας ευρύτατα διαδεδομένος μετασχηματισμός για την δημιουργία αντιπροσωπευτικών και αποτελεσματικών υποχώρων είναι η Ανάλυση Κυρίων Συνιστωσών (Principal Component Analysis - PCA). Οι Turk και Pentland [53] εισήγαγαν τον PCA στο πεδίο της αναγνώρισης προσώπων. Είναι μια μέθοδος που χρησιμοποιείται ως στατιστική προσέγγιση αναγνώρισης μοτίβων, επεξεργασίας σημάτων και μείωσης του όγκου των δεδομένων (συμπύεση) όπως και κατά την εξαγωγή χαρακτηριστικών ενός σήματος.

Δεδομένου ενός συνόλου από εικόνες με το πρόσωπο του ίδιου ατόμου, ανακτώνται οι κύριες συνιστώσες της κατανομής (τα περισσότερα ισχυρά ιδιοδιανύσματα ή ιδιοπρόσωπα – eigenfaces), σχηματίζοντας τη βάση του νέου υποχώρου του ιδιοχώρου ή προσωποχώρου (face space). Όταν μια καινούρια εικόνα πρέπει να ταξινομηθεί, προβάλλεται στον προσωποχώρο μέσω των ιδιοπροσώπων και στη συνέχεια ανακατασκευάζεται πίσω στο χώρο εισόδου. Το σφάλμα αυτής της ανακατασκευής μπορεί να χρησιμοποιηθεί σαν μία μέτρηση της εγγύτητας του εικονιζόμενου προσώπου με αυτό από του οποίου ανακατασκευάστηκε ο προσωποχώρος. Αυτή η τεχνική επεκτάθηκε στην ανίχνευση των χαρακτηριστικών του προσώπου και σε αυτήν καθαυτή ανίχνευση προσώπων [54], χρησιμοποιώντας αυτά τα χαρακτηριστικά ως φόρμες συσχέτισης. Στην εργασία [55], εκτός από τις κύριες συνιστώσες της κατανομής του προσώπου, λήφθηκαν υπ' όψιν και τα ορθογώνια συμπληρώματα για την αντιμετώπιση της μεταβολής του φωτισμού.

Αποτελεί μια εφαρμογή του μετασχηματισμού Karhunen-Loeve . Οι κύριες συνιστώσες ορίζονται ως τα ιδιοδιανύσματα (eigenfaces) του πίνακα διασποράς C

$$C = \sum_{i=1}^n (x_i - \mu)(x_i - \mu)^T \quad (12)$$

Όπου  $x_i$  το διάνυσμα κάθε εικόνας,  $\mu$  η μέση τιμή των διανυσμάτων εικόνας και  $n$  το πλήθος των εικόνων. Η προβολή των εικόνων γίνεται στα ιδιοδιανύσματα που αντιστοιχούν στις  $m$  μεγαλύτερες ιδιοτιμές (eigenvalues). Οι κύριες συνιστώσες έχουν ίδια διάσταση με τις αρχικές εικόνες και αναφέρονται συχνά ως eigenfaces .

Η τεχνική των Eigenfaces αναπτύχθηκε από το Τεχνολογικό Ινστιτούτο της Μασαχουσέτης (MIT) και παρακινήθηκε από μια τεχνική που αναπτύχθηκε από τους Sirovich και Kirby το 1987 για αποτελεσματική αναπαράσταση των εικόνων προσώπων χρησιμοποιώντας την ανάλυση πρωτευουσών συνιστωσών PCA). Παραλλαγές της τεχνικής των eigenfaces χρησιμοποιούνται συχνά ως βάση άλλων μεθόδων αναγνώρισης προσώπου. Έχει υποστηριχτεί ότι αυτή η τεχνική δεν κατέχει οποιαδήποτε ομοιότητα στον τρόπο που οι άνθρωποι αναγνωρίζουν και μετρούν την ομοιότητα μεταξύ των προσώπων. Σημείωση : οι εικόνες προς εκπαίδευση και οι προς εξέταση εικόνες θα πρέπει να έχουν όλες το ίδιο μέγεθος, αλλά και να κανονικοποιηθούν ώστε τα μάτια και το στόμα να ευθυγραμμιστούν.

Αφού γίνει η προβολή, το διάνυσμα εισόδου διάστασης  $N$ , δηλαδή το πρόσωπο, αναπαρίσταται με ένα διάνυσμα του  $m$ - διάστατου υποχώρου. Οι ιδιοτιμές αποτελούν ένδειξη της αναλογίας μεταβλητότητας των δεδομένων τα οποία αφορά κάθε ιδιοδιάνυσμα. Η πρώτη κύρια συνιστώσα δείχνει την κατεύθυνση της μέγιστης μεταβλητότητας, η δεύτερη την κατεύθυνση μέγιστης μεταβλητότητας που είναι ορθογώνια ως προς την πρώτη κ.ο.κ.

Η απομάκρυνση ορισμένων ιδιοδιανυσμάτων μπορεί να βελτιώσει την απόδοση της μεθόδου με την απομάκρυνση του θορύβου. Για το λόγο αυτό, η συνήθης προσέγγιση της μεθόδου περιλαμβάνει επιλογή των ιδιοδιανυσμάτων που αντιστοιχούν στις μεγαλύτερες ιδιοτιμές. Η επιλογή αυτή οδηγεί ουσιαστικά σε επιλογή των διευθύνσεων στις οποίες οι εικόνες διαφέρουν περισσότερο. Στις περισσότερες εφαρμογές, τα ιδιοδιανύσματα που αντιστοιχούν σε μικρές ιδιοτιμές θεωρούνται θόρυβος και δεν λαμβάνεται υπόψη στη διαδικασία ταυτοποίησης.

#### 4.1.6.1 Μαθηματική προσέγγιση για την εύρεση Eigenfaces:

Έστω λοιπόν μία συλλογή από  $N$  πρόσωπα με  $m \times n$  διαστάσεις το καθένα.



**Εικόνα 14:** Εικόνες εκπαίδευσης αλγορίθμου

Η κάθε εικόνα  $A_i$ , όπου  $i = 1, 2, 3, \dots, N$  μπορεί να αναπαρασταθεί ως ένα διάνυσμα  $\Gamma_i$  διάστασης  $1 \times (mn)$ . Από αυτά τα διανύσματα αφαιρούμε το μέσο διάνυσμα:

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

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

έπειτα σχηματίζουμε τον πίνακα αυτοσυσχέτισης :

$$C = \frac{1}{N} \sum_{i=1}^M \Phi_i \Phi_i^T \quad (13)$$

Επόμενο βήμα είναι ο υπολογισμός των ιδιοδιανυσμάτων του. Εδώ όμως υπάρχει ένα πρόβλημα, ο πίνακας  $C$  είναι ένας πολύ μεγάλος πίνακας ( $(mn) \times (mn)$  διαστάσεων). Έστω  $A = [\Phi_1, \Phi_2, \dots, \Phi_N]$  τότε ο  $C$  γράφεται ως  $AA^T$

Υπολογίζω τα ιδιοδιανύσματα  $V_i$  του  $A^T A$  (πίνακας  $N \times N$ ):

$$A^T A v_i = \mu_i v_i \quad (14)$$

Μπορώ όμως από τα ιδιοδιανύσματα  $v_i$  να υπολογίσω τα ιδιοδιανύσματα  $v_i$  του πίνακα  $C$ ;

$$A^T A v_i = \mu_i v_i \rightarrow A A^T A v_i = A \mu_i v_i$$

$$A A^T A v_i = \mu_i A v_i \rightarrow C A v_i = \mu_i A v_i$$

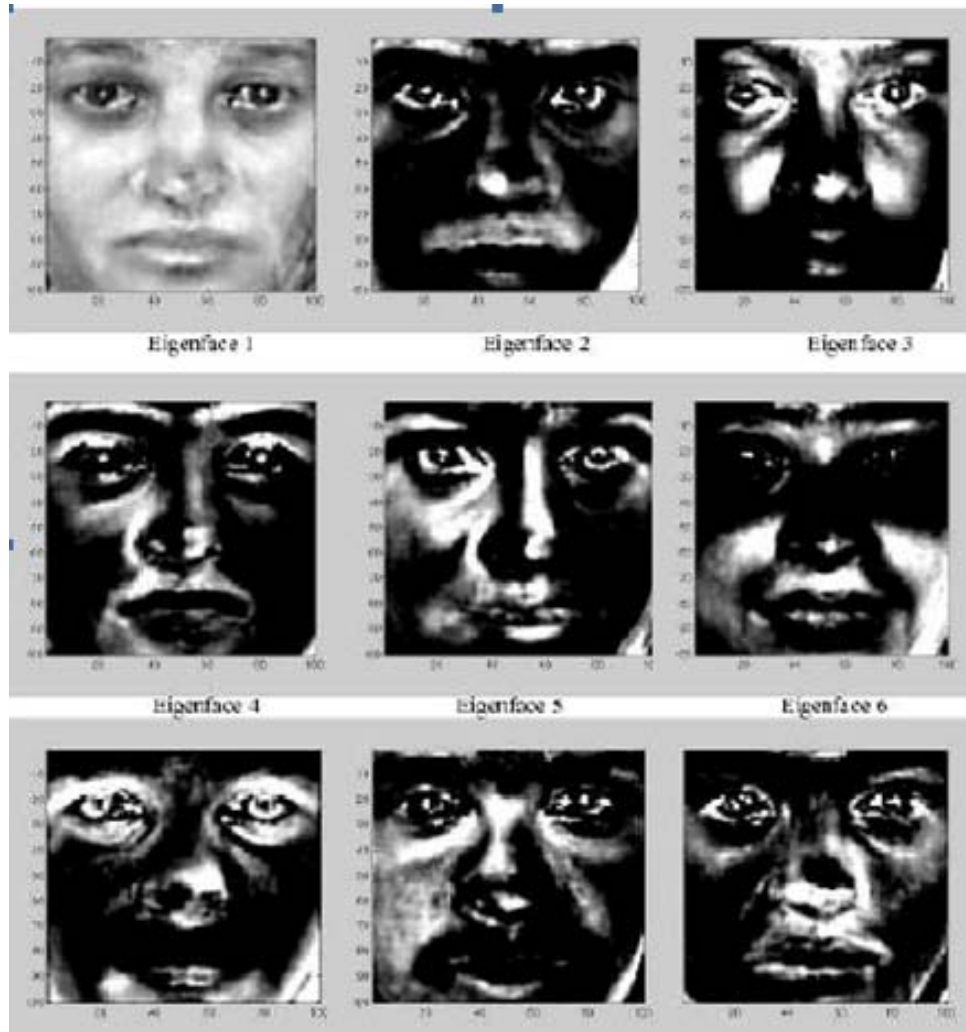
Άρα,

$$C u_i = \mu_i u_i \quad (15)$$

Όπου  $u_i = A v_i$

Αυτό υποδηλώνει ότι ο  $C$  και ο  $A^T A$  έχουν τις ίδιες ιδιοτιμές και τα ιδιοδιανύσματα τους σχετίζονται κατά αυτή την σχέση:  $u_i = A v_i$ .

Σημείωση ο πίνακας  $C$  έχει μέχρι  $mn$  ιδιοτιμές και ιδιοδιανύσματα, και ο  $A^T A$  έχει μέχρι  $N$  ιδιοτιμές και ιδιοδιανύσματα. Οι  $N$  ιδιοτιμές του πίνακα  $A^T A$  αντιστοιχούν στις  $N$  σατανικότερες (μεγαλύτερες) ιδιοτιμές του  $C$ . Σε αυτό το σημείο θα κρατήσουμε μόνο  $K$  ιδιοδιανύσματα, τα οποία ανταποκρίνονται στις  $K$  μεγαλύτερες ιδιοτιμές.

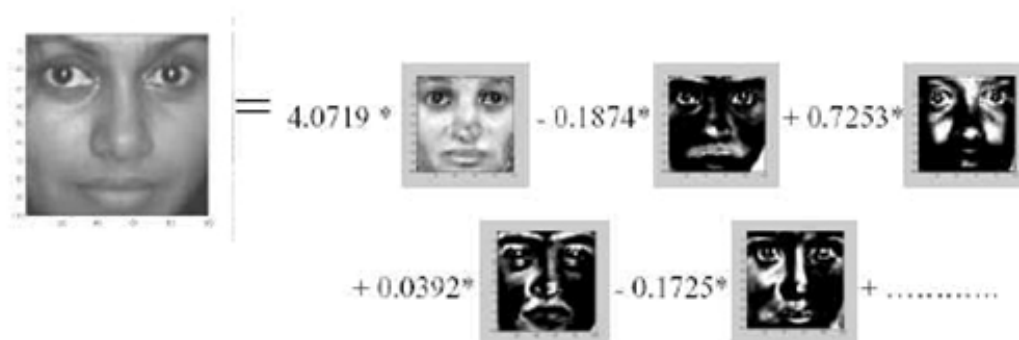


Σχήμα 4: Εικόνες ιδιοδιανυσμάτων (Eigenfaces)

Κάθε πρόσωπο  $\Phi_i$  μπορεί να αναπαρασταθεί σαν ένας γραμμικός συνδυασμός των  $K$  σημαντικότερων ιδιοδιανυσμάτων:

$$\Phi_i = \sum_{j=1}^K u_j^T \Phi_i u_j \quad (16)$$

Έτσι κάθε εικόνα εκπαίδευσης  $\Phi_i$  αναπαριστάται σε αυτή την βάση με ένα διάνυσμα:



Σχήμα 5: Ανακατασκευή του προσώπου από τις εικόνες ιδιοδιανυσμάτων

$$\Omega_i = (w_1^i, w_2^i, \dots, w_k^i)^T \quad i=1,2,\dots,N \quad (17)$$

Όπου  $w_i = u_j^T \Phi_i$

Έπειτα δοθέντος ενός προσώπου εξέτασης  $\Gamma^t$  το οποίο είναι κεντραρισμένο όσον αφορά τα μάτια και το στόμα και ίδιων διαστάσεων με τις εικόνες εκπαίδευσης, ακολουθούμε τα εξής βήματα:

Αφαιρούμε την μέση εικόνα  $\Psi: \Phi^t = \Gamma^t - \Psi$

Έπειτα την προβάλλουμε στον χώρο των ιδιοδιανυσμάτων  $\Phi^t = \sum_{j=1}^K u_j^T \Phi^t u_j$

Αναπαριστάμε το  $\Phi^t$  ως  $\Omega^t = [w_1^t, w_2^t, \dots, w_k^t]$  και βρίσκουμε την ελάχιστη απόσταση από τις εικόνες εκπαίδευσης:

$$e_r = \min_i (| |\Omega^t - \Omega^i| | ) \quad (18)$$

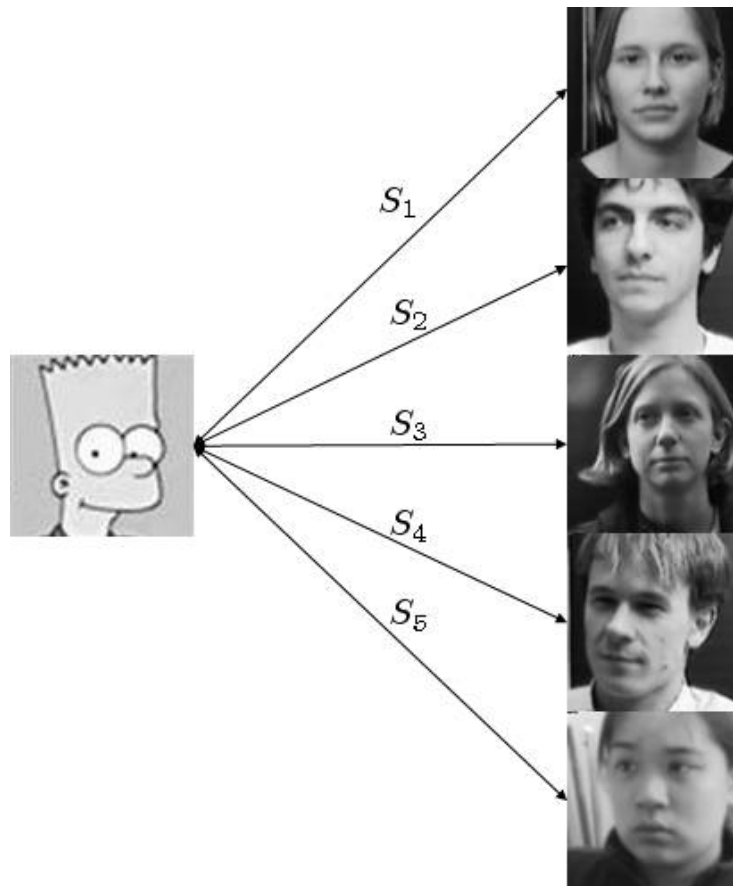
Αν το  $e_r$  είναι μικρότερο από ένα κατώφλι  $Tr$  και αντιστοιχεί στην απόσταση από το πρόσωπο εκπαίδευσης  $r$  τότε το πρόσωπο προς εξέταση αναγνωρίστηκε ως το πρόσωπο  $r$ .

Εάν όμως το  $e_r$  είναι μεγαλύτερο από το κατώφλι  $Tr$  τότε το πρόσωπο προς εξέταση δεν ανήκει στη βάση δεδομένων. Παρακάτω θα δούμε πως πρέπει το όριο να πρέπει να επιλεγεί.

#### **Η απόφαση σχετικά με το όριο:**

Γιατί είναι το όριο, σημαντικό;

Σκεφτείτε για απλότητα έχουμε ΜΟΝΟ 5 εικόνες στο σύνολο εκπαίδευσης. Και μία εικόνα που δεν είναι στο σύνολο εκπαίδευσης δίνεται για αναγνώριση. Η βαθμολογία για κάθε μία από τις εικόνες 5 θα βρεθεί έξω με την εισερχόμενη εικόνα. Και ακόμη και αν μια εικόνα δεν είναι στη βάση δεδομένων, θα εξακολουθούν να λένε ο ανιχνευτής αναγνωρίζεται ως η εικόνα κατάρτισης με την οποία η βαθμολογία της είναι η χαμηλότερη. Είναι σαφές ότι αυτή είναι μια ανωμαλία που πρέπει να εξετάσουμε. Είναι για το σκοπό αυτό που θα αποφασίσει το όριο.



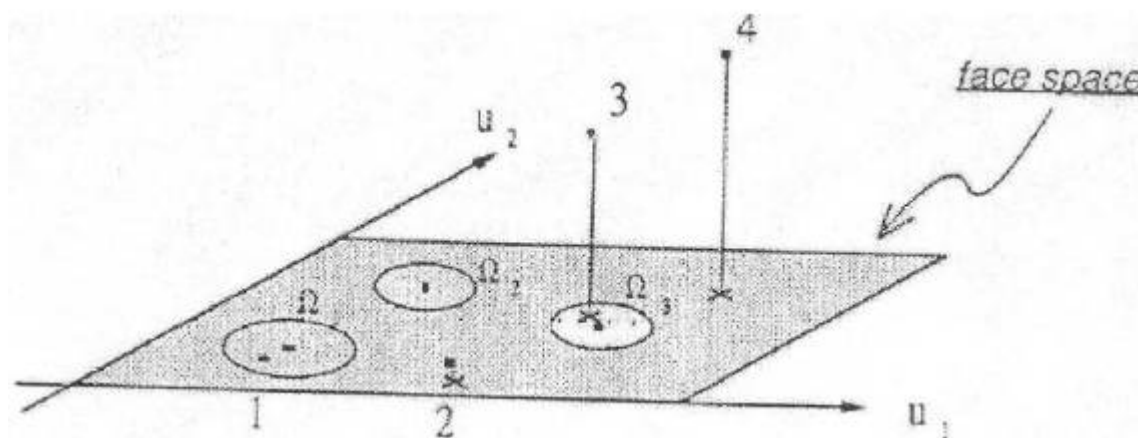
**Εικόνα 15:** Εικόνα προς εξέταση

Λαμβάνοντας υπ' όψιν μια εικόνα Simpson ως μη-πρόσωπο εικόνα, αυτή η εικόνα θα συγκριθεί με κάθε μία από τις εικόνες της κατάρτισης. Ας πούμε ότι  $S_4$  είναι το χαμηλότερο σκορ από όλους. Αλλά η εικόνα ανίχνευσης σαφώς δεν ανήκει στη βάση δεδομένων. Για να ορίσουμε το κατώφλι επιλέγουμε ένα μεγάλο σύνολο από τυχαίες εικόνες (τόσο για το πρόσωπο και για το μη-πρόσωπο), και στη συνέχεια υπολογίζουμε τα αποτελέσματα για τις εικόνες των ανθρώπων στη βάση δεδομένων και επίσης για αυτό το τυχαίο σύνολο ρυθμίζει το κατώφλι αναλόγως.



### Περισσότερα για το Face space :

Εδώ είναι μια σύντομη συζήτηση σχετικά με το χώρο πρόσωπο.



Εικόνα 16: Face space

Θεωρείστε μια απλοποιημένη αναπαράσταση του προσωποχώρου, όπως φαίνεται στο παραπάνω σχήμα. Οι εικόνες από ένα πρόσωπο και ιδίως τα πρόσωπα στο σύνολο της κατάρτισης θα πρέπει να βρίσκονται κοντά στον προσωποχώρο. Κάθε πρόσωπο στο σύνολο εκπαίδευσης μετασχηματίζεται εντός του χώρου, το πρόσωπο και τα συστατικά του είναι αποθηκευμένα στη μνήμη. Ένα πρόσωπο εισόδου δίνεται στο σύστημα, και ακολούθως προβάλλεται επί του προσωποχώρου. Το σύστημα υπολογίζει την απόστασή του από όλα τα αποθηκευμένα τα πρόσωπα. Ωστόσο, τα δύο θέματα πρέπει να εξεταστούν προσεκτικά:

1. Τι θα συμβεί αν η εικόνα που παρουσιάζουν στο σύστημα δεν είναι ένα πρόσωπο;
2. Τι θα συμβεί αν το πρόσωπο που παρουσιάζονται στο σύστημα δεν έχει ήδη μάθει, δηλαδή, δεν αποθηκεύεται ως γνωστό πρόσωπο;

Το πρώτο ελάττωμα εύκολα αποφεύγεται δεδομένου ότι το πρώτο eigenface είναι μια καλή όψη φίλτρου οποία μπορεί να εξετάσει αν κάθε εικόνα σχετίζεται σε μεγάλο βαθμό με τον εαυτό της. Οι εικόνες με χαμηλή συσχέτιση μπορούν να απορριφθούν. Η μήπως τα δύο αυτά θέματα αντιμετωπίζονται με κατηγοριοποίηση σε τέσσερις διαφορετικές περιοχές:

Κοντά στο προσωποχώρο και αποθηκεύονται πρόσωπο  $\Rightarrow$  γνωστά πρόσωπα .

Αυτό συμβαίνει όταν η εικόνα δεν είναι τίποτα, αλλά η εικόνα του προσώπου ενός ατόμου είναι γνωστή (γνωστό = εικόνα του προσώπου αυτού είναι ήδη στη βάση δεδομένων).

Κοντά στο προσωποχώρο, αλλά όχι κοντά σε ένα γνωστό πρόσωπο  $\Rightarrow$  άγνωστα πρόσωπα

Αυτό συμβαίνει όταν η εικόνα είναι καθετήρας ενός ατόμου (δηλαδή την εικόνα του προσώπου), αλλά δεν ανήκει σε κανέναν στη βάση δεδομένων δηλαδή μακριά από κάθε κατηγορία προσώπου.

Μακριά από το προσωποχώρο και κοντά σε ένα πρόσωπο κλάση  $\Rightarrow$  μη-πρόσωπα

Αυτό συμβαίνει όταν η εικόνα καθετήρα δεν είναι από ένα πρόσωπο, ωστόσο μοιάζει ακόμα μια συγκεκριμένη κατηγορία προσώπου αποθηκεύονται στη βάση δεδομένων.

Μακριά από το προσωποχώρο και όχι κοντά σε κλάση =  $\Rightarrow$  μη-πρόσωπα

Από τους τέσσερις, η τύπου 3 είναι υπεύθυνη για τα περισσότερα ψευδώς θετικά αποτελέσματα. Δεδομένου ότι ένα πρόσωπο έχει καλά εκπροσωπηθεί από τον προσωποχώρο, η ανοικοδόμηση του θα πρέπει να είναι παρόμοια με την αρχική, επομένως το σφάλμα ανασυγκρότησης θα είναι μικρό. Μη πρόσωπο εικόνες θα έχουν μια μεγάλη ανοικοδόμηση σφάλμα το οποίο είναι μεγαλύτερο από ένα ορισμένο κατώφλι  $thr$ . Η  $er$  απόσταση προσδιορίζει αν η εικόνα εισόδου είναι κοντά σε ένα γνωστό πρόσωπο.

#### 4.1.7 Independent Component Analysis (ICA)

Η μέθοδος PCA ανήκει στην κατηγορία των στατιστικών μεθόδων που βρίσκουν ένα σύνολο από εικόνες βάσης και αναπαριστούν τα πρόσωπα ως ένα γραμμικό συνδυασμό των εικόνων αυτών. Διαχωρίζει τις συσχετίσεις των δεδομένων εισόδου αλλά δεν επεκτείνεται στις υψηλής τάξης εξαρτήσεις όπως οι σχέσεις μεταξύ τριών και περισσότερων pixels [57].

Σε μια διεργασία, όμως, όπως η αναγνώριση προσώπου ή αναγνώριση εκφράσεων, στην οποία σημαντική πληροφορία μπορεί να περιέχεται στις υψηλής τάξης ροπές. Η Ανάλυση Ανεξάρτητων Συνιστωσών (Independent Component Analysis – ICA) που αποτελεί μια γενίκευση της μεθόδου PCA, είναι μία τέτοια μέθοδος η οποία προσπαθεί να υπολογίσει μία στατιστικά ανεξάρτητη δυναμική βάση.

Για εφαρμογές επεξεργασίας εικόνων προσώπου έχουν αναπτυχθεί δύο αρχιτεκτονικές βασισμένες στην Ανάλυση Ανεξάρτητων Συνιστωσών [58]. Η πρώτη αρχιτεκτονική (ARCHITECTURE I) παράγει στατιστικά ανεξάρτητες εικόνες βάσεις ενώ η δεύτερη (ARCHITECTURE II) παράγει στατιστικά ανεξάρτητους συντελεστές. Στην εργασία η αναπαράσταση των εικόνων προσώπου έγινε με βάση την πρώτη αρχιτεκτονική.

Σύμφωνα με την αρχιτεκτονική αυτή, η είσοδος του αλγορίθμου είναι ένας πίνακας  $X$  στις γραμμές του οποίου έχουν τοποθετηθεί  $n$  εικόνες, δηλαδή οι εικόνες του συνόλου εκπαίδευσης, με τη μορφή διανυσμάτων μήκους  $N$ , όπου  $N$  το πλήθος pixels των εικόνων. Στόχος της μεθόδου ICA είναι να υπολογίσει μια στατιστικά ανεξάρτητη διανυσματική βάση  $U$  και να παραστήσει τις εικόνες προσώπου του  $X$  ως ένα γραμμικό συνδυασμό της βάσης αυτής. Η διαδικασία συνοψίζεται στην εξίσωση

$$U = WX \quad (19)$$

όπου  $W$  ένας  $n \times n$  αντιστρέψιμος πίνακας. Η διανυσματική βάση  $U$ , διάστασης  $n \times N$  και ο πίνακας  $W$  υπολογίζονται από τη μέθοδο ICA.

Για την υλοποίηση της μεθόδου, δηλαδή των υπολογισμών των πινάκων  $W$  και  $U$ , χρησιμοποιήθηκε ο αλγόριθμος Infomax που προτάθηκε από τους Bell και Sejnowski [59] και προέρχεται από την αρχή βελτίωσης μεταφοράς πληροφορίας σε νευρώνες με σιγμοειδείς συναρτήσεις μεταφοράς.

Ο αλγόριθμος συνίσταται ως εξής: Θεωρούμε τους πίνακες  $X$  και  $W$ . Που αναφέρθηκαν προηγουμένως και συνδέονται με τη σχέση 19 και τον πίνακα  $Y = f(U)$  που αναπαριστά τις εξόδους

των  $n$  – νευρώνων και έχει ίδιες διαστάσεις με τον  $X$ . Κάθε στοιχείο της  $f=(f_1, \dots, f_n)$  είναι μια αντιστρέψιμη συνάρτηση που απεικονίζει πραγματικούς αριθμούς στο διάστημα  $[0,1]$ . Τυπικά η συνάρτηση που υλοποιείται είναι η

$$f(U) = \frac{1}{1+e^{-1}} \quad (20)$$

Οι γραμμές του  $U$ , δηλαδή τα διανύσματα  $u_1, \dots, u_n$ , είναι γραμμικοί συνδυασμοί των εισόδων και μπορούν να ερμηνευθούν ως προσυναπτικοί ενεργοποιητές των  $n$  – νευρώνων. Οι γραμμές του  $Y$ , τα διανύσματα  $y_1, \dots, y_n$ , μπορούν να ερμηνευθούν ως μετασυναπτικοί βαθμοί ενεργοποίησης και περιορίζονται στο διάστημα  $[0,1]$ . Σκοπός του αλγορίθμου είναι να μεγιστοποιηθεί η κοινή πληροφορία μεταξύ της εισόδου και της εξόδου του νευρωνικού δικτύου  $Y$  [60]. Αυτό επιτυγχάνεται με τη βαθμωτή άνοδο της εντροπίας της εξόδου σε σχέση με τον πίνακα βαρών  $W$ . Ο αλγόριθμος είναι επαναληπτικός και ο κανόνας εκπαίδευσης του πίνακα βαρών  $W$  προκύπτει από τη σχέση

$$\Delta W \propto \nabla W H(Y) W^T W = (I + Y^t U^T) W \quad (21)$$

όπου  $I$  ο μοναδιαίος πίνακας. Η χρήση της λογιστικής συνάρτησης 20 δίνει  $v_i = (1 - 2y_i)$ .

Όταν υπάρχουν πολλαπλές εισοδοί και εξοδοί, η μεγιστοποίηση της κοινής εντροπίας των εξόδων  $Y$  ενθαρρύνει τις εξόδους στον  $U$  να κινηθούν προς την στατιστική ανεξαρτησία. Όταν η μορφή των μη γραμμικών συναρτήσεων μεταφοράς  $f$  είναι ίδια με τις αθροιστικές συναρτήσεις πυκνότητας των Independent Components (ICs), μεγιστοποιώντας την κοινή εντροπία των εξόδων  $Y$  ελαχιστοποιείται η κοινή πληροφορία μεταξύ των εξόδων του  $U$  [60].

Ο αλγόριθμος περιλαμβάνει και μια διαδικασία ‘sphering’ πριν τη διαδικασία εκπαίδευσης. Οι μέσοι όροι γραμμών αφαιρούνται από το σύνολο δεδομένων εισόδου  $X$  και έπειτα ο  $X$  διέρχεται από ένα μηδενικής φάσης whitening φίλτρο  $W_Z$  το οποίο ισούται με το διπλάσιο της τετραγωνικής ρίζας του πίνακα συμμεταβλητότητας

$$W_Z = 2(XX^T)^{1/2} \quad (22)$$

Η διαδικασία αυτή αφαιρεί τις πρώτης και τις δεύτερης τάξης στατιστικές των δεδομένων καθώς οι διακυμάνσεις αντισταθμίζονται. Όταν οι εισοδοί στον αλγόριθμο ICA έχουν υποβληθεί σε διαδικασία sphering, ο πίνακας του μετασχηματισμού της εξίσωσης 19 είναι ο  $W_I$  ο οποίος προκύπτει από το γινόμενο του πίνακα  $W_Z$  και του πίνακα που εκπαιδεύτηκε κατά τη διαδικασία ICA

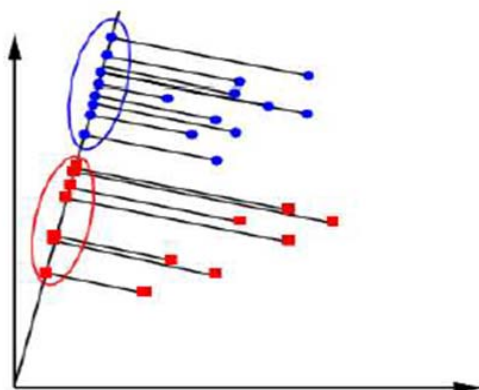
$$W_I = W W_Z \quad (23)$$

Επομένως η σχέση 19 αντικαθίσταται από την

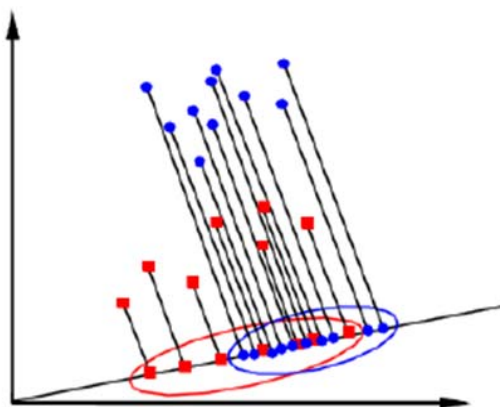
$$U = W_I X \quad (24)$$

#### 4.1.8 Γραμμική Διακρίνουσα Ανάλυση (Linear Discriminant Analysis, LDA)

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 [61]. 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 δισδιάστατο χώρο σε ένα χώρο  $C$  διαστάσεων, όπου  $C_0$  αριθμός των κλάσεων των εικόνων (ή διαφορετικά ο αριθμός των διαφορετικών ατόμων στη βάση των προσώπων). Για παράδειγμα, έστω ότι δύο σύνολα σημείων του δισδιάστατου χώρου προβάλλονται σε μία ευθεία. Από την κλίση της ευθείας αυτής εξαρτάται αν τα σημεία θα διαχωρίζονται πλήρως όπως φαίνεται στην Εικόνα 17 ή θα υπάρχουν κάποια που θα συμπίπτουν (Εικόνα 18).



**Εικόνα 17:** Σημεία που διαχωρίζουν πλήρως σε δυο κλάσεις κατά την προβολή τους σε ευθεία



**Εικόνα 18:** Σημεία που συμπίπτουν χωρίς να είναι εφικτός ο ξεκάθαρος διαχωρισμός τους σε κλάσεις όταν προβάλλονται σε ευθεία

Η γραμμική διακρίνουσα του Fisher επιδιώκει να βρει εκείνη την ευθεία που διαχωρίζει τα σημεία με το βέλτιστο δυνατό τρόπο. Έτσι προκειμένου να αναγνωριστεί μία εικόνα, συγκρίνεται η προβολή αυτής με κάθε μία από τις προβολές των εικόνων του συνόλου εκπαίδευσης (training set). Έπειτα, η εικόνα προς αναγνώριση θεωρείται «όμοια» με εκείνη του συνόλου εκπαίδευσης που παρουσιάζει τη μικρότερη απόσταση μεταξύ των προβολών τους.

Ένας εντός κλάσης πίνακας διασποράς (within-class scatter matrix) αναπαριστά διαφοροποιήσεις στην εμφάνιση του ίδιου προσώπου εξαιτίας είτε του διαφορετικού φωτισμού είτε της έκφρασης του προσώπου. Ο εντός κλάσης πίνακας διασποράς  $S_w$  ορίζεται ως:

$$S_w = \sum_{j=1}^C \sum_{i=1}^{N_j} (\Gamma_i^j - \mu_j)(\Gamma_i^j - \mu_j)^T \quad (25)$$

όπου  $\Gamma_i^j$  είναι το  $i$ -οστό δείγμα (διάνυσμα εικόνας προσώπου) της κλάσης  $j$ ,  $\mu_j$  ο μέσος όρος της κλάσης  $j$ ,  $C_0$  αριθμός των κλάσεων και  $N_j$  αριθμός των εικόνων- δειγμάτων της κλάσης  $j$ . Από την άλλη μεριά, ο ανάμεσα στις κλάσεις πίνακας διασποράς (between-class scatter matrix) περιγράφει τις διαφορές που υπάρχουν λόγω της διαφορετικής ταυτότητας του προσώπου. Ο ανάμεσα στις κλάσεις πίνακας διασποράς  $S_b$  δίδεται από την σχέση:

$$S_b = \sum_{j=1}^C (\mu_j - \mu)(\mu_j - \mu)^T \quad (26)$$

όπου  $\mu$  ο μέσος όρος όλων των κλάσεων.

Εφαρμόζοντας αυτή τη μέθοδο, επιδιώκεται η εύρεση εκείνης της διεύθυνσης προβολής που αφενός μεγιστοποιεί την απόσταση μεταξύ των εικόνων που ανήκουν σε διαφορετικές κλάσεις και αφετέρου ελαχιστοποιεί την απόσταση των εικόνων που είναι μέλη της ίδιας κλάσης. Με άλλα λόγια, ζητούμενο είναι η μεγιστοποίηση του ανάμεσα στις κλάσεις πίνακα διασποράς  $S_b$  (Εξίσωση 26) και η ταυτόχρονη ελαχιστοποίηση του εντός κλάσης πίνακα διασποράς  $S_w$  (Εξίσωση 25) στον υποχώρο.

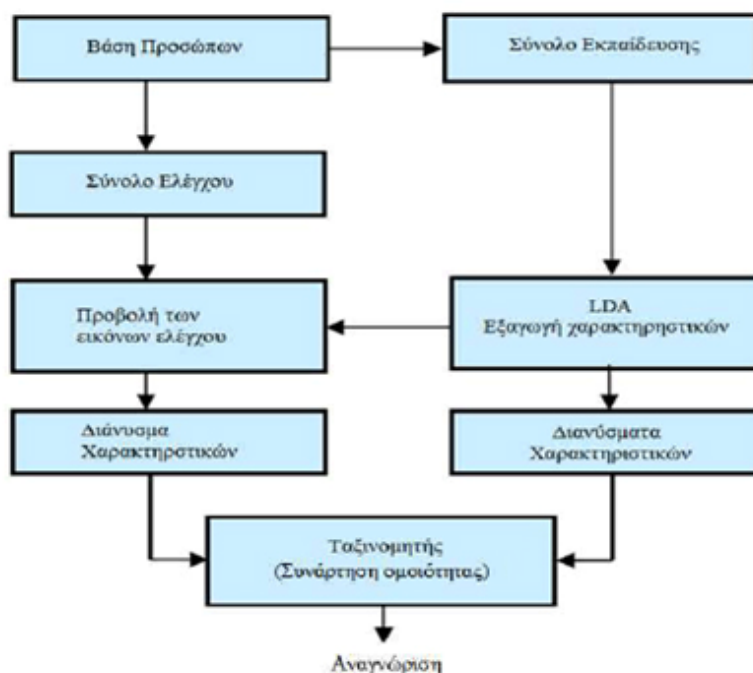
Η Εικόνα 19 που ακολουθεί δείχνει έναν καλό και ένα κακό διαχωρισμό κλάσεων.



**Εικόνα 19:** (α) Καλός διαχωρισμός κλάσεων. (β) Κακός διαχωρισμός κλάσεων

Με τη μέθοδο LDA η χρήση γραμμικών μεθόδων σε κλάσεις για τη μείωση της διάστασης του χώρου καθώς και απλών ταξινομητών στον υποχώρο των χαρακτηριστικών, μπορεί να επιτύχει καλά αποτελέσματα στην αναγνώριση προσώπων σε σημαντικά μικρό χρόνο [62]. Συνεπώς, η Γραμμική Διακρίνουσα του Fisher [63] είναι ένα παράδειγμα μίας μεθόδου που χρησιμοποιεί κλάσεις και προσπαθεί να διαμορφώσει έτσι τον πίνακα διασποράς τους ώστε να επιτευχθεί καλύτερη ταξινόμηση. Το κριτήριο που προτάθηκε από τον Fisher, προσπαθεί να μεγιστοποιήσει το λόγο της ορίζουσας του ανάμεσα στις κλάσεις πίνακα διασποράς (between-class scatter matrix) προς την ορίζουσα του εντός κλάσης πίνακα διασποράς (within-class scatter matrix). Πιο αναλυτικά, όπως και στην περίπτωση των ιδιοπροσώπων (PCA), οι εικόνες του συνόλου εκπαίδευσης προβάλλονται σε έναν υποχώρο. Στον ίδιο υποχώρο προβάλλονται και οι εικόνες προς αναγνώριση ενώ

χρησιμοποιείται κάποια συνάρτηση ομοιότητας προκειμένου να επιτευχθεί η αναγνώριση. Αυτό που διαφέρει στις δύο προσεγγίσεις είναι ο τρόπος υπολογισμού αυτού του υποχώρου. Η εικόνα υπό αναγνώριση θεωρείται ότι αντιστοιχεί σε εκείνη του συνόλου εκπαίδευσης που η μεταξύ τους απόσταση είναι η ελάχιστη. Τα παραπάνω συνοψίζονται στην Εικόνα 25 που ακολουθεί και περιγράφει σχηματικά τον τρόπο λειτουργίας του αλγορίθμου γραμμικής διακρίνουσας ανάλυσης [64].



**Εικόνα 20:** Αρχή λειτουργίας του αλγορίθμου γραμμικής διακρίνουσας ανάλυσης (LDA)

Σε αντίθεση με τη μέθοδο της ανάλυσης σε Κύριες Συνιστώσες (Principal Component Analysis, PCA) που εξαγάγει χαρακτηριστικά με σκοπό να αναπαραστήσει με βέλτιστο τρόπο τις εικόνες, η μέθοδος της γραμμική διακρίνουσας ανάλυσης (LDA) προσπαθεί να βρει εκείνο τον υποχώρο που διαχωρίζει καλύτερα τις διαφορετικές κλάσεις προσώπων όπως φαίνεται στην Εικόνα 17.

Ο LDA βρίσκει ένα σύνολο  $M$  βασικών διανυσμάτων  $\{w_m\}_{m=1}^M$ , έτσι ώστε ο λόγος της ορίζουσας του ανάμεσα στις κλάσεις πίνακα διασποράς (between-class scatter matrix), προς την ορίζουσα του εντός κλάσης πίνακα διασποράς (within-class scatter matrix) να είναι μέγιστος. Το πρόβλημα αυτό ορίζεται από την σχέση (27):

$$W_{opt} = \operatorname{argmax}_W \left| \frac{W^T S_b W}{W^T S_w W} \right| = [w_1 \dots w_M] \quad (27)$$

Υποθέτοντας ότι ο εντός κλάσης πίνακας διασποράς,  $S_w$ , είναι αντιστρέψιμος τα διανύσματα  $\{W_i | i=1 \dots M\}$  μπορούν να υπολογιστούν ως τα πρώτα  $M$  ιδιοδιανύσματα με τις μεγαλύτερες ιδιοτιμές του  $S_w^{-1} S_b$ , τα οποία ονομάζονται διαφορετικά fisherfaces. Ο εντός κλάσης πίνακας διασποράς «συλλαμβάνει» τον τρόπο με τον οποίο κατανέμονται μέσα στην εκάστοτε κλάση οι εικόνες των προσώπων. Αναλυτικότερα οι διαφορετικές εικόνες του προσώπου πρέπει να κατανέμονται πολύ κοντά η μία με την άλλη. Από την άλλη μεριά, ο ανάμεσα στις κλάσεις πίνακας διασποράς αποτυπώνει τον τρόπο με τον οποίο διαχωρίζονται οι κλάσεις μεταξύ τους. Όταν οι εικόνες προβάλλονται στα διανύσματα  $W$ , θα πρέπει να κατανέμονται όσο το δυνατόν πιο κοντά μέσα στις κλάσεις και να διαχωρίζονται ανάμεσα στις κλάσεις [65].

Έτσι η αναπαράσταση των εικόνων από τον αρχικό χώρο στον υποχώρο των fisherfaces επιτυγχάνεται με την προβολή τους στον υποχώρο ο οποίος αποτελείται από τα  $M$  ιδιοδιανύσματα. Παρόλα αυτά μπορεί ο  $S_w$  να είναι μη αντιστρέψιμος εξαιτίας του γεγονότος ότι ο αριθμός των δειγμάτων εκπαίδευσης είναι συνήθως μικρότερος των διαστάσεων των εικόνων αυτών. Το πρόβλημα αυτό είναι γνωστό ως «μικρού αριθμού δειγμάτων» (small sample size problem, SSS) και συναντάται συχνά στην αναγνώριση προσώπων [65], [66], [67]. Μία πιθανή λύση του προβλήματος αυτού είναι να προηγηθεί η εφαρμογή ενός PCA βήματος με σκοπό την «αφαίρεση» του μηδενοχώρου (null space) του εντός κλάσης πίνακα διασποράς,  $S_w$  πριν την εύρεση των διανυσμάτων  $W$  με τη βοήθεια της (30s) [65], [68].

Μία παρόμοια μέθοδος αναπτύχθηκε από τον Friedman [69] και είναι αυτή της κανονικοποιημένης διακρίνουσας ανάλυσης δευτέρου βαθμού. Σε αυτή την περίπτωση κάθε πρόβλημα του πίνακα διασποράς για την εκάστοτε κλάση,  $S_i$ , μπορεί να είναι λανθασμένα τοποθετημένη (ill-posed). Η λύση που προτάθηκε από τον Friedman είναι η εισαγωγή ενός όρου κανονικοποίησης. Συγκεκριμένα είναι πολλαπλάσιο του μοναδιαίου πίνακα,  $\gamma \cdot I$ , έτσι ώστε  $S_i' = S_i + \gamma \cdot I$ , όπου  $\gamma$  ο όρος κανονικοποίησης και  $I$  ο μοναδιαίος πίνακας. Αυτή η κανονικοποίηση έχει σαν αποτέλεσμα τη μείωση της τιμής των μεγαλύτερων ιδιοτιμών και την αύξηση των μικρότερων εξομαλύνοντας έτσι την πόλωση. Ένα άλλο πλεονέκτημά της είναι η σταθεροποίηση των μικρότερων ιδιοτιμών. Επιπρόσθετα, πρέπει να αναφερθεί ότι στη μέθοδο αυτή ο εντός κλάσης πίνακας διασποράς  $S_w$  είναι ισοδύναμος με το μέσο όρο των πινάκων διασποράς όλων των κλάσεων  $S_w = \frac{1}{C} \sum_{i=1}^C S_i$

Εξαιτίας της επιτυχίας της μεθόδου του Friedman, προτάθηκε μία παραλλαγή της η οποία και υλοποιείται στην παρούσα εργασία. Στην παραλλαγή αυτή το κριτήριο του Fisher λαμβάνει την εξής μορφή:

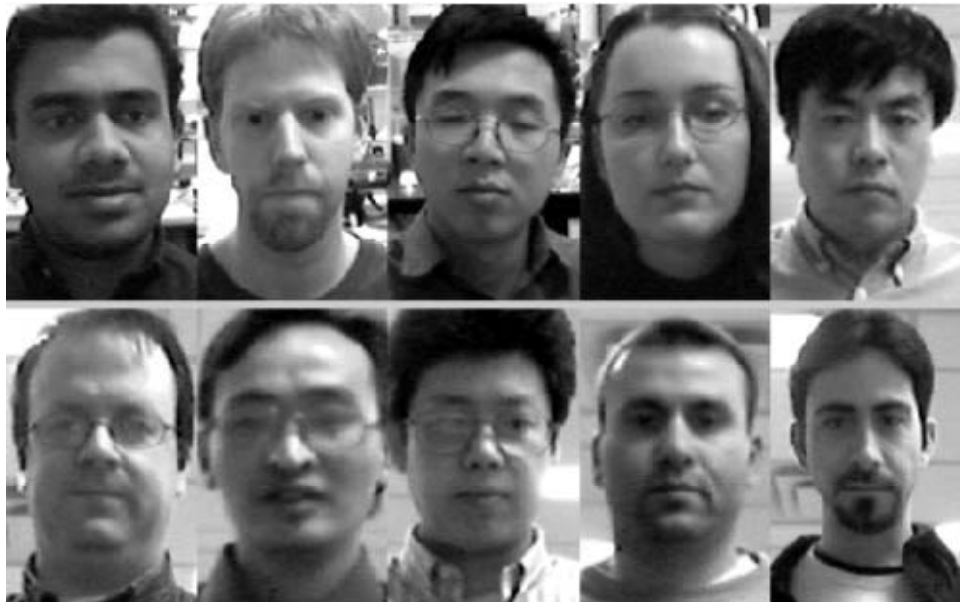
$$w = \operatorname{argmax}_w \frac{|w^T S_b w|}{|\eta (w^T S_b w) + (w^T S_w w)|} \quad (28)$$

όπου  $0 \leq \eta \leq 1$  είναι ο όρος κανονικοποίησης. Παρότι η εξίσωση (28) φαίνεται εντελώς διαφορετική από αυτή της σχέσης (27) είναι δυνατό να αποδειχτεί ότι είναι ισοδύναμες [70].

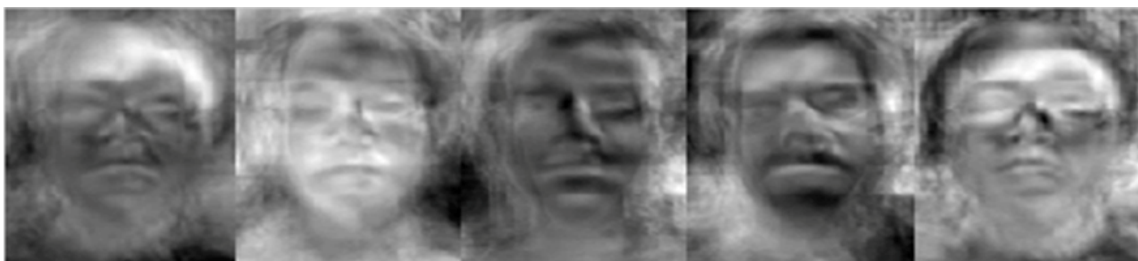
Το τροποποιημένο κριτήριο του Fisher είναι μία συνάρτηση του όρου  $\eta$ , ο οποίος ελέγχει τη δύναμη της κανονικοποίησης. Όταν  $\eta=0$ , το τροποποιημένο κριτήριο του Fisher μετατρέπεται στο τυπικό και συνηθισμένο κριτήριο που δεν εισάγει την έννοια της κανονικοποίησης. Αντίθετα όταν  $\eta=1$ , η κανονικοποίηση είναι η μεγαλύτερη δυνατή.

Με βάση τα παραπάνω προτάθηκε η κανονικοποιημένη γραμμική διακρίνουσα ανάλυση (regularized LDA, r-LDA), που προσπαθεί να βελτιστοποιήσει το κανονικοποιημένο κριτήριο του Fisher όπως αποτυπώνεται από την εξίσωση (28) [70]. Η r-LDA μέθοδος αρχικά βρίσκει το συμπληρωματικό χώρο του  $S_b$ . Έστω  $U_m = [u_1, \dots, u_m]$  είναι τα ιδιοδιανύσματα του  $S_b$  που αντιστοιχούν στις πρώτες  $m$  μεγαλύτερες μη μηδενικές ιδιοτιμές  $\Lambda_b$ , όπου  $m \leq C - 1$ . Ο συμπληρωματικός χώρος  $A'$  απαρτίζεται από τα  $U_m$  και κλιμακώνεται (scale) από  $H = U_m \Lambda_b^{-1/2}$  έτσι ώστε  $H^T S_b H = I$ , όπου  $I$  είναι ο μοναδιαίος πίνακας διαστάσεων  $m \times m$ . Έτσι, μπορεί να αποδειχτεί ότι ο παρονομαστής της εξίσωσης (28) μετασχηματίζεται στην κανονικοποιημένη έκφραση του Friedman,  $\eta I + H^T S_w H$ , στον υποχώρο που οριοθετείται από τα  $H$ .

Έπειτα ζητούμενο είναι η εύρεση των χαρακτηριστικών βάσης, που ελαχιστοποιούν τον κανονικοποιημένο παρονομαστή. Δεν είναι δύσκολο να παρατηρηθεί ότι τα ζητούμενα χαρακτηριστικά αντιστοιχίζονται στα  $M (\leq m)$  ιδιοδιανύσματα του  $H^T S_w H, P_M = [p_1, \dots, p_M]$ , με τις μικρότερες ιδιοτιμές  $\Lambda_w$ . Συνδυάζοντας τα αποτελέσματα αυτά μπορούμε να εξαγάγουμε την ζητούμενη λύση  $W = H P_M (\eta I + \Lambda_w)^{-1/2}$  που αποτελεί το σύνολο των διανυσμάτων βάσης με τη βέλτιστη διαχωριστικότητα. Για το δείγμα των εικόνων εκπαίδευσης της Εικόνας 21(α) τα διανύσματα αυτά φαίνονται στην Εικόνα 21(β).



(α)



(β)

**Εικόνα 21:** (α)(β) Τα πέντε πρώτα διανύσματα βάσης του r-LDA



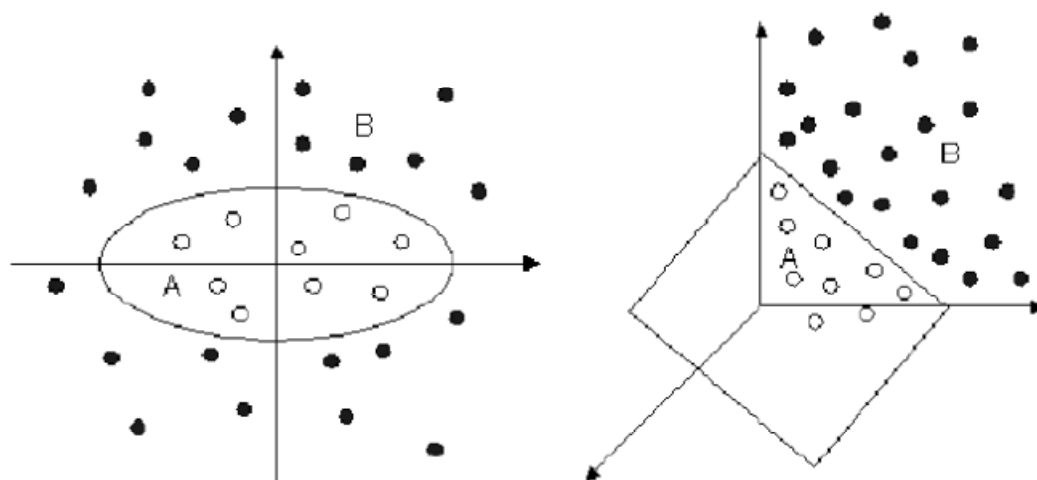
Τέλος, εφόσον ο υποχώρος που ορίζεται από τα  $W$  μπορεί να εμπεριέχει την τομή  $(A \cap B)$ , είναι πιθανό να υπάρχουν μηδενικές ή πολύ μικρές ιδιοτιμές στο  $\Lambda_w$ , που να έχουν υψηλή διακύμανση κατά την προβολή σε περιβάλλοντα όπου είναι εμφανές το πρόβλημα «μικρού αριθμού δειγμάτων» [69]. Αυτό έχει σαν αποτέλεσμα, κάθε πόλωση (bias) που προκύπτει από τα ιδιοδιανύσματα που αντιστοιχούν σε αυτές τις ιδιοτιμές να μεγαλώνει εξαιτίας της κανονικοποίησης ( $P_M \Lambda_w^{-1/2}$ ) στη σχέση  $W = HP_M(\eta I + \Lambda_w)^{-1/2}$ . Αντίθετα, η εισαγωγή της κανονικοποίησης βοηθά στη μείωση της σημαντικότητας αυτών των ιδιαίτερα ασταθών ιδιοδιανυσμάτων και συνεπώς μειώνεται η συνολική διακύμανση. Επιπλέον, αν ο παράγοντας κανονικοποίησης είναι  $\eta=0$ , τότε οι μηδενικές ιδιοτιμές  $\Lambda_w$  μπορεί να δημιουργήσουν πρόβλημα εφόσον χρησιμοποιούνται ως διαιρέτες. Το πρόβλημα επιλύεται εύκολα στη περίπτωση της κανονικοποιημένης γραμμικής διακρίνουσας ανάλυσης (r-LDA) θέτοντας στον παράγοντα κανονικοποίησης τιμή μεγαλύτερη του μηδενός.

#### 4.1.9 Αλγόριθμοι πυρήνα (Kernel Algorithms)

Οι προσεγγίσεις των αλγορίθμων PCA και LDA κωδικοποιούν την πληροφορία χρησιμοποιώντας εξαρτήσεις δευτέρας τάξεως όπως για παράδειγμα η συνδιακύμανση μεταξύ δύο εικονοστοιχείων (pixel) ενώ δεν λαμβάνουν υπόψη τους τις εξαρτήσεις που ενυπάρχουν ανάμεσα σε περισσότερα των δύο εικονοστοιχείων των εικόνων-δειγμάτων. Οι εξαρτήσεις υψηλότερης τάξης σε μία εικόνα συμπεριλαμβάνουν μη γραμμικές συσχετίσεις στις τιμές της έντασης των εικονοστοιχείων. Παράδειγμα αποτελούν οι σχέσεις τριών ή περισσότερων εικονοστοιχείων που παρατηρούνται σε ακμές ή καμπύλες και εμπεριέχουν χρήσιμη πληροφορία για την αναγνώριση. Πολλοί ερευνητές εικάζουν ότι οι στατικές υψηλής τάξης (higher order statistics, HOS) είναι καίριας σημασίας κατά την αναπαράσταση σύνθετων μοτίβων – σχεδίων. Πρόσφατα μάλιστα στατιστικές υψηλότερης τάξης έχουν εφαρμοστεί σε προβλήματα οπτικής εκμάθησης (visual learning).

Στην [71] οι συγγραφείς χρησιμοποίησαν στατιστικές υψηλότερης τάξης σε εικόνες ενός αντικειμένου-στόχου προκειμένου να λάβουν μία καλύτερη εκτίμηση της άγνωστης διασποράς του. Πειράματα σε αναγνώριση προσώπων έδειξαν συγκρίσιμα, αν όχι καλύτερα, αποτελέσματα από άλλους αλγορίθμους που στηρίζονται στον PCA. Στην [72] επεκτάθηκε ο τυπικός PCA στην λεγόμενη ανάλυση σε Κύριες Συνιστώσες του πυρήνα (Kernel PCA, KPCA). Εμπειρικά αποτελέσματα τόσο στην αναγνώριση ψηφίων όσο και στην αναγνώριση αντικειμένων έδειξαν ότι ο Kernel PCA είναι ικανός να εξαγάγει μη γραμμικά χαρακτηριστικά και συνεπώς να οδηγήσει σε καλύτερα αποτελέσματα. Πρόσφατα οι [73], [74] και [75] εφάρμοσαν τεχνάσματα του πυρήνα (kernel tricks) στην γραμμική διακρίνουσα ανάλυση δημιουργώντας έτσι τον KFLD. Τα πειράματά τους απέδειξαν ότι αυτός ο αλγόριθμος δύναται να εξαγάγει εκείνα τα χαρακτηριστικά που διαχωρίζουν τον χώρο με τον καλύτερο δυνατό τρόπο. Αυτό είναι ισοδύναμο με την εύρεση των μη γραμμικών χαρακτηριστικών που εμπεριέχουν σημαντική πληροφορία για το διαχωρισμό του αρχικού χώρου.

Η βασική ιδέα των αλγορίθμων του πυρήνα μπορεί να γίνει καλύτερα κατανοητή αν χρησιμοποιήσουμε το ακόλουθο παράδειγμα (Εικόνα 27), όπου δισδιάστατα δείγματα εισόδου, έστω  $z = [z_1, z_2]$ , προβάλλονται στον τρισδιάστατο «χώρο χαρακτηριστικών» μέσω μίας μη γραμμικής συνάρτησης:  $\varphi: z = [z_1, z_2] \leftarrow \varphi(z) = [x_1, x_2, x_3] := [z_1^2, \sqrt{2} z_1 z_2, z_2^2]$ . [76]. Η Εικόνα 27 δείχνει ότι στο χώρο εισόδου ένα μη γραμμικό ελλειψοειδές όριο απόφασης χρειάζεται για το διαχωρισμό των κλάσεων A και B. Αντίθετα, οι δύο κλάσεις διαχωρίζονται γραμμικά στον υψηλότερης διάστασης χώρο των χαρακτηριστικών F.



**Εικόνα 22:** Ένα παράδειγμα κατηγοριοποίησης προτύπων σε δύο κλάσεις [76]. Στα αριστερά: δείγματα στον διδιάστατο χώρο εισόδου, όπου χρειάζεται ένα μη γραμμικό ελλειψοειδές όριο απόφασης (decision boundary) για το διαχωρισμό των κλάσεων A και B. Στα δεξιά: δείγματα που προβάλλονται στον τρισδιάστατο χώρο χαρακτηριστικών όπου ένα γραμμικό υπερεπίπεδο μπορεί να διαχωρίσει τις κλάσεις.

Προκειμένου να επιτευχθεί ο στόχος του γραμμικού διαχωρισμού, πρέπει η διάσταση του χώρου των χαρακτηριστικών να είναι ιδιαίτερα μεγάλη και πιθανώς άπειρη. Θετικό όμως είναι το γεγονός ότι δεν είναι απαραίτητη η εύρεση της  $\varphi(z)$  καθώς ο χώρος των χαρακτηριστικών μπορεί να υπολογιστεί έμμεσα με τη βοήθεια μηχανών πυρήνα (kernel machine). Το τέχνασμα που χρησιμοποιείται σε αυτές τις μεθόδους είναι η αντικατάσταση των εσωτερικών γινομένων στον χώρο των χαρακτηριστικών με μία συνάρτηση πυρήνα στο χώρο εισόδου. Για να γίνει αυτό πλήρως κατανοητό θα χρησιμοποιηθεί και πάλι το παράδειγμα της Εικόνας 22, όπου ο χώρος των χαρακτηριστικών αποτελείται από τα μονώνυμα δεύτερης τάξης του χώρου εισόδου. Έστω  $z_i \in \mathbb{R}^2$  και  $z_j \in \mathbb{R}^2$  είναι δύο δείγματα του χώρου εισόδου. Το εσωτερικό γινόμενο των διανυσμάτων χαρακτηριστικών  $\varphi(z_i) \in \mathbb{F}$  και  $\varphi(z_j) \in \mathbb{F}$  μπορεί να υπολογιστεί με την ακόλουθη συνάρτηση πυρήνα,  $k(z_i, z_j)$  που ορίζεται στο  $\mathbb{R}^2 \times \mathbb{R}^2$ :

$$\begin{aligned} \varphi(z_i) \cdot \varphi(z_j) &= [z_{i1}^2, \sqrt{2}z_{i1}z_{i2}, z_{i2}^2]^T [z_{j1}^2, \sqrt{2}z_{j1}z_{j2}, z_{j2}^2] = \\ &= ([z_{i1}z_{i2}]^T [z_{j1}z_{j2}])^2 = (z_i \cdot z_j)^2 =: k(z_i, z_j) \end{aligned} \quad (29)$$

Από το παράδειγμα αυτό φαίνεται ότι η κεντρική ιδέα για τη γενίκευση ενός αλγορίθμου γραμμικής εκμάθησης στην αντίστοιχη έκδοση του πυρήνα είναι η μετατροπή όλων των υπολογισμών του αλγορίθμου στο χώρο των χαρακτηριστικών έτσι ώστε να χρησιμοποιούνται μόνο εσωτερικά γινόμενα.

#### 4.1.10 Ανάλυση σε Κύριες Συνιστώσες πυρήνα (Kernel PCA)

Δοθέντος ενός συνόλου από  $m$  κεντρικά (centered) δείγματα  $x_k = [x_{k1}, x_{k2} \dots \dots \dots x_{kn}]^T$

$\in \mathbb{R}^N$ , ο αλγόριθμος PCA στοχεύει στην εύρεση εκείνων των διευθύνσεων προβολής που μεγιστοποιούν τη διακύμανση,  $C$ , των  $x_k$  κάτι που είναι ισοδύναμο με την εύρεση των ιδιοτιμών του πίνακα συνδιακύμανσης των δειγμάτων:

$$\lambda_{\chi} = C_w \quad (30)$$

όπου  $\lambda \geq 0$  οι ιδιοτιμές και  $w \in \mathbb{R}^n$  τα ιδιοδιανύσματα.

Στην ανάλυση Κύριων Συνιστωσών του πυρήνα, κάθε διάνυσμα  $x$  προβάλλεται από τον αρχικό χώρο, σε έναν χώρο χαρακτηριστικών υψηλής τάξης,  $\mathbb{R}^f$ , με τη βοήθεια μία μη γραμμικής συνάρτησης αντιστοίχισης (mapping function):  $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^f, f \gg n$ . Στην περίπτωση του χώρου  $\mathbb{R}^n$  το πρόβλημα εύρεσης ιδιοδιανυσμάτων ορίζεται ως εξής:

$$\lambda w^\Phi = C^\Phi w^\Phi \quad (31)$$

όπου  $C^\Phi$  ο πίνακας συνδιακύμανσης. Όλες οι λύσεις  $w^\Phi$  με  $\lambda \neq 0$  λαμβάνουν τιμές στο  $\Phi(x_1), \dots, \Phi(x_m)$ , και υπάρχουν συντελεστές  $a_i$  τέτοιοι ώστε:

$$W^\Phi = \sum_{i=1}^m a_i \Phi(x_i) \quad (32)$$

Θεωρώντας έναν  $m \times m$  πίνακα  $K$  τέτοιο ώστε:

$$K_{ij} = k(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j) \quad (33)$$

το πρόβλημα της ανάλυσης σε Κύριες Συνιστώσες του πυρήνα μετασχηματίζεται ως

εξής:

$$m \lambda K \alpha = K^2 \alpha \quad (34)$$

$$m \lambda \alpha = K \alpha \quad (35)$$

όπου  $\alpha$  ένα διάνυσμα στήλη με τιμές  $a_1, \dots, a_m$ . Οι παραπάνω σχέσεις βασίζονται στην υπόθεση ότι όλα τα δείγματα  $\Phi(x)$  που προβάλλονται είναι κεντρικά (centered) στο  $\mathbb{R}^f$ . Στην [72] έχει πραγματοποιηθεί ανάλυση του τρόπου με τον οποίο τα διανύσματα  $\Phi(x)$  μπορούν να μετατραπούν σε κεντρικά στο  $\mathbb{R}^f$ .

Σημειώνεται σε αυτό το σημείο ότι ο αλγόριθμος PCA είναι μία ειδική περίπτωση του kernel PCA με πολυωνυμικό πυρήνα πρώτης τάξης. Με άλλα λόγια, ο kernel PCA είναι μία γενίκευση της τυπικής ανάλυσης Κύριων Συνιστωσών εφόσον διαφορετικοί πυρήνες μπορούν να χρησιμοποιηθούν για να παράγουν διάφορες μη γραμμικές προβολές.

Είναι εφικτό λοιπόν να ληφθούν οι προβολές των διανυσμάτων του  $\mathbb{R}^f$  σε ένα χώρο μικρότερης διάστασης που ορίζεται από τα ιδιοδιανύσματα  $W^\Phi$ . Για παράδειγμα έστω ένα δείγμα  $x$  του οποίου η προβολή είναι  $\Phi(x)$  στο  $\mathbb{R}^f$ . Τότε η προβολή του  $\Phi(x)$  στα ιδιοδιανύσματα  $l$  είναι οι μη γραμμικές Κύριες Συνιστώσες που αντιστοιχούν στο  $\Phi$ :

$$W^\Phi \Phi(x) = \sum_{i=1}^m a_i (\Phi(x_i) \cdot \Phi(x)) = \sum_{i=1}^m a_i k(x_i, x) \quad (36)$$

Για κάποιες επιλογές πυρήνα  $k(x, y)$ , μπορεί να αποδειχτεί με τη βοήθεια της συναρτησιακής ανάλυσης ότι υπάρχει συνάρτηση αντιστοίχισης  $\Phi$  σε κάποιο χώρο εσωτερικών γινομένων  $F$

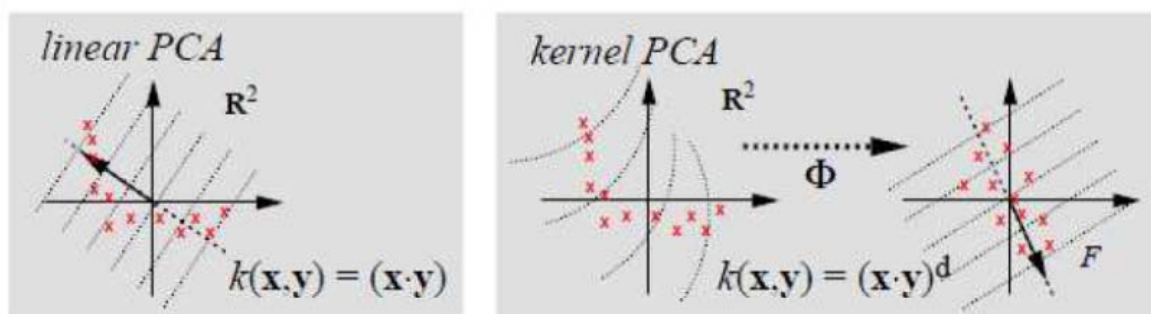
(πιθανόν άπειρης διάστασης) έτσι ώστε ο πυρήνας  $k$  να μπορεί να υπολογίσει τα εσωτερικά γινόμενα του χώρου  $F$ .

Οι πυρήνες που επιτυχώς χρησιμοποιήθηκαν σε μηχανές υποστήριξης διανύσματος (Support Vector Machines) [77] συμπεριλαμβάνουν τους πολυωνυμικούς  $k(x, y) = (xy)^d$ , τις συναρτήσεις ακτινωτής βάσης  $k(x, y) = \exp(-\frac{\|x-y\|^2}{2\sigma^2})$  και τους σιγμοειδείς  $k(x, y) = \tanh(\kappa(x \cdot y) + \theta)$ . Μπορεί να αποδειχτεί ότι οι πολυωνυμικοί πυρήνες βαθμού  $d$  είναι άρρηκτα συνδεδεμένοι με μία αντιστοίχιση  $\Phi$  στο χώρο των χαρακτηριστικών που αποτελείται από όλα τα εσωτερικά γινόμενα  $d$  παραγόντων ενός μοτίβου εισόδου. Για παράδειγμα αν το πλήθος των παραγόντων είναι δύο (2) και ο πολυωνυμικός βαθμός επίσης δύο (2) έχουμε:

$$(xy)^2 = (x_1^2, x_1x_2, x_2, x_2, x_1x_2^2)(y_1^2, y_1, y_2, y_2, y_1y_2^2)^T \quad (37)$$

Αν τα μοτίβα εισόδου αντιστοιχούν σε εικόνες, μπορεί να χρησιμοποιηθεί ο χώρος όλων των εσωτερικών γινομένων  $d$  εικονοστοιχείων και συνεπώς να ληφθούν υπόψη υψηλότερης τάξης στατιστικές όταν εφαρμόζεται ο αλγόριθμος PCA.

Έτσι αν αντικαταστήσουμε όλες τις εμφανίσεις του εσωτερικού γινομένου  $\Phi(x_i)\Phi(x)$  με συναρτήσεις πυρήνα λαμβάνουμε τον αλγόριθμο Kernel PCA η κύρια ιδέα του οποίου φαίνεται στην Εικόνα 28 που ακολουθεί.



**Εικόνα 28:** Βασική Ιδέα του αλγορίθμου Kernel PCA. Χρησιμοποιώντας μία μη γραμμική συνάρτηση  $k$  αντί του τυπικού εσωτερικού γινομένου είναι δυνατή η εφαρμογή του PCA σε ένα χώρο  $F$  μεγάλης διάστασης που δεν είναι γραμμικά συσχετιζόμενος με τον χώρο εισόδου. Οι διακεκομμένες γραμμές είναι οι ισοϋψείς καμπύλες των σταθερών τιμών των χαρακτηριστικών.

Αρχικά, κατά την εφαρμογή του αλγορίθμου kernel PCA, υπολογίζεται ο  $m \times m$  πίνακας  $K$  (33), έπειτα επιλύεται η (35)  $m\lambda a = Ka$  μετατρέποντας τον πίνακα  $K$  σε διαγώνιο (πίνακας με μη μηδενικά στοιχεία μόνο στη διαγώνιο), στη συνέχεια κανονικοποιούνται οι συντελεστές  $a_i$  και τέλος εξάγονται οι Κύριες Συνιστώσες (principal components) υπολογίζοντας τις προβολές στα ιδιοδιανύσματα με βάση την (36) [78].

Με άλλα λόγια, είναι δυνατό να εξαχθούν οι πρώτοι  $q$  ( $1 \leq q \leq m$ ) μη γραμμικές Κύριες Συνιστώσες (ιδιοδιανύσματα  $w^{(q)}$ ) χρησιμοποιώντας τη συνάρτηση πυρήνα χωρίς να είναι απαραίτητη η εκτέλεση της ιδιαίτερα δαπανηρής πράξης της προβολής των δειγμάτων στον χώρο υψηλής διάστασης  $\mathbb{R}^F$ . Οι πρώτες  $q$  Συνιστώσες αντιστοιχούν στις πρώτες  $q$  ιδιοτιμές της  $m\lambda a = Ka$ . Στην αναγνώριση

προσώπων όπου κάθε  $x$  αναπαριστά μία εικόνα οι εξαγόμενοι μη γραμμικές Κύριες Συνιστώσες καλούνται ιδιοπρόσωπα του πυρήνα (kernel eigenfaces).

#### 4.1.10 Γραμμική Διακρίνουσα Ανάλυση του πυρήνα (Kernel Fisher LDA)

Όπως και στην περίπτωση του kernel PCA, γίνεται η υπόθεση ότι τα δείγματα  $\Phi(x)$  που προβάλλονται είναι κεντρικά (centered) στον χώρο  $\mathbb{R}^f$ . Οι εξισώσεις διατυπώνονται με τέτοιο τρόπο ώστε να χρησιμοποιούνται αποκλειστικά εσωτερικά γινόμενα κατά τη γραμμική διακρίνουσα ανάλυση. Έστω ότι  $\tilde{S}_w$  και  $\tilde{S}_b$  είναι ο εντός στις κλάσεις πίνακας διασποράς και ο ανάμεσα στις κλάσεις πίνακας διασποράς αντίστοιχα. Τότε ζητούμενο είναι να βρεθούν οι ιδιοτιμές του  $S_b$  που δίδεται από τη σχέση:

$$\tilde{S}_b = \sum_{i=1}^C \left( \sqrt{\frac{c_i}{N}} (\bar{\varphi}_i - \bar{\varphi}) \right) \left( \sqrt{\frac{c_i}{N}} (\bar{\varphi}_i - \bar{\varphi}) \right)^T = \sum_{i=1}^C \tilde{\varphi}_i \tilde{\varphi}_i^T = \Phi_b \Phi_b^T \quad (38)$$

όπου,  $\tilde{\varphi}_i = \sqrt{\frac{c_i}{N}} (\bar{\varphi}_i - \bar{\varphi})$  και  $\Phi_b = [\tilde{\varphi}_1, \dots, \tilde{\varphi}_c]$ . Εφόσον η διάσταση του χώρου των χαρακτηριστικών  $J'$  μπορεί να είναι ιδιαίτερα μεγάλη και πιθανώς άπειρη, είναι αδύνατο να υπολογιστούν απευθείας τα ιδιοδιανύσματα του  $J' \times J'$  πίνακα  $\tilde{S}_b$ . Όμως είναι εφικτό να βρεθούν με έμμεσο τρόπο, τα πρώτα  $m$  ( $\leq C - 1$ ) σημαντικότερα ιδιοδιανύσματα του  $\tilde{S}_b$  που αντιστοιχούν σε μη μηδενικές ιδιοτιμές από τα ιδιοδιανύσματα του πίνακα  $\Phi_b^T \Phi_b$  μεγέθους  $(C \times C)$ .

Στη συνέχεια θεωρούμε ότι υπάρχει μία συνάρτηση πυρήνα  $k(z_i, z_j) = \varphi(z_i) \cdot \varphi(z_j)$  για κάθε  $\varphi(z_i), \varphi(z_j) \in F$ . Έπειτα ορίζεται ο  $N \times N$  πίνακας  $K$  εσωτερικών γινομένων ως εξής:

$$K = (k_{ih})_{h=1, \dots, C}^{i=1, \dots, C} \quad \text{όπου } k_{ih} = (k_{ij})_{h=1, \dots, C}^{i=1, \dots, C} \quad (39)$$

όπου  $k_{ij} = k(z_{li}, z_{hj}) = \varphi_{li} \cdot \varphi_{hj}$ ,  $\varphi_{li} = \varphi(z_{li})$  και  $\varphi_{hj} = \varphi(z_{hj})$ . Ο πίνακας  $K$  επιτρέπει τη μετατροπή του  $\Phi_b^T \Phi_b$  ως εξής [38]:

$$\Phi_b^T \Phi_b = \frac{1}{N} B \cdot (A_{NC}^T \cdot K \cdot A_{NC} - \frac{1}{N} (A_{NC}^T \cdot K \cdot \mathbf{1}_{NC}) - \frac{1}{N} (\mathbf{1}_{NC}^T \cdot K \cdot A_{NC}) + \frac{1}{N^2} (\mathbf{1}_{NC}^T \cdot K \cdot \mathbf{1}_{NC})) \cdot B \quad (40)$$

Όπου  $B = \text{diag}[\sqrt{C_1}, \dots, C_c]$ ,  $\mathbf{1}_{NC}$  είναι ένας  $N \times C$  πίνακας του οποίου τα στοιχεία είναι όλα ίσα με τη μονάδα,  $A_{NC} = \text{diag}[a_{c_1}, \dots, a_{c_c}]$  είναι ένας  $N \times C$  διαγώνιος (block diagonal) πίνακας και  $a_{c_i}$  είναι ένα διάνυσμα  $C_i \times 1$  που όλοι του οι όροι είναι ίσοι με  $\frac{1}{c_i}$ .

Έστω ότι  $\lambda_i$  και  $\tilde{e}_i$  ( $i = 1, \dots, C$ ) είναι η  $i$ -οστή ιδιοτιμή και το αντίστοιχο ιδιοδιάνυσμα του

$\Phi_b^T \Phi_b$ . Τα ιδιοδιανύσματα διατάσσονται χρησιμοποιώντας τις αντίστοιχες ιδιοτιμές που έχουν ταξινομηθεί με φθίνουσα σειρά. Εφόσον ισχύει  $(\Phi_b \Phi_b^T)(\Phi_b \tilde{e}_i) = \lambda_i (\Phi_b \tilde{e}_i)$ , το  $\tilde{v}_i = \Phi_b \tilde{e}_i$  είναι ένα

ιδιοδιάνυσμα του  $\tilde{S}_b$ . Προκειμένου να αφαιρεθεί ο μηδενωχώρας του  $\tilde{S}_b$ , χρησιμοποιούνται μόνο τα πρώτα  $m$  ( $\leq C - 1$ ) ιδιοδιανύσματα. Συγκεκριμένα,  $\tilde{V} = [\tilde{v}_1, \dots, \tilde{v}_m = \Phi_b E_m]$ , όπου  $E_m = [\tilde{e}_1, \dots, \tilde{e}_m]$ , είναι τα ιδιοδιανύσματα οι ιδιοτιμές των οποίων είναι μεγαλύτερες του μηδενός. Μπορεί εύκολα να παρατηρηθεί ότι  $V^T \tilde{S}_b \tilde{V} = \Lambda_b$  όπου  $\Lambda_b = \text{diag}[\lambda_1^2, \dots, \lambda_m^2]$  ένας ( $m \times m$ ) διαγώνιος πίνακας.

Έστω  $\tilde{U} = \tilde{V} \Lambda_b^{-1/2}$ , κάθε στήλη του οποίου είναι ένα διάνυσμα του χώρου των χαρακτηριστικών  $F$ . Προβάλλοντας τόσο τον ανάμεσα στις κλάσεις πίνακα διασποράς,  $\tilde{S}_b$ , όσο και τον εντός κλάσης πίνακα διασποράς,  $\tilde{S}_w$ , στον υποχώρο που οριοθετείται από τον  $\tilde{U}$ , μπορεί εύκολα να παρατηρηθεί ότι

$\tilde{U}^T \tilde{S}_b \tilde{U} = I$ , όπου  $I$  ένας μοναδιαίος πίνακας  $m \times m$  διαστάσεων. Επιπλέον  $\tilde{U}^T \tilde{S}_w \tilde{U}$  μπορεί να αναλυθεί ως εξής:

$$\tilde{U}^T \tilde{S}_w \tilde{U} = (E_m \Lambda_b^{-1/2})^T (\Phi_b^T \tilde{S}_w \Phi_b) (E_m \Lambda_b^{-1/2}) \quad (41)$$

Χρησιμοποιώντας τον πίνακα του πυρήνα  $K$ , το γινόμενο  $\Phi_b^T \tilde{S}_w \Phi_b$  μπορεί να μετασχηματιστεί ως εξής:

$$\Phi_b^T \tilde{S}_w \Phi_b = \frac{1}{N^2} B \cdot (A_{NC}^T \cdot K \cdot A_{NC} - \frac{1}{N} (A_{NC}^T \cdot K \cdot \mathbf{1}_{NC})) -$$

$$\frac{1}{N} (\mathbf{1}_{NC}^T \cdot K \cdot A_{NC}) + \frac{1}{N^2} (\mathbf{1}_{NC}^T \cdot K \cdot \mathbf{1}_{NC}) \cdot B \quad (42)$$

όπου  $\tilde{K} = K(1 - W) \cdot K$  με  $W = \text{diag}[w_1, \dots, w_c]$  να είναι ένας  $N \times N$  διαγώνιος (block diagonal) πίνακας και  $W_i$  ένας  $C_i \times C_i$  πίνακας όλα τα στοιχεία του οποίου είναι ίσα με  $\frac{1}{c_i}$ .

Μετατρέποντας σε διαγώνιο πίνακα το γινόμενο  $\tilde{U}^T \tilde{S}_w \tilde{U}$ , προκύπτει ένας πίνακας διαστάσεων  $m \times m$ . Έστω  $\tilde{p}_i$  το  $i$ -οστό ιδιοδιάνυσμα του  $\tilde{U}^T \tilde{S}_w \tilde{U}$ , όπου  $i = 1, \dots, m$ . Τα ιδιοδιανύσματα αυτά έχουν διαταχθεί σύμφωνα με τις αντίστοιχες ιδιοτιμές τους,  $\lambda_i$ , οι οποίες έχουν ταξινομηθεί κατά αύξουσα σειρά. Στο σύνολο των ταξινομημένων ιδιοδιανυσμάτων αυτά που αντιστοιχούν στις μικρότερες ιδιοτιμές είναι και αυτά που ελαχιστοποιούν τον παρονομαστή του κανονικοποιημένου κριτηρίου του Fisher που λαμβάνει την παρακάτω μορφή στην περίπτωση της κανονικοποιημένης γραμμικής διακρίνουσας ανάλυσης του πυρήνα (εξίσωση 43). Τα ιδιοδιανύσματα αυτά μπορούν να θεωρηθούν ως τα χαρακτηριστικά που μεγιστοποιούν τη διαχωριστικότητα.

$$\tilde{W} = \operatorname{argmax}_{\tilde{W}} \frac{|\tilde{W}^T \tilde{S}_b \tilde{W}|}{|\eta(\tilde{W}^T \tilde{S}_b \tilde{W}) + (\tilde{W}^T \tilde{S}_w \tilde{W})|} \quad (43)$$

όπου  $0 \leq \eta \leq 1$  είναι η παράμετρος κανονικοποίησης [49].

Έστω  $\tilde{P}_M = [\tilde{p}_1, \dots, \tilde{p}_M]$  και  $\Lambda_w = \text{diag}[\lambda_1, \dots, \lambda_M]$  είναι τα πρώτα  $M$  ( $\leq m$ ) ιδιοδιανύσματα και οι αντίστοιχες ιδιοτιμές τους. Τότε η ζητούμενη λύση μπορεί να επιτευχθεί μέσω  $\Gamma = \tilde{U} \tilde{P}_M (\eta I + \Lambda_w)^{-1/2}$ , που αποτελεί το σύνολο των βέλτιστων μη γραμμικών χαρακτηριστικών βάσης.

Για κάθε μοτίβο εισόδου  $z$ , η προβολή στον υποχώρο που ορίζεται από το σύνολο των βασικών χαρακτηριστικών  $\Gamma$ , μπορεί να υπολογιστεί από:

$$\mathbf{y} = \tilde{\Gamma}^T \boldsymbol{\varphi}(\mathbf{z}) = \left( \tilde{\mathbf{E}}_m \cdot \tilde{\Lambda}_b^{-1/2} \cdot \tilde{\mathbf{P}}_M \cdot (\eta \mathbf{I} + \tilde{\Lambda}_\omega)^{-1/2} \right)^T (\tilde{\Phi}_b^T \boldsymbol{\varphi}(\mathbf{z})) \quad (44)$$

$$\text{Όπου } \tilde{\Phi}_b^T \boldsymbol{\varphi}(\mathbf{z}) = [\tilde{\phi}_1 \dots \tilde{\phi}_c] \boldsymbol{\varphi}(\mathbf{z})$$

Εισάγεται ένα  $(N \times 1)$  διάνυσμα πυρήνα,

$$\mathbf{v}(\boldsymbol{\varphi}(\mathbf{z})) = [\boldsymbol{\varphi}_{11}^T \boldsymbol{\varphi}(\mathbf{z}) \boldsymbol{\varphi}_{12}^T \boldsymbol{\varphi}(\mathbf{z}) \dots \boldsymbol{\varphi}_{c(c-1)}^T \boldsymbol{\varphi}(\mathbf{z}) \boldsymbol{\varphi}_{cc}^T \boldsymbol{\varphi}(\mathbf{z})]^T \quad (45)$$

που ορίζεται από το εσωτερικό γινόμενο του  $\boldsymbol{\varphi}(\mathbf{z})$  με κάθε χαρτογραφημένο δείγμα εκπαίδευσης  $\boldsymbol{\varphi}(z_{ij})$  στο χώρο των χαρακτηριστικών  $F$ . Αναδιαμορφώνοντας την εξίσωση 44 με τη χρήση του διανύσματος πυρήνα (45) παράγεται η σχέση:

$$\mathbf{y} = \boldsymbol{\theta} \cdot \mathbf{v}(\boldsymbol{\varphi}(\mathbf{z})) \quad (46)$$

Όπου

$$\boldsymbol{\theta} = \frac{1}{\sqrt{N}} (\tilde{\mathbf{E}}_m \cdot \tilde{\Lambda}_b^{-\frac{1}{2}} \cdot \tilde{\mathbf{P}}_M \cdot (\eta \mathbf{I} + \tilde{\Lambda}_\omega)^{-1/2})^T \cdot \mathbf{B} \cdot (\mathbf{A}_{NC}^T - \frac{1}{N} \mathbf{1}_{NC}^T) \quad (47)$$

είναι ένας  $(M \times N)$  πίνακας. Έτσι από την εξίσωση εισάγεται μία μη γραμμική αναπαράσταση ( $\mathbf{y}$ ) μικρής διάστασης των  $\mathbf{z}$  με βελτιωμένη διαχωριστική δυνατότητα.

## 4.2 Υλοποίηση Αλγορίθμων

### 4.2.1 Εισαγωγή

Σε αυτή την ενότητα θα εξεταστούν τέσσερεις αλγόριθμοι με τη βάση ORL Face Database . 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. Αυτή η συμπεριφορά του συστήματος κωδικοποιείται στη χαρακτηριστική λειτουργίας δέκτη (Receiver Operating Characteristic-ROC) που δείχνει την τιμή του ποσοστού εσφαλμένης αποδοχής (False Accept Rate-FAR) και του ποσοστού εσφαλμένης απόρριψης (False Reject Rate-FRR) για κάθε τιμή του κατωφλίου.

### 4.2.2 Βάση ORL Face

Η βάση που χρησιμοποιήθηκε για τη σύγκριση των αλγορίθμων είναι Βάση ORL Face. Από τη βάση προσώπων, επιλέχθηκαν μέρος από τις εικόνες της κάθε κλάσης για τη δημιουργία της βάσης εκπαίδευσης των αλγορίθμων ενώ όλες οι υπόλοιπες εικόνες σχηματίζουν τη βάση ελέγχου.

Αποτελείται από 40 διαφορετικές κλάσεις ,της μορφής pgm ,με 10 φωτογραφίες σε κάθε μία κλάση. Για μερικές κλάσεις , οι εικόνες που έχουν ληφθεί είναι σε διαφορετικές χρονικές στιγμές, μεταβάλλοντας το φωτισμό, τις εκφράσεις του προσώπου (ανοικτά / κλειστά μάτια, χαμογελά / δεν χαμογελά) και με λεπτομέρειες προσώπου (με γυαλιά /χωρίς γυαλιά). Όλες οι εικόνες έχουν ληφθεί σε σκούρο φόντο με ομοιογενή θέματα, σε όρθια, μετωπική θέση όπως φαίνεται στην Εικόνα 23.



Εικόνα 23: Παράδειγμα κλάσης της ORL

### 4.2.3 Υλοποίηση και εφαρμογή αλγορίθμων

Η υλοποίηση των αλγορίθμων και η διαδικασία εφαρμογής και ελέγχου πραγματοποιήθηκε με τη χρήση του λογισμικού MATLAB. Αναλυτικά οι αλγόριθμοι που υλοποιήθηκαν παρατίθενται στον Πίνακα 8:



Όνομασία Αλγορίθμων	Περιγραφή
1. PCA	Ανάλυση σε Κύριες Συνιστώσες (Principal Component Analysis) - Μέθοδος Ιδιοπροσώπων
2. LDA	Γραμμική διακρίνουσα ανάλυση (Linear Discriminant Analysis, LDA)
3.KPCA	Kernel principal component analysis (kernel PCA)
4.KFA	Γραμμική Διακρίνουσα Ανάλυση του πυρήνα (Kernel Fisher LDA)

**Πίνακας 8:**Πίνακας αλγορίθμων

Η εφαρμογή των αλγορίθμων πραγματοποιήθηκε με σκοπό τη μεταξύ τους σύγκριση ως προς την ικανότητα ταυτοποίησης των προσώπων ελέγχου με μία από τις κλάσεις που συμμετέχουν στην βάση εκπαίδευσης.

Επιπλέον στόχος της υλοποίησης των αλγορίθμων αποτελεί η αξιολόγηση των διαφόρων συναρτήσεων ομοιότητας και αποστάσεων ως προς την απόδοση τους για κάθε αλγόριθμο. Οι αποστάσεις που χρησιμοποιήθηκαν, όπως περιγράφηκαν αναλυτικά στο κεφάλαιο 4, είναι η Ευκλείδεια απόσταση, η Cityblock απόσταση, η Ομοιότητα συνημίτονου (Cosine), η απόσταση Mahalanobis.

Για τη βάση που χρησιμοποιήσαμε επιλέξαμε από κάθε κλάση οι πρώτες 5 φωτογραφίες να χρησιμοποιηθούν για να εκπαιδύσουμε τους αλγορίθμους μας και τις υπόλοιπες τις χρησιμοποιήσαμε για να τους δοκιμάσουμε. Οι αλγόριθμοι που χρησιμοποιήσαμε βρίσκονται στο φάκελο FaceRecognition.rar. Περιλαμβάνει το PhD tool που είναι μια συλλογή του Matlab με λειτουργίες και τα σενάρια που προορίζεται να βοηθήσει τους ερευνητές που εργάζονται στον τομέα της αναγνώρισης προσώπου. Η εργαλειοθήκη έχει παραχθεί ως υποπροϊόν της ερευνητικής εργασίας.[79]

Η διαδικασία των μετρήσεων βασίζεται σε τρία (3) βήματα:

1. Δημιουργία βάσης εκπαίδευσης και ελέγχου .
2. Εφαρμογή αλγορίθμων στη βάση εκπαίδευσης και υπολογισμός των διανυσμάτων των αποστάσεων των εικόνων ελέγχου από τις εικόνες της βάσης εκπαίδευσης.
3. Ανάλυση μετρήσεων, υπολογισμός ποσοστών επιτυχίας αναγνώρισης. Επιπλέον έγινε σύγκριση της απόδοσης των συναρτήσεων ομοιότητας σε κάθε αλγόριθμο.

Τα στατιστικά στοιχεία που εξήχθησαν είναι τα ποσοστά αναγνώρισης κάθε συνάρτησης ομοιότητας για κάθε αλγόριθμο αναγνώρισης, τα διανύσματα των τιμών των συναρτήσεων ομοιότητας των εικόνων ελέγχου με τις εικόνες της βάσης εκπαίδευσης καθώς και τα ποσοστά των FAR αναγνώρισεων.

Ως FAR (False Accept Rate= ποσοστό ψευδούς ταυτοποίησης) ορίζεται το ποσοστό το σύστημα μας να ταιριάζει μια εικόνα εισόδου σε λάθος κλάση, και το FRR (False Rejection Rate=

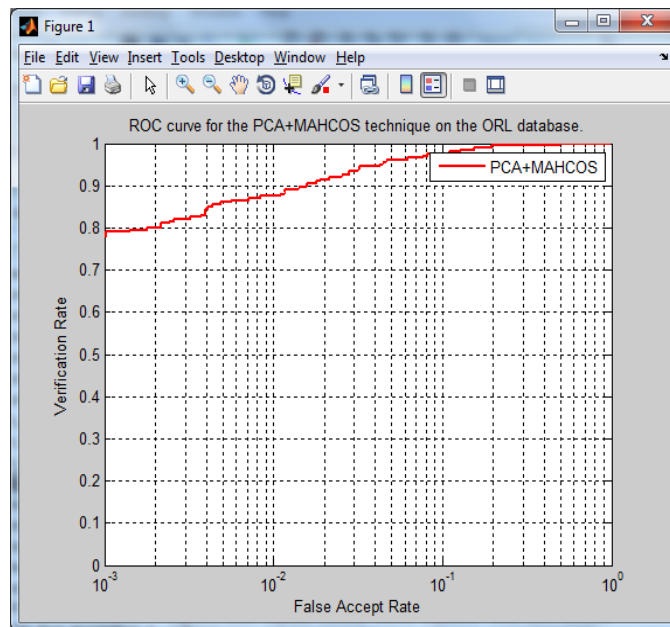
ποσοστό ψευδούς απόρριψης / μη ταυτοποίησης) ορίζεται το ποσοστό το σύστημα μας να μη ταιριάζει την εικόνα εισόδου σε μια κλάση που προϋπάρχει στη βάση δεδομένων.

Για συστήματα ασφαλείας τα οποία χρησιμοποιούν αναγνώριση προσώπου, είναι σημαντικό να ελαχιστοποιηθεί το ποσοστό κατά το οποίο ένα πρόσωπο μίας κλάσης μπορεί να ταυτοποιηθεί με κάποιο άλλο.

Τα αποτελέσματα που ακολουθούν είναι τα ποσοστά αναγνώρισης βάση των αποστάσεων Ευκλείδειας, της Cityblock απόστασης, της Ομοιότητας συνημίτονου και της απόσταση Mahalanobis.

#### 4.2.3.1 Ποσοστά αναγνώρισης του PCA στην βάση ORL για διάφορες συναρτήσεις ομοιότητας

*Αποτελέσματα απόστασης Mahalanobis*



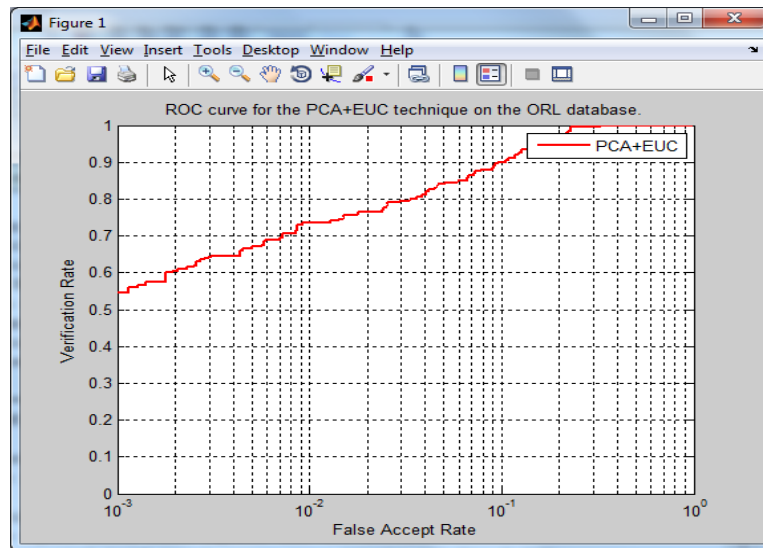
**Εικόνα 24:** Απόδοση ταυτοποίησης PCA με χρήση Mahalanobis

The verification rate at 1% FAR equals (in %): 87.50%

The verification rate at 0.1% FAR equals (in %): 78.00%

The verification rate at 0.01% FAR equals (in %): 64.50%

Αποτελέσματα Ευκλείδειας απόστασης



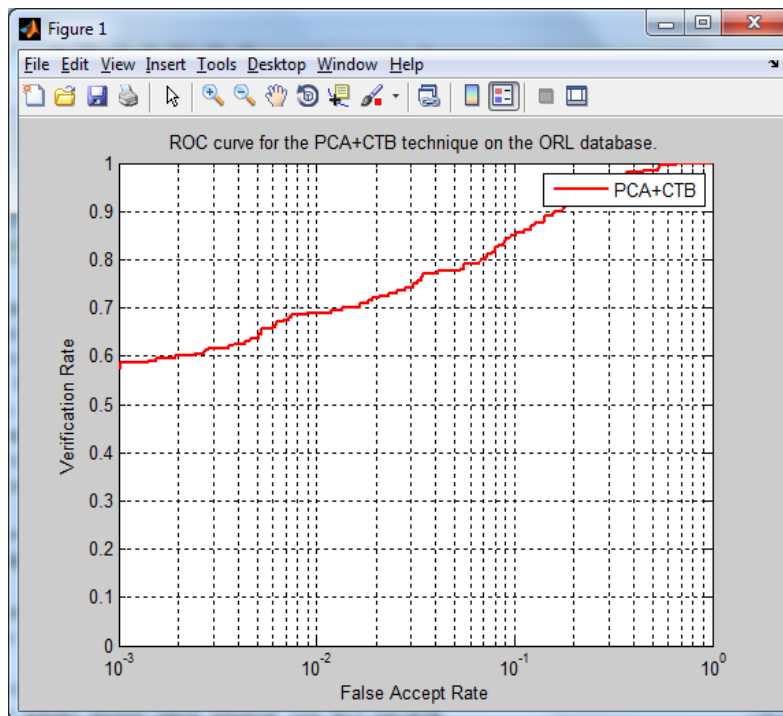
Εικόνα 25: Απόδοση ταυτοποίησης PCA με χρήση Ευκλείδειας

The verification rate at 1% FAR equals (in %): 73.50%

The verification rate at 0.1% FAR equals (in %): 54.50%

The verification rate at 0.01% FAR equals (in %): 37.50%

Αποτελέσματα απόστασης Cityblock



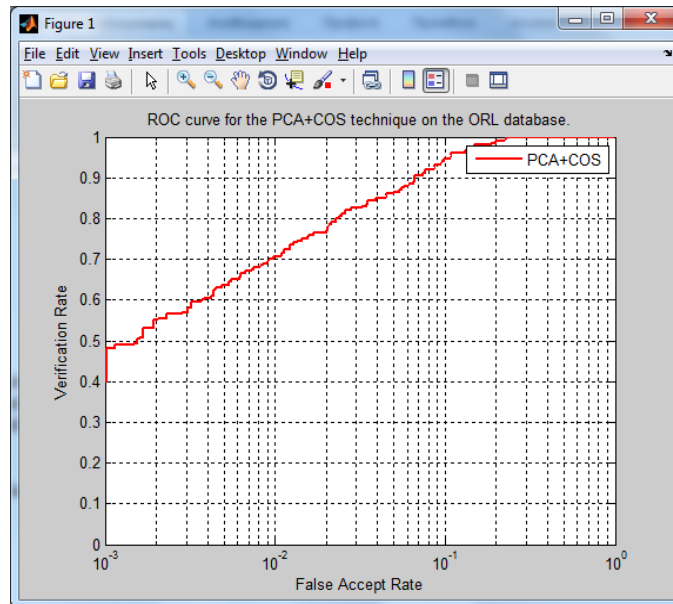
Εικόνα 26: Απόδοση ταυτοποίησης PCA με χρήση Cityblock

The verification rate at 1% FAR equals (in %): 69.00%

The verification rate at 0.1% FAR equals (in %): 57.50%

The verification rate at 0.01% FAR equals (in %): 39.00%

*Αποτελέσματα απόστασης Cosine*



**Εικόνα 27:** Απόδοση ταυτοποίησης PCA με χρήση Cosine

The verification rate at 1% FAR equals (in %): 70.50%

The verification rate at 0.1% FAR equals (in %): 40.00%

The verification rate at 0.01% FAR equals (in %): 17.00%

Στον Πίνακα 9 που ακολουθεί καταγράφονται τα ποσοστά επιτυχίας του αλγόριθμου PCA για κάθε συνάρτηση ομοιότητας.

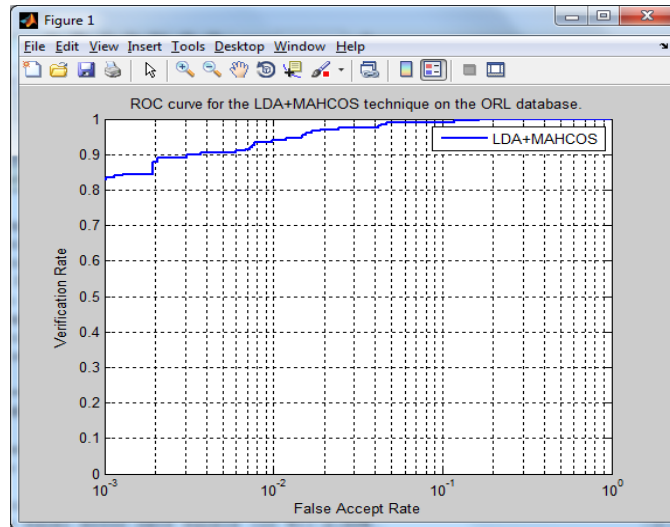
	<b>0.01% FAR</b>	<b>0.1% FAR</b>	<b>1% FAR</b>
<b>Mahalanobis</b>	64.50 %	78.00 %	87.50 %
<b>Cityblock</b>	69.00 %	57.50 %	69.00 %
<b>Cosine</b>	17.00 %	40.00 %	70.50 %
<b>Euclidean</b>	37.50 %	54.50 %	73.50 %

**Πίνακας 9:** Ποσοστά αναγνώρισης PCA

Όπως φαίνεται στον παραπάνω πίνακα ο αλγόριθμος PCA παρουσιάζει μεγαλύτερα ποσοστά αναγνώρισης με την χρήση της απόστασης Cityblock.

#### 4.2.3.2 Ποσοστά αναγνώρισης του LDA στην βάση ORL για διάφορες συναρτήσεις ομοιότητας

Αποτελέσματα απόστασης Mahalanobis



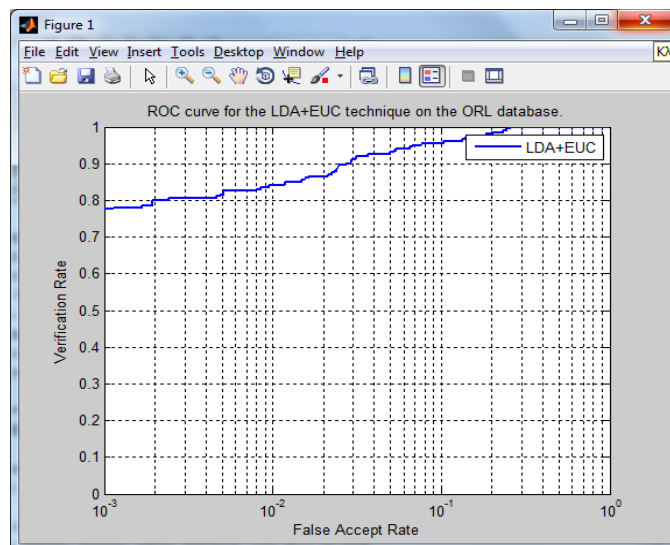
Εικόνα 28: Απόδοση ταυτοποίησης LDA με χρήση Mahalanobis

The verification rate at 1% FAR equals (in %): 94.00%

The verification rate at 0.1% FAR equals (in %): 83.00%

The verification rate at 0.01% FAR equals (in %): 78.50%

Αποτελέσματα της Ευκλείδειας απόστασης



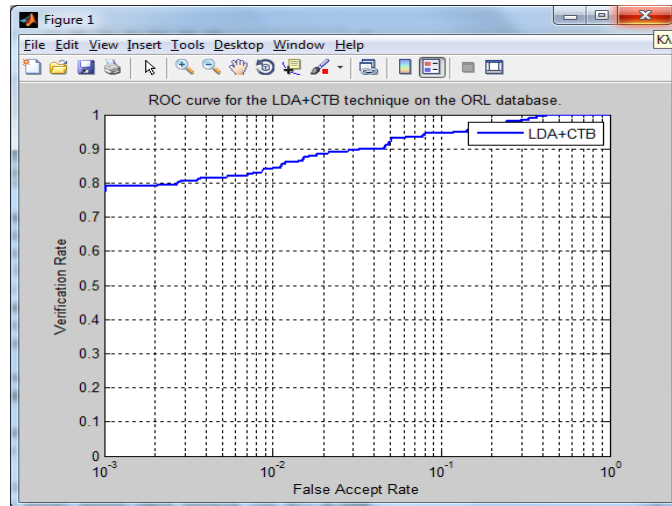
Εικόνα 29: Απόδοση ταυτοποίησης LDA με χρήση Ευκλείδειας

The verification rate at 1% FAR equals (in %): 84.00%

The verification rate at 0.1% FAR equals (in %): 77.50%

The verification rate at 0.01% FAR equals (in %): 70.00%

*Αποτελέσματα της απόστασης Cityblock*



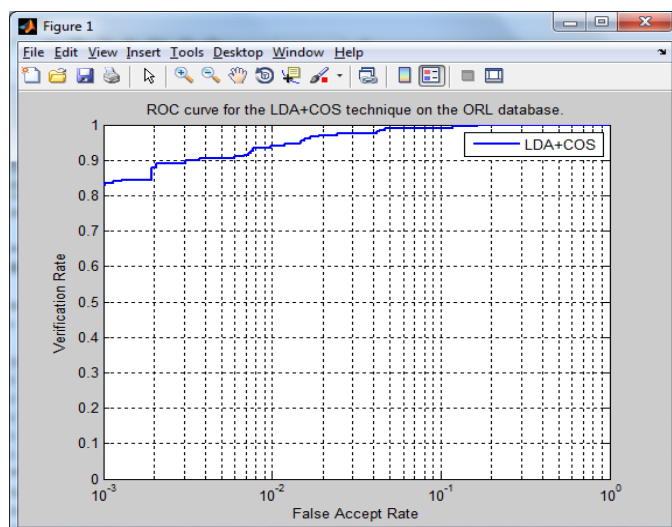
**Εικόνα 30:** Απόδοση ταυτοποίησης LDA με χρήση Cityblock

The verification rate at 1% FAR equals (in %): 84.50%

The verification rate at 0.1% FAR equals (in %): 77.50%

The verification rate at 0.01% FAR equals (in %): 63.00%

*Αποτελέσματα της απόστασης Cosine*



**Εικόνα 31:** Απόδοση ταυτοποίησης LDA με χρήση Cosine

The verification rate at 1% FAR equals (in %): 94.00%

The verification rate at 0.1% FAR equals (in %): 83.00%

The verification rate at 0.01% FAR equals (in %): 78.50%

Στον Πίνακα 10 που ακολουθεί καταγράφονται τα ποσοστά επιτυχίας του αλγόριθμου LDA για κάθε συνάρτηση ομοιότητας.

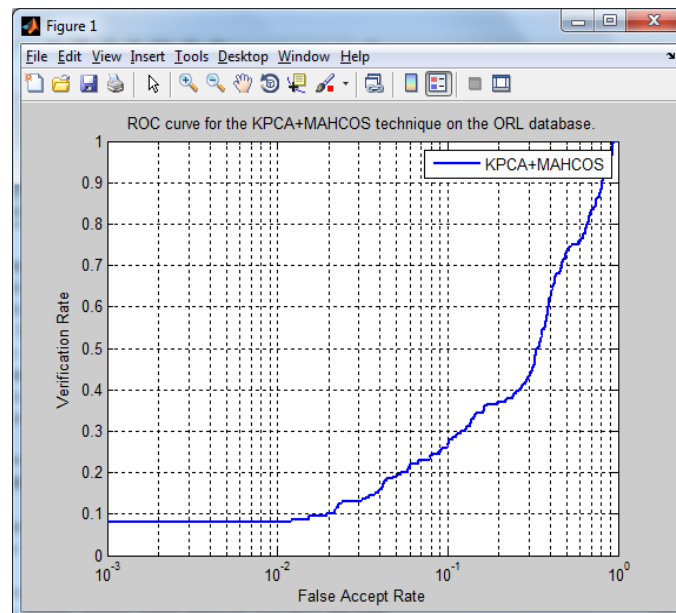
	<b>0.01% FAR</b>	<b>0.1% FAR</b>	<b>1% FAR</b>
<b>Mahalanobis</b>	78.50 %	83.00 %	94.00 %
<b>Cityblock</b>	63.00 %	77.50 %	84.50 %
<b>Cosine</b>	78.50 %	83.00 %	94.00 %
<b>Euclidean</b>	70.00 %	77.50 %	84.00 %

**Πίνακας 10:** Ποσοστά αναγνώρισης LDA

Όπως φαίνεται στον παραπάνω πίνακα ο αλγόριθμος LDA παρουσιάζει μεγαλύτερα ποσοστά αναγνώρισης με την χρήση των αποστάσεων Mahalanobis και Cosine.

#### 4.2.3.3 Ποσοστά αναγνώρισης του Kernel PCA στην βάση ORL για διάφορες συναρτήσεις ομοιότητας

Αποτελέσματα απόστασης Mahalanobis



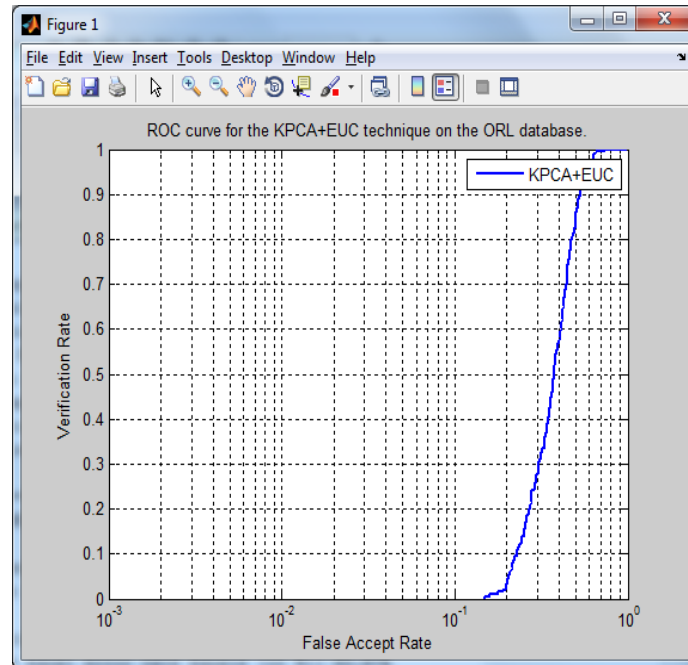
**Εικόνα 32:** Απόδοση ταυτοποίησης KPCA με χρήση Mahalanobis

The verification rate at 1% FAR equals (in %): 8.00%

The verification rate at 0.1% FAR equals (in %): 8.00%

The verification rate at 0.01% FAR equals (in %): 5.50%

*Αποτελέσματα της Ευκλείδειας απόστασης*



**Εικόνα 33:** Απόδοση ταυτοποίησης KPCA με χρήση Ευκλείδειας

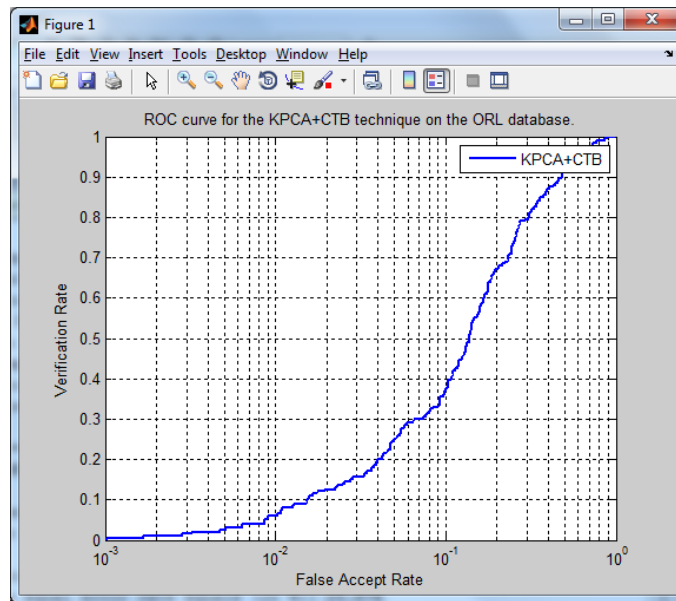
The verification rate at 1% FAR equals (in %): 0.50%

The verification rate at 0.1% FAR equals (in %): 0.50%

The verification rate at 0.01% FAR equals (in %): 0.50%



Αποτελέσματα απόστασης Cityblock



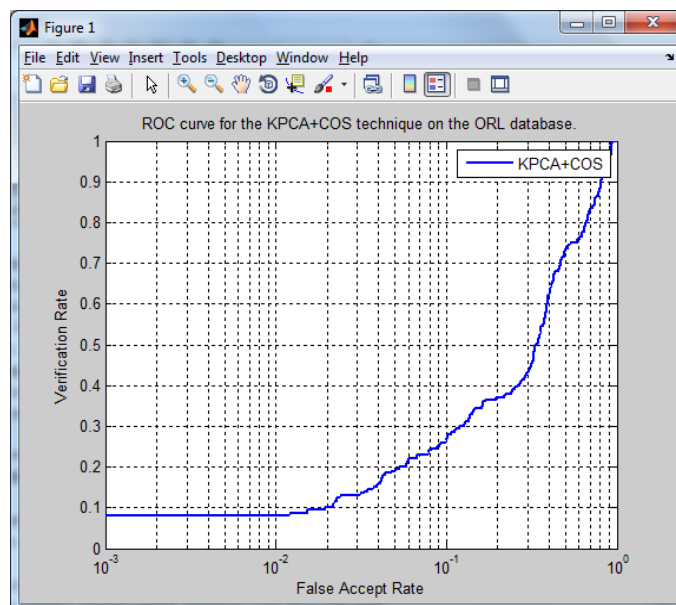
Εικόνα 34: Απόδοση ταυτοποίησης KPCA με χρήση Cityblock

The verification rate at 1% FAR equals (in %): 6.00%

The verification rate at 0.1% FAR equals (in %): 0.50%

The verification rate at 0.01% FAR equals (in %): 0.50%

Αποτελέσματα της απόστασης συνημίτονου



Εικόνα 35: Απόδοση ταυτοποίησης KPCA με χρήση Cosine

The verification rate at 1% FAR equals (in %): 8.00%

The verification rate at 0.1% FAR equals (in %): 8.00%

The verification rate at 0.01% FAR equals (in %): 5.50%

Στον Πίνακα 11 που ακολουθεί καταγράφονται τα ποσοστά επιτυχίας του αλγόριθμου kernel PCA για κάθε συνάρτηση ομοιότητας.

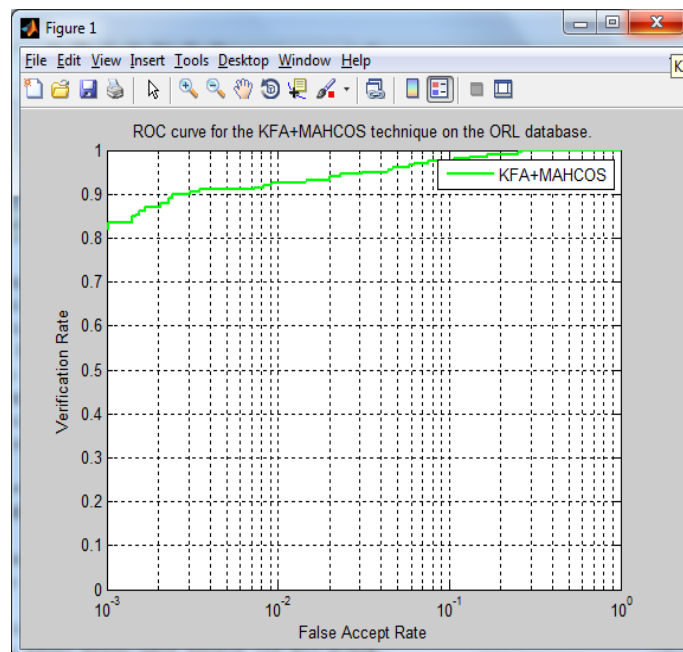
	0.01% FAR	0.1% FAR	1% FAR
<b>Mahalanobis</b>	5.50 %	8.00 %	8.00 %
<b>Cityblock</b>	0.50 %	0.50 %	6.00 %
<b>Cosine</b>	5.50 %	8.00 %	8.00 %
<b>Euclidean</b>	0.50 %	0.50 %	0.50 %

**Πίνακας 11:** Ποσοστά αναγνώρισης kernel PCA

Όπως φαίνεται στον παραπάνω πίνακα ο αλγόριθμος kernel PCA παρουσιάζει μεγαλύτερα ποσοστά αναγνώρισης με την χρήση των αποστάσεων Mahalanobis και Cosine.

#### 4.2.3.1 Ποσοστά αναγνώρισης του KFA στην βάση ORL για διάφορες συναρτήσεις ομοιότητας

*Αποτελέσματα απόστασης Mahalanobis*



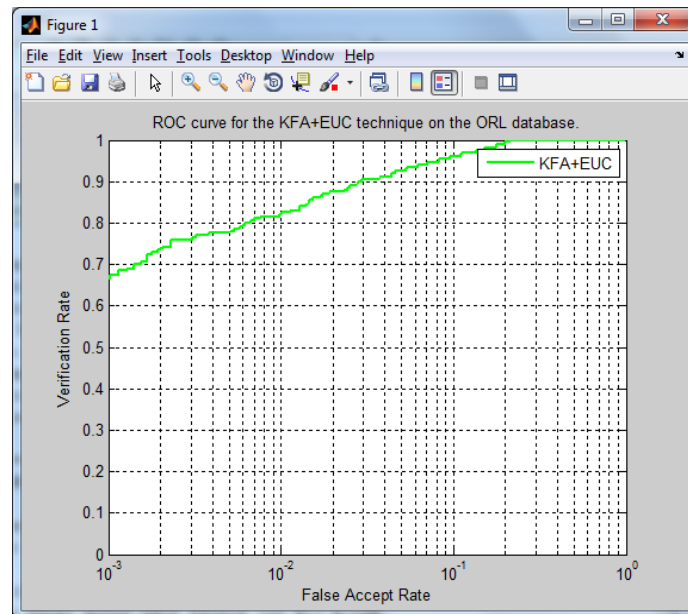
**Εικόνα 36:** Απόδοση ταυτοποίησης KFA με χρήση Mahalanobis

The verification rate at 1% FAR equals (in %): 92.50%

The verification rate at 0.1% FAR equals (in %): 82.00%

The verification rate at 0.01% FAR equals (in %): 68.00%

*Αποτελέσματα Ευκλείδειας απόστασης*



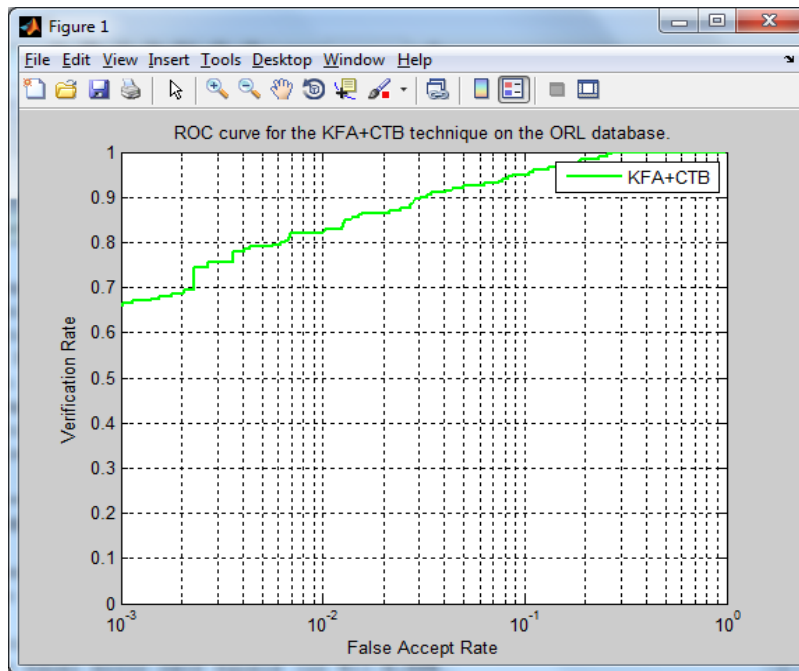
**Εικόνα 37:** Απόδοση ταυτοποίησης KFA με χρήση Ευκλείδειας

The verification rate at 1% FAR equals (in %): 81.50%

The verification rate at 0.1% FAR equals (in %): 66.50%

The verification rate at 0.01% FAR equals (in %): 54.50%

Αποτελέσματα της απόστασης Cityblock



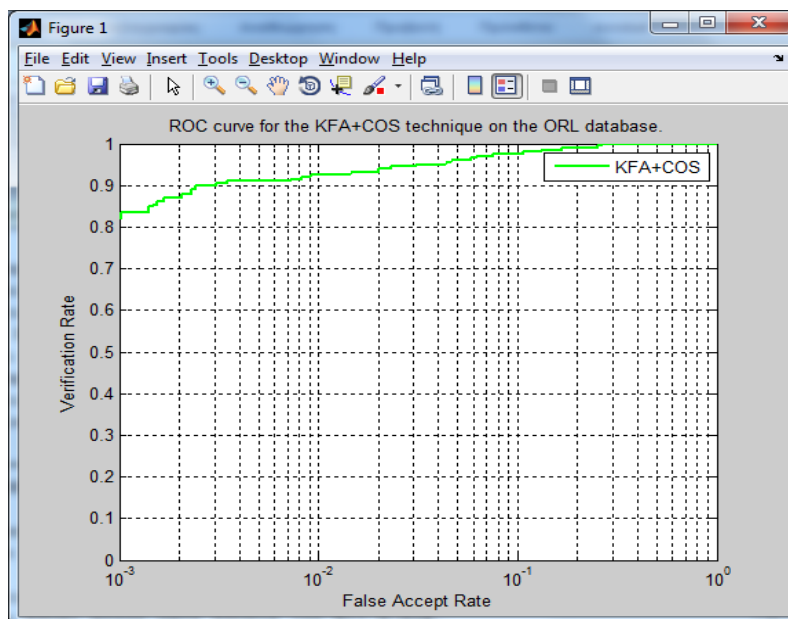
Εικόνα 38: Απόδοση ταυτοποίησης KFA με χρήση Cityblock

The verification rate at 1% FAR equals (in %): 82.00%

The verification rate at 0.1% FAR equals (in %): 66.00%

The verification rate at 0.01% FAR equals (in %): 55.50%

Αποτελέσματα της απόστασης Cosine



Εικόνα 39: Απόδοση ταυτοποίησης KFA με χρήση Cosine

The verification rate at 1% FAR equals (in %): 92.50%

The verification rate at 0.1% FAR equals (in %): 82.00%

The verification rate at 0.01% FAR equals (in %): 68.00%

Στον Πίνακα 12 που ακολουθεί καταγράφονται τα ποσοστά επιτυχίας του αλγόριθμου KFA για κάθε συνάρτηση ομοιότητας.

	<b>0.01% FAR</b>	<b>0.1% FAR</b>	<b>1% FAR</b>
<b>Mahalanobis</b>	68.00 %	82.00 %	92.50 %
<b>Cityblock</b>	55.50 %	66.00 %	82.00 %
<b>Cosine</b>	68.00 %	82.00 %	92.50 %
<b>Euclidean</b>	54.50 %	66.50 %	81.50 %

**Πίνακας 12:** Ποσοστά αναγνώρισης KFA

Όπως φαίνεται στον παραπάνω πίνακα ο αλγόριθμος KFA παρουσιάζει μεγαλύτερα ποσοστά αναγνώρισης με την χρήση των αποστάσεων Mahalanobis και Cosine.

#### 4.2.4 Συγκεντρωτικά Αποτελέσματα

Στον Πίνακα 13 που ακολουθεί καταγράφονται τα ποσοστά επιτυχίας των αλγορίθμων με το μικρότερο ποσοστό ψευδούς ταυτοποίησης (0.01%) για κάθε αλγόριθμο αναγνώρισης προσώπου και κάθε απόσταση ομοιότητας. Αποτυπώνεται η ικανότητα των αλγορίθμων για ταυτοποίηση των εικόνων ελέγχου με μία από τις κλάσεις που σχηματίζουν την βάση εκπαίδευσης. Με έντονη γραφή σημειώνεται η απόσταση ομοιότητας με το βέλτιστο ποσοστό αναγνώρισης (ταυτοποίησης) για κάθε αλγόριθμο.

	Euclidean	Cosine	Cityblock	Mahalanobis
PCA	37.50	17.00	39.00	<b>64.50</b>
LDA	70.00	<b>78.50</b>	63.00	<b>78.50</b>
KPCA	0.50	<b>5.50</b>	0.50	<b>5.50</b>
KLDA (KFA)	54.50	<b>68.00</b>	55.50	<b>68.00</b>

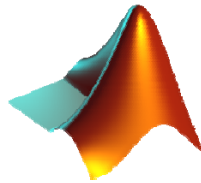
**Πίνακας 13:** Συγκεντρωτικός πίνακας αποστάσεων

Όπως φαίνεται στον παραπάνω πίνακα οι αλγόριθμοι LDA και KFA παρουσιάζουν μεγαλύτερα ποσοστά αναγνώρισης με την χρήση των Cosine και Mahalanobis, ο αλγόριθμος PCA ευνοείται από τη χρήση της Mahalanobis, ενώ ο αλγόριθμος KPCA παρουσιάζει μεγαλύτερα ποσοστά αναγνώρισης με τη χρήση των Cosine και Mahalanobis αν και τα αποτελέσματά τους δεν είναι καθόλου ικανοποιητικά.

Συγκρίνοντας τα βέλτιστα αποτελέσματα για κάθε αλγόριθμο ανεξάρτητα από την απόσταση ομοιότητας, το μεγαλύτερο ποσοστό αναγνώρισης το εμφάνισε ο αλγόριθμος LDA με ποσοστό 78.50% με χρήση της απόσταση Cosine και με χρήση της απόσταση Mahalanobis και μικρότερο ποσοστό ο KPCA με ποσοστό 5.50% με την χρήση της Cosine και Mahalanobis.

## 5. Πειραματικό Μέρος

### 5.1 Η ιστορία του Matlab



Το **MATLAB** δημιουργήθηκε προς το τέλος της δεκαετίας του '70 από τον Cleve Moler, πρόεδρο του τμήματος πληροφορικής του πανεπιστημίου του New Mexico. Ο σκοπός σχεδιασμού του ήταν η πρόσβαση των σπουδαστών στις βιβλιοθήκες LINPACK και EISPACK χωρίς να έχουν γνώσεις Fortran. Σύντομα διαδόθηκε σε άλλα πανεπιστήμια και βρήκε ισχυρό ακροατήριο εντός της κοινότητας των εφαρμοσμένων μαθηματικών.

Είναι ένα σύγχρονο ολοκληρωμένο μαθηματικό λογισμικό πακέτο που χρησιμοποιείται σε πανεπιστημιακά μαθήματα αλλά και ερευνητικές και άλλες εφαρμογές με επιστημονικούς υπολογισμούς (scientific computing). Το όνομά του προέρχεται από τα αρχικά γράμματα των λέξεων **MATrix LABoratory** (εργαστήριο πινάκων). Το MATLAB είναι ένα διαδραστικό (interactive) πρόγραμμα για αριθμητικούς υπολογισμούς και οπτικοποίηση δεδομένων (datavizualization) με δυνατότητες προγραμματισμού που το καθιστούν ένα ισχυρό και χρήσιμο εργαλείο στις μαθηματικές και φυσικές επιστήμες. Σε αντίθεση με τα λογισμικά Maple και Mathematica, το MATLAB στις αρχικές του εκδόσεις δεν έκανε συμβολικούς υπολογισμούς. Στις νεότερες εκδόσεις του, το πακέτο περιλαμβάνει εργαλείοι που επιτρέπουν συμβολικούς υπολογισμούς.

Όπως υποδηλώνεται και από το όνομά του, το MATLAB είναι ειδικά σχεδιασμένο για υπολογισμούς με πίνακες, όπως η επίλυση γραμμικών συστημάτων, η εύρεση ιδιοτιμών και ιδιοδιανυσμάτων, η αντιστροφή τετραγωνικών πινάκων κλπ. Επιπλέον το πακέτο αυτό είναι εφοδιασμένο με πολλές επιλογές για γραφικά (δηλ. την κατασκευή γραφικών παραστάσεων) και προγράμματα γραμμένα στη δική του γλώσσα προγραμματισμού για την επίλυση άλλων προβλημάτων όπως η εύρεση των ριζών μη γραμμικής εξίσωσης, η επίλυση μη γραμμικών συστημάτων, η επίλυση προβλημάτων αρχικών τιμών με συνήθεις διαφορικές εξισώσεις κα.

Η γλώσσα προγραμματισμού του MATLAB δίνει την ευχέρεια στον χρήστη να το επεκτείνει με δικά του προγράμματα. Συχνά θα γράφουμε η MATLAB (εννοώντας τη γλώσσα προγραμματισμού) και όχι το (πακέτο) MATLAB.

Το MATLAB είναι σχεδιασμένο για την αριθμητική επίλυση προβλημάτων σε αριθμητική πεπερασμένης ακρίβειας (finite-precision arithmetic), δηλαδή δεν βρίσκει την ακριβή αλλά μια προσεγγιστική λύση ενός προβλήματος. Αυτή είναι και η βασική του διαφορά από τα συστήματα συμβολικών υπολογισμών όπως η Maple και το Mathematica.

Ας σημειωθεί ότι ο καλύτερος (και ουσιαστικά ο μόνος) τρόπος εκμάθησης της MATLAB είναι η συστηματική ενασχόληση με αυτή και η διερεύνησή της από τον ίδιο τον χρήστη. Το πακέτο είναι εφοδιασμένο με ένα εκτενές σύστημα βοήθειας όπου κάθε εντολή επεξηγείται αναλυτικά και με αντιπροσωπευτικά παραδείγματα. Η πιο σημαντική εντολή της MATLAB είναι η help (βοήθεια)! Επίσης, στην επίσημη ιστοσελίδα της MATLAB: <http://www.mathworks.com> μπορεί κάποιος να βρει μια πληθώρα πληροφοριών τόσο για αρχάριους όσο και προχωρημένους.

### 5.1.1 Τι είναι μια γραφική διεπαφή χρήστη

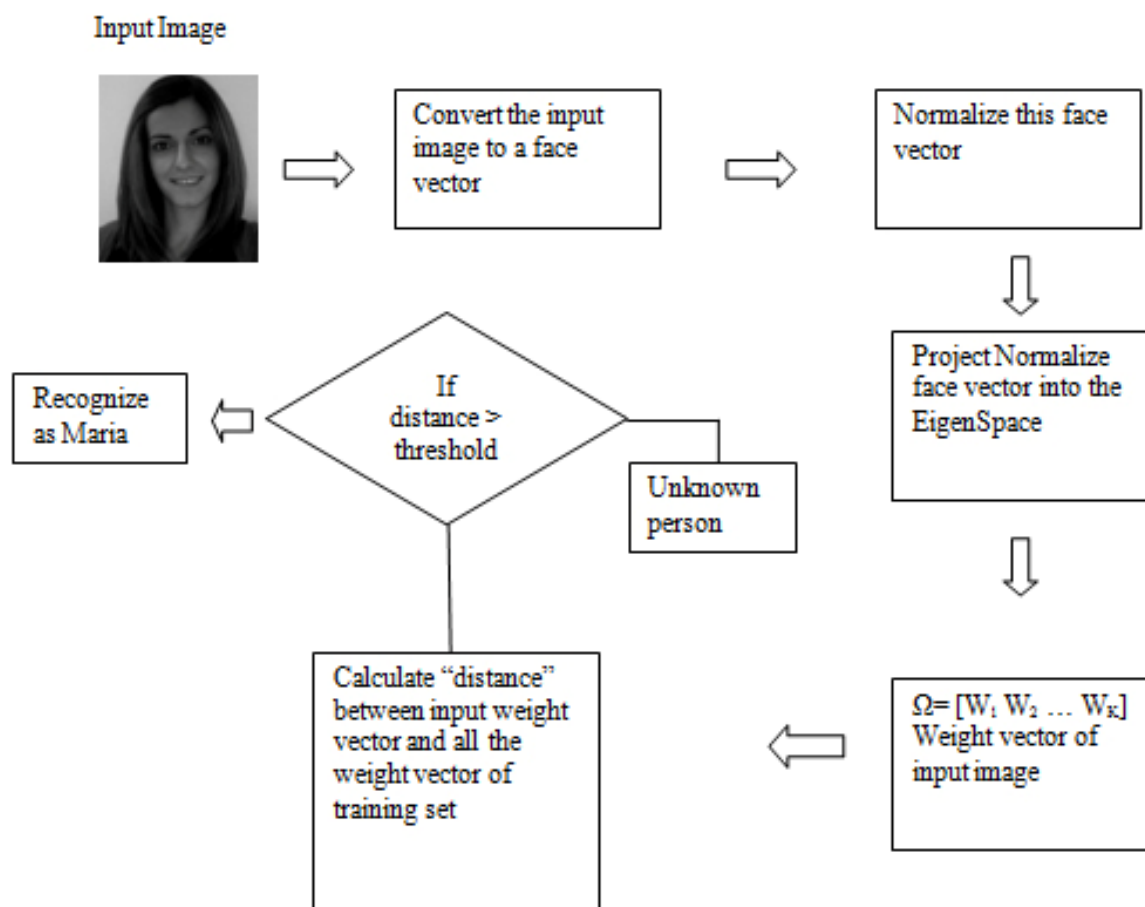
Μια γραφική διεπαφή χρήστη (Graphical User Interface - GUI) επιτρέπει στους ανθρώπους να αλληλεπιδρούν με ηλεκτρονικές συσκευές. Περιέχει εικόνες γραφικών και οπτικούς δείκτες, ετικέτες εντολών ή ένα αντικείμενο καθοδήγησης τα οποία δίνουν όλες τις απαραίτητες πληροφορίες και όλες τις δυνατές πράξεις του χρήστη. Τα περισσότερα σύγχρονα προγράμματα και λειτουργικά συστήματα υπολογιστών, προσφέρουν στους χρήστες τους κάποιο GUI γιατί αυτός ο τρόπος αλληλεπίδρασης με τον υπολογιστή ταιριάζει αρκετά στην ανθρώπινη εμπειρία και φύση. Σωστά σχεδιασμένα γραφικά προσφέρουν ένα όμορφο, εύχρηστο και λειτουργικό περιβάλλον εργασίας.

## 5.2 Αναγνώριση Προσώπου Με τη Χρήση Matlab

### 5.2.1 Υλοποίηση του PCA

Στη παρακάτω εφαρμογή χρησιμοποιήσαμε τον PCA αλγόριθμος βασισμένος στα Eigenface. Εδώ είναι το πώς λειτουργεί η αναγνώριση προσώπου βάση του αλγορίθμου των Eigenfaces, όταν εισάγουμε μέσα στο σύστημα μας μια εικόνα προς ταυτοποίηση.

f

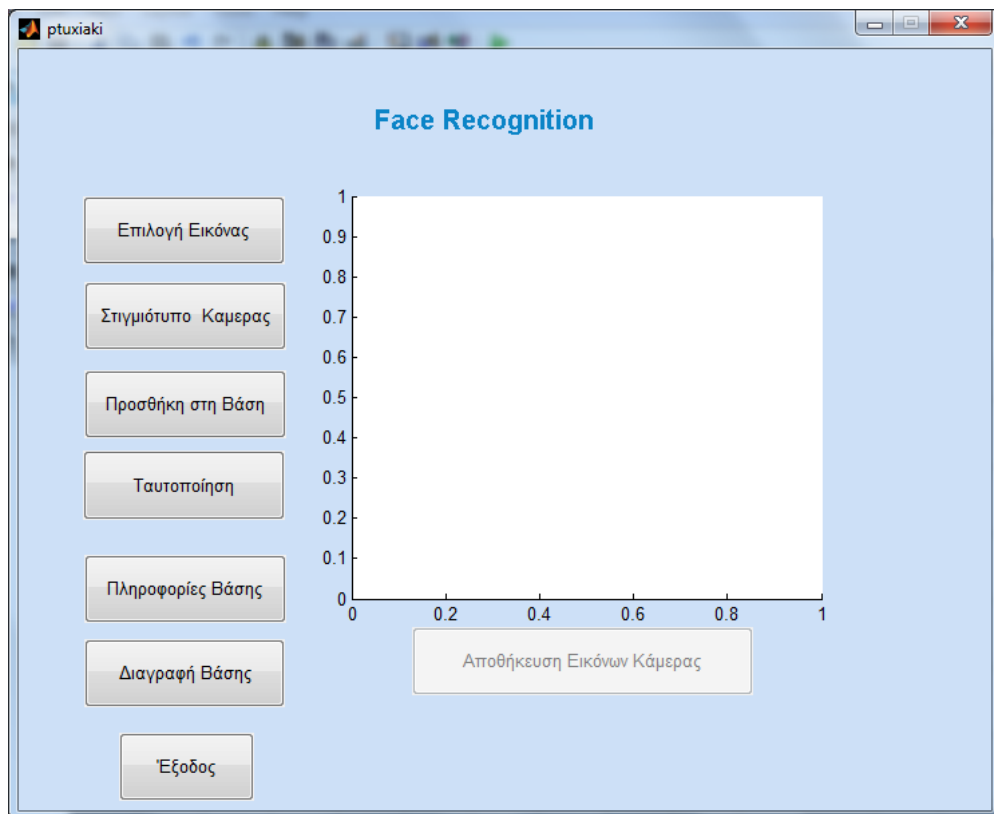


Εικόνα 40: Διάγραμμα Λειτουργίας του PCA

Η βάση που δημιουργήσαμε περιέχει 6 άτομα. Οι εικόνες έχουν ληφθεί στιγμές σε διάφορες όψεις, με εμπρόσθια όψη, προφίλ. Οι εκφράσεις προσώπου είναι αυθόρμητες και καθοδηγούμενες. Επίσης οι εικόνες έχουν καταγραφεί κάτω από διαφορετικές συνθήκες φωτισμού και είναι της μορφής pgm.

1. Μετατρέπουμε την εικόνα σε face vector
2. Κάνουμε κανονικοποίηση της εικόνας δηλαδή αφαιρούμε από την εικόνα το μέσο πρόσωπο που βρήκαμε από την εκπαίδευση των δεδομένων μας.
3. Υπολογίζουμε τα βάρη των eigenfaces έτσι να μπορέσουμε να αναπαραστήσουμε την εικόνα.
4. Βρίσκουμε την απόσταση των βαρών της εικόνας από των εικόνων εκπαίδευσης.
5. Αν η απόσταση αυτή ξεπερνάει το κατώφλι που έχει οριστεί δεν αναγνωρίζει το πρόσωπο διαφορετικά το αναγνωρίζει.

Στόχος της εργασίας μας είναι η επιτυχής έκβαση του αλγορίθμου αναγνώρισης προσώπου. Ο αλγόριθμος που χρησιμοποιήσαμε για την υλοποίηση είναι ο PCA ο οποίος βασίζεται στα Eigenfaces. Το interface που δημιουργήσαμε φαίνεται παρακάτω.

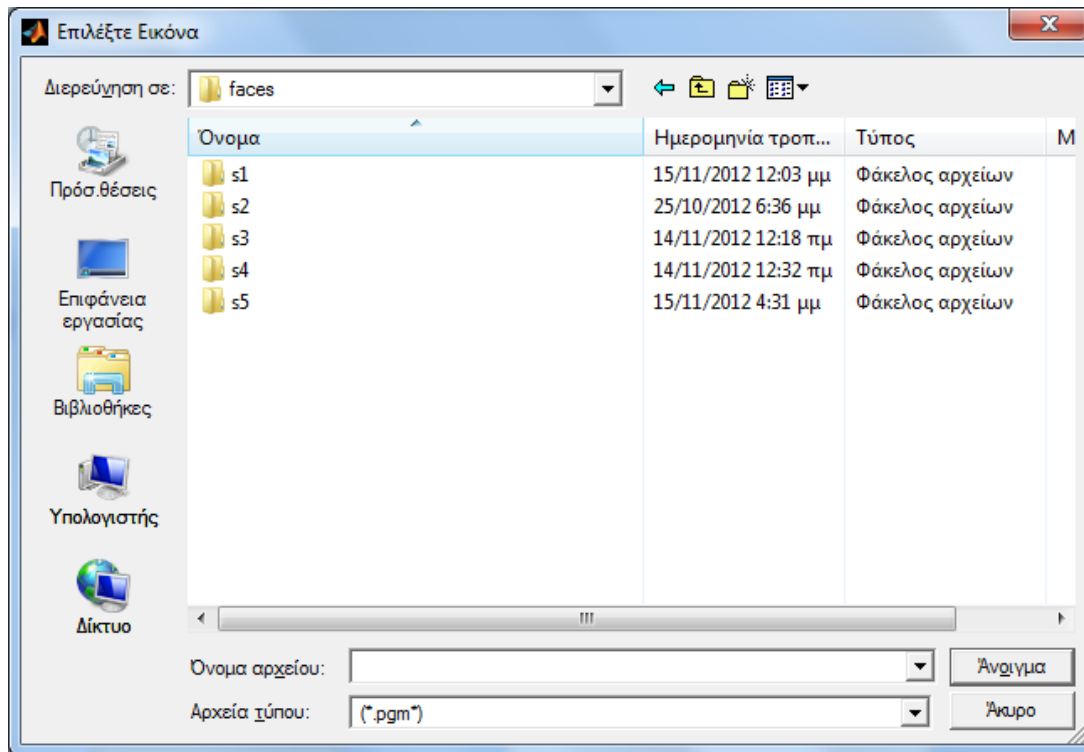


**Εικόνα 41:** Εκκίνηση διεπαφής

Όπως φαίνεται η διεπαφή μας απαρτίζεται από Push Buttons και Edit text.

Η επιλογή εικόνας μπορεί να γίνει με δύο τρόπους, μέσω του button Επιλογή Εικόνας ή Στιγμιότυπο Κάμερας. Επιλέγοντας Επιλογή Εικόνας ανοίγει ένα νέο παράθυρο μέσω του οποίου φορτώνουμε μια φωτογραφία. Η επιλογή μιας εικόνας γίνεται από το φάκελο faces που βρίσκονται όλα τα πρόσωπα της βάσης και τοποθετείται πάνω στο axes του παραθύρου.

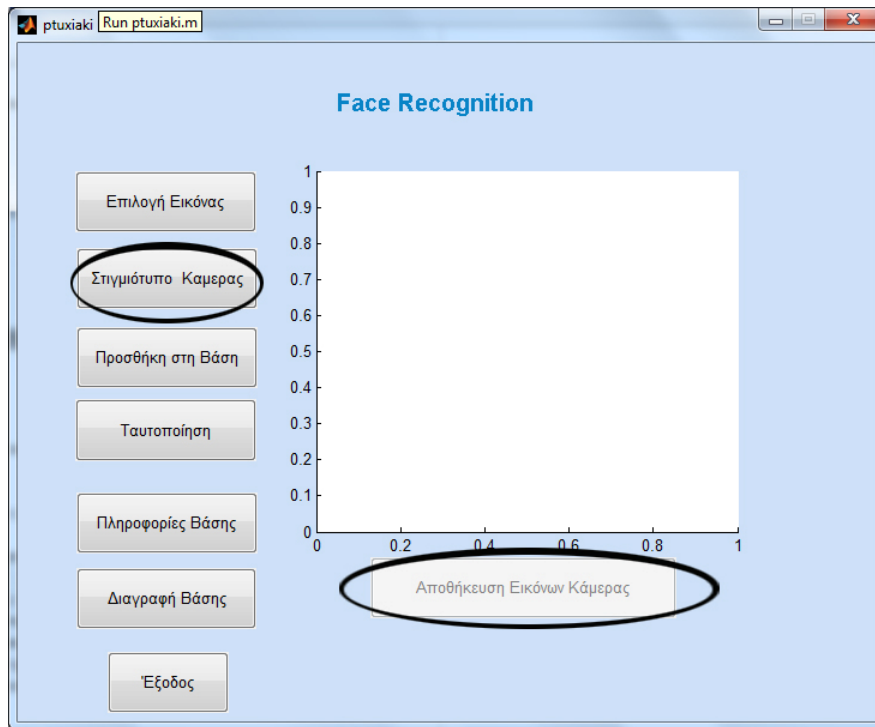




**Εικόνα 42:** Επιλογή Εικόνας μέσα από φάκελο

Όλες οι εικόνες είναι σε μορφή pgm ,διότι αν ήταν έγχρωμες αυτό θα είχε ως αποτέλεσμα να χρησιμοποιούμε τρία bytes για την αναπαράσταση του εικονοστοιχείου. Για μείωση όμως του χώρου που καταλαμβάνουν τα δεδομένα, όπως και για την μείωση του υπολογιστικού κόστους , οι εικόνες από έγχρωμες μεταφέρθηκαν στη κλίμακα του γκρι. Με αυτό τον τρόπο αποθηκεύουμε το κάθε εικονοστοιχείο σε ένα byte, στο ένα τρίτο του χώρου που θα χρειαζόμασταν αρχικά.

Δεδομένου ότι οι εικόνες μας θα πρέπει να είναι όλες στο ίδιο μέγεθος, έτσι ώστε να υλοποιηθεί σωστά ο αλγόριθμος, η λήψη των εικόνων που χρησιμοποιήθηκαν έχει γίνει μέσω της WebCam. Πατώντας το button Στιγμιότυπο Κάμερας μπορούμε να πάρουμε ένα στιγμιότυπο και το τοποθετεί στους άξονες.



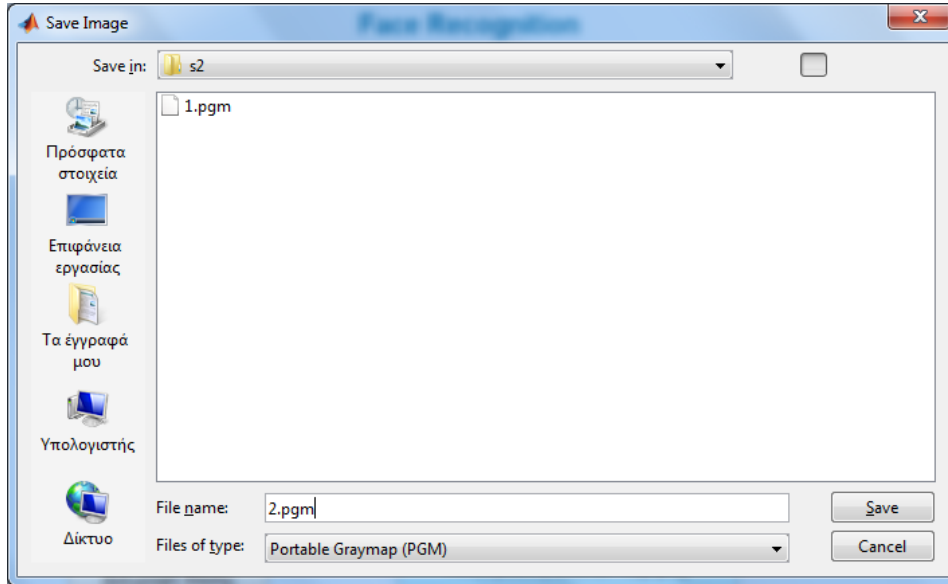
**Εικόνα 43:** Κουμπί Στιγμιότυπο Κάμερας

Παρατηρούμε ότι δεν μπορούμε να πατήσουμε το κουμπί Αποθήκευση Εικόνας εφόσον δεν έχουμε πατήσει για λήψη φωτογραφίας και σε οποιαδήποτε άλλη περίπτωση.



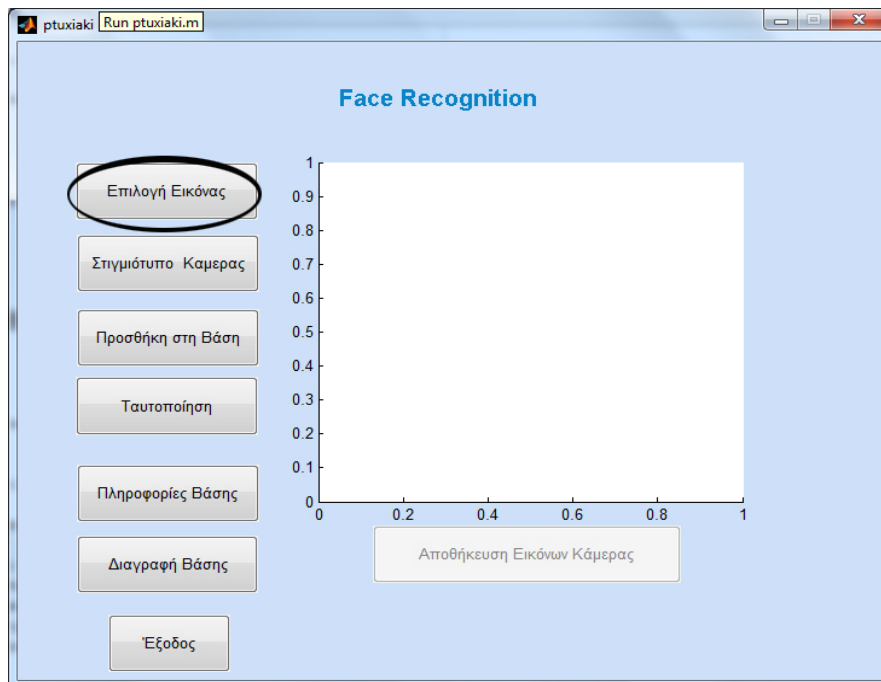
**Εικόνα 44:** Επιλογή εικόνας μέσω κάμερας

Επιλέγοντας αποθήκευση, αποθηκεύουμε την εικόνα στον αντίστοιχο φάκελο που έχουμε δημιουργήσει για κάθε άτομο ξεχωριστά. Μέσα σε αυτό το φάκελο όλες οι εικόνες είναι διαφορετικές, δηλαδή έχουν ληφθεί υπό διαφορετικές συνθήκες. Μπορούμε επίσης να χρησιμοποιήσουμε την εικόνα αυτή για να προσθέσουμε εικόνα στη βάση επιλέγοντας σε ποιο άτομο ανήκει, εκπαιδευοντας έτσι τον αλγόριθμο με περισσότερες εικόνες .

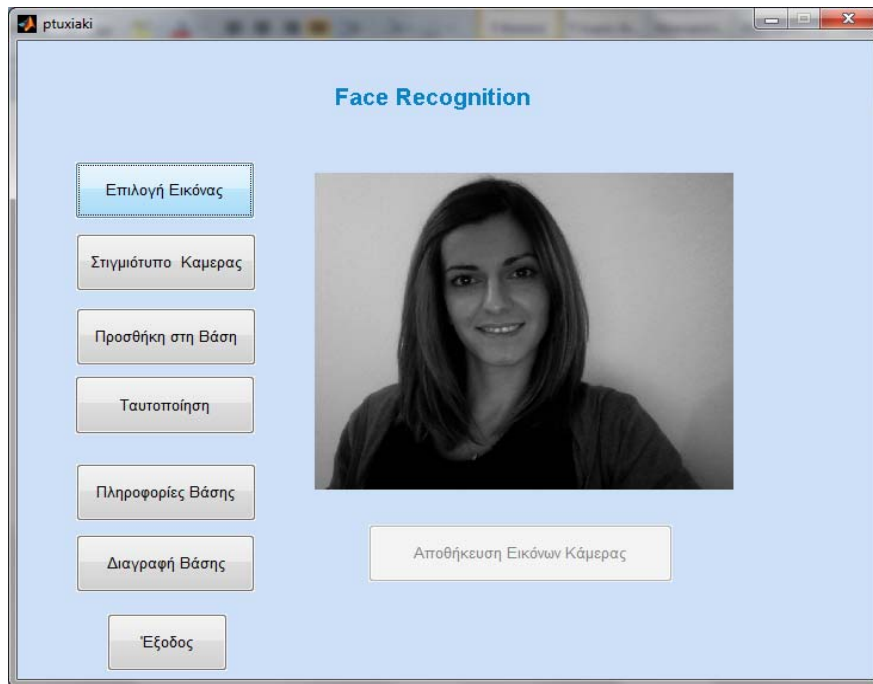


**Εικόνα 45:** Αποθήκευση Εικόνας μέσω Κάμερας

Ένας άλλος τρόπος που μπορούμε να επιλέξουμε μια εικόνα είναι πατώντας το κουμπί Επιλογή Εικόνας παρουσιάζοντας την ομοίως στους άξονες.

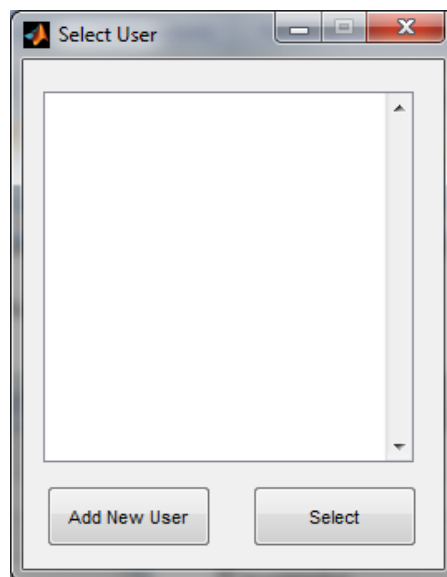


**Εικόνα 46:** Επιλογή Εικόνας



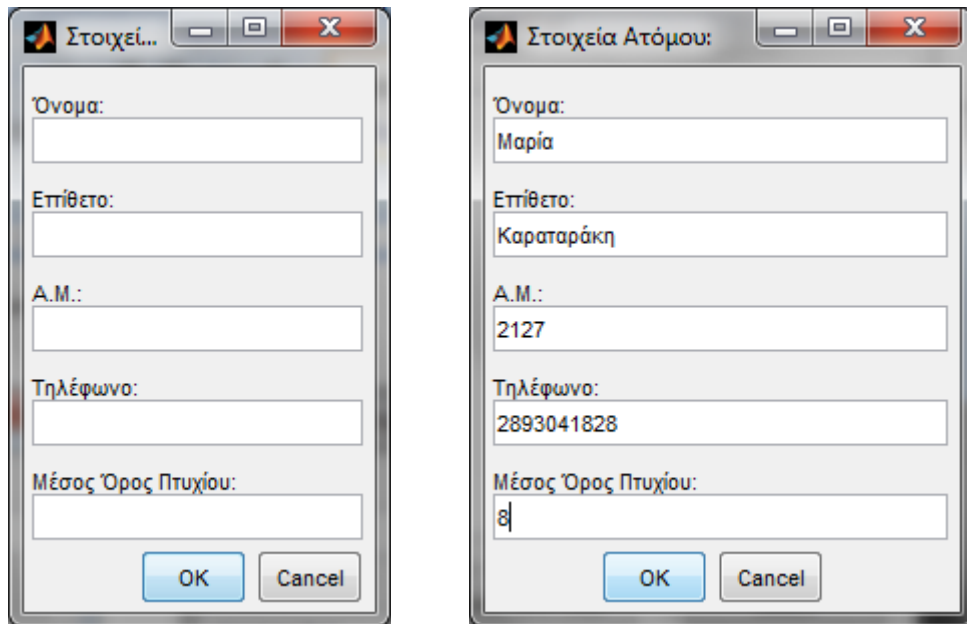
**Εικόνα 47:** Επιλογή εικόνας από φάκελο

Επιλέγοντας Προσθήκη στη Βάση αποθηκεύουμε την εικόνα στη βάση μέσω της οποίας θα γίνει η σύγκριση των προσώπων για ταυτοποίηση. Προτού την αποθηκεύσουμε μας ζητάει να του πούμε σε ποια κλάση θα ανήκει αυτή η εικόνα.



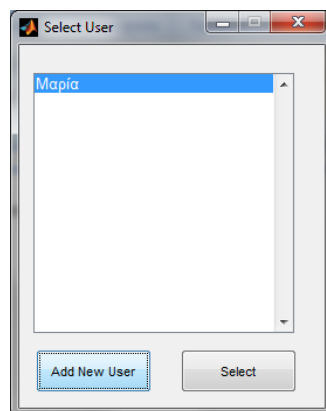
**Εικόνα 48:** Παράθυρο επιλογής ή προσθήκης χρήστη

Ο διαχωρισμός των κλάσεων γίνεται βάση ενός συγκεκριμένου γνωρίσματος, εμείς επιλέξαμε αυτό να είναι το όνομα. Για κάθε νέο πρόσωπο στη βάση, έχουμε το δικαίωμα προσθήκης κάποιων πληροφοριών πατώντας Add New User.



**Εικόνα 49:** Καταχώρηση γνωρισμάτων

Αποθηκεύοντας τα στοιχεία κατω από την εικόνα εμφανίζεται σε ποιά κλάση εντάσσεται η κάθε φωτογραφία καθώς επίσης και αν η βάση δεν υπάρχει τότε μας ενημερώνει ότι δημιουργήθηκε επιτυχώς. Επίσης μας δείχνει όλα τα ονόματα που έχουν καταχωρηθεί στη βάση κάθε φορά που επιλέγουμε προσθήκη στη βάση.





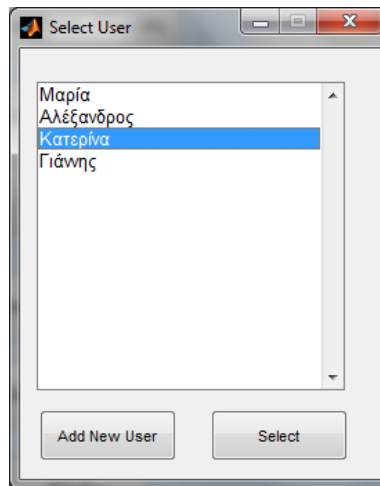
**Εικόνα 50:** Προσθήκη εικόνας στη βάση

Επιλέγοντας μια άλλη εικόνα του ίδιου ατόμου και συγκρίνοντας την με ήδη υπάρχουσες της βάσης ,επιλέγοντας Ταυτοποίηση μας εμφανίζει ποια είναι η κοντινότερη κλάση. Το αποτέλεσμα φαίνεται παρακάτω.



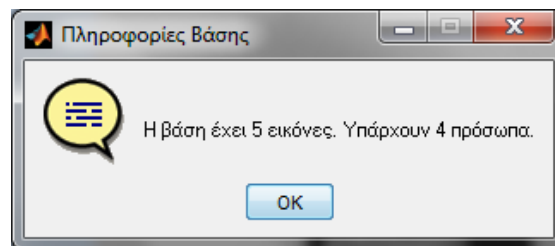
**Εικόνα 51:** (α)Αναγνώριση προσώπου (β)Στοιχεία ατόμου

Επίσης αν θέλουμε να προσθέσουμε περισσότερες από μία φωτογραφίες του ίδιου ατόμου στην ίδια κλάση, επιλέγουμε Προσθήκη στη Βάση και επιλογή του ονόματος ατόμου που ανήκει που η φωτογραφία. Έτσι ο αλγόριθμός μας θα έχει μεγαλύτερο ποσοστό επιτυχίας.



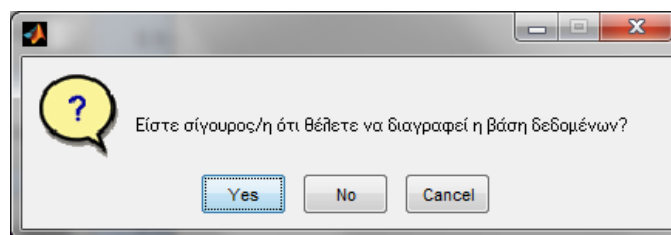
**Εικόνα 52:** Προσθήκη περισσότερων από μια φωτογραφιών στη βάση

Για να δούμε τι περιέχει η Βάση επιλέγουμε Πληροφορίες Βάσης και βλέπουμε πόσες εικόνες και πόσα άτομα την απαρτίζουν.

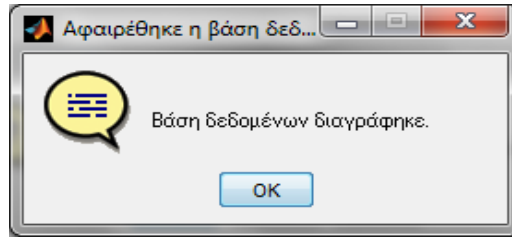


**Εικόνα 53:** Πληροφορίες Βάσης

Αν θελήσουμε να διαγράψουμε τη Βάση Δεδομένων μας έχουμε την επιλογή Διαγραφή Βάσης, όπου μας ρωτάει αν θέλουμε σίγουρα να τη διαγράψουμε και πατώντας OK η Βάση Δεδομένων μας έχει διαγραφεί ενημερώνοντας μας με ένα αντίστοιχο μήνυμα.



**Εικόνα 54:** Ερώτηση επιβεβαίωσης διαγραφής



Εικόνα 55: Επιβεβαίωση διαγραφής

### 5.2.2 Πειραματικά Αποτελέσματα

Στο σημείο αυτό παρουσιάζονται πειραματικά αποτελέσματα που πήραμε κατά την εφαρμογή του αλγορίθμου. Ο αλγόριθμος μας αποτελείται από πέντε κλάσεις. Παρακάτω ακολουθεί μια φωτογραφία του κάθε ατόμου.

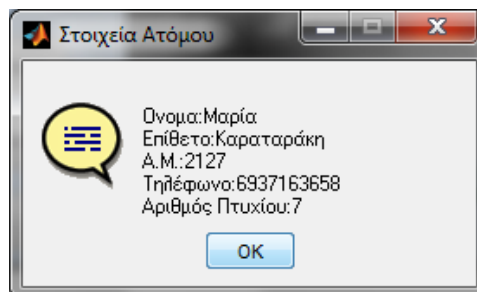
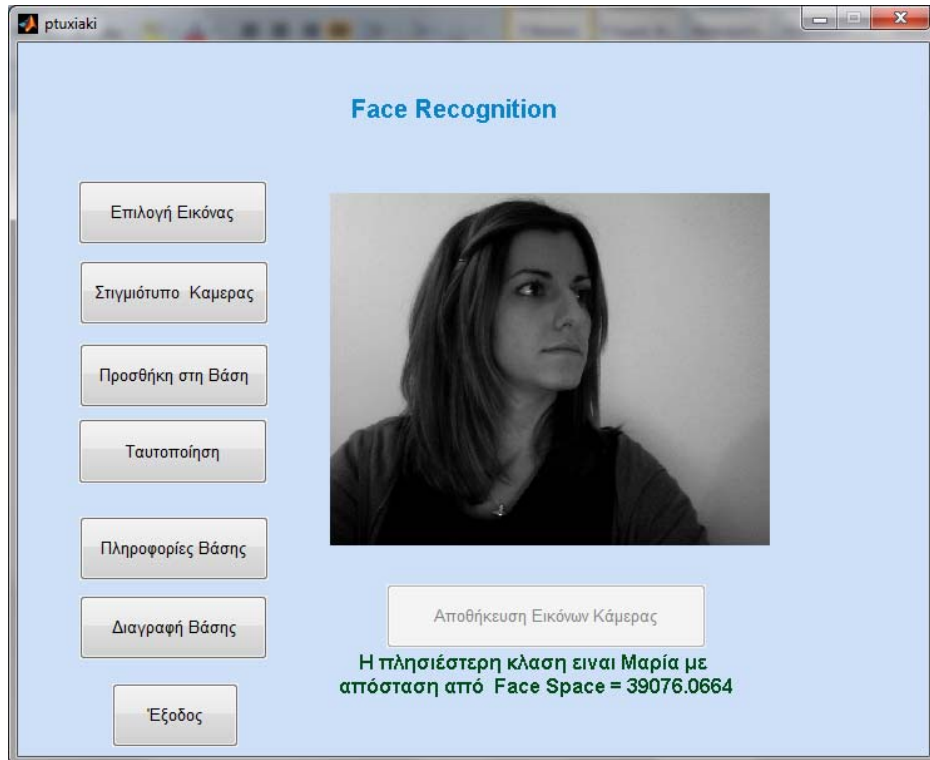


Εικόνα 56: Πρόσωπα Βάσης

Ο αλγόριθμός μας έχει εκπαιδευτεί με τις τρεις πρώτες φωτογραφίες από κάθε κλάση και με τις υπόλοιπες φωτογραφίες ελέγχεται η αποτελεσματικότητά του.



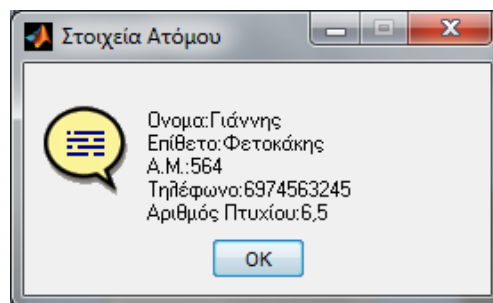
### Δοκιμή 1η: Εικόνα με πλάγια όψη



Εικόνα 57: Εικόνα με πλάγια όψη

Στο πρώτο πείραμα ελέγξαμε την αποτελεσματικότητα του αλγορίθμου με μια εικόνα σε πλάγια όψη. Το αποτέλεσμα που πήραμε είναι ότι το πρόσωπο 'Μαρία' είναι πιο κοντά στην εικόνα εισόδου. Παρατηρούμε ότι το 1 πείραμα είναι επιτυχημένο.

## Δοκιμή 2<sup>η</sup>: Εικόνα Προσώπου με γυαλιά

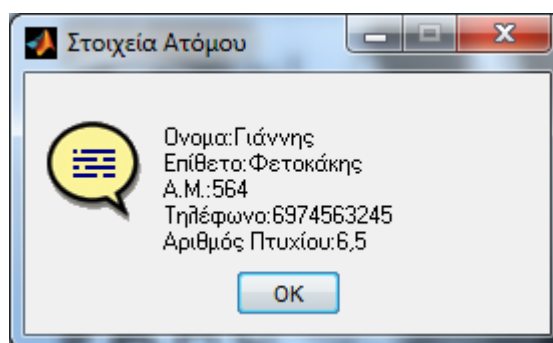
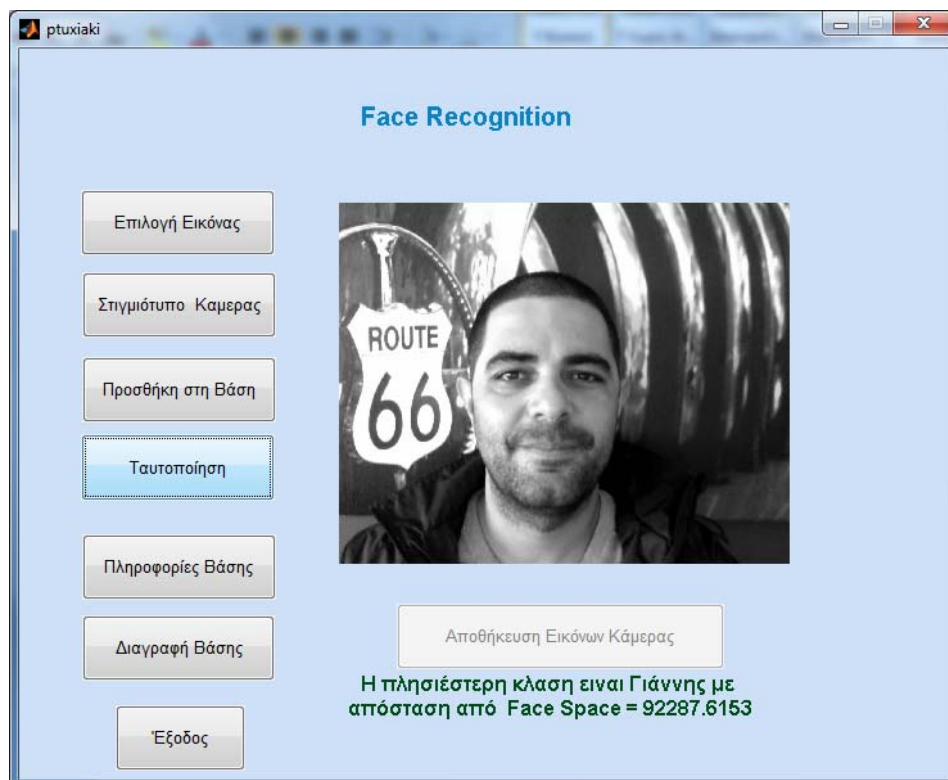


**Εικόνα 58:** Εικόνα Προσώπου με γυαλιά

Στο δεύτερο πείραμα δοκιμάσαμε ταυτοποίηση με χρήση γυαλιών. Παρατηρούμε ότι γίνεται σωστή ταυτοποίηση με τη χρήση γυαλιών.

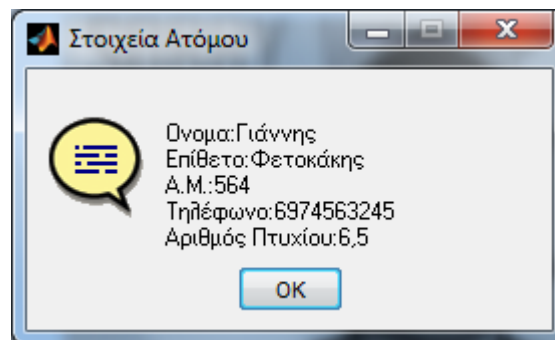
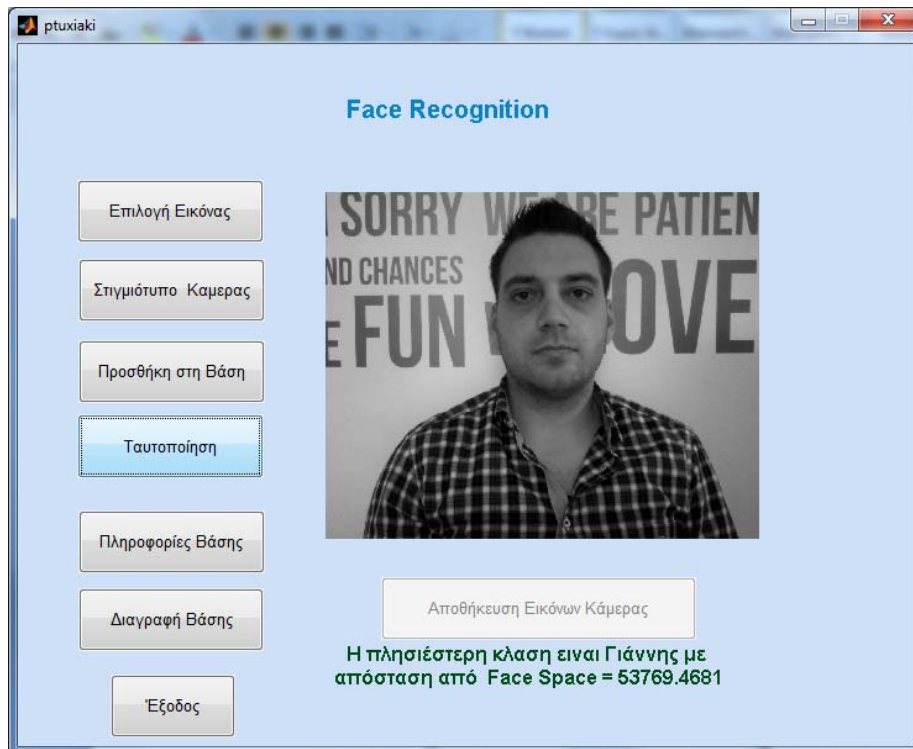
### Δοκιμή 3<sup>η</sup>: Σε διαφορετικό Φόντο

A' περίπτωση



Εικόνα 59: Λανθασμένη αναγνώριση σε ανομοιόμορφο φόντο

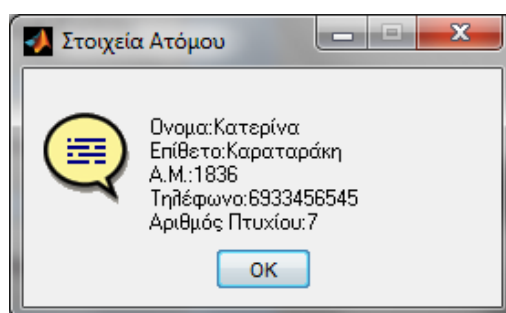
B' περίπτωση



**Εικόνα 60:** Σωστή αναγνώριση σε ανομοιόμορφο φόντο

Παρατηρούμε ότι με χρήση ανομοιόμορφου φόντου, στη Α' περίπτωση έχουμε λανθασμένη ταυτοποίηση, ενώ στη περίπτωση Β' έχουμε σωστή αναγνώριση, οπότε συμπεραίνουμε ότι υπάρχουν πολλές πιθανότητες για μη έγκυρα αποτελέσματα.

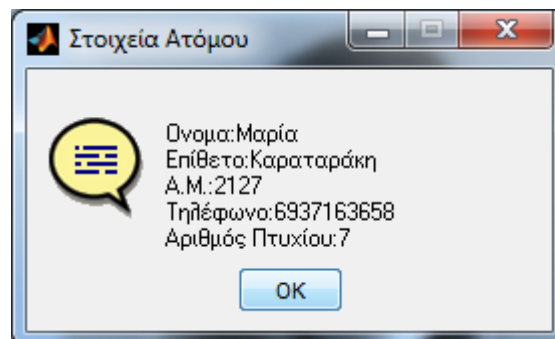
#### Δοκιμή 4<sup>η</sup>:Χρήση Καπέλου



Εικόνα 61: Σωστή αναγνώριση με χρήση καπέλου

Με χρήση καπέλου ο αλγόριθμος ταυτοποιηθεί σωστά το πρόσωπο που ελέγχουμε .

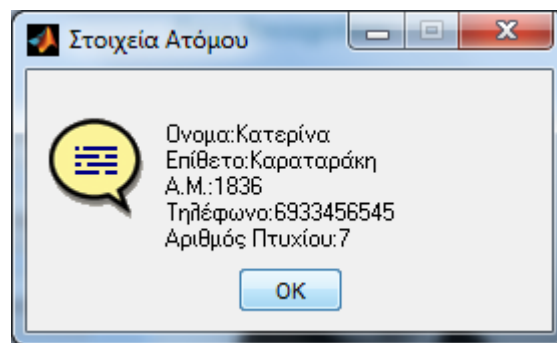
### Δοκιμή 5<sup>η</sup>: Αλλαγή φωτισμού



**Εικόνα 62:** Σωστή αναγνώριση με αλλαγή φωτισμού

Με αλλαγή του φωτισμού η απόδοση του αλγορίθμου δεν επηρεάζεται πραγματοποιώντας σωστή ταυτοποίηση.

**Δοκιμή 5<sup>η</sup>: Συνδυασμός κλίσης κεφαλιού και έκφρασης προσώπου**



**Εικόνα 63:** Σωστή αναγνώριση με κλίση του κεφαλιού σε συνδυασμό με έκφραση προσώπου

Η εφαρμογή ταυτοποιεί το πρόσωπα ανεξάρτητα των εκφράσεων του προσώπου σε συνδυασμό με κλίση κεφαλιού.

### **Μελλοντική Εξέλιξη**

Μια πρώτη πιθανή μελλοντική εφαρμογή των συστημάτων ταυτοποίησης προσώπου είναι η χρήση τους σε χώρους εστίασης για την αναγνώριση του προσωπικού κατά την προσέλευση και αποχώρησή τους από το χώρο εργασίας τους. Αυτό θα εξυπηρετούσε στην πιο άμεση και έγκυρη καταγραφή του συνολικού χρόνου εργασίας τους με αποφυγή τυχόν λανθασμένης ή εσκεμμένης εισαγωγής εργαζομένου.

Μία δεύτερη πιθανή εφαρμογή είναι η ταυτοποίηση των χρηστών ενός προγράμματος εστίασης έτσι ώστε ανάλογα τα δικαιώματα που θα έχει αναθέσει ο διαχειριστής ανά χρήστη, το πρόγραμμα αυτόματα θα ρυθμίζει τις δικαιοδοσίες που θα έχει ο εκάστοτε χρήστης.

Τέλος, οι αλγόριθμοι αναγνώρισης προσώπου αν επεκταθούν έτσι ώστε να μπορούν να αναγνωρίζουν εκφράσεις προσώπου όπως επίσης και κινήσεις προσώπου, θα μπορούσαν να λειτουργήσουν ως μία συσκευή εισόδου η οποία δουλεύει μόνο με πιστοποιημένους χρήστες. Αυτό θα μπορούσε να βοηθήσει άτομα με ειδικές ανάγκες τα οποία έχουν κινητικά προβλήματα να πιστοποιούνται και να χειρίζονται διάφορες εφαρμογές όπως να πλοηγούνται στο Internet, να έχουν πρόσβαση σε υπηρεσίες άμεσης ανάγκης, κλπ.



## Συμπεράσματα

Συμπεριλαμβανομένων των προαναφερθέντων, τα συστήματα αναγνώρισης προτύπων είναι ακόμα σε πολύ χαμηλό επίπεδο, σχετικά με αυτό που έχει στο μυαλό του κάποιος που ακούει το όνομά τους. Παρόλο που τα αποτελέσματά τους είναι έγκυρα και με πολύ μικρά ποσοστά σφάλματος υπό ιδανικές συνθήκες, δυσκολεύονται ιδιαίτερα όταν αυτές δεν πληρούνται. Για παράδειγμα, ο ερευνητής Ralph Gross του Ινστιτούτου Carnegie Mellon Robotics στο Pittsburgh, δηλώνει (-σε ελεύθερη μετάφραση-) "Η αναγνώριση προσώπου πηγαίνει πολύ καλά με πρόσωπα στραμμένα στο φακό έως και 20 μοίρες διαφορετικά, αλλά όσο μετακινείται προς το προφίλ, υπάρχουν προβλήματα".

Παρόμοια προβλήματα εμφανίζονται όταν δεν υπάρχει επαρκής φωτισμός, όταν τα πρόσωπα φέρουν γυαλιά ηλίου, όταν τα μαλλιά του προσώπου εμπλέκονται στο πρόσωπο και γενικά οτιδήποτε μπορεί να ελαττώσει τα χαρακτηριστικά που αντιλαμβάνεται το πληροφοριακό σύστημα για το πρόσωπο. Μειώνοντας τα χαρακτηριστικά, το σύστημα χάνει την ευστάθειά του και δεν μπορεί να λάβει απόφαση με την ίδια σιγουριά, όπως αν είχε όλες τις παραμέτρους.

## Βιβλιογραφία

- [1] M. Tekalp, "Face and 2-D Mesh Animation in MPEG-4," Tutorial Issue On The MPEG-4 Standard, Image Communication Journal, Elsevier, 1999.
- [2] Durand Andre (2002); Three Tiers of Identity; Digital Identity World, March 16, 2002, available at: <http://www.digitalidworld.com/print.php?sid=26>
- [3] Hildebrandt Mireille James Backhouse Eds. (2005); "Descriptive analysis and inventory of profiling practices", FIDIS deliverable 7.2, June 2005.
- [4] ΕΦΗΜΕΡΙΣ ΤΗΣ ΚΥΒΕΡΝΗΣΕΩΣ ΤΗΣ ΕΛΛΗΝΙΚΗΣ ΔΗΜΟΚΡΑΤΙΑΣ, ΤΕΥΧΟΣ ΔΕΥΤΕΡΟ Αρ Φύλλου 650, 9 ΑΠΡΙΛΙΟΥ 2009
- [5] Thierry Nabeth, Mirielle Hildebrandt, D 2.1: Inventory of topics and clusters, 2005.
- [6] Janeen Renaghan. "Etched in Stone," Zoogoer, August 1997, (Smithsonian National Zoological Park, 26 January 2005).
- [7] "Dermatoglyphics," Hand Analysis, International Institute of Hand Analysis, 24 January 2005.
- [8] Z. McMahon, Biometrics: History, Indiana University, Indiana University Computer Science Department, 24 January 2005, available at: <http://www.cs.indiana.edu/~zcmahon/biometrics-history.htm>
- [9] Alphonse Bertillone, access on 4<sup>th</sup> May 2009, available at: <http://en.wikipedia.org/wiki/Bertillon>
- [10] Bonsor, K.. "How Facial Recognition Systems Work" access on 4th May 2009, available at: <http://electronics.howstuffworks.com/facial-recognition.htm>
- [11] V. Govindaraju, D.B.Sher, R.K.Srihari and S.N.Srihari, "Locating of human faces in newspapers photographs," Proc. of CVBR, pp. 549-554, 1989.
- [12] V. Govindaraju, S.N.Srihari and D.B.Sher, "A computational model for face location," Proc. of 3rd Int. Conf. on Computer Vision, pp. 718-721, 1990.
- [13] R.A.Campell, S. Cannon, G. Jones and N. Morgan, "Individual face identification by computer vision," Proc. Conf. Modeling Simulation Microcomputer, pp. 62-63, 1987.
- [14] Sir F. Galton, "Personal identification and description - I," Nature, pp.173-177, June 1888.
- [15] Sir F. Galton, "Numeralized profile for classification and recognition," Nature 83, pp.127-130, March 1910.
- [16] T. Sakai, M. Nagao and S. Fujibayshi, "Line extraction and pattern recognition in a photograph," Pattern Recognition 1, pp. 233-248, 1969.
- [17] M. Kelly, "Visual identification of people by computer," Technical Report AI-130, Stanford, CA, 1970
- [18] I. Craw, H. Ellis and J.R.Lishman, "Automatic extraction of face-features," Pattern Recognition Letters 5, pp.183-187, 1987.
- [19] I. Craw, D. Tock and A. Benett, "Finding face features," Proc. of 2nd European Conference on Computer Vision, pp.92-96, 1992.
- [20] J. Sepherd, "An interactive computer system for retrieving faces," Ellis (Eds), Dordrecht: Nijhoff, pp. 398- 409, 1985.
- [21] S. Sirohey, "Human face segmentation and identification," Technical Report CAR-TR-695, University of Maryland, MD, 1993.
- [22] J.L.Perry and J.M.Carney, "Human face recognition using a multilayer perceptron," Proc. of Int. Conf. on Neural Networks II, pp. 413, January 1990.
- [23] A.L.Yulie, D.S.Cohen and P.W.Hallinan, "Feature extraction of faces using deformable templates," Proc. Of CVBR, pp. 104-109, 1989.
- [24] M. Nixon, "Eye spacing measurements for facial recognition," SPIE Proc. 575, Applications of Digital Image Processing VIII, pp. 279-285, 1985.
- [25] L.D.Harmon, "Automatic recognition of human face profiles," Proc. of 3rd Int. Joint Conf. on Pattern Recognition, pp. 183-188, 1976.
- [26] L.D.Harmon, S.C.Kuo, P.F.Ramig and U. Raudkivi, "Identification of human face profile by computers," Pattern Recognition 10, pp. 301-312, 1978.

- [27] L.D.Harmon, M.K.Khan, R. Lasch and P.F.Ramig, "Machine identification of human faces," *Pattern Recognition* 13, pp. 97-110, 1981.
- [28] G.J.Kaufman K.J.Breeding, "The automatic recognition of human faces from profile silhouettes," *IEEE Trans. on System Man and Cybernetics* 6, pp. 113-121, 1976.
- [29] A.J.Goldstein, L.D.Harmon and A.B.Lesk, "A basic study on human face recognition," *Proc. of IEEE* 59, pp. 748-760, May 1971.
- [30] Y. Kaya and K. Kobayashi, "Identification of human faces," *Frontiers of Pattern Recognition*, pp. 265-289, Academic Press, New York 1971.
- [31] R. Buhr, "Analyse und klassifikation von gesichtsbildern," *ntzArchiv* 8, pp. 245-256, 1986.
- [32] K.H.Wong, H.M.Law and P.W.M.Tsang, "A system for recognizing human faces," *Proc. of ICASSP*, pp. 1638-1642, 1989.
- [33] R.J.Baron, "Mechanisms of human facial recognition," *Int. Journal of Man Machine Studies* 15, pp. 137-178, January 1990.
- [34] K. Preston Jr, "Computing at the speed of light," *Electronics* 38, pp. 72-83, 1965.
- [35] W.K.Taylor, "Machine learning and recognition of faces," *Electronic Letters* 3, pp. 436-437, 1967.
- [36] Xiaoguang Lu, "Image Analysis for Face Recognition," Michigan State University, Dept. of Computer Science & Engineering, Michigan, East Lansing, MI, 48824,.
- [37] W Bledsoe, "The Model Method in Facial Recognition," Panoramic Research, Inc, Palo Alto, California., Technical Report PRI 15, 1964.
- [38] W. Bledsoe and H. Chan, "A Man-Machine Facial Recognition System-Some Preliminary Results," Panoramic Research, Inc, Palo Alto, California., Technical Report PRI 19A, 1965.
- [39] W. Bledsoe, "Semiautomatic Facial Recognition," Stanford Research Institute, Menlo Park, California, Technical Report SRI Project 6693, 1968.
- [40] Takeo Kanade, *Picture Processing by Computer Complex and Recognition of Human Faces*. Kyoto , Japan: doctoral dissertation, 1973.
- [41] W. Zhao, R. Chellappa, A. Rosenfeld, and P.J. Phillips, "Face Recognition: A Literature Survey," in *ACM Computing Survey*, 2003, pp. 399-458.
- [42] L. Wiscott, J.M. Fellous, N. Kuiger, and C. von Der Malrsburg, "Face recognition by elastic bunch graph matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence In Pattern Analysis and Machine Intelligenc*, vol. 19, no. 7, pp. 775-779, July 1997.
- [43] F. Samaria and S. Young, "HMM based architecture for face identification," *Image and Computer Vision*, vol. 12, pp. 537-583, October 1994.
- [44] V. Nefian and H. Hayes III, "Hidden Markov Models for face recognition," in *IEE International Conference on Acoustic speech and Signal Processing*, 1998.
- [45] L. Sirovich and M. Kirby, "Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces," *IEEE Transactions on Pattern Analysis 149 and Machine Intelligence*, vol. 4, no. 3, pp. 519-524, March 1987.
- [46] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscence*, vol. 3, no. 1, pp. 71-86, 1991.
- [47] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, July 1997.
- [48] W. Zhao, R. Chellappa, A. Rosenfeld, and P.J. Phillips, "Face Recognition: A Literature Survey," in *ACM Computing Survey*, 2003, pp. 399-458.
- [49] D.L. Swets and J.J. Weng, "Using Discriminant Eigenfeatures for Image Retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 8, pp. 831-836, 1996.
- [50] M.S. Bartlett, "Independent component representations for face recognition," in *Symposium on Electronic Imaging: Science and Technology—Conference on Human Vision and Electronic Imaging III*, San Jose, 1998.
- [51] Marian S Bartlett, Javier R. Movellan, and Terrence J. Sejnowski, *IEEE Transactions on Neural Networks*, vol. 13, no. 6, pp. 1450-1463, November 2002.
- [52] Anthony J Bell and Terrence J Sejnowski, "An Information-Maximization Approach to Blind Separation and Blind Deconvolution," *Neural Comput*, vol. 7, no. 6, pp. 1129-1159, 1995.

- [53] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.
- [54] A. Pentland, B. Moghaddam, and T. Starner. View-based and modular eigenspaces for face recognition. In *Proceedings of Fourth IEEE International Conference on Computer Vision*, pages 84–91, 1994.
- [55] B. Moghaddam and A. Pentland. Probabilistic visual learning for object recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):696–710, 1997.
- [56] A Papoulis and S U Pillai, *Probability, Random Variables and Stochastic Processes with Errata Sheet*, 4th ed.: McGraw-Hill, 2002.
- [57] M. S. Bartlett : *Face Image Analysis by Unsupervised Learning and Redundancy Reduction*, 1998, pp. 1–51.
- [58] M. S. Bartlett, H. Lades, and T. J. Sejnowski : Independent component representations for face recognition, in *Proc. SPIE Symp. Electron. Imaging: Science Technology Human Vision and Electronic Imaging III*, vol. 3299, T. Rogowitz and B. Pappas, Eds., San Jose, CA, 1998, pp. 528–539.
- [59] A. J. Bell and T. J. Sejnowski : n informationmaximization approach to blind separation and blind deconvolution, *Neural Comput.*, vol. 7, no. 6, 1995, pp. 1129– 1159.
- [60] M. S. Bartlett, J. R. Movellan and T. J. Sejnowski : Face recognition by independent component analysis, *IEEE Transactions on Neural Networks*, vol.13, 2002, pp. 1450–1464.
- [61] K. Etemad and R. Chellappa, "Discriminant Analysis for Recognition of Human Face Images," *Journal of the Optical Society of America A*, vol. 14, no. 8, pp. 1724-1733, August 1997.
- [62] Mathworks Inc., "Image processing toolbox," MATLAB 7.,.
- [63] A Papoulis and S U Pillai, *Probability, Random Variables and Stochastic Processes with Errata Sheet*, 4th ed.: McGraw-Hill, 2002.
- [64] Alaa Eleyan and Hasan Demirel, "PCA and LDA Based Face Recognition Using Feedforward Neural Network Classifier," in *Multimedia Content Representation, Classification and Security.*: Springer Berlin / Heidelberg, 2006, pp. 199-206.
- [65] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, 1997.
- [66] F. L. Chen, H. Y. Mark Liao, M. T. Ko, J. C. Lin, and G. J. Yu, "A new LDA based face recognition system which can solve the small sample size problem," *Pattern Recognition*, vol. 33, pp. 1713-1726, 2000.
- [67] H. Yu and J. Yang, "A direct lda algorithm for high-dimensional data with application to face recognition," *Pattern Recognition*, vol. 34, pp. 2067-2070, 2001.
- [68] Juwei Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Face Recognition Using LDA Based Algorithms," *IEEE Trans. on Neural Networks*, vol. 14, no. 1, pp. pp. 195-200, January 2003.
- [69] J.H. Friedman, "Regularized discriminant analysis," *Journal of the American Statistical Association*, vol. 84, pp. 165-175, 1989.
- [70] J. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Regularization Studies of Linear Discriminant analysis in small sample size scenarios with application to face recognition," *Pattern Recognition Letters*, vol. 26, no. 2, pp. 181-191, January 2005.
- [71] A. N. Rajagopalan, K. S. Kumar, J. Karlekar, R. Manivasakan, and M. M. Patil, "Finding faces in photographs," in *IEEE Int'l Conf. on Computer Vision*, 1998, pp. 640-645.
- [72] B. Scholkopf, A. Smola, and K. R. Muller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, vol. 10, no. 5, pp. 1299- 1319, 1998.
- [73] S. Mika et al., "Invariant feature extraction and classification in kernel spaces," *NIPS 12*, pp. 526-532, 2000.
- [74] V. Roth and V. Steinhage, "Nonlinear discriminant analysis using kernel functions," *NIPS 12*, pp. 268-574, 2000.

[75] G. Baudat and F. Anouar, "Generalized discriminant analysis using a kernel approach," *Neural Computation*, vol. 12, pp. 2385-2404, 2000.

[76] B. Scholkopf and A. J. Smola, Cambridge: MIT Press, 2001.

[77] B. Scholkopf, C. Burges, and V. Vapnik, "Extracting support data for a given task," in *First International Conference on Knowledge Discovery & Data Mining*, Menlo Park, CA, 1995.

[78] B. Scholkopf, A. J. Smola, and K. R. Muller, "Kernel Principal Component Analysis," *Neural Information Processing Systems*, vol. 1327, no. 3, pp. 327- 352, 1999.

[79] <http://www.mathworks.com/matlabcentral/fileexchange/35106-the-phd-face-recognition-toolbox>