

Πολυτεχνείο Κρήτης
Τμήμα Ηλεκτρονικών Μηχανικών και
Μηχανικών Ηλεκτρονικών Υπολογιστών



Διπλωματική εργασία

«Αναγνώριση προσώπων μέσω προβολών»

Εφραίμης Λευτέρης

Επιβλέπων Καθηγητής: Ζερβάκης Μιχάλης (καθηγητής)
Εξεταστική Επιτροπή: Πετράκης Ευριπίδης (αναπληρωτής καθηγητής)
Μπάλας Κωνσταντίνος (αναπληρωτής καθηγητής)

Χανιά, Μάρτιος 2011

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον Επιβλέποντα καθηγητή της εργασίας κ. Ζερβάκη Μιχάλη για τη συνεργασία και την κατανόηση καθ' όλη τη διάρκεια της εργασίας.

Αφιερώνω αυτή τη διπλωματική στους γονείς μου και στους φίλους που με βοήθησαν να επιστρέψω και να ολοκληρώσω τον κύκλο μου στα Χανιά.

Περιεχόμενα

Περιεχόμενα.....	3
Εικόνες.....	5
1 Εισαγωγή.....	6
1.1 Βιομετρική ταυτοποίηση.....	6
1.2 Τύποι βιομετρικών στοιχείων.....	7
1.2.1 Δακτυλικά αποτυπώματα.....	7
1.2.2 Αναγνώριση Ίριδας.....	8
1.2.3 Ταυτοποίηση υπογραφής.....	8
1.2.4 Αναγνώριση φωνής.....	8
1.2.5 Αναγνώριση προσώπου.....	9
1.3 Ανάλυση πρωτευόντων χαρακτηριστικών PCA.....	11
1.4 Στοιχεία σύγκρισης για ταχύτητα.....	13
1.5 Σκοπός της εργασίας.....	15
1.6 Τομείς πιθανών εφαρμογών.....	17
2 Μεθοδολογία.....	19
2.1 Εισαγωγή.....	19
2.2 Χώρος προσώπου και αριθμός διαστάσεων του.....	20
2.3 Χώρος Εικόνας Vs χώρος προσώπου.....	20
2.4 Πρωτεύοντα χώρος και βασικές συναρτήσεις.....	21
2.5 Ανάλυση κύριων συνιστωσών.....	21
2.5.1 Εξαγωγή χαρακτηριστικών PCA.....	21
2.5.2 Σημασία των διαφορετικών ιδιοτιμών.....	22
2.5.3 Πιθανή Προεπεξεργασία.....	27
2.5.4 Αντικρουόμενοι στόχοι ανάλυσης.....	27
2.6 Υπολογισμός Ιδιοπροσώπων.....	27
2.7 Κατηγοριοποίηση εικόνων προσώπων με χρήση Ιδιοπροσώπων.....	31
2.8 Ανάλυση με εικόνα σε στήλες vs ανάλυση με εικόνα σε γραμμές του πίνακα συνδιακύμανσης.....	32
2.9 Περίληψη της αναγνώρισης με Ιδιοπρόσωπα.....	33
3 Υλοποίηση.....	34
3.1.1 Προετοιμασία βάσης δεδομένων.....	34
3.1.2 Εκπαίδευση.....	35
3.1.3 Έλεγχος.....	36
4 Αποτελέσματα.....	38
4.1 Ορισμός τεστ και χειρισμός του συνόλου δεδομένων.....	38
4.2 Αριθμός διατηρούμενων ιδιοδιανυσμάτων.....	39
4.2.1 Πειραματικά αποτελέσματα που προέκυψαν από την διακύμανση της ανάλυσης.....	46

4.2.2	Πειραματικά αποτελέσματα που προέκυψαν από την διακύμανση του αριθμού των διατηρούμενων ιδιοτιμών	46
4.2.3	Πειραματικά αποτελέσματα που προέκυψαν από την διακύμανση το μέγεθος ποιότητα των εικόνων, για δεδομένη ανάλυση.....	48
4.2.4	Πειραματικά αποτελέσματα που πήραμε από την αφαίρεση των πρώτων ιδιοτιμών (illumination compensation)	49
4.2.5	Κανονικοποίηση	50
4.3	Επεξεργαστική Ισχύς και χρονική ανάλυση	51
4.4	Περιορισμοί του αλγορίθμου	53
5	Συμπεράσματα	54
6	Μελλοντικές προεκτάσεις	55
	Βιβλιογραφία	57
	Παράρτημα Α.....	58
	Παράρτημα Β.....	72
	Βοηθητικές Μαθηματικές έννοιες και παράδειγμα PCA.....	72
6.1.1	Στατιστική	72
6.1.2	Μέσος όρος.....	72
6.1.3	Τυπική απόκλιση	72
6.1.4	Διακύμανση	73
6.1.5	Συνδιακύμανση	73
6.1.6	Πίνακας συνδιακύμανσης	74
6.1.7	Άλγεβρα πινάκων.....	75
6.1.8	Ιδιοτιμές	77

Εικόνες

Εικόνα 1: Παράδειγμα κανονικοποιημένης εικόνας προσώπου από τη βάση FERET που χρησιμοποιήθηκε στη δημοσίευση [9].....	14
Εικόνα 2: κλειδαριά αναγνώρισης προσώπου	17
Εικόνα 3 παράδειγμα εικόνας και μεγέθυνση τμήματος σε επίπεδο ορατού pixel.	19
Εικόνα 4 Αριστερά αρχικές εικόνες και ιδιοδιανύσματα που διατηρούν το 99,9% της αρχικής ενέργειας. Δεξιά τροποποιημένο ως προς το ιστόγραμμα data set και ιδιοδιανύσματα που περιέχουν το ίδιο ποσοστό ενέργειας.....	23
Εικόνα 5: Το μέσο πρόσωπο ψ	28
Εικόνα 6: Παράδειγμα γ	28
Εικόνα 7: Παράδειγμα $\phi=\gamma-\psi$	28
Εικόνα 8: Παράδειγμα συνόλου προσώπων εκπαίδευσης.....	30
Εικόνα 9: Παράδειγμα ιδιοπροσώπων	30
Εικόνα 10: μέθοδος δικυβικής παρεμβολής (bicubic interpolation).....	38
Εικόνα 11: Πίνακας δεδομένων με εφαρμογή PCA και χρήση 2 ιδιοδιανυσμάτων.....	83
Εικόνα 12: Δεδομένα μετά από μετασχηματισμό PCA με διατήρηση μόνο του πρώτου ιδιοδιανύσματος.	83
Εικόνα 13: Ανακατασκευασμένα αρχικά δεδομένα με τη χρήση του ελαττωμένου διανύσματος χαρακτηριστικών.	84

1 Εισαγωγή

Στις κοινωνικές επαφές μας το πρόσωπο αποτελεί το κύριο σημείο ενδιαφέροντος, παίζει σημαντικό ρόλο στην μεταφορά ταυτότητας αλλά και συναισθήματος. Μπορούμε να αναγνωρίσουμε εκατοντάδες πρόσωπα που μάθαμε στη διάρκεια της ζωής μας, και να αναγνωρίζουμε οικεία πρόσωπα με μια ματιά ακόμα και μετά από χρόνια απουσίας. Αυτό είναι εφικτό παρά τις μεγάλες αλλαγές στην οπτική διέγερση που μπορεί να οφείλονται σε οπτικές συνθήκες όπως κακός φωτισμός, έκφραση, γήρανση ή περιφερικές αποσπάσεις όπως αλλαγές στα μαλλιά ή στα μούσια ή γυαλιά. Η δυνατότητα αναγνώρισης του ατόμου ενισχύεται όσο μεγαλώνει ο βαθμός οικειότητας με το συγκεκριμένο πρόσωπο.

Τα υπολογιστικά μοντέλα αναγνώρισης προτύπων, είναι ενδιαφέροντα επειδή μπορούν να συνεισφέρουν όχι μόνο στη θεωρητική βαθύτερη κατανόηση αλλά και σε πρακτικές εφαρμογές. Υπολογιστές που αναγνωρίζουν πρόσωπα μπορούν να βρουν εφαρμογή σε μια ευρεία γκάμα προβλημάτων, συμπεριλαμβανομένης της αναγνώρισης εγκληματιών σε δημόσιους χώρους, σε συστήματα ασφαλείας, επεξεργασία εικόνας και video και στην ανθρώπινη αλληλεπίδραση με τον υπολογιστή. Δυστυχώς η ανάπτυξη ενός υπολογιστικού μοντέλου αναγνώρισης προσώπων είναι αρκετά δύσκολη, επειδή τα πρόσωπα είναι πολύπλοκα, πολυδιάστατα οπτικά ερεθίσματα που μεταφέρουν σημασιολογική πληροφορία.

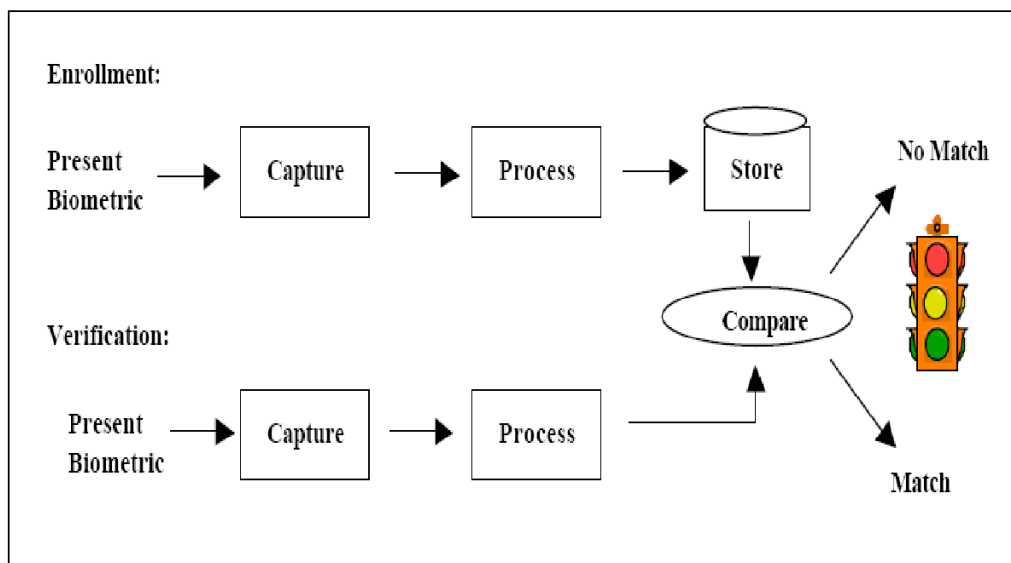
Η πρόκληση έγκειται στην υλοποίηση ενός συστήματος με δυνατότητα έγκαιρης και προληπτικής αναγνώρισης προτύπων που δεν θα εξαρτάται από τρισδιάστατη πληροφορία ή γεωμετρική ακρίβεια. Πρέπει λοιπόν να αναπτύξει ένα υπολογιστικό μοντέλο που να είναι ακριβές, γρήγορο και απλό στην υλοποίηση. Επίσης πρέπει να μπορεί να είναι εφικτή η εκμάθηση νέων προσώπων χωρίς να χρειάζεται επίβλεψη αλλά και η αναγνώριση αυτών αργότερα. Η αναγνώριση κάτω από ευρεία γκάμα συνθηκών επιτυγχάνεται μέσα από την εκπαίδευση με πεπερασμένο αριθμό χαρακτηριστικών εικόνων (“en face”, προφίλ και 45° μοίρες).

1.1 Βιομετρική ταυτοποίηση

Η βιομετρική ταυτοποίηση είναι η μέθοδος αυτοματοποιημένης ταυτοποίησης ενός προσώπου με κριτήριο κάποιο φυσικό ή συμπεριφορικό χαρακτηριστικό του προσώπου. Παραδείγματα φυσικών χαρακτηριστικών που χρησιμοποιούνται ευρέως για ταυτοποίηση είναι τα δακτυλικά αποτυπώματα ή το πλήρες αποτύπωμα παλάμης, η ίριδα του ματιού ή επιμέρους χαρακτηριστικά του προσώπου.

Η βιομετρική ταυτοποίηση απαιτεί την σύγκριση ενός καταχωρημένου δείγματος (βιομετρικό προσδιοριστικό) έναντι ενός πρόσφατα αποτυπωμένου βιομετρικού δείγματος (για παράδειγμα, αποτύπωση φωτογραφίας κατά τη διάρκεια εισόδου στο σύστημα). Κατά τη διάρκεια της διαδικασίας ταυτοποίησης, όπως φαίνεται και στο Διάγραμμα 1, ένα δείγμα βιομετρικού γνωρίσματος αποτυπώνεται, επεξεργάζεται από ένα υπολογιστή και αποθηκεύεται για μετέπειτα σύγκριση. Η βιομετρική αναγνώριση μπορεί να χρησιμοποιηθεί στην περίπτωση, όπου το βιομετρικό σύστημα αναγνωρίζει ένα πρόσωπο ανάμεσα σε ένα πληθυσμό αναζητώντας μια ταυτότητα, βασισμένη αποκλειστικά στη βιομετρία.

Ορισμένες φορές η αναγνώριση καλείται αντιστοίχιση από «ένα προς πολλά». «1 προς N» ο χρήστης αναγνωρίζεται ως μέλος μια ομάδας. Επίσης μπορεί να χρησιμοποιηθεί και σε κατάσταση ταυτοποίησης. Όπου η ταυτοποίηση ενός υποτιθέμενου προσώπου γίνεται με τη σύγκριση ενός προεγγεγραμμένου στιγμιότυπου. Αυτό καλείται και αντιστοίχιση «1 προς 1». Συνήθως στα περισσότερα υπολογιστικά συστήματα ή σε δικτυακό περιβάλλον η ταυτοποίηση χρησιμοποιεί την παρακάτω απλή διαδικασία.



Διάγραμμα 1: Σχηματική αναπαράσταση μιας απλής διαδικασίας ταυτοποίησης

1.2 Τύποι βιομετρικών στοιχείων

1.2.1 Δακτυλικά αποτυπώματα

Τα δακτυλικά αποτυπώματα είναι μοναδικά σε κάθε άνθρωπο, διαφέρουν ακόμα και στα μονοζυγωτικά δίδυμα. Η πιο διαδεδομένη εφαρμογή βιομετρικής αναγνώρισης είναι η αναγνώριση δακτυλικού αποτυπώματος για χρήση σε αστυνομικές ταυτότητες και

διαβατήρια. Πιο σύγχρονη, εμπορική εφαρμογή είναι η χρήση τους, σε σταθερούς και φορητούς σταθμούς εργασίας, η οποία λόγω της μαζικής παραγωγής και της πληθώρας κατασκευαστών που δραστηριοποιούνται στο χώρο παρέχεται με χαμηλό κόστος. Με τη χρήση αυτών των συσκευών δεν απαιτείται πλέον η πληκτρολόγηση συνθηματικών κωδικών αντί αυτού απλά ένα άγγιγμα δίνει άμεση πρόσβαση στο σύστημα. Τα συστήματα δακτυλικών αποτυπωμάτων μπορούν να χρησιμοποιηθούν και στην ταυτοποίηση προσώπων. Στις Ηνωμένες πολιτείες Αμερικής το δακτυλικό αποτύπωμα χρησιμοποιείται σε δημόσιες υπηρεσίες για να διασφαλίσουν ότι οι δικαιούχοι δεν απολαμβάνουν παραπλανητικά παροχές υπό άλλο όνομα.

1.2.2 Αναγνώριση Ίριδας

Αυτή η μέθοδος βασίζεται στην ίριδα του ματιού, που είναι το χρωματιστό μέρος που περιβάλλει την κόρη του ματιού. Τα σχέδια που την αποτελούν θεωρούνται μοναδικά. Αυτά τα σχέδια λαμβάνονται από μια εικόνα που συλλαμβάνεται από σύστημα video. Συσκευές ανάγνωσης ίριδας χρησιμοποιούνται σε εφαρμογές ταυτοποίησης προσώπων εδώ και αρκετά χρόνια. Τα συστήματα αναγνώρισης ίριδας διαρκώς γίνονται πιο προσιτά και αυτή η τάση θα διατηρηθεί και στο προσεχές μέλλον.

Η τεχνολογία αυτή μπορεί να χρησιμοποιηθεί τόσο σε ταυτοποίηση προσώπων όσο και κατηγοριοποίηση (σε συστήματα που κάνουν αντιστοίχιση «1 προς πολλά» ή «1 προς 1» σε μια βάση δεδομένων). Τα σύγχρονα συστήματα είναι τόσο μικρά που μπορούν να ενσωματωθούν ακόμα και σε φακούς επαφής ή σε γυαλιά οράσεως. Η τεχνολογία αυτή δεν επίσης δεν είναι παρεμβατική. Δεν απαιτεί φυσική επαφή με το σαρωτή. Και έχει αποδειχθεί ότι λειτουργεί και με άτομα από διαφορετικές εθνικότητες και φυλές.[8]

1.2.3 Ταυτοποίηση υπογραφής

Αυτή η τεχνική χρησιμοποιεί μια δυναμική ανάλυση της υπογραφής για να ταυτοποιήσει το άτομο. Η τεχνολογία βασίζεται στην ταχύτητα, την πίεση και την γωνία που χρησιμοποιεί το άτομο όταν βάζει μια υπογραφή. Μια πρακτική εφαρμογή είναι η χρήση δικτυακών υπογραφών σε εφαρμογές τηλε-εμπορίου όπου η υπογραφή μπορεί να αποτελέσει κριτήριο ταυτοποίησης του προσώπου. Το σύστημα ταυτοποίησης υπογραφής χρησιμοποιείται και από τις τράπεζες που διατηρούν ηλεκτρονικό αρχείο χειρόγραφων δειγμάτων υπογραφών με σκοπό την μελλοντική ταυτοποίηση κάποιου πελάτη- εντολέα.

1.2.4 Αναγνώριση φωνής

Η αναγνώριση φωνής έχει ιστορία 40 χρόνων περίπου. όταν με χρήση αναλογικών φίλτρων υπολογιζόταν ο μέσος όρος στο χρόνο του φωνητικού σήματος προκειμένου να συγκριθεί με κάποιο προεγγεγραμμένο δείγμα. Η αναγνώριση φωνής ή ομιλητή πιο σωστά χρησιμοποιεί γνωρίσματα της φωνής που έχει βρεθεί ότι διαφέρουν από άτομο σε

άτομο. Αυτά τα ακουστικά γνωρίσματα αντικατοπτρίζουν τόσο την ανατομία (μέγεθος και σχήμα στόματος και λαιμού) όσο και επίκτητα συμπεριφερικά χαρακτηριστικά, χροιά ή τρόπο ομιλίας.

Η απόδοση μπορεί να επηρεαστεί από αλλαγές στα συμπεριφερικά χαρακτηριστικά της φωνής και από τη διαφορετική απόδοση της αρχικής εγγραφής και του δείγματος αν έχει για παράδειγμα χρησιμοποιηθεί διαφορετικό μικρόφωνο ή τηλέφωνο αν πρόκειται για απομακρυσμένη ταυτοποίηση. Η αλλαγή φωνής λόγω γήρανσης πρέπει επίσης να προϋπολογιστεί από τα συστήματα αναγνώρισης. Η διαδικασία σύλληψης της φωνής επίσης δεν είναι παρεμβατική. Η τεχνολογία χρειάζεται καλύτερο εξοπλισμό από τα υπάρχοντα μικρόφωνα και συστήματα μετάδοσης φωνής, πιστότερης καταγραφής και αναπαραγωγής ή μεταφοράς του ηχητικού σήματος για να επιτρέψει την αναγνώριση φωνής ακόμα και από απόσταση μέσω κοινών τηλεφώνων (ενσύρματων ή ασύρματων).

1.2.5 Αναγνώριση προσώπου

Η ταυτοποίηση ενός ατόμου σύμφωνα με την εικόνα του προσώπου του μπορεί να γίνει με διάφορες τεχνικές όπως την αποτύπωση του σε ορατό φάσμα, χρησιμοποιώντας μια φθηνή κάμερα ή χρησιμοποιώντας το υπέρυθρο αποτύπωμα της θερμότητας που αποδίδει το πρόσωπο του. Η αναγνώριση προσώπου στο ορατό φάσμα τυπικά μοντελοποιεί καίρια χαρακτηριστικά στο κέντρο της εικόνας ενός προσώπου.

Χρησιμοποιώντας μια ευρεία γκάμα καμερών, το σύστημα ορατού φάσματος εξάγει χαρακτηριστικά από τις αποτυπωμένες εικόνες που δεν αλλάζουν με την πάροδο του χρόνου ενώ αποφεύγουν επιφανειακά, παροδικά χαρακτηριστικά όπως η έκφραση ή τα μαλλιά. Μερικές από τις μεθόδους που εφαρμόζονται για τη μοντελοποίηση εικόνων προσώπου στο ορατό φάσμα είναι:

- ανάλυση πρωτευόντων χαρακτηριστικών,
- ανάλυση τοπικών χαρακτηριστικών,
- νευρωνικά δίκτυα,
- θεωρία ελαστικών γράφων και
- πολυεπίπεδη ανάλυση.

Οι βασικές δυσκολίες της αναγνώρισης στο ορατό φάσμα είναι η επίδραση του μεταβαλλόμενου σε ένταση ή και σε γωνία φωτισμού. Η μέθοδος της αναγνώρισης προσώπων γενικά έχει να ανταπεξέλθει και με το πρόβλημα της πόζας αλλά και της έκφρασης του ατόμου. Και τα δύο αυτά θέματα σχετίζονται με τις αποστάσεις μεταξύ των χαρακτηριστικών του ατόμου και διαφοροποιούν την εικόνα του.

Μερικά συστήματα αναγνώρισης προσώπου απαιτούν στατική εικόνα ή συγκεκριμένη θέση του προσώπου προκειμένου να αποτυπώσουν το πρόσωπο, ενώ άλλα συστήματα

χρησιμοποιούν μια διαδικασία σε πραγματικό χρόνο για να αναγνωρίσουν το κεφάλι ενός προσώπου και απομονώνουν το πρόσωπο αυτόματα. Στα πλεονεκτήματα της αναγνώρισης προσώπου συγκαταλέγεται και η ευκολία με την οποία λαμβάνεται το δείγμα. Δεν είναι παρεμβατική, δεν απαιτεί επαφή, είναι αδιάλειπτη και γενικά εύκολα αποδεκτή από τους υποψήφιους χρήστες της.

Η αναγνώριση προσώπων θα μπορούσε να οριστεί διαφορετικά ως η διαδικασία ταυτοποίησης κάποιου ατόμου με κριτήριο την εικόνα του προσώπου του. Οι προσεγγίσεις αναγνώρισης προσώπων με στατικές εικόνες μπορούν γενικά να χωριστούν σε δύο κατηγορίες:

1. ολιστικές μεθόδους και
2. στηριζόμενες σε επιμέρους χαρακτηριστικά.

Οι ολιστικές μέθοδοι χρησιμοποιούν όλο το πρόσωπο ακατέργαστο σαν δεδομένο εισόδου, ενώ οι υπόλοιπες βασίζονται σε εξαγωγή τοπικών χαρακτηριστικών και χρησιμοποιούν γεωμετρικές και εμφανισιακές ιδιότητες του προσώπου. Στην προσπάθεια να βελτιωθούν οι υπάρχουσες τεχνικές μελετήθηκε η ανθρώπινη λειτουργία της όρασης και της μνήμης αυτό οδήγησε στο συμπέρασμα ότι τα χαρακτηριστικά του προσώπου επεξεργάζονται ολιστικά από τον άνθρωπο.

Πιο αναλυτικά, ο ανθρώπινος εγκέφαλος αντιμετωπίζει τα χαρακτηριστικά του προσώπου ολιστικά: μπορούν τα προσωπικά χαρακτηριστικά (μάτια, μύτη, στόμα, φρύδια) να επεξεργαστούν ανεξάρτητα από το υπόλοιπο πρόσωπο. Τα πρόσωπα μπορούν συνήθως να αναγνωριστούν από πολύ λίγη πληροφορία, έχει αποδειχθεί ότι μόνο ένα χαρακτηριστικό (όπως τα μάτια ή τα φρύδια) μπορεί να επαρκέσει για αναγνώριση αρκετών διάσημων προσώπων. Ωστόσο αν το πάνω μισό από ένα πρόσωπο συνδυαστεί με το κάτω μισό ενός άλλου οι δύο διακριτές προσωπικότητες είναι πολύ δύσκολο να αναγνωριστούν.

Το ολιστικό πλαίσιο φαίνεται να επηρεάζει το πως επεξεργάζονται τα επιμέρους χαρακτηριστικά. Όταν δύο μισά του προσώπου είναι λάθος ευθυγραμμισμένα , προφανώς εμποδίζεται η κανονική ολιστική αντιμετώπιση, οι δύο μισές ταυτότητες αναγνωρίζονται εύκολα. Αυτά τα αποτελέσματα φανερώνουν ότι όταν λαμβάνονται υπόψη ανεξάρτητα , ορισμένες φορές είναι επαρκή για αναγνώριση του προσώπου. Στο πλαίσιο του προσώπου, ωστόσο, η γεωμετρική σχέση μεταξύ ενός χαρακτηριστικού και του υπόλοιπου προσώπου μπορεί να ξεπεράσει την αναγνωριστικότητα, την ικανότητα να χαρακτηρίζει, αυτού του χαρακτηριστικού. Σε μια πρόσφατη εργασία ανακάλυψαν πως κάποιος μπορεί να μάθει να χρησιμοποιεί την ολιστική πληροφορία του προσώπου Και αυτή η ολιστική επεξεργασία συνδέεται με την έκφραση του προσώπου [2].

1.3 Ανάλυση πρωτευόντων χαρακτηριστικών PCA

Η Principal Component Analysis (PCA), είναι πιθανότατα η πιο διαδεδομένη στατιστική τεχνική πολλών μεταβλητών και χρησιμοποιείται σχεδόν από όλους τους μαθηματικούς κλάδους. Είναι επίσης μάλλον η παλαιότερη μέθοδος πολλών μεταβλητών. Στην πραγματικότητα η αρχή της μπορεί να αποδοθεί στον Pearson (1901) αλλά με τη σύγχρονη μορφή της τυποποιήθηκε από τον Hotelling (1933) που επίσης χρησιμοποίησε πρώτος τον όρο *principal component* [4]. Οι Sirovich and Kirby ήταν οι πρώτοι που περιέγραψαν πρόσωπα με τη χρήση της μεθόδου PCA. Το πρώτο πλήρως αυτοματοποιημένο σύστημα αναγνώρισης προσώπων αναπτύχθηκε από τον Kanade (1973) και παραμένει ένα από τα λίγα, που ακολουθούν μια από την κορυφή προς τη βάση μέθοδο κατευθυνόμενη από ένα μοντέλο αναμενόμενων χαρακτηριστικών ιδιοτήτων.

Ο Burt (1988a, b) χρησιμοποιεί μια «έξυπνα σκεφτόμενη» μέθοδο βασισμένη σε ταυτοποίηση προτύπου σε πολλά επίπεδα ανάλυσης. Αυτή η μέθοδος, από το γενικό προς το ειδικό, χρησιμοποιεί έναν εξειδικευμένο υπολογιστή σχεδιασμένο να υπολογίζει πυραμίδες εικόνων πολλών αναλύσεων γρήγορα, και έχει γίνει επίδειξη όπου ταυτοποιούσε ανθρώπους σε σχεδόν πραγματικό χρόνο. Δουλεύει καλά κάτω από περιορισμένες συνθήκες αλλά υποφέρει από τα τυπικά προβλήματα που πλήττουν τις μεθόδους που βασίζονται σε ταυτοποίηση συσχέτισης με βασικότερα το μέγεθος της εικόνας και το θόρυβο [3].

Το 1987, οι Sirovich and Kirby χρησιμοποίησαν την PCA ώστε να επιτύχουν μια μειωμένη αναπαράσταση εικόνων προσώπου. Το έργο των οποίων συνέχισαν οι Turk and Pentland το 1991, χρησιμοποίησαν προβολές PCA σαν διανύσματα προβολών των χαρακτηριστικών για να επιλύσουν το πρόβλημα της αναγνώρισης προσώπου με τη χρήση Ευκλείδειας απόστασης σαν κριτήριο ομοιότητας. Αυτή η μέθοδος ονομάστηκε αργότερα Ιδιοπρόσωπα. Η δημοσίευσή τους αυτή είναι η πρώτη σε παραπομπές στον δικτυακό τόπο face-reg.org με πολύ μεγάλη διαφορά από την δεύτερη δημοσίευση και θεωρείται de facto πρότυπο και σημείο σύγκρισης ως προς την απόδοση στον τομέα της αναγνώρισης προσώπων (Moghaddam et al., 2000). Αυτή η τεχνική βασίζεται στον μετασχηματισμό Karhunen-Loeve, που είναι γνωστός και ως principal component analysis (PCA) ανάλυση πρωτευόντων χαρακτηριστικών.

Η PCA πλεονεκτεί έναντι άλλων μεθόδων λόγω της ταχύτητας και της απλής ικανότητας εκμάθησης [3].

Άλλες μέθοδοι αναγνώρισης ατόμων μέσω προβολών είναι η Fisher Linear Discriminant (FLD) και η Evolutionary Pursuit (EP).

Η Fisher Linear Discriminant αναλύει με διαφορετικό τρόπο τις εικόνες προσώπων, ψάχνει να βρει νέους άξονες για να εκφράσει τα δεδομένα, τέτοιους ώστε η απόσταση από πρόσωπο σε πρόσωπο να είναι μέγιστη. Προσπαθεί να αυξήσει δηλαδή τη διαχωρισιμότητα των κλάσεων ενώ ταυτόχρονα όλες οι εικόνες των κλάσεων να είναι όσο το δυνατόν πιο κοντά μεταξύ τους.

Η κύρια διαφορά της PCA και της FLD είναι ότι η PCA προσπαθεί να βρει μια προβολή που περιγράφει βέλτιστα τα δεδομένα, υπό την έννοια των ελαχίστων τετραγώνων (καλή αναπαράσταση), ενώ η FLD ψάχνει μια προβολή που διαχωρίζει βέλτιστα τα δεδομένα με τρόπο ελαχίστων τετραγώνων (καλή διακριτικότητα). Η FLD απαιτεί πολλές εικόνες ανά κλάση για να βελτιωθεί αλλά αυτός είναι και ο βασικότερος παράγοντας αποτυχίας της. Οδηγείται σε υπερεκπαίδευση με αποτέλεσμα να χάνει την γενικότητα της αναγνωρισιμότητας για κάθε άτομο. Η PCA ευνοείται από το συνολικό αριθμό εικόνων γενικά ενώ η FLD από τον αριθμό των εικόνων ανά κλάση.

Η μέθοδος EP Evolutionary Pursuit προσπαθεί να ξεπεράσει τα προβλήματα των PCA και FLD κάνοντας χρήση μιας δυναμικής στρατηγικής. Προσπαθεί δυναμικά να επιλέγει τους άξονες που περιγράφουν καλύτερα τα δεδομένα αλλά ταυτόχρονα υπολογίζει την διακριτική ικανότητα αλλά και τη γενικότητα της μεθόδου. Η επίλυση αυτού του προβλήματος οδηγεί σε πολυδιάστατους χώρους, γίνεται χρήση ενός γενετικού αλγορίθμου ο οποίος απαιτεί κάποιους αρχικούς άξονες. Αυτούς τους αρχικούς άξονες τους λαμβάνει μέσω της PCA μεθόδου.

Τα αποτελέσματα των μεθόδων είναι άμεσα συγκρίσιμα σε ποσοστό επιτυχίας Παραθέτονται και σε επόμενο πίνακα (**Error! Reference source not found.**)

Περιγραφή σκοπού

Περιγραφή με όρους θεωρίας πληροφορίας,

Θέλουμε να εξάγουμε την σχετική πληροφορία σε ένα πρόσωπο, να την κωδικοποιήσουμε όσο πιο αποτελεσματικά, και τέλος να συγκρίνουμε την κωδικοποιημένη εικόνα προσώπου με μια βάση προσωποπροτύπων κωδικοποιημένων παρόμοια. Να συλλάβουμε με κάποιο τρόπο την διαφορετικότητα ενός συνόλου εικόνων προσώπου, ανεξαρτήτως από οποιαδήποτε κριτική χαρακτηριστικών, ολιστικά, και να χρησιμοποιήσουμε αυτή την πληροφορία για να κωδικοποιήσουμε και να συγκρίνουμε ανεξάρτητες εικόνες προσώπων.

Περιγραφή με όρους μαθηματικών,

Θέλουμε να εντοπίσουμε τις βασικές συνιστώσες της κατανομής των προσώπων, ή τα ιδιοδιανύσματα του πίνακα συνδιακύμανσης του συνόλου των εικόνων προσώπων, διαχειριζόμενοι κάθε εικόνα ως ένα διάνυσμα σε ένα χώρο πολλών διαστάσεων. Κάθε ανεξάρτητο πρόσωπο μπορεί να περιγραφεί ακριβώς ως γραμμικός συνδυασμός των ιδιοπροσώπων. Κάθε πρόσωπο μπορεί επίσης να προσεγγιστεί χρησιμοποιώντας μόνο τα «καλύτερα» ιδιοπρόσωπα, αυτά που έχουν τις μεγαλύτερες ιδιοτιμές, και που λόγω αυτού εμπεριέχουν την περισσότερη διακύμανση μέσα στο σύνολο των εικόνων των προσώπων.[3]

Η κλασική προσέγγιση που βασίζεται σε ιδιοπρόσωπα προσεγγίζει τα διανύσματα προσώπου με (προσωποεικόνες) με διανύσματα λιγότερων διαστάσεων. Η κλάση με το πιο κοντινό διάνυσμα είναι το αποτέλεσμα της διαδικασίας αναγνώρισης, δηλαδή η ταυτότητα του προσώπου. Επιπροσθέτως, χρησιμοποιείται ένα σύστημα απόρριψης για τα άγνωστα πρόσωπα αν το κριτήριο ομοιότητας απέχει περισσότερο από το κατώφλι που έχουμε ορίσει. [1]

Η προσέγγιση έχει πλεονεκτήματα σε σχέση με άλλες μεθόδους αναγνώρισης προσώπου σε επίπεδο ταχύτητας απλότητας μαθησιακής ικανότητας και στην ανοχή σε μικρές ή σταδιακές αλλαγές των προσώπων. Στην εργασία αυτή αναδιατυπώσαμε το πρόβλημα παρεμβάλλοντας την ανίχνευση πόζας ως πρώτο στάδιο και αυτό απέδωσε. Βασική προϋπόθεση για να δουλέψει αυτή η προσέγγιση είναι να υπάρχουν εικόνες του προσώπου σε όλες τις κλάσεις πόζας.

Η υπολογιστική προσέγγιση αυτού του συστήματος στηρίζεται τόσο σε πληροφορίες φυσιολογίας όσο και στη θεωρία πληροφορίας καθώς επίσης και από την πρακτική ανάγκη για ακρίβεια αποτελέσματος σε σχεδόν πραγματικό χρόνο. Αυτή η προσέγγιση διαχειρίζεται το πρόβλημα της αναγνώρισης προσώπου ως ένα θεμελιωδώς δισδιάστατο πρόβλημα αναγνώρισης, προκειμένου να μην κάνει 3D γεωμετρική απεικόνιση, εκμεταλλευόμενη το γεγονός ότι τα πρόσωπα είναι συνήθως σε όρθια θέση και έτσι μπορούν να περιγραφούν με ένα μικρό αριθμό δισδιάστατων χαρακτηριστικών όψεων.

1.4 Στοιχεία σύγκρισης για ταχύτητα

Η διαδικασία της αναγνώρισης που υλοποιήθηκε στην δημοσίευση [3] παίρνει περίπου 400 ms, τρέχει ιδιαίτερα αργά ανεπαρκώς για ανάλυση σε πραγματικό χρόνο σε Lisp on Sun4 χρησιμοποιώντας μια εικόνα μεγέθους 128x128. Με κάποιο εξειδικευμένο εξοπλισμό, η παρούσα έκδοση θα μπορούσε να τρέξει περίπου σε ταχύτητα πραγματικού video (33ms) [3] Η παραπάνω εικόνα δίνει διανύσματα $128 \times 128 = 16384$ pixels η μέθοδος που υλοποιήσαμε στα σημερινά μηχανήματα εκτελεί το σκέλος της αναγνώρισης σε

χρόνο πολύ μικρότερο από το όριο του πραγματικού Video (33ms). Χρειάζεται 4 ms για την εξέταση εικόνων $200 \times 150 = 30000$ pixels και 15ms για μια εικόνα $400 \times 300 = 120000$ pixels. Η ταχύτητα αυτή κάνει εφικτό τον έλεγχο ατόμων από εικόνες που λαμβάνονται σε πραγματικό χρόνο από κάποια κάμερα ασφαλείας αφού αφήνει 20ms τουλάχιστον για την αναγνώριση προσώπου και τυχόν απαιτούμενη προεπεξεργασία. Θα μπορούσε ακόμη να γίνεται έλεγχος πολλαπλών όψεων του ατόμου καθώς αυτό κινείται στο χώρο κάλυψης της κάμερας.

Στη δημοσίευση [9] δίνεται το παρακάτω πινακάκι χρόνων

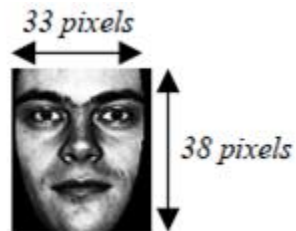
Computation times are shown in Tab. 2

TABLE II. COMPUTATION TIME

Platforms	Processing Time
PC 3GHz	3 ms
DSP 225 MHz	5 ms
NIOS Soft Floating point operators	62 s
NIOS Hard Floating point operators	28 s
NIOS fixed point	15s
IP @50 MHz	60 ms

Πίνακας 1: χρόνοι εκτέλεσης αλγορίθμου δημοσίευσης [9] σε διαφόρων τύπων επεξεργαστές

Για διάφορες μεθόδους υλοποίησης του αλγορίθμου αναγνώρισης που κάνει χρήση της PCA μεθόδου. Τα νούμερα αναφέρονται σε εικόνες 33×38 pixels



Εικόνα 1: Παράδειγμα κανονικοποιημένης εικόνας προσώπου από τη βάση FERET που χρησιμοποιήθηκε στη δημοσίευση [9]

Που έχουν δημιουργηθεί με wavelet συμπίεση των εικόνων. Από τα νούμερα δεν μπορούμε να κάνουμε άμεση σύγκριση για την ταχύτητα η συγκεκριμένη μέθοδος φαίνεται πιο αργή από την μέθοδο που υλοποιεί η εργασία αλλά αυτό δεν μπορεί να πιστοποιηθεί. Εξετάζει εικόνες 7 φορές μικρότερες, σε πλήθος pixels, από τη μικρότερη εφαρμογή του αλγορίθμου της διπλωματικής αλλά ταυτόχρονα εφαρμόζεται σε ένα

σύνολο δεδομένων 4 φορές μεγαλύτερο. Ο χρόνος επεξεργασίας είναι τόσο μικρός που δεν επιτρέπει μεγάλη ακρίβεια στις μετρήσεις και αυστηρά συμπεράσματα.

Από την ίδια μελέτη για την λειτουργία της ανθρώπινης διαδικασίας της αναγνώρισης έχουμε τα παρακάτω στοιχεία για την επεξεργαστική «ισχύ» του ανθρώπινου εγκεφάλου: Η καθυστέρηση απόκρισης σε πρόσωπα στον εγκεφαλικό φλοιό φτάνει τα 120ms, υποδηλώνοντας μεγάλους προκαθορισμένους feedforward υπολογισμούς : Οι άνθρωποι παρατηρητές μπορούν να εκτελέσουν διεργασίες οπτικής αναγνώρισης πολύ γρήγορα. Τα συμπεριφερικά RTs είναι ήδη αρκετά γρήγορα και αντιπροσωπεύουν μια ενδεχομένως υπερεκτίμηση του χρόνου που χρειάζεται για την αναγνώριση λόγω του μηχανισμού που σηματοδοτεί την απάντηση.

Όντως, όταν ένας νευρωνικός δείκτης αναγνώρισης χρησιμοποιείται, σε τόσο σύνθετες διεργασίες όπως το να αποφασισθεί η παρουσία ενός ζώου σε μια σκηνή φύσης φαίνεται να απαιτεί μόλις 50 ms. Πρόσφατα αποδείχθηκε ότι αυτή η διεργασία (ζώο- όχι ζώο) δείχνει αρκετά πολύπλοκη, μπορεί να επιλύεται χρησιμοποιώντας απλοϊκές οπτικές αναπαραστάσεις. Αυτό δείχνει, ότι υπάρχουν νευροφυσιολογικά στοιχεία ειδικευμένα σε αυτό και ότι οι πραγματικά πολύπλοκες διεργασίες όπως η αναγνώριση προσώπων, μπορεί να πραγματοποιηθούν σ' ένα εκπληκτικά σύντομο χρονικό διάστημα. Πιο συγκεκριμένα έχει σημειωθεί ότι υπάρχουν ορισμένοι νευρώνες κύτταρα μέσα στο IT cortex που είναι επιλεκτικά σε πρόσωπα. Επιπροσθέτως, η καθυστέρηση απόκρισης σε αυτά τα κύτταρα κυμαίνεται στην περιοχή των 80–160ms. Πιο πρόσφατα αποτελέσματα δείχνουν ότι η διακριτική ικανότητα μεγάλης ευκρίνειας της αναγνώρισης προσώπου είναι εφικτό να γίνει σε περίπου 50 ms μετά την έκθεση .

Η υπολογιστική σχετικότητα των αποτελεσμάτων είναι ότι η αναγνώριση όπως υλοποιείται ως το επίπεδο του IT cortex πιθανότατα απαιτεί μόνο ένα feedforward υπολογισμό. Οι φτωχές ποιοτικά εικόνες θα χρειάζονται λίγο παραπάνω επαναληπτική επεξεργασία (και γι' αυτό περισσότερο επεξεργαστικό χρόνο), οι σχετικά καθαρές εικόνες μπορούν να εξεταστούν πολύ γρήγορα. Αυτός είναι ένας πολύ σοβαρός περιορισμός για τους αλγορίθμους αναγνώρισης, καθώς καθορίζει ότι επαρκής πληροφορία μπορεί να εξαχθεί άμεσα από την εικόνα χωρίς να χρειάζεται να καθαριστεί αργότερα [2].

1.5 Σκοπός της εργασίας

Η εργασία επικεντρώνεται στην αναγνώριση προσώπου από όλα τα βιομετρικά χαρακτηριστικά. Παρακάτω αναλύεται και περιγράφεται η αναγνώριση με χρήση εικόνων προσώπου. Αφορμή για την εργασία αυτή έδωσε η δημοσίευση των Mathew Turk Alex Pentland EigenFaces for recognition που αποτελεί 20 και πλέον χρόνια μετά

την παρουσίαση της το βασικό σημείο σύγκρισης όλων των νέων μεθόδων. Ο συγκεκριμένος αλγόριθμος επιλέχθηκε λόγω της απλότητας αλλά και της απόδοσης του.

Ο αλγόριθμος αυτός θα μπορούσε να αποτελέσει το δεύτερο στάδιο μετά την ανίχνευση προσώπου, σε κάποιο σύστημα παρακολούθησης video ή στην είσοδο μιας κάμερας ασφαλείας κτηρίου. Σκοπός της εφαρμογής είναι ο έλεγχος της ταυτότητας μιας μικρής προεγγεγραμμένης ομάδας ατόμων. Με βάση τις απαιτήσεις που απορρέουν από τις εφαρμογές είναι σημαντικό να λειτουργεί σε πραγματικό χρόνο ή και ταχύτερα, εδώ τίθεται το άνω όριο των 33ms ανά εικόνα .

Γίνεται προσπάθεια να βελτιωθεί ο αλγόριθμος τροποποιώντας την μέθοδο του υπολογισμού των αποστάσεων, του κριτηρίου ομοιότητας, των εικόνων αλλά και να μελετηθούν οι υπόλοιποι παράμετροι του προβλήματος όπως η ανάλυση των εικόνων και η ποιότητα αυτών ή το πλήθος των διατηρούμενων ιδιοτιμών.

Επαναδιατυπώθηκε το πρόβλημα της αναγνώρισης με στόχο τη μελέτη της αποτελεσματικότητας του αλγορίθμου διαιρώντας το αρχικό πρόβλημα της αναγνώρισης ατόμου σε δυο μικρότερα. Αυτό της αναγνώρισης πόζας και εκ των υστέρων στην αναγνώριση του ατόμου μέσα στο σύνολο των εικόνων της επιλεγμένης πόζας. Η αδυναμία που εμφανίστηκε κατά την εφαρμογή στη βάση δεδομένων μας είναι η έλλειψη εικόνων από τους φακέλους πόζας. Δεν διέθεταν όλα τα άτομα πλήρη κατάλογο εικόνων με αποτέλεσμα το σύστημα να οδηγείτε σε σωστή πόζα και στη συνέχεια να μην υπάρχει το άτομο στην πόζα για να επιλεγεί.

Ακόμη, δοκιμάστηκαν δύο τρόποι για να υπολογισθεί η απόσταση της εξεταζόμενης εικόνας. Απόσταση από το μέσο πρότυπο της κλάσης και η ελάχιστη απόσταση της κάθε κλάσης. Επίσης σε επίπεδο βάσης έγιναν δύο οργανώσεις αρχείων.

Προκειμένου να ελαττωθεί το πλήθος των διαστάσεων χρησιμοποιείται η μέθοδος PCA. Η μέθοδος αυτή εκτελεί γραμμική προβολή της αρχικής εικόνα. Το σύστημα ενεργεί προβάλλοντας τις εικόνες προσώπων σε ένα νέο χώρο χαρακτηριστικών που τονίζει τις σημαντικές διαφορές ανάμεσα σε γνωστές εικόνες προσώπων. Τα σημαντικά χαρακτηριστικά, γνωστά ως ιδιοχαρακτηριστικά δεν ανταποκρίνονται αναγκαστικά σε χαρακτηριστικά του προσώπου όπως διαισθητικά τα αντιλαμβανόμαστε, μάτια, μύτη ή φρύδια. Η μέθοδος παρέχει την δυνατότητα να εκπαιδευτεί το σύστημα και αργότερα να αναγνωρίζει πρόσωπα χωρίς να απαιτείται επίβλεψη. Είναι γρήγορη, σχετικά εύκολη και δουλεύει επαρκώς αποτελεσματικά σε περιορισμένο περιβάλλον, για παράδειγμα σε μια εταιρία ή κτίριο.

1.6 Τομείς πιθανών εφαρμογών

Έλεγχος πρόσβασης

- Είσοδος σε χώρους τραπεζών, εφαρμόζεται ήδη ως απλή καταγραφή των εισερχόμενων προσώπων, δεν γίνεται ταυτοποίηση. Η μέθοδος απαιτεί “en face” εικόνα του προσώπου για να επιτρέψει την είσοδο στο χώρο.
- ATM,θα μπορούσε να αντικαταστήσει τη χρήση κωδικού ή και της κάρτας αναλήψεων
- Αεροδρόμια ,πραγματοποιείται ήδη εγγραφή των προσώπων κατά τη διαδικασία του Check In και του ελέγχου των διαβατηρίων, δεν γνωρίζουμε αν γίνεται ταυτόχρονα αναζήτηση για εγκληματίες σε κάποια βάση δεδομένων της ασφάλειας.



Εικόνα 2: κλειδαριά αναγνώρισης προσώπου

Ψυχαγωγία

- Παιχνίδια υπολογιστών, μπορούν να υλοποιηθούν κλειδώματα που να χρειάζονται την ταυτοποίηση του προσώπου του χρήστη για να επιτραπεί η χρήση του προγράμματος.
- Αλληλεπίδραση ανθρώπου υπολογιστή, αναγνωρίζοντας σωστά άτομα και εκφράσεις μπορεί ο υπολογιστής να ανταποκριθεί σε περισσότερες εντολές.
- Human-Robotics, το ανθρώπινο πρόσωπο ή ταυτότητα και η έκφραση του, θα μπορούσε να αποτελεί την είσοδο πιθανών εντολών για κάποιο ρομποτικό μηχανισμό. Τα ρομπότ που έρχονται σε επαφή με ανθρώπους πρέπει να αναλύσουν και να κατανοήσουν το άτομο ως προς την ταυτότητα αλλά και την έκφραση του προκειμένου να αντιδράσουν όσο το δυνατόν φυσικότερα και αποτελεσματικά.
- Οικογενειακά φωτογραφικά άλμπουμ, θα μπορούσε ο αλγόριθμος αυτός να εφαρμοστεί ως δεύτερο στάδιο μετά από έναν ανιχνευτή προσώπων και να ευρετηριοποιεί τις φωτογραφίες αυτόματα με βάση τα αναγνωριζόμενα πρόσωπα που συμμετέχουν σε αυτές.

Εξυπνες κάρτες

- Διαβατήρια , Διπλώματα οδήγησης ,ή εκλογικά βιβλιάρια με chip που περιέχουν αποθηκευμένες και πρόσφατα ενημερωμένες κάθε φορά φωτογραφίες μπορεί να αποτρέπουν τη χρήση πλαστών εγγράφων με την επιτόπου ταυτοποίηση του ατόμου
- Ασφαλιστικά βιβλιάρια , αυτόματη αναγνώριση του ασφαλισμένου με χρήση κάμερας για αποφυγή κακόβουλης χρήσης παροχών.

Ασφάλεια πληροφοριών

- Γονικός έλεγχος περιεχομένου, αντικατάσταση του κωδικού μέσω του τηλεκοντρόλ που μπορεί να υποκλαπεί με ταυτοποίηση προσώπου.
- Πρόσβαση υπολογιστικού περιβάλλοντος, αντικατάσταση των συνθηματικών εισόδου στο σύστημα με ταυτοποίηση προσώπου, θα μπορούσε να λειτουργήσει τόσο σε επίπεδο εταιρίας μέσω δικτύου είτε σε επίπεδο φορητού σταθμού εργασίας με αποθηκευμένο φωτογραφικό προφίλ του συγκεκριμένου χρήστη.
- πρόσβαση σε προσωπικές συσκευές κινητά σημειωματάρια, κάτι πολύ εύκολο σε υλοποίηση μιας και οι περισσότερες συσκευές διαθέτουν ήδη κάμερες το κόστος των ποίων είναι αμελητέο πλέον.
- πρόσβαση σε ιατρικό υλικό/υπηρεσίες, θα μπορούσε να συμπληρώσει και να αυτοματοποιήσει ένα ολοκληρωμένο πακέτο μηχανοργάνωσης των ταμείων που θα προστάτευε τόσο τα ταμεία όσο και τους ασφαλιζόμενους από την ανεξέλεγκτη χρήση του ιατρικού ιστορικού .
- πρόσβαση στο διαδίκτυο, διαβαθμισμένη πρόσβαση ανάλογα με το επίπεδο που επιτρέπεται στο κάθε αναγνωρισμένο άτομο.

Επιβολή νόμου και παρακολούθηση

- Προηγμένη παρακολούθηση Video , εφαρμογή ανιχνευτή προσώπου μπορεί να τροφοδοτεί σε πραγματικό χρόνο με δεδομένα των αλγόριθμο αναγνώρισης προσώπων

Εφαρμογές πολυμέσων

- διαχείριση πολυμέσων , χρησιμοποιείται και στην αναζήτηση της βάσης προσώπων.
- Αυτόματη προσθήκη ετικετών σε ένα φωτογραφικό Λεύκωμα.

Εμπορικές εφαρμογές

- βιομετρικά συστήματα αναγνώρισης προσώπων, ταυτοποίηση προσώπων για όλους τους παραπάνω σκοπούς-εφαρμογές.
- Home Robotics. Ανάλυση προσώπων και ανθρωπόμορφη διεπαφή συστημάτων με τον άνθρωπο χρήστη.

2 Μεθοδολογία

2.1 Εισαγωγή

Οι εικόνες προσώπων αναπαρίστανται σαν πολυδιάστατοι πίνακες από Pixels.



Εικόνα 3 παράδειγμα εικόνας και μεγέθυνση τμήματος σε επίπεδο ορατού pixel.

Για κάθε συνιστώσα χρώματος διατηρούμε στο αντίστοιχο κελί του πίνακα με δείκτες τις συντεταγμένες του pixel την τιμή της έντασης της κάθε συνιστώσας. Για να περάσουμε σε ένα δυσδιάστατο πίνακα χρησιμοποιούμε συγκεκριμένους τύπους για την συμμετοχή κάθε χρωματικής συνιστώσας, υπολογίζουμε έτσι την εικόνα σε κλίμακα του γκρι η συνάρτηση `rg2gray` του MATLAB χρησιμοποιεί για παράδειγμα τον τύπο:

$$0.2989 * R + 0.5870 * G + 0.1140 * B$$

Το πρόσωπο συχνά ανήκει σε ένα δομημένο υποχώρο χαμηλής διαστασιμότητας. Η αναγνώριση προσώπων και η τεχνητή όραση γενικά, δείχνουν αυξανόμενο ενδιαφέρον σε τεχνικές που κάνουν χρήση αυτής της παρατήρησης, και εφαρμόζουν αλγεβρικές και στατιστικές μεθόδους για την εξαγωγή της χαμηλότερης πολλαπλότητας, σε πλήθος διαστάσεων, του πραγματικού χώρου.

Ζητήματα προεπεξεργασίας όπως ο προσανατολισμός του προσώπου, μασκάρισμα, αντιστάθμιση φωτεινότητας, διατηρούνται κοινά σε όλες τις διαφορετικές υλοποιήσεις της βιβλιογραφίας. Στην παρούσα εργασία οι εικόνες χρησιμοποιήθηκαν όπως υπήρχαν στη βάση του M.I.T. Εφαρμόζοντας κανονικοποίηση ως προς τη θέση και το μέγεθος των προσώπων η μέθοδος έδωσε πολύ χειρότερα αποτελέσματα.

Επίσης έγιναν και αλλαγές στο φωτισμό με δυο τρόπους

1. με επέμβαση στο ιστόγραμμα της εικόνας, τυχαία μετατόπιση (Photoshop)
 2. με προσθήκη τοπικού κατευθυντικού φωτισμού `gradient` και `spot masks` (Lightroom).
- Σκοπός ήταν να μελετηθεί η επίδραση της μεταβολής του φωτισμού στα ιδιοδιανύσματα που περιγράφουν τον προσωποχώρο.

2.2 Χώρος προσώπου και αριθμός διαστάσεων του.

Η υπολογιστική ανάλυση εικόνων προσώπου ασχολείται με οπτικά σήματα (αντανάκλαση φωτός πάνω στην επιφάνεια του προσώπου) που συλλέγονται από έναν αισθητήρα με τη μορφή ενός πίνακα τιμών. Οι τιμές του πίνακα μπορεί να εμπεριέχουν πληροφορία χρώματος ή μόνο την τιμή της έντασης. Μετά από κατάλληλη προεπεξεργασία, κανονικοποίηση και προσαρμογή μεγέθους σε ένα προκαθορισμένο $M \times N$ σχήμα, ο πίνακας με τα pixels μπορεί να αναπαρασταθεί σαν ένα διάνυσμα σε έναν $(M \times N)$ -διάστατο χώρο εικόνας απλά γράφοντας τις τιμές των pixels με προκαθορισμένη σειρά, raster. Τυπική διαδικασία κανονικοποίησης είναι η αφαίρεση της μέσης τιμής φωτεινότητας ενός συνόλου εικόνων από κάθε εικόνα.

Ακόμα και οι χαμηλής ανάλυσης εικόνες δημιουργούν τεραστίων διαστάσεων χώρους χαρακτηριστικών (20.000 διαστάσεις στην περίπτωση μια εικόνας προσώπου μόλις 100×200 pixel). Ωστόσο, η εγγενής διαστασιμότητα του προσωποχώρου είναι πολύ χαμηλότερη από το χώρο εικόνων, αφού τα πρόσωπα είναι όμοια μεταξύ τους σχετικά σε εμφάνιση και παρουσιάζουν σημαντική στατιστική κανονικότητα [1].

Κρίσιμο σημείο στην ανάλυση τέτοιων πολυδιάστατων δεδομένων είναι ο καθορισμός του πλήθους των αναγκαίων διαστάσεων που χρειάζονται για να περιγραφεί κάθε διάνυσμα.

2.3 Χώρος Εικόνας Vs χώρος προσώπου

Προκειμένου να καθορίσουμε μια τυχαία εικόνα στο χώρο εικόνων χρειάζεται να περιγράψουμε την τιμή του κάθε pixel. Κατά συνέπεια η ονομαστική διαστασιμότητα του χώρου υπαγορεύεται από την απεικόνιση του κάθε pixel, και είναι $M \times N$ πολύ υψηλός αριθμός ακόμα και για εικόνες μέτριου μεγέθους. Ωστόσο, μεγάλο μέρος της επιφάνειας του προσώπου είναι λείο και έχει συγκεκριμένη υφή.

Επομένως η ανά pixel δειγματοληψία είναι στην πράξη πλεονάζουσας πυκνότητας. Η τιμή ενός pixel είναι τυπικά υψηλά εξαρτώμενη από τις τιμές των γειτονικών pixels. Επιπροσθέτως, η εμφάνιση των προσώπων είναι ιδιαίτερα περιορισμένη, για παράδειγμα οποιαδήποτε “en face” όψη προσώπου είναι κατά προσέγγιση σχεδόν συμμετρική, έχει δύο μάτια στις πλευρές, μύτη στη μέση κλπ.

Η πλειοψηφία των σημείων στο χώρο εικόνας δεν αντιπροσωπεύει πιθανά φυσικά πρόσωπα. Κατά συνέπεια οι φυσικοί περιορισμοί υπαγορεύουν ότι οι εικόνες προσώπου μπορούν να περιοριστούν σε έναν υποσύνολο του χώρου που καλείται προσωποχώρος.

2.4 Πρωτεύοντας χώρος και βασικές συναρτήσεις

Είναι σύνηθες να μοντελοποιείται ο προσωποχώρος σαν ένας, πιθανώς αποσυνδεδεμένος, πρωτεύοντας χώρος, ενσωματωμένος σε έναν πολυδιάστατο χώρο. Η εγγενής διαστασιμότητα του καθορίζεται από τους βαθμούς ελευθερίας μέσα στο χώρο των προσώπων. Σκοπός της ανάλυσης είναι να επιλεγεί αυτός ο αριθμός και να εξαχθούν οι κυριότερες συνιστώσες του πρωτεύοντα χώρου. Οι κυριότερες συνιστώσες υπολογίζονται σε συνάρτηση με την τιμή pixel και αναφέρονται σαν συναρτήσεις βάσης του πρωτεύοντα χώρου .

Για να γίνουν τα παραπάνω περισσότερο κατανοητά, Έστω ότι μια ευθεία γραμμή στο R^3 , που διέρχεται από την αρχή των αξόνων και παράλληλη προς το διάνυσμα $a=[a_1;a_2;a_3]^T$. Οποιοδήποτε σημείο της γραμμής περιγράφεται από 3 συντεταγμένες. Ωστόσο, ο υποχώρος που αποτελείται από όλα τα σημεία πάνω στη γραμμή έχει ένα βαθμό ελευθερίας με βασική συνιστώσα που σχετίζεται με τη μετατόπιση κατά την διεύθυνση του a , για να αναπαραστήσει τα σημεία στον υποχώρο απαιτεί μια απλή βασική συνάρτηση:

$$\Phi(x_1, x_2, x_3) = \sum_{j=1}^3 a_j x_j$$

Ο παραλληλισμός εδώ βρίσκεται ανάμεσα στη γραμμή και τον προσωποχώρο και μεταξύ του R^3 και του χώρου εικόνων.

2.5 Ανάλυση κύριων συνιστωσών

Η Principal Components Analysis (PCA), είναι μία μέθοδος αναγνώρισης προτύπων στα δεδομένα με τέτοιο τρόπο ώστε να ενισχύονται οι ομοιότητες και οι διαφορές τους. Επειδή τα πρότυπα των δεδομένων είναι δύσκολο να ανιχνευθούν σε χώρους πολλών διαστάσεων, όπου η δυνατότητα της γραφικής αναπαράστασης είναι αδύνατη, η μέθοδος PCA είναι ισχυρό εργαλείο ανάλυσης δεδομένων. Το άλλο κύριο πλεονέκτημα της PCA είναι ότι αφού βρεθούν τα πρότυπα που κρύβονται στα δεδομένα, μπορούν να συμπιεστούν ελαττώνοντας τις διαστάσεις, χωρίς σημαντική απώλεια πληροφορίας. Αυτή η τεχνική χρησιμοποιείται και για τη συμπίεση εικόνων. Η PCA είναι μια χρήσιμη στατιστική μέθοδος που εφαρμόζεται και σε άλλους επιστημονικούς κλάδους εκτός της αναγνώρισης προσώπων όπως η συμπίεση εικόνας και είναι μια μέθοδος ανίχνευσης προτύπων σε πολυδιάστατα δεδομένα.

2.5.1 Εξαγωγή χαρακτηριστικών PCA

Η ανάλυση πρωτευόντων χαρακτηριστικών (PCA) είναι μια τεχνική ελάττωσης της διαστασιμότητας που βασίζεται στην εξαγωγή επαρκούς αριθμού κύριων

χαρακτηριστικών των πολυδιάστατων δεδομένων. Σκοπός της PCA είναι να περιορίσει τον αριθμό των διαστάσεων του χώρου δεδομένων (παρατηρούμενων μεταβλητών) σε ένα λιγότερων διαστάσεων εγγενές υποχώρο χαρακτηριστικών (ανεξάρτητων μεταβλητών), και να περιγράψει τα δεδομένα οικονομικότερα. Αυτή είναι η περίπτωση όπου υπάρχει ισχυρή συσχέτιση μεταξύ των παρατηρούμενων μεταβλητών. Το πρώτο πρωτεύων χαρακτηριστικό είναι γραμμικός συνδυασμός των αρχικών διαστάσεων που έχουν την μεγαλύτερη διακύμανση ; το n -οστό πρωτεύων χαρακτηριστικό είναι ο γραμμικός συνδυασμός με την μεγαλύτερη διακύμανση , υπό τον όρο να είναι ορθογώνιο προς τα προηγούμενα $n - 1$ πρωτεύοντα χαρακτηριστικά.

Με αυτή την έννοια, τα χαρακτηριστικά που εξάγονται από την PCA είναι τα πιο εκφραστικά. Ωστόσο αυτό δεν σημαίνει απαραίτητα ότι είναι και τα πιο «διαχωρίσιμα χαρακτηριστικά» [9].

Στόχοι της PCA

- (a) Εξαγωγή των πιο σημαντικών δεδομένων από τον πίνακα δεδομένων,
- (b) Περιορισμό του μεγέθους των δεδομένων διατηρώντας μόνο την πιο σημαντική πληροφορία.
- (c) Απλοποίηση της περιγραφής του συνόλου των δεδομένων
- (d) Ανάλυση της δομής των μεταβλητών

Η μέθοδος pca υπολογίζει νέες μεταβλητές που ονομάζονται *principal components* που λαμβάνονται ως γραμμικός συνδυασμός των αρχικών μεταβλητών. Το πρώτο principal component απαιτείται να έχει την μεγαλύτερη δυνατή διακύμανση (το δεύτερο υπολογίζεται με τον περιορισμό να είναι ορθογώνιο στο πρώτο και να έχει την μεγαλύτερη δυνατή αδράνεια [4].

2.5.2 Σημασία των διαφορετικών ιδιοτιμών

Συγκεκριμένα στην περίπτωση της αναγνώρισης προσώπων που βασίζεται στην PCA ανάλυση, μερικοί συγγραφείς ισχυρίζονται ότι η κανονικοποίηση του φωτισμού μπορεί να επιτευχθεί με την απόρριψη των τριών πρώτων πρωτευόντων χαρακτηριστικών PCA components [7]. Η παραπάνω μεθοδολογία συμφωνεί και με την παρατήρηση των (Valentin,abdi Edelman & O'toole, in press), ότι διαφορετικά ιδιοχαρακτηριστικά συλλαμβάνουν διαφορετικό είδος πληροφορίας. Με τις μεγάλες ιδιοτιμές να περιέχουν πληροφορία σχετική με τον προσανατολισμό. Επιπρόσθετα στην ανάθεση κατηγοριών (e.g. φύλο, φυλή) των προσώπων. Αυτά τα ιδιοχαρακτηριστικά είναι δυνατά και μπορούν να υπολογιστούν από ένα μικρό σύνολο προσώπων. Σε αντίθεση με τα ιδιοχαρακτηριστικά μικρών ιδιοτιμών που περιέχουν πληροφορία αναγνώρισης ατόμου.

Επιλέχθηκαν 16 φωτογραφίες “en face” και έγιναν test που απέδειξαν ότι οι διαφορές φωτισμού συγκεντρώνονται κυρίως στα πρώτα πέντε ιδιοδιανύσματα. Αλλοιώσεις ακόμα και σε μια εικόνα του συνόλου, ως προς το φωτισμό, άλλαζαν τελείως τα πρώτα ιδιοδιανύσματα. Μελετήθηκαν τρία διαφορετικά τροποποιημένα σύνολα εικόνων εκπαίδευσης και συγκρίθηκαν με το αρχικό ανεπηρέαστο.

1. Με επέμβαση στο ιστόγραμμα της εικόνας (Photoshop)
2. Με χρήση gradient φίλτρου για ομαλή μετάβαση φωτισμού (Lightroom)
3. Με χρήση τοπικού φωτισμού όπως θα προέκυπτε από κάποια λάμπα (Lightroom)

Και οι 3 μέθοδοι έδωσαν διαφορετικά μεταξύ τους αλλά και διαφορετικά με το αρχικό σύνολο εικόνων ιδιοδιανύσματα, με τη μέγιστη επίδραση να εμφανίζεται στις πρώτες πέντε ιδιοτιμές.



Εικόνα 4 Αριστερά αρχικές εικόνες και ιδιοδιανύσματα που διατηρούν το 99,9% της αρχικής ενέργειας. Δεξιά τροποποιημένο ως προς το ιστόγραμμα data set και ιδιοδιανύσματα που περιέχουν το ίδιο ποσοστό ενέργειας.

Αλγόριθμος

Ένα σημαντικό και κυρίως άλυτο πρόβλημα στην μείωση διαστάσεων είναι η επιλογή της k -οστής εσωτερικής διάστασης του πρωτεύοντα χώρου. Καμία αναλυτική παραγωγή αυτού του αριθμού για ένα σύνθετο οπτικό σήμα δεν είναι διαθέσιμη μέχρι σήμερα. Προκειμένου να απλοποιήσουμε το πρόβλημα, είναι σύνηθες να υποθέτουμε ότι στο προς εξέταση θορυβώδες ενσωματωμένο σήμα (στην περίπτωση μας, ένα σημείο επιλεγμένο από το χώρο προσώπου) σε ένα πολυδιάστατο χώρο, ο λόγος σήματος προς θόρυβο είναι υψηλός. Στατιστικά, αυτό σημαίνει ότι η απόκλιση των δεδομένων από τις πρωτεύουσες συνιστώσες του χώρου είναι υψηλή σε σχέση με την διακύμανση μέσα στο συμπληρωματικό χώρο.

Αυτή η υπόθεση αφορά το ιδιοφάσμα του πίνακα συνδιακύμανσης των δεδομένων. Η i -οστή ιδιοτιμή είναι ίση με την διακύμανση κατά μήκος του i -οστού πρωτεύοντος χαρακτηριστικού, κατά συνέπεια ένας λογικός αλγόριθμος για να ανιχνεύσουμε το K είναι να ψάξουμε για την τοποθεσία στο ελλατούμενο ιδιοφάσμα όπου η τιμή των λ_i πέφτει σημαντικά.

Αφού τα βασικά διανύσματα που υπολογίζονται με την PCA είχαν την ίδια διάσταση με τις αρχικές εικόνες προσώπων ονομάστηκαν ιδιοπρόσωπα "Eigenfaces".

Η PCA είναι μια προσέγγιση, της θεωρίας πληροφορίας, της κωδικοποίησης και αποκωδικοποίησης εικόνων προσώπων και μπορεί να δώσει διορατικότητα στο πληροφοριακό περιεχόμενο των εικόνων προσώπου, δίνοντας έμφαση στα τοπικά αλλά και ολικά σημαντικά χαρακτηριστικά.

Αυτά τα χαρακτηριστικά μπορεί να είναι, ίσως όμως και όχι, χαρακτηριστικά με την συμβατική έννοια όπως μάτια, μύτη χείλη ή μαλλιά.

Μια απλή προσέγγιση για να εξάγουμε την πληροφορία που περιέχεται σε μια εικόνα προσώπου είναι κάπως να συλλέξουμε την απόκλιση που περιέχει μια εικόνα σε σχέση με μια συλλογή εικόνων, ανεξάρτητη από οποιοδήποτε χαρακτηριστικό σύγκρισης, και να χρησιμοποιήσουμε αυτή την πληροφορία για να κωδικοποιήσουμε αλλά και να συγκρίνουμε με ανεξάρτητες εικόνες προσώπων.

Τα ιδιοδιανύσματα μπορούμε να τα θεωρήσουμε ως ομάδες χαρακτηριστικών που όλα μαζί χαρακτηρίζουν την διακύμανση μεταξύ των εικόνων προσώπου. Κάθε τοποθεσία της εικόνας συνεισφέρει περισσότερο ή λιγότερο σε κάθε ιδιοδιάνυσμα, έτσι ώστε να μπορούμε να προβάλλουμε τα ιδιοδιανύσματα σαν μία σκιά αποτύπωση προσώπου που αποκαλείται **ιδιοπρόσωπο *eigenface***.

Κάθε ανεξάρτητο πρόσωπο μπορούμε να το αναπαραστήσουμε ακριβώς ως γραμμικό συνδυασμό ιδιοπροσώπων. Κάθε πρόσωπο μπορεί να προσεγγιστεί χρησιμοποιώντας μόνο τα «καλύτερα» ιδιοπρόσωπα, αυτά που έχουν τις μεγαλύτερες ιδιοτιμές και που γι' αυτό συνεισφέρουν την περισσότερη διακύμανση μέσα στο σύνολο των εικόνων προσώπων. Τα καλύτερα ιδιοπρόσωπα αποτελούν ένα M -διάστατο υποχώρο που αποκαλείται χώρος προσώπων "face space" και αποτελείται από όλες τις πιθανές εικόνες προσώπων.

Αυτή η μέθοδος αναγνώρισης προσώπου απαιτεί κάποιες εργασίες αρχικοποίησης:

1. Λήψη ενός συνόλου αρχικών εικόνων προσώπου (σύνολο εκπαίδευσης).
2. Υπολογισμός των ιδιοπροσώπων από το εκπαιδευτικό σύνολο, κρατώντας τις M εικόνες που αντιστοιχούν στις μεγαλύτερες ιδιοτιμές. Αυτές οι M εικόνες ορίζουν τον προσωποχώρο. Καθώς νέα πρόσωπα προκύπτουν, τα ιδιοπρόσωπα μπορούν να προσαρμοστούν ή να επαναυπολογιστούν.
3. Υπολογισμός της αντίστοιχης κατανομής σε M -διάστατο χώρο για κάθε γνωστό άτομο προβάλλοντας το πρόσωπο του στο χώρο των προσώπων "face space".

Αφού Αρχικοποιηθεί το σύστημα, ακολουθούνται τα παρακάτω βήματα για να αναγνωριστούν οι νέες εικόνες προσώπων:

1. Υπολογισμός ενός συνόλου βαρών βασισμένων στην εικόνα εισόδου και στα M ιδιοπρόσωπα προβάλλοντας την εικόνα εισόδου μέσα σε κάθε Ιδιοπρόσωπο.
2. Απόφαση αν η εικόνα περιγράφει πρόσωπο με έλεγχο του πόσο όμοια είναι στο χώρο των προσώπων "face space". Ευκλείδεια απόσταση από το μέσο πρόσωπο.
3. Αν πρόκειται για πρόσωπο, καθορίζουμε το αν πρόκειται για γνωστό ή άγνωστο πρόσωπο, ως προς τη υπάρχουσα βάση.
4. (Προαιρετικά) Ενημερώνονται τα ιδιοπρόσωπα και τα αντίστοιχα βάρη προτύπων.
5. (Προαιρετικά) Αν το ίδιο άγνωστο πρόσωπο εμφανισθεί αρκετές φορές, υπολογίζονται τα βάρη των χαρακτηριστικών προτύπων και το ενσωματώνονται στα γνωστά πρόσωπα.

Μπορεί να αποδειχθεί ότι στην περίπτωση ενός PCA χώρου η απόσταση Mahalanobis είναι ισοδύναμη με την ευκλείδεια απόσταση, για τις συνηθισμένες υλοποιήσεις τα κριτήρια ομοιότητας Ευκλείδεια απόστασης και τα συνημητονοειδή δίνουν συγκρίσιμα αποτελέσματα. Από τα κριτήρια ομοιότητας πάλι, εκείνες που δεν χρησιμοποιούν τεχνικές whitening υπερτερούν ως προς την απόδοση των αντίστοιχων κριτηρίων με χρήση whitening [1].

STANDARD EIGENSPACE (YALE DATABASE)

projection method	images per class	axes	Euclidean	cos()	SOM	FFC	whitening Euclidean	whitening cos()	whitening SOM	whitening FFC
PCA	6	49	95,7 2,7	95,8 2,7	94,2 2,8	81,8 5,4	83,3 5,9	89,3 4,1	88,8 3,8	81,8 5,4
FLD		14	94,6 2,1	95,2 2,5	93,5 2,4	85,9 5,2	97,2 2,2	97,0 2,5	96,7 3,5	90,8 5,5
EP		21	89,8 4,1	94,3 4,0	92,7 4,2	85,2 3,8	-	-	-	-
PCA	5	42	94,0 2,5	94,1 2,5	92,5 3,1	76,8 10,4	82,2 7,2	87,7 5,6	87,3 5,8	76,8 10,4
FLD		14	94,0 3,2	94,3 2,6	92,8 2,8	87,3 5,8	95,0 3,9	94,2 4,7	93,7 4,5	87,7 6,1
EP		16	93,7 3,2	93,9 2,6	92,1 3,3	88,2 3,6	-	-	-	-
PCA	4	35	93,4 1,9	93,4 2,1	91,6 2,3	78,7 5,5	85,4 4,0	88,0 4,0	79,2 5,3	78,7 5,5
FLD		14	92,9 2,4	93,5 2,4	93,8 2,5	84,7 3,8	94,4 2,1	92,9 3,9	93,1 4,2	85,1 5,7
EP		14	92,3 2,6	92,9 2,5	91,8 2,6	85,3	-	-	-	-
PCA	3	28	91,9 2,5	92,4 2,2	88,5 2,7	78,6 6,8	84,2 4,0	86,0 4,5	84,0 5,6	78,6 6,8
FLD		14	89,8 4,5	90,9 4,4	85,7 4,6	81,6 5,5	93,0 2,2	92,0 2,8	92,1 2,8	83,9 5,0
EP		14	84,2 3,5	91,2 4,4	88,5 4,4	82,1 5,1	-	-	-	-
PCA	2	21	88,9 5,0	88,9 5,0	79,1 6,1	75,5 6,9	83,4 6,3	85,1 4,0	75,2 6,6	75,5 6,9
FLD		14	88,5 3,4	88,1 4,2	77,2 4,2	79,6 6,1	89,9 4,1	88,1 2,8	86,6 3,4	77,2 7,2
EP		14	77,9 4,8	88,4 5,1	78,5 3,9	78,4 5,3	-	-	-	-

Πίνακας 2: Αποτελέσματα αλγορίθμων PCA,FLD,EP στη βάση YALE BASE για διάφορα κριτήρια ομοιότητας και πλήθος εικόνων ανά κλάση.

Από τον πίνακα βλέπουμε ότι η FLD μέθοδος προβολής υπερτερεί ελαφρώς σε απόδοση έναντι της PCA αλλά είναι δυσκολότερη σε υλοποίηση και αρχικοποίηση. Επίσης για την μέθοδο PCA οι απλές τεχνικές υπερτερούν των whitening αντίστοιχων.

STANDARD EIGENSPACE (YALE DATABASE)

projection method	images per class	axes	Euclidean	cos()	SOM	FFC	whitening Euclidean	whitening cos()	whitening SOM	whitening FFC
PCA	6	49	95,7 2,7	95,8 2,7	94,2 2,8	81,8 5,4	83,3 5,9	89,3 4,1	88,8 3,8	81,8 5,4
	5	42	94,0 2,5	94,1 2,5	92,5 3,1	76,8 10,4	82,2 7,2	87,7 5,6	87,3 5,8	76,8 10,4
	4	35	93,4 1,9	93,4 2,1	91,6 2,3	78,7 5,5	85,4 4,0	88,0 4,0	79,2 5,3	78,7 5,5
	3	28	91,9 2,5	92,4 2,2	88,5 2,7	78,6 6,8	84,2 4,0	86,0 4,5	84,0 5,6	78,6 6,8
	2	21	88,9 5,0	88,9 5,0	79,1 6,1	75,5 6,9	83,4 6,3	85,1 4,0	75,2 6,6	75,5 6,9

Πίνακας 2: PCA μέθοδος και απόδοση συναρτήσει πλήθος εικόνων ανά κλάση

Τέλος αναφερόμενοι μόνο στην PCA προβολή που μας απασχόλησε βλέπουμε ότι τα κριτήρια ομοιότητας για βάση δεδομένων μέχρι 5 εικόνες ανά άτομο διαφέρουν στο πρώτο δεκαδικό του ποσοστού επί της εκατό. Διαφορά αμελητέα.

2.5.3 Πιθανή Προεπεξεργασία

- **Προσανατολισμός προσώπου** γίνεται με το κεντράρισμα των ματιών στην ίδια σχετική θέση.
- **Αλλαγή διάστασης παραθύρου** μεταβάλλουμε την αρχική εικόνα προσώπου χρησιμοποιώντας προκαθορισμένες αναλογίες για να πάρουμε εικόνες προσώπου 100×200 κουκίδων.
- **Καθάρισμα εικόνας** αφαιρώντας τα όρια προκειμένου να αποφύγουμε την επεξεργασία στοιχείων/γραμμών που δεν ανήκουν στην εικόνα του προσώπου.
- **Αντιστάθμιση απόκλισης φωτισμού** - Αφαιρούμε μια καμπύλη βέλτιστης προσέγγισης φωτεινότητας από κάθε εικόνα. Αυτή η μέθοδος αυξάνει την ανοχή στις υψηλές σκιές που προκαλούνται
- **Κανονικοποίηση** για να γίνουν όλα τα διανύσματα εικόνων της ίδια ενέργειας. Στην συγκεκριμένη μελέτη έγινε μόνο crop και resize των εικόνων με σκοπό να μελετηθεί όσο το δυνατόν πιο γενικά το πρόβλημα. Έγινε και κανονικοποίηση ως προς το μέγεθος των προσώπων που εξετάστηκε αλλά αυτό είχε αρνητική επίδραση στην απόδοση του συστήματος.

2.5.4 Αντικρουόμενοι στόχοι ανάλυσης

Η μέθοδος PCA έχει να αντιμετωπίσει το πρόβλημα της γενικότητας της αναγνώρισης και της ακρίβειας ταυτόχρονα, σε συγκεκριμένα χρονικά περιθώρια. Για να αυξηθεί η ακρίβεια χρειάζονται περισσότερες εικόνες συνολικά αλλά και ανά πρόσωπο. Ωστόσο όσο πιο πολλές εικόνες συμμετέχουν στην εκπαίδευση τόσο πιο συγκεκριμένο γίνεται το προφίλ της αναμενόμενης εικόνας, με αποτέλεσμα πρόσωπα που διαφέρουν αρκετά από το μέσο πρόσωπο να απορρίπτονται ως άγνωστα ή ακόμη και ως μη πρόσωπα.

2.6 Υπολογισμός Ιδιοπροσώπων.

Οι εικόνες προσώπων είναι αρκετά όμοιες μεταξύ τους κατά κανόνα και δεν καταλαμβάνουν μεγάλο χώρο και έτσι μπορούν να περιγραφούν από ένα χώρο σχετικά λίγων διαστάσεων. Η κεντρική ιδέα της PCA είναι να υπολογίσει τα διανύσματα που περιγράφουν καλύτερα την κατανομή των εικόνων προσώπου μέσα στο συνολικό χώρο εικόνων. Αυτά τα διανύσματα ορίζουν τον υποχώρο των εικόνων προσώπου, τον οποίο θα αποκαλούμε προσωποχώρο "face space". Κάθε διάνυσμα είναι μήκους N^2 , περιγράφει μια $N \times N$ εικόνα, και είναι ένας γραμμικός συνδυασμός των αρχικών εικόνων προσώπου, και επειδή μοιάζουν με πρόσωπα εμφανισιακά, αναφερόμαστε σε αυτά ως ιδιοπρόσωπα "eigenfaces".

Προβολή στον ιδιοχώρο

Τα βάρη που περιγράφουν κάθε πρόσωπο υπολογίζονται με την προβολή της εικόνας του προσώπου μέσα σε κάθε ιδιοεικόνα [3]. Αν υποθεθεί ότι το σύνολο των εικόνων αποτελείται από τις $\gamma_1, \gamma_2, \gamma_3, \gamma_4, \dots, \gamma_M$. Το μέσο πρόσωπο του συνόλου αυτού ορίζεται ως:

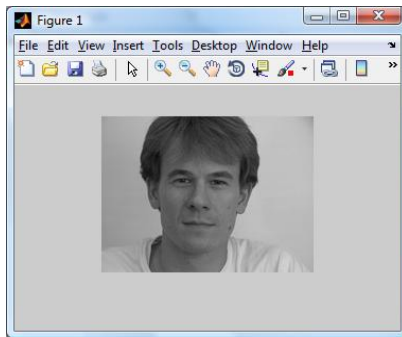
$$\psi = \frac{1}{M} \sum_{n=1}^M \gamma_n \quad (1)$$



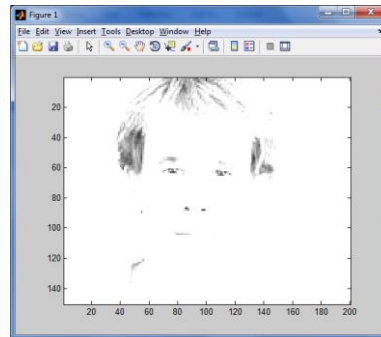
Εικόνα 5: Το μέσο πρόσωπο ψ

Κάθε πρόσωπο i διαφέρει από το μέσο πρόσωπο κατά ένα διάνυσμα

$$\phi_i = \gamma_i - \psi. \quad (2)$$



Εικόνα 6: Παράδειγμα γ



Εικόνα 7: Παράδειγμα $\phi = \gamma - \psi$

Στην εικόνα φαίνεται ένα παράδειγμα του συνόλου εκπαίδευσης, καθώς επίσης και το μέσο πρόσωπο ψ . Αυτό το σύνολο πολύ μεγάλων διανυσμάτων θα αποτελέσει αντικείμενο της PCA, η οποία αναζητά ένα σύνολο M ορθογώνιων διανυσμάτων u_n , που περιγράφουν με τον καλύτερο τρόπο την κατανομή των δεδομένων. Το k -οστο διάνυσμα,

U_n , επιλέγεται έτσι ώστε: $\lambda_k = \frac{1}{M} \sum_{n=1}^M (u_k^T \phi_n)^2 \quad (3)$

γίνεται μέγιστο όταν: $u_l^T u_k = \delta_{lk} = \begin{cases} 1, & \text{για } l = k \\ 0, & \text{αλλιού} \end{cases} \quad (4)$

Τα διανύσματα u_k και οι τιμές λ_k είναι τα ιδιοδιανύσματα και οι ιδιοτιμές αντίστοιχα του πίνακα συνδιακύμανσης:

$$C = \frac{1}{M} \sum \varphi_n \varphi_n^T \quad (5) = A A^T$$

όπου ο πίνακας $A = [\varphi_1 \varphi_2 \varphi_3 \dots \varphi_M]$.

Ο πίνακας C , ωστόσο, είναι $N \times N$, και καθορίζει τα N^2 ιδιοδιανύσματα και ιδιοτιμές και είναι δύσκολο να υπολογιστεί ακόμα και για τυπικού μεγέθους εικόνες.

Χρειάζεται ένας υπολογιστικά εφικτός αλγόριθμος για να υπολογιστούν αυτά τα ιδιοδιανύσματα.

Αν τα pixels των εικόνων είναι λιγότερα από τις διαστάσεις του χώρου ($M < N$) θα υπάρχουν μόνο $M-1$, αντί για N^2 , σημαντικά ιδιοδιανύσματα. Τα υπόλοιπα ιδιοδιανύσματα θα έχουν μηδενικές ιδιοτιμές. Μπορεί να επιλυθεί για τις N^2 διαστάσεις επιλύοντας πρώτα για τα ιδιοδιανύσματα ενός $M \times M$ πίνακα π.χ. επίλυση ενός 16×16 πίνακα αντί για έναν $16.384 \times 16.384 = 128^2$, πίνακα και μετά να πάρουμε τους κατάλληλους γραμμικούς συνδυασμούς των εικόνων προσώπων.

Έστω τα ιδιοδιανύσματα v_i του πίνακα $A^T A$ τέτοια ώστε:

$$A^T A v_i = \mu_i v_i \quad (6)$$

Αν πολλαπλασιαστούν και τα δύο μέλη της εξίσωσης (6) από αριστερά με A και προκύπτει:

$$A A^T A v_i = \mu_i A v_i \quad (7)$$

Από την οποία φαίνεται ότι τα $A v_i$ είναι τα ιδιοδιανύσματα του

$$C = A A^T.$$

Ακολουθώντας αυτή την ανάλυση, μπορεί να κατασκευαστεί ένας $M \times M$ πίνακας $L = A^T A$, όπου $L_{mn} = \varphi_m^T \varphi_n$ και να βρούμε τα M ιδιοδιανύσματα, v_i του L . Αυτά τα διανύσματα καθορίζουν το γραμμικό συνδυασμό των M εικόνων εκπαίδευσης από τα ιδιοπρόσωπα u_i

$$u_i = \sum_{k=1}^M v_{ik} \varphi_k \quad i=1, \dots, M \quad (8)$$

Με αυτή την ανάλυση μειώνονται σημαντικά οι υπολογισμοί, από την τάξη μεγέθους των συνολικών pixels των εικόνων (N^2) στην τάξη μεγέθους του αριθμού των προς εκπαίδευση εικόνων (M). Οι σχετιζόμενες ιδιοτιμές επιτρέπουν την ταξινόμηση των

ιδιοδιανυσμάτων σύμφωνα με την συμμετοχή τους, χρησιμότητα, στον χαρακτηρισμό της διαφορετικότητας ανάμεσα στις εικόνες.



Εικόνα 8: Παράδειγμα συνόλου προσώπων εκπαίδευσης



Εικόνα 9: Παράδειγμα ιδιοπροσώπων

2.7 Κατηγοριοποίηση εικόνων προσώπων με χρήση Ιδιοπροσώπων:

Ένα νέο πρόσωπο (γ) μετασχηματίζεται στο αντίστοιχο ιδιοπρόσωπο (προβαλλόμενο στον προσωποχώρο) με μια απλή πράξη:

$$\omega = u_k^T (\gamma - \psi) \quad (9) \text{ για } k=1, \dots, M.$$

Αυτή περιγράφει ένα σύνολο σημείου προς σημείου της εικόνας πολλαπλασιασμών και αθροισμάτων, πράξεων που εκτελούνται ταχύτερα από τον κατώφλι του real time (33ms) για εικόνες μέχρι 400x300 και το υπάρχον hardware.

Τα βάρη σχηματίζουν ένα διάνυσμα $\Omega T = [\omega_1 \ \omega_2 \dots \omega_M]$ που περιγράφει την συνεισφορά κάθε ιδιοπροσώπου στην αναπαράσταση της εικόνας εισόδου, μεταχειρίζεται το ιδιοπρόσωπο σαν μια σύνολο βάσης έκφρασης για τις εικόνες προσώπων. Το διάνυσμα αυτό μπορεί να χρησιμοποιηθεί σε μια τυπική διαδικασία αναγνώρισης προτύπων για να αποφασίσουμε ποια κλάση από τις προκαθορισμένες περιγράφει καλύτερα το πρόσωπο, αν κάποια το προσεγγίζει. Ο απλούστερος τρόπος για να αποφασιστεί κάτι τέτοιο είναι να βρεθεί η κλάση που ελαχιστοποιεί την Ευκλείδεια απόσταση:

$$\varepsilon_k^2 = \|(\Omega - \Omega_k)\|^2 \quad (10)$$

όπου Ω_k είναι το διάνυσμα που περιγράφει την K-οστή κλάση. Οι κλάσεις προσώπων Ω_i υπολογίζονται από το μέσο όρο των αποτελεσμάτων ενός μικρού συνόλου των εικόνων που αναφέρονται στο i άτομο.

Ένα πρόσωπο ανήκει σε μια κλάση K όταν η απόσταση ε_k από αυτή βρίσκεται κάτω από το επιλεγμένο κατώφλι Θ_ε . Αλλιώς το πρόσωπο χαρακτηρίζεται ως «άγνωστο» και θεωρητικά θα μπορούσε να δημιουργήσει μια νέα κλάση ατόμου.

Επειδή δημιουργώντας το διάνυσμα των βαρών είναι ταυτόσημο με το να προβληθεί η αρχική εικόνα προσώπου στον λίγων διαστάσεων χώρο, αρκετές εικόνες θα προβληθούν στον ίδιο διάνυσμα προτύπου.

Η απόσταση ανάμεσα στην εικόνα και στον προσωποχώρο είναι απλά το τετράγωνο της απόστασης της κανονικοποιημένης αρχικής εικόνας $\Phi = \gamma - \psi$ και $\Phi_f = \sum_{i=1}^M \omega_i u_i$ και της προβολής στον προσωποχώρο:

$$\varepsilon^2 = \|\Phi - \Phi_f\|^2 \quad (11)$$

Κατά συνέπεια υπάρχουν τέσσερις περιπτώσεις για μια εικόνα εισόδου και το διάνυσμα προτύπου :

- Κοντά στο χώρο προσώπων και κοντά σε μια κλάση προσώπων,
- Κοντά στο χώρο προσώπων αλλά όχι κοντά σε γνωστή κλάση προσώπων,
- Μακριά από το χώρο προσώπων και κοντά σε κλάση προσώπου και
- Μακριά από το χώρο προσώπου και μακριά από κάθε γνωστή κλάση προσώπου.

Στην πρώτη περίπτωση αναγνωρίστηκε ένα άτομο και ταυτοποιήθηκε.

Στην δεύτερη περίπτωση αναγνωρίστηκε ένα άγνωστο άτομο. Οι δύο τελευταίες περιπτώσεις δείχνουν ότι δεν πρόκειται για κάποιο άτομο στην εικόνα.

2.8 Ανάλυση με εικόνα σε στήλες vs ανάλυση με εικόνα σε γραμμές του πίνακα συνδιακύμανσης

Αν σχηματιστεί ο πίνακας συνδιακύμανσης τοποθετώντας τις εικόνες ως διανύσματα στήλες τότε το αποτέλεσμα της ανάλυσης είναι ο υπολογισμός νέων διανυσμάτων ίδιου μεγέθους με τις αρχικές εικόνες που μοιάζουν με πρόσωπα και ονομάζονται ιδιοπρόσωπα. Τα ιδιοπρόσωπα αποτελούν τα διανύσματα βάσης, με γραμμικό συνδυασμό των οποίων μπορεί να περιγραφθεί κάθε νέα εικόνα εισόδου. Τα διανύσματα αυτά εξαρτώνται πληροφοριακά από την κάθε εικόνα του συνόλου εκπαίδευσης ξεχωριστά και κάθε σημείο τους εξαρτάται ισχυρά από τα γειτονικά σημεία της αρχικής εικόνας.

Αν ο πίνακας συνδιακύμανσης σχηματιστεί με τις εικόνες σε γραμμές αντί για στήλες και ακολουθηθεί η ίδια ανάλυση τότε υπάρχει αποδέσμευση της πληροφορίας του κάθε σημείου από τα γειτονικά του, μέσα στην εικόνα και συσχετίζεται με το ίδιο σημείο των υπόλοιπων εικόνων του συνόλου. Αυτό οδηγεί σε χωρική ανεξαρτησία των παραγόμενων δεδομένων που αναμένουμε να σχετίζονται με τοπικά χαρακτηριστικά που περιγράφουν το σύνολο των εικόνων και όχι κάθε εικόνα ξεχωριστά. Τελικά οι εικόνες που μελλοντικά μεταφέρονται στο νέο ελαττωμένο χώρο εκφράζονται ως γραμμικός συνδυασμός χαρακτηριστικών ανεξάρτητων μεταξύ τους.

Αυτή η μέθοδος ανάλυσης χρησιμοποιείται κυρίως στην Independent component Analysis ICA με σκοπό να υπολογιστεί ένας παραγοντικός κώδικας, που περιγράφει τα δεδομένα χωρίς χωρικές συσχετίσεις.

2.9 Περίληψη της αναγνώρισης με Ιδιοπρόσωπα.

Συνοψίζοντας την προσέγγιση της αναγνώρισης προσώπων με ιδιοπρόσωπα με τα παρακάτω βήματα:

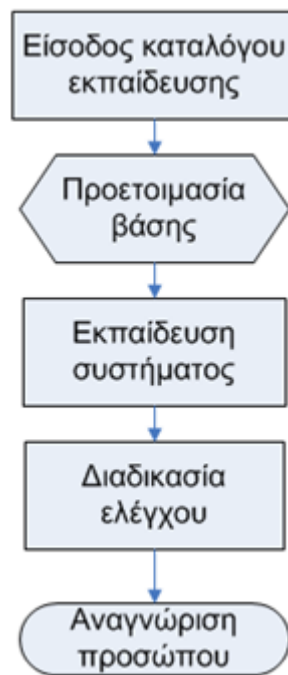
- Συλλογή συνόλου χαρακτηριστικών εικόνων προσώπων των γνωστών ατόμων. Αυτό το σύνολο περιλαμβάνει έναν αριθμό εικόνων για κάθε άτομο, με κάποιες αποκλίσεις σε έκφραση και στον φωτισμό . (θα μπορούσε να είναι από πέντε έξι εικόνες για δέκα άτομα , $M = 59$)
- Υπολογίζεται ο (59×59) πίνακας L , βρίσκονται οι ιδιοτιμές και τα ιδιοδιανύσματα, και επιλέγονται τα M' ιδιοδιανύσματα που σχετίζονται με τις μεγαλύτερες ιδιοτιμές. (Εστω $M' = 10$ στο παράδειγμα).
- Συνδυάζονται οι κανονικοποιημένες εικόνες σύμφωνα με την εξίσωση (6) για να παράγουμε τα $(M' = 10)$ ιδιοπρόσωπα u_k .
- Για κάθε γνωστό άτομο, υπολογίζεται το διάνυσμα κλάσης Ω_k με τη χρήση του μέσου διανύσματος προτύπου Ω [από την εξ. (8)] υπολογίζεται από τις αρχικές εικόνες του ατόμου. επιλέγεται ένα κατώφλι Θ_e που ορίζει την μέγιστη επιτρεπόμενη απόσταση από οποιαδήποτε κλάση και ένα κατώφλι $\varepsilon < \Theta_e$ που ορίζει την μέγιστη επιτρεπόμενη απόσταση από τον προσωποχώρο .
- Για κάθε νέα εικόνα που αναγνωρίζεται , υπολογίζεται το διάνυσμα προτύπου Ω , την απόσταση ε_i από κάθε γνωστή κλάση , και την απόσταση ε από το χώρο των προσώπων. Αν η ελάχιστη απόσταση $\varepsilon_k < \Theta_e$ και $\varepsilon < \Theta_e$, ταξινομεί την εικόνα εισόδου σαν ένα άτομο σχετιζόμενο με το διάνυσμα της κλάσης Ω_k . Αν η ελάχιστη απόσταση $\varepsilon_k = \Theta_e$ αλλά η απόσταση $\varepsilon < \Theta_e$, τότε η εικόνα μπορεί να ταξινομηθεί ως άγνωστο πρόσωπο, και ενδεχομένως να χρησιμοποιηθεί για την έναρξη μια νέας κλάσης ατόμου.
- Αν η νέα εικόνα ταξινομηθεί ως γνωστό άτομο, μπορεί να προστεθεί στο αρχικό σύνολο γνωστών εικόνων προσώπου, και τα ιδιοπρόσωπα μπορεί να επαναυπολογιστούν. Αυτό δίνει την δυνατότητα να διαφοροποιηθεί ο προσωποχώρος καθώς το σύστημα συναντά περισσότερα στιγμιότυπα γνωστών προσώπων και να μένει πάντα επίκαιρος.

3 Υλοποίηση

3.1 Υλοποίηση αναγνώρισης προσώπων με χρήση PCA σε MATLAB

Ολόκληρη η διαδικασία εκπαίδευσης και ελέγχου είναι σειριακή και μπορεί να περιγραφεί με τα παρακάτω βήματα:

1. Προετοιμασία βάσης δεδομένων
2. Εκπαίδευση
3. Έλεγχος



Διάγραμμα 2 : Διάγραμμα ροής διαδικασίας υλοποίησης

3.1.1 Προετοιμασία βάσης δεδομένων

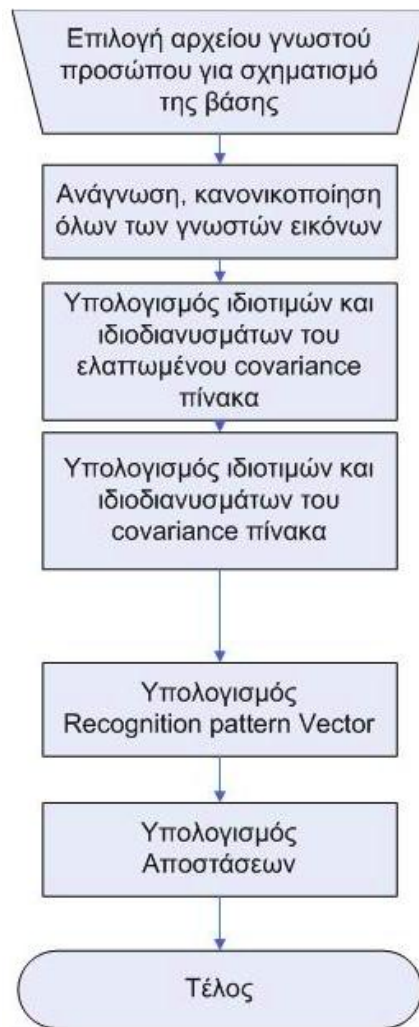
Η βάση εφοδιάστηκε με 4-6 φωτογραφίες ανά άτομο σε διαφορετικές γωνίες θέασης και με διάφορες εκφράσεις προσώπου. Υπάρχουν 10 άτομα στην βάση. Σαν βάση θεωρείται ο κατάλογος person που περιέχει υποφακέλους για κάθε πρόσωπο και περιέχει όλες τις φωτογραφίες του.

Η βάση προετοιμάζεται με όλες τις «γνωστές» φωτογραφίες όλων των ατόμων, 5-6 φωτογραφίες των 10 ατόμων σε διάφορες εκφράσεις και γωνίες θέασης αλλά σε

παρόμοιες συνθήκες φωτισμού και φόντου με χρήση κάμερας υψηλής ανάλυσης. Επιλέχθηκαν τυχαία 8άδες και 6άδες εικόνων που θα αποτελέσουν τα τεστ υποσύνολα και τα αποθηκεύτηκαν στον υποφάκελο test.

3.1.2 Εκπαίδευση

1. Επιλέγεται οποιοδήποτε (.jpg) αρχείο από τον φάκελο της βάσης χρησιμοποιώντας τον dialog box open file.
 2. Χρησιμοποιώντας αυτό διαβάζονται όλες τις φωτογραφίες που ανήκουν στην ίδια ρίζα, όλες οι εικόνες, όψεις των ατόμων.
 3. Κανονικοποιούνται τα πρόσωπα αφαιρώντας το μέσο όρο όλων των εικόνων.
 4. Υπολογίζονται τα σημαντικά ιδιοδιανύσματα του ελαττωμένου πίνακα συνδιακύμανσης.
 5. Υπολογίζονται τα ιδιοδιανύσματα του πίνακα συσχέτισης .
 6. Υπολογίζονται τα πρότυπα διανύσματα για κάθε εικόνα και το μέσο RPV για κάθε άτομο.
 7. Για κάθε άτομο υπολογίζεται η μέγιστη από όλες τις αποστάσεις από όλες τις εικόνες του RPVs από το μέσο RPV του προσώπου αυτού.
- 7β. Για κάθε άτομο υπολογίζεται εναλλακτικά και η ελάχιστη απόσταση κάθε κλάσης από την προς εξέταση εικόνα αντί για την απόσταση από το μέσο όρο της κλάσης.



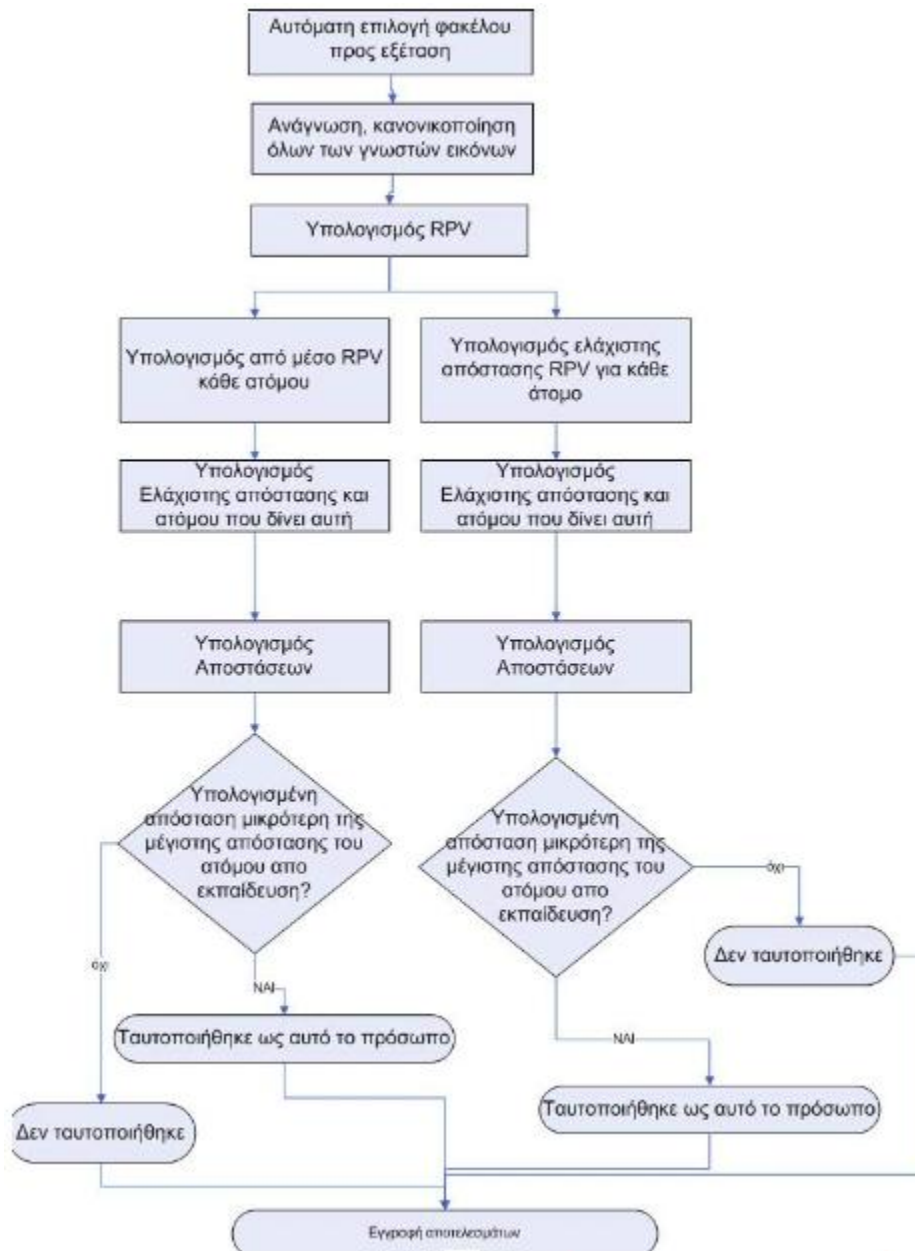
Διάγραμμα 3 :Διάγραμμα ροής Εκπαίδευσης.

3.1.3 Έλεγχος

1. Ο έλεγχος υλοποιείται με τα παρακάτω βήματα:
2. Επιλέγει το σύστημα αυτόματα τον υποφάκελο test της που βρίσκεται ένα επίπεδο ψηλότερα στο path από την εικόνα που διαλέξαμε προς εκπαίδευση και εκτελεί επαναληπτικά τα παρακάτω βήματα για όλες τις εικόνες που περιέχει.
3. Η εικόνα διαβάζεται και κανονικοποιείται.
4. Υπολογίζεται το RPV κάθε εικόνας χρησιμοποιώντας τα ιδιοδιάνυσμα του πίνακα συνδιακύμανσης.
5. Υπολογίζεται η απόσταση του RPV της εικόνας εισόδου
 - a. από το μέσο RPV κάθε ατόμου
 - b. ή εναλλακτικά την ελάχιστη απόσταση από κάθε εικόνα του κάθε ατόμου.

7. Επισημαίνεται το άτομο με το οποίο υπάρχει η ελάχιστη απόσταση. Εναλλακτικά εξετάζεται και η δεύτερη τρίτη και τέταρτη ελάχιστη απόσταση.

8. Αν η ελάχιστη απόσταση είναι μικρότερη από την μέγιστη απόσταση που υπολογίστηκε από το άτομο κατά τη διάρκεια της εκπαίδευσης τότε η εικόνα ταυτοποιείται σαν το συγκεκριμένο άτομο.



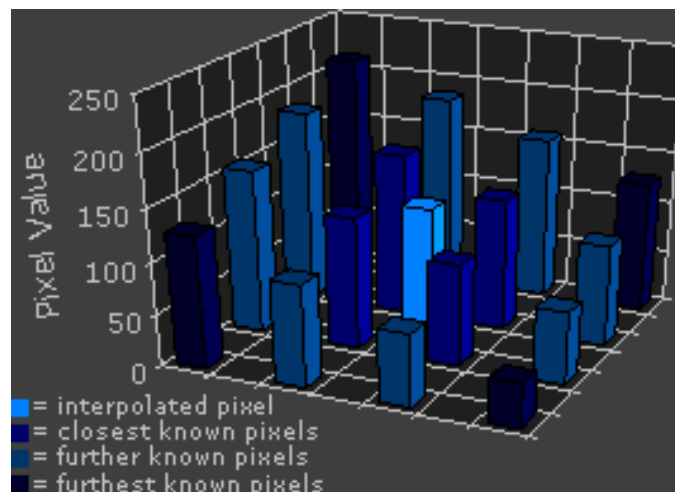
Διάγραμμα 4 : Διάγραμμα ροής Ελέγχου

4 Αποτελέσματα

4.1 Ορισμός τεστ και χειρισμός του συνόλου δεδομένων

Το σύνολο που χρησιμοποιήθηκε ήταν η βάση προσώπου του M.I.T. CBCL FACEREC DATABASE που αποτελείται από δέκα άτομα με περίπου 5 εικόνες, πόζες ανά άτομο, 59 εικόνες συνολικά. Αυτές χωρίστηκαν σε 8άδες ή και 6άδες μέχρι εξαντλήσεως του συνόλου και αποτελούσαν το προς εξέταση δείγμα. Οι υπόλοιπες εικόνες 51 ή 53 αποτελούσαν το σύνολο εκπαίδευσης. Αυτό το μοτίβο επαναλαμβανόταν μέχρι να εξεταστούν όλες οι φωτογραφίες και οι 59 αυτές δοκιμές αποτελούσαν ένα τεστ κάθε φορά. Απόκλιση από την παραπάνω μεθοδολογία υπήρξε μόνο στις περιπτώσεις που χρειαζόταν οπτική παρατήρηση των ιδιοδιανυσμάτων. Σε αυτές τις περιπτώσεις χρησιμοποιήθηκαν ως σύνολο εκπαίδευσης οι “en face” εικόνες του συνόλου και σε αυτές έγινε κανονικοποίηση μεγέθους και θέσης με σκοπό το όσο το δυνατόν πιο καλό οπτικό αποτέλεσμα, δεν έχει νόημα το μέσο πρόσωπο για παράδειγμα όταν αθροίζονται εικόνες που αφορούν διαφορετικές πόζες. Η ίδια τακτική ακολουθήθηκε και στην περίπτωση εξέτασης της διαφοροποίησης του φωτισμού.

Τέλος οι εικόνες έγιναν crop και resize για να κανονικοποιηθούν αλλά και να γίνουν τα συγκριτικά ανάλυσης και ποιότητας με τη χρήση του Photoshop και πιο συγκεκριμένα με χρήση δικυβικού αλγορίθμου interpolation. Στη διαδικασία της σμίκρυνσης δηλαδή λαμβάνονται υπόψη οι 4x4 γείτονες με βάρη αντιστρόφως ανάλογα της απόστασης από το υπολογιζόμενο pixel.



Εικόνα 10: μέθοδος δικυβικής παρεμβολής (bicubic interpolation)

4.2 Αριθμός διατηρούμενων ιδιοδιανυσμάτων

Στη δημοσίευση των Pentland και Turk αναφέρεται πως για ένα δείγμα σαράντα φωτογραφιών, τέσσερις όψεις για δέκα άτομα, χρειάστηκαν μόλις 7 ιδιοδιανύσματα για να υλοποιήσουν την αναγνώριση με επαρκή ακρίβεια.

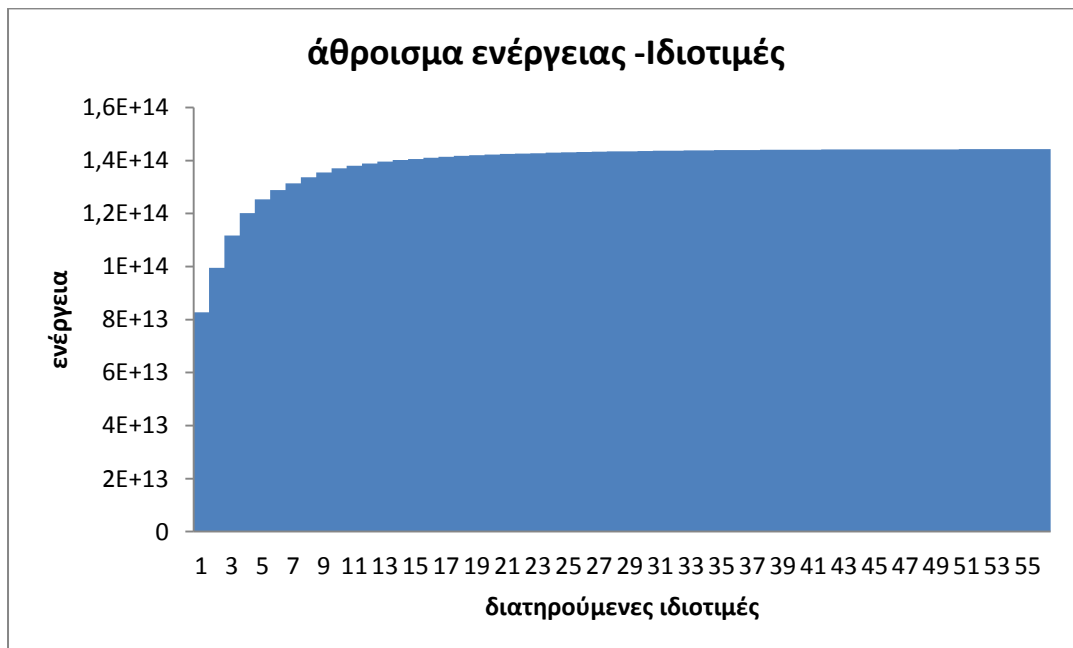
Στην Υλοποίηση αυτή έγινε προσέγγιση του θέματος με χρήση της ενέργειας του σήματος. Κρατήθηκαν τόσα ιδιοδιανύσματα ώστε να διατηρείται ποσοστό μεγαλύτερο του 97,5% της συνολικής ενέργειας. Αυτό οδήγησε σε δυναμική διατήρηση όσων ιδιοδιανυσμάτων χρειαζόμασταν κάθε φορά ανάλογα με το Data Set.

Για να γίνει αυτό αθροίζονται τα τετράγωνα των ιδιοτιμών και υπολογίζεται το επιθυμητό ποσοστό (97,5%) επαναλαμβάνεται το μερικό άθροισμα στον πίνακα φθίνουσας σημασίας, ιδιοτιμής, μέχρι να ξεπεραστεί το κατώφλι.

Από παραδείγματα για μέχρι είκοσι οχτώ φωτογραφίες τρεις όψεις για κάθε άτομο χρειάστηκαν μόλις πέντε ιδιοδιανύσματα για ~97,5% της ολικής ενέργειας. Για πενήντα τέσσερις εικόνες χρειάστηκαν επτά ιδιοδιανύσματα και για τριανταπέντε φωτογραφίες χρειάστηκαν 6 ιδιοδιανύσματα. Κάτι που ήταν αναμενόμενο με βάση και τη βιβλιογραφία.

Τα παρακάτω διαγράμματα παρουσιάζουν την ενέργεια του μερικού αθροίσματος τετραγώνων ιδιοτιμών για διάφορες περιπτώσεις πλήθους εικόνων εκπαίδευσης.

55 εικόνες



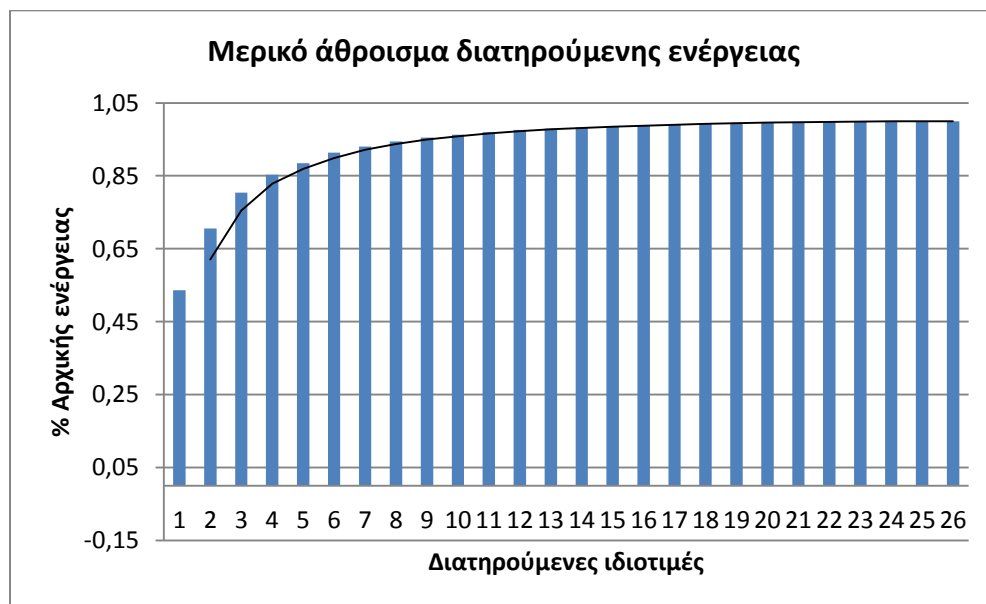
Διάγραμμα 5: Μερικό άθροισμα μέχρι την i-οστή ιδιοτιμή- διατηρούμενη ιδιοτιμή



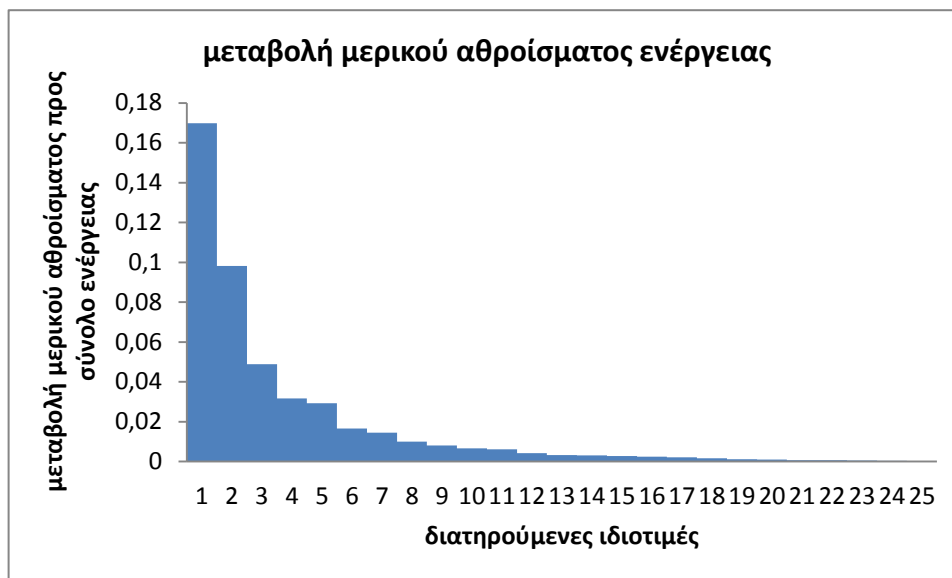
Διάγραμμα 6: Ρυθμός μεταβολής μερικού αθροίσματος συναρτήσει πλήθους ιδιοτιμών.

Ο αλγόριθμος επέλεξε να κρατήσει 12 ιδιοτιμές για 99% της ενέργειας ή 8 τιμές για 97,5% της ενέργειας. Κάτι που επιβεβαιώνεται διαισθητικά και από τα γραφήματα. Ο ρυθμός αλλαγής μετά το 12 σχεδόν μηδενίζεται, αντίθετα το μερικό άθροισμα παραμένει σχεδόν σταθερό.

28 εικόνες εκπαίδευσης

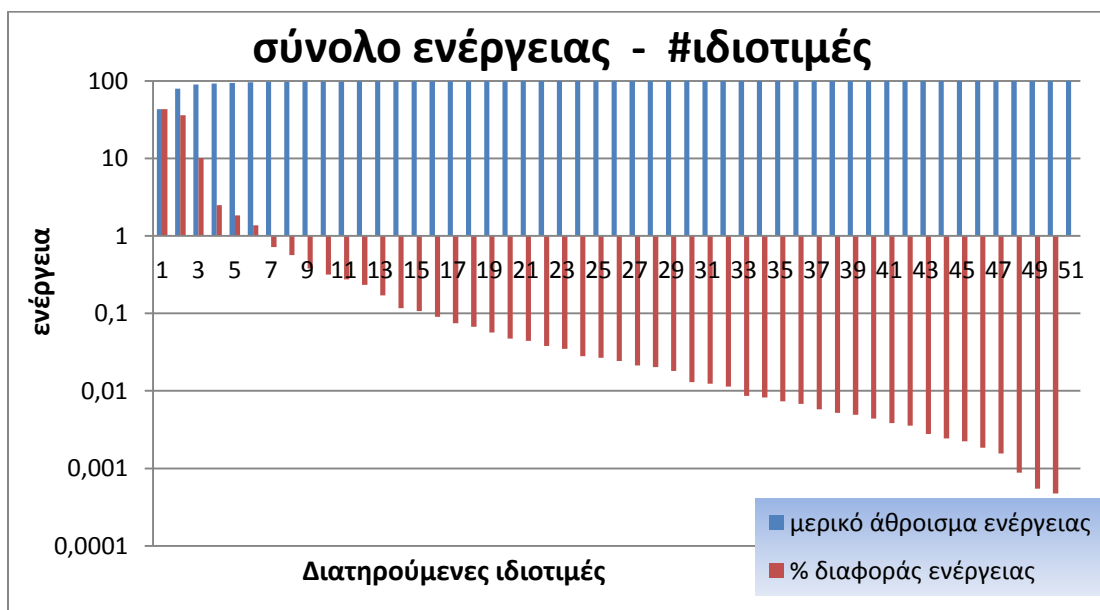


Διάγραμμα 7: Μερικό αθροίσμα ενέργειας συναρτήσει αριθμού ιδιοτιμών που συμμετέχουν σε αυτό



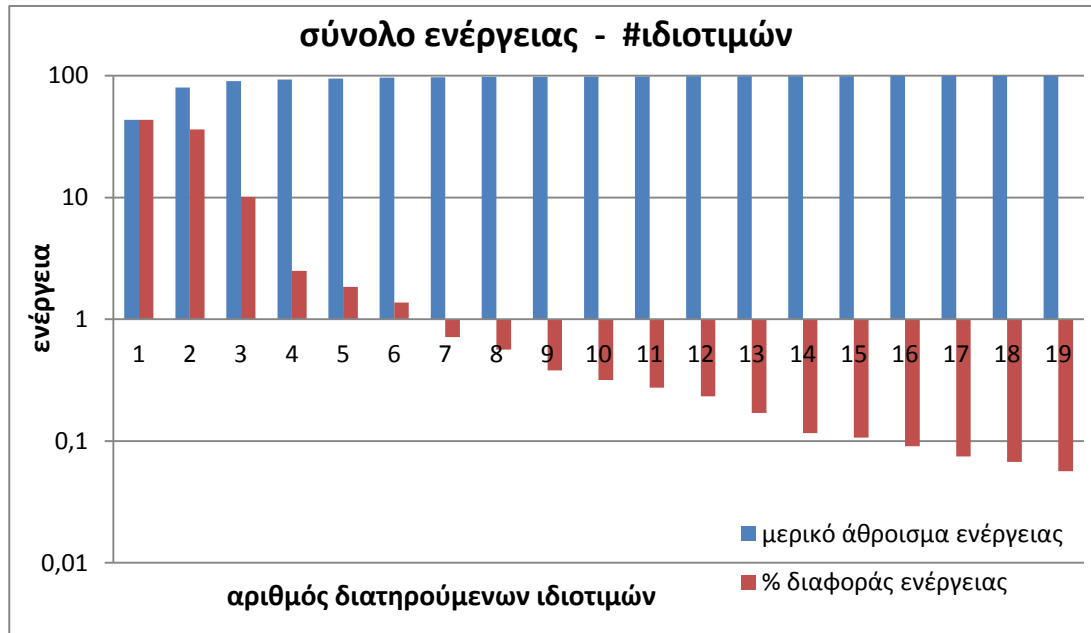
Διάγραμμα 8: Ρυθμός μεταβολής ποσοστού ενέργειας συναρτήσει πλήθους ιδιοτιμών.

Η μέθοδος με το ποσοστό ενέργειας επέλεξε 5 ιδιοτιμές για διατήρηση του 95% της ενέργειας.



Διάγραμμα 9: Σύνολο ενέργειας συναρτήσει αριθμού διατηρούμενων ιδιοτιμών.

Εστιάζοντας στις πρώτες 19 ιδιοτιμές



Διάγραμμα 10: Μερικό άθροισμα, ρυθμός μεταβολής ενέργειας- ιδιοτιμές (εστίαση στις 19 πρώτες ιδιοτιμές)

Γίνεται αντιληπτό ότι με βάση την % διαφορά θα έπρεπε να διατηρηθούν μόλις 6 ιδιοτιμές για το τεστ με τις 51 εικόνες αφού σε αυτή η ενέργεια ξεπερνά το 95%, παράλληλα στην ιδιοτιμή 6 η αύξηση της ενέργειας είναι της τάξεως του 0,5% και πέφτει.

Ωστόσο από τις πειραματικές μετρήσεις φάνηκε πως υπήρχε περαιτέρω βελτίωση του ποσοστού επιτυχίας για ποσοστό 97,5% της συνολικής ενέργειας και πλήθος ιδιοτιμών-ιδιοδιανυσμάτων 13. Περαιτέρω αύξηση των ιδιοδιανυσμάτων που διατηρήθηκαν οδήγησε σε αύξηση της απόδοσης της μεθόδου τόσο σε μεγάλης όσο και σε μικρής ανάλυσης εικόνες μέχρι το ποσοστό του 99,5% που αντιστοιχεί σε περίπου 18 ιδιοτιμές.

Αυτό ήταν και το πραγματικό όριο πέρα από το οποίο δεν υπήρχε αύξηση της απόδοσης, ούτε σε πραγματικές επιτυχίες ούτε και σε εν δυνάμει επιτυχίες¹.

¹ Εν δυνάμει επιτυχίες είναι οι επιτυχίες που βρίσκονται μέσα στην δεύτερη, τρίτη και τέταρτη ελάχιστη απόσταση, πέρα δηλαδή από την απολύτως ελάχιστη απόσταση που το σύστημα υποδεικνύει ως σωστή.

Αναγνώριση με βάση την πόζα Vs αναγνώριση με βάση το άτομο.

Τα τεστ χωρίστηκαν σε δύο μεγάλες ομάδες ανάλογα με τη διαδικασία που ακολουθούσαν για να καταλήξουν σε άτομο.

- A. Σε αυτά που δούλευαν με πρώτο κριτήριο την ανίχνευση της πόζας ως κλάση και μετά στην πόζα αυτή εντόπιζαν το άτομο

Ο αλγόριθμος αυτός δοκιμάστηκε με τη χρήση φακέλων χειροκίνητα διαλεγμένων φωτογραφιών και ταξινομημένων ως προς της πόζα του ατόμου. Το αποτέλεσμα ήταν επιτυχία που άγγιζε το 88% λαμβάνοντας ως σωστά τα αποτελέσματα που το σύστημα επέλεξε τη σωστή πόζα και δεν υπήρχε εικόνα του ατόμου στον συγκεκριμένο φάκελο. Εδώ αξίζει να σχολιαστούν 5-6 περιπτώσεις από τα 51 τεστ σε δυο ανεξάρτητους κύκλους τεστ που εκτελέστηκαν όπου το σύστημα κατέληξε σε εσφαλμένη ή αμφισβητούμενη πόζα στην οποία όμως υπήρχε το άτομο σε αντίθεση με την πόζα που αναμενόταν να επιλέξει και που δεν περιείχε το άτομο. Μένει να μελετηθεί για τις περιπτώσεις αυτές ποιές ιδιοτιμές ήταν αυτές που προκαλούσαν την μετατόπιση της απόφασης στο φάκελο που περιείχε το άτομο.

Στο τεστ αυτό μετρήθηκε η διαφορετικότητα της μέσης πόζας με την προς εξέταση νέα εικόνα. Παραλληλίστηκε δηλαδή η κλάση προσώπου με την κλάση πόζας και η κλάση του μέσου προσώπου ή αλλιώς πιθανού προσώπου, με την κλάση όλων των πιθανών όψεων κάποιου ατόμου. Επίσης υπολογίστηκε και με την μέθοδο της ελάχιστης απόστασης από κάθε εικόνα της κάθε πόζας, η μέθοδος αυτή έδωσε ταυτοτικά αποτελέσματα με την μέθοδο απόστασης από το μέσο της κλάσης.

Το ποσοστό επιτυχίας ήταν για την πόζα 79-86% ανάλογα αν οι ενδιάμεσες πόζες θα ληφθούν ως σωστές ή όχι και για το πρόσωπο περίπου 40% αν λάβουμε υπόψη μας τα σωστά πρόσωπα χωρίς να προσμετρούνται τα αμφισβητούμενα αποτελέσματα λόγω έλλειψης υλικού(πόζας). Τα μη λάθη έφταναν περίπου στα 91% οπότε μένει να οριστεί τελικά η επιτυχία.

Τα πλησιέστερα σε μορφή αποτελέσματα αναφέρονται στη δημοσίευση των Xiaozheng Zhang*, YongshengGao Face recognition across pose: A review [8] όπου μελετήθηκε η επίδραση που είχε στην απόδοση του αλγορίθμου το πλήθος διαφορετικής πόζας και η γωνία της πόζας. Η μέθοδος PCA, για δείγμα περίπου 50 εικόνων με 5 πόζες όπως υπήρχαν και στη βάση αυτής της εργασίας από 30% έως 89% ποσοστό συγκρίσιμο με τα αποτελέσματα που παρουσιάστηκαν παραπάνω.

Επιπροσθέτως οι Pendland και Turk στην αρχική τους δημοσίευση [1] είχαν δώσει τα παρακάτω στοιχεία:

85% correct averaged over orientation variation ,πολύ κοντά στο 86%-91% που πέτυχε η υλοποίηση αυτής της εργασίας όταν εξέτασε την πόζα και μετά το άτομο.

Experiments and performances of face recognition algorithms across pose on USF-3D database, ORL database, Bern University database, XM2VTS database, MIT database, Asian database, and MVU database.

Database	No. of faces	Pose variations	Gallery/probe	Approach	Result (%)
ORL	40	10 random poses within $\pm 20^\circ$ in yaw and tilt	5 random/5 remaining	Eigenfaces [47]	89.5
				SOM+CN [47]	96.2
				PDBNN [53]	96.0
Bern Univ.	30	5 poses: 0° , $\pm 20^\circ$ in yaw and tilt	1 random/5 remaining random	Eigenfaces [47]	61.4
				SOM+CN [47]	70.0
				Eigenfaces [27]	65.12
				LEM [27]	72.09
				DCP [28]	68.61
XM2VTS	125	5 poses: 0° , $\pm 30^\circ$ in yaw and tilt	1 (0°)/4 remaining	Cylindrical 3D pose recovery [26]	80
				Fisherfaces [42]	46
				Expert fusion [42]	70
USF-3D (synthetic images)	100	3 poses: 0° , $\pm 90^\circ$ in yaw	1 (0°)/2 ($\pm 90^\circ$)	TFA [63]	91
				KPCA [50]	43.3
				GDA [50]	36.0
WVU	40	2 poses: 0° , 24° in yaw	1 (0°)/1 (24°)	Correlation filter [50]	79.7
				Linear shape models [76]	100
				Mosaicing [71]	97.84
MIT	62	10 poses within $\pm 40^\circ$ in yaw and $\pm 20^\circ$ in tilt	1 (15°)/9 remaining	Parallel deformation [10]	82.2
Asian	46	5 poses: 0° , $\pm 15^\circ$, $\pm 25^\circ$ in yaw	1 (0°)/4 remaining	Eigenfaces [68]	31.5
				AAM [68]	68

Pose angles are approximate. The citations indicate the papers reporting the results.

Πίνακας 3 αποτελέσματα διαφόρων αλγορίθμων αναγνώρισης προσώπων σε βάσεις δεδομένων διαφορετικού αριθμού πόζας και γωνιών.

Χαρακτηριστικό παράδειγμα αποτυχίας ήταν η κοπέλα του φακέλου 02 (κινεζικής καταγωγής) το προφίλ της οποίας απέχει από όλα τα υπόλοιπα (μικρή μύτη και οδηγούσε το σύστημα σε en face όψη άλλου ατόμου)

Η βασική αδυναμία της μεθόδου αυτής είναι ότι δεν μπορεί να στηριχθεί στις υπάρχουσες πόζες ενός ατόμου προκειμένου να το αναγνωρίσει σε μια τυχαία/ ενδιάμεση των υπάρχόντων, διότι απομονώνει μια όψη, την πλησιέστερη και αναζητά μέσα στο φάκελο της το πλησιέστερο πρόσωπο.

Αντίθετα η ανάλυση με κλάση -τις εικόνες ενός προσώπου- πλεονεκτεί στο ότι καταλήγοντας στην κλάση του ατόμου ο μέσος όρος θα μοιάζει με οποιαδήποτε εικόνα του ατόμου κατά κανόνα.

B. Σε αυτά που ενεργούσαν κατευθείαν με κλάσεις που αφορούσαν άτομα.

Αντίθετα τα τεστ ως προς άτομα έδειξαν πολύ καλύτερη συμπεριφορά ως προς το τελικό αποτέλεσμα την αναγνώριση του προσώπου. Ακόμα και στις περιπτώσεις που η υπό εξέταση εικόνα δεν ήταν μέρος του συνόλου εκπαίδευσης (ενδιάμεσες πόζες για παράδειγμα)

Επίσης χρησιμοποιήθηκαν και δύο μέθοδοι μέτρησης της απόστασης. Αver η μέθοδος που μετράει την απόσταση της εικόνας που εξετάζεται από το μέσο όρο κάθε κλάσης και

ελάχιστης απόστασης που μετράει την ελάχιστη απόσταση κάθε κλάσης από την εξεταζόμενη εικόνα. Οι δύο μέθοδοι φαίνεται αρχικά να αποδίδουν το ίδιο καλά. Με ελάχιστη διαφορά σε επιτυχίες.

Στην προσπάθεια μας να αυξήσουμε την απόδοση του συστήματος, διατηρήθηκαν λίστες με περισσότερες από μία αποστάσεις, πέραν της ελάχιστης. Διατηρώντας λίστες για κάθε πιθανό αποτέλεσμα με την δεύτερη, τρίτη και τέταρτη «ελάχιστη» απόσταση είχαμε αύξηση της απόδοσης. Ορίστηκε ένα νέο μέγεθος εδώ η *δυναμική απόδοση = πραγματική απόδοση + πιθανά σωστά* το οποίο αυξάνεται παράλληλα με τις ιδιοτιμές. Ταυτόχρονα βλέπουμε και μια μετακίνηση από τα εν δυνάμει σωστά στα απολύτων σωστά με την αύξηση των ιδιοτιμών που διατηρήθηκαν για να περιγράψουν τις εικόνες.

Εδώ εμφανίστηκε και το παράδοξο με την κοινή λογική της ανάλυσης της εικόνας. Η απόδοση μοιάζει να είναι να είναι ανεξάρτητη της ανάλυσης και ίσως αυτό να οφείλεται στο ότι οι εικόνες χαμηλής ανάλυσης διατηρούν πληροφορία στα ΣΗΜΑΝΤΙΚΑ χαρακτηριστικά των προσώπων.

Τα ανώτερα όρια απόδοσης δεν μεταβλήθηκαν κάτι που σημαίνει πως η μέθοδος δεν χρειάζεται αλλά ούτε και ωφελείται από μεγάλου μεγέθους σε (MB) εικόνες για δεδομένη ανάλυση.

Οι μέθοδοι *aver* και ελάχιστης απόστασης ανά κλάση είχαν τα ίδια ποσοστά επιτυχίας σε όλες τις ομάδες τεστ που τρέξαμε ανεξαρτήτου άλλων παραμέτρων , ανάλυση, ποιότητα, πλήθος ιδιοτιμών.

Ωστόσο ο διαφορετικός τρόπος που προσεγγίζουν το πρόβλημα αναδεικνύει μια μεγάλη διαφορά μεταξύ τους. Η μέθοδος *aver* έχει τη διπλάσια δυνατότητα επιτυχίας ανάμεσα στις αποτυχίες ανεξαρτήτου ιδιοτιμών ή ανάλυσης.

Σε ό,τι συνθήκες και να τρέξαμε τα τεστ η *aver* έδινε περίπου την ίδια απόδοση στην πρώτη ελάχιστη απόσταση αλλά έχει τις διπλάσιες εν δυνάμει επιτυχίες μέσα στα λάθη. Να περιέχεται το σωστό άτομο στην δεύτερη τρίτη ή ελάχιστη απόσταση.

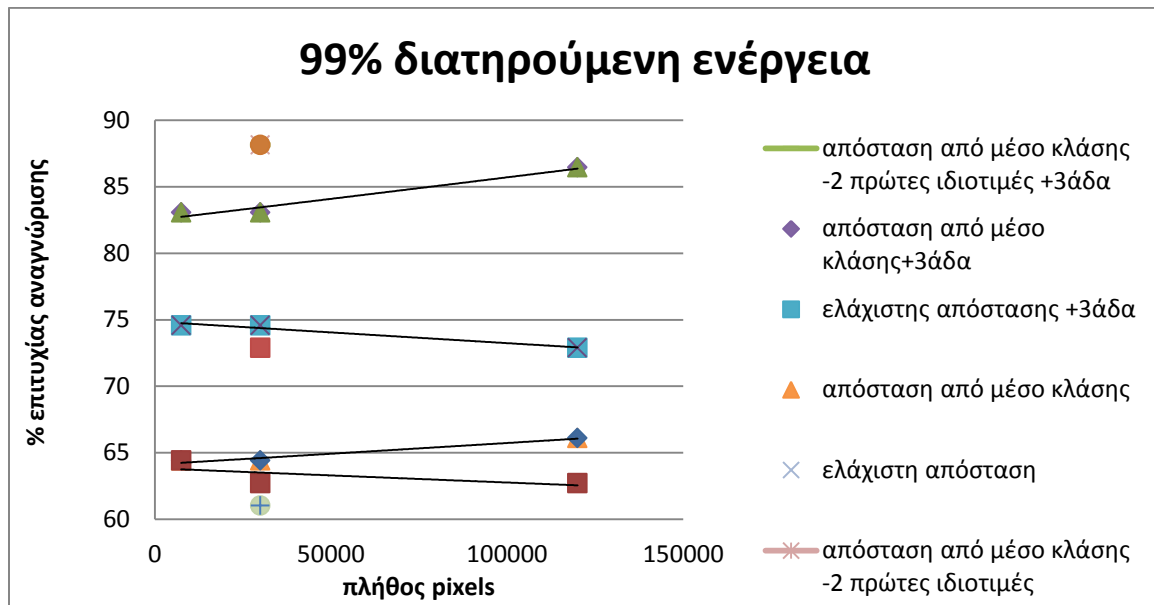
Τα παρακάτω πειράματα έχουν γίνει μόνο για την μέθοδο ανάλυσης με βάση το άτομο, χωρίς το ενδιάμεσο στάδιο της πόζας, λόγω της ανοχής που παρουσίαζε η μέθοδος στην έλλειψη εικόνων , στην δυνατότητα της δηλαδή να καταλήγει σε σωστό άτομο χωρίς να απαιτεί την ύπαρξη ίδιας πόζας.

4.2.1 Πειραματικά αποτελέσματα που προέκυψαν από την διακύμανση της ανάλυσης

Από τον πίνακα των αποτελεσμάτων εξάγεται το συμπέρασμα ότι η διάσταση της υπό εξέταση εικόνας δεν είναι βασικός παράγοντας για τη μέθοδο.

Δοκιμάστηκαν αναλύσεις 400x300 , 200x150 και 100x75 pixels

Τα αποτελέσματα δείχνουν τις αναλύσεις με βάση τα απόλυτα επιτυχημένα τεστ να είναι ανακατεμένες και εξαρτώμενες περισσότερο από τις ιδιοτιμές που διατηρήσαμε.



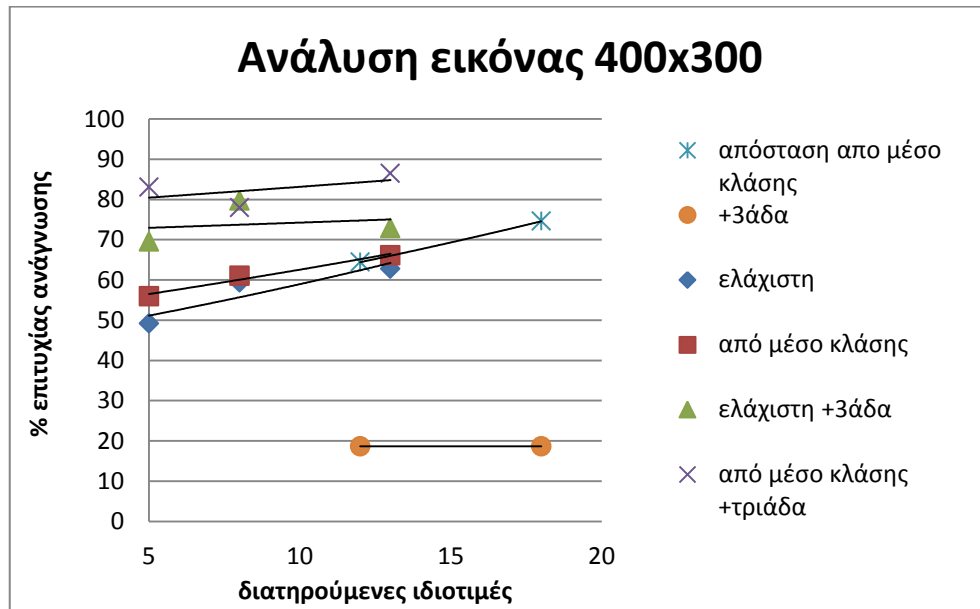
Διάγραμμα 11: Διάγραμμα μεταβολής απόδοσης - ανάλυσης εικόνας

Από την παραπάνω γραφική φαίνεται πως η απόκλιση για οποιαδήποτε μέθοδο δεν ξεπερνά το 2-3% ποσοστό που δεν μπορεί να χαρακτηριστεί ασφαλές για συμπεράσματα. Η ανάλυση δεν φαίνεται να επηρεάζει την απόδοση. Επίσης φαίνεται η καθαρή υπεροχή της μεθόδου μέτρησης απόστασης από το μέσο όρο της κλάσης που έδωσε βελτίωση απόδοσης περίπου 20% και υπερίσχυσε της μεθόδου ελαχίστου απόστασης κατά 10 ποσοστιαίες μονάδες.

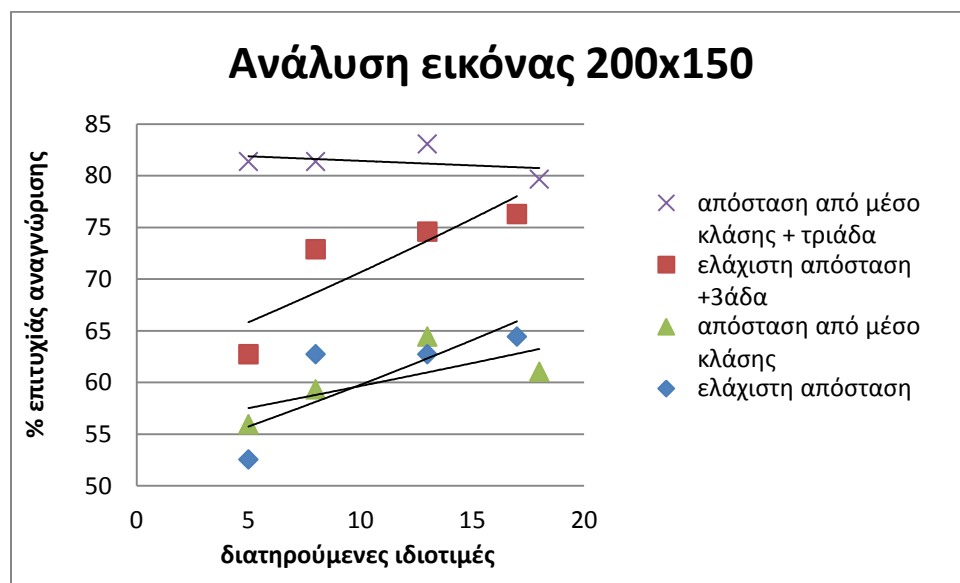
4.2.2 Πειραματικά αποτελέσματα που προέκυψαν από την διακύμανση του αριθμού των διατηρούμενων ιδιοτιμών

Για κάθε ανάλυση ξεχωριστά υπήρξε αύξηση της απόδοσης τόσο για την μέθοδο ελάχιστης απόστασης από κλάση όσο και για τη μέθοδο της απόστασης από το μέσο όρο της κλάσης με την αύξηση των ιδιοτιμών.

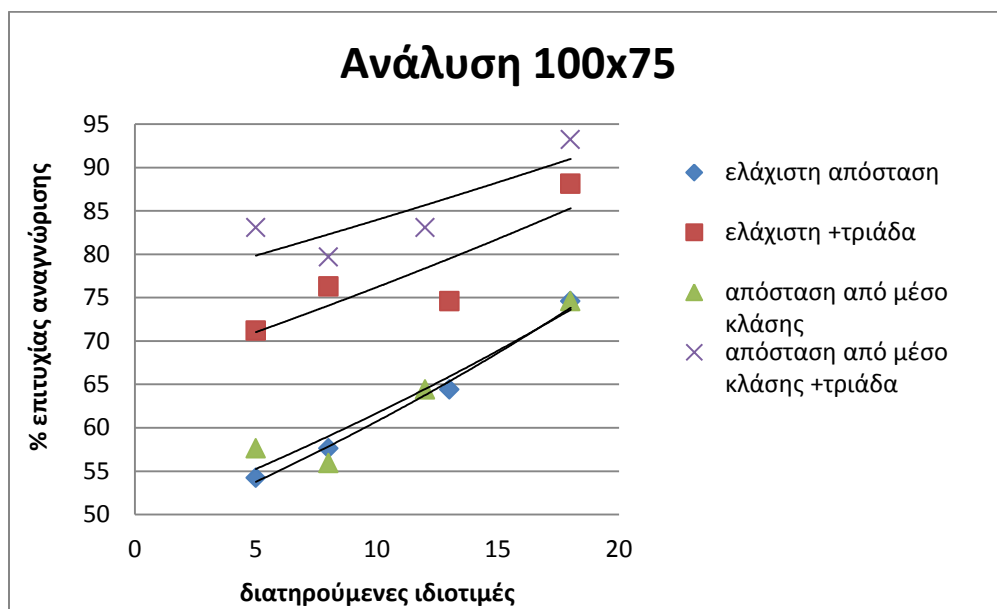
Διατηρώντας δηλαδή σταθερή ανάλυση εικόνων, αύξηση των ιδιοτιμών οδηγεί σε αύξηση της απόδοσης. Αυτό το φαινόμενο διατηρείται μέχρι το ποσοστό της διατηρούμενης ενέργειας να ξεπεράσει το 97,5% ποσοστό που σε πλήθος ιδιοτιμών ισοδυναμεί σε περίπου 18. Περαιτέρω αύξηση σε όσα τεστ δοκιμάστηκε δεν συνοδεύονταν ούτε από αύξηση των σωστών ούτε των εν δυνάμει σωστών (+τριάδα) αποτελεσμάτων.



Διάγραμμα 12: Ανάλυση 400x300 % απόδοσης – διατηρούμενες ιδιοτιμές



Διάγραμμα 13: Ανάλυση 200x150 % απόδοσης – διατηρούμενες ιδιοτιμές



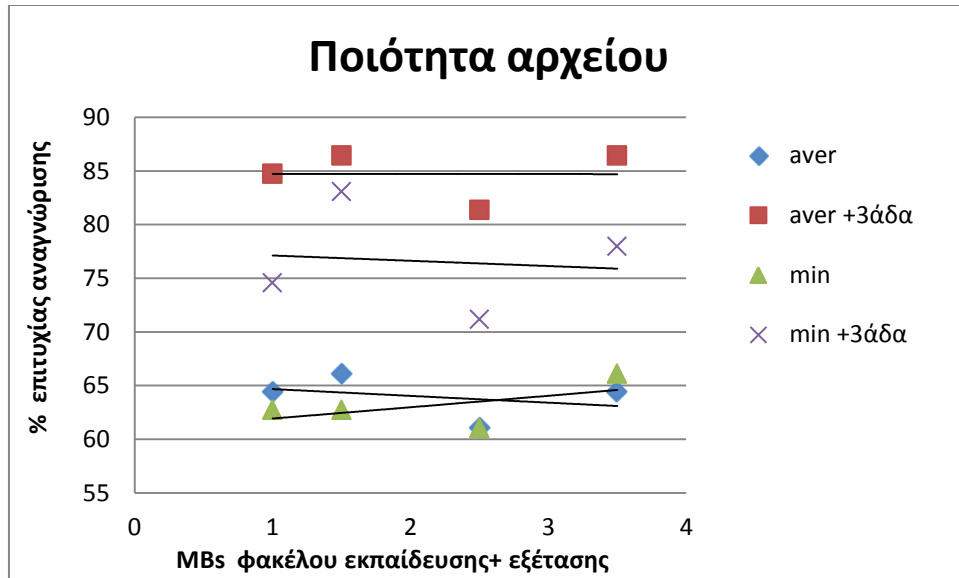
Διάγραμμα 14: Ανάλυση 100x75 % απόδοσης – διατηρούμενες ιδιοτιμές

Το συμπέρασμα και από αυτά τα διαγράμματα είναι ότι η μέθοδος που χρησιμοποιεί την απόσταση από το μέσο όρο της κλάσης λειτουργεί σωστότερα μέσα στις αρχικά αποτυχημένες προσπάθειες.

4.2.3 Πειραματικά αποτελέσματα που προέκυψαν από την διακύμανση το μέγεθος ποιότητα των εικόνων, για δεδομένη ανάλυση

Ένα ακόμα χαρακτηριστικό που τέθηκε υπό εξέταση είναι η ποιότητα της εικόνας. Για το σκοπό αυτό έγιναν με τη χρήση Photoshop* επέμβαση στην διατηρούμενη ποσότητα πληροφορίας που κρατάει η μετατροπή από jpg σε jpg και από τις αρχικές εικόνες 400x300 με συνολικό μέγεθος φακέλου 3,5MB δημιουργήσαμε σύνολα τεστ με μέγεθος 2,5MB 1,5MB και ένα υβριδικό τεστ με εικόνες ποιότητας 3,5MB για εκπαίδευση και χαμηλής ποιότητας προς εξέταση εικόνες από το φάκελο μεγέθους 1,5MB.

* το Photoshop εκτελεί συμπίεση δύο σταδίων 1^ο στάδιο Discrete Cosine Transform (DCT) ανάλυση με block size 8x8 Και μετά σε κάθε block επανασυμπίεζει απορρίπτοντας τιμές συντελεστών κάτω ενός κατωφλίου που ο χρήστης ορίζει και στρογγυλοποιεί τις τιμές που υπερβαίνουν το κατώφλι.

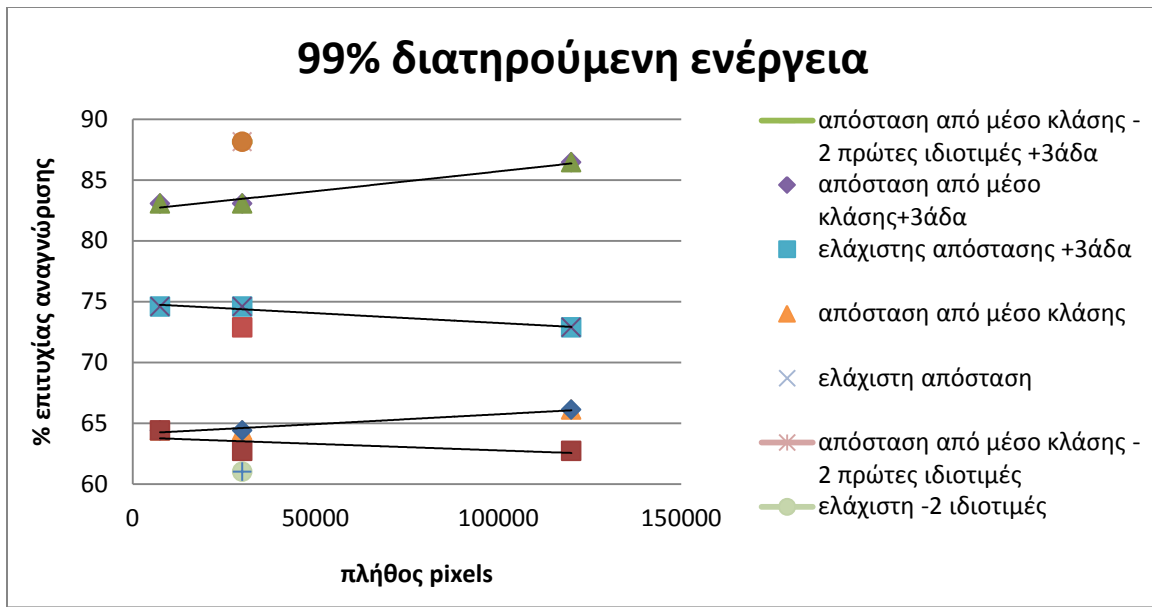


Διάγραμμα 15: Μεταβολή της απόδοσης συναρτήσει της ποιότητας της εικόνας.

Στο παραπάνω διάγραμμα φαίνεται πως το μέγεθος της εικόνας σε MB δεν σχετίζεται άμεσα με την απόδοση της μεθόδου, με κοινή συμπεριφορά για όλες τις μεθόδους μέτρησης απόστασης.

4.2.4 Πειραματικά αποτελέσματα που πήραμε από την αφαίρεση των πρώτων ιδιοτιμών (illumination compensation)

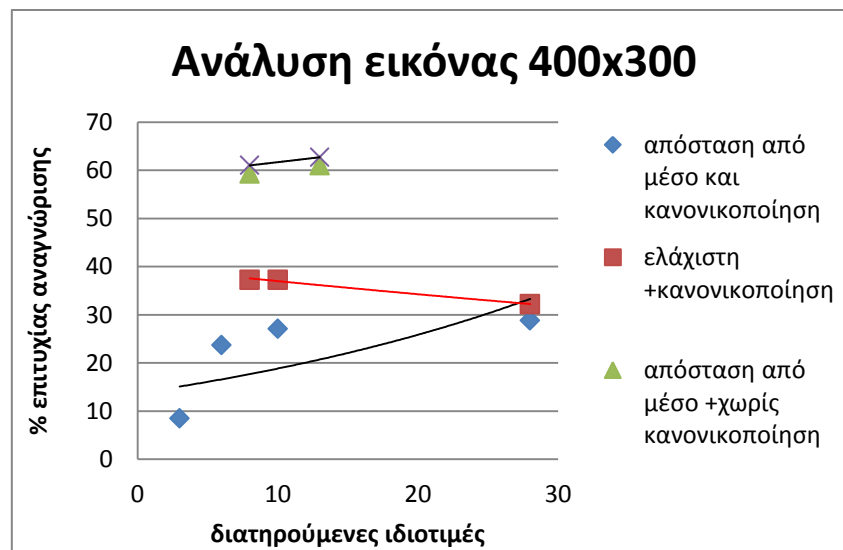
Εξετάστηκε η αφαίρεση της πρώτης, των δύο πρώτων και τις τρεις πρώτες ιδιοτιμές σύμφωνα με τις θεωρητικές αναφορές άλλων δημοσιεύσεων. Τα αποτελέσματα έδειξαν πως αφαιρώντας μέχρι δύο ιδιοτιμές δεν επηρεάζεται το ποσοστό επιτυχίας του συστήματος δεδομένου ότι έχουν διατηρηθεί αρκετές ιδιοτιμές αρχικά. Αυτό ήταν αναμενόμενο αν εξετασθούν και οι δημοσιεύσεις που αναφέρουν ότι οι πρώτες ιδιοτιμές περιέχουν πληροφορία για τον προσανατολισμό των προσώπων, τη φυλή και ότι στις μικρότερες ιδιοτιμές και τα αντίστοιχα ιδιοδιανύσματα βρίσκουμε ιδιοχαρακτηριστικά που προσδιορίζουν τα άτομα.



Διάγραμμα 16: % επίδραση των πρώτων ιδιοτιμών στην απόδοση της μεθόδου.

4.2.5 Κανονικοποίηση

Αντίθετα με ότι από βιβλιογραφία αναφερόταν και αναμέναμε η κανονικοποίηση των εικόνων έδωσε πολύ χειρότερα αποτελέσματα από τις τυχαίες εικόνες του αρχικού σετ. Για να κανονικοποιήσουμε τις εικόνες έγινε crop από αυτές με σκοπό τα πρόσωπα να καταλαμβάνουν περίπου το ίδιο μέρος του καρέ και να είναι ίδιων διαστάσεων.



Διάγραμμα 17: Επίδραση της κανονικοποίησης στην απόδοση της μεθόδου.

Εδώ θα πρέπει να αναφερθεί ότι όλες οι βιβλιογραφικές αναφορές που εξετάζουν ζητήματα κανονικοποίησης κάνουν λόγο για διεστιακή απόσταση ματιών καθώς και κέντρο στόματος, κάτι που παρέχεται μαζί με το σύνολο των εικόνων της βάσης προσώπων FERET. Αυτό σημαίνει πως αυτόματα απορρίπτουν όλες τις άλλες πόζες και εξετάζουν μόνο την “en face” πόζα των ατόμων. Γι’αυτό και τα αποτελέσματα δεν είναι άμεσα συγκρίσιμα.

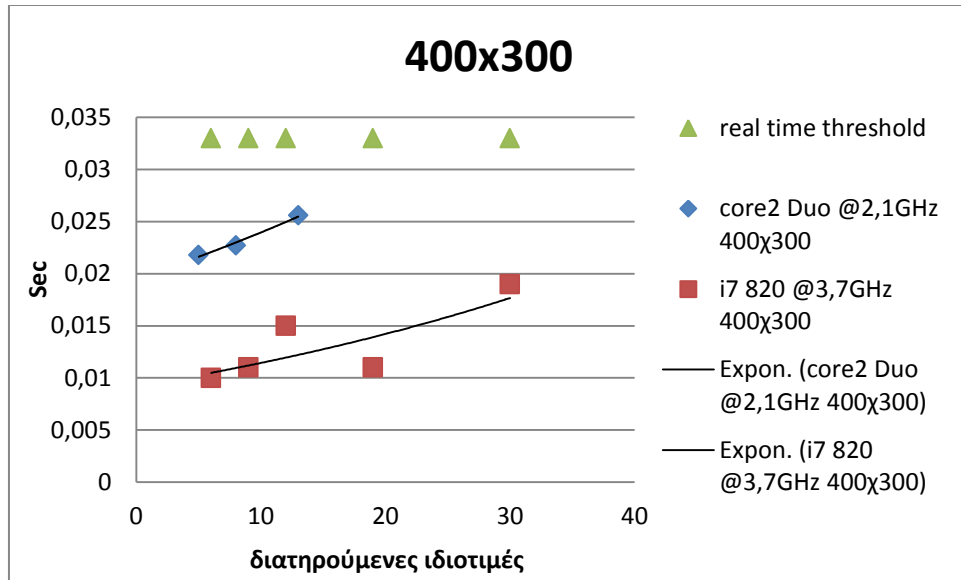
4.3 Επεξεργαστική Ισχύς και χρονική ανάλυση

Σε ένα φορητό σύστημα Intel Core2Duo με χρονισμό 2.1GHz ,2 GB φυσικής μνήμης , λειτουργικό win Vista x32 το σύστημα τοποθετείται εντός της real time έννοιας με χρόνους επεξεργασίας που δεν ξεπερνούν τα 25 msec για ανάλυση φωτογραφιών μεγέθους 400x300 και πέφτουν αισθητά για μικρότερες αναλύσεις εικόνων.

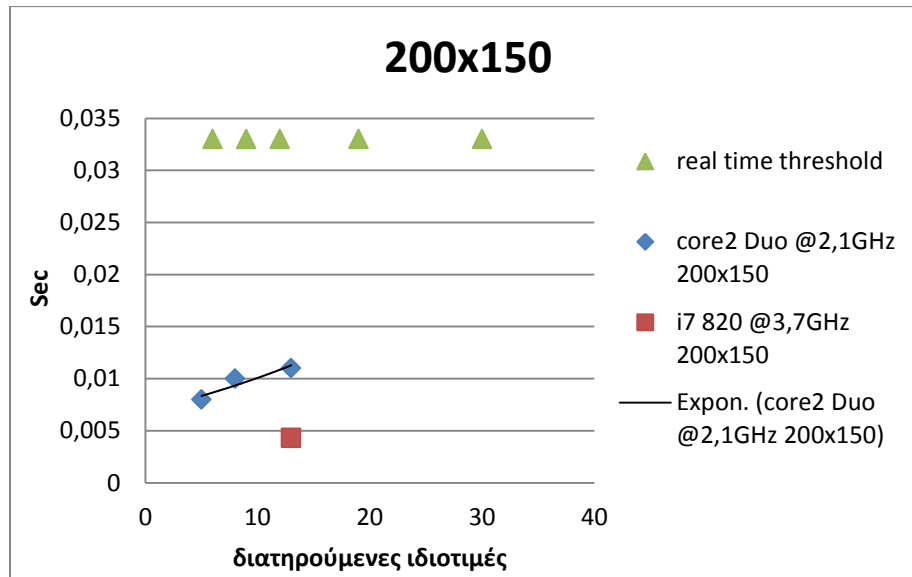
Ιδιοτιμές /Ανάλυση	5	6	8	9	12	13	19
400x300 A*	0,0218		0,0227			0,0256	
400x300 B**		0,01		0,011	0,015	0,011	0,019
200x150 A	0,008		0,01			0,011	
200x150 B						0,0043	
100x75 A	0,0068		0,0071			0,0068	
100x75 B						0,003	

Πίνακας 4: Ενδεικτικοί χρόνοι επεξεργασίας. * A= intel core 2 Duo @2,1GHZ,2GB φυσική μνήμη, 7200rpm Disk, Win Vista x32 **B= intel i7 860 @ 3,8GHZ,8GB φυσική μνήμη, Solid State Disk, Win 7 x64. Χρόνοι σε Sec

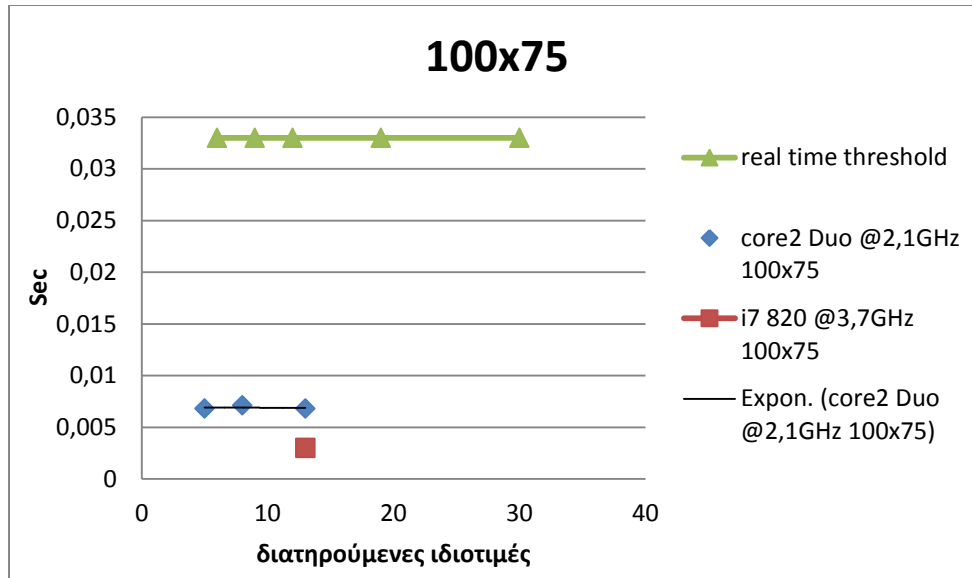
Αυτό υποδηλώνει ξεκάθαρα πως ούτε οι ιδιοτιμές ούτε η αύξηση της ανάλυσης δεν είναι εμπόδιο στην υλοποίηση του αλγορίθμου. Και μάλιστα υπάρχει περιθώριο και για face detection ή προεπεξεργασία της τάξης των 20ms ανά καρέ video.



Διάγραμμα 18: Ανάλυση 400x300 επίδραση αριθμού ιδιοτιμών στην επεξεργαστική πολυπλοκότητα.



Διάγραμμα 19: Ανάλυση 200x150 επίδραση αριθμού ιδιοτιμών στην επεξεργαστική πολυπλοκότητα.



Διάγραμμα 20: Ανάλυση 100x75 επίδραση αριθμού ιδιοτιμών στην επεξεργαστική πολυπλοκότητα.

4.4 Περιορισμοί του αλγορίθμου

Η επιτυχία ενός πρακτικού συστήματος αναγνώρισης προσώπων με εικόνες που συλλαμβάνονται σε πραγματικό χρόνο εξαρτάται από διάφορους εξωτερικούς παράγοντες όπως:

- Κανονικοποίηση του μεγέθους προσώπου
- Τυχαία θέση προσώπου (3D θέση, κίνηση κεφαλιού)
- Ανοχή στην έκφραση προσώπου και στο παρουσιαστικό (γένια, μαλλιά)
- Διαφορετικότητα στις συνθήκες φωτισμού (εσωτερικός φωτισμός, φυσικό φως και γωνίες αυτού)
- Μερική απόκρυψη αλλοίωση προσώπου (γυαλιά ηλίου, καπέλα, μαντήλια)
- Σταθερότητα στη γήρανση.

Χρησιμοποιήθηκε μια βάση προσώπων που είχε κανονικοποιημένο φωτισμό αλλά δεν εξασφάλιζε ούτε την πόζα ούτε το σχετικό μέγεθος του κεφαλιού. Οι φωτογραφίες φαίνεται να έχουν ληφθεί όλες μαζί για κάθε άτομο, οπότε στοιχεία γήρανσης ή αλλαγές σε περιφερικά χαρακτηριστικά δεν μελετήθηκαν.

5 Συμπεράσματα

Πολλές μέθοδοι που εκτελούν αναγνώριση προσώπων περιορίζονται από τη χρήση μοντέλων προτύπων προσώπου ή περιγραφής χαρακτηριστικών (σύγκριση απλών αποστάσεων), υποθέτοντας πως ένα πρόσωπο δεν είναι τίποτα παραπάνω από το άθροισμα των επιμέρους ανεξάρτητων στοιχείων του. Κατά κανόνα αγνοούν ένα μεγάλο μέρος της πληροφορίας που κρύβεται στα βάρη των προτύπων και κάνουν δύσκολη την προσαρμογή ή την εκτίμηση μέρους της υλοποίησης.

Η συγκεκριμένη μέθοδος που κάνει χρήση της PCA για την αναγνώριση προσώπων δημιουργήθηκε από την θεωρία πληροφορίας και οδηγεί σε ένα μικρό σύνολο χαρακτηριστικών εικόνων που περιγράφει καλύτερα το σύνολο των γνωστών προσώπων, χωρίς αυτές να ανταποκρίνονται στα επιμέρους χαρακτηριστικά όπως τα έχουμε στο μυαλό μας.

Τα ιδιοπρόσωπα ως προσέγγιση είναι μια μέθοδος που ταιριάζει πολύ καλά στην περιγραφή του προβλήματος της αναγνώρισης προσώπων. Είναι γρήγορη, σχετικά απλή στην υλοποίηση και λειτουργεί επαρκώς καλά σε ένα κλειστό, ελεγχόμενο περιβάλλον.

Μερικά θέματα που πρέπει να επιλυθούν είναι η αδυναμία προσαρμογής στις αλλαγές φωτισμού καθώς επίσης και στον αριθμό των εκπαιδευόμενων εικόνων συνολικά.

Παράγοντες όπως η ποιότητα της εικόνας ή η ανάλυση σύμφωνα με τα πειράματα δεν είχαν επίδραση στην απόδοση της εικόνας. Η μέθοδος που υλοποιήθηκε απέδωσε το ίδιο καλά ακόμα και σε εικόνες 100x75 , χωρίς κανονικοποίηση. Η κανονικοποίηση μπερδέψε περισσότερο το σύνολο των εικόνων της κλάσης και αυτό ίσως να εξηγεί τα κατά πολύ χειρότερα αποτελέσματα των συγκεκριμένων πειραμάτων.

Ο πρώτος παράγοντας σε βαρύτητα που επηρεάζει την απόδοση της μεθόδου είναι ο αριθμός των διατηρούμενων ιδιοδιανυσμάτων, πρόβλημα άλυτο για τη βιβλιογραφία. Έχουν εφαρμοστεί ο κανόνας του αγκώνα, το ποσοστό της διατηρούμενης ενέργειας ~95% και η απόρριψη του 40% των μικρότερων ιδιοτιμών καμία από τις παραπάνω μεθόδους δεν περιέγραψε σωστά την κατάσταση της βάσης δεδομένων που μελετήθηκε στην εργασία αυτή. Ωστόσο λόγω της μικρής πολυπλοκότητας που προσθέτει η αύξηση των ιδιοτιμών μπορούν να διατηρηθούν περισσότερες ιδιοτιμές και να είναι βέβαιο ότι θα επιτευχθεί η βέλτιστη απόδοση.

6 Μελλοντικές προεκτάσεις

Ενδιαφέρον θα ήταν να μπορέσουν να μετατραπούν οι εν δυνάμει επιτυχίες με κάποιο τρόπο σε πραγματικές επιτυχίες. Ο αλγόριθμός που υλοποιεί η μέθοδος σε συνεργασία με την μέτρηση απόστασης από το μέσο της κλάσης φαίνεται να έχει περιθώρια βελτίωσης. Παρατηρήθηκε το εξής φαινόμενο: Με τη διατήρηση του 95% της αρχικής ενέργειας των ιδιοδιανυσμάτων, που αντιστοιχεί σε περίπου 5-7 ιδιοτιμές για τη βάση που εξετάστηκε, υπήρχαν αρκετές σωστές προβλέψεις. Υπήρχαν όμως και αρκετές προβλέψεις στην δεύτερη τρίτη και τέταρτη ελάχιστη απόσταση. Αύξηση του αριθμού των διατηρούμενων ιδιοτιμών μετακινούσε τις προβλέψεις κατά μία θέση πιο κοντά στην επιλογή του αλγορίθμου για παράδειγμα από την τρίτη εναλλακτική ελάχιστη στη δεύτερη εναλλακτική ελάχιστη απόσταση ή από την δεύτερη εναλλακτική ελάχιστη στην πρώτη εναλλακτική ελάχιστη. Περαιτέρω αύξηση των διατηρούμενων ιδιοτιμών οδηγούσε εν δυνάμει επιτυχίες να μετατραπούν σε πραγματικές επιτυχίες. Από το σημείο που γινόταν χρήση του 99,5% της αρχικής ενέργειας δεν υπήρχε μετακίνηση των εν δυνάμει σωστών προβλέψεων, ενώ εξακολουθούσαν να υπάρχουν αρκετά. Η δέσμευση αυτής της έξτρα πληροφορίας και η μετατροπή της σε επιτυχία θα μπορούσε ίσως να συνδυαστεί με το μοτίβο κίνησης των εν δυνάμει επιτυχιών, τη θέση τους δηλαδή στην πρώτη τετράδα ελαχίστων αποστάσεων.

Αυτή η εργασία βασίζεται στην αναγνώριση ιδιοπροσώπων που επιτυγχάνει μέγιστη απόδοση 92.5%. Οι προσαρμοζόμενοι αλγόριθμοι μπορεί να επιτύχουν καλύτερο ελάχιστο ποσοστό απόδοσης. Υπάρχει η δυνατότητα βελτίωσης του αλγορίθμου με τη χρήση νευρωνικών δικτύων που υπερσχύει της ανάλυσης με τα ιδιοπρόσωπα. Με τη χρήση νευρωνικών δικτύων μπορούμε να επιτύχουμε αύξηση της απόδοσης αντί να έχουμε μια σταθερή τιμή κατωφλίου μπορεί αυτή να είναι δυναμική και να ακολουθεί τις συνθήκες της εκάστοτε βάσης δεδομένων έτσι ώστε να μεγιστοποιείται η ακρίβεια.

Επιπλέον πρέπει να δοκιμαστεί ο αλγόριθμος και με άλλη βάση δεδομένων που να περιέχει διαφοροποιήσεις στο φωτισμό για να φανεί αν πράγματι η απόδοση χωρίς τα δύο πρώτα ιδιοδιανύσματα παραμένει σταθερή. Τα πειράματα έδειξαν ότι ο φωτισμός επηρεάζει τα πρώτα πέντε ιδιοδιανύσματα αλλά με την αφαίρεση του τρίτου ιδιοδιανύσματος είχαμε πτώση της απόδοσης. Χωρίς αλλαγή στο φωτισμό όταν αφαιρέθηκε και το τρίτο ιδιοδιάνυσμα έπεσε δραστικά η απόδοση.

Τέλος η προσέγγιση με τη μέθοδο του ενδιάμεσου βήματος της πόζας μένει να εξεταστεί και με τη χρήση μιας άλλης βάσης δεδομένων πληρέστερης από τη βάση του MIT που χρησιμοποιήθηκε, έτσι ώστε στην διαδικασία που χωρίζονται τυχαία τα τεστ να υπάρχει εικόνα για κάθε άτομο σε κάθε πόζα, κάτι που δεν κρίνεται υπερβολικό δεδομένου το

σκοπό της εφαρμογής, αναζήτηση ταυτότητας σε μικρό προεγγεγραμμένο δείγμα χρηστών.

Βιβλιογραφία

1. **Eigenspace-Based Face Recognition: A Comparative Study of Different Approaches**, Javier Ruiz-del-Solar, Member, IEEE, and Pablo Navarrete, Student Member, IEEE IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS, VOL. 35, NO. 3, AUGUST 2005.
2. **Face recognition by humans: 19 results all computer vision researchers should know about**, Pawan Sinha, Benjamin Balas, Yuri Ostrovsky, and Richard Russell Proceedings of the IEEE | Vol. 94, No. 11, pages 1948-1959 November 2006
3. M. Turk and A. Pentland, “**Eigenfaces for recognition**,” *J. Cogn. Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.
4. **Principal Component Analysis**, Herve Abdi & Lynne J. Williams (in press, 2010). *Wiley Interdisciplinary Reviews: Computational Statistics*, 2.
5. **Illumination compensation and normalization in eigenspace-based face recognition: A comparative study of different pre-processing approaches**. Javier Ruiz-del-Solar *, Julio Quinteros Pattern Recognition Letters 29 (2008) 1966–1979.
6. **Elementary Linear Algebra 5e**” by Howard Anton, Publisher John Wiley & Sons Inc, ISBN 0-471-85223-6.
7. **Face Recognition Across Pose and Illumination**, Ralph Gross, Simon Baker, Iain Matthews, and Takeo Kanade
8. **Face recognition across pose: A review**, Xiaozheng Zhang*, Yongsheng Gao Computer Vision and Image Processing Lab, Institute for Integrated and Intelligent Systems, Griffith University, Australia , Pattern Recognition 42 pages 2876 – 2896 ,2009
9. **A Comparative Implementation of PCA Face Recognition Algorithm**, Nicolas Morizet, Frédéric Amiel, Insaf Dris Hamed, Thomas Ea Department of Electronics, Institut Supérieur d’Electronique de Paris (I.S.E.P.), Paris, France, Stan Z.Li & Anil K. Jain “ ” Springer publications
10. <http://www.face-rec.org>
11. <http://www.alglib.net>
12. <http://math.fullerton.edu/mathews/n2003/JacobiMethodProg.html>
13. www.biometrics.gov
14. www.biometrics.org

Παράρτημα Α

Κώδικας Matlab

```
%pcacovPrincipal component analysis on covariance matrix
%pcaresResiduals from principal component analysis
%princompPrincipal component analysis (PCA) on data
%mahal Mahalanobis distance

%improved Speed for Calculating N-D Euclidean Distance Transforms
with the bwdist Function
%A new algorithm improves the speed and reduces the memory
footprint for the bwdist function.

%training
clc;
clear all; % Remove items from workspace, freeing up system
memory
close all;
global db_max;

person.name = []; %name of persons in DB
person.path = []; %location = path name containing face data
person.desired = [];
person.obtained = [];
stats.test=[];
stats.idiotimes=[];
stats.dist=[]'
stats.losest=[];
[~, p1]=uigetfile('*.jpg','epelexe ena arxeio apo ti basi'); %to
~ to proteine mono tou **FIX**
tic; %arxizei i metrisi xronoy gia tin fasi tis ekpaideysis
title='The system is learning';
msg_handle = msgbox( title, 'Please Wait...' );
ext = '*.jpg';

%diatrexoume ti basi
%ka8e katalogos idiou epipedou me ton arxiko periexei ena proswpo
tis basis

op = pwd; %kratame to trexwn path =pwd
cd(p1 ); %periexei apo uigetfile to path tou epilegmenou arxeiou
tis basis
cd .. %anebainoume ena epipedo ston katalogo pou ze i basi mas
```

```

p2 = pwd; %p2 apo8ikeyoume ti riza tis basis
cd ../;
testpath=pwd;
p6=strcat(testpath,'\test\');
cd( op ); %epanerxomaste ston arxiko katalogo
d = dir( p2); %diabazoume tous fakelous gia na 3eroume ta pros
ekpaideysi proswpa
[s x] = size( d );
j = 1; %j kataei ton ari8mo ton proswpwn pou fortwnontai sti basi
%3ekinaei apo 1 giati xrisimopoieitai parakatw ws deiktis se
pinaka

for i = 3:s %prwta yparxoun oi fakeloi . kai .. akolou8oun oi
kanonikoi
    if( d(i).isdir == 1 )%an einai fakelos
        name d(i).name;
        person(j).name = name;
        l = size( name );
        person(j).name_length = l( 2 );
        name1=strcat(p2,'\ ',name,'\ ');

        person(j).path = name1;
        j = j + 1;
    end
end
persons = j - 1; %epeidi to arxikopoisame se 1 afairoume 1

disp( 'diabastikan ta onomata ton proswpwn' );

sms=fprintf(' i basi periexei %d proswpa.',persons);
%sms=sprintf(' i basi periexei %d proswpa.',persons);
%disp(sms);

%fortwnoume ta proswpa apo to disko fdb stin basi proswpwn fdb
fdb = [];
nof = 0; %synolikos ari8mos eikonwn proswpwn sti basi
%nof ari8mos teilwn fdb ,ka8e eikona kataligei mia stili
for i = 1:persons % 'persons' ari8mos atomwn sti basi
    p2 = [ person(i).path ext ];

    d = dir( p2 ); %d = n by 1 struct; n = no of faces
    faces = size( d ); %faces = no of faces of ith person in DB
    for j = 1:faces
        p2 = [ person(i).path d(j).name ];
        %disp(p2); im = imread( p2 );
        infostruc=imfinfo(p2);
        width=infostruc.Width; %krataw tis diastaseis gia to reshape
argotera

```

```

height=infostruc.Height;

%im = rgb2ray(im);
% % rgb2gray converts RGB values to grayscale values
% % by forming a weighted sum of the R, G, and B components:
% % 0.2989 * R + 0.5870 * G + 0.1140 * B

s1 = size(im); % returns the sizes of each dimension of im.
l = length( s1 ); % returns the size of the longest dimension of
s.
if( l == 3 ) % 3=>colour picture.
c = rgb2gray(im);
else c = im;
end

t = double( c(:) ); % convert matrix 'c' in column matrix.
% % image(t);
person(i).fd{j} = t;
fdb = [fdb t]; %edw eskage me to k10 poy eixe alles diastaseis
%kai me alles diastaseis
nof = nof + 1;
end
person(i).faces = faces; %faces = ari8mwn eikonwn gia to i-osto
face sti basi
end

% % %paizei
% % for i=1:persons*person(i).faces(:,1); %emfanizei ola ta atoma
oles tis opseis
% %
% % facel=reshape(fdb(:,i),height,width);
% %
% % max_a=max(max(facel));%megisto synoliko(megisto ana stili))
% %
% % facel=facel./max_av*255; %kanonikopoiisi me max value to 255
gia grayscale
% %
% % figure(;
% % image(facel);%emfanisi tou pinaka ws eikona
% % colormap(gray);
% %
% % end

dfdb = [];
average_face = mean(fdb, 2); %epistrefei to meso(mean) ka8e
grammis kai se morfi stilis

```

```

% % % %ayto douleyei
% % % einai to meso proswpo
% % % % % %
% % % avface=reshape(avrage_face,height,width);
% % %
% % %
% % % max_av=max(max(avface));%megisto synoliko(megisto ana
stili))
% % %
% % % avface=avface./max_av*255; %kanonikopoiisi me max value to
255 gia grayscale
% % %
% % % figure();
% % % image(avface);%emfanisi tou pinaka ws eikona
% % % colormap(gray);
% % %

for i = 1:nof

    %dfdb = [ dfdb double(fdb(:,i:i))-average_face ];
    dfdb = [ dfdb ( fdb(:,i) ) - average_face ];
end

% % % %typwnei ta kanonikopoimena proswpa (face-average face)
% % % for i = 1:nof
% % % facce=reshape(fdb(:,i),height,width);
% % % figure;
% % % image(facce);
% % % colormap(gray);
% % % end

%clear fdb;
dfdb = dfdb / sqrt( nof );
%actual covarince matrix = cm = dfdb * dfdb' -> meaglwn
diastasewn pinakas
%apo 8ewria ypologizoume dfdb' * dfdb = reduced cm -> M diastasi
anti NxN
%reduced cm = rcm -> diastasewn M=nof x nof
rcm = dfdb' * dfdb; %rcm = reduced covariance matrix
[rvector, rvalue] = eig( rcm ); % idiodianysmata kai idiotimes
tou rcm
%Reduced covariance Matrix
%clear rcm;
disp( 'Calculated Eigen Vectors & Eigen Values of Reduced CM' );

```

```

%normalise eigen vectors of rcm
for i = 1:persons
    t1 = rvector( :, i ); % edw skaei prwti fora me megales eikones
    % ??? Index exceeds matrix dimensions.
    % Error in ==> ind_1 at 100
    % t1 = rvector( :, i );
    s1 = sqrt( sum( t1.^2 ) );
    rvector( : i ) = rvector( :, i )./s1;
end

%calculate eigen vector & eigen value of cm = evector & evalue
evector = dfdb * rvector;
evalue = diag( rvalue );

%clear rvector rvalue;
disp( 'Calculated Eigen Vectors & Eigen Values of CM' );
[sorted_evalue, index] = sort( evalue ); %sorted in acending
order
sorted_evalue = flipud( sorted_evalue ); %rearranged in decending
order ***flipud-->flip upside down

index = flipud( index ); %rearranged corresponding indies also
%Now rearrange eigenvectors in the order of rearranged eigen
values
evector( :, 1:nf ) = evector ( :, index );
% % % % % % % % smallest_evlp = 0.1 * sorted_evalue(1);% to
obtain the effective eigen values

% edw 8a kanoume sum ta tetragwna tou evalue
% kai meta 8a kratame tosa wste to meriko sum na pernaei to 95%
toy synolou
% ekei 8a topo8eteitai to index=i;

total_energy_sum=(sum(sorted_evalue.^2));
preserved_energy_factor=0.975;
total_energy_sum=preserved_energy_factor*total_energy_sum;

part_energy=0;

for i=1:size(sorted_evalue);
    part_energy=sorted_evalue(i)^2+part_energy;
    if part_energy>total_energy_sum

        break;
    end
end

%tsekaristike me 28 eikones,<9 proswpa*3> kai me ypologismo mesw
excel ta tetragwna tw n 5 prwtwn idiotimwn

```

```

%periexoun possto megalytero apo 95% tis synolikis energieias
%idio apotelesma kai gia 9 eikones mia ana proswpo

%stis 54 eikones xreiaslike alles fores 6 kai alles fores 7
idiodianysmata
%opote paizei ontws dynamika

% % % % % % % % % for i = 1:nof
% % % % % % % % % if( sorted_evalue(i) < smallest_evlp )
% % % % % % % % % break;
% % % % % % % % % end
% % % % % % % % % end

index = i;idiotimes=index;
stats.idiotimes=idiotimes;
%display(idiotimes);
% kratame osa idiodianysmata pernane to kritirio katwfliwsis
% ***sorted_evalue(i) < smallest_evlp
evector = evector( :, 4:index );
evalue = sorted_evalue( 4:index );
index=index-3;
%normalise eigen vectors of cm
for i = 1:index
    t1 = evector( :, i );
    s1 = sqrt( sum( t1.^2 ) );
    evector( :, i ) = evector( :, i )./s1;
end

%Douleiei pleon
[~,size_e=size(evector); % ~ epeidi den me endiaferei na kratismw
to size tis
% tis prwtis metablitis
% % % for i =1:size_e
% % % idioproswpo=evector(:,i);
% % % idioproswpo=reshape(idioproswpo,150,200);
% % %
% % % max_av=max(idioproswpo);%megisto ana stili
% % % max_av=max(max_av);%megisto synolika
% % %
% % % idioproswpo=idiproswpo./max_av*255; %kanonikopoiisi me max
value to 255 gia grayscale
% % % figure;
% % % image(idioproswpo);

```

```

% % % %display('1');
% % % colormap(gray);
% % % end

%%Calculating RPV

s = size( average_face );
for i = 1:persons
    p2 = [ person(i).path ext ];
    d = dir( p2 );

    sum_rpv = zeros( index, 1 ); %rpv = recognising pattern vector
    for j = 1:person(i).faces

        t = person(i).fd{j} - average_face;
        rpv = evector' * t; %rpv = recognising pattern vector
        rpv = rpv / s(1,1); %kanonikopoiisi tou rpv
        person(i).rpv{j} = rpv;
        person(i).ido{j} = i;
        sum_rpv = sum_rpv + rpv;
    end

    person(i).average_rpv{1} = sum_rpv / person(i).faces(1,1);
    distance = zeros( person(i).faces(1, 1), 1);

    for j = 1:person(i).faces
        p2 = [ person(i).path d(j).name ];
        distance( j, 1 ) = norm(person(i).rpv{j} -
person(i).average_rpv{1} ); %=[ alla epitrepoun nested pra3eis
keliwn
    end
    [min_dist, index_dist ] = min( distance );
    [max_dist, index_dist ] = max (distance);
    person(i).min_dist = min_dist;
    person(i).max_dist = max_dist;
end

for k= 1:persons
    minmum(k) = person(k).min_dist;
    maxmum(k) = person(k).max_dist;
end

db_min=min(minmum);
db_max=max(maxmum);
%display(db_max);
disp( 'Calclated Recognising Pattern Vectors' );

```

```

close(msg_handle); %close dlg box showing "Learning..."

t = toc; %stop counting time to estimate how long the training
takes
t = seconds2human( t ); %metafrazei se readable ton xrono pou
eixame se secs
ms = 'Training Completed Successfully in ';
ms = [ms t];

msgbox( ms, 'Completed' );

t=0;

read_test_files;

%[~, p6]=uigetfile('*.jpg','epelexe ena arxeio apo ton katalogo
e3etasis');
ext = '*.jpg';

d = dir( p6 ); %diabazoume tous fakelous gia na 3eroume ta pros
ekpaideysi proswpa
[s x] = size( d );

%fortwnoume ta proswpa pros e3etasi

tic_or=0;
tic_alt=0;

faces = size( d ); %faces = no of faces of ith person in DB
test_face_num=(faces(1)-2);
for j = 3:faces
p5 = [ d(j,1).name];
%dis(p5);

t11=0;
ti=0;
tests;
tests_alt;
end

```

```

% fprintf(' 8a e3etasoume %d proswpa.',faces);

msgbox(' Done ' );

%%Testing

% diabazoume mia eikona ,
% an einai egxrwmi metatrepoume se grayscale kai metatrepoume se
pinaka stili.
% an einai grayscale metatrepoume se pinaka stili.
% % ypologizoume to rpv kai ta Distance kai elaxisti apostasi

%select the desired file to be tested
f=p5;%, p1]=uigetfile('*.','Select a file for Testing');
tic
file_name =p5;
file_name=strcat( p6, f );
stats.test=file_name;
% Read the image.
im = imread( file_name );
s = size(im); % returns the sizes of each dimension of im.
l = length( s ); % returns the size of the longest dimension of
s.
if( l == 3 ) % 3=>colour picture.
    c = rgb2gray(im);
else
    c = im;
end
t = ( c(:) ); % convert matrix 'c' in column matrix.

%%

si = size( aveage_face );
fi = double(t) - aveage_face; %f = difference of the given face
from average.
x_rpv_m = ((evector)' * fi)/ s(1,1); %recognising pattern vector.

sum_sum=zeros(size(person(1).average_rpv{1}));
for i = 1:persons
    sum_sum=sum_sum+person(i).average_rpv{1};
end
min_dis_avg=norm(x_rpv_m - sum_sum./persons);

```

```

%disp(min_dis_avg);
db_max_avg=80; %<=====THRESHOLD 2

if(min_dis_avg > db_max_avg)
% title = 'NOT A FACE,';
% msg = [title ' DIFFERENT ITEM ' ];
% h = msgbox(msg );

else
% msg = 'PROBABLY A FACE';
% h = msgbox(msg );

    s = size( average_face ); %%%..%% giati na mi diairesoume me
to max(average_face); TO MAX(AVERAGEFACE) einai tis taksis tou
150 enw to size to average face einai *1000 peripoy, tris takseis
panw, giati na to kaneis auto??
    f = double(t) - average_face;
    x_rpv = (evector)' * f; %recognising pattern vector.
    x_rpv = x_rpv s(1,1);

    distance = zeros( persons, 1 ); % for dist of given face from
faces in DB.

    for i = 1:persons

        distance( i, 1 ) = norm( x_rpv - person(i).average_rpv{1} );
        test_dist(i,1)=distance( i, 1 );
    end

    [sorted_test_dst, dindex] = sort( test_dist); %sorted in
acending order

    [ min_dist, index ] = min( distance ); % minimum distance in all
distances.

    for q = 1:size(person(index).rpv,2) %.. code
        distance2q, 1 ) = norm( x_rpv - person(index).rpv{q} );
    %#ok<SAGROW> %.. code
    end
    [ min_dist2, index2 ] = min( distance2 ); % minimum distance in
all dstances %.. code

    r.index = index;
    r.min_dist = min_dist;
    r.rpv = x_rpv;

    if(r.min_dist > db_max)
% % title = ' UNKNOWN FACE';

```

```

% % msg = [title ' - not in database ' ];
% % h = msgbox(msg );

else
% % title = 'Closest Match';
% % msg = [title ' = ' person(r.index).name];
% % h msgbox(msg);
stats.closest=person(r.index).name;
stats.dist=r.min_dist;

% fprintf('i poza pou teriazei einai i %s
',person(r.index).name); %.. code
% fprintf('kai nomizo oti einai o typos Number %d tou
fakelou',index2); %.. code
sms11=sprintf('fakelos %s ',person(r.index).name); %.. code
sms12=sprintf('deiktis %d tou fakelou',index2); %.. code
stats.tag1=[sms11];
stats.tag2=[sms12];

end
end
ti=toc;
% open the ile with append permission
tic_or=tic_or+ti;
fid =
fopen('test200x150_unnormalized_975_8ades_person_or_2MB5_idiotimi3
.txt', 'a');
fprintf(fid, '%s %s %f %f %s %s %f \r\n', stats.test,
stats.closest, stats.idiotimes, stats.dist, stats.tag1,
stats.tag2,ti);
% % % % % % % % fprintf(fid, '%s %s %f %f %s %s %f \r\n',
stats.test, stats.closest, stats.idiotimes, stats.dist,
stats.tag1, stats.tag2, t1);
for i=2:4

fprintf(fid,'%s %s %s %s \r\n','person ',
person(1,dindex(i)).name ,'apostasi', sorted_test_dist(i));

end

fprintf(fid,'%f \r\n',(tic_or/test_face_num));
fprintf(fid,'\r\n');
%fprintf(, stats.idiotimes, stats.dist, stats.tag1, stats.tag2,
t1);
fclose(fid);

%%Testing

% diabazoume mia eikona ,

```

```

% an einai egxrwmi metatrepoume se grayscale kai metatrepoume se
pinaka stili.
% an einai grayscale metatrepoume se pinaka stili.
% % ypologizoume to rpv kai ta Distance kai elaxisti apostasi

%select
('*.','Select a file for Testing');
tic
file_name =p5;
file_name=strcat( p6, f );
stats.test=file_name;
% Read the image.
im = imread( file_name );
s = size(im); % returns the sizes of each dimension of im.
l = length( s ); % returns the size of the longest dimension of
s.
if( l == 3 ) % 3=>colour picture.
    c = rgb2gray(im);
else
    c = im;
end
t = ( c(:) ); % convert matrix 'c' in column matrix.

%%

si = size( average_face );
fi = double(t) - average_face; %f = difference of the given face
from average.
x_rpv_m = ((evector)' * fi)/ s(1,1); %recognising pattern vector.

sum_sum=zeros(size(person(1).average_rpv{1}));
for i = 1:persons
    sum_sum=
min_dis_avg=norm(x_rpv_m - sum_sum./persons);

%disp(min_dis_avg);
db_max_avg=80; %<=====THRESHOLD 2

if(min_dis_avg > db_max_avg)
% title = ' NOT A FACE,';
% msg = [title ' DIFFERENT ITEM ' ];
% h = msgbox(msg );

else
% msg = 'PROBABLY A FACE';
% h = msgbox(msg );

```

```

s = size( average_face ); %%%...%% f = double(t) -
average_face;
x_rpv = (
s(1,1);

distance = zeros( persons, 1 ); % for dist of given face from
faces in DB.

for i = 1:persons
%edw prepei na alla3ei o dist calc

%%fid = fopen('distance.txt', 'a');
for j = 1:person(i).faces
rm(x_rpv-person(i).rpv{j}); %=[ alla epitrepoun nested pra3eis
keliwn
% open the file with append permission

test_dist_alt(j,1)=distanze( j, 1 );
[sorted_test_distalt, dindexalt] = sort( distanze);

end
[min_dist, index_dist ] = min( distanze );
[max_dist , index_dist ] = max (distanze);
person(
_dizt = max_dist;
%%f fprintf(fid, '%f %f %f %f %s %s \r\n', i, min_dist,
max_dist);
%%f fclose(fid);
distance( i, 1 ) = person(i).min_dizt ;
end
[ min_dist, index ] = min( distance ); % minimum distance in all
dstances.

for q = 1:size(person(index).rpv,2) %.. code

[ min_dist2, index2 ] = min( distance2 ); % minimum distance in
all dstances %.. code

r.index = index;
r.min_dist = min_dist;
r.rpv = x_rpv;

if(r.min_dist > db_max)
% % title = ' UNKNOWN FACE';
% % msg = [title ' - not in database ' ];
% % h = msgbox(msg );

```

```

else
% % title = 'Closest Match';
% % msg = [title ' = ' person(r.index).name];
% % h = msgbox(msg);
stats.closest=person(r.index).name;
stats.dist=r.min_dist;

oti einai o typos Number %d tou fakelou',index2); %.. code
sms11=sprintf('fakelos %s ',person(r.index).name); %.. code
sms12=sprintf('deiktis %d tou fakelou',index2); %.. code
stats.tag1=[sms11];
stats.tag2=[sms12];
end

end
t11=toc;
% open the file with append permission
fid =
fopen('test200x150_unnormalized_975_8ades_person_alt_2MB5_idiotimi
3.txt', 'a');
fprintf(fid, '%s %s %f %f %s %s %f \r\n', stats.test,
stats.closest, stats.idiotimes, stats.dist, stats.tag1,
stats.tag2 t11);
tic_alt=tic_alt+t11;
for i=2:4

    fprintf(fid,'%s %s %s %s \r\n','person ',
person(1,dindexalt(i)).name , 'apostasi', sorted_test_distalt(i));

end

fprintf(fid,'%f \r\n',(tic_alt/test_face_num));
fprintf(fid,'\r\n');

fclose(fid);

```

Παράρτημα Β

Βοηθητικές Μαθηματικές έννοιες και παράδειγμα PCA

Η PCA είναι μια χρήσιμη στατιστική μέθοδος που εφαρμόζεται και σε άλλους επιστημονικούς κλάδους εκτός της αναγνώρισης προσώπων όπως η συμπίεση εικόνας και είναι μια μέθοδος ανίχνευσης προτύπων σε πολυδιάστατα δεδομένα. Θα αναφερθούμε στην μέθοδο αφού πρώτα εξηγήσουμε βασικές μαθηματικές και στατιστικές έννοιες που χρησιμοποιούνται για τον υπολογισμό της.

Σε αυτό το κομμάτι θα ασχοληθούμε με τα στοιχειώδη μαθηματικά που εμπλέκονται στην PCA προκειμένου να κατανοήσουμε την διαδικασία της Principal Components Analysis.

Οι έννοιες αναλύονται ανεξάρτητα και παρέχονται παραδείγματα. Είναι λιγότερο σημαντική η απομνημόνευση ακριβώς του μηχανισμού με τον οποίο γίνεται κάτι από το να κατανοηθεί ο λόγος για τον οποίο μια τέτοια τεχνική χρησιμοποιείται, και πως εξηγείται η σημασία του αποτελέσματος της πράξης για τα εκάστοτε δεδομένα. Δεν χρησιμοποιούνται όλες οι τεχνικές στην PCA αλλά αυτές παρέχουν τις βάσεις για την κατανόηση της.

6.1.1 Στατιστική

Όλη η ιδέα γύρω από την στατιστική είναι πως θα αναλυθεί ένα μεγάλο σύνολο δεδομένων, με τρόπο που να φανερώνει τις συσχετίσεις μεταξύ των ανεξάρτητων δειγμάτων του.

Σύνολο δεδομένων $X = [1 \ 2 \ 4 \ 6 \ 12 \ 15 \ 25 \ 45 \ 68 \ 67 \ 65 \ 98]$

6.1.2 Μέσος όρος

Είναι το άθροισμα των στοιχείων ενός συνόλου προς το πλήθος των στοιχείων του συνόλου αυτού.

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

6.1.3 Τυπική απόκλιση

Περιγράφει το πόσο μοιρασμένα είναι τα δεδομένα του συνόλου γύρω από το μέσο όρο. Τυπική απόκλιση εννοούμε το τι εύρος καταλαμβάνουν τα δεδομένα (SD). Πως υπολογίζεται αυτό? « Η μέση απόσταση από το μέσο όρο του δείγματος σε ένα δείγμα». Υπολογίζεται το τετράγωνο της απόστασης κάθε στοιχείου από το μέσο, αθροίζονται οι αποστάσεις, και διαιρούνται με το πλήθος ελαττωμένο κατά 1 και διατηρείται η θετική τετραγωνική ρίζα αυτού.

$$s = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(n-1)}}$$

Διαιρείται το πλήθος αντί για πλήθος ελαττωμένο κατά 1 όταν πρόκειται για ένα μέρος(dataset) ευρύτερου συνόλου, αν πρόκειται για ολόκληρο τον πληθυσμό τότε η διαίρεση με το πλήθος δίνει ακριβέστερο αποτέλεσμα για την τυπική απόκλιση Standard Deviation

<http://mathcentral.uregina.ca/RR/database/RR.09.95/weston2.html>

6.1.4 Διακύμανση

Είναι ένα άλλο μέτρο που χαρακτηρίζει το πόσο απλωμένα είναι τα δεδομένα ενός συνόλου. Είναι σχεδόν ταυτοτικό με την τυπική απόκλιση .ο τύπος που τη δίνει είναι

$$s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(n-1)}$$

Στην ουσία είναι το τετράγωνο της τυπικής απόκλισης και συμβολίζεται με (s^2) θα χρειαστεί για τον ορισμό της συνδιακύμανσης παρακάτω.

6.1.5 Συνδιακύμανση

Τα δύο τελευταία μέτρα που ορίστηκαν ήταν μονοδιάστατα , τι γίνεται όμως αν το σύνολο που πρέπει να περιγραφθεί είναι περισσότερων διαστάσεων; Όταν υπάρχουν περισσότερες διαστάσεις το ενδιαφέρον εστιάζεται στη μεταξύ τους σχέση, αν αυτή υπάρχει. Για παράδειγμα μπορεί να μελετηθεί το ύψος των μαθητών μια τάξης και οι βαθμοί τους και να γίνει προσπάθεια ανίχνευσης αν υπάρχει σχέση μεταξύ ύψους και βαθμών. Η Standard deviation και η διακύμανση επιδρούν στη μια διάσταση μόνο, έτσι μπορούν να υπολογίσουν την τυπική απόκλιση σε κάθε διάσταση του συνόλου δεδομένων ανεξάρτητα των άλλων διαστάσεων. Ωστόσο είναι χρήσιμο το πόσο διαφέρουν οι διαστάσεις από το μέσο λαμβάνοντας υπόψη και τις άλλες.

Αυτό κάνει η συνδιακύμανση, υπολογίζεται πάντα ανάμεσα σε δύο διαστάσεις. Αν υπολογιστεί η συνδιακύμανση μιας διάστασης και του εαυτού της καταλήγουμε στην διακύμανση. Σε ένα σύστημα 3 διαστάσεων (X,Y,Z) μπορεί να υπολογισθεί η συνδιακύμανση μεταξύ των X,Y και X,Z και Y,Z και μεταξύ των X,X, Y,Y και Z,Z που καταλήγουν στην διακύμανση των αντίστοιχων διαστάσεων

Ο τύπος της συνδιακύμανσης μοιάζει πολύ με της διακύμανσης

Διακύμανση, αν αναπτυχθεί το τετράγωνο $var(X) = \frac{\sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})}{(n-1)}$
 Και αντίστοιχα η συνδιακύμανση μετά την περιγραφή που δόθηκε

$$cov(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{(n-1)}$$

Για κάθε σύνολο δεδομένων πολλαπλασιάζεται η διαφορά του κάθε x από το μέσο του X και του κάθε Y από το μέσο Y , αθροίζονται αυτά τα γινόμενα και δικαιούνται με το πλήθος ελαττωμένο κατά ένα.

Η απόλυτη τιμή δεν παίζει τόσο ρόλο, όσο το πρόσημο του αποτελέσματος.
 Αν είναι θετικό τότε ότι οι δύο διαστάσεις μεγαλώνουν ταυτόχρονα.
 Αν είναι αρνητικό όταν μεγαλώνει η μία διάσταση ελαττώνεται η άλλη
 Και τέλος αν είναι 0 οι δυο διαστάσεις είναι ανεξάρτητες μεταξύ τους.

Για δύο διαστάσεις μπορεί να οπτικοποιηθεί σε ένα διάγραμμα 2 αξόνων τα δεδομένα μας και να προκύψει συμπέρασμα αλλά η μέθοδος της συνδιακύμανσης επιτρέπει να εξαρθούν συμπεράσματα και σε περισσότερες διαστάσεις που δεν υπάρχει η δυνατότητα της γραφικής αναπαράστασης. Η συνδιακύμανση είναι αντιμεταθετική $Cov(x,y)=Cov(y,x)$ διότι στον πολλαπλασιασμό του παραπάνω τύπου δεν έχει σημασία με ποια σειρά θα μπουν οι διαφορές $(X - X_{\text{μέσο}}) (Y - Y_{\text{μέσο}})$

6.1.6 Πίνακας συνδιακύμανσης

Η συνδιακύμανση περιορίζεται στις 2 διαστάσεις
 Αν υπάρχουν περισσότερες, για κάθε διάσταση υπάρχουν παραπάνω από μια τιμές

συνδιακύμανσης. Στην πραγματικότητα για n διαστάσεις υπάρχουν $\frac{n!}{(n-2)!*2}$ τιμές συνδιακύμανσης.

Ένας χρήσιμος τρόπος για να υπολογιστούν όλες τις τιμές συνδιακύμανσης ανάμεσα σε όλες τις διαστάσεις είναι να υπολογιστούν και να τις ομαδοποιηθούν σε ένα πίνακα.

Έτσι ο ορισμός του πίνακα συνδιακύμανσης για ένα σύνολο δεδομένων n διαστάσεων

$$C^{n \times n} = (c_{i,j}, c_{i,j} = cov(Dim_i, Dim_j)),$$

είναι :

Όπου $C^{n \times n}$ είναι πίνακας με n γραμμές και n στήλες, και Dim_x είναι η x -οστή διάσταση. Από τον τύπο βγαίνει το συμπέρασμα ότι υπάρχει η $n \times n$ τετραγωνικός πίνακας για ένα n -διάστατο σύνολο δεδομένων, και κάθε σημείο στον πίνακα είναι το αποτέλεσμα της συνδιακύμανσης ανάμεσα σε δύο ξεχωριστές διαστάσεις. Για παράδειγμα το στοιχείο της

δεύτερης σειράς και τρίτης στήλης $C_{2,3}$ είναι η τιμή της συνδιακύμανσης υπολογισμένης μεταξύ της δεύτερης και τρίτης διάστασης.

Πίνακας συνδιακύμανσης υποθετικού συνόλου δεδομένων τριών διαστάσεων x, y, z

$$C = \begin{pmatrix} cov(x, x) & cov(x, y) & cov(x, z) \\ cov(y, x) & cov(y, y) & cov(y, z) \\ cov(z, x) & cov(z, y) & cov(z, z) \end{pmatrix}$$

Ο πίνακας έχει ως στοιχεία της διαγωνίου την διακύμανση της κάθε διάστασης $Cov(x, x)$ και ότι είναι συμμετρικός ως προς την διαγώνιο αφού $Cov(x, z) = Cov(z, x)$

6.1.7 Άλγεβρα πινάκων

Αυτό το κομμάτι αποτελεί το υπόβαθρο για την άλγεβρα πινάκων που απαιτεί η PCA. Πιο συγκεκριμένα θα ορισθούν τα ιδιοδιανύσματα και οι ιδιοτιμές ενός δοσμένου πίνακα .

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 11 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \times \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 12 \\ 8 \end{pmatrix} = 4 \times \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

Παράδειγμα ενός όχι ιδιοδιανύσματος και ενός ιδιοδιανύσματος.

$$2 \times \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} \times \begin{pmatrix} 6 \\ 4 \end{pmatrix} = \begin{pmatrix} 24 \\ 16 \end{pmatrix} = 4 \times \begin{pmatrix} 6 \\ 4 \end{pmatrix}$$

Παράδειγμα του πως ένα πολλαπλασιασμένο ιδιοδιάνυσμα παραμένει ιδιοδιάνυσμα

Είναι γνωστό ότι μπορούν να πολλαπλασιαστούν δύο πίνακες μεταξύ τους δεδομένου ότι έχουν συμβατό μέγεθος. Τα ιδιοδιανύσματα είναι μια υποπερίπτωση αυτού. Παραπάνω δίνονται δύο περιπτώσεις πολλαπλασιασμού ανάμεσα σε πίνακα και διάνυσμα.

Στο πρώτο παράδειγμα, το διάνυσμα δεν είναι ακέραιο πολλαπλάσιο του αρχικού διανύσματος, ενώ στο δεύτερο παράδειγμα είναι το αποτέλεσμα είναι 4 φορές το αρχικό διάνυσμα. Γιατί συμβαίνει αυτό; Το διάνυσμα είναι ένας δισδιάστατος χώρος. Το διάνυσμα $(3 \ 2)$ (από το δεύτερο παράδειγμα) αντιπροσωπεύει ένα βέλος που ξεκινάει από το σημείο $(0,0)$ και καταλήγει στο σημείο $(3,2)$. Ο άλλος πίνακας ο τετραγωνικός μπορεί να θεωρηθεί σαν ένας πίνακας μετασχηματισμού. Αν πολλαπλασιάσουμε από αριστερά με ένα διάνυσμα, παράγει ένα άλλο διάνυσμα που είναι μετασχηματισμένο από την αρχική του θέση. Είναι η φύση του μετασχηματισμού από όπου πηγάζουν τα ιδιοδιανύσματα. Υπάρχει ένας πίνακας μετασχηματισμού που πολλαπλασιαζόμενος από αριστερά δίνει το συμμετρικό διάνυσμα ως προς την ευθεία $y=x$. Αν ένα διάνυσμα είναι πάνω στην ευθεία $y=x$, το συμμετρικό του είναι το ίδιο το διάνυσμα. Αυτό το διάνυσμα (και όλα τα πολλαπλάσια του, δεν παίζει ρόλο το μήκος του), θα είναι ιδιοδιάνυσμα ενός πίνακα μετασχηματισμού. Τι γνωρίσματα έχουν τα ιδιοδιανύσματα; Πρώτα από όλα ιδιοδιανύσματα έχουν μόνο οι τετραγωνικοί πίνακες, και δεν έχουν όλοι οι τετραγωνικοί πίνακες ιδιοδιανύσματα. Αν ένας τετραγωνικός πίνακας διαστάσεων n έχει ιδιοδιανύσματα τότε αυτά είναι n σε πλήθος. Ένας 3×3 πίνακας, έχει 3 ιδιοδιανύσματα. Ακόμα και αν διαιρεθεί με έναν αριθμό το ιδιοδιάνυσμα πριν τον πολλαπλασιασμό θα το αποτέλεσμα θα είναι ο ίδιος πολλαπλασιαστικός συντελεστής, όπως στο παραπάνω παράδειγμα.

Αυτό γίνεται γιατί με την κλιμάκωση του διανύσματος επηρεάζεται μόνο το μήκος και όχι ο προσανατολισμός του. Τέλος όλα τα ιδιοδιανύσματα ενός πίνακα είναι κάθετα ή ορθογώνια μεταξύ τους ασχέτως του πλήθους των διαστάσεων που έχουμε. Αυτό είναι σημαντικό γιατί μπορούν να εκφραστούν τα δεδομένα σε σχέση με όρους που αναφέρονται σε αυτά τα ορθογώνια ιδιοδιανύσματα, αντί να εκφράζονται ως προς τους X και Y άξονες. Θα χρειαστεί αυτό αργότερα στην PCA.

Επίσης οι μαθηματικοί χρησιμοποιούν τα ιδιοδιανύσματα που το μήκος τους είναι ακριβώς ένα, κι αυτό γιατί το μήκος δεν παίζει ρόλο για το αν είναι ή όχι ιδιοδιανύσματα παρά μόνο ο προσανατολισμός τους. Έτσι προκειμένου τα ιδιοδιανύσματα να παραμένουν σταθερά, όποτε βρίσκεται ένα ιδιοδιάνυσμα κανικοποιείται ώστε να έχει μήκος ένα, έτσι ώστε όλα τα ιδιοδιανύσματα να έχουν το ίδιο μήκος. Παράδειγμα

$(3 \ 2)$ κάθετο

Είναι ιδιοδιάνυσμα ,με μήκος $\sqrt{(3^2+2^2)} = \sqrt{13}$ έτσι διαιρείται το αρχικό διάνυσμα με το μέτρο του και αποκτά μήκος 1.

$$(3 \ 2) / \sqrt{13} = (3/\sqrt{13} \ 2/\sqrt{13})$$

Ο υπολογισμός ιδιοδιανύσματος είναι εύκολος για σχετικά μικρούς πίνακες μέχρι 3×3 . Από εκεί και μετά η διαδικασία γίνεται πιο σύνθετη και επαναληπτική. Υπάρχουν αρκετές μαθηματικές βιβλιοθήκες για όλα τα προγραμματιστικά περιβάλλοντα που υλοποιούν αυτές τις συναρτήσεις.

Ένα χρήσιμο μαθηματικό πακέτο (newmat) μπορεί να βρεθεί στη διεύθυνση <http://webnz.com/robert/>.

Περισσότερες πληροφορίες για τα ιδιοδιανύσματα ,πως τα υπολογίζουμε και την ορθογωνιότητα τους στο [8]

6.1.8 Ιδιοτιμές

Οι ιδιοτιμές είναι άμεσα συνδεδεμένες με τα ιδιοδιανύσματα, στην πράξη υπολογίστηκε μια ιδιοτιμή σε προηγούμενο παράδειγμα. Ο συντελεστής που κλιμακώνει το διάνυσμα μετά τον πολλαπλασιασμό είναι ίδιος ίσος με 4, αυτή είναι η ιδιοτιμή που συνδέεται με το ιδιοδιάνυσμα. Άσχετα από το τι πολλαπλάσιο του ιδιοδιανύσματος υπολογίστηκε πριν πολλαπλασιαστεί με τον τετραγωνικό πίνακα, πάντα θα προκύπτει 4 φορές μεγαλύτερο διάνυσμα όπως στο παράδειγμα .

Τα ιδιοδιανύσματα και οι ιδιοτιμές έρχονται πάντα σε ζεύγη. Συνήθως η ίδια βιβλιοθήκη που υλοποιεί τα ιδιοδιανύσματα, υπολογίζει και τις ιδιοτιμές τους.

Μέθοδος PCA

Βήμα πρώτο : Λαμβάνονται τα δεδομένα

Στο παράδειγμα που ακολουθεί θα γίνει χρήση απλών δεδομένων δύο διαστάσεων, προκειμένου να μπορούν να γίνουν γραφικές παραστάσεις ώστε να φανεί τι κάνει η μέθοδος PCA σε κάθε βήμα. Τα δεδομένα του παραδείγματος φαίνονται στον επόμενο πίνακα. Η αριστερή στήλη έχει τα αρχικά δεδομένα και η δεξιά αφού αφαιρεθεί ο μέσος όρος. Ακολουθεί γραφική παράσταση.

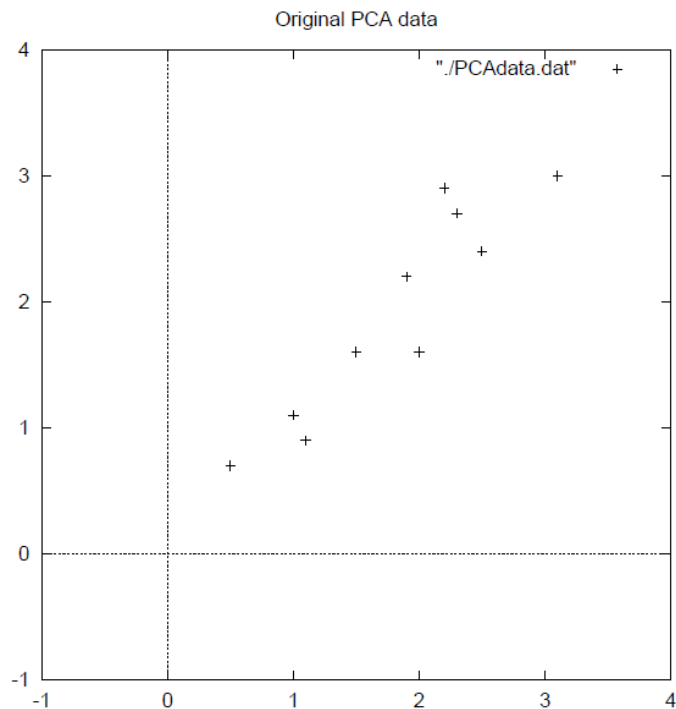
Βήμα δεύτερο : Αφαίρεση του μέσου

Για να δουλέψει η μέθοδος σωστά πρέπει να αφαιρεθεί ο μέσος από κάθε διάσταση των δεδομένων.

Ο μέσος που αφαιρείται είναι ο μέσος όρος κάθε διάστασης. Έτσι από όλες τις τιμές έχουν αφαιρεθεί (Ο μέσος όρος των τιμών των δεδομένων) από όλες τις τιμές. Αυτό παράγει ένα σύνολο δεδομένων που έχει μέσο μηδέν.

x		y		
Data =	2.5	2.4	.69	.49
	0.5	0.7	-1.31	-1.21
	2.2	2.9	.39	.99
	1.9	2.2	.09	.29
	3.1	3.0	1.29	1.09
	2.3	2.7	.49	.79
	2	1.6	.19	-.31
	1	1.1	-.81	-.81
	1.5	1.6	-.31	-.31
	1.1	0.9	-.71	-1.01
		x		
		y		

Πίνακας 5 Αριστερά: αρχικά δεδομένα , δεξιά: αφού αφαιρέθηκε ο μέσος όρος



Διάγραμμα 21 γραφική απεικόνιση των παραπάνω αρχικών δεδομένων.

Βήμα τρίτο: Υπολογισμός του πίνακα συνδιακύμανσης

Αφού τα δεδομένα είναι δύο διαστάσεων ο πίνακας θα είναι 2 x2. Δίνετε απευθείας το αποτέλεσμα. :

$$cov = \begin{pmatrix} .616555556 & .615444444 \\ .615444444 & .716555556 \end{pmatrix}$$

Έτσι αφού τα μη διαγώνια στοιχεία το πίνακα είναι θετικά αναμενόταν οι μεταβλητές στον X και Y άξονα να μεγαλώνουν ταυτόχρονα.

Βήμα τέταρτο : Υπολογισμός των ιδιοδιανυσμάτων και ιδιοτιμών του πίνακα συνδιακύμανσης

Αφού ο πίνακας συνδιακύμανσης είναι τετραγωνικός μπορεί να υπολογιστούν οι ιδιοτιμές και τα ιδιοδιανύσματα του. Είναι σημαντικά, αφού περιέχουν σημαντική πληροφορία για τα δεδομένα. Παρατίθενται τα ιδιοδιανύσματα και οι ιδιοτιμές.

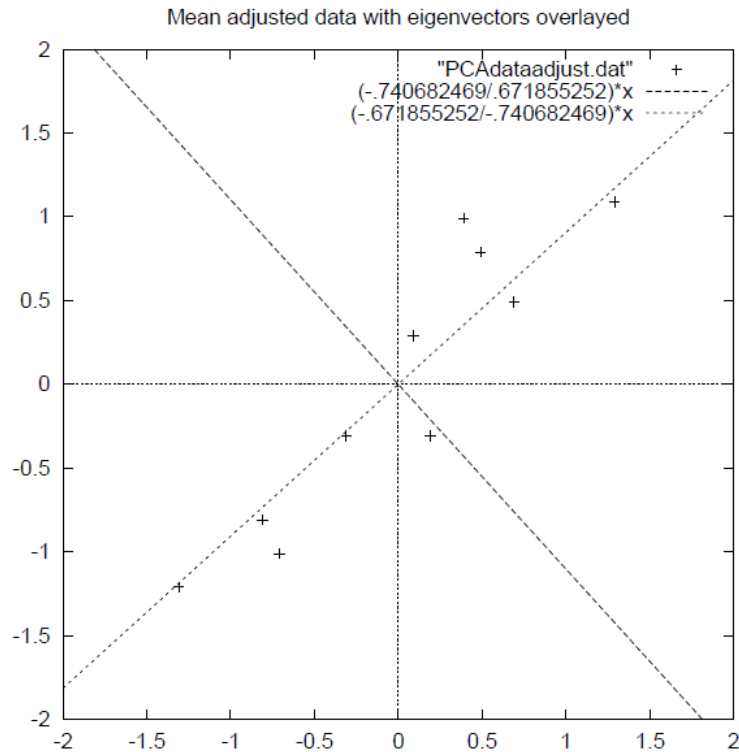
$$eigenvalues = \begin{pmatrix} .0490833989 \\ 1.28402771 \end{pmatrix}$$

$$eigenvectors = \begin{pmatrix} -.735178656 & -.677873399 \\ .677873399 & -.735178656 \end{pmatrix}$$

Ας σημειωθεί εδώ ότι τα ιδιοδιανύσματα είναι και τα δύο μοναδιαία, με μήκος 1. Αυτό είναι πολύ σημαντικό για την υλοποίηση της PCA, τα περισσότερα πακέτα μαθηματικών επιστρέφουν ως λύσεις μοναδιαία ιδιοδιανύσματα. Προκύπτει ότι τα δεδομένα ακολουθούν σε μεγάλο βαθμό ένα πρότυπο. Όπως αναμενόταν από τον πίνακα συνδιακύμανσης, οι δύο μεταβλητές μεγαλώνουν ταυτόχρονα. Πάνω στα δεδομένα έχουν αναπαρασταθεί και τα ιδιοδιανύσματα. Εμφανίζονται ως διακεκομμένες γραμμές διαγώνιες πάνω στην γραφική παράσταση. Τα ιδιοδιανύσματα, είναι κάθετα το ένα στο άλλο. Αλλά το σημαντικότερο, παρέχουν πληροφορία σχετικά με τα πρότυπα των δεδομένων. Το ένα ιδιοδιάνυσμα περνάει από το μέσο των δεδομένων, σα να πρόκειται για γραμμή βέλτιστης προσέγγισης. Αυτό το ιδιοδιάνυσμα δείχνει πως συσχετίζονται τα δεδομένα γύρω από τη γραμμή. Το δεύτερο ιδιοδιάνυσμα δίνει την λιγότερο σημαντική πληροφορία, πως όλα τα δεδομένα ακολουθούν την κύρια γραμμή αλλά απέχουν από την γραμμή κατά κάποιο ποσό. Έτσι από την διαδικασία με τα ιδιοδιανύσματα και τον πίνακα συνδιακύμανσης υπολογίστηκαν γραμμές που χαρακτηρίζουν τα δεδομένα. Τα υπόλοιπα βήματα περιλαμβάνουν μετασχηματισμούς των δεδομένων ώστε να εκφράζονται σε σχέση με τις γραμμές αυτές.

Βήμα πέμπτο: Επιλέγονται οι συνιστώσες που σχηματίζουν το διάνυσμα χαρακτηριστικών.

Εδώ φαίνεται πως συνδέεται η έννοια της συμπίεσης δεδομένων και της ελαττωμένης διαστασιμότητας . Αν εξεταστούν τα ιδιοδιανύσματα και οι ιδιοτιμές από την προηγούμενη παράγραφο ,



Εικόνα: Διάγραμμα των κανονικοποιημένων (αφαιρέθηκε ο μέσος όρος) με το ιδιοδιάνυσμα του πίνακα συνδιακύμανσης σχεδιασμένο από πάνω.

Οι ιδιοτιμές είναι ιδιαίτερα διαφορετικές ως τιμές. Στην πραγματικότητα αποδεικνύεται ότι το ιδιοδιάνυσμα με την *μεγαλύτερη* ιδιοτιμή είναι η κυριότερη συνιστώσα *principle component* του συνόλου των δεδομένων. Στο παράδειγμα το ιδιοδιάνυσμα με την μεγαλύτερη ιδιοτιμή ήταν αυτό που υποδεικνύει το μέσο των δεδομένων. Είναι η σημαντικότερη σχέση μεταξύ των διαστάσεων των δεδομένων. Γενικά, αφού υπολογιστούν τα ιδιοδιανύσματα του πίνακα της συνδιακύμανσης ταξινομούνται με κριτήριο την τιμή της ιδιοτιμής τους, από την υψηλότερη προς την χαμηλότερη.

Αυτό δίνει τις συνιστώσες με σειρά σημαντικότητας. Είναι δυνατόν να απορριφθούν οι συνιστώσες χαμηλότερης σημασίας και αν αυτές έχουν μικρές ιδιοτιμές η πληροφορία που χάνουμε είναι σχετικά μικρή. Αγνοώντας κάποιες συνιστώσες μειώνονται οι διαστάσεις των δεδομένων. Αν αρχικά υπήρχαν n διαστάσεις στα δεδομένα, και υπολογιστούν p ιδιοδιανύσματα και ιδιοτιμές, τότε επιλέγονται μόνο τα πρώτα p ιδιοδιανύσματα, τελικά το τελικό σύνολο δεδομένων θα έχει μόνο p διαστάσεις. Αυτό που μένει να γίνει είναι να σχηματίσουμε ένα διάνυσμα χαρακτηριστικών, που είναι απλά ένας πίνακας διανυσμάτων. Κατασκευάζεται με στοιχεία στήλες τα ιδιοδιανύσματα που θα διατηρηθούν.

$$FeatureVector = (eig_1 \ eig_2 \ eig_3 \ \ eig_n)$$

Στο παράδειγμα υπάρχουν 2 ιδιοδιανύσματα και μπορεί να δημιουργηθεί ένας πίνακας 2 x 2 ή ένα πίνακα διάνυσμα 2 x1 :

$$\begin{pmatrix} -.677873399 & -.735178656 \\ -.735178656 & .677873399 \end{pmatrix} \begin{pmatrix} -.677873399 \\ -.735178656 \end{pmatrix}$$

Παρακάτω θα φανεί το αποτέλεσμα της κάθε επιλογής.

Βήμα έκτο: Σχηματίζοντας το νέο σύνολο δεδομένων

Είναι το τελευταίο βήμα της PCA και είναι και το απλούστερο. Αφού επιλεγούν οι συνιστώσες (ιδιοδιανύσματα) που θα διατηρηθούν και σχηματιστεί το διάνυσμα χαρακτηριστικών, λαμβάνεται ο ανάστροφος του διανύσματος και πολλαπλασιάζεται από τα αριστερά των αρχικών δεδομένων, αφού αυτά αναστραφούν.

$$FinalData = RowFeatureVector \times RowDataAdjust,$$

Όπου Row Feature Vector είναι ο πίνακας με τα ιδιοδιανύσματα στις στήλες ανεστραμμένος έτσι ώστε τα ιδιοδιανύσματα να είναι σε γραμμές και αυτό με την μεγαλύτερη σημασία (ιδιοτιμή) να βρίσκεται στην κορυφή, και Row Data Adjust είναι τα μέσο κανονικοποιημένα δεδομένα ανεστραμμένα. Για παράδειγμα τα δεδομένα είναι σε κάθε στήλη, με κάθε γραμμή να διατηρεί μια ξεχωριστή διάσταση.

Τα **τελικά δεδομένα** είναι το τελικό σύνολο δεδομένων, με τα δεδομένα σε στήλες και τις διαστάσεις σε γραμμές. Τι δίνει αυτό; Δίνει τα αρχικά δεδομένα εκφρασμένα αποκλειστικά *συναρτήσει των διανυσμάτων που διατηρήθηκαν*. Τα αρχικά δεδομένα ήταν εκφρασμένα ως προς X,Y αφού αυτοί ήταν οι άξονες. Είναι εφικτό να εκφραστούν τα ίδια δεδομένα ως προς οποιουδήποτε άξονες χρειάζεται. Αν αυτοί οι άξονες είναι κάθετοι μεταξύ τους, τότε η έκφραση είναι η πιο αποδοτική. Αυτό εξηγεί γιατί ήταν σημαντικό τα ιδιοδιανύσματα να είναι κάθετα πάντα μεταξύ τους. Άλλαξε έτσι την έκφραση ως προς X,Y σε έκφραση ως προς τα δύο ιδιοδιανύσματα που επιλέχθηκαν και το νέο σύνολο δεδομένων έχει ελαττωμένη διαστασιμότητα. Για να φανεί αυτό στα δεδομένα έγινε ο τελικός μετασχηματισμός με καθένα από τα δύο πιθανά διανύσματα χαρακτηριστικών. Υπολογίστηκε ο ανάστροφος πίνακας των τελικών δεδομένων ώστε να εμφανιστεί με την ίδια μορφή πίνακα με τα αρχικά δεδομένα. Επίσης έγινε η γραφική παράσταση του για να γίνει συσχέτιση με τις συνιστώσες στην περίπτωση που διατηρήθηκαν και τα δυο ιδιοδιανύσματα προκύπτουν τα δεδομένα και η γραφική 1. Αυτή η γραφική περιέχει τα αρχικά δεδομένα στραμμένα έτσι ώστε τα ιδιοδιανύσματα να είναι οι νέοι άξονες.

Αυτό είναι κατανοητό αφού δεν υπάρχει χαμένη πληροφορία κατά την αναπαράσταση αυτή.

Ο άλλος μετασχηματισμός που μπορεί να γίνει είναι να ληφθεί υπόψη μας μόνο το ιδιοδιάνυσμα με την μεγαλύτερη ιδιοτιμή. Τα αποτελέσματα του φαίνονται στον πίνακα 2. Όπως ήταν αναμενόμενο έχει μόνο μία διάσταση. Αν συγκριθούν τα δεδομένα αυτά με

τα δεδομένα του μετασχηματισμού 2 διαστάσεων θα διαπιστωθεί ότι είναι ακριβώς η πρώτη στήλη από τις δύο.

Έτσι αν γίνει η γραφική παράσταση θα είναι μονοδιάστατη και όλα τα στοιχεία θα βρίσκονται κατά μήκος του X άξονα και στα σημεία που βρίσκονταν και στο προηγούμενο διάγραμμα (ως προς το X άξονα της γραφικής 1) και έχει παραληφθεί εντελώς ο άξονας Y που αντιστοιχεί στο δεύτερο ιδιοδιάνυσμα.

Οπότε τι έχει επιτευχθεί εδώ; Βασικά έχουν μετασχηματιστεί τα δεδομένα έτσι ώστε να εκφράζονται ως προς τις μεταξύ τους πρότυπα, όπου τα πρότυπα είναι οι γραμμές που καλύτερα προσεγγίζουν τις σχέσεις μεταξύ των δεδομένων.

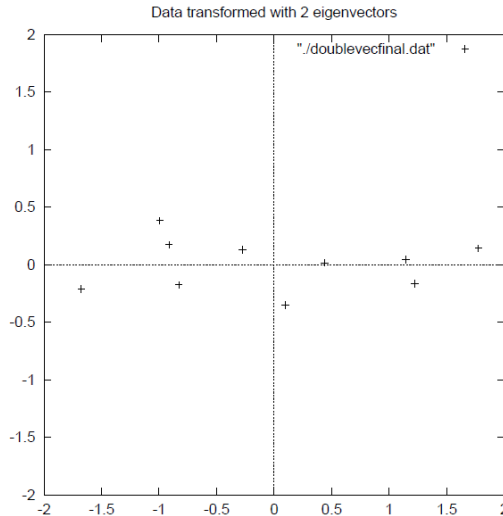
Αυτό βοηθάει γιατί έχουμε να κατηγοριοποιήσουμε τα σημεία ως ένα συνδυασμό της συνεισφοράς της κάθε γραμμής. Αρχικά υπήρχαν οι απλοί άξονες X και Y. Οι τιμές στους X και Y άξονες των δεδομένων δεν παρέχουν πληροφορία για τα μεταξύ των στοιχείων πρότυπα. Τώρα οι τιμές των δεδομένων περιγράφουν που ακριβώς βρίσκονται (πάνω ή κάτω) οι γραμμές τάσης (ή οι γραμμές ελαχίστων τετραγώνων) τα στοιχεία δεδομένων. Στην περίπτωση του μετασχηματισμού με τα δύο ιδιοδιανύσματα, έχουν απλά μεταβληθεί τα δεδομένα ώστε να είναι εκφρασμένα ως προς τα ιδιοδιανύσματα αντί για τους συνηθισμένους άξονες. Αλλά η ανάλυση με το ένα ιδιοδιάνυσμα έχει αφαιρέσει την συνεισφορά του άξονα με την μικρότερη πληροφορία και το σύνολο των δεδομένων εκφράζεται σε σχέση μόνο του πρώτου ιδιοδιανύσματος.

Ανάκτηση των αρχικών δεδομένων

Η ανάκτηση των αρχικών δεδομένων είναι μεγάλης σημασίας ειδικά σε περιπτώσεις όπως η χρήση του μετασχηματισμού PCA για συμπίεση δεδομένων. Το παρακάτω περιεχόμενο παρέχεται από

<http://www.vision.auc.dk/sig/Teaching/Flerdim/Current/hotelling/hotelling.html>

	<i>x</i>	<i>y</i>
Transformed Data=	-.827970186	-.175115307
	1.77758033	.142857227
	-.992197494	.384374989
	-.274210416	.130417207
	-1.67580142	-.209498461
	-.912949103	.175282444
	.0991094375	-.349824698
	1.14457216	.0464172582
	.438046137	.0177646297
	1.22382056	-.162675287



Εικόνα 11: Πίνακας δεδομένων με εφαρμογή PCA και χρήση 2 ιδιοδιανυσμάτων

Και γραφική απεικόνιση αυτών των σημείων στους νέους άξονες(ιδιοδιανύσματα).

Transformed Data (Single eigenvector)

x
-.827970186
1.77758033
-.992197494
-.274210416
-1.67580142
-.912949103
.0991094375
1.14457216
.438046137
1.22382056

Εικόνα 12: Δεδομένα μετά από μετασχηματισμό PCA με διατήρηση μόνο του πρώτου ιδιοδιανύσματος.

Πώς θα ανακτηθούν τα αρχικά δεδομένα? Πριν γίνει αυτό, υπενθυμίζεται ότι ΜΟΝΟ ΑΝ έχουν διατηρηθεί όλα τα ιδιοδιανύσματα κατά το μετασχηματισμό μπορεί να ληφθούν ακριβώς τα ίδια αρχικά δεδομένα. Αν αφαιρεθεί τμήμα των ιδιοδιανυσμάτων στον τελικό μετασχηματισμό, τότε τα ανακτημένα δεδομένα θα έχουν μερική απώλεια πληροφορίας. Υπενθυμίζεται ότι ο τελικός μετασχηματισμός είναι :

$$FinalData = RowFeatureVector \times RowDataAdjust,$$

Που μπορεί να αντιστραφεί έτσι ώστε να ανακτήσουμε τα αρχικά δεδομένα.

$$RowDataAdjust = RowFeatureVector^{-1} \times FinalData$$

Όπου $RowFeatureVector^{-1}$ είναι ο αντίστροφος του $\square RowFeatureVector$. Ωστόσο, όταν λαμβάνονται όλα τα ιδιοδιανύσματα στον σχηματισμού του διανύσματος

χαρακτηριστικών, αποδεικνύεται ότι ο αντίστροφος του διανύσματος χαρακτηριστικών είναι ίσος με τον ανάστροφο του διανύσματος αυτού. Αυτό είναι αληθές μόνο επειδή τα στοιχεία του πίνακα είναι όλα τα μοναδιαία ιδιοδιανύσματα του συνόλου δεδομένων. Αυτό κάνει την διαδικασία επιστροφής στο αρχικό σύνολο πολύ πιο εύκολη επειδή η εξίσωση μπορεί να γίνει:

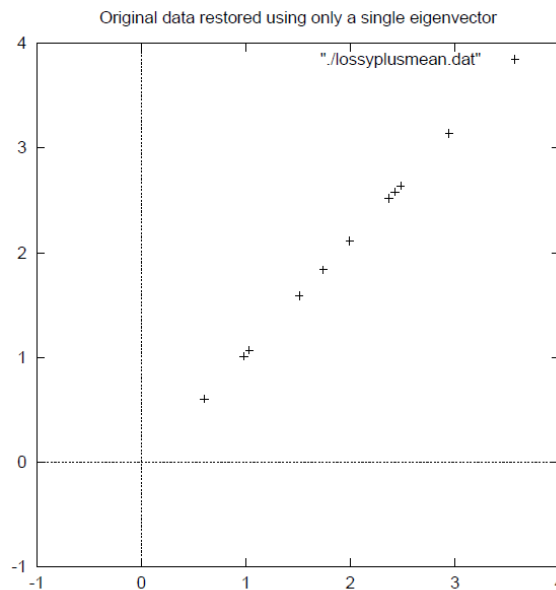
$$RowDataAdjust = RowFeatureVector^T \times FinalData$$

Αλλά για να προκύψουν τα πραγματικά αρχικά δεδομένα πίσω πρέπει να προστεθεί ο μέσος αυτών των αρχικών δεδομένων. Έτσι ο τύπος γίνεται τελικά:

$$RowOriginalData = (RowFeatureVector^T \times FinalData) + OriginalMean$$

Αυτή η εξίσωση εφαρμόζεται επίσης και όταν δεν έχουν χρησιμοποιήσει όλα τα ιδιοδιανύσματα για το σχηματισμό του διανύσματος χαρακτηριστικών. Ο τύπος παράγει το σωστό αποτέλεσμα αλλά τα αρχικά δεδομένα ανακτώνται μερικώς μόνο. Παρακάτω υπολογίζονται τα αρχικά δεδομένα χρησιμοποιώντας το πλήρες διάνυσμα χαρακτηριστικών και το ελαττωμένο διάνυσμα χαρακτηριστικών για να γίνει ορατή η ποσότητα της πληροφορίας που χάθηκε.

Στην εικόνα 6 φαίνεται η αντίστοιχη γραφική παράσταση



Εικόνα 13: Ανακατασκευασμένα αρχικά δεδομένα με τη χρήση του ελαττωμένου διανύσματος χαρακτηριστικών.

Συγκρίνοντας αυτή την γραφική παράσταση σε σχέση με την αρχική γραφική παρατηρούμε ότι η διακύμανση κατά τον άξονα της κύρια συνιστώσας εξακολουθεί να υφίσταται, η διακύμανση κατά τον άξονα της δεύτερης συνιστώσας έχει χαθεί.