

Ατομική Διπλωματική Εργασία

**ΜΕΛΕΤΗ ΤΕΧΝΙΚΩΝ ΑΝΑΓΝΩΡΙΣΗΣ ΚΑΙ ΑΝΙΧΝΕΥΣΗΣ
ΠΡΟΣΩΠΩΝ ΚΑΙ ΥΛΟΠΟΙΗΣΗ ΕΦΑΡΜΟΓΩΝ ΑΝΙΧΝΕΥΣΗΣ
ΠΡΟΣΩΠΩΝ ΓΙΑ ΤΟ ΣΥΣΤΗΜΑ BIGBLUEBUTTON**

Μαρία Μιχαήλ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ



ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

Μάιος 2010

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

ΜΕΛΕΤΗ ΤΕΧΝΙΚΩΝ ΑΝΑΓΝΩΡΙΣΗΣ ΚΑΙ ΑΝΙΧΝΕΥΣΗΣ
ΠΡΟΣΩΠΩΝ ΚΑΙ ΥΛΟΠΟΙΗΣΗ ΕΦΑΡΜΟΓΩΝ ΑΝΙΧΝΕΥΣΗΣ
ΠΡΟΣΩΠΩΝ ΓΙΑ ΤΟ ΣΥΣΤΗΜΑ BIGBLUEBUTTON

Μαρία Μιχαήλ

Επιβλέπων καθηγητής
Δρ. Γιώργος Παπαδόπουλος

Η Ατομική Διπλωματική Εργασία υποβλήθηκε προς μερική εκπλήρωση των απαιτήσεων απόκτησης του πτυχίου Πληροφορικής του Τμήματος Πληροφορικής του Πανεπιστημίου Κύπρου

Μάιος 2010

Ευχαριστίες

Η διπλωματική εργασία αποτελεί μέρος ολοκλήρωσης του προπτυχιακού μου προγράμματος Γενικής Κατεύθυνσης Πληροφορικής. Κατ' αρχάς θα ήθελα να εκφράσω τις ευχαριστίες μου στον κ. Γιώργο Παπαδόπουλο για την εμπιστοσύνη που μου έδειξε, την καθοδήγηση του, και τις άψογες συνθήκες συνεργασίας μαζί του, αλλά και την ευκαιρία που μου έδωσε να ασχοληθώ με ένα τόσο ενδιαφέρον αντικείμενο μελέτης.

Αξίζει να δώσω τις ευχαριστίες μου στις φίλες μου Μαρία και Μαρίνα για τη στήριξη που μου έδιναν κατά τη διάρκεια εκπόνησης της πτυχιακής μου εργασίας, αλλά και καθ' όλη τη διάρκεια των σπουδών μου στο Πανεπιστήμιο Κύπρου.

Και πάνω απ' όλα, ένα μεγάλο ευχαριστώ στους γονείς μου για την ολόψυχή τους αγάπη, την συμπαράσταση και την κατανόηση καθώς επίσης και για την ηθική και υλική υποστήριξη τους όλα αυτά τα χρόνια.

Η παρούσα πτυχιακή αφιερώνεται στη μνήμη της αδερφής μου Χλόης Μιχαήλ.

Περίληψη

Κατά τη διάρκεια των τελευταίων δέκα ετών, η ανίχνευση και αναγνώριση προσώπου έχουν γίνει ένας δημοφιλής τομέας έρευνας στην όραση υπολογιστών και από τις πιο επιτυχημένες εφαρμογές. Λόγω της φύσης του προβλήματος, ενδιαφέρον δείχνουν όχι μόνο οι ερευνητές πληροφορικής, αλλά και επιστήμονες νευρολογίας και ψυχολογίας επίσης.

Η έρευνα στην ανάλυση εκφράσεων έχει έλθει ξανά στο προσκήνιο τα τελευταία χρόνια. Οι μέθοδοι αυτόματης ανάλυσης εκφράσεων αποτελούν μια ενεργή περιοχή έρευνας που βρίσκει προοπτικές εφαρμογές στην γενικότερη ερευνητική περιοχή της αλληλεπίδρασης ανθρώπου - υπολογιστή. Οι εκφράσεις του προσώπου αντανακλούν όχι μόνο ανθρώπινα συναισθήματα αλλά και νοητικές λειτουργίες και ψυχολογικά σήματα και κατέχουν έτσι ένα μεγάλο μέρος της κοινωνικής αλληλεπίδρασης. Αυτός είναι ο λόγος για τον οποίο η αυτόματη αναγνώριση εκφράσεων αποτελεί βασικό τομέα της αλληλεπίδρασης ανθρώπου - υπολογιστή.

Σκοπός αυτής της διπλωματικής εργασίας είναι η παρουσίαση ενός συστήματος ανίχνευσης και παρακολούθησης προσώπου, καθώς επίσης και αυτόματης αναγνώρισης συναισθημάτων (χαρά, θυμό και έκπληξη) καθώς το πρόσωπο θα βρίσκεται μπροστά από μια κάμερα. Τέτοιου είδους συστήματα μπορούν μελλοντικά να αποτελέσουν τμήμα μιας πληθώρας εφαρμογών που θα αναβαθμίσουν την ποιότητα πολλών δραστηριοτήτων.

Περιεχόμενα

Κεφάλαιο 1	Εισαγωγή.....	7
1.1	Γενικά	7
1.2	Σκοπός της Διπλωματικής Εργασίας	7
1.3	Δομή της Διπλωματικής Εργασίας	8
Κεφάλαιο 2	Αναγνώριση προσώπων.....	10
2.1	Εισαγωγή	10
2.2	Προβλήματα αυτοματοποίησης αναγνώρισης προσώπων	11
2.3	Εφαρμογές	13
Κεφάλαιο 3	Ανίχνευση προσώπων.....	16
3.1	Γενικά	16
3.2	Προκλήσεις στη διαδικασία ανίχνευσης προσώπων	17
3.3	Ιστορικό των κυριότερων μεθόδων Ανίχνευσης προσώπων	18
3.4	Μέθοδοι ανίχνευσης/εντοπισμού Προσώπου	20
3.5	Μέθοδοι βασισμένες στην εμφάνιση	24
3.5.1	Δείγματα προσώπων και μη προσώπων	24
3.5.2	Ταξινομητές	26
3.5.3	Αναπαράσταση προσώπων	27
3.5.4	Προεπεξεργασία	27
3.5.5	Μετεπεξεργασία	28
Κεφάλαιο 4	Ανίχνευση προσώπου κατά Viola & Jones.....	30
4.1	Περιγραφή του προβλήματος	31
4.2	Ολοκληρωτική εικόνα	31
4.3	Χαρακτηριστικά Haar	33
4.4	Εξαγωγή χαρακτηριστικών με τον αλγόριθμο Adaboost	35
4.5	Ταξινομητής Cascade	38

Κεφάλαιο 5	Ανίχνευση εκφράσεων προσώπου.....	40
5.1	Εισαγωγή	40
5.2	Καθορισμός του προβλήματος	41
5.3	Τεχνικές αναγνώρισης των εκφράσεων	42
5.4	Εφαρμογές	43
Κεφάλαιο 6	Υλοποίηση συστήματος	45
6.1	Γενικά	45
6.2	Βιβλιοθήκες που έχουν ελεγχθεί	46
6.3	Kinect SDK 1.5	46
6.4	Βιβλιοθήκη OpenCV	51
6.5	Βιβλιοθήκη BRF	52
6.6	Εφαρμογή 1	52
6.7	Εφαρμογή 2	53
6.7.1	Αναγνώριση χαμόγελου	56
6.7.2	Πώς να αποφασίσουμε ένα σωστό κατώφλι	57
6.7.3	Αναγνώριση στρογγυλού σχήματος χειλιών	58
6.7.4	Αναγνώριση ανασήκωσης φρυδιών	58
6.7.5	Αναγνώριση συνοφρύωσης φρυδιών	59
6.8	Αποτελέσματα	61
6.9	Προβλήματα που αντιμετωπίσαμε	61
6.9.1	Περιορισμός του προσώπου εντός ενός πλαισίου	61
6.9.2	Διαφορετικότητα στα πρόσωπα	61
6.9.3	Διαφορετικότητα στις εκφράσεις των προσώπων	62
Κεφάλαιο 7	Μελλοντική Εργασία.....	62
7.1	Επίλογος	62
Βιβλιογραφία		64

Κεφάλαιο 1

Εισαγωγή

1.1 Γενικά	1
1.2 Σκοπός Διπλωματικής Εργασίας	4
1.3 Δομή Διπλωματικής Εργασίας	10

1.1 Γενικά

Η χρήση της τεχνολογίας σε όλους σχεδόν τους τομείς της ζωής, από την ιατρική μέχρι την ψυχαγωγία, έχει καταστήσει αναγκαία την ανάπτυξη έξυπνων και φιλικών προς το χρήστη εφαρμογών, γεγονός που οδήγησε την επιστημονική κοινότητα στην αναζήτηση μεθόδων βελτίωσης του συστήματος επικοινωνίας ανθρώπου-υπολογιστή. Το ιδανικό μοντέλο για το σχεδιασμό ενός τέτοιου συστήματος δεν είναι άλλο από αυτό της επικοινωνίας μεταξύ των ανθρώπων, η οποία περιλαμβάνει πολλά περισσότερα από την απλή μετάδοση ενός μηνύματος διαμέσου ήχου. Η ανθρώπινη επικοινωνία βασίζεται σε ένα πλήθος μέσων έκφρασης, καθοριστικών για την κατανόηση της συναισθηματικής κατάστασης και των προθέσεων του συνομιλητή, όπως είναι κυρίως ο τόνος της φωνής και οι εκφράσεις του προσώπου, οι οποίες θα μας απασχολήσουν στη συγκεκριμένη διπλωματική εργασία. Συστήματα που θα είναι σε θέση να αναγνωρίζουν τις εκφράσεις στο πρόσωπο του χρήστη μπορούν να αποτελέσουν μέρος ενός διαδραστικού περιβάλλοντος που θα κάνει πιο άμεση και <<φυσική>> την αλληλεπίδραση ανθρώπου-μηχανής.

Είναι σημαντικό να αναφέρουμε ότι η ανίχνευση προσώπων (Face Detection) γίνεται ολοένα και πιο σημαντική σε πολλές εφαρμογές και έχει αποδοθεί μεγάλη σημασία σε τέτοιες εφαρμογές τα τελευταία χρόνια. Ένα από τα πιο ενδιαφέροντα χαρακτηριστικά του ανθρώπινου συστήματος όρασης είναι η ικανότητα να ανιχνεύει και να εντοπίζει πρόσωπα στο οπτικό πεδίο του ματιού. Αποτελεί το πρώτο βήμα στη διαδικασία αναγνώρισης προσώπου κατά το οποίο θα προσδιοριστεί η ταυτότητά του, θα εντοπιστούν τα επιμέρους χαρακτηριστικά του προσώπου του και θα βρεθεί η θέση και η κατεύθυνση του στον τρισδιάστατο χώρο. Έτσι, μπορεί επίσης να ξεκινήσει μια παρακολούθηση του προσώπου ή

και των επιμέρους χαρακτηριστικών του στο χρόνο, ώστε να προσδιοριστεί η συνολική του κίνηση ή οι επιμέρους κινήσεις των χαρακτηριστικών του για εφαρμογές στην επίβλεψη σκηνών, ή στην αναγνώριση ομιλίας και των ανθρώπινων εκφράσεων. Η αξιοπιστία του συστήματος ανίχνευσης παίζει κυρίαρχο ρόλο στην απόδοση και χρησιμότητα ολόκληρου του συστήματος αναγνώρισης προσώπου.

1.2 Σκοπός Διπλωματικής Εργασίας

Σκοπός της παρούσας διπλωματικής εργασίας είναι ο σχεδιασμός και η αξιολόγηση ενός συστήματος ανίχνευσης προσώπου και αναγνώρισης ανθρώπινων εκφράσεων. Πρόκειται για μία εφαρμογή πραγματικού χρόνου μέσα από την οποία παρακολουθούμε τις κινήσεις του προσώπου ενός ατόμου που βρίσκεται μπροστά από μια κάμερα και ανιχνεύουμε αρχικά τα επιμέρους χαρακτηριστικά του (μάτια, μύτη, στόμα). Στη συνέχεια, επεκτείνουμε τη λειτουργία της εφαρμογής αναγνωρίζοντας κάποιες από τις ανθρώπινες εκφράσεις του προσώπου καθώς επίσης και ορισμένες λέξεις.

Ξεκινώντας από την εισαγωγή στατικών εικόνων, επεκτείναμε την εφαρμογή σε πραγματικό χρόνο ώστε να γίνει πιο δημιουργική και ταυτόχρονα πιο ενδιαφέρουσα. Στόχος της συγκεκριμένης διπλωματικής εργασίας είναι να δείξουμε την εξέλιξη και πως μπορεί να χρησιμοποιηθεί και να εφαρμοστεί μελλοντικά σε μεγαλύτερα συστήματα εφαρμογών. Το CVN project που βασίζεται στο BigBlueButton έχει κατασκευάσει μια συσκευή και έχει υλοποιήσει ένα σύστημα επικοινωνίας το οποίο βασίζεται στην αληθοφάνεια της επικοινωνίας. Επομένως, συγκεκριμένη εφαρμογή μπορεί να είναι άμεσα αξιοποιήσιμη σε ένα τέτοιο σύστημα.

1.3 Δομή Διπλωματικής Εργασίας

Η παρούσα πτυχιακή εργασία αποτελείται από επτά κεφάλαια.

Στο κεφάλαιο 2 γίνεται αναφορά στην αναγνώριση προσώπων όπου περιγράφονται τα προβλήματα αυτοματοποίησης των συγκεκριμένων εφαρμογών, καθώς και πιθανές χρήσεις εφαρμογών στο μέλλον.

Στο κεφάλαιο 3 περιγράφεται λεπτομερώς η ανίχνευση και παρακολούθηση προσώπων η οποία καλύπτει και το μεγαλύτερο μέρος της υλοποίησης του συστήματος. Αναλύονται εκτενέστερα οι μέθοδοι ανίχνευσης και εντοπισμού των προσώπων, παρουσιάζεται μία ιστορική αναδρομή και γίνεται αναφορά στους διάφορους ταξινομητές που χρησιμοποιούνται.

Στο κεφάλαιο 4 γίνεται αναφορά στον αλγόριθμο Viola & Jones ο οποίος χρησιμοποιείται μέσα από τη βιβλιοθήκη OpenCv που επιλέξαμε για την υλοποίηση του συστήματος. Η συγκεκριμένη μέθοδος ουσιαστικά σαρώνει την εικόνα εισόδου σε όλες τις πιθανές θέσεις και κλίμακες και στη συνέχεια ταξινομεί τα παράθυρα ως πρόσωπα και μη-πρόσωπα. Επιπλέον παρουσιάζονται οι γνωστοί αλγόριθμοι Adaboost και Cascade οι οποίοι αναλύονται περαιτέρω.

Στο κεφάλαιο 5 γίνεται αναφορά στην ανίχνευση ανθρώπινων εκφράσεων προσώπου που αποτελεί σημαντικό μέρος της υλοποίησης του συστήματος και παρουσιάζονται κάποιες εφαρμογές στις οποίες θα μπορούσε μελλοντικά να χρησιμοποιηθεί.

Στο κεφάλαιο 6 περιγράφεται λεπτομερώς η υλοποίηση της εφαρμογής και η λειτουργία της, ενώ παρουσιάζονται επίσης και οι οθόνες (screenshots) που παρουσιάζονται στο χρήστη κατά τη χρήση της εφαρμογής.

Στο κεφάλαιο 7 παρουσιάζονται οι μελλοντικές εργασίες στις οποίες θα μπορούσε να χρησιμοποιηθεί η εφαρμογή που έχει υλοποιηθεί.

Κεφάλαιο 2

ΑΝΑΓΝΩΡΙΣΗ ΠΡΟΣΩΠΩΝ

2.1 Εισαγωγή

2.2 Προβλήματα αυτοματοποίησης αναγνώρισης προσώπων

2.3 Εφαρμογές

2.1 Εισαγωγή

Η αναγνώριση προσώπων έχει γνωρίσει τα τελευταία χρόνια ιδιαίτερη ανάπτυξη, γεγονός που οφείλεται κυρίως στο μεγάλο εύρος των εφαρμογών που χρησιμοποιούν συστήματα αναγνώρισης προσώπων και τη διαθεσιμότητα κατάλληλης τεχνολογίας μετά από πολλά χρόνια έρευνας. Το πρόβλημα της αναγνώρισης εικόνων του ανθρώπινου προσώπου από τον υπολογιστή εξακολουθεί να προσελκύει ερευνητές που δραστηριοποιούνται στις περιοχές της επεξεργασίας εικόνας, της αναγνώρισης προτύπων, των νευρωνικών δικτύων, της υπολογιστικής όρασης των γραφικών του υπολογιστή, κ.α. Λόγω της φύσης του προβλήματος, ενδιαφέρον δείχνουν όχι μόνο οι ερευνητές πληροφορικής, αλλά και επιστήμονες νευρολογίας και ψυχολογίας επίσης.

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

Το πρόσωπο είναι ένα μοναδικό χαρακτηριστικό για τα ανθρώπινα όντα. Ακόμα και τα πρόσωπα δύο ομοζυγωτικών διδύμων διαφέρουν. Εντοπίζουμε, αναγνωρίζουμε και διακρίνουμε τα πρόσωπα με χαρακτηριστική ευκολία και απλή παρατήρηση. Η ευκολία αναγνώρισης των προσώπων από τον άνθρωπό ήταν ένας από τους λόγους για την επιλογή γραμματοσήμων τα οποία περιλάμβαναν πρόσωπα: ήταν πολύ δυσκολότερο να γίνουν πλαστογραφίες γραμματοσήμων γιατί αυτές γίνονταν εύκολα αντιληπτές. Η μοναδικότητα

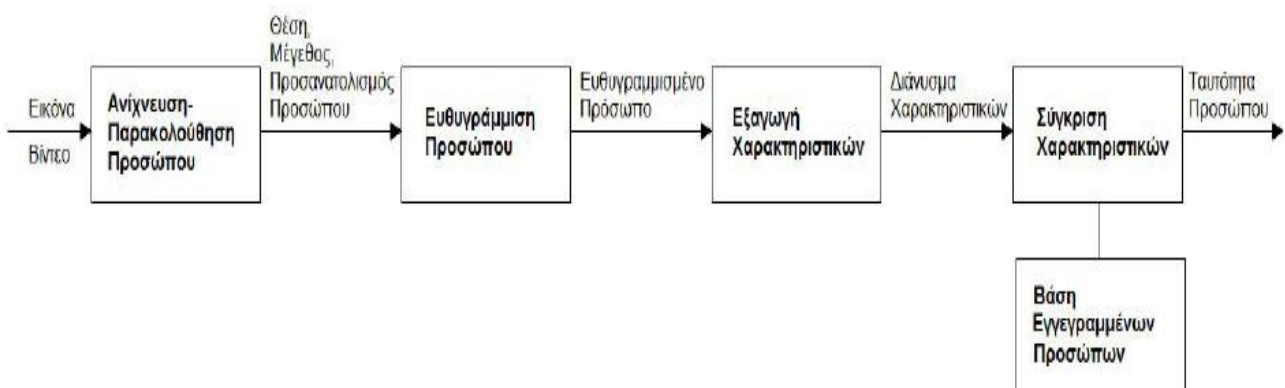
των προσώπων είναι επίσης ο βασικός λόγος για την ευρεία χρήση του σε εφαρμογές όπου η ταυτοποίηση ανθρώπων είναι ιδιαίτερα σημαντική.

Στο σημείο αυτό πρέπει να γίνει διάκριση μεταξύ ταυτοποίησης προσώπων και πιστοποίησης. Κατά την ταυτοποίηση το σύστημα καλείται να βρει την ταυτότητα του ατόμου της συλλογής στο οποίο ανήκει η εικόνα δοκιμής. Το σύστημα απαντά στο ερώτημα "ποιος είναι " ο εικονιζόμενος. Επιπλέον, χρησιμοποιείται αποκλειστικά το οπτικό σύστημα του ανθρώπου. Στην γενικότερη ταυτοποίηση ανθρώπων χρησιμοποιούνται και άλλα στοιχεία, όπως το βάδισμα, η στάση του σώματος, η φωνή ή το ντύσιμο.

Κατά την πιστοποίηση, το σύστημα αναγνώρισης προσώπων καλείται να εξακριβώσει αν η ταυτότητα που επικαλείται κάποιος χρήστης είναι αληθής ή όχι, καλείται δηλαδή να απαντήσει στο ερώτημα "είναι πράγματι αυτός που ισχυρίζεται ότι είναι " κι έτσι να διακρίνει αν πρόκειται για εξουσιοδοτημένη χρήση (legitimate user) ή για πλαστοπροσωπία (impostor). [1,13]

Το πρόβλημα της αναγνώρισης προσώπων σε γενικές γραμμές τίθεται ως εξής: *Δοθέντων φωτογραφιών ή ακολουθιών βίντεο από μια σκηνή, αναγνώρισε ένα ή περισσότερα άτομα χρησιμοποιώντας μια υπάρχουσα βάση δεδομένων από πρόσωπα.* Συμπληρωματικές πληροφορίες, όπως η ηλικία, το φύλο, η φυλή, μπορεί να είναι διαθέσιμες για κάθε αποθηκευμένο στη βάση πρόσωπο, ούτως ώστε να περιορίζεται η έρευνα.

Ένα σύστημα αναγνώρισης προσώπου αποτελείται γενικά από τέσσερα υποσυστήματα [10], όπως φαίνεται στην εικόνα 2.1: την ανίχνευση, την ευθυγράμμιση, την εξαγωγή χαρακτηριστικών και τη σύγκριση, όπου ο εντοπισμός και η κανονικοποίηση των προσώπων, μέσω της ανίχνευσης και της ευθυγράμμισης, προηγούνται της αναγνώρισης, μέσω της εξαγωγής των χαρακτηριστικών και της σύγκρισης με τη βάση των εγγεγραμμένων προσώπων.



Εικ 2.1 : Διάγραμμα διαδικασίας αναγνώρισης προσώπου [10]

2.2 Προβλήματα αυτοματοποίησης αναγνώρισης προσώπων

Διάφορα προβλήματα πηγάζουν από την αυτοματοποίηση της παραπάνω διαδικασίας:

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

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

Ο τρόπος με τον οποίο γίνεται η αναγνώριση προσώπων από τον άνθρωπο έχει απασχολήσει τους τομείς της Ψυχοφυσικής και της Νευροφυσιολογίας για περισσότερα από 20 χρόνια. Η μοναδικότητα των προσώπων, το ερώτημα αν η αναγνώριση γίνεται ολιστικά ή με χρήση επιμέρους χαρακτηριστικών, η ανάλυση των εκφράσεων του προσώπου για σκοπούς αναγνώρισης, η οργάνωση της μνήμης για τα πρόσωπα και διάφορα άλλα θέματα απασχόλησαν τις παραπάνω επιστήμες. Οι μηχανικοί έδειξαν ενδιαφέρον για την έρευνα αυτή με σκοπό την υλοποίηση ενός αποδοτικού συστήματος αναγνώρισης. Μερικά από τα αποτελέσματα είναι πράγματι ενδιαφέροντα αν και τελικά η αναγνώριση προσώπων από μηχανές αναπτύχθηκε ανεξάρτητα από τις μελέτες της Ψυχοφυσικής και Νευροφυσιολογίας. Σημαντική προσπάθεια στον τομέα της αναγνώρισης προσώπων από μηχανές είχε γίνει τα πρώτα χρόνια της δεκαετίας του '70. Κατά τη διάρκεια της δεκαετίας του '80 υπήρξε μια σχετική στασιμότητα. Από τις αρχές της δεκαετίας του '90 παρατηρείται μια σημαντική δραστηριότητα και ανανεωμένο ενδιαφέρον, το οποίο εκφράζεται με τη δημοσίευση σε τακτική βάση άρθρων σε περιοδικά όπως το *IEEE Transactions on PAMI*, και τη διοργάνωση ειδικών παρουσιάσεων - sessions - σε μεγάλα διεθνή συνέδρια (*IEEE ICIP, IEEE ICASSP, IEEE Multimedia*) [17].

2.3 Εφαρμογές

Ένα σύστημα εντοπισμού και ταυτοποίησης προσώπων θα ήταν χρήσιμο σε μια πλειάδα εφαρμογών όπως αναγνώριση εγκληματιών, πρόσβαση σε συστήματα ασφαλείας όπως ταμεία τραπεζών (ATM) όπου πραγματοποιούνται ηλεκτρονικές συναλλαγές και γενικά σε κάθε εφαρμογή που απαιτεί πρόσβαση βάσει κάποιου αποδεικτικού στοιχείου ταυτότητας. Σε τέτοιου είδους εφαρμογές η αναγνώριση προσώπου υποκαθιστά κλασικούς τρόπους επαλήθευσης της ταυτότητας του χρήστη όπως οι προσωπικοί αριθμοί ή κωδικοί πρόσβασης (PIN, password), οι οποίοι μπορούν εύκολα να υποκλαπούν ή να ξεχαστούν. Η πρόσβαση βασισμένη σε αναγνώριση προσώπων θα μπορούσε να είναι περισσότερο αξιόπιστη και πιο απλή. Εντούτοις για εφαρμογές που απαιτούν υψηλή αξιοπιστία όσον αφορά την ταυτοποίηση, η αναγνώριση προσώπου δεν επαρκεί –τουλάχιστον με τα υπάρχοντα δεδομένα. Αντίθετα σε εφαρμογές όπως τα συστήματα επικοινωνίας ανθρώπου μηχανής, ανάκλησης στοιχείων από βάσεις δεδομένων με βάση το περιεχόμενο, τηλεδιάσκεψη κ.λ.π, η αναγνώριση προσώπου είναι πιο εύκολα υλοποιήσιμη και σε πολλές περιπτώσεις απαραίτητη. Η ανάπτυξη ενός υπολογιστικού μοντέλου αναγνώρισης προσώπων είναι μια πρόκληση για την ανθρώπινη εφευρετικότητα και φιλοδοξία, αφού οι συνεχείς τεχνολογικές εξελίξεις ιδιαίτερα στη χώρα της περιβάλλουσας νοημοσύνης, καθιστούν την ανάπτυξη αξιόπιστων συστημάτων αναγνώρισης προσώπου, μία επιτακτική ανάγκη. Εφαρμογές ενός τέτοιου μοντέλου, ενδεχομένως σε μελλοντικό χρόνο, μπορούν να χρησιμοποιηθούν για ασφάλεια κτιρίων και σταθμών εργασίας, για αναγνώριση κακοποιών, επαλήθευση πιστωτικών καρτών, ακόμα για βελτίωση της επικοινωνίας ανθρώπου – μηχανής.

Αναλυτικότερα, εμπορικές και νομικής φύσεως εφαρμογές της αναγνώρισης προσώπου περιγράφονται παρακάτω, όπως και άλλες εφαρμογές κοινωνικής φύσεως.

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

Γενικότερα, μπορεί κάποιος να ταξινομήσει τις τεχνικές και τα επιμέρους προβλήματα σε δύο κατηγορίες:

- **στατικής ταύτισης:** δεν υπάρχει δυνατότητα παρακολούθησης κίνησης
- **δυναμικής ταύτισης:** υπάρχει δυνατότητα παρακολούθησης της χρονικής μεταβολής των δεδομένων.

Ακόμη και στο πλαίσιο των πιο πάνω γενικών κατηγοριών υπάρχουν σημαντικές διαφοροποιήσεις ανάλογα με τη συγκεκριμένη εφαρμογή. Οι διαφορές αφορούν στην ποιότητα της εικόνας δηλαδή στη χωρική ανάλυση, στην πολυπλοκότητα του φόντου η οποία επηρεάζει αποφασιστικά την ευκολία με την οποία μπορεί να εντοπιστεί το πρόσωπο μέσα στην εικόνα, στην ύπαρξη ενός καλά ορισμένου κριτηρίου ταύτισης καθώς και στο ποσοστό, στο είδος και στην υφή της παρέμβασης από τον άνθρωπο, όπως στις εφαρμογές 4 και 5 του Πίνακα 1.

Σε μερικές εφαρμογές, όπως στον εντοπισμό των μεταβολών που επήλθαν με την γήρανση απαιτείται η εύρεση των μετασχηματισμών εκείνων που μετά την εφαρμογή τους οι νέες εικόνες θα είναι παρόμοιες με τις αναμενόμενες.

Σύμφωνα με όσα έχουμε προαναφέρει, προκύπτουν τρία διαφορετικά προβλήματα που αφορούν στις διάφορες εφαρμογές. Συγκεκριμένα πρόκειται για την ταύτιση, την ανίχνευση ομοιότητας και την εφαρμογή μετασχηματισμών. Οι εφαρμογές 1, 2, 3, 10 και 11 θέτουν κυρίως προβλήματα ταύτισης μιας εικόνας προσώπου με κάποια άλλη. Οι εφαρμογές 4-7 αναφέρονται στη δημιουργία μιας εικόνας προσώπου όμοιας με αυτήν που θυμούνται ή περιγράφουν αυτόπτες μάρτυρες. Τέλος οι εφαρμογές 8 και 9 εμπλέκουν την εφαρμογή μετασχηματισμών και τροποποιήσεων σε μια εικόνα με βάση στοιχεία που είναι χρήσιμα σε άλλες εφαρμογές. Κάθε μια από της εφαρμογές που αναφέρονται στον Πίνακα 1 θέτει διαφορετικές απαιτήσεις και περιορισμούς στη διαδικασία της αναγνώρισης προσώπου. Η ταύτιση απαιτεί όπως η προς ταυτοποίηση εικόνα ανήκει σε ένα σύνολο από εικόνες οι οποίες έχουν επιλεγεί από το υπολογιστικό σύστημα. Η ανίχνευση ομοιότητας εκτός από την ταύτιση απαιτεί επιπλέον οι επιλεγείσες εικόνες να ομοιάζουν με αυτές που αυτόπτες μάρτυρες ανακαλούν στη μνήμη τους. Η τελευταία απαίτηση θέτει ένα σοβαρό περιορισμό στο κριτήριο ταύτισης το οποίο πρέπει να είναι αντίστοιχο με αυτό που χρησιμοποιούν οι άνθρωποι. Οι εφαρμογές μετασχηματισμού από την άλλη πλευρά πρέπει να προσομοιάζουν τον τρόπο με τον οποίο λειτουργεί η γήρανση -ή τουλάχιστον αυτό που αναμένουν οι άνθρωποι να συμβεί ως συνέπεια της παρόδου του χρόνου.

Για το πρόβλημα της αναγνώρισης προσώπου, έχουν προταθεί πολλές τεχνικές τα τελευταία χρόνια εκ των οποίων οι τρεις δημοφιλέστερες είναι [4,11]:

- Αναγνώριση προσώπου με Ανάλυση Πρωτευουσών Συνιστωσών γνωστή και ως η τεχνική των eigenfaces.
- Αναγνώριση προσώπου με Γραμμική διακριτή Ανάλυση (LDA).
- Αναγνώριση προσώπου με μεθόδους ταιριάσματος Ελαστικών Γράφων (EGM).

Πίνακας 1: Συνήθειες εφαρμογές εμπορικής και νομικής φύσεως της αναγνώρισης προσώπων

A/A	Εφαρμογές	Πλεονεκτήματα	Μειονεκτήματα
1(α)	Πιστωτικές κάρτες, ΑΤΜ μηχανήματα, διαβατήρια	Καλή ποιότητα εικόνων, ελεγχόμενες συνθήκες λήψης, εύκολος εντοπισμός προσώπου	Μη υπάρχουσες βάσεις, μεγάλο μέγεθος βάσεων, σπάνια
1(β)	Mug shots ταύτιση	Ελεγχόμενες συνθήκες λήψης, περισσότερες από μια φωτογραφίες για κάθε πρόσωπο	χρησιμοποιούμενη αναζήτηση
2	Ασφάλεια τραπεζών/καταστημάτων	Υψηλή ζήτηση, γεωγραφικά και τοπικά Περιορισμένη αναζήτηση	Μη ελεγχόμενο φόντο, μικρός αριθμός φωτογραφιών
3	Επισκόπηση - παρακολούθηση πλήθους	Υψηλή ζήτηση, μικρό μέγεθος αρχείων, εκμετάλλευση κίνησης	Μη ελεγχόμενο φόντο, χαμηλή ποιότητα εικόνων, εκτέλεση σε πραγματικό χρόνο
4	Ταυτοποίηση από ειδικούς	Υψηλή ζήτηση, δυνατότητα αναβάθμισης φωτογραφιών	Χαμηλή ποιότητα εικόνων, απαραίτητη η άδεια από τις αστυνομικές αρχές
5	Ανακατασκευή προσώπου με βάση τις μαρτυρίες	Όρια που τίθενται από τους αυτόπτες μάρτυρες	Αμφίβολη ομοιότητα
6	Ηλεκτρονικό βιβλίο από mug shots	Όρια που τίθενται από την περιγραφή	Κούραση παρατηρητή
7	Ηλεκτρονική παράταξη υπόπτων	Όρια που τίθενται από την περιγραφή	Κούραση παρατηρητή
8	Ανακατασκευή προσώπου από υπολειπόμενα στοιχεία	Υψηλή ζήτηση	Απαιτείται η χρήση δεδομένων από τη φυσιολογία
9	Υπολογισμός αλλαγών που επήλθαν με την πάροδο του χρόνου (aging)	Υψηλή ζήτηση	Υπολογισμός αλλαγών που επήλθαν με την πάροδο του χρόνου
10	Υποδιαίρεση ειδήσεων σε θεματικές ενότητες.	Χρήση χρώματος	Ετερογενείς βάσεις πολυμέσων
11	Υποδιαίρεση ειδήσεων σε θεματικές ενότητες.	Χρήση χρώματος	Ανομοιογένεια συνθηκών studio και εξωτερικών λήψεων

Κεφάλαιο 3

ΑΝΙΧΝΕΥΣΗ ΠΡΟΣΩΠΩΝ

3.1 Γενικά	16
3.2 Προκλήσεις στη διαδικασία ανίχνευσης προσώπων	17
3.3 Ιστορικό των κυριότερων μεθόδων Ανίχνευσης προσώπων	18
3.4 Μέθοδοι ανίχνευσης/εντοπισμού Προσώπου	20
3.5 Μέθοδοι βασισμένες στην εμφάνιση	24
3.5.1 Δείγματα προσώπων και μη-προσώπων	24
3.5.2 Ταξινομητές	26
3.5.3 Αναπαράσταση προσώπων	27
3.5.4 Προεπεξεργασία	27
3.5.5 Μετεπεξεργασία	28

3.1 Γενικά

Η ανίχνευση προσώπου (face detection) είναι το πρώτο βήμα στη διαδικασία αναγνώρισης προσώπου (face recognition). Μετά από την επιτυχημένη ανίχνευση του ανθρώπινου προσώπου μπορεί να ακολουθήσει η αναγνώριση του ώστε να προσδιοριστεί η ταυτότητά του, να ανιχνευθούν τα επιμέρους χαρακτηριστικά του προσώπου του (όπως τα μάτια, στόμα, μύτη), να βρεθεί η θέση του στο χώρο ή ακόμα να εντοπισθούν ανθρώπινες εκφράσεις. Η αξιοπιστία του συστήματος ανίχνευσης παίζει κυρίαρχο ρόλο στην απόδοση και χρησιμότητα του συστήματος αναγνώρισης προσώπου.

Το πρόβλημα της ανίχνευσης προσώπου μπορεί να διατυπωθεί ως: *“Δεδομένου μιας εικόνας, ζητείται η ανίχνευση και ο ακριβής εντοπισμός της θέσης ενός αγνώστου εκ των προτέρων αριθμού προσώπων που πιθανόν να υπάρχουν στην εικόνα”*. [20] Αν δοθεί μια φωτογραφία ο ιδανικός ανιχνευτής προσώπου πρέπει να είναι ικανός να ανακαλύπτει και να εντοπίζει όλα τα πρόσωπα που είναι παρόντα, ανεξάρτητα από τις συνθήκες φωτισμού, τη θέση, την κλίμακα, τον προσανατολισμό και τους μορφασμούς των προσώπων. Μετά από αυτό το

στάδιο, μπορούμε να αποκωδικοποιήσουμε και τα συναισθήματα που μας παρουσιάζονται μέσω της έκφρασης των προσώπων που βλέπουμε.

3.2 Προκλήσεις στη διαδικασία ανίχνευσης προσώπου

Οι προκλήσεις που συνδέονται με την ανίχνευση προσώπου συνήθως οφείλονται κυρίως στους εξής παράγοντες [20]:

- **Στάση/πόζα του προσώπου:** Οι εικόνες ενός προσώπου διαφέρουν λόγω της στάσης της φωτογραφικής μηχανής (μετωπική, ανάποδα, προφίλ, 45 μοίρες) και κάποια χαρακτηριστικά του προσώπου τα οποία μπορεί να γίνουν μερικώς ή πλήρως αποφραγμένα.
- **Ύπαρξη δομικών στοιχείων:** Μπορούν να παρουσιάζονται ή όχι χαρακτηριστικά του προσώπου όπως γένια, μουστάκια και γυαλιά τα οποία μπορούν εύκολα να μεταβάλλουν το σχήμα, το χρώμα και το μέγεθος του προσώπου.
- **Εκφράσεις προσώπου (ύπαρξη μορφασμών):** Η εμφάνιση του προσώπου ενός ατόμου επηρεάζεται άμεσα από τις εκφράσεις του.
- **Προσανατολισμός της εικόνας:** Οι εικόνες του προσώπου μπορούν να διαφέρουν ανάλογα με τις περιστροφές γύρω από τον οπτικό άξονα της φωτογραφικής μηχανής και ανάλογα με το επίπεδο της εικόνας.
- **Συνθήκες εικόνας:** Όταν η εικόνα σχηματίζεται, παράγοντες όπως ο φωτισμός και οι συνθήκες αποτύπωσης της εικόνας (ευαισθησία της κάμερας και ανάλυση) μπορούν να επηρεάσουν την εμφάνιση ενός προσώπου.
- **Θέση και μέγεθος της εικόνας**
- **Περιεχόμενο της εικόνας:** Σε μια εικόνα προσώπου είναι πιθανή η ύπαρξη εμποδίων και άλλων προσώπων.

Όλα τα παραπάνω αποτελούν τις βασικές αιτίες της μεγάλης δυσκολίας που εμφανίζει το πρόβλημα της ανίχνευσης ενός προσώπου. Υπάρχουν πολλά αλληλένδετα προβλήματα της ανίχνευσης προσώπου. Ο εντοπισμός του προσώπου έχει στόχο να εντοπίσει τη θέση του μέσα σε μία εικόνα, καθώς και τη θέση των χαρακτηριστικών του όπως μάτια, μύτη, στόμα, αυτιά, κλπ. Αυτές οι μεγάλες παραλλαγές στην εμφάνιση του προσώπου, στην στάση, στην έκφραση και τον φωτισμό κάνουν πολύ πολύπλοκο το πολύπτυχο του χώρου των προσώπων.

Στην εικόνα 3.1 συναντούμε πρόσωπα με ποικίλες εκφράσεις προσώπου μπορεί να είναι μερικά επικαλυπτόμενα ή ο φωτισμός μπορεί να μην είναι ομοιόμορφα διάχυτος στην περιοχή του προσώπου. Όλες αυτές οι περιπτώσεις θα μπορούσαν να θεωρηθούν ως παραμορφώσεις του σήματος εισόδου, τις οποίες το σύστημα ανίχνευσης θα πρέπει να είναι αρκετά ικανό να χειριστεί με επιτυχία.



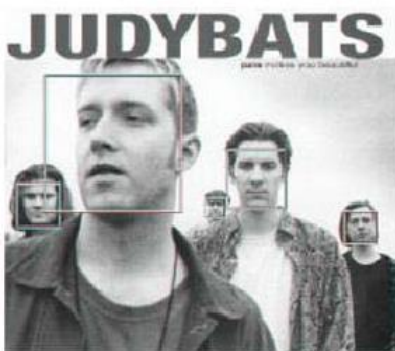
Εικ 3.1 : Προκλήσεις στην ανίχνευση προσώπων

3.3 Ιστορικό των κυριότερων μεθόδων Ανίχνευσης προσώπων

Παρακάτω αναφέρουμε τις κυριότερες μεθόδους υποδειγμάτων και εμφάνισης για την ανίχνευση προσώπων, που επηρέασαν ουσιαστικά την εξέλιξη της έρευνας στο συγκεκριμένο τομέα:

- 1992: Craw et al., Χρήση Υποδειγμάτων Μορφών (Shape Templates) [2]
- 1995: Lanitis et al., Χρήση Μοντέλων Ενεργών Μορφών (Active Shape Models) [5]
- 1997: Osuna et al., Χρήση μη γραμμικών Μηχανών Ανυσματικής Στήριξης (SVM)
- 1998: Papageorgiou and Poggio, Χρήση κυματιδίων Haar (Haar wavelets) με SVM ταξινομητή [12]
- 1998: Sung and Poggio, Χρήση Γκαουσιανών κατανομών και πολυστρωματικών αντιλήπτρων (multilayer perceptron) [16]
- 1998: Rowley et al., Χρήση συνόλου νευρωνικών δικτύων (neural networks) [14]

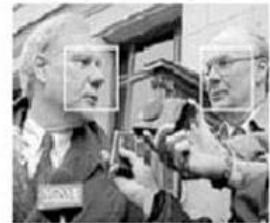
- 2000: Schneiderman and Kanade, Χρήση στατιστικής για την τοπική εμφάνιση και εκμάθηση με AdaBoost [15]
- 2001: Viola and Jones., Χρήση των κυματιδίων τύπου Haar σε ταξινομητές ενός-μόνο χαρακτηριστικού, που εκπαιδεύονται με τον AdaBoost αλγόριθμο και οργανώνονται σε μια διαδοχή αποφάσεων [19]
- 2002: Lienhart et al., Χρήση ενός εκτεταμένου συνόλου περιστραμμένων χαρακτηριστικών Haar για την αντιμετώπιση των στροφών στο επίπεδο εικόνας [9]
- 2002: Li et al., Χρήση ενός επεκτεταμένου συνόλου περιστραμμένων χαρακτηριστικών Haar και ενός πυραμιδοειδούς ανιχνευτή για την αντιμετώπιση των στροφών εκτός του επιπέδου της εικόνας [8]



• Viola & Jones



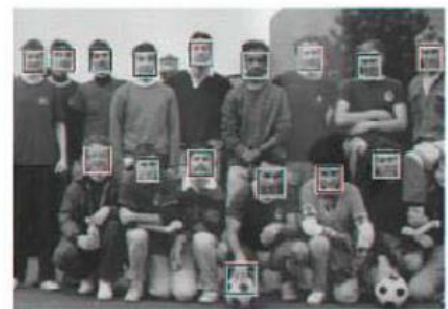
• Li et al.



• Schneiderman et al.



• Rowley et al.



Εικ 3.2 : Παραδείγματα ανίχνευσης προσώπων

Η προσέγγιση της Adaboost εκμάθησης χαρακτηριστικών τύπου Haar των Viola & Jones έχει επιτύχει την καλύτερη απόδοση μέχρι σήμερα και ως προς την ακρίβεια και ως προς την ταχύτητα.

3.4 Μέθοδοι ανίχνευσης/εντοπισμού Προσώπου

Παρακάτω αναλύονται οι υπάρχουσες τεχνικές για την ανίχνευση ή τον εντοπισμό προσώπων σε μονόχρωμες ή έγχρωμες εικόνες:

➤ **Μέθοδοι βασισμένες στη γνώση (knowledge-based methods)**

Αυτές οι μέθοδοι κωδικοποιούν την ανθρώπινη γνώση σε κανόνες αναφορικά με το τι συνιστά ένα τυπικό πρόσωπο. Συνήθως κωδικοποιούνται από τον ερευνητή οι σχέσεις (θέσεις, αποστάσεις) ανάμεσα στα χαρακτηριστικά του προσώπου. Οι δυσκολίες της μεθόδου οφείλονται στη μετατροπή των σχέσεων των χαρακτηριστικών σε συγκεκριμένους κανόνες. Αυτές οι μέθοδοι έχουν σχεδιαστεί για τον εντοπισμό του προσώπου.

➤ **Προσεγγίσεις αναλλοίωτων χαρακτηριστικών (feature invariant approaches)**

Στόχος αυτών των μεθόδων είναι να βρεθούν τα δομικά χαρακτηριστικά ενός προσώπου που υπάρχουν ακόμα και όταν η στάση, η θέση ή και οι συνθήκες φωτισμού ποικίλουν. Με βάση τα δομικά χαρακτηριστικά (μάτια, στόμα, μύτη, φρύδια, κλπ.) χτίζονται στατιστικά μοντέλα μορφών που στη συνέχεια μπορούν να χρησιμοποιηθούν για τον εντοπισμό του προσώπου. Οι δυσκολίες της μεθόδου οφείλονται στην απόκρυψη των χαρακτηριστικών ή στην ύπαρξη πρόσθετων αντικειμένων που αλλοιώνουν τις μορφές.

➤ **Μέθοδοι συνταιριάσματος υποδειγμάτων (template matching methods)**

Σε αυτή τη μέθοδο χρησιμοποιούνται διάφορα τυπικά υποδείγματα ώστε να περιγράψουν γενικότερα το πρόσωπο ή καθένα από τα χαρακτηριστικά του ξεχωριστά. Για την ανίχνευση υπολογίζεται η συσχέτιση μεταξύ της εικόνας εισόδου και των αποθηκευμένων υποδειγμάτων. Οι μέθοδοι αυτές χρησιμοποιούνται για ανίχνευση αλλά και για εντοπισμό των προσώπων. Οι δυσκολίες της μεθόδου οφείλονται στην αντιμετώπιση των παραλλαγών στην κλίμακα, τη μορφή και τον προσανατολισμό, δεδομένου ότι τα υποδείγματα είναι προκαθορισμένα.

➤ **Μέθοδοι βασισμένες στην εμφάνιση (appearance-based methods)**

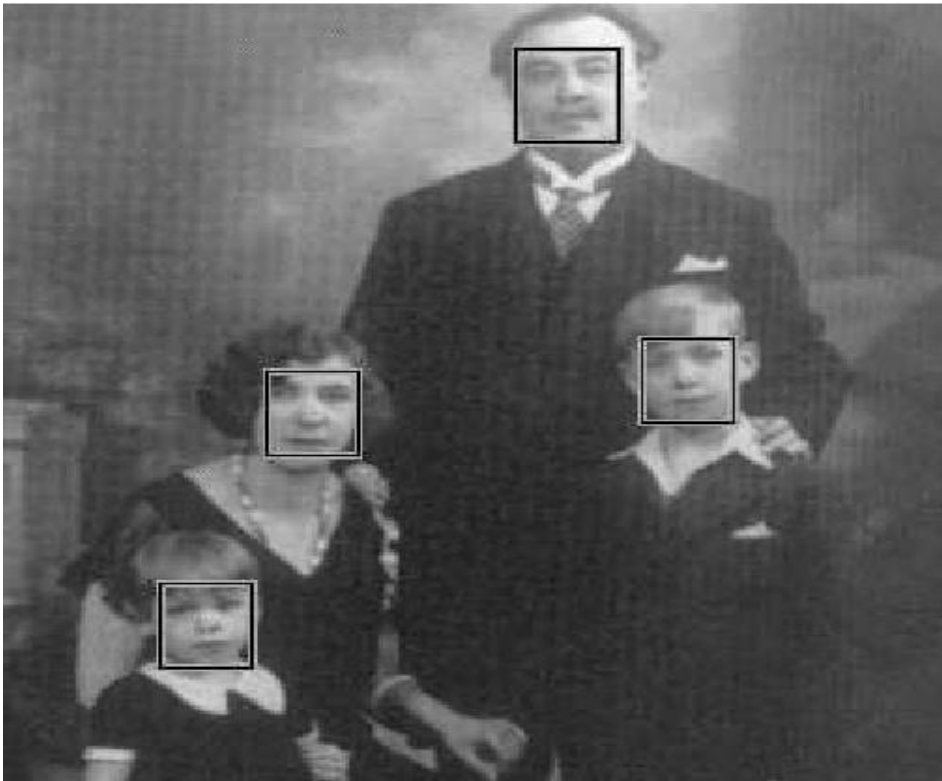
Εδώ τα μοντέλα ή υποδείγματα προσώπου, σε αντίθεση με τη μέθοδο των υποδειγμάτων όπου τα υποδείγματα είναι προκαθορισμένα, μαθαίνονται από ένα σύνολο εικόνων εκπαίδευσης που είναι αντιπροσωπευτικές της ποικιλότητας της

εμφάνισης των προσώπων. Αυτά τα μοντέλα από εκμάθηση χρησιμοποιούνται στη συνέχεια για την ανίχνευση των προσώπων. Οι μέθοδοι αυτές χρησιμοποιούνται κυρίως για ανίχνευση αλλά και για εντοπισμό των προσώπων.

Οι πιο βασικές και επιτυχημένες τεχνικές για ανίχνευση προσώπου είναι οι μέθοδοι που βασίζονται στην εμφάνιση (appearance-based). Ο λόγος είναι ότι αυτές βασίζονται αποκλειστικά στην εμφάνιση των προσώπων χωρίς να χρησιμοποιούνται άλλα ευρήματα. Η διαδικασία ανίχνευσης έχει ως εξής:

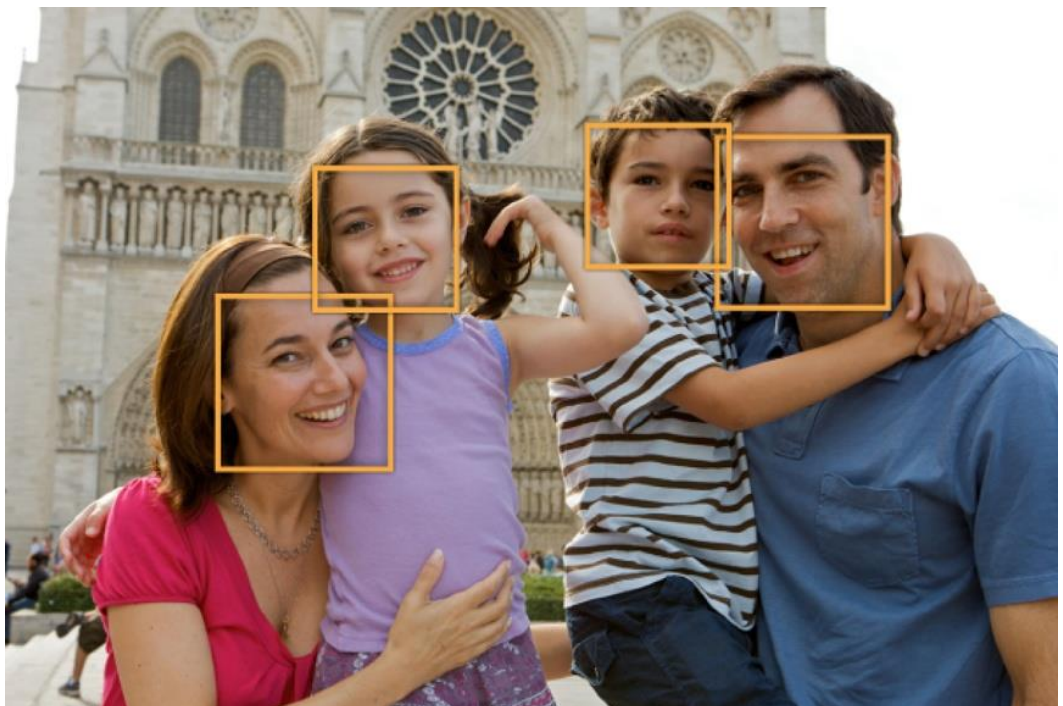
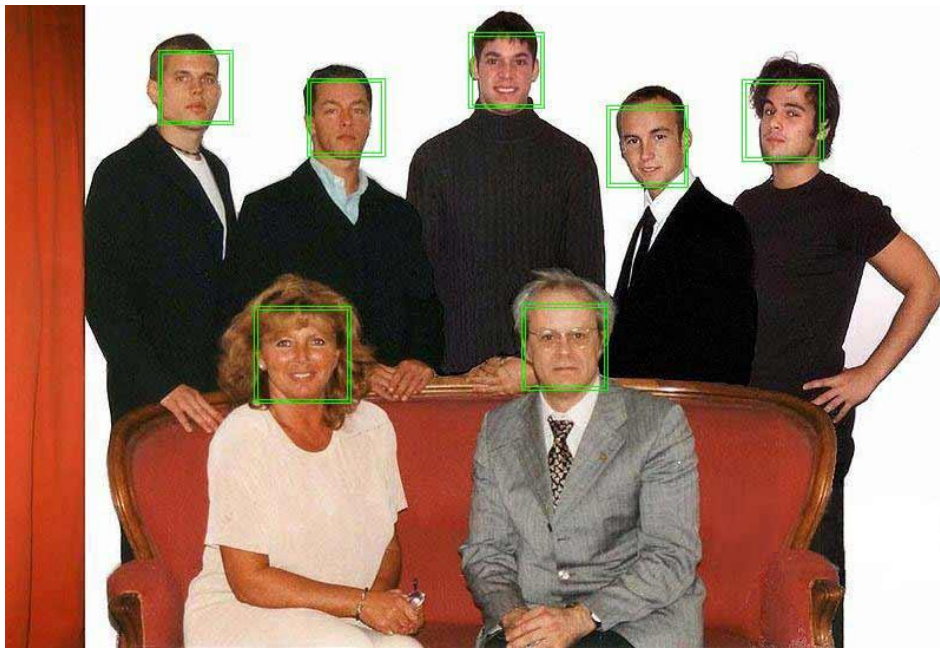
Μια εικόνα εισόδου σαρώνεται σε όλες τις δυνατές θέσεις και κλίμακες από ένα υπο-παράθυρο ανίχνευσης. Η ανίχνευση προσώπου αποφασίζεται από την ταξινόμηση του δείγματος στο υπο-παράθυρο σαν πρόσωπο ή μη-πρόσωπο. Ο ταξινομητής προσώπων/μη προσώπων εκπαιδεύεται από ένα εκπαιδευτικό σύνολο δειγμάτων από πρόσωπα και μη-πρόσωπα χρησιμοποιώντας στατιστικές μεθόδους εκπαίδευσης. Η ανίχνευση προσώπου έγκειται στην ανακάλυψη και εντοπισμό των ανθρώπινων προσώπων σε μια εικόνα ανεξαρτήτως οποιοδήποτε προκλήσεων αναφέρθηκαν πιο πάνω. Έχει γίνει σημαντική ερευνητική προσπάθεια για τη δημιουργία σύνθετων και γρήγορων ταξινομητών και από το 1990 έχει γίνει σημαντική πρόοδος στο θέμα αυτό. Έρευνες απέδειξαν ότι μέθοδοι ανίχνευσης και αλγόριθμοι εκμάθησης έχουν επιδείξει εξαιρετικά αποτελέσματα. Έχουν επίσης θεσπιστεί πολλές μετρικές για την αξιολόγηση των αλγορίθμων όπως η εκμάθηση του χρόνου καθώς και ο χρόνος εκτέλεσης τους.

Μια περιοχή εικόνας χαρακτηρίζεται ως πρόσωπο από έναν ταξινομητή και θεωρείται ότι μπορεί να εντοπιστεί σωστά αν η περιοχή της εικόνας καλύπτει περισσότερο από ένα ορισμένο ποσοστό του προσώπου στην εικόνα. Γενικά θεωρείται πολύ δύσκολο για ένα μοναδικό ανιχνευτή να χειριστεί με επιτυχία όλες τις περιπτώσεις που μπορεί να βρεθεί ένα πρόσωπο στην εικόνα. Οι ανιχνευτές μπορούν να κάνουν δύο τύπους σφαλμάτων: ψευδώς αρνητικά αποτελέσματα σε πρόσωπα που λείπουν με αποτέλεσμα χαμηλά ποσοστά ανίχνευσης και ψευδώς θετικά αποτελέσματα όπου μια εικόνα κηρύσσεται ως πρόσωπο ενώ στην πραγματικότητα δεν είναι [20]. Είναι προφανές ότι δεν μπορούμε να εμπιστευτούμε έναν ταξινομητή ο οποίος θα μας δώσει πολλές εσφαλμένες ειδοποιήσεις έτσι ώστε να ανιχνεύει τα ένα ή δύο πρόσωπα που μπορεί να υπάρχουν στην εικόνα. Γενικεύοντας, υπάρχει πάντα ένας συμψηφισμός (trade) μεταξύ του ρυθμού των εσφαλμένων ειδοποιήσεων και του ποσοστού αληθών ανιχνεύσεων, στην οποία ένας ικανοποιητικός συμβιβασμός απαιτείται. Μέθοδοι που αποδίδουν το υψηλότερο δυνατό ποσοστό ανίχνευσης ενώ διατηρούν το ρυθμό εσφαλμένων ειδοποιήσεων είναι περισσότερο προτιμητέες.



Εικ 3.3 : Ένα τυπικό αποτέλεσμα της διαδικασίας ανίχνευσης προσώπου

Μια τυπική έξοδος της διαδικασίας ανίχνευσης προσώπων δίνεται στην εικόνα 3.3 πιο πάνω, όπου και τα 4 πρόσωπα έχουν ανιχνευθεί και εντοπιστεί με ακρίβεια ενώ καμία άλλη περιοχή στην εικόνα δεν έχει χαρακτηριστεί ως ‘πρόσωπο’ (εσφαλμένη ειδοποίηση – false alarm). Τα εντοπισμένα πρόσωπα αυτής της εικόνας είναι σχετικά τετριμμένες περιπτώσεις με καθαρά μετωπική θέα, ομοιόμορφη φωτεινότητα, χωρίς την επιρροή θορύβου κτλ.



Εικ. 3.4 : Παραδείγματα εικόνων στα οποία ανιχνεύονται τα ανθρώπινα πρόσωπα χρησιμοποιώντας ένα πλαίσιο σε σχήμα τετραγώνου.

3.5 Μέθοδοι βασισμένες στην εμφάνιση

Σε αντίθεση με τη μέθοδο συνταιριάσματος υποδειγμάτων όπου τα πρότυπα είναι προκαθορισμένα από τους ειδικούς, τα “πρότυπα” στις μεθόδους που είναι βασισμένες στην εμφάνιση, βασίζονται σε παραδείγματα από εικόνες. Γενικότερα, οι συγκεκριμένες μέθοδοι, στηρίζονται σε τεχνικές στατιστικές ανάλυσης και μηχανικής μάθησης ώστε να βρούμε τα χαρακτηριστικά μιας εικόνας προσώπου ή μη-προσώπου. Τα χαρακτηριστικά μάθησης είναι υπό τη μορφή των μοντέλων κατανομής ή διακρίνουσων λειτουργιών που χρησιμοποιούνται για το πρόσωπο ανίχνευσης. Παράλληλα, η μείωση διαστάσεων πραγματοποιείται συνήθως για λόγους υπολογιστικής αποδοτικότητας και αποτελεσματικότητας της ανίχνευσης. [20]

Γενικότερα, οι μέθοδοι που είναι βασισμένες στην εμφάνιση παρουσιάζουν ορισμένα κοινά χαρακτηριστικά κατά τη διαδικασία ανίχνευσης των προσώπων:

- Σύνολο εκπαίδευσης θετικών και αρνητικών δειγμάτων προσώπων
- Εκπαίδευση και χρήση ενός ταξινομητή
- Συγκεκριμένο τρόπο αναπαράστασης των προσώπων
- Προεπεξεργασία
- Συγκεκριμένη στρατηγική ανίχνευσης/αναζήτησης στο χώρο και την κλίμακα
- Μετεπεξεργασία

Τα παραπάνω χαρακτηριστικά θα επεξηγηθούν αναλυτικότερα στη συνέχεια.

2.5.1 Δείγματα προσώπων και μη προσώπων

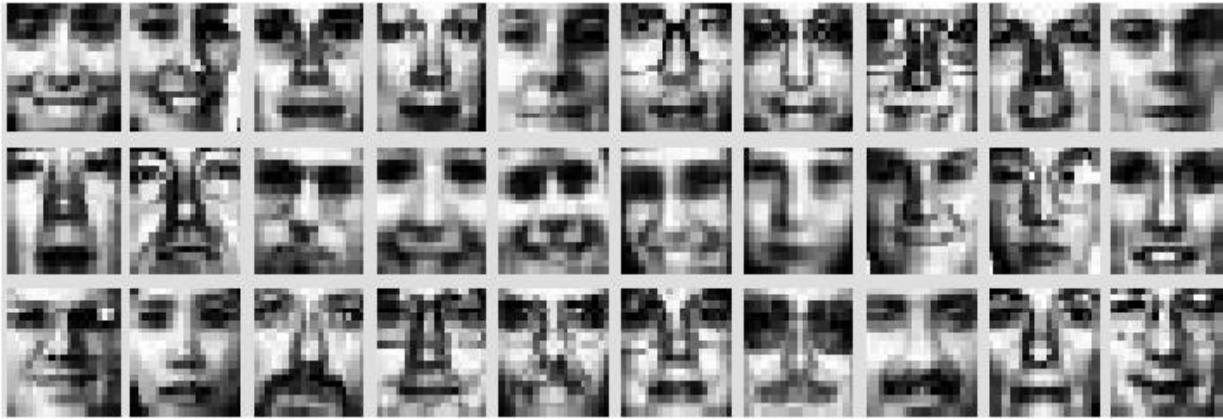
Όπως αναφέραμε πιο πάνω, ένα πρόσωπο ανίχνευσης μπορεί να θεωρηθεί ως δύο τάξεων πρόβλημα αναγνώρισης στο οποίο μια περιοχή εικόνας μπορεί να ταξινομηθεί ως “πρόσωπο” και “μη-πρόσωπο”. Η διαδικασία ταξινόμησης χρησιμοποιώντας κατηγοριοποιημένα δεδομένα μπορεί να γίνει με στατιστικές ή αλγοριθμικές μεθόδους, όπου συνδυάζοντας διάφορα αποτελέσματα μπορούμε να καταλήξουμε σε πρόσωπο ή μη-πρόσωπο. Οι ταξινομητές μπορούν να εκπαιδευτούν σε όμοια ή διαφορετικά σύνολα δεδομένων.

Στη δική μας περίπτωση, για την εκπαίδευση του ταξινομητή χρειάζεται να χρησιμοποιηθούν θετικά και αρνητικά δείγματα προσώπων:

➤ **Θετικά δείγματα:**

Απαιτείται ποικιλία δειγμάτων ώστε να έχουμε όσο το δυνατόν μεγαλύτερη ορθότητα στο αποτέλεσμα και να γίνει ακριβής ταξινόμηση του δείγματος σε πρόσωπο ή μη-πρόσωπο. Οι εικόνες που θα εισάγονται στον ταξινομητή θα πρέπει να έχουν όλες το

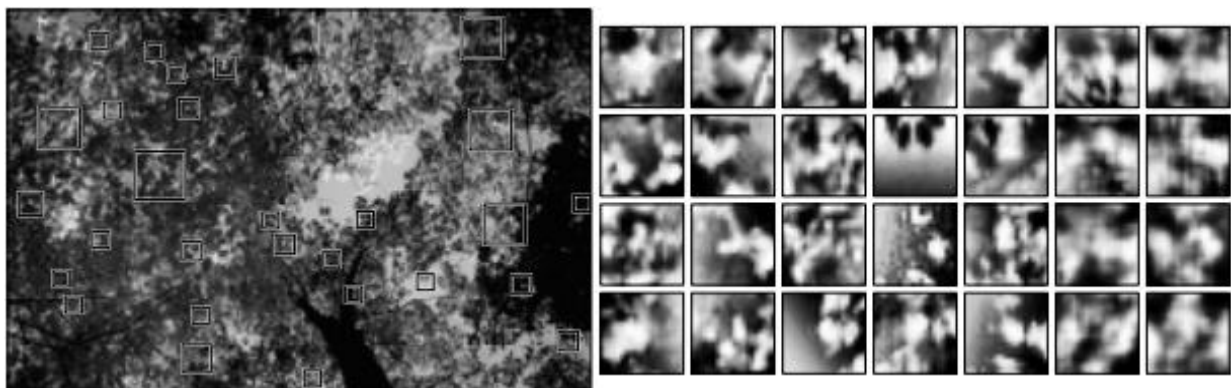
ίδιο μέγεθος, οπότε χρειάζεται προεπεξεργασία και κανονικοποίηση των εικόνων σε ένα συγκεκριμένο μέγεθος των 20x20 pixels.



Εικ. 3.5 : Παραδείγματα προσώπων που χρησιμοποιούνται ως θετικά δείγματα

➤ *Αρνητικά δείγματα*

Τα αρνητικά δείγματα περιέχουν οποιαδήποτε εικόνα στην οποία δεν παρουσιάζεται κάποιο πρόσωπο. Τα συγκεκριμένα δείγματα δέχονται επίσης κάποια προεπεξεργασία και κανονικοποιούνται όπως τα προηγούμενα στο ίδιο μέγεθος (20x20 pixels).



Εικ 3.6 : Παραδείγματα μη-προσώπων που χρησιμοποιούνται ως αρνητικά δείγματα [14]

Το σύστημα αρχικά εκπαιδεύεται σε εικόνες οι οποίες δεν περιέχουν πρόσωπα. Κάθε φορά που μια περιοχή στην εικόνα ανιχνεύεται εσφαλμένα ως πρόσωπο, αποτελεί δείγμα μη-προσώπου και προστίθεται στην κατηγορία των αρνητικών προτύπων εκπαίδευσης.

➤ **Εικονικά θετικά δείγματα**

Το σύνολο των δειγμάτων εκπαίδευσης πολλαπλασιάζεται, δημιουργώντας με τεχνητό τρόπο νέα δείγματα εικόνων από τα ήδη υπάρχοντα. Με κατοπτρισμό, μετατοπίσεις και κλιμακώσεις των προσώπων, μπορούμε να αυξήσουμε τα δείγματα εκπαίδευσης ώστε να έχουμε λιγότερα σφάλματα και να γίνει πιο ακριβής ταξινόμηση όπως φαίνεται στην εικόνα 3.7.



Εικ. 3.7 : Παραδείγματα προσώπων με κατοπτρισμό, μετατοπίσεις και κλιμακώσεις που χρησιμοποιούνται ως εικονικά θετικά δείγματα

3.5.2 Ταξινομητές

Έχει γίνει σημαντική ερευνητική προσπάθεια για τη δημιουργία σύνθετων και γρήγορων ταξινομητών και από το 1990 έχει γίνει σημαντική πρόοδος στο θέμα αυτό. Έρευνες απέδειξαν ότι αλγόριθμοι εκμάθησης έχουν επιδείξει εξαιρετικά αποτελέσματα. Έχουν επίσης θεσπιστεί πολλές μετρικές για την αξιολόγηση των αλγορίθμων όπως η εκμάθηση του χρόνου καθώς και ο χρόνος εκτέλεσης τους.

Παρακάτω παρουσιάζονται οι κυριότεροι ταξινομητές/αλγόριθμοι που έχουν χρησιμοποιηθεί σε εργασίες ανίχνευσης προσώπων:

- Αλγόριθμοι προσαρμοστικής ενίσχυσης (Adaptive Boosting - Adaboost)
- Νευρωνικά Δίκτυα (Neural Networks)
- Απλοϊκός Ταξινομητής Bayes (Naïve Bayes classifier)
- Μηχανές Ανυσματικής Στήριξης (Support Vector Machines - SVM)
- Μέθοδοι βασισμένες σε Κατανομές (Distribution – based methods)

- Ανάλυση Θεμελιωδών Συνιστωσών (Principal Component Analysis - PCA)

3.5.3 Αναπαράσταση προσώπων

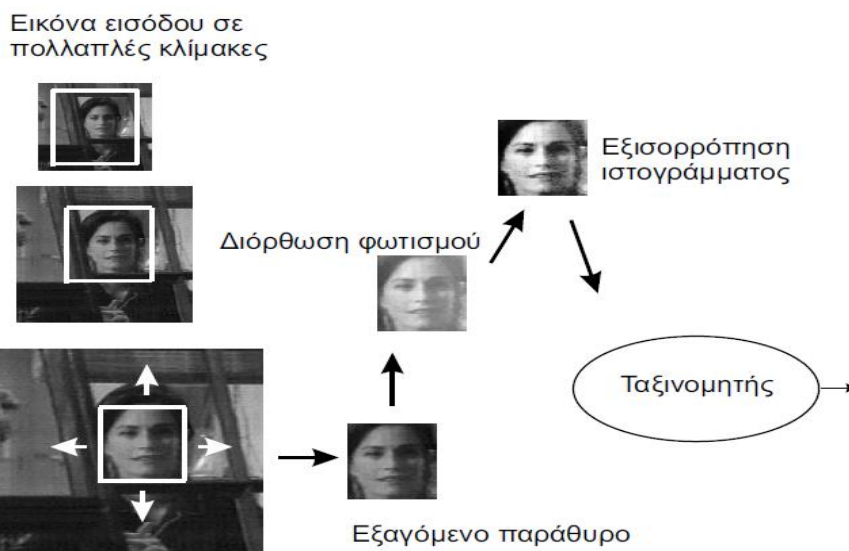
Συγκεκριμένοι τρόποι αναπαράστασης των προσώπων έχουν χρησιμοποιηθεί σε διαφορετικές εργασίες, ανάλογα με τα χαρακτηριστικά (features) που έχουν επιλεγεί:

- Ολιστική (Holistic): Κάθε εικόνα προσώπου χρησιμοποιείται ολόκληρη και αναπαρίσταται από ένα διάνυσμα τιμών έντασης.
- Βασισμένη σε μπλοκ (Block-based): Αποσύνθεση κάθε εικόνας προσώπου σε ένα σύνολο επικαλυμμένων ή μη επικαλυμμένων περιοχών (μπλοκ) όπως στην Εικ. 9.
- Απεικόνιση σε πολλαπλές κλίμακες
- Επιπλέον επεξεργασία με Διανυσματική Κβάντιση (Vector Quantization), Ανάλυση Θεμελιωδών Συνιστωσών (Principal Component Analysis), Κυματίδια (Wavelets) κλπ.
- Χρήση χαρακτηριστικών γνωρισμάτων (features).

3.5.4 Προεπεξεργασία

Η προεπεξεργασία με κανονικοποίηση της έντασης των εικονοστοιχείων (pixels) βοηθάει στη διόρθωση των αποκλίσεων των παραμέτρων αποτύπωσης των εικόνων μέσω των καμερών, όπως επίσης και των παραλλαγών στις συνθήκες φωτισμού. Τα εικονοστοιχεία της εικόνας δίνονται απευθείας σε μία εκπαιδευμένη μηχανή ταξινόμησης, χωρίς την παρεμβολή εξωτερικών παραγόντων. Για να ανιχνευτούν πρόσωπα σε διαφορετικές τοποθεσίες και μεγέθη (κλίμακες), πολλαπλές περιοχές (παράθυρα) εξάγονται από την εικόνα και δίνονται στη μηχανή ταξινόμησης. Αυτό το σενάριο απεικονίζεται στο σχήμα 9. Το μέγεθος του παραθύρου αυτού διαφέρει από μέθοδο σε μέθοδο, αλλά μία ευρύτατα αποδεκτή επιλογή είναι περίπου 20x20 εικονοστοιχεία.

Η εικόνα 3.8 μας δίνει τα διάφορα βήματα πριν δώσουμε το δεδομένο είσοδο στον ταξινομητή. Ένα παράθυρο σαρώνει την εικόνα εισόδου σε διάφορες κλίμακες και τοποθεσίες. Σε όλες τις περιπτώσεις που εξάγουμε το παράθυρο, τα εικονοστοιχεία του υφίστανται διόρθωση φωτισμού και εξισορρόπηση ιστογράμματος.



Εικ. 3.8 : Βήματα προεργασίας

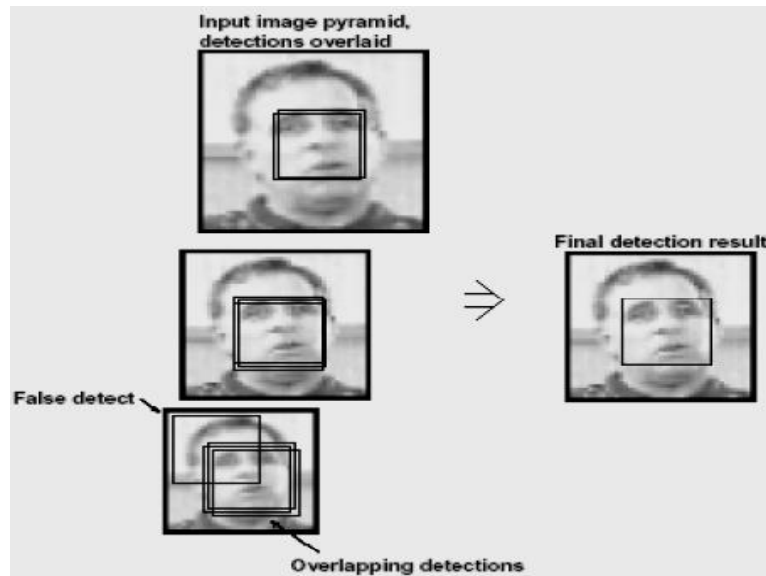
Είναι ένας καλός συμβιβασμός ανάμεσα σε ένα χώρο εισόδου με σχετικά χαμηλή διάσταση και της ικανοποιητικής διαύγειας του προσώπου μέσα στο παράθυρο. Γενικότερα, τα βήματα προ-επεξεργασίας στο υπο-παράθυρο της εικόνας και κατά τη διαμόρφωση των δειγμάτων των προσώπων είναι τα εξής: [14,16]

- Τροποποίηση μεγέθους εικόνας: Τροποποίηση όλων των προτύπων προσώπων σε ένα συγκεκριμένο μέγεθος (π.χ. σε 20x20 pixels)
- Μασκάρισμα: περιορισμός των μη επιθυμητών δομών του υπόβαθρου (θόρυβος) κοντά στα όρια του πλαισίου ενός προτύπου προσώπου
- Επιδιόρθωση διαβαθμισμένου φωτισμού: ανεύρεση μιας γραμμικής συνάρτησης η οποία να ταιριάζει στη διαβάθμιση των τιμών έντασης του παραθύρου του προσώπου και της οποίας οι τιμές αφού αφαιρεθούν, επιδιορθώνουν ακραίες αναβαθμίσεις φωτισμού (σκιές που δημιουργούνται από ακραίες γωνίες φωτισμού)
- Εξισορρόπηση Ιστογράμματος: αντισταθμίζει τις επιδράσεις φωτισμού εξαιτίας διαφορετικών συνθηκών φωτισμού και διαφορετικών καμπυλών απόκρισης της κάθε κάμερας λήψης και βελτιώνει την ποιότητα της εισόδου
- Εξομάλυνση ακμών: αφαιρεί κάποιον πιθανόν ενοχλητικό θόρυβο.

3.5.5 Μετεπεξεργασία

Από τη στιγμή που ο ανιχνευτής δεν επηρεάζεται από μικρές μεταβολές στη μετατόπιση και την κλίμακα, πολλαπλές ανιχνεύσεις μπορεί να συμβούν σε κοινές θέσεις ή σε πολλαπλές κλίμακες κατά τη σάρωση μιας εικόνας. Εσφαλμένες ανιχνεύσεις μπορεί επίσης να συμβούν,

αλλά συνήθως με λιγότερη συχνότητα από ότι πολλαπλές ανιχνεύσεις προσώπων. Όπως φαίνεται στην εικόνα 3.9, είναι χρήσιμο να γίνει μετεπεξεργασία των θετικών παραθύρων ανίχνευσης, προκειμένου να συνδυαστούν οι επικαλυπτόμενες ανιχνεύσεις σε μια μόνο ανίχνευση και να περιοριστούν οι εσφαλμένες ανιχνεύσεις.



Εικ. 3.9 : Συγχώνευση πολλαπλών ανιχνεύσεων και απαλοιφή των λανθασμένων

Ο αριθμός των πολλαπλών επικαλυπτόμενων ανιχνεύσεων στη γειτονιά μιας θέσης μπορεί να χρησιμοποιηθεί σαν μια αποτελεσματική ένδειξη για την ύπαρξη ενός προσώπου σε αυτό το σημείο, ενώ σε αντίθετη περίπτωση να απορριφθεί ως εσφαλμένη ανίχνευση.

Για τον χειρισμό των πολλαπλών ανιχνεύσεων μπορούν να χρησιμοποιηθούν οι παρακάτω τεχνικές [14,7,15,9]:

- Ανεύρεση επικαλυπτόμενων ανιχνεύσεων (η τομή τους να ξεπερνά ένα κατώφλι, π.χ. 80%)
- Συγχώνευση πολλαπλών ανιχνεύσεων (ο αριθμός των πολλαπλών ανιχνεύσεων να ξεπερνά ένα κατώφλι, π.χ. 4)
- Καθορισμός του τελικού παραθύρου (επιλογή του παραθύρου ανίχνευσης με τη μέγιστη εμπιστοσύνη ταξινόμησης $H(x)=\max$, ή εύρεση του μέσου παράθυρου)

Κεφάλαιο 4

ΑΝΙΧΝΕΥΣΗ ΠΡΟΣΩΠΟΥ ΚΑΤΑ VIOLA & JONES

4.1 Περιγραφή του προβλήματος	30
4.2 Ολοκληρωτική εικόνα	31
4.3 Χαρακτηριστικά Haar	31
4.4 Εξαγωγή χαρακτηριστικών με τον αλγόριθμο Adaboost	33
4.5 Ταξινομητής Cascade	35
4.6 Μέθοδος Haar+Adaboost : Συμπεράσματα	38

4.1 Περιγραφή του προβλήματος

Όπως έχουμε προαναφέρει η μέθοδος ανίχνευσης προσώπων Viola & Jones έχει επιτύχει την καλύτερη απόδοση μέχρι σήμερα και ως προς την ακρίβεια και ως προς την ταχύτητα. Η μέθοδος ουσιαστικά σαρώνει την εικόνα εισόδου σε όλες τις πιθανές θέσεις και κλίμακες και στη συνέχεια ταξινομεί τα παράθυρα ως πρόσωπα και μη-πρόσωπα. Οι ιδέες της μεθόδου αυτής είναι οι εξής [18,19]:

- Χρήση απλών χαρακτηριστικών τύπου Haar για απόκτηση γνώσης από δεδομένα μάθησης, που αποτελούνται από εικόνες προσώπων και μη-προσώπων.
- Χρήση μιας νέας εικόνας αναπαράστασης των προσώπων που ονομάζεται Εικόνα Ολοκλήρωμα (Integral Image), και που επιτρέπει τα χαρακτηριστικά που χρησιμοποιούνται από τον ανιχνευτή να υπολογίζονται πολύ γρήγορα.
- Χρήση ενός αλγόριθμου μάθησης, που στηρίζεται στον AdaBoost, ο οποίος επιλέγει ένα μικρό αριθμό από κρίσιμα οπτικά χαρακτηριστικά και αποδίδει άκρως αποτελεσματικούς ταξινομητές.
- Συνδυασμός των ταξινομητών σε διάταξη καταρράκτη (cascade) , που επιτρέπει περιοχές υποβάθρου της εικόνας να απορρίπτονται γρήγορα, αναλώνοντας περισσότερο υπολογιστικό χρόνο σε περιοχές που μοιάζουν περισσότερο σε πρόσωπα.

4.2 Ολοκληρωτική εικόνα

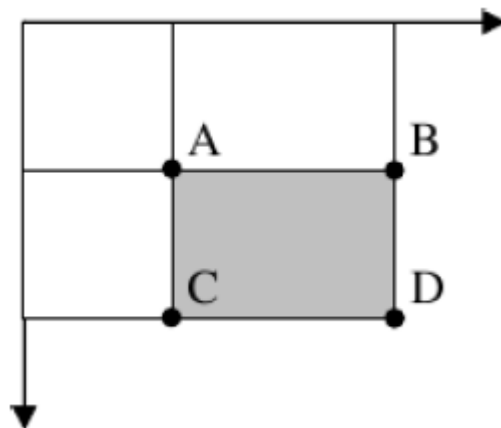
Ο στόχος είναι η δημιουργία ενός ανιχνευτή προσώπου που θα είναι γρήγορος και ακριβής.

Τα ορθογώνια χαρακτηριστικά των Viola & Jones μπορούν να υπολογιστούν πολύ αποτελεσματικά με τη χρήση μιας βοηθητικής εικόνας, που αναφέρεται ως “εικόνα ολοκλήρωμα” (integral image) και ορίζεται ως

$$I(x, y) = \sum i(x', y'), \quad x' \leq x, y' \leq y$$

Άρα η τιμή της ολοκληρωτικής εικόνας σε ένα σημείο (x, y) είναι το άθροισμα των τιμών όλων των σημείων πάνω και αριστερά από το σημείο αυτό, συμπεριλαμβανομένου του σημείου (x, y) . Η χρήση της ολοκληρωτικής εικόνας διευκολύνει τις υπολογιστικές διαδικασίες. Συγκεκριμένα, το άθροισμα των σημείων ενός ορθογωνίου της ολοκληρωτικής εικόνας μπορεί εύκολα να υπολογιστεί από τα τέσσερα ακρινά του σημεία. Για το παράδειγμα του σχήματος 4.1, όπου το ορθογώνιο ορίζεται από τα σημεία A, B, C και D ισχύει

$$S_{ABCD} = I_{int}(D) - [I_{int}(B) + I_{int}(C)] + I_{int}(A)$$

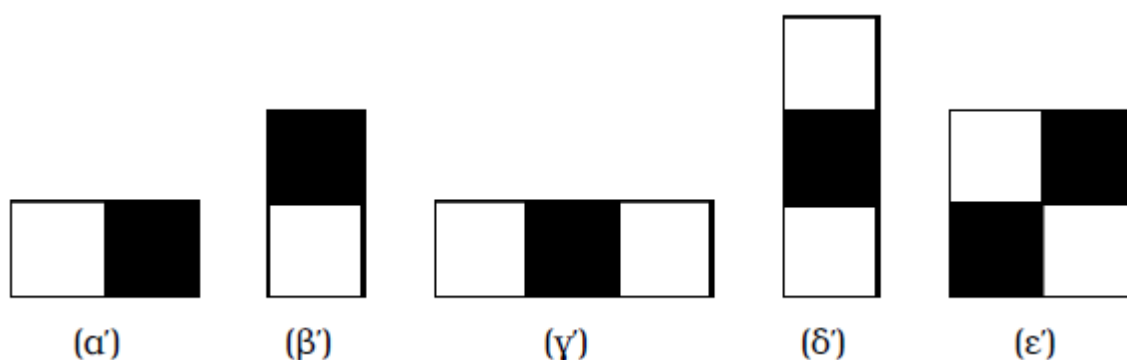


Εικ. 4.1 : Ολοκληρωτική εικόνα

4.3 Χαρακτηριστικά Haar

Αν ο ανιχνευτής προσώπου σάρωνε κάθε ένα pixel ξεχωριστά, τότε το υπολογιστικό κόστος θα ήταν πολύ μεγάλο. Αντίθετα, η χρησιμοποίηση ενός συνόλου χαρακτηριστικών (feature set) των τιμών των pixels και ο συνδυασμός τους κάνει τον αλγόριθμο πολύ ταχύτερο. Έτσι, το σύνολο χαρακτηριστικών που χρησιμοποιείται χωρίζει την εικόνα σε ορθογώνιες περιοχές και υπολογίζει το άθροισμα των pixels των τετραγώνων μέσω της ολοκληρωτικής εικόνας(integral image). Μάλιστα αυτή η τακτική είναι ιδιαίτερα συμφέρουσα αν συνυπολογιστεί το γεγονός ότι τα χαρακτηριστικά περιέχουν μεγαλύτερη πληροφορία για την περιοχή σε σύγκριση με τα μεμονωμένα pixels.

Τα χαρακτηριστικά που πρότειναν οι Viola & Jones είναι ιδιαίτερα απλά ορθογώνια σχήματα όπως φαίνεται και στο Σχήμα 4.2. Η τιμή του κάθε χαρακτηριστικού ορίζεται ως η διαφορά του αθροίσματος των pixels της λευκής περιοχής από αυτά της μαύρης περιοχής. Φυσικά η θέση και το μέγεθος των χαρακτηριστικών κάθε φορά διαφέρουν και εξαρτώνται από το παράθυρο που σαρώνει την εικόνα εισόδου.



Εικ 4.2 : Viola & Jones χαρακτηριστικά που μοιάζουν με Haar

Κάθε χαρακτηριστικό συνοδεύεται από κάποιες παραμέτρους και η τιμή τους εξαρτάται από την τιμή της υπολογιζόμενης σταθμισμένης ανάλογα με το εμβαδόν διαφοράς των αθροισμάτων των εντάσεων των εικονοστοιχείων πάνω σε ορθογώνιες περιοχές. Για παράδειγμα, το χαρακτηριστικό της εικόνας 12(γ') έχει τέσσερις παραμέτρους: τη θέση μέσα στο παράθυρο σάρωσης, το μέγεθος της άσπρης (ή θετικής) περιοχής, το μέγεθος της μαύρης (ή αρνητικής) περιοχής και το ύψος.

Έχει αποδειχθεί ότι μια καλή ανάλυση για τον ανιχνευτή είναι 24x24 pixels. Δίνοντας όλα τα πιθανά μεγέθη και όλες τις πιθανές θέσεις στα χαρακτηριστικά του σχήματος 4.2, προκύπτει ένα σύνολο 160000 περίπου διαφορετικών χαρακτηριστικών που είναι ένα υπερπλήρες

σύνολο. Τα χαρακτηριστικά αυτά, αν και απλοϊκά στη μορφή, επιτυγχάνουν εξαιρετικά αποτελέσματα με πολύ μικρό υπολογιστικό κόστος.

4.4 Εξαγωγή χαρακτηριστικών με τον αλγόριθμο Adaboost

Στο γενικότερο πρόβλημα της ενίσχυσης (boosting), συνδυάζεται ένα μεγάλο σύνολο λειτουργιών ταξινόμησης, αποδίδοντας μεγαλύτερο βάρος σε κάθε καλή λειτουργία ταξινόμησης και μικρότερο βάρος σε κάθε χειρότερη λειτουργία. Οι ταξινομητές που συνδυάζονται, ώστε να αποτελέσουν τον τελικό *ισχυρό ταξινομητή* (strong classifier), ονομάζονται *αδύναμοι ταξινομητές* (weak classifiers), και αρκεί να αποφασίζουν λίγο καλύτερα από την τυχαιότητα. Η διαδικασία συνδυασμού αδύναμων ταξινομητών για τη δημιουργία ενός δυνατού ταξινομητή χρησιμοποιείται από την ανθρώπινη λογική κατά τη λήψη αποφάσεων. Ένας αδύναμος ταξινομητής εκτελείται πολλαπλές φορές και για κάθε εκτέλεση ανανεώνεται μια κατανομή βάρη, η οποία υποδεικνύει πόσο σημαντικά είναι τα δείγματα για τη διαδικασία της ταξινόμησης. Ο πιο δημοφιλής αλγόριθμος ενίσχυσης είναι ο αλγόριθμος *προσαρμοστικής ενίσχυσης* (adaptive boosting) AdaBoost, ο οποίος ονομάζεται και *διακριτός* (discrete) AdaBoost, μιας και αποδίδει διακριτές τιμές εξόδου [24]. Όπως έχουμε ήδη αναφέρει πιο πάνω, το πλήθος των ορθογώνιων χαρακτηριστικών που δημιουργούνται είναι υπερβολικά μεγάλο, ώστε να απαιτείται η επιλογή των πλέον αποτελεσματικών χαρακτηριστικών που θα χρησιμοποιηθούν στον τελικό ταξινομητή. Θεωρώντας μια αντιστοίχιση μεταξύ αδύναμων ταξινομητών και χαρακτηριστικών οι Viola & Jones, χρησιμοποίησαν τον αλγόριθμο AdaBoost [25] σαν μια αποτελεσματική διαδικασία για την ανεύρεση ενός μικρού αριθμού "καλών" χαρακτηριστικών που επιπλέον είναι σημαντικά διαφοροποιημένα. Μία απλή και πρακτική μέθοδος για την ολοκλήρωση αυτής της αντιστοίχισης είναι ο περιορισμός του αδύναμου ταξινομητή σε λειτουργίες ταξινόμησης που εξαρτώνται από ένα μόνο χαρακτηριστικό [7, 8].

Έτσι, η μέθοδος AdaBoost στοχεύει στην επίλυση των παρακάτω 3 θεμελιωδών προβλημάτων [7]:

- (1) εκμάθηση των πιο αποτελεσματικών χαρακτηριστικών από ένα μεγάλο σύνολο χαρακτηριστικών,
- (2) κατασκευή αδύναμων ταξινομητών, καθένας από τους οποίους στηρίζεται σε ένα μόνο από τα δημιουργημένα χαρακτηριστικά, και
- (3) συνδυασμός των αδύναμων ταξινομητών για την κατασκευή ενός ισχυρού ταξινομητή.

Input: Εικόνες εκπαίδευσης $(x_1, \dots, x_n)$ και οι αντίστοιχες ετικέτες $(y_1, \dots, y_n)$ όπου $y_i \in \{0,1\}$ για θετικά και αρνητικά δείγματα αντίστοιχα. Ο αριθμός των αρνητικών δειγμάτων είναι m και αυτός των θετικών δειγμάτων $l = n - m$

1: Αρχικοποίηση των n βαρών ως $(2m)^{-1}$ για $y_i = 0$ και $(2l)^{-1}$ για $y_i = 1$

2: **for** $t = 1, \dots, T$ **do**

3: Κανονικοποίηση των βαρών $W_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}}$

4: Επιλογή του καλύτερου αδύναμου ταξινομητή ως προς το σφάλμα με βάρη $\epsilon_t = \min \sum_i w_i |h(x_i, f, p, \theta) - y_i|$

5: Ορισμός της $h_t(x) = h(x, f_t, p_t, \theta_t)$ όπου f_t, p_t, θ_t οι τιμές των μεταβλητών που ελαχιστοποιούν το ϵ_t

6: Ανανέωση των βαρών $w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$ όπου $e_i = 0$ αν το δείγμα x_i ταξινομείται σωστά, $e_i = 1$ αν το δείγμα x_i ταξινομείται λανθασμένα και $\beta_t = \frac{\epsilon_t}{1-\epsilon_t}$

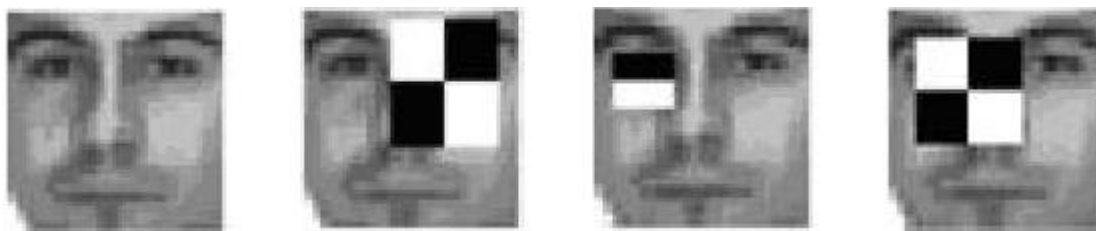
7: **end for**

8: Ο δυνατός ταξινομητής με $a_t = \log \frac{1}{\beta_t}$ ορίζεται ως

$$h_x = \begin{cases} 1, & \text{αν } \sum_{t=1}^T a_t, h_t(x) \geq \frac{1}{2} \sum_{t=1}^T a_t \\ 0, & \text{αλλιώς} \end{cases}$$

Σύμφωνα με τους Viola & Jones [18,19] κατά την εφαρμογή του AdaBoost:

- Σε κάθε πέρασμα t εκπαιδεύεται ένας νέος αδύναμος ταξινομητής h_t που προστίθεται στο σύνολο με μεγαλύτερο συντελεστή a_t όσο μικρότερο σφάλμα ταξινόμησης ϵ δίνει.
- Το βάρος κάθε δείγματος είναι η βασική παράμετρος στο συγκεκριμένο αλγόριθμο και ενημερώνεται έτσι ώστε σε κάθε επόμενο πέρασμα τα δείγματα που ταξινομούνται σωστά να έχουν μικρότερη βαρύτητα κατά β_t .
- Ο τελικός ισχυρός ταξινομητής αποτελείται από T αδύναμους ταξινομητές που αντιστοιχούν στα ισχυρότερα χαρακτηριστικά, με βάρη αντιστρόφως ανάλογα προς το σφάλμα ταξινόμησης. Στο σχήμα 4.3 φαίνονται κάποια από τα ορθογώνια χαρακτηριστικά που εφαρμόστηκαν σε ένα τυπικό πρόσωπο.
- Όπως αποδεικνύεται από τους Freund & Schapire [26] ο ισχυρός ταξινομητής μπορεί να πετύχει αυθαίρετα υψηλό ρυθμό ορθών ταξινομήσεων με αυθαίρετα χαμηλό ρυθμό εσφαλμένων ταξινομήσεων, αρκεί το πλήθος των αδύναμων ταξινομητών να είναι αρκετά μεγάλο.



Εικ 4.3 : Τα κυριότερα χαρακτηριστικά εφαρμοσμένα σε ένα τυπικό πρόσωπο

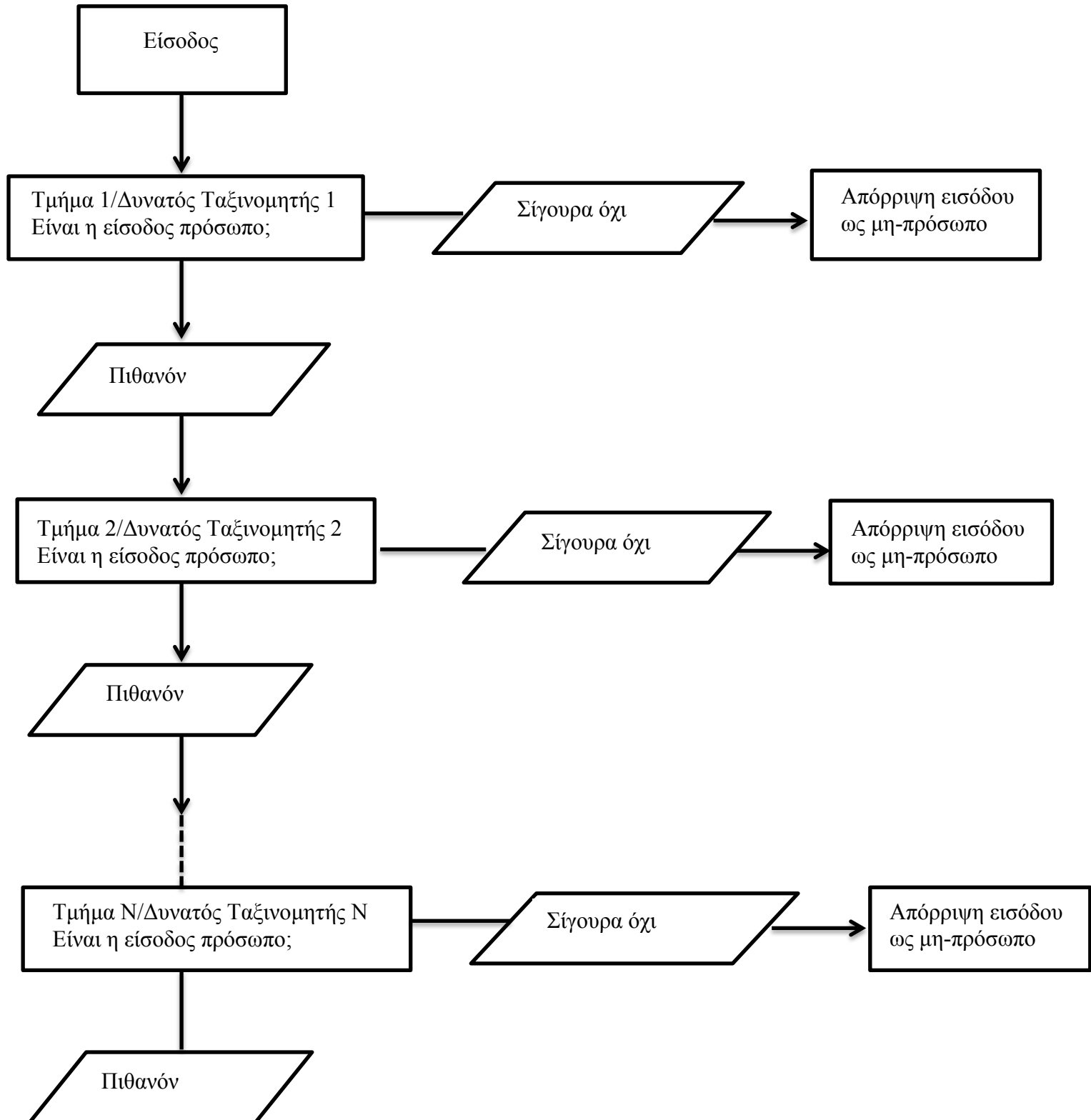
4.5 Ταξινομητής Cascade

Ο αλγόριθμος Adaboost απαιτεί μεγάλο αριθμό δειγμάτων και καταλήγει σε μεγάλο πλήθος αδύναμων ταξινομητών που απαιτούν μεγάλο υπολογιστικό κόστος. Για αυτό το λόγο οι Viola & Jones εισήγαγαν την έννοια του καταρράκτη ταξινομητών (cascade of classifiers). Όπως έχουμε προαναφέρει, η βασική αρχή της μεθόδου ανίχνευσης προσώπου των Viola & Jones είναι η πολλαπλή σάρωση της εικόνας εισόδου με τον ανιχνευτή, κάθε φορά με διαφορετικό μέγεθος. Ακόμα και αν μια εικόνα περιλαμβάνει ένα ή περισσότερα πρόσωπα, ο μεγαλύτερος αριθμός των παραθύρων υπολογισμού θα έδινε αρνητική απάντηση, δηλαδή παρουσία μη-προσώπων. Επομένως, ακολουθείται μια διαφορετική λογική, κατά την οποία αντί ο αλγόριθμος να βρίσκει πρόσωπα, μπορεί να απορρίπτει μη-πρόσωπα. Η λογική αυτή βασίζεται στο γεγονός ότι η διαδικασία της απόρριψης μη-προσώπων είναι ταχύτερη από τη διαδικασία εύρεσης προσώπων. Επομένως, γίνεται ξεκάθαρο ότι η χρήση του cascade ταξινομητή είναι αποτελεσματική, αφού η χρονική πολυπλοκότητα παραμένει ίδια οποιαδήποτε και αν είναι η είσοδος.

Ο ταξινομητής cascade αποτελείται από σειριακά τμήματα ταξινομητών στη μορφή καταρράκτη (cascade), καθένα από τα οποία είναι ένας δυνατός ταξινομητής. Η μέθοδος αυτή επιτυγχάνει απόδοση στην ανίχνευση και μειώνει ριζικά το χρόνο υπολογισμού. Το κάθε τμήμα καθορίζει κατά πόσο ένα παράθυρο εισόδου είναι σίγουρα μη-πρόσωπο ή πιθανό πρόσωπο. Αν το παράθυρο ταξινομηθεί ως μη-πρόσωπο από κάποιο τμήμα (δυνατό ταξινομητή), τότε απορρίπτεται. Αν όμως ένα παράθυρο ταξινομηθεί ως πιθανό πρόσωπο, τότε συνεχίζει στο επόμενο τμήμα του ταξινομητή cascade. Προφανώς όσα περισσότερα τμήματα καταφέρει να περάσει ένα παράθυρο χωρίς να απορριφθεί, τόσο μεγαλύτερη είναι η πιθανότητα να περιλαμβάνει τελικά πρόσωπο. Η διαδικασία αυτή αναπαριστάται στο σχήμα 4.4.

Ένα σημαντικό πλεονέκτημα του ταξινομητή cascade είναι η μείωση των false-positive εικόνων (λανθασμένα θετικών). Με τη χρήση ενός δυνατού ταξινομητή, η μείωση αυτή

επιτυγχάνεται με αποδοχή των false-negative εικόνων ώστε να διορθωθεί η αναλογία με τις false-positive. Ωστόσο αυτή η τακτική δεν είναι σωστή. Με τη χρήση όμως ενός ταξινομητή πολλών σταδίων, οι false-positive εικόνες δεν αποτελούν πρόβλημα καθώς αναμένεται να απορριφθούν σε κάποιο στάδιο και να έχουν μείνει πολύ λίγες στο τελικό στάδιο.



Εικ 4.4 : Σχηματική αναπαράσταση μιας ανίχνευσης με καταρράκτη ταξινομητών

Παρακάτω φαίνονται τα βήματα του cascade ταξινομητή. Ο αλγόριθμος απαιτεί σαν είσοδο τον ελάχιστο βαθμό αποδεκτής ανίχνευσης $\text{minimum acceptable detection rate } d$ και το μέγιστο βαθμό false positive εικόνων f . Ακόμη τονίζεται ότι κατά την εκπαίδευση ενός τμήματος (δυνατού ταξινομητή) n , τα αρνητικά δείγματα πρέπει να είναι false-negative δείγματα που προέρχονται από το τμήμα $n - 1$.

Η εκπαίδευση του καταρράκτη ταξινομητή γίνεται χρησιμοποιώντας τον Adaboost και καθορίζει:

- Τον αριθμό των επιπέδων του καταρράκτη ταξινομητή
- Τον αριθμό των χαρακτηριστικών σε κάθε επίπεδο
- Το κατώφλι σε κάθε επίπεδο

ώστε να ελαχιστοποιείται ο αριθμός των χρησιμοποιούμενων χαρακτηριστικών, δεδομένων των f , d και F_{target} .

Αλγόριθμος 1 Αλγόριθμος Adaboost

Input: Βαθμός false-positive f , βαθμός ανίχνευσης d και στόχος βαθμού false-positive F_{target} . Σύνολο P θετικών δειγμάτων (προσώπων) και σύνολο N αρνητικών δειγμάτων (μη-προσώπων).

1: Αρχικοποίηση $F_0 = 1$, $D_0 = 1$, $i = 0$

2: **while** $F_i > F_{\text{target}}$ και $n_i < N$ **do**

3: $i = i + 1$, $n_i = 0$ και $F_i = F_{i-1}$

4: **while** $F_i > F_{i-1} - 1$ **do**

5: $n_i = n_i + 1$

6: Εκτέλεση αλγορίθμου Adaboost με P , N και n_i χαρακτηριστικά

7: Υπολογισμός του τρέχοντος ταξινομητή για το σύνολο που προέκυψε ώστε να καθοριστούν τα F_i και D_i

8: Μείωση της τιμής κατωφλίου του i -οστού ταξινομητή ώστε ο βαθμός ανίχνευσης να γίνει τουλάχιστον d . D_{i-1}

9: **end while**

10: $N = 0$

11: **if** $F_i > F_{\text{target}}$ **then**

12: Υπολογισμός του τρέχοντος ανιχνευτή cascade για τις εικόνες μη-προσώπου και τοποθέτηση τυχόντων λανθασμένων ανιχνεύσεων στο σύνολο N

13: **end if**

4.6 Μέθοδος Haar+Adaboost : Συμπεράσματα

Αναφορικά με την προσέγγιση Haar+AdaBoost [18, 19, 7, 8], μπορούμε να εξάγουμε τα παρακάτω συμπεράσματα [10] σχετικά με τα σύνολα των χαρακτηριστικών, τους αλγόριθμους boosting, τους αδύναμους ταξινομητές, τα μεγέθη των υποπαραθύρων, και τα μεγέθη του συνόλου εκπαίδευσης:

- Οι AdaBoost μέθοδοι εκπαίδευσης στην ανίχνευση προσώπων είναι οι πιο αποτελεσματικές μέχρι σήμερα. Αναφορικά με τα επίπεδα ανίχνευσης και σφαλμάτων, συγκρίνονται με την μέθοδο των νευρωνικών δικτύων του Rowley κ.ά. [14], αλλά σε μερικές περιπτώσεις είναι αρκετές φορές ταχύτερες.
- Ένα υπερπλήρες σύνολο από χαρακτηριστικά τύπου Haar είναι αποτελεσματικό για ανίχνευση προσώπου. Η χρήση της μεθόδου εικόνας ολοκληρώματος κάνει τον υπολογισμό αυτών των χαρακτηριστικών εφικτό και ανεξάρτητο από την κλίμακα.
- Τα εκτεταμένα χαρακτηριστικά τύπου Haar βοηθούν στην ανίχνευση των περιστραμμένων προσώπων.
- Η εκπαίδευση AdaBoost μπορεί να επιλέξει το καλύτερο υποσύνολο από ένα ευρύ σύνολο χαρακτηριστικών και να κατασκευάσει έναν ισχυρό μη γραμμικό ταξινομητή.
- Η διάταξη καταρράκτη βελτιώνει σημαντικά την ταχύτητα ανίχνευσης και μειώνει αποτελεσματικά τα σφάλματα με μικρό κόστος στους χρόνους ανίχνευσης.
- Πιο σύνθετοι αδύναμοι ταξινομητές όπως μικρά δένδρα CART μπορούν να μοντελοποιήσουν δεύτερης ή/και τρίτης τάξης εξαρτήσεις των χαρακτηριστικών, και μπορούν να είναι επωφελείς σε μη γραμμική επεξεργασία της ανίχνευσης προσώπου.
- Το ιδανικό μέγεθος του υπο-παραθύρου για την επεξεργασία ανίχνευσης προσώπων φαίνεται να είναι $20 \times 20 \sim 24 \times 24$ pixels.
- Πιθανές βελτιώσεις μπορεί να είναι εφικτές με τον σχεδιασμό επιπλέον χαρακτηριστικών συμπληρωματικά στα ήδη υπάρχοντα, που υιοθετούν πιο προχωρημένες τεχνικές εκπαίδευσης, και που θα μπορούσαν να καταλήξουν σε πιο σύνθετους ταξινομητές, αποφεύγοντας το πρόβλημα της υπερπροσαρμογής (overfitting).
- Γρήγορο και σχεδόν εύρωστο σύστημα που τρέχει σε πραγματικό χρόνο
- Αρνητικά στοιχεία :
 - Απαιτείται η αναζήτηση στο χώρο και την κλίμακα

- Απαιτεί αρκετά θετικά και αρνητικά δείγματα
- Ανάλωση πολύ χρόνου στη φάση της εκπαίδευσης (μπορεί να απαιτεί μέρες εκπαίδευσης)
- Περιορισμένη προσέγγιση των περιπτώσεων διαφορετικού προσανατολισμού των προσώπων
- Απαιτεί αρκετή εργασία υλοποίησης.

Κεφάλαιο 5

ΑΝΙΧΝΕΥΣΗ ΕΚΦΡΑΣΕΩΝ ΠΡΟΣΩΠΟΥ

5.1 Περιγραφή του προβλήματος	40
5.2 Καθορισμός του προβλήματος	41
5.3 Τεχνικές αναγνώρισης των εκφράσεων	42
5.4 Εφαρμογές	43

5.1 Περιγραφή του προβλήματος

Η έρευνα στην ανάλυση εκφράσεων έχει έλθει ξανά στο προσκήνιο τα τελευταία χρόνια. Η συνεχής ανάπτυξη των εφαρμογών πολυμέσων κατέστησε την αναγνώριση προσώπων αλλά και την ανάλυση εκφράσεων εξαιρετικά σημαντικές εφαρμογές στον τομέα της επικοινωνίας ανθρώπου μηχανής. Οι μέθοδοι αυτόματης ανάλυσης εκφράσεων αποτελούν μια ενεργή περιοχή έρευνας που βρίσκει προοπτικές εφαρμογές στην γενικότερη ερευνητική περιοχή της αλληλεπίδρασης ανθρώπου -υπολογιστή. Για να γίνει όμως αυτό, είναι αναγκαίο να μπορεί ο υπολογιστής να αλληλεπιδρά φυσικά με το χρήστη με έναν τρόπο παρόμοιο με αυτόν που αλληλεπιδρούν οι άνθρωποι μεταξύ τους. Οι εκφράσεις του προσώπου αντανακλούν όχι μόνο ανθρώπινα συναισθήματα αλλά και νοητικές λειτουργίες και ψυχολογικά σήματα και κατέχουν έτσι ένα μεγάλο μέρος της κοινωνικής αλληλεπίδρασης. Αυτός είναι ο λόγος για τον οποίο η αυτόματη αναγνώριση εκφράσεων αποτελεί βασικό τομέα της αλληλεπίδρασης ανθρώπου-υπολογιστή.

Οι εκφράσεις του προσώπου είναι ένα από τα πιο ισχυρά μέσα με τα οποία οι άνθρωποι εξωτερικεύουν τα συναισθήματα τους, τις προθέσεις τους και τις απόψεις τους ο ένας στον άλλον. Εκτός από τις πληροφορίες όμως που αφορούν τη συναισθηματική κατάσταση του ανθρώπου, οι εκφράσεις του προσώπου παρέχουν και πληροφορίες για την νοητική κατάσταση, όπως για το αν κατακλύζεται από ενδιαφέρον, πλήξη, σύγχυση ή άγχος αλλά και εναλλακτικά μηνύματα με πληροφορίες για την έμφαση στο λόγο και τη σύνταξη.



Εικ 5.1 : Οι 7 εκφράσεις: ουδέτερη, θυμός, απέχθεια, φόβος, χαρά, λύπη και έκπληξη

Ο πρωταγωνιστικός ρόλος του προσώπου στη διαπροσωπική επικοινωνία έγινε έναυσμα, τις τελευταίες δεκαετίες, για εκτενείς μελέτες στον τομέα της ψυχολογίας με αντικείμενο τη αναγνώριση συναισθημάτων από τις εκφράσεις του προσώπου, δηλαδή τις διαφοροποιήσεις σε σχέση με την ουδέτερη έκφραση. Ο ψυχολόγος Paul Ekman, μετά από έρευνα ετών σε διάφορους πληθυσμούς σε όλο τον κόσμο, κατέληξε στο συμπέρασμα ότι οι εκφράσεις που αντιπροσωπεύουν τα συναισθήματα χαράς, λύπης, φόβου, θυμού, έκπληξης και απέχθειας είναι παγκόσμιες ή αλλιώς καθολικές, γίνονται δηλαδή κατανοητές και ερμηνεύονται με τον ίδιο τρόπο παγκοσμίως, ανεξαρτήτως πολιτισμικών παραγόντων [3].

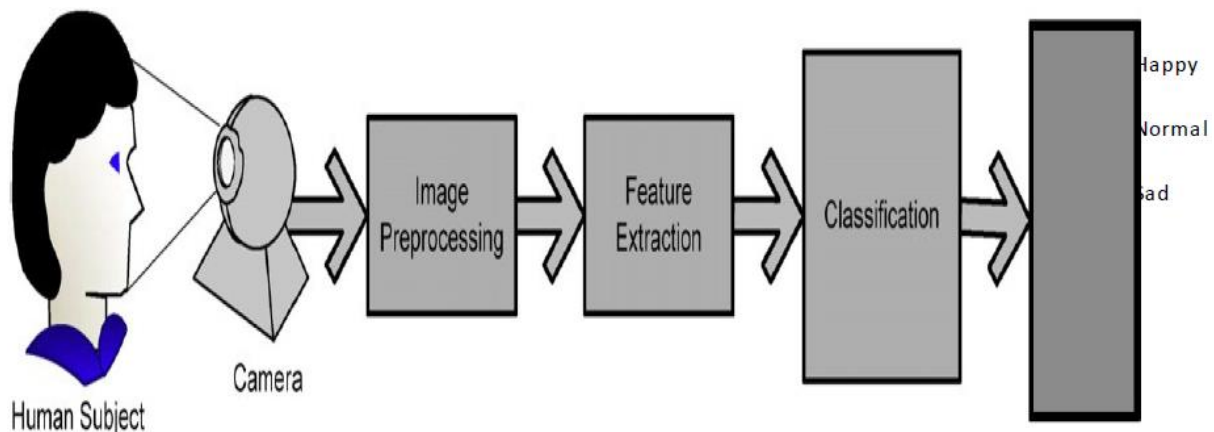
Η αυτόματη αναγνώριση εκφράσεων αποτελεί μια πολυσύνθετη διεργασία καθώς οι φυσιογνωμίες των προσώπων διαφέρουν σημαντικά από άτομο σε άτομο εξαιτίας πολλών παραγόντων, όπως είναι η ηλικία, η εθνικότητα, το φύλο, η τριχοφυΐα του προσώπου, τα διακοσμητικά και καλλυντικά προϊόντα και αντικείμενα που ενδεχομένως κρύβουν σημεία του προσώπου όπως τα γυαλιά μυωπίας ή ακόμη και τα μαλλιά. Επιπλέον τα πρόσωπα παρουσιάζονται ανόμοια λόγω των διαφορετικών συνθηκών φωτισμού. Η έρευνα, όμως της αναγνώρισης συναισθημάτων είναι σε έξαρση και αναπτύσσονται πολλές διαφορετικές προσεγγίσεις μεταξύ των οποίων κάποιες επιτυγχάνουν ιδιαίτερα ενθαρρυντικά αποτελέσματα.

5.2 Καθορισμός του προβλήματος

Ενώ, για τους ανθρώπους, η ανίχνευση του προσώπου επιτυγχάνεται σχεδόν αβίαστα και η ερμηνεία των εκφράσεων σχετικά εύκολα, η ανάπτυξη ενός αυτόματου συστήματος που θα εντοπίζει και θα ταξινομεί με επιτυχία τις εκφράσεις του προσώπου, σε πραγματικό χρόνο, δεν είναι εύκολη υπόθεση.

Ένα τυπικό σύστημα αναγνώρισης του προσώπου έκφρασης εμφανίζεται στο σχήμα 5.2 παρακάτω. Η διαδικασία αναγνώρισης αρχίζει με την απόκτηση της εικόνας χρησιμοποιώντας μια συσκευή λήψεων εικόνων όπως μια φωτογραφική μηχανή ή μια WebCam. Στην συνέχεια υλοποιούνται τα εξής βήματα:

- Αναγνώριση Προσώπου (Face Recognition) στην εικόνα και προεπεξεργασία των προσώπων (σε grey value, προσαρμογή μεγέθους, μετατροπή κ.τ.λ.)
- Κατάτμηση, έτσι ώστε να βρεθούν τα πιο σημαντικά τμήματα του προσώπου από τον αλγόριθμο.
- Υπολογισμός των μεμονωμένων γνωρισμάτων στην εικόνα. Αυτά τα γνωρίσματα είναι αμετάβλητα κατά την κλιμάκωση.
- Ταξινόμηση αυτών των βασικών γνωρισμάτων και καταγραφή των υπάρχων “εκπαιδευμένων” δεδομένων.



Εικ 5.2 : Διάγραμμα συστήματος αναγνώρισης ανθρώπινων εκφράσεων

5.3 Τεχνικές αναγνώρισης εκφράσεων

Στην παράγραφο αυτή παρουσιάζονται μερικές από τις τεχνικές οι οποίες έχουν εμφανιστεί και πραγματεύονται τους τρόπους με τους οποίους ο υπολογιστής μπορεί να ανακτήσει πληροφορίες σχετικά με τη συναισθηματική κατάσταση κάποιου ατόμου μέσα από τις εκφράσεις του προσώπου του. Οι προσεγγίσεις σχετικά με την ανάλυση εκφράσεων διακρίνονται χονδρικά σε τρεις κατηγορίες ανάλογα με την πηγή της πληροφορίας που χρησιμοποιείται:

- **στατικές:** χρήση εικόνων προσώπων που απεικονίζουν κάποια έκφραση,
- **ημιστατικές:** χρήση δύο εικόνων μια με το πρόσωπο σε ουδέτερη κατάσταση και μία με το πρόσωπο στη κορύφωση της έκφρασης και

- **δυναμικές:** χρήση ακολουθίας βίντεο στην οποία απεικονίζεται η εξέλιξη της έκφρασης. Με τη συγκεκριμένη προσέγγιση θα ασχοληθούμε στην παρούσα διπλωματική, αφού ο χρήστης θα βρίσκεται μπροστά από τη κάμερα και θα ανιχνεύουμε τις εκφράσεις του.

5.4 Εφαρμογές

Ένα σύστημα με τα προαναφερθέντα χαρακτηριστικά, μπορεί να αποτελέσει τμήμα μιας πληθώρας εφαρμογών που θα αναβαθμίσουν την ποιότητα πολλών δραστηριοτήτων. Ενδεικτικά, αναφέρονται μερικοί τρόποι αξιοποίησης:

- Βοήθημα ανθρώπων με ιδιαιτερότητες, όπως ένα σύστημα ενίσχυσης της αναγνώρισης συναισθημάτων από άτομα με προβλήματα όρασης, ή ως ένα μέσο υποστήριξης και εκπαίδευσης παιδιών με αυτισμό.
- Βοήθημα ανθρώπων που παρουσιάζουν κάποια μορφή αναπηρίας. Ο παραδοσιακός τρόπος επικοινωνίας ανθρώπου και υπολογιστή πραγματοποιείται με τη χρήση πληκτρολογίου και ποντικιού. Εναλλακτικά μπορεί να πραγματοποιηθεί με τη χρήση μικροφώνου και με διαδικασίες λεκτικής ανάλυσης. Για ανθρώπους με σοβαρά κινητικά προβλήματα αλλά και με προβλήματα ομιλίας και ακοής, η επικοινωνία μέσω κινήσεων του κεφαλιού ή των περιοχών του προσώπου όπως τα χείλη, τα μάτια και τα βλέφαρα είναι ιδιαίτερα επιθυμητή.
- Πρόσθετο εργαλείο στην ιατρική φροντίδα, για παράδειγμα σε ένα ολοκληρωμένο σύστημα νοσηλευτικής παρακολούθησης χρόνιων ασθενών, ώστε να ανιχνεύεται και να αντιμετωπίζεται αναλόγως τυχόν αρνητική συναισθηματική αλλαγή (αίσθημα πανικού ή δυσανασχέτησης).
- Μέσο βελτίωσης των εφαρμογών εκπαίδευσης από απόσταση αλλά και των προγραμμάτων εκμάθησης μέσω υπολογιστή, δίνοντας τη δυνατότητα προσαρμογής του τρόπου και του ρυθμού διδασκαλίας σύμφωνα με τις αντιδράσεις του χρήστη.
- Λοιπές εφαρμογές: βελτίωση των λογισμικών προσωπικών υπολογιστών, των διαδραστικών ηλεκτρονικών παιχνιδιών, των μεθόδων μάρκετινγκ και διαφήμισης κ.ά.

Ένα σύστημα επικοινωνίας ανθρώπου υπολογιστή θα πραγματοποιηθεί στα πλαίσια της διπλωματικής εργασίας με τη χρήση μιας απλής βιντεοκάμερας. Στο σύστημα αυτό εμφανίζονται κάποια επιμέρους προβλήματα τα οποία απαιτούν αξιόπιστη λύση και αφορούν το τμήμα σύλληψης και επεξεργασίας της ακολουθίας βίντεο. Η επίλυση συγκεκριμένων

θεμάτων όπως ο εντοπισμός της θέσης και η παρακολούθηση της κίνησης του προσώπου, ο εντοπισμός των βασικών χαρακτηριστικών του προσώπου και η παρακολούθηση της κίνησης τους καθώς και η εύρεση της θέσης συγκεκριμένων σημείων στην επιφάνεια του, αποτελούν τον κορμό του ανωτέρω σχήματος επικοινωνίας. Στην ανάλυση εκφράσεων τα χαρακτηριστικά του προσώπου είναι ιδιαίτερα σημαντικά και η αποτελεσματική ανίχνευση τους καθίσταται επιτακτική. Για το σκοπό αυτό τα χαρακτηριστικά πρέπει να είναι ιδιαίτερα ευδιάκριτα.

Η κατανόηση των συναισθημάτων είναι επίσης ένα εξαιρετικά σημαντικό θέμα στην επικοινωνία ανθρώπου - μηχανής. Δεδομένης της δυσκολίας με την οποία οι άνθρωποι ερμηνεύουν την εσωτερική κατάσταση των συνανθρώπων τους είναι αντιληπτό από όλους ότι η κατανόηση των ανθρωπίνων συναισθημάτων από τον υπολογιστή αποτελεί μια μεγάλη πρόκληση. Το αντικείμενο της εργασίας αυτής αποτελούν οι εκφράσεις της χαράς, της λύπης, της έκπληξης και της ουδέτερης έκφρασης.

Κεφάλαιο 6

ΥΛΟΠΟΙΗΣΗ ΣΥΣΤΗΜΑΤΟΣ ΑΝΙΧΝΕΥΣΗΣ ΠΡΟΣΩΠΟΥ

6.1 Γενικά	45
6.2 Βιβλιοθήκες που έχουν ελεγχθεί	46
6.3 Kinect SDK 1.5	46
6.4 Βιβλιοθήκη OpenCV	51
6.5 Βιβλιοθήκη BRF	52
6.6 Εφαρμογή 1	52
6.7 Εφαρμογή 2	53
6.7.1 Αναγνώριση χαμόγελου	56
6.7.2 Πώς να αποφασίσουμε ένα σωστό κατώφλι	57
6.7.3 Αναγνώριση στρογγυλού σχήματος χειλιών	58
6.7.4 Αναγνώριση ανασήκωσης φρυδιών	58
6.7.5 Αναγνώριση συνοφρύωσης φρυδιών	59
6.8 Αποτελέσματα	
6.9 Προβλήματα που αντιμετωπίσαμε	61
6.9.1 Περιορισμός του προσώπου εντός ενός πλαισίου	61
6.9.2 Διαφορετικότητα στα πρόσωπα	61
6.9.3 Διαφορετικότητα στις εκφράσεις των προσώπων	62

6.1 Γενικά

Σε αυτή την διπλωματική εργασία θα ασχοληθούμε με face tracking και πώς αυτό μπορεί να χρησιμοποιηθεί σε εφαρμογές ψυχαγωγίας αλλά και σε εφαρμογές ανίχνευσης συναισθημάτων. Τέτοιες εφαρμογές μπορεί να είναι είτε ως πρόσθετο εργαλείο στην ιατρική φροντίδα, είτε ως μέσο βελτίωσης εφαρμογών όπως έχουν επεξηγηθεί στο κεφάλαιο 5.4. Αυτές οι εφαρμογές μπορούν να αλλάξουν τον τρόπο που οι άνθρωποι παίζουν παιχνίδια και την εμπειρία ψυχαγωγίας. Επίσης να μεταμορφώσει το πώς οι άνθρωποι αλληλεπιδρούν με τους υπολογιστές, περιφερειακά και με άλλες συσκευές ελεγχόμενης κίνησης. Δίνουν τη

δυνατότητα στους χρήστες να αλληλεπιδρούν με φυσικό τρόπο με τους υπολογιστές με απλές χειρονομίες.

Αποφασίσαμε να επεκτείνουμε και τις δύο δυνατότητες του face tracking, που μπορεί να είναι είτε στο χώρο της ψυχαγωγίας είτε στο χώρο των e-learning συστημάτων με ανίχνευση συναισθημάτων.

Έτσι δημιουργήσουμε δύο ξεχωριστές εφαρμογές, εφόσον η χρησιμότητα τους διαφέρει για να είναι πιο ξεκάθαρη η διαφοροποίηση στη χρήση τους. Η πρώτη εφαρμογή θα χρησιμοποιήσει την τεχνολογία face tracking για ανίχνευση προσώπου και πρόσθεση διαφόρων αντικειμένων πάνω από το πρόσωπο όπως γυαλιά, καπέλο κλπ. Τέτοιου είδους εφαρμογή υπάγεται στην κατηγορία ψυχαγωγίας όπου ο χρήστης μπορεί να δει το πρόσωπο του στην οθόνη του υπολογιστή μέσω της webcam με πρόσθετα διάφορα αντικείμενα, για παράδειγμα να φορέσει εικονικά γυαλιά ηλίου. Ενώ η δεύτερη εφαρμογή ασχολείται με την ανίχνευση εκφράσεων και συναισθημάτων και υπάγεται στην κατηγορία των e-learning συστημάτων. Γι' αυτό και προτιμήσαμε να διαχωρίσουμε τις δύο εφαρμογές εφόσον είναι δύο διαφορετικές πτυχές του face tracking αλλά εξίσου σημαντικές και ενδιαφέρουσες.

6.2 Βιβλιοθήκες που έχουν ελεγχθεί

Για τις δύο εφαρμογές έχει χρησιμοποιηθεί η ίδια τεχνολογία (βιβλιοθήκη) για την ανίχνευση προσώπου. Έχουμε δοκιμάσει διάφορες βιβλιοθήκες γνωστές στο χώρο για face tracking όπως Kinect SDK 1.5, Open CV, BRF.

6.3 Kinect SDK 1.5

Το Kinect είναι μια συσκευή ανίχνευσης κινήσεων από την Microsoft για τη βιντεοκονσόλα Xbox360 και για Windows PCs. Με βάση γύρω από μια περιφερειακή κάμερα, επιτρέπει στους χρήστες να ελέγχουν και να αλληλεπιδρούν με το Xbox360 μέσα από φυσικές χειρονομίες και προφορικές εντολές, χωρίς να χρειάζεται να αγγίζουν κάποιο game controller. Παρέχει τη δυνατότητα στους προγραμματιστές να χτίσουν εφαρμογές σε C++, C# ή Visual Basic και παρέχει τις εξής δυνατότητες:

1. Ανίχνευση σκελετού: Η δυνατότητα να παρακολουθεί τον σκελετό ενός ή δύο ατόμων που διακινούνται εντός του οπτικού πεδίου της Kinect για εφαρμογές που είναι gesture-driven.

2. Ακατέργαστη ροή αισθητήρων: επιστρέφει ακατέργαστες πληροφορίες σχετικές με την κίνηση, το βάθος και χρώμα

3. Προηγμένες δυνατότητες ήχου: δυνατότητες επεξεργασίας ήχου για την αναγνώριση της τρέχουσας πηγής ήχου, και την ενσωμάτωση με την ομιλία αναγνώρισης των Windows API.

Ενώ το Kinect SDK 1.5 είναι πολύ δυνατό στο χώρο του face tracking και γενικά των gesture-driven εφαρμογών, εντούτοις το κύριο μειονέκτημα του εναπόκειται στο ότι είναι απαραίτητο να έχεις τη συσκευή Kinect για να μπορέσεις να το δουλέψεις. Εφόσον μας παρέχει πολύ περισσότερα από όσα χρειαζόμαστε και υπάρχουν άλλες βιβλιοθήκες εξίσου καλές που δεν προαπαιτούν τη χρήση οποιασδήποτε συσκευής πέραν από ένα υπολογιστή και μια κάμερα, αποφασίσαμε να μη χρησιμοποιήσουμε αυτή τη βιβλιοθήκη.

6.4 Βιβλιοθήκη OpenCV

Η "Open Source Computer Vision Library" (OpenCV) είναι μία ελεύθερα διαθέσιμη βιβλιοθήκη, ανοικτού κώδικα συλλογή από ρουτίνες σε C++, που αφορούν την τεχνητή όραση. Η OpenCV παρέχει μια στέρεη υποδομή για την τεχνητή όραση που επιτρέπει στους ειδικούς να δουλέψουν σε ένα υψηλότερο επίπεδο από το στοιχειώδες. Με την τύπου BSD (Berkeley Software Distribution) άδεια χρήσης της, η OpenCV παρέχεται από την Intel για ελεύθερη εμπορική και ερευνητική χρήση.

Η υποστήριξη της OpenCV στην τεχνητή όραση είναι εκτεταμένη. Υποστηρίζει ρουτίνες για την ανάγνωση, απεικόνιση, επεξεργασία και αποθήκευση απλών φωτογραφικών εικόνων αλλά και κινηματογραφικών ταινιών. Παρέχονται ένα πλήθος από ρουτίνες για την επεξεργασία εικόνας και υποστηρίζονται ρουτίνες υπολογιστικής γεωμετρίας, αναγνώρισης προτύπων και αλγόριθμοι ανίχνευσης αντικειμένων.

Το πιο γνωστό project μέχρι στιγμής που έχει υλοποιηθεί με την βιβλιοθήκη OpenCV είναι το "Marilena" project που ανιχνεύει πρόσωπα σε στατικές φωτογραφίες. Αυτό είναι το κυριότερο μειονέκτημα της OpenCV για τη δική μας περίπτωση εφόσον εμείς θέλουμε αναγνώριση προσώπου από webcam και όχι από στατικές φωτογραφίες.

Κάτι τέτοιο έγινε από την βιβλιοθήκη BRF, όπου ο σκοπός ήταν να μεταφερθεί η OpenCV σε Flash Professional, AS3 με την επέκταση ότι θα δέχεται είσοδο από το video της webcam και θα γίνεται συνεχής face tracking.

6.5 Βιβλιοθήκη BRF

Η βιβλιοθήκη αυτή (που είναι διαθέσιμη δωρεάν για ακαδημαϊκούς σκοπούς) μας παρέχει τις δυνατότητες της OpenCV αλλά γραμμένη σε Flash AS3 και με την επέκταση ότι οι υπολογισμοί γίνονται συνεχώς για την ανίχνευση και αναγνώριση του προσώπου, ενώ αυτό βρίσκεται σε κίνηση (βίντεο μέσω της webcam). Είναι αρκετά καλά τεκμηριωμένη με πολύ καλό API επομένως θεωρήσαμε την καλύτερη επιλογή από τις διαθέσιμες βιβλιοθήκες.

Η BRF μας παρέχει τα εξής:

1. Αναγνώριση Προσώπου - βρίσκει την τραχεία θέση ενός προσώπου σε μια εικόνα ή σε webcam stream
2. Εκτίμηση Προσώπου - Ανάλυση του προσώπου που βρέθηκε και υπολογισμός των χαρακτηριστικών του προσώπου / χαρακτηριστικά (πχ μάτια, χείλη, φρύδια)
3. Εκτίμηση Πόζας - υπολογίζει τη 3D θέση του κεφαλιού του προσώπου που βρέθηκε
4. Face Tracking - Ακολουθεί το πρόσωπο καθώς αυτό κινείται σε ένα webcam stream

Εμείς έχουμε επεκτείνει τις δυνατότητες της BRF για την υλοποίηση εφαρμογών ψυχαγωγίας και για σκοπούς e-learning. Παίρνοντας το αποτέλεσμα που μας παρέχει για το πρόσωπο που εντοπίστηκε, υπολογίζουμε την περιοχή που μας ενδιαφέρει και φορτώνουμε ένα 3D μοντέλο με την βιβλιοθήκη FLARE3D και το τοποθετούμε στο πρόσωπο. Για παράδειγμα, υπολογίζουμε την περιοχή των ματιών και το μέγεθος τους και τοποθετούμε το μοντέλο μας, σε αυτή την περίπτωση είναι γυαλιά ηλίου, στην ανάλογη θέση και στο ανάλογο μέγεθος. Το αντίστοιχο θα συμβεί αν αλλάξει μοντέλο τότε θα γίνουν οι ανάλογες ρυθμίσεις για να τοποθετηθεί σωστά στο πρόσωπο.

Στην περίπτωση ανίχνευσης εκφράσεων και συναισθημάτων έχει γίνει πιο εκτεταμένη μελέτη των σημείων του προσώπου και πώς αυτά μετακινούνται αναλόγως της έκφρασης. Από την βιβλιοθήκη μας επιστρέφεται ένας πίνακας με σημεία του προσώπου συνολικά 68 σημεία (0-67). Τα σημεία που επιστρέφονται είναι γύρω από την περιοχή των ματιών, των φρυδιών, της μύτης, του στόματος και του περιγράμματος του προσώπου. Πιο κάτω θα γίνει πιο εκτεταμένη επεξήγηση των δύο εφαρμογών και των αλγορίθμων τους.

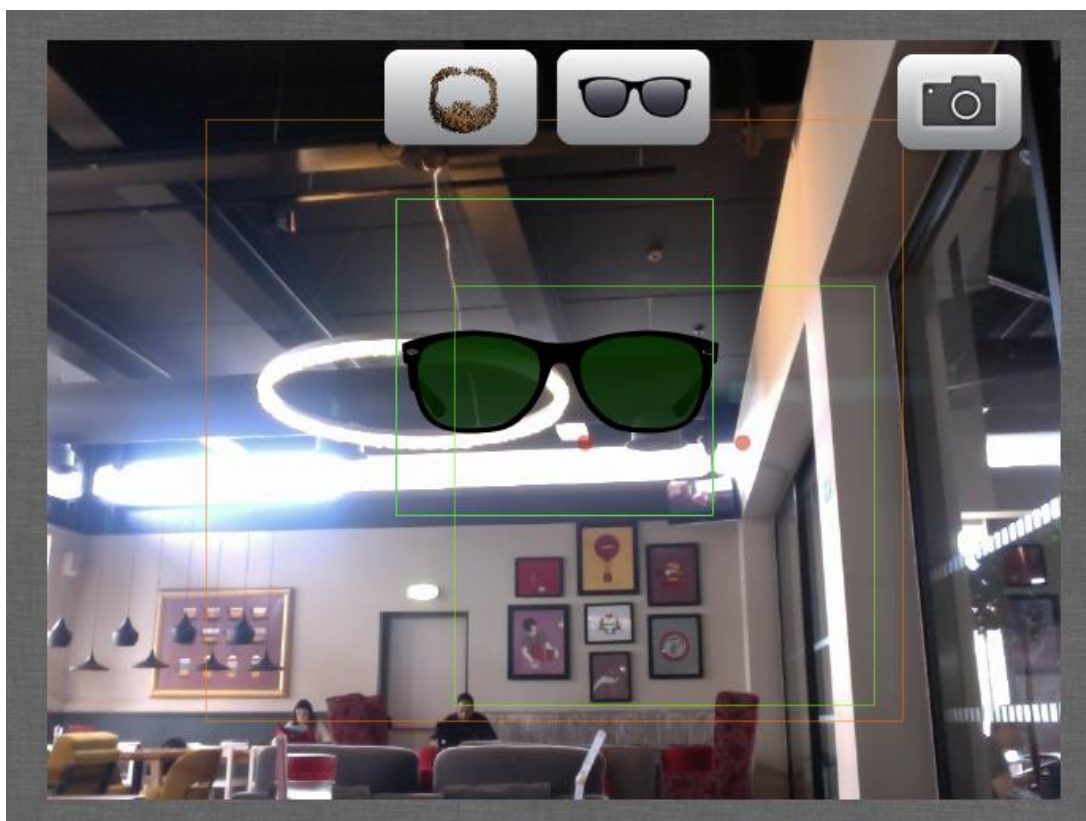
Για την χρήση της βιβλιοθήκης BRF χρειαζόμαστε:

1. Βιντεοκάμερα συνδεδεμένη με τον υπολογιστή (webcam)
2. Υπολογιστή με Flash Player 11
3. Καλά φωτισμένο δωμάτιο

Εφόσον οι πιο πάνω συνθήκες ισχύουν τότε μπορούμε να χρησιμοποιήσουμε σωστά την εφαρμογή.

6.6 Εφαρμογή 1

Στην πρώτη εφαρμογή, ο χρήστης τοποθετεί το πρόσωπο του στο κέντρο της βιντεοκάμερας και της οθόνης μέχρι να γίνει εύρεση του προσώπου εντός της εικόνας της βιντεοκάμερας. Αν βρέθηκε πρόσωπο γίνεται ανίχνευση των συντεταγμένων (x,y,z) του προσώπου και στη συνέχεια ανίχνευση των χαρακτηριστικών του προσώπου (μάτια, χείλη, μύτη, φρύδια, περίγραμμα προσώπου). Εφόσον το πρόσωπο κινείται στο χώρο της βιντεοκάμερας, γίνεται συνεχής υπολογισμός της κατεύθυνσης και της θέσης του προσώπου. Μόλις εντοπιστεί το πρόσωπο, φορτώνεται το πρώτο μοντέλο που είναι τα γυαλιά ηλίου και εφαρμόζεται στο πρόσωπο. Αν πατήθηκε κουμπί αλλαγής μοντέλου, τότε γίνεται αλλαγή μοντέλου και η αντίστοιχη εφαρμογή στο πρόσωπο. Ο χρήστης έχει τη δυνατότητα να φωτογραφίσει τον εαυτό του με το μοντέλο και να φυλάξει τη φωτογραφία στον υπολογιστή του.



Εικ. 6.1 : Η εικόνα μόλις ξεκινήσει η εφαρμογή

Η πιο πάνω εικόνα εμφανίζεται μόλις ξεκινήσει η εφαρμογή, κατά την οποία πρέπει να τοποθετηθεί το πρόσωπο στα συγκεκριμένα πλαίσια και να πάρει την κατάλληλη θέση ώστε να προσαρμοστούν τα γυαλιά στο πρόσωπο. Ο χρήστης τοποθετεί το πρόσωπο του στο κέντρο ανάμεσα στα πράσινα τετράγωνα και γίνεται η ανίχνευση προσώπου. Πρώτα φορτώνεται το μοντέλο με τα γυαλιά και αντίστοιχα μπορεί να επιλέξει τα γυαλιά ή το μούσι. Ο χρήστης μπορεί να μετακινηθεί στο χώρο και τα αντικείμενα να συνεχίσουν να ακολουθούν τη διαδρομή του προσώπου. Ανάλογα με τα μοντέλα θα πρέπει να γίνουν και οι αντίστοιχες μετατροπές. Όπως για παράδειγμα, τα γυαλιά πρέπει να τοποθετηθούν στην περιοχή των ματιών ενώ το μουστάκι στην περιοχή των χειλιών. Επομένως για κάθε αντικείμενο γίνονται οι ανάλογες μετατροπές στην τοποθεσία και μέγεθος του μοντέλου. Τα αντικείμενα μας είναι 3D και μπορούν να ακολουθήσουν μέχρι και 30 μοίρες κλίση στο κεφάλι του χρήστη. Για την υλοποίηση 3D έχει χρησιμοποιηθεί η βιβλιοθήκη FLARE 3D που είναι δωρεάν και πιο απλή σε σχέση με άλλες βιβλιοθήκες του Flash (όπως η Papervision ή Away 3D).

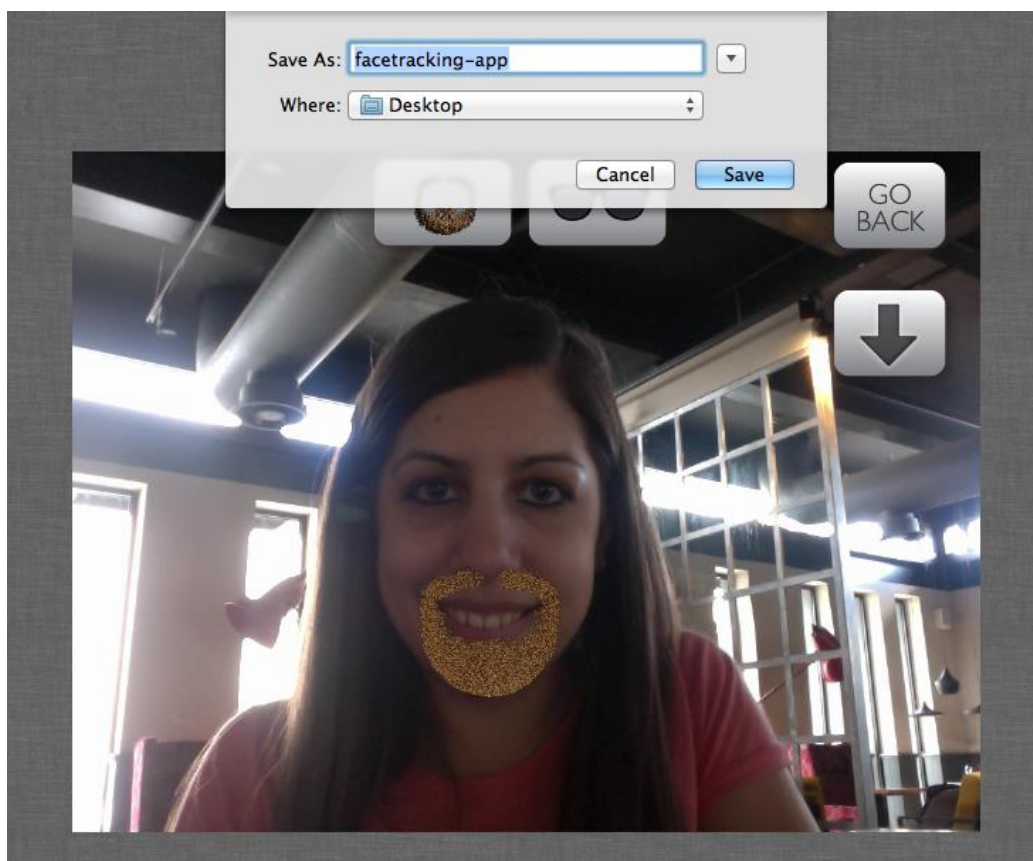


Εικ 6.2: Τοποθέτηση του μοντέλου των γυαλιών στο πρόσωπο του χρήστη

Για να προσθέσουμε περισσότερο χαρακτήρα σε θέμα ψυχαγωγίας της εφαρμογής, προσθέσαμε ένα κουμπί για να βγάζουμε φωτογραφία τον εαυτό μας με ένα αντικείμενο και στη συνέχεια να το αποθηκεύουμε στον υπολογιστή μας, όπως φαίνεται στις εικόνες 6.3 και 6.4.



Εικ. 6.3 : Τοποθέτηση δεύτερου μοντέλου στο πρόσωπο του χρήστη και λήψη φωτογραφίας



Εικ 6.4 : Ο χρήστης μπορεί να φωτογραφήσει τον εαυτό του και να αποθηκεύσει τη φωτογραφία στον υπολογιστή του

Ένα από τα προβλήματα αυτή της εφαρμογής είναι η διαφορετικότητα στα πρόσωπα διαφορετικών ανθρώπων. Αυτό έχει ως αποτέλεσμα το μοντέλο να μην είναι ακριβώς τοποθετημένο στα σωστά σημεία στον κάθε άνθρωπο. Η εφαρμογή μπορεί να επεκταθεί σε μελλοντική εργασία έτσι ώστε να προστεθούν controllers για ρύθμιση του μεγέθους και των x, y, z συντεταγμένων του μοντέλου έτσι ώστε να εφαρμόζεται καλύτερα στο πρόσωπο του κάθε ανθρώπου. Να μπορεί δηλαδή ο χρήστης να μετακινήσει το μοντέλο για να το τοποθετήσει ή να το μεγαλώσει αναλόγως.

Τέτοια εφαρμογή είναι κατάλληλη για σκοπούς διαφήμισης, όπως για παράδειγμα:

1. Σε ηλεκτρονικό κατάστημα πώλησης γυαλιών οράσεως ή μυωπίας όπου ο χρήστης θα μπορεί να δοκιμάσει τα εικονικά γυαλιά στο πρόσωπο του.
2. Σε κομμωτήριο όπου ο χρήστης θα δοκιμάζει διάφορα χτενίσματα, χρώματα και στυλ στο πρόσωπο του.

Ή ακόμη για σκοπούς ψυχαγωγίας, όπως σε κάμερες κατά τη διάρκεια που ο χρήστης βγάζει φωτογραφία τον εαυτό του να μπορέσει να προσθέσει αντικείμενα στο πρόσωπο του.

6.7 Εφαρμογή 2

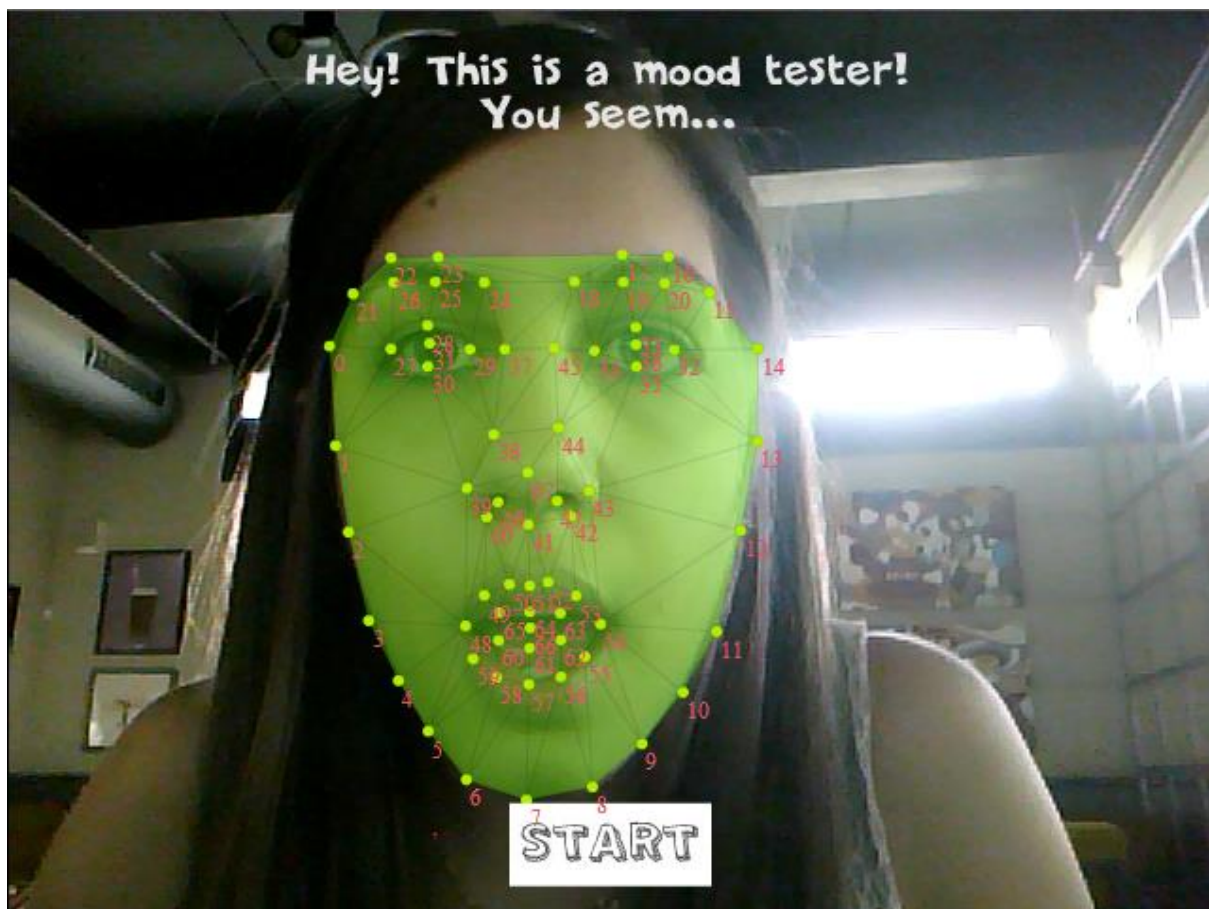
Σε αυτή την εφαρμογή υλοποιήσαμε το κομμάτι ανίχνευσης εκφράσεων του προσώπου. Οι πιο βασικές εκφράσεις του προσώπου είναι χαρούμενος, έκπληκτος, θυμωμένος, αηδιασμένος, λυπημένος και φοβισμένος όπως φαίνονται στη φωτογραφία πιο κάτω:



Εικ. 6.5 Ανθρώπινες Εκφράσεις

Στη συγκεκριμένη διπλωματική εργασία θα ασχοληθούμε με τις πρώτες τρεις εκφράσεις που είναι χαρούμενος, έκπληκτος και θυμωμένος. Παρακολουθώντας τα πρόσωπα μπορούμε να βγάλουμε κάποια συμπεράσματα σχετικά με την κίνηση των χαρακτηριστικών. Όταν κάποιος είναι χαρούμενος παρακολουθούμε ότι χαμογελά ενώ το υπόλοιπο πρόσωπο βρίσκεται σε ηρεμία, ενώ όταν είναι έκπληκτος ή θυμωμένος συμμετέχουν και άλλα χαρακτηριστικά στην έκφραση. Όταν είναι έκπληκτος υπάρχει ανασήκωση φρυδιών και άνοιγμα του στόματος ενώ όταν είναι θυμωμένος υπάρχει συνοφρύωμα φρυδιών και κλείσιμο στόματος. Δηλαδή για να δηλώσουμε κάποια συναισθήματα έχουμε συνδυασμό εκφράσεων. Γι'αυτό αποφασίσαμε να δημιουργήσουμε εφαρμογή αναγνώρισης συναισθημάτων που είναι ένας συνδυασμός εκφράσεων.

Στην αναγνώριση προσώπου προστίθεται μια μάσκα πάνω από το πρόσωπο που αποτελείται από πολύγωνα. Το σημεία αυτά είναι αριθμημένα έτσι ώστε να μπορούμε να χειριστούμε το κάθε σημείο ξεχωριστά και να γνωρίζουμε τις συντεταγμένες x, y του κάθε σημείου ανά πάσα στιγμή. Στη φωτογραφία 6.6 μπορούμε να δούμε τα σημεία που επιστρέφονται στην οθόνη. Τα σημεία ξεκινάνε από το 0 που είναι στο δεξί μάγουλο του περιγράμματος και καταλήγουν στη μέση της μύτης με αριθμό 67, δηλαδή 68 σημεία στο σύνολο τους. Μπορούμε επίσης να παίρνουμε σημεία της κάθε περιοχής, όπως για παράδειγμα τα σημεία γύρω από το κάτω χείλος, άνω χείλος, αριστερό μάτι, δεξί μάτι, αριστερό φρύδι, δεξί φρύδι.



Εικ 6.6 : Τοποθέτηση μάσκας και αρίθμηση όλων των σημείων του προσώπου

Για την αναγνώριση εκφράσεων έχει γίνει προσεκτική μελέτη των σημείων και πώς αυτά μετακινούνται αναλόγως των κινήσεων, όπως για παράδειγμα ο τρόπος με τον οποίο μετακινούνται τα σημεία όταν το πρόσωπο χαμογελά. Οι κινήσεις με τις οποίες ασχοληθήκαμε είναι η αναγνώριση χαμόγελου, ανασήκωση φρυδιών, συνοφρύωμα φρυδιών

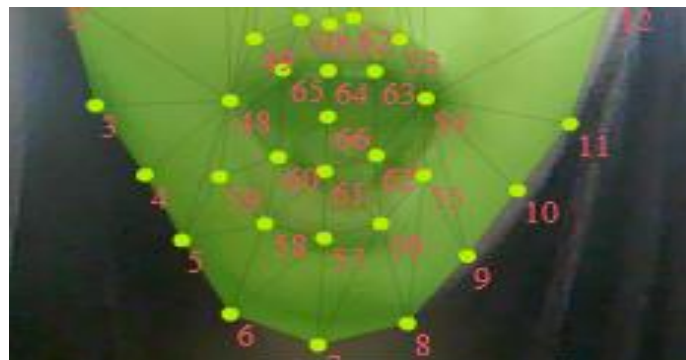
και στρογγυλός σχηματισμός χειλιών (“ο”) που μπορούν να μας βοηθήσουν να αναγνωρίσουμε τις εκφράσεις που αναφέραμε πιο πάνω.

Πριν ξεκινήσει η αναγνώριση εκφράσεων γίνεται ένας υπολογισμός των σταθερών σημείων και αποστάσεων του προσώπου όταν αυτό βρίσκεται σε ηρεμία, όπως:

- Χαμόγελο: Η απόσταση μεταξύ της αριστερής και δεξιάς άκρης των χειλιών στον άξονα x. Σε αυτή την περίπτωση δε μας ενδιαφέρει η ψ απόσταση εφόσον ο χρήστης μπορεί να μετακινηθεί στον άξονα ψ .



- Σχηματισμός στρογγυλού σχήματος χειλιών: Η ψ απόσταση μεταξύ των μέσων σημείων του άνω και κάτω χείλους.



- Για ανασήκωση και συνοφρύωμα φρυδιών: Τα κορυφαία σημεία του αριστερού και δεξιού φρυδιού στον άξονα ψ .



Ο υπολογισμός των σταθερών τιμών γίνεται έτσι ώστε να μπορούμε να υπολογίσουμε πότε το πρόσωπο είναι σε ηρεμία και πότε εκτελεί μια έκφραση. Όπως επίσης και για το λόγο ότι ο κάθε άνθρωπος έχει διαφορετικές αποστάσεις στο πρόσωπο του, έτσι να μπορεί να γίνει νέος υπολογισμός των αποστάσεων για καινούργιο πρόσωπο.

6.7.1 Αναγνώριση χαμόγελου

Για την αναγνώριση χαμόγελου ακολουθήσαμε τον εξής αλγόριθμο [22]:

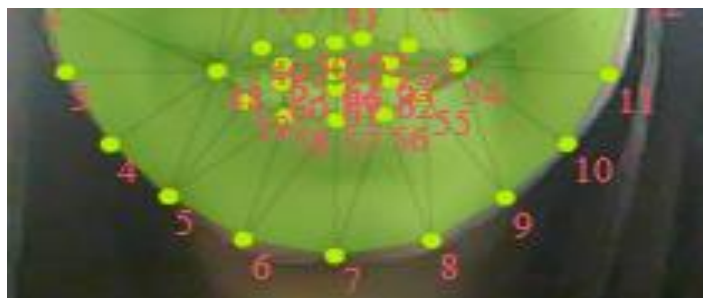
1. Αναγνώριση του ανθρώπινου προσώπου στο πρώτο frame και αναγνώριση και τοποθέτηση των 68 σημείων που αναφέραμε πιο πριν. Υπολογισμός της σταθερής απόστασης D_{std} μεταξύ των άκρων του στόματος,
2. Σε κάθε image frame, ακολουθούμε τις θέσεις των δύο ακρών των χειλιών, την δεξιά και αριστερή άκρη και υπολογίζουμε την απόσταση μεταξύ των δύο σημείων D_x .
3. Αν η D_x απόσταση ανάμεσα στην αριστερή και δεξιά άκρη του στόματος, είναι μεγαλύτερη από την σταθερή απόσταση D_{std} συν ένα κατώφλι T_{smile} τότε θεωρούμε ότι υπάρχει χαμόγελο.

$$D_x > D_{std} + T_{smile}$$

=> τότε έχουμε χαμόγελο

4. Επαναλαμβάνουμε από το βήμα 2 και 3.

Θεωρούμε ότι η x απόσταση ανάμεσα στο δεξή και αριστερό σημείο του στόματος, παίζει σημαντικό ρόλο στην ανθρώπινη δράση του χαμόγελου. Αντίθετα δεν μας επηρεάζει η απόσταση y δεδομένου ότι ο χρήστης μπορεί να μετατοπιστεί λίγο πάνω ή κάτω ή να υπάρξει περιστροφή της κεφαλής και τα οποία θα σημάνουν ψευδώς τον ανιχνευτή μας.



6.7.2 Πώς να αποφασίσουμε ένα σωστό κατώφλι

Εάν έχουμε μέση απόσταση μεταξύ της αριστερής και τη δεξιάς γωνιάς του στόματος. Σε κάθε frame έχουμε D_{mean} που είναι η απόσταση μεταξύ των δύο γωνιών του στόματος. Τότε αν η D_{mean} είναι μεγαλύτερη από την απόσταση $D_{std} + T_{smile}$, τότε θεωρείται ότι έχουμε χαμόγελο, αλλιώς όχι. Έχουμε δοκιμάσει διάφορες τιμές για το βέλτιστο για να δούμε ποια τιμή είναι η καλύτερη [22].

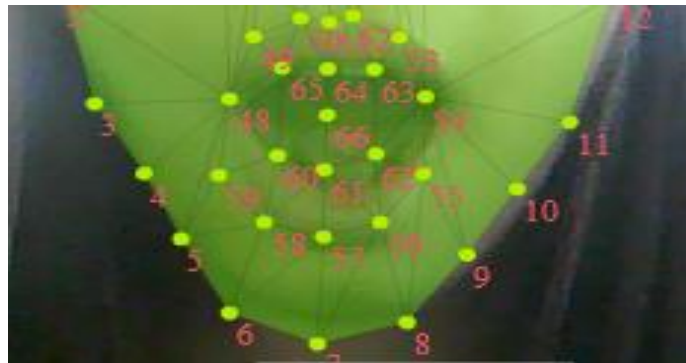
Κατώφλι Tsmile	Αληθής Ανίχνευση (10)
0.02 * Dstd	6
0.03 * Dstd	9
0.04 * Dstd	8
0.06 * Dstd	6
0.08 * Dstd	4

Πίνακας για εύρεση κατωφλίου

Όσο πιο μικρό το κατώφλι τόσο πιο ευαίσθητο είναι το σύστημα στην ανίχνευση χαμόγελου με αποτέλεσμα να ανιχνεύει χαμόγελο ακόμη και σε μικρές κινήσεις. Ενώ όσο πιο μεγάλο το κατώφλι τόσο πιο δύσκολη γίνεται η ανίχνευση άρα τόσο πιο μεγάλη και έντονη πρέπει να είναι η κίνηση των χειλιών στο χαμόγελο για να γίνει αναγνώριση. Παρατηρούμε ότι με κατώφλι 0.02 και 0.06 το ποσοστό σωστής ανίχνευσης είναι το ίδιο, όμως υπάρχει διαφορά στην ευαισθησία ανίχνευσης. Για παράδειγμα, για κατώφλι 0.02 το σύστημα είναι πολύ ευαίσθητο και ανιχνεύει χαμόγελο ακόμη και με πολύ μικρές κινήσεις του στόματος. Ενώ αντίθετα με κατώφλι 0.06 υπάρχει μεγαλύτερη αντίσταση στην ανίχνευση χαμόγελου, και πρέπει το χαμόγελο να είναι πολύ έντονο. Αποφασίσαμε ότι κατώφλι 0.03 μας παρέχει τα καλύτερα αποτελέσματα.

6.7.3 Αναγνώριση στρογγυλού σχήματος χειλιών

Όπως και στην περίπτωση του χαμόγελου, έτσι και εδώ χρησιμοποιήσαμε τον ίδιο αλγόριθμο με τη μόνη διαφορά ότι σε αυτή την περίπτωση ελέγχουμε την ψ απόσταση ανάμεσα στα δύο μεσαία σημεία του στόματος. Αν η ψ απόσταση των σημείων 51 και 57 όπως φαίνεται στην εικόνα 6.5 έχει αυξηθεί σε σχέση με τη σταθερή απόσταση των σημείων όταν το πρόσωπο βρίσκεται σε ηρεμία τότε έχουμε άνοιγμα των χειλιών και σχηματισμό στρογγυλού σχήματος του στόματος.



6.7.4 Αναγνώριση ανασήκωσης φρυδιών

Για την αναγνώριση φρυδιών χρησιμοποιούμε τα δύο σημεία που βρίσκονται στη μέση του αριστερού και στη μέση του δεξιού φρυδιού δηλαδή τα σημεία 22 και 16 στην φωτογραφία 6.5. Κατά την έναρξη της εφαρμογής αναγνωρίζονται οι αρχικές θέσεις των σημείων αυτών, ενόσω το πρόσωπο βρίσκεται σε ηρεμία. Σε κάθε image frame όπου υπάρχει αλλαγή στο video, γίνεται υπολογισμός των καινούργιων αποστάσεων και αν οι αποστάσεις αυτές είναι μεγαλύτερες από τις προκαθορισμένες τότε υπάρχει ανασήκωση φρυδιών. Τα φρύδια θεωρούνται ανασηκωμένα εάν οποιοδήποτε από τα δύο σημεία, δεξιού ή αριστερού φρυδιού, βρίσκεται πιο ψηλά από τις αρχικές τους θέσεις [23].



Οι μετατοπίσεις του προσώπου στον άξονα ψ, οδηγούσαν σε λάθος ανίχνευση της έκφρασης εφόσον οι θέσεις των φρυδιών μετακινούνταν μαζί με την μετατόπιση του προσώπου. Για να αποτρέψουμε τέτοιες λανθασμένες ανιχνεύσεις έχουμε προσθέσει τον έλεγχο ότι το κεντρικό σημείο του προσώπου που είναι η μύτη, δεν έχει μεγάλη απόκλιση από την αρχική του θέση. Αυτός ο έλεγχος μπορεί να βελτιωθεί και άλλο, με υπολογισμό του ύψους της απόστασης του προσώπου εφόσον η ανασήκωση φρυδιών προκαλεί αύξηση σε αυτό, ενώ η απλή μετατόπιση δεν επηρεάζει το ύψος του προσώπου.

Όπως και στην περίπτωση του χαμόγελου με το κατώφλι έχουμε δώσει ένα πολύ μικρό περιθώριο στην ανασήκωση φρυδιών έτσι ώστε να μη γίνεται λανθασμένη ανίχνευση σε πολύ μικρές κινήσεις. Στην περίπτωση όμως του χαμόγελου έχουμε μεγαλύτερο φάσμα κινήσεων εφόσον το στόμα μπορεί να μετακινηθεί κατά τη διάρκεια ομιλίας κλπ. Ενώ στην περίπτωση των φρυδιών δεν έχουμε φάσμα κινήσεων, άρα η τιμή του κατωφλίου δεν πρέπει να είναι πολύ μεγάλη.

6.7.5 Αναγνώριση συνοφρύωσης φρυδιών

Για αυτή την έκφραση χρησιμοποιείται ο ίδιος αλγόριθμος και λογική όπως στην περίπτωση ανασήκωσης φρυδιών, με τη διαφορά ότι τα φρύδια βρίσκονται πιο χαμηλά από τις αρχικές τους θέσεις.

6.8 Αποτελέσματα

Για να ελέγξουμε τα αποτελέσματα του προγράμματος έχουμε κάνει μετρήσεις για την σωστή επιτυχία ανίχνευσης εκφράσεων και συναισθημάτων. Για την εύρεση του ποσοστού επιτυχίας έχουμε τρέξει την εφαρμογή και μετρήσαμε τις πετυχημένες φορές που έγινε ορθή αναγνώριση. Τα αποτελέσματα φαίνονται στον πιο κάτω πίνακα:

	Χαρούμενος	Έκπληκτος	Θυμωμένος	Σε ηρεμία
Ποσοστό Επιτυχίας (%)	89%	92%	83%	95%

Πίνακας αποτελεσμάτων

Παρατηρούμε ότι η έκφραση “Εκπληκτος” είχε τα ψηλότερα ποσοστά ορθής επιτυχίας ενώ η έκφραση “Θυμωμένος” τα χαμηλότερα ποσοστά. Αυτό οφείλεται στην καθαρότητα έκφρασης όταν κάποιος μένει εκπληκτος, όπου το πρόσωπο κάνει πιο μεγάλες και καθαρές κινήσεις. Ενώ αντίθετα στην περίπτωση “Θυμωμένος” υπήρχε αρκετά συχνά λάθος εκτίμηση και πρόκληση σύγχυσης με την έκφραση “Χαρούμενος”. Ενώ η ανίχνευση της έκφρασης “Χαρούμενος” είχε σχετικά καλά αποτελέσματα και γινόταν με περισσότερη ακρίβεια.

Από τα πιο πάνω συμπεραίνουμε ότι όσο πιο μεγάλες και καθαρές είναι οι κινήσεις στην έκφραση του προσώπου, τόσο πιο ψηλά είναι τα ποσοστά επιτυχίας στην ανίχνευση εκφράσεων. Ενώ όσο πιο κλειστές είναι οι κινήσεις όπως στην περίπτωση “Θυμωμένος” τόσο πιο δύσκολος είναι ο διαχωρισμός και η καθαρότητα της έκφρασης.

6.9 Προβλήματα που αντιμετωπίσαμε

6.9.1 Περιορισμός του προσώπου εντός ενός πλαισίου

Κατά τη διάρκεια ανίχνευσης χαμόγελου προσέξαμε ότι ένας σημαντικός παράγοντας για λανθασμένη αναγνώριση χαμόγελου, είναι η μετακίνηση του προσώπου κοντά ή μακριά από την κάμερα. Όσο πιο κοντά μετακινείται ο χρήστης στην κάμερα τόσο πιο μεγάλη είναι η D_{mean} , ενώ όσο πιο μακριά τόσο πιο μικρή η απόσταση D_{mean} . Έτσι ο χρήστης πρέπει να παραμένει σε σταθερή απόσταση από την κάμερα έτσι ώστε να γίνεται σωστός υπολογισμός της σταθερής απόστασης μεταξύ των άκρων του στόματος, D_{mean} . Και σε κάθε αλλαγή απόστασης να γίνεται καινούργιος υπολογισμός της D_{mean} .

Για αυτό θα πρέπει ο χρήστης να παραμένει εντός ενός πλαισίου και να μη κάνει έντονες κινήσεις μπρός-πίσω στην κάμερα. Ο υπολογισμός εκφράσεων θεωρείται έγκυρος μόνο εάν ο χρήστης βρίσκεται εντός του πλαισίου.

6.9.2 Διαφορετικότητα στα πρόσωπα

Ακόμη ένα πρόβλημα που αντιμετωπίσαμε είναι η διαφορετικότητα στα πρόσωπα των ανθρώπων, για παράδειγμα κάποιοι άνθρωποι μπορούν να ανασηκώσουν τα φρύδια τους πιο ψηλά ενώ κάποιοι άλλοι να μην μπορούν, με αποτέλεσμα να μην ξεχωρίζει σαν διαφορετική έκφραση. Για το λόγο αυτό πρέπει να γίνεται μια αρχική σάρωση του προσώπου και

υπολογισμός των μέσων αποστάσεων, όπως η μέση απόσταση των δύο ακρών του στόματος, η αρχική θέση των φρυδιών ενόσω αυτά βρίσκονται σε ηρεμία.

Για να το αντιμετωπίσουμε τα πιο πάνω και για σωστό υπολογισμό των αποστάσεων, έχουμε προσθέσει ένα κουμπί στην οθόνη (κουμπί “Start”) το οποίο θα πρέπει κάθε φορά που αλλάζει αποστάσεις ο χρήστης ή αλλάζει γενικά ο χρήστης της εφαρμογής, να ξαναγίνουν οι υπολογισμοί των αποστάσεων και των αρχικών θέσεων.

6.9.3 Διαφορετικότητα στις εκφράσεις ανθρώπων

Κατά την αναγνώριση εκφράσεων, αντιμετωπίζουμε το πρόβλημα ότι οι εκφράσεις στα πρόσωπα των ανθρώπων δεν είναι ποτέ ξεκάθαρες. Για παράδειγμα, κάποιος μπορεί κατά τη διάρκεια που χαμογελά να ανασηκώνει τα φρύδια του ενώ κάποιος άλλος να κλείνει τα μάτια. Αυτό οδηγεί σε εσφαλμένη αναγνώριση εκφράσεων λόγω του ότι μπορούν ταυτόχρονα να ισχύουν διάφορες εκφράσεις.

Η παρούσα διπλωματική εργασία μπορεί να χρησιμοποιηθεί και να εφαρμοστεί μελλοντικά σε μεγαλύτερα συστήματα εφαρμογών. Το CVN project που βασίζεται στο BigBlueButton έχει κατασκευάσει μια συσκευή και έχει υλοποιήσει ένα σύστημα επικοινωνίας το οποίο βασίζεται στην αληθοφάνεια της επικοινωνίας. Επομένως, η συγκεκριμένη εφαρμογή μπορεί να είναι άμεσα αξιοποιήσιμη σε ένα τέτοιο σύστημα. Η ενσωμάτωση διαδικασιών Face Detection σε ένα σύστημα όπως το BigBlueButton ή και άλλα παρόμοια συστήματα θα είχε σημαντικά πλεονεκτήματα. Για παράδειγμα, στο BigBlueButton θα μπορούσε το σύστημα να ανιχνεύει αν ο χρήστης βρίσκεται μπροστά από την κάμερα, αν κινείται η όχι, ακόμα και συγκεκριμένες εκφράσεις του χρήστη ώστε να συμπεραίνει αν υπάρχει κάποιο πρόβλημα στον χρήστη, είτε από άποψη ασφάλειας (ο χρήστης λυποθύμισε), είτε ευχρηστίας και συναισθημάτων (αν είναι ευχαριστημένος από το σύστημα κλπ.).

Η εφαρμογή θα μπορούσε να χρησιμοποιηθεί σε οπτικούς οίκους όπου οι πελάτες θα μπορούν να δοκιμάζουν τα γυαλιά μέσω της εφαρμογής και σε on line καταστήματα.

Στην εφαρμογή 2, θα μπορούσαν μελλοντικά να γίνουν βελτιώσεις σε προβλήματα που αντιμετωπίστηκαν κατά τη διάρκεια της υλοποίησης της εφαρμογής. Θα μπορούσε να τοποθετηθεί ένας ελεγκτής ο οποίος θα βοηθούσε στη μετακίνηση του προσώπου. Έτσι, δε θα χρειαζόταν ο χρήστης να παραμένει συνεχώς εντός ενός πλαισίου και θα μπορούσε να κάνει οποιεσδήποτε κινήσεις, χωρίς να επηρεάζεται η εγκυρότητα των αποτελεσμάτων. Η βελτιστοποίηση του αλγορίθμου θα βοηθούσε ώστε να έχουμε καλύτερα και πιο έγκυρα ποσοστά ανίχνευσης εκφράσεων.

Έτσι, θα μπορούσε να χρησιμοποιηθεί σαν βοήθημα ανθρώπων με ιδιαιτερότητες, όπως σε ένα σύστημα ενίσχυσης της αναγνώρισης συναισθημάτων από άτομα με προβλήματα όρασης. Ταυτόχρονα, θα ήταν σημαντικό βοήθημα για ανθρώπους που παρουσιάζουν κάποια μορφή αναπηρίας. Ο παραδοσιακός τρόπος επικοινωνίας ανθρώπου και υπολογιστή πραγματοποιείται με τη χρήση πληκτρολογίου και ποντικιού. Εναλλακτικά μπορεί να πραγματοποιηθεί με τη χρήση μικροφώνου και με διαδικασίες λεκτικής ανάλυσης. Για ανθρώπους με σοβαρά κινητικά προβλήματα αλλά και με προβλήματα ομιλίας και ακοής, η

επικοινωνία μέσω κινήσεων του κεφαλιού ή των περιοχών του προσώπου όπως τα χείλη, τα μάτια και τα βλέφαρα είναι ιδιαίτερα επιθυμητή.

Τέλος θα μπορούσε να χρησιμοποιηθεί ως πρόσθετο εργαλείο στην ιατρική φροντίδα, για παράδειγμα σε ένα ολοκληρωμένο σύστημα νοσηλευτικής παρακολούθησης χρόνιων ασθενών, ώστε να ανιχνεύεται και να αντιμετωπίζεται αναλόγως τυχόν αρνητική συναισθηματική αλλαγή (αίσθημα πανικού ή δυσανασχέτησης).

7.1 Επίλογος

Μέσα στα πλαίσια της ολοκλήρωσης της διπλωματικής μου εργασίας κατάφερα να αποκομίσω πολύ σημαντικά οφέλη τα οποία θα μου φανούν σίγουρα χρήσιμα στην μετέπειτα μου πορεία. Έμαθα λοιπόν, τη σωστή μεθοδολογία για την εκπόνηση μιας εργασίας ξεκινώντας από το ερευνητικό επίπεδο. Στη συνέχεια, ο στόχος υλοποίησης μιας εφαρμογής που θα είχε τη δυνατότητα να ανιχνεύει και να αναγνωρίζει ανθρώπινες εκφράσεις επιτεύχθηκε, με την ελπίδα ότι θα χρησιμοποιηθεί στο μέλλον σε μεγαλύτερες εφαρμογές ώστε να βοηθήσει όσους έχουν ανάγκη.

Βιβλιογραφία

- [1] D. Blackburn, J. Bone, and P. Phillips, "Face Recognition Vendor Test 2000, Evaluation Report," Tech. Rep., February 2001.
- [2] I. Craw, D. Tock and A. Benett. Finding Face Features. In Proc. Second European Conf. Computer Vision, 1992.
- [3] P. Eckman, Emotion in the Human Face, Cambridge University Press, 1982
- [4] C. Havran, L. Hupet, J. Czyz, J. Lee, L. Vandendorpe, M. Verleysen Independent Component Analysis for face authentication, Universite catholique de Louvain, Electricity Dept., Belgium, 2002
- [5] A. Lanitis, C. J. Taylor and T. Cootes. An Automatic Face Identification System Using Flexible Appearance Models. In Image and Vision Computing, 1995.
- [6] Seung-Ik Lee, Joon-Hyun Ahn, and Sung-Bae Cho, Exploiting diversity of Neural ensembles
- [7] S. Z. Li and A. K. Jain Handbook of Face Recognition, Springer 2005
- [8] S. Z. Li, L. Zhu, Z. Q. Zhang, A. Blake, H. Zhang, and H. Shum. Statistical Learning of Multi-view Face Detection. In Proceedings Of the European Conference on Computer Vision, 2002.
- [9] R. Lienhart, A. Kuranov, and V. Pisarevsky. Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection. MRL Technical Report, Intel Labs, 2002.
- [10] X. Lu and A. Jain. Multimodal Facial Feature Extraction for Automatic 3D Face Recognition. Technical Report MSU-CSE-05-22. Michigan State University, 2005.
- [11] Xiaoguang Lu : Image Analysis for Face Recognition, Dept. of Computer Science & Engineering Michigan State University, 2003.

- [12] C. P. Papageorgiou, M. Oren, and T. Poggio. A General Framework for Object Detection. In Proceedings of IEEE International Conference on Computer Vision, 1998.
- [13] P. Phillips, P. Grother, R. Micheals, D. Blackburn, E. Tabassi, and J. Bone, "Face Recognition Vendor Test 2002, Evaluation Report," Tech. Rep., March 2003.
- [14] H. A. Rowley, S. Baluja, and T. Kanade. Neural Network-Based Face Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998.
- [15] H. Schneiderman and T. Kanade. A Statistical Approach to 3D Object Detection Applied to Faces and Cars. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2000.
- [16] K-K. Sung and T. Poggio. Example-Based Learning for View-Based Human Face Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998.
- [17] M. Tekalp, "Face and 2-D Mesh Animation in MPEG-4," Tutorial Issue On The MPEG-4 Standard, Image Communication Journal, Elsevier, 1999.
- [18] P. Viola and M. Jones. Rapid Object Detection using a Boosted Cascade of Simple Features. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001.
- [19] P. Viola and M. Jones. Robust Real-time Object Detection. In IEEE ICCV Workshop on Statistical and Computational Theories of Vision, 2001.
- [20] Ming-Hsuan Yang, Member, IEEE, David J. Kriegman, Senior Member, IEEE and Narendra Ahuja, Fellow, IEEE, Detecting Faces in images : A Survey.
- [21] Erik Hjelmas, Boon Kee Low Face detection: A survey. Computer Vision and Image Understanding
- [22] Yu-Hao Huang, Chiou-Shann Fuh, FACE DETECTION AND SMILE DETECTION

- [23] Dr. Nikolaos Bourbakis ATRC Wright State University AIIS, A Face Detection and Facial Expression Recognition Method
- [24] S. Theodoridis, K. Koutroumbas. *Pattern Recognition*, 3d Ed. Academic Press 2006.
- [25] Y. Freund and R. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting, AT&T Labs, 1995.
- [26] Y. Freund and R. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting, AT&T Labs, 1995.

Παράρτημα

```
/*
* EYEBROWS FROWNED
* Detects if the position of the eyebrows is lower than their standard position
*****/
function detectEyebrowsFrowned(points : Vector.<Point>) :Boolean{

    //trace(points[22].y, points[16].y);
    //trace("standard", _standardBrowsPositionLeft, _standardBrowsPositionRight);

    if(_standardBrowsPositionLeft <= (points[22].y-_eyebrowFrownThreshold) || _standardBrowsPositionRight <= (points[16].y-
_eyebrowFrownThreshold))
    {
        //_browsFrown = true;
        return true;

    }
    else if (_standardBrowsPositionLeft > (points[22].y-_eyebrowFrownThreshold) || _standardBrowsPositionRight > (points[16].y-
_eyebrowFrownThreshold))
    {
        //_browsFrown = false;
        return false;
    }
    return false;
}
```

```
/******  
* DETECT "O" SHAPE OF MOUTH :-o  
* if y distance between the two middle points of mouth  
* is larger than the standard distance then return true  
*****/  
function detectOShape(points : Vector.<Point>) : Boolean{  
  
    if(_standardMouthDistanceY < Math.abs(points[51].y-points[57].y)){  
  
        //trace("o detected");  
        return true;  
    }  
  
    return false;  
}
```

```

/*****
 * DETECT SMILE :)
 * If x direction distance between the tracked left
 * mouth corner and right mouth corner is larger than
 * the standard distance plus a threshold Tsmile, then we claim a smile detected
 *****/
function detectSmile(points : Vector.<Point>) : Boolean{

    // returns the points around the lower lip
    // returns an array of 10 points, 0 - 9 where 0 is the right corner of the lips and 6 the left corner of the lips
    // and the rest follow between point 0 - 6 and 6 - 9 making an outline of the lower lip
    //var points : Vector.<Point> = _brfManager.faceShape.shapePoints;
    var _currentMouthDistance : Number = 0;

    // the distance between the two corners of the lips
    _currentMouthDistance = Math.abs(points[48].x - points[54].x);
    //trace(_currentMouthDistance);

    if(_currentMouthDistance != 0){

        if(_currentMouthDistance >= (_standardMouthDistance + _smileThreshold))
        {

            //trace("smile, ", _currentMouthDistance);
            // _layout._mcSmile.visible = true;
            // _smileIsDetected = true;

            return true;
        }
        else if(_currentMouthDistance < (_standardMouthDistance + _smileThreshold) || faceHasMoved())
        {

```

```

        //_smileIsDetected = false;
        //_layout._mcSmile.visible = false;
        return false;
    }
}

return false;
}

/*****/
function detectLiftedBrows(points : Vector.<Point>) : Boolean{

    if(_standardBrowsPositionLeft > (points[22].y+_eyebrowThreshold) || _standardBrowsPositionRight > (points[16].y+_eyebrowThreshold)
    && !faceHasMoved())
    {
        //_browsLifted = true;
        return true;
        //trace("Brows Lifted");
    }
    else if (_standardBrowsPositionLeft < (points[22].y+_eyebrowThreshold) || _standardBrowsPositionRight <
(points[16].y+_eyebrowThreshold) )
    {
        //_browsLifted = false;
        return false;
    }
    return false;
}

```

```
/******
```

```
* DETECT STANDARD DISTANCES
```

```
******/
```

```
function detectStandardDistances(event : MouseEvent = null):void{
```

```
    /* lower lip distances */
```

```
    var points : Vector.<Point> = _brfManager.faceShape.shapePoints;
```

```
    if(points[48].x != 0 && points[54].x != 0)
```

```
    {
```

```
        _standardMouthDistance = Math.abs(points[48].x - points[54].x);
```

```
        _smileThreshold = _standardMouthDistance * _smileThresVal;
```

```
        //trace("_standardMouthDistance: ", _standardMouthDistance);
```

```
        //trace("_smileThreshold ", _smileThreshold);
```

```
    }
```

```
    /* left eyebrow distance */
```

```
    //points = _brfManager.faceShape.pointsLeftBrow;
```

```
    _standardBrowsPositionLeft = points[22].y;
```

```
    _standardBrowsPositionRight = points[16].y;
```

```
    _eyebrowThreshold = _standardBrowsPositionLeft * _eyebrowThresVal;
```

```
    _eyebrowFrownThreshold = _standardBrowsPositionLeft * _eyebrowFrownThresVal;
```

```
    //trace("_eyebrowThreshold: ", _eyebrowThreshold);
```

```
    /* nose distance */
```

```
    points = _brfManager.faceShape.shapePoints;
```

```
    _nosePointx = points[67].x;
```

```
    _nosePointy = points[67].y;
```

```
    /* closed mouth distance in y axis, to help us detect if mouth is open */
```

```
_standardMouthDistanceY = Math.abs(points[51].y-points[57].y);
```

```
_standardDistanceDetected = true;
```

```
//trace("Standard Mouth Distance: ", _standardMouthDistance);  
//trace("Smile Threshold: ", _smileThreshold);
```

