



Πανεπιστήμιο Πειραιώς – Τμήμα Πληροφορικής
Πρόγραμμα Μεταπτυχιακών Σπουδών
«Προηγμένα Συστήματα Πληροφορικής»

Μεταπτυχιακή Διατριβή

Τίτλος Διατριβής	Αναγνώριση ανθρώπων σε θερμογραφικές εικόνες
Όνοματεπώνυμο Φοιτητή	Νικόλαος Βιολάτος
Πατρώνυμο	Διονύσιος
Αριθμός Μητρώου	ΜΠΣΠ/10035
Επιβλέπων	Γεώργιος Τσιχριντζής, Καθηγητής

Οκτώβριος 2012

Τριμελής Εξεταστική Επιτροπή

ΣΥΝΟΨΗ

Τα βιομετρικά χαρακτηριστικά του ανθρώπου, όπως είναι το πρόσωπο, η ομιλία, τα χέρια, το σώμα, τα δακτυλικά αποτυπώματα, το χρώμα της ίριδας των ματιών κ.α., τον χαρακτηρίζουν μοναδικά. Το πιο δημοφιλές από αυτά είναι το πρόσωπο και υπάρχουν αντίστοιχες τεχνικές εντοπισμού και αναγνώρισης του, οι οποίες είναι πιο εύκολες και αποτελεσματικές από τεχνικές αναγνώρισης άλλων βιομετρικών χαρακτηριστικών. [1][2][3]

Ακόμα και ο εξοπλισμός δεν είναι ιδιαίτερα ακριβός για την υλοποίηση ενός συστήματος αναγνώρισης ανθρώπων σε κλασικές φωτογραφίες. Χρειάζεται ένας απλός υπολογιστής και μια φωτογραφική μηχανή. Για την αναγνώριση σε θερμογραφικές εικόνες απαιτείται θερμική κάμερα που όμως έχει ιδιαίτερα υψηλό κόστος, το οποίο κυμαίνεται ανάλογα τις δυνατότητες.

Η ταυτότητα του προσώπου έχει γίνει πιθανόν ένα πεδίο έρευνας που σχετίζεται με την αναγνώριση. Η αναγνώριση προσώπου διαφέρει από την ταυτότητα του προσώπου επειδή η πρώτη έχει ως «καθήκον» να προσδιορίσει την ταυτότητα ενός αντικειμένου, ενώ η δεύτερη χρειάζεται για επαληθευτεί η ταυτότητα ενός χρήστη. Στην παρούσα εργασία θα δοθεί βάρος στην αναγνώριση ανθρώπου σε εικόνες προσώπου, δηλαδή στην αναγνώριση προσώπου.

Οι εικόνες που χρησιμοποιήθηκαν δεν θα είναι οι κλασικές έγχρωμες ή ασπρόμαυρες στην κλίμακα του γκρι, αλλά θερμογραφικές. Τα πλεονεκτήματα των θερμογραφικών εικόνων είναι ότι δεν επηρεάζονται από τις συνθήκες φωτισμού του χώρου και δεν αλλοιώνονται ιδιότητες και χαρακτηριστικά αν το άτομο που φωτογραφίζεται φοράει καπέλο, γυαλιά ή έχει μεταμφιεστεί. Στις θερμογραφικές εικόνες εξετάζουμε θερμοκρασίες σε αντίθεση με τις κλασικές φωτογραφίες που αποτυπώνουν χρωματική πληροφορία με τιμές φωτεινότητας.

Οι αλγόριθμοι και οι τεχνικές που χρησιμοποιήθηκαν επιλέχθηκαν με τέτοιο τρόπο ώστε να μπορούν να εφαρμοστούν και σε κλασικές εικόνες αλλά και το προτεινόμενο σύστημα να γίνει όσο το δυνατόν επεκτάσιμο.

ABSTRACT

The biometric characteristics of humans, such as face, speech, hands, body, fingerprints, iris color of eyes, etc., characterize unique. The most popular of these is the face and corresponding detection and recognition techniques, which are easier and more effective than other recognition techniques of other biometric features. [1][2][3]

Even the equipment is less expensive to implement a system for the recognition of people in classic photos. It is needed a simple computer and a camera. To identify on thermographic images required a thermal camera, which has a very high cost and varies the possibilities.

The identity of the person has become a field of research likely associated with the recognition. Facial recognition differs from the identity of the person because the first has as a "duty" to determine the identity of an object, while the latter need to verify the identity of a user. In this work we give weight to recognize human on facial images, i.e. face recognition.

The images used will not be the classic color, black and white or grayscale ones, but thermographic. The advantages of thermographic images that are not affected by ambient light conditions and are not altered properties and attributes if the person being photographed is wearing a hat, glasses or masquerading. In thermographic images look temperatures unlike conventional photographs illustrating color information with brightness values.

The algorithms and techniques used were selected in such a way that they can be applied to conventional images and the proposed system to be as scalable as possible.

*Στον Ανδρέα,
τον φίλο και θείο που έφυγε νωρίς*

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1 ΕΙΣΑΓΩΓΗ	8
ΚΕΦΑΛΑΙΟ 2 ΔΟΜΗ ΕΡΓΑΣΙΑΣ.....	9
ΚΕΦΑΛΑΙΟ 3 ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΚΑΙ ΕΞΑΓΩΓΗ ΧΑΡΤΗ ΘΕΡΜΟΚΡΑΣΙΩΝ.....	12
ΚΕΦΑΛΑΙΟ 4 ΕΝΤΟΠΙΣΜΟΣ ΠΡΟΣΩΠΟΥ	13
4.1 Συστήματα εντοπισμού προσώπου	13
4.2 Συστήματα εντοπισμού ανθρώπινων κινήσεων πραγματικού χρόνου	13
4.3 Συστήματα εντοπισμού προσώπου πραγματικού χρόνου.....	14
4.4 Εντοπισμός δέρματος προσώπου	14
ΚΕΦΑΛΑΙΟ 5 ΕΞΑΓΩΓΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ	17
5.1 Ερευνητικές προσεγγίσεις.....	17
5.2 Πλέγμα προσώπου.....	18
5.3 Κατανομή θερμοκρασιών της περιοχής προσώπου	19
5.4 Περίμετρος περιγράμματος προσώπου.....	20
5.5 Λόγος μήκος/πλάτος.....	20
5.6 Διαφορές θερμοκρασίας στα μάγουλα και το πηγούνι	21
ΚΕΦΑΛΑΙΟ 6 ΑΝΑΓΝΩΡΙΣΗ ΠΡΟΣΩΠΟΥ	22
6.1 Αλγόριθμος PCA (Principal component analysis).....	22
6.2 Αλγόριθμος ICA (Independent Components Analysis).....	23
6.3 Τεχνητά νευρωνικά δίκτυα.....	24
6.4 Αναγνώριση προσώπου σε θερμογραφικές εικόνες.....	25
ΚΕΦΑΛΑΙΟ 7 ΒΕΛΤΙΩΣΕΙΣ – ΕΠΕΚΤΑΣΕΙΣ -ΣΥΜΠΕΡΑΣΜΑΤΑ	35
ΒΙΒΛΙΟΓΡΑΦΙΑ	37

ΠΕΡΙΕΧΟΜΕΝΑ ΕΙΚΟΝΩΝ ΚΑΙ ΠΙΝΑΚΩΝ**ΕΙΚΟΝΕΣ***

Εικόνα 2.1: Η αρχιτεκτονική του προτεινόμενου συστήματος	11
Εικόνα 3.2: Αφαίρεση παρασκηνίου με την μέθοδο Otsu	12
Εικόνα 4.3: Χωρισμός της εικόνας σε περιοχές	15
Εικόνα 4.4: Επιλογή της μεγαλύτερης περιοχής	16
Εικόνα 4.5: Επιλογή της περιοχής του προσώπου	16
Εικόνα 5.6: Χωρισμός του προσώπου σε περιοχές (πλέγμα προσώπου)	18
Εικόνα 5.7: Η περίμετρος είναι $2*(Y2-Y)+2*(X2-X)$	20
Εικόνα 5.8: Ο λόγος μήκος/πλάτος είναι $(X2-X)/(Y2-Y)$	21
Εικόνα 5.9: Σημεία και περιοχές ενδιαφέροντος στο πλέγμα προσώπου	22
Εικόνα 6.10: Βιολογικό νευρωνικό δίκτυο	24
Εικόνα 6.11: Τεχνητό νευρωνικό δίκτυο	25
Εικόνα 6.12: Πολυστρωματικό τεχνητό νευρωνικό δίκτυο	25
Εικόνα 6.13: Κατωφλίωση Otsu για την εικόνα 1	26
Εικόνα 6.14: Τμηματοποίηση της εικόνας 1 σε περιοχές (με ανοικτό μπλε η μεγαλύτερη περιοχή)	26
Εικόνα 6.15: Επιλογή της μεγαλύτερης περιοχής έπειτα από την τμηματοποίηση της εικόνας 1	27
Εικόνα 6.16: Επιλογή της περιοχής του προσώπου για την εικόνα 1	27
Εικόνα 6.17: Χωρισμός του προσώπου σε περιοχές (πλέγμα προσώπου) για την εικόνα 1	28
Εικόνα 6.18: Κατωφλίωση Otsu για την εικόνα 11 (ελέγχου)	29
Εικόνα 6.19: Τμηματοποίηση της εικόνας 11 σε περιοχές (με ανοικτό μπλε η μεγαλύτερη περιοχή)	29
Εικόνα 6.20: Επιλογή της μεγαλύτερης περιοχής έπειτα από την τμηματοποίηση της εικόνας 11	30
Εικόνα 6.21: Επιλογή της περιοχής του προσώπου για την εικόνα 11	30
Εικόνα 6.22: Χωρισμός του προσώπου σε περιοχές (πλέγμα προσώπου) για την εικόνα 11	31
Εικόνα 6.23: Κατωφλίωση Otsu για την εικόνα 10 (ελέγχου)	32
Εικόνα 6.24: Τμηματοποίηση της εικόνας 10 σε περιοχές (με ανοικτό μπλε η μεγαλύτερη περιοχή)	32
Εικόνα 6.25: Επιλογή της μεγαλύτερης περιοχής έπειτα από την τμηματοποίηση της εικόνας 10	33
Εικόνα 6.26: Επιλογή της περιοχής του προσώπου για την εικόνα 10	33

Εικόνα 6.27: Χωρισμός του προσώπου σε περιοχές (πλέγμα προσώπου) για την εικόνα 10 **34**

ΠΙΝΑΚΕΣ**

Πίνακας 5.1: Κατανομή θερμοκρασιών σε 12 ομάδες	19
Πίνακας 6.2: Οι εικόνες εκπαίδευσης	26
Πίνακας 6.3: Οι εικόνες ελέγχου για το πρώτο πείραμα	29
Πίνακας 6.4: Οι εικόνες ελέγχου για το δεύτερο πίνακα	31

* *Εικόνα X.Y: όπου X είναι το κεφάλαιο που βρίσκεται και Y ο αύξων αριθμός των εικόνων*

** *Πίνακας A,B: όπου A είναι το κεφάλαιο που βρίσκεται και B ο αύξων αριθμός των πινάκων*

ΚΕΦΑΛΑΙΟ 1 ΕΙΣΑΓΩΓΗ

Τα συστήματα ασφαλείας συνεχώς βελτιώνονται και χρησιμοποιούν όλο και περισσότερες καινοτόμες τεχνικές για να ανταποκρίνονται στις σύγχρονες απαιτήσεις. Πλέον, συστήματα για εξουσιοδοτημένη πρόσβαση με μαγνητικές κάρτες, bar codes και ηλεκτρονικές ταυτότητες, δεν ανταποκρίνονται στις απαιτήσεις αυτές. Ο τομέας της έρευνας για τα συστήματα ασφαλείας έχει στραφεί στη Βιομετρική.

Η φιλοσοφία των συστημάτων που στηρίζονται στη Βιομετρική είναι να εξάγουν φυσικά χαρακτηριστικά και προσωπικά χαρακτηριστικά από τα άτομα που ζητούν πρόσβαση. Τα φυσικά αυτά χαρακτηριστικά μπορεί να είναι δακτυλικά αποτυπώματα, χαρακτηριστικά προσώπου (χρώμα ίριδας ματιών, απόσταση ματιών, περίμετρος προσώπου κτλ), χαρακτηριστικά χεριών (μήκη δακτύλων, πάχος δακτύλων, μέγεθος παλάμης κτλ) κ.α. Προσωπικά χαρακτηριστικά είναι η υπογραφή, η αρτηριακή πίεση (ήχος φλεβών) αλλά είναι λιγότερο ασφαλή γι' αυτό πρέπει να χρησιμοποιούνται συνδυαστικά με τα φυσικά χαρακτηριστικά.

Στη συνέχεια τα χαρακτηριστικά που εξάγονται συγκρίνονται με τα αντίστοιχα χαρακτηριστικά ατόμων που τους έχει δοθεί το δικαίωμα πρόσβασης. Κατά την σύγκριση και με βάση τις αποκλίσεις που υπάρχουν κρίνεται αν θα δοθεί η πρόσβαση στα άτομα που την ζήτησαν.

Η συλλογή των παραπάνω χαρακτηριστικών γίνεται μέσα από εικόνες στατικές ή στιγμιότυπα βίντεο που τραβήχτηκαν κατά την αίτηση πρόσβασης. Οι εικόνες αυτές μπορεί να είναι έγχρωμες, ασπρόμαυρες στην κλίμακα του γκρι ή θερμογραφικές. Η παρούσα εργασία κατά το μεγαλύτερο μέρος της βασίστηκε στην εργασία των Yuan-Tsung Chen και Ming-Shi Wang για την αναγνώριση προσώπου σε θερμογραφικές εικόνες. [4] Οι θερμογραφικές εικόνες που χρησιμοποιήθηκαν προέρχονται από εργαστήριο «Imaging, Robotics, and Intelligent Systems» του Πανεπιστημίου του Tennessee. [5]

Η τεχνολογία των θερμογραφικών εικόνων και των αντίστοιχων καμερών βασίζεται στο γεγονός ότι κάθε αντικείμενο με θερμοκρασία μεγαλύτερη του απόλυτου μηδέν (-273°C) εκπέμπει ηλεκτρομαγνητική ακτινοβολία (φυσική ή θερμική). Όλη η προσπίπτουσα ακτινοβολία απορροφάται από ένα μαύρο σώμα στο συνεχές φάσμα (νόμος Planck [6]).

$$W(\lambda, T) = \frac{2\pi hc^2 / \lambda^5}{\exp(hc / \lambda kT) - 1}$$

όπου W η μονοχρωματική εκπέμπουσα δύναμη, c η ταχύτητα φωτός στο κενό ($2,9979 \times 10^8$ m/sec), h η σταθερά Planck (6.625×10^{-34} W.sec²), k η σταθερά Boltzmann (1.3805×10^{-23} WSec⁰/K), λ το μήκος κύματος (σε μέτρα), T η απόλυτη θερμοκρασία (σε °K)

Σύμφωνα με το νόμο του Planck η συνολική ακτινοβολία μέλανος σώματος είναι:

$$W(T) = \int_0^{\infty} W(\lambda, T) d\lambda = \sigma T^4$$

όπου, σ η σταθερά Stefan-Boltzmann (5.670×10^{-8} W/ m²°K⁴)

(Το μήκος κύματος της μέγιστης εκπεμπόμενης ακτινοβολίας σε σχέση με την θερμοκρασία μελανού σώματος σύμφωνα με το Νόμο του Wein [7] δίνεται από τη σχέση $\lambda_{\max} = 2898/T$). Το

ανθρώπινο σώμα εκπέμπει ακτινοβολία σε 2 ζώνες τη μέση (3-5 μm) και την υπέρυθρη (8-12 μm).

Οι δυνατότητες που παρέχει μια θερμική εικόνα είναι πολύ μεγάλες και μπορούν να βγουν πιο ασφαλή συμπεράσματα για το πρόσωπο που απεικονίζουν σε σχέση με τις έγχρωμες ή τις ασπρόμαυρες εικόνες. Οι θερμογραφικές εικόνες, όπως το λέει και το όνομά τους βασίζονται στην θερμοκρασία. Δείχνουν τα επίπεδα θερμοκρασίας που εκπέμπονται από τα αντικείμενα και τους ανθρώπους που μπορεί να υπάρχουν σε μια σκηνή. Αντίθετα, στις έγχρωμες και τις ασπρόμαυρες εικόνες η βασική πληροφορία που παρέχεται είναι τα επίπεδα φωτεινότητας, που όμως επηρεάζονται και από εξωτερικούς παράγοντες όπως είναι τα επίπεδα φωτισμού.

Η αλλαγή της ροής στα επιφανειακά αγγεία του σώματος προκαλεί αλλαγές της τοπικής θερμοκρασίας του δέρματος. Οι αλλαγές αυτές είναι πιο εμφανείς στο πρόσωπο όπου το δέρμα είναι πολύ λεπτότερο. Στις θερμογραφικές εικόνες οι πληροφορίες για την κατανομή θερμοκρασίας δίνουν την δυνατότητα να εντοπιστεί ο ανθρώπινος παράγοντας, το σώμα ή και το πρόσωπο και στη συνέχεια να αντληθούν διάφορες πληροφορίες για τον άνθρωπο που απεικονίζεται. Αξίζει να σημειωθεί ότι από μια θερμογραφική εικόνα ενός προσώπου μπορεί να γίνει μια εκτίμηση για τη συναισθηματική κατάσταση του ατόμου το οποίο απεικονίζει (αναγνώριση συναισθημάτων), τα επίπεδα άγχους κλπ.

Προφανώς, ο τομέας των συστημάτων ασφαλείας είναι ένας χώρος που φαίνεται ξεκάθαρα η χρησιμότητα των θερμογραφικών εικόνων για τον εντοπισμό και την αναγνώριση ανθρώπων. Υπάρχουν όμως και άλλοι τομείς που χρησιμοποιούνται, όπως είναι η ιατρική για τον εντοπισμό και την αναγνώριση βλαβών σε ανθρώπινους ιστούς και για την διάγνωση ασθενειών. Επίσης, στην βιομηχανία οι θερμογραφικές εικόνες είναι χρήσιμες για τον εντοπισμών σφαλμάτων σε κυκλώματα ηλεκτρονικών συστημάτων κατά την διάρκεια παραγωγής.

ΚΕΦΑΛΑΙΟ 2 ΔΟΜΗ ΕΡΓΑΣΙΑΣ

Το Κεφάλαιο 3 περιγράφει την διαδικασία προεπεξεργασίας των θερμογραφικών εικόνων που θα εισαχθούν στο προτεινόμενο σύστημα. Η προεπεξεργασία περιλαμβάνει τον διαχωρισμό του παρασκήνιου από το προσκήνιο με την μέθοδο κατωφλίωσης Otsu. Στην συνέχεια τις θερμογραφικές εικόνες, οι οποίες στην πραγματικότητα είναι έγχρωμες εικόνες με αποχρώσεις του γκρι, τις μετατρέπουμε στην κλίμακα του γκρι. Τέλος, γίνεται εξαγωγή των τιμών θερμοκρασίας (χάρτες θερμοκρασιών) για τις εικόνες που εισήχθησαν με βάση ένα γενικό τύπο.

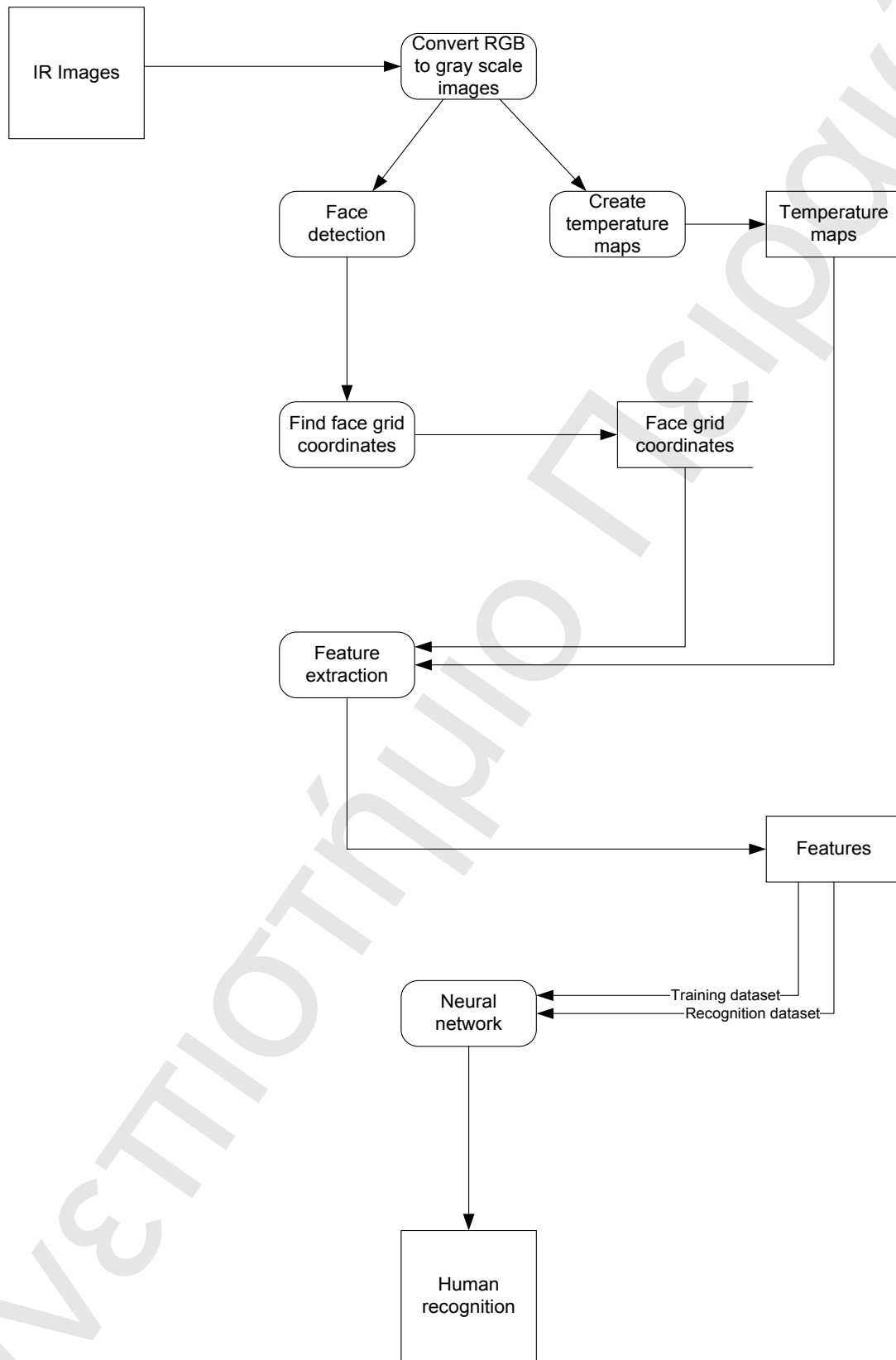
Στο Κεφάλαιο 4 γίνεται μια αναλυτική αναφορά στις μεθόδους και τους αλγόριθμους εντοπισμού προσώπου πάνω σε εικόνες. Στην τελευταία παράγραφο του κεφαλαίου γίνεται παρουσίαση του τρόπου εντοπισμού προσώπου σε θερμογραφικές εικόνες για το προτεινόμενο σύστημα. Η διαδικασία που ακολουθήθηκε βασίστηκε στον εντοπισμό δέρματος και προσαρμόστηκε για θερμογραφικές εικόνες. Στην πραγματικότητα οι περιοχές της εικόνας που έχουν κατανομές θερμοκρασίες σύμφωνες με τις κατανομές θερμοκρασίας ενός πρότυπου προσώπου, είναι αυτές που ανήκουν στο πρόσωπο.

Στο Κεφάλαιο 5 προηγείται μια βιβλιογραφική έρευνα για διάφορους τύπους χαρακτηριστικών που μπορούν να εξαχθούν από μια εικόνα προσώπου και κάποιοι μέθοδοι εξαγωγής. Έπειτα, παρουσιάζεται η διαδικασία εξαγωγής χαρακτηριστικών για κάθε είδος χαρακτηριστικών που χρησιμοποιείται για την εκμάθηση του συστήματος αναγνώρισης. Τα χαρακτηριστικά που χρησιμοποιούνται είναι 12 κατανομές θερμοκρασίας της περιοχής του προσώπου (το πλήθος των pixels που ανήκουν σε 12 ομάδες με συγκεκριμένα εύρη θερμοκρασιών), η περίμετρος του προσώπου (η περίμετρος του ελάχιστου περικλειόντος παραλληλογράμμου του προσώπου), ο λόγος μήκος/πλάτος του ελάχιστου περικλειόντος

παραλληλογράμμου και οι διαφορές T1-T2, T1-T3, T2-T3 των αντίστοιχα τριών χαμηλότερων θερμοκρασιών [T1, T2, T3] από τις περιοχές των μάγουλων και του πηγουνιού.

Στο Κεφάλαιο 6 γίνεται παρουσίαση τεχνικών και αλγορίθμων αναγνώρισης προσώπου. Για το υλοποιημένο σύστημα χρησιμοποιήθηκε ένα τεχνητό νευρωνικό δίκτυο που εκπαιδεύτηκε από τα χαρακτηριστικά που εξάχθηκαν. Το πλήθος τιμών των χαρακτηριστικών είναι 17 και συγκεκριμένα είναι οι 12 κατανομές θερμοκρασίας, η περίμετρος του προσώπου, ο λόγος μήκος/πλάτος του προσώπου, και οι τρεις τιμές διαφορών θερμοκρασίας στα 2 μάγουλα και το πηγούνι. Έτσι λοιπόν, όσο αφορά την αρχιτεκτονική του τεχνητού νευρωνικού δικτύου, οι νευρώνες εισόδου είναι 17, ένας νευρώνας εξόδου που θα μας δείξει το ποσοστό επιτυχίας αναγνώρισης του προσώπου και στο κρυφό επίπεδο το πλήθος των νευρώνων διαμορφώνεται με βάση τον τύπο (αριθμός νευρώνων εισόδου + αριθμός νευρώνων εξόδου)/2 με στρογγυλοποίηση προς τα πάνω, δηλαδή 9. Στην συνέχεια δίδονται ενδεικτικά παραδείγματα εκτέλεσης με εικόνες διαφόρων συνθηκών φωτισμού και μεταμπίεσης των προσώπων.

Σχεδόν κάθε τμήμα του προτεινόμενου συστήματος παρουσιάζεται σε ξεχωριστό κεφάλαιο και η αρχιτεκτονική του συνοψίζεται στο παρακάτω διάγραμμα.



Εικόνα 2.1: Η αρχιτεκτονική του προτεινόμενου συστήματος

Τέλος, στο Κεφάλαιο 7 αναφέρονται κάποια συμπεράσματα που βγήκαν από την παρούσα εργασία, προτεινόμενες μελλοντικές επεκτάσεις με προσθήκη περισσότερων δυνατοτήτων και βελτιώσεις στο υλοποιημένο σύστημα ώστε να έχει υψηλότερες επιδόσεις.

ΚΕΦΑΛΑΙΟ 3 ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΚΑΙ ΕΞΑΓΩΓΗ ΧΑΡΤΗ ΘΕΡΜΟΚΡΑΣΙΩΝ

Το προτεινόμενο σύστημα διαβάζει τις θερμογραφικές εικόνες που προορίζονται για την ανίχνευση κάποιου προσώπου. Το πρώτο στάδιο προεπεξεργασίας αφορά την αφαίρεση του παρασκηνίου από τις εικόνες με την μέθοδο Otsu.



Εικόνα 3.2: Αφαίρεση παρασκηνίου με την μέθοδο Otsu

Εν συνεχεία, επειδή οι θερμογραφικές εικόνες είναι στην πραγματικότητα έγχρωμες εικόνες (κλίμακα RGB) με γκρι χρώματα, πρέπει να τις μετατρέψουμε στην κλίμακα του γκρι. Για τους υπολογισμούς μας αυτό δεν αρκεί, αφού οι θερμοκρασίες των περιοχών του προσώπου είναι το βασικό στοιχείο για την εξαγωγή συμπερασμάτων. Για την μετατροπή των τιμών φωτεινότητας των ρixel στην κλίμακα του γκρι σε τιμές θερμοκρασίας, εφαρμόζουμε τον τύπο [4] (ο τύπος αυτός ενδέχεται να αλλάζει για κάποιες θερμογραφικές κάμερες):

$$\text{Temperature} = (\text{gray-level} - 128) \times \text{Sensitivity} / 8 + \text{level}$$

όπου Temperature είναι η τιμή θερμοκρασίας που αντιστοιχεί στο ρixel με τιμή φωτεινότητας gray-level-128 (αφού η εικόνα είναι στην κλίμακα του γκρι-128 στάθμες).

Η ευαισθησία Sensitivity τίθεται στο 0.3 και η σταθερά level στο 29 (οι τιμές αυτές ενδέχεται να αλλάζουν για κάποιες θερμογραφικές κάμερες).

ΚΕΦΑΛΑΙΟ 4 ΕΝΤΟΠΙΣΜΟΣ ΠΡΟΣΩΠΟΥ

Το πρώτο βήμα πριν την επεξεργασία των πληροφοριών που λαμβάνονται από ένα πρόσωπο είναι ο εντοπισμός της περιοχής του προσώπου, δηλαδή, με απλά λόγια η αναγνώριση όλων των περιοχών που περιέχουν το ανθρώπινο πρόσωπο. Το πρόβλημα της εύρεσης των προσώπων σε μια εικόνα θα πρέπει να αντιμετωπιστεί ανεξάρτητα από την πολυπλοκότητα της σκηνής της εικόνας, την γωνία λήψης, τις συνθήκες φωτισμού κ.α. .

4.1 Συστήματα εντοπισμού προσώπου

Πολλά συστήματα και τεχνικές έχουν αναπτυχθεί για την ανίχνευση προσώπου σε στατικές εικόνες, ωστόσο, τα περισσότερα από αυτά μπορούν να ανιχνεύσουν πρόσωπα μόνο σε πρόσθια προβολή ή σχεδόν πρόσθια και πολλές φορές απαιτείται να έχει γίνει κάποια προεπεξεργασία ή κανονικοποίηση. Στα περισσότερα χρησιμοποιείται ένα πολυστρωματικό νευρωνικό δίκτυο για να βρουν τις περιοχές της εικόνας που ανήκουν στο πρόσωπο με βάση τις φωτεινότητες της εικόνας και χωρικές σχέσεις των pixels.

Το 1997 αναπτύχθηκε μια μέθοδος [8] χωρίς επίβλεψη και οπτική εκμάθηση. Στόχος της τεχνικής αυτής είναι η εκτίμηση της πυκνότητας σε ένα μεγάλων διαστάσεων χώρο με αποσύνθεση ιδιοχώρου.

Το 1998 προτάθηκε μια μέθοδος [9] στην οποία με χρήση ενός νευρωνικού δικτύου στόχος ήταν να βρεθεί μια διακρίνουσα συνάρτηση που να διαχωρίζει τις περιοχές που ανήκουν στο πρόσωπο και στις περιοχές που δεν ανήκουν, χρησιμοποιώντας μέτρα αποστάσεων. Δύο είναι οι τύποι εκτίμησης πυκνότητας που παράγονται για την μοντελοποίηση των δεδομένων της εκπαίδευσης: μια πολυμεταβλητή Gaussian (για μονοτροπικές κατανομές) και ένα μίγμα Gaussian μοντέλων (για πολυτροπικές κατανομές). Αυτές οι πυκνότητες πιθανότητας χρησιμοποιούνται στη συνέχεια για να εκτιμηθεί το πλαίσιο μέγιστης πιθανότητας για οπτική αναζήτηση και τον εντοπισμό και αναγνώριση αντικειμένων και κωδικοποίηση. Η τεχνική εκμάθησης που εφαρμόζεται αφορά πιθανολογική οπτική μοντελοποίηση, εντοπισμό και αναγνώριση ανθρώπινων προσώπων και κινούμενων περιοχών, όπως τα χέρια.

4.2 Συστήματα εντοπισμού ανθρώπινων κινήσεων πραγματικού χρόνου

Έχουν δημιουργηθεί αρκετά συστήματα που εντοπίζουν ανθρώπινες κινήσεις σε πραγματικό χρόνο και συνεχώς βελτιώνονται. Χαρακτηριστικό παράδειγμα είναι ένα σύστημα που αναπτύχθηκε το 1994. [10] Πρόκειται για ένα framework πραγματικού χρόνου για την παρακολούθηση των εκφράσεων του προσώπου χρησιμοποιώντας μεθόδους παρεμβολής και νευρωνικά δίκτυα.

Μια άλλη μέθοδος που μπορεί να χειριστεί κινήσεις του κεφαλιού είναι μια στατιστική μέθοδος για την ανίχνευση αντικειμένων 3D που προτάθηκε το 2000. [11] Τα στατιστικά στοιχεία που συλλέγονται αφορούν την εμφάνιση τόσο των αντικειμένων όσο και των και "μη αντικειμένων" χρησιμοποιώντας τα αντίστοιχα ιστογράμματα. Κάθε ιστόγραμμα αντιπροσωπεύει τις κοινές στατιστικές ενός υποσυνόλου των συντελεστών κυματιδίων και τη θέση τους πάνω στο αντικείμενο. Η προσέγγιση που ακολουθείται είναι να χρησιμοποιούνται πολλά τέτοια ιστογράμματα αναπαριστώντας μια μεγάλη γκάμα οπτικών ιδιοτήτων. Η μέθοδος αυτή χρησιμοποιεί έναν αλγόριθμο που εντοπίζει αξιόπιστα τα ανθρώπινα πρόσωπα και είναι ο πρώτος που μπορεί αξιόπιστα να ανιχνεύσει τα επιβατικά αυτοκίνητα από ένα ευρύ φάσμα γωνιών και στάσεων λήψης φωτογραφίας.

Άλλες τέτοιες μέθοδοι, έχουν προταθεί πρόσφατα και οι περισσότερες από αυτές τις μεθόδους δίνουν βαρύτητα σε στατιστικές τεχνικές μάθησης και χρησιμοποιούν χαρακτηριστικά εμφάνισης.

4.3 Συστήματα εντοπισμού προσώπου πραγματικού χρόνου

Αναμφισβήτητα, ένα θέμα που απασχολεί την επιστημονική κοινότητα πάρα πολύ είναι η αυτόματη ανίχνευση προσώπου σε πραγματικό χρόνο και μια μέθοδος προτάθηκε για πρώτη φορά από τους Viola και Jones το 2004. [12]

Στην εργασία αυτή περιγράφεται μια προσέγγιση μηχανής μάθησης για την οπτική ανίχνευση αντικειμένων που μπορεί να επεξεργαστεί εικόνες πολύ γρήγορα και επιτυγχάνει υψηλά ποσοστά ανίχνευσης. Αυτό έργο χαρακτηρίζεται από τρία καινοτόμα στοιχεία. Το πρώτο είναι η χρησιμοποίηση μιας αναπαράστασης της εικόνας που ονομάζεται "Integral" [13] και προτάθηκε το 1984. Ο τύπος αυτός εικόνας επιτρέπει τον πολύ γρήγορο υπολογισμό των ιδιοτήτων που χρησιμοποιούνται από τον ανιχνευτή. Το δεύτερο είναι ότι ο ανιχνευτής αποτελείται από ένα σύνολο ταξινομητών που εκπαιδεύονται κάνοντας χρήση του αλγόριθμου AdaBoost [14] και ο οποίος επιλέγει έναν μικρό αριθμό κρίσιμων οπτικών χαρακτηριστικών από ένα μεγαλύτερο σύνολο και είναι εξαιρετικής απόδοσης και αποτελεσματικότητας [15].

Κάθε ταξινομητής περιλαμβάνει εσωτερικά τα λεγόμενα «box filters» που θυμίζουν τις λειτουργίες των μητρών του Haar [16] και μπορεί να κάνει υπολογισμούς πολύ γρήγορα σε οποιαδήποτε θέση και κλίμακα. Αυτό είναι ουσιαστικής σημασίας για την ταχύτητα του ανιχνευτή.

Το τρίτο σημαντικό στοιχείο της εργασίας των Viola και Jones είναι ότι χρησιμοποιείται μια μέθοδος τύπου «καταρράκτη» για τον συνδυασμό ολοένα και πιο περίπλοκων ταξινομητών. Η μέθοδος αυτή επιτρέπει την γρήγορη απόρριψη περιοχών του παρασκηνίου, ενώ ξοδεύει περισσότερη υπολογιστική ισχύ σε περιοχές που μοιάζουν να περιέχουν αντικείμενα. Ο καταρράκτης μπορεί να θεωρηθεί ως ένας μηχανισμός που εστιάζει περισσότερο σε συγκεκριμένα αντικείμενα και ο οποίος σε αντίθεση με προηγούμενες προσεγγίσεις, παρέχει στατιστικές εγγυήσεις ότι οι περιοχές που απορρίπτονται είναι απίθανο να περιέχουν αντικείμενα ενδιαφέροντος. Στον τομέα της ανίχνευσης προσώπου, το σύστημα αποδίδει με υψηλά ποσοστά ανίχνευσης συγκριτικά με καλύτερα παλαιότερα συστήματα. Ο ανιχνευτής λειτουργεί με 15 καρτέ ανά δευτερόλεπτο, χωρίς να καταφεύγει στη μελέτη των διαφορών μεταξύ των εικόνων ή την ανίχνευση χρώματος του δέρματος και χρησιμοποιείται σε εφαρμογές πραγματικού χρόνου.

Μια εξίσου σημαντική παραλλαγή του συστήματος των Viola-Jones υλοποιήθηκε ένα χρόνο μετά το 2005 [17] και χρησιμοποιεί τον αλγόριθμο GentleBoost [18] αντί για τον AdaBoost. Το σύστημα αυτό βασίζεται σε ένα πιθανολογικό μοντέλο εικόνας και χρησιμοποιήθηκαν αλγόριθμοι αποφάσεων και συμπερασμού για την εύρεση αντικειμένων και χαρακτηριστικών τους. Το μοντέλο αυτό αντιλαμβάνεται την εικόνα ως ένα σύνολο κομματιών ακαθόριστου μεγέθους, όπου το καθένα μπορεί να περιέχει κάποιο αντικείμενο ή να είναι μέρος του παρασκηνίου. Το μοντέλο με τη σειρά του υιοθετεί άλλα πιθανοκρατικά μοντέλα με δείκτες και για να αποφασιστεί αν ένα κομμάτι είναι αντικείμενο ή ανήκει στο παρασκήνιο προηγείται κάποια εκπαίδευση. Η εκπαίδευση βασίζεται σε αλγόριθμους ενίσχυσης (Boosting algorithms [19])

Το σύστημα έχει ως κύρια λειτουργία την εύρεση προσώπων και ματιών σε πραγματικό χρόνο και παραμένει αναλλοίωτο σε αλλαγές στο φωτισμό, στη δομή των προσώπων που εισάγονται, στις εκφράσεις τους ή αν φορούν γυαλιά. Τέλος, το σύστημα αυτό μπορεί ταυτόχρονα να παρακολουθεί πολλά άτομα και να εντοπίζει τα μάτια τους.

4.4 Εντοπισμός δέρματος προσώπου

Στην παρούσα εργασία ο εντοπισμός προσώπου βασίζεται στην μέθοδο του εντοπισμού δέρματος (skin detection [20]). Το βασικό στοιχείο για τον εντοπισμό είναι η χρωματική πληροφορία κάτι που είναι αρκετά σύνηθες στις κλασικές φωτογραφίες. Στις θερμογραφικές εικόνες που δεν έχουμε χρώμα δέρματος αλλά θερμοκρασία σώματος κάνουμε την εξής εβώρηση. Οποιαδήποτε περιοχή της εικόνας έχει θερμοκρασίες άνω των 30.5 βαθμών της

κλίμακας Κελσίου ενδέχεται να είναι ανθρώπινο σώμα. Ειδικά στην περίπτωση που έχουμε εικόνες προσώπου είναι ακόμα πιο εύκολος ο εντοπισμός αφού και να υπάρχει κάποιο αντικείμενο ή κάτι το οποίο εκπέμπει θερμότητα θα είναι κατά πολύ μικρότερο. Έτσι λοιπόν σε περίπτωση που βρεθούν και άλλες μικρότερες περιοχές της εικόνας με θερμοκρασίες άνω του 30.5, απλά αγνοούνται.

Όσο αφορά τώρα τον ακριβή εντοπισμό της περιοχής που αντιστοιχεί στο πρόσωπο, προτού γίνει η επιλογή της μεγαλύτερης περιοχής πραγματοποιείται μια σειρά από μορφολογικές πράξεις (διάβρωση, κλείσιμο, διαστολή, γέμισμα) και χωρίζεται η εικόνα σε περιοχές. Η μεγαλύτερη περιοχή που αντιστοιχεί στο πρόσωπο είναι με χρώμα ανοιχτό μπλε.



Εικόνα 4.3: Χωρισμός της εικόνας σε περιοχές

Οι πράξεις αυτές γίνονται ώστε να επιλεγεί μια περιοχή ικανοποιητικά μεγαλύτερη από αυτή που αρχικά εντοπίστηκε. Ο λόγος είναι να καλυφθούν τυχόν κενά ή λόγω σφαλμάτων κάποια σημεία που δεν επιλέχθηκαν και έτσι να μην μείνει κανένα τμήμα του προσώπου εκτός επιλογής. Ένα απλό παράδειγμα είναι τα μάτια και τα φρύδια που έχουν γενικά τις μικρότερες θερμοκρασίες σε σχέση με το υπόλοιπο πρόσωπο και πολλές φορές αρκετά πιο κάτω από 30.5° C. Έτσι στην αρχική επιλογή προσώπου τα τμήματα αυτά δεν υπάρχουν, γι' αυτό γίνονται οι μορφολογικές πράξεις.

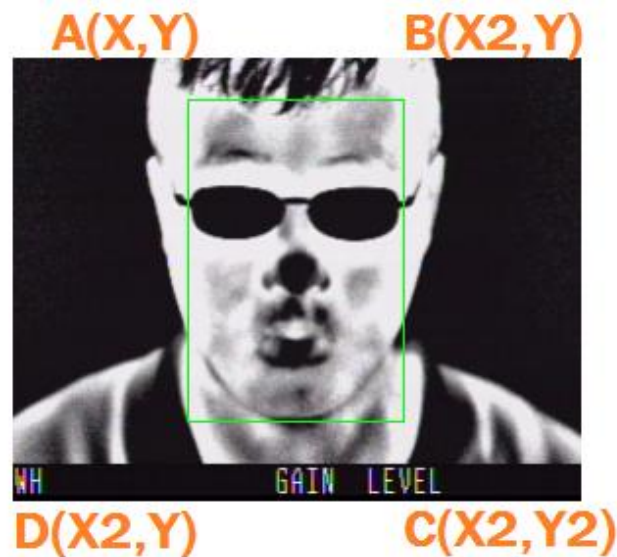
Στη συνέχεια ευρίσκεται το ελάχιστον περικλείον παραλληλόγραμμο για την περιοχή που επιλέχθηκε προηγούμενα. Σε περίπτωση που οι περιοχές είναι περισσότερες από μία, επιλέγεται η μεγαλύτερη και αυτό γίνεται σε αυτό το βήμα αφού με το ελάχιστον περικλείον παραλληλόγραμμο είναι υπολογιστικά εφικτό να βρεθεί η μεγαλύτερη περιοχή. Το σύστημα απλά κρατάει τις τιμές των $X, Y, X2, Y2$.



Εικόνα 4.4: Επιλογή της μεγαλύτερης περιοχής

Το ελάχιστο περικλείον παραλληλόγραμμο περιλαμβάνει μια περιοχή γύρω από το κεφάλι έτσι είμαστε σίγουροι ότι έχει επιλεγεί ολόκληρη η χρήσιμη πληροφορία και μαζί βεβαίως υπάρχει και άχρηστη πληροφορία. Οπότε πρέπει να ακολουθήσει μία διαδικασία για να επιλεγεί μόνο η περιοχή του προσώπου και να απαλλαχτούμε από την άχρηστη πληροφορία.

Η μέθοδος για τον ακριβή εντοπισμό του προσώπου χωρίς περιττή πληροφορία χρησιμοποιεί την θέση των 4 κορυφών του ελάχιστου περικλείοντος παραλληλογράμμου, δηλαδή τις συντεταγμένες των 4 εικονοστοιχείων μέσα στην εικόνα. Έτσι λοιπόν, μετακινούμε και τα 4 σημεία κατά ένα ρixel προς το κέντρο του παραλληλογράμμου ώστε να συρρικνωθεί η περιοχή. Στην συνέχεια ελέγχουμε τα 3 από τα 4 σημεία αν είναι σημεία του προσώπου, εξετάζοντας την αντίστοιχη θερμοκρασία από τον χάρτη θερμοκρασιών που παράχθηκε σε προηγούμενο βήμα. Αν δεν ανήκουν στο πρόσωπο, τότε μετακινούμε ξανά τα 4 σημεία κατά ένα ρixel μέχρι να ικανοποιηθεί η συνθήκη. Μετά το πέρας της διαδικασίας το σύστημα κρατά τις νέες τιμές για τα $X, Y, X2, Y2$.



Εικόνα 4.5: Επιλογή της περιοχής του προσώπου

ΚΕΦΑΛΑΙΟ 5 ΕΞΑΓΩΓΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ

5.1 Ερευνητικές προσεγγίσεις

Μετά τον εντοπισμό προσώπου, το επόμενο βήμα είναι η εξαγωγή χαρακτηριστικών δηλαδή πληροφοριών που σχετίζονται με το πρόσωπο και είναι τριών ειδών σύμφωνα με [21]:

- (α) τα χρονικά αναλλοίωτα χαρακτηριστικά (στατικά)
- (β) τα ολιστικά χαρακτηριστικά (αφορούν σε ολόκληρο το πρόσωπο)
- (γ) τα αναλυτικά χαρακτηριστικά (αφορούν τμήματα του προσώπου).
- (δ) τα χαρακτηριστικά που απλά είναι ορατά ή βασίζονται στον όγκο και την ποσότητα της πληροφορίας που καταλαμβάνουν (2D ή 3D).

Οι πιο διαδεδομένες προσεγγίσεις αναγνώρισης προσώπου υιοθετούν ολιστικά και στατικά στο χρόνο (2D) χαρακτηριστικά. Αντίθετα, πολλές άλλες προσεγγίσεις για την αυτόματη ανάλυση της έκφρασης του προσώπου υιοθετούν (2D) αναλυτικά και χωροχρονικά χαρακτηριστικά προσώπου. [22] Η διαπίστωση αυτή είναι επίσης σύμφωνη με ευρήματα από την ψυχολογική έρευνα που υποδηλώνει ότι οι διεργασίες του εγκεφάλου λειτουργούν ολιστικά και όχι σε τοπικό επίπεδο, ενώ επεξεργάζονται τις εκφράσεις του προσώπου σε τοπικό επίπεδο [23].

Οι περισσότεροι ερευνητές στον τομέα ανάλυσης εκφράσεων προσώπου προσανατολίζονται κυρίως στην εξαγωγή 2D χωροχρονικών χαρακτηριστικών προσώπου. Τα χαρακτηριστικά του προσώπου που συνήθως εξάγονται είναι είτε γεωμετρικά χαρακτηριστικά, όπως τα σχήματα των τμημάτων του προσώπου (μάτια, στόμα, κ.λπ.), είτε θέσεις βασικών σημείων του προσώπου (γωνίες των ματιών, θέση στόματος, απόσταση ματιών κλπ.) ή χαρακτηριστικά που αντιπροσωπεύουν την υφή του δέρματος του προσώπου, συμπεριλαμβάνοντας ρυτίδες, εξογκώματα, αυλάκια. Τυπικά παραδείγματα των γεωμετρικών χαρακτηριστικών που βασίζονται οι μέθοδοι αυτοί υπάρχουν στο [24], όπου χρησιμοποιήθηκε ένα πλέγμα 19 σημείων προσώπου που βασίστηκε στην εργασία των Edwin Chang και Odest Chadwicke [25], οι οποίοι χρησιμοποίησαν ένα μοντέλο σχήματος που ορίζεται από 58 σημεία προσώπου.

Τυπικά παραδείγματα των υβριδικών, γεωμετρικών και μεθόδων βασισμένων σε χαρακτηριστικά εμφάνισης, είναι αυτά που ο Ying-li Tian [26] χρησιμοποίησε: μοντέλα βασισμένα στο σχήμα των ματιών, των φρυδιών, του στόματος και μοντέλα βασισμένα σε περιοδικά χαρακτηριστικά, όπως οι ρυτίδες. Οι Yongmian Zhang και Qiang Ji [27] σε άλλη εργασία, χρησιμοποίησαν 26 σημεία του προσώπου γύρω από τα μάτια, τα φρύδια, και το στόμα και τα ίδια περιοδικά χαρακτηριστικά όπως στην εργασία [26]. Συνεπώς, με βάση τις εργασίες που έχουν πραγματοποιηθεί για την εξαγωγή χαρακτηριστικών, η συλλογή τόσο γεωμετρικών όσο και χαρακτηριστικών εμφάνισης είναι απαραίτητη για αποτελεσματικότερη αναγνώριση προσώπων [28].

Έχει αναφερθεί ότι οι μέθοδοι που βασίζονται σε γεωμετρικά χαρακτηριστικά έχουν συχνά καλύτερες επιδόσεις από εκείνες που βασίζονται σε χαρακτηριστικά εμφάνισης χρησιμοποιώντας, π.χ., κυματίδια Gabor [29] ή eigenfaces [30]. Βέβαια, αυτό μπορεί να εξαρτάται από τη μέθοδο ταξινόμησης και από μια προσέγγιση μηχανής μάθησης που θα λαμβάνει τα χαρακτηριστικά ως είσοδο. Πρόσφατες μελέτες, όπως εκείνη των Pantic & Patras [31], καθώς και εκείνες που παρουσιάζονται στο παρόν κεφάλαιο, δείχνουν ότι σε ορισμένες περιπτώσεις μπορεί τα γεωμετρικά χαρακτηριστικά να υπερτερούν από τα χαρακτηριστικά που βασίζονται στην εμφάνιση. Ωστόσο, φαίνεται ότι η χρήση τόσο γεωμετρικών όσο και χαρακτηριστικών βασισμένων στην εμφάνιση, μπορεί να είναι η καλύτερη επιλογή στην περίπτωση ορισμένων εκφράσεων του προσώπου.

Αρκετές προσεγγίσεις για την αυτόματη ανάλυση της έκφρασης του προσώπου, βασισμένες σε 3D μοντέλα προσώπου, έχουν πρόσφατα προταθεί. Ο Gokturk [32] πρότεινε μια μέθοδο για την αναγνώριση κινήσεων προσώπου, (π.χ. το βλέφαρο παίζει, χαμόγελο το άτομο κ.α.) και βασίζεται σε 3D παραμορφώσεις του προσώπου. Η μέθοδος αυτή εφαρμόζεται σε ρεύματα (streams) στερεοφωνικής εικόνας χρησιμοποιώντας ένα 19-σημείων πρότυπο των

ματιών του πρόσωπου και κοινές οπτικές τεχνικές ροής. Η εργασία του Cohen [33] επικεντρώνεται στο σχεδιασμό δικτύου ταξινομητών Bayes για την αναγνώριση συναισθημάτων από βίντεο με βάση τα χαρακτηριστικά του προσώπου και παρακολουθείται από μια μέθοδο που ονομάζεται Piecewise Bezier Volume Deformation tracking [34]. Αυτός ο tracker χρησιμοποιεί ένα 3D wireframe μοντέλο που αποτελείται από 16 επιφανειακά μπαλώματα, ενσωματωμένα σε Bezier όγκους [35]. Ο Cohen [36] έδωσε έμφαση στην αυτόματη ανάλυση των κινήσεων των φρυδιών και της κεφαλής από βίντεο και βασίστηκε σε ένα κυλινδρικό μοντέλο κεφαλιού για την εκτίμηση των 6 βαθμών ελευθερίας κίνησης της κεφαλής [37].

Το 3D modeling του πρόσωπου είναι εξαιρετικά επίκαιρο στους σημερινούς ερευνητές, λόγω της δυνατότητάς του να παράγει view independent συστήματα αναγνώρισης σήματος προσώπου. Οι κυριότερες αδυναμίες των σημερινών μεθόδων αφορούν την ανάγκη ενός μεγάλου ποσού χειροκίνητης εκπαίδευσης επισήμασμένων δεδομένων και σχεδόν πάντα απαιτείται χειροκίνητη επιλογή των σημείων του προσώπου στο πρώτο καρέ του βίντεο εισόδου, βάσει της οποίας το μοντέλο προσώπου θα ταιριάζει στο πρόσωπο.

5.2 Πλέγμα προσώπου

Όπως αναφέρθηκε στην προηγούμενη παράγραφο σε κάποιες εργασίες (πλέγμα 19 σημείων προσώπου, του Chang [25]) έχει χρησιμοποιηθεί πλέγμα προσώπου για την υποβοήθηση στην διαδικασία εξαγωγής χαρακτηριστικών από εικόνες προσώπων.

Στην παρούσα εργασία, αφού ολοκληρωθεί ο εντοπισμός προσώπου και η εύρεση της ελάχιστης περιοχής που το περικλείει, γίνεται διαίρεση της περιοχής αυτής σε 12 υποπεριοχές (3X4 μπλοκ για την ακρίβεια). Χωρίζοντας το πρόσωπο σε αυτά τα τμήματα, έχουμε την δυνατότητα να εξετάσουμε κάποια τοπικά χαρακτηριστικά που παρουσιάζουν ιδιαίτερο ενδιαφέρον και παρουσιάζονται στη συνέχεια. Ο υπολογισμός των συντεταγμένων των εσωτερικών παραλληλογράμμων (μπλοκ) γίνεται με βάση τις συντεταγμένες του ελάχιστου περικλείοντος παραλληλογράμμου του προσώπου.



Εικόνα 5.6: Χωρισμός του προσώπου σε περιοχές (πλέγμα προσώπου)

5.3 Κατανομή θερμοκρασιών της περιοχής προσώπου

Σε προηγούμενο βήμα έγινε η εξαγωγή του χάρτη θερμοκρασιών για κάθε θερμογραφική εικόνα που εισήχθη στο σύστημα. Αφού προηγήθηκε εντοπισμός της περιοχής του προσώπου, θα ευρεθεί η κατανομή των θερμοκρασιών της περιοχής του προσώπου. Η κατανομή θα γίνει σε 12 ομάδες θερμοκρασιών. Στο αρχικό πείραμα η κατανομή έγινε σε 9 ομάδες, οι 8 από 30.5°C έως 33.5 °C με βήμα 0.375 και η 9^η για θερμοκρασίες άνω των 33.5°C.

Αναλυτικά οι ομάδες ήταν:

- 1) 30.500°C - 30.874°C, 2) 30.875°C - 31.249°C, 3) 31.250°C - 31.624°C,
- 4) 31.625°C - 31.999°C, 5) 32.000°C - 32.374°C, 6) 32.375°C - 32.749°C,
- 7) 32.750°C - 33.124°C, 8) 33.125°C - 33.499°C, 9) 33.5°C και άνω.

Τελικά επιλέχθηκε να γίνει η κατανομή σε 12 ομάδες θερμοκρασίες, αφού τα αποτελέσματα της αναγνώρισης ήταν καλύτερα. Οι 11 ομάδες περιλαμβάνουν θερμοκρασίες από 30.5 έως 33.5 με βήμα 0.2792 και η 12^η για θερμοκρασίες άνω των 33.5.

Επομένως, οι 12 ομάδες θερμοκρασίας είναι οι εξής:

- 1) 30.50000°C - 30.77272°C, 2) 30.77273°C - 31.04542°C,
- 3) 31.04543°C - 31.31812°C, 4) 31.31813°C - 31.59082°C,
- 5) 31.59083°C - 31.86352°C, 6) 31.86353°C - 32.13622°C,
- 7) 32.13623°C - 32.40892°C, 8) 32.40893°C - 32.68162°C,
- 9) 32.68163°C - 32.95432°C, 10) 32.95433°C - 33.22702°C,
- 11) 33.22703°C - 33.49999°C, 12) 33.5°C και άνω.

Στη συνέχεια υπολογίζεται ο αριθμός των pixels που έχουν αντίστοιχη θερμοκρασία μέσα στα όρια της κάθε ομάδας. Σε ένα τυπικό παράδειγμα εκτέλεσης του αλγορίθμου για την εξαγωγή χαρακτηριστικών για την εικόνα 2 λάβαμε τις παρακάτω τιμές:

Ομάδα θερμοκρασίας	Πλήθος pixels
Group1	0
Group2	307
Group3	0
Group4	0
Group5	0
Group6	428
Group7	0
Group8	0

Group9	0
Group10	673
Group11	0
Group12	14198

Πίνακας 5.1: Κατανομή θερμοκρασιών σε 12 ομάδες

5.4 Περίμετρος περιγράμματος προσώπου

Το περίγραμμα του προσώπου εντοπίστηκε σε προηγούμενο βήμα με την εύρεση του ελάχιστου περικλείοντος παραλληλογράμμου του προσώπου. Ένα σημαντικό χαρακτηριστικό για την αναγνώριση ενός προσώπου είναι και η περίμετρος του. Η περίμετρος είναι μια παράμετρος που εξαρτάται από την οπτική γωνία και την απόσταση μεταξύ κάμερας και αντικειμένου. Για την αποφυγή αποκλίσεων θα πρέπει οι εικόνες που χρησιμοποιούνται να είναι κανονικοποιημένες (κατά κάποιο τρόπο) και ίδιου μεγέθους.



Εικόνα 5.7: Η περίμετρος είναι $2*(Y2-Y)+2*(X2-X)$

5.5 Λόγος μήκος/πλάτος

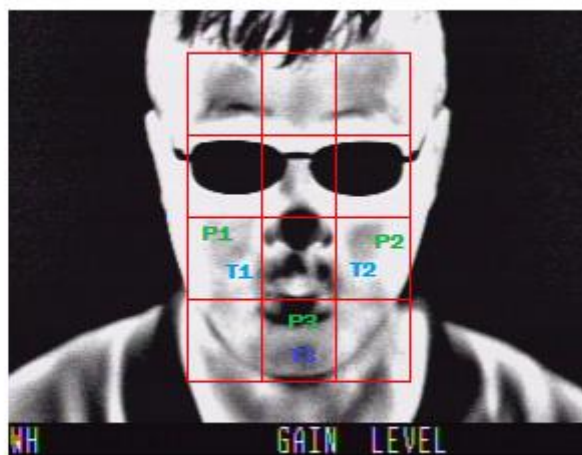
Ο λόγος μήκος/πλάτος του ελάχιστου περικλείοντος παραλληλογράμμου του προσώπου είναι ένα χαρακτηριστικό που περιγράφει το σχήμα του προσώπου.



Εικόνα 5.8: Ο λόγος μήκος/πλάτος είναι $(X2-X)/(Y2-Y)$

5.6 Διαφορές θερμοκρασίας στα μάγουλα και το πηγούνι

Ιδιαίτερου ενδιαφέροντος και πολύ σημαντικές είναι οι πληροφορίες που αντλούμε από τις 3 υπο-περιοχές των δύο μάγουλων (P1: περιοχή αριστερού μάγουλου, P2: περιοχή δεξιού μάγουλου) και του πηγουνιού (περιοχή P3) από το πλέγμα του προσώπου. Για κάθε μια από αυτές τις περιοχές θα βρούμε τις 3 μικρότερες θερμοκρασίες (T1, T2, T3 αντίστοιχα για τις περιοχές P1, P2, P3) από τον χάρτη θερμοκρασιών. Οπότε, στα χαρακτηριστικά που θα εξάγουμε για την αναγνώριση θα είναι ένας πίνακας με τις διαφορές θερμοκρασιών T1-T2, T1-T3 και T2-T3.



Εικόνα 5.9: Σημεία και περιοχές ενδιαφέροντος στο πλέγμα προσώπου

ΚΕΦΑΛΑΙΟ 6 ΑΝΑΓΝΩΡΙΣΗ ΠΡΟΣΩΠΟΥ

Για την αναγνώριση ανθρώπων και ειδικότερα προσώπων σε εικόνες ή στιγμιότυπα βίντεο, έχουν αναπτυχθεί διάφορες τεχνικές τα τελευταία 30 χρόνια, εξαιτίας και των απαιτήσεων που υπάρχουν από αντίστοιχες εφαρμογές.

6.1 Αλγόριθμος PCA (Principal component analysis)

Ένας πολύ ευέλικτος αλγόριθμος που προέρχεται από τον χώρο της Στατιστικής είναι ο PCA και προσαρμόστηκε για την χρήση του σε εικόνες. Οι Kirby και Sirovich χρησιμοποίησαν πρώτοι τον αλγόριθμο PCA για την αναγνώριση προσώπων και απέδειξαν ότι ο PCA είναι ένα βέλτιστος αλγόριθμος συμπίεσης που ελαχιστοποιεί το μέσο τετραγωνικού σφάλμα μεταξύ των αρχικών εικόνων και των ανακατασκευών τους για κάθε δεδομένο επίπεδο συμπίεσης [38]. Οι Turk & Pentland [39] διέδωσαν τη χρήση του PCA για την αναγνώριση προσώπου.

Ο αλγόριθμος PCA βασίζεται στην ιδέα ότι η αναγνώριση ενός προσώπου μπορεί να επιτευχθεί από μια μικρή ομάδα χαρακτηριστικών που αντιπροσωπεύει συνολικά τις ήδη αναγνωρισμένες εικόνες γι' αυτό το πρόσωπο. Η εφαρμογή του PCA ξεκινά με μια σειρά από εικόνες γνωστών ανθρώπων που χρησιμοποιούνται ως εικόνες εκπαίδευσης. Οι υπόλοιπες εικόνες στις οποίες θέλουμε να γίνει αναγνώριση είναι οι εικόνες ελέγχου. Από την ανάλυση των εικόνων εκπαίδευσης, εξάγεται ένα σύνολο κύριων συστατικών. Στην συνέχεια οι προβολές των εικόνων ελέγχου πάνω στα συστατικά αυτά χρησιμοποιούνται για τον υπολογισμό των αποκλίσεων μεταξύ των εικόνων ελέγχου και των εικόνων εκπαίδευσης. Έπειτα, οι αποστάσεις αυτές χρησιμοποιούνται για να γίνουν προβλέψεις σχετικά με τα πρόσωπα που είναι στις εικόνες ελέγχου (αν αναγνωρίζονται ή όχι).

Θεωρούμε μια μήτρα $D \times K$ διαστάσεων που περιλαμβάνει τα συνολικά δεδομένα για ένα πρόσωπο, όπου το D αντιπροσωπεύει τον αριθμό των pixels της κάθε εικόνας (οι εικόνες είναι ίδιου μεγέθους) και το K το πλήθος των εικόνων αυτών. Επίσης, έστω XX^T είναι ο πίνακας συνδιακύμανσης για τις εικόνες εκπαίδευσης, ενώ οι κύριες συνιστώσες του πίνακα συνδιακύμανσης υπολογίζονται από την επίλυση της παρακάτω εξίσωσης:

$$R^T (XX^T) R = \Lambda$$

όπου Λ είναι ο διαγώνιος πίνακας των ιδιοτιμών και R είναι η μήτρα των ορθοκανονικών ιδιοδιανυσμάτων. Γεωμετρικά, το R είναι ένας πίνακας περιστροφής που περιστρέφεται το αρχικό σύστημα συντεταγμένων πάνω στα ιδιοδιανύσματα. και όπου το ιδιοδιάνυσμα συνδέεται με τη μεγαλύτερη ιδιοτιμή είναι ο άξονας της μέγιστης διακύμανσης. Το ιδιοδιάνυσμα που συνδέεται με τη δεύτερη μεγαλύτερη ιδιοτιμή είναι ο ορθογώνιος άξονας με τη δεύτερη ανώτατη τιμή διακύμανσης, κ.ο.κ. Συνήθως, μόνο τα ιδιοδιανύσματα M που συνδέονται με τις M μεγαλύτερες ιδιοτιμές χρησιμοποιούνται για να καθορίσουν τον υπόχωρο, όπου το M είναι το επιθυμητό μέγεθος διάστασης του υποχώρου.

6.2 Αλγόριθμος ICA (Independent Components Analysis)

Ο αλγόριθμος ICA είναι επίσης, μια στατιστική τεχνική για την επεξεργασία σήματος και είναι πολύ στενά συνδεδεμένη με τη μέθοδο που ονομάζεται Blind Source Separation (BSS) ή Blind Signal Separation [40]. Η βασική ιδέα του αλγορίθμου ICA είναι ότι χρησιμοποιεί συναρτήσεις βάσης για να αντιπροσωπεύσει ένα σύνολο τυχαίων μεταβλητών, όπου οι συνιστώσες είναι στατιστικώς ανεξάρτητες ή ανεξάρτητες όσο είναι δυνατόν.

Έστω s το διάνυσμα άγνωστων σημάτων εισόδου και x είναι το διάνυσμα των παρατηρούμενων μειγμάτων. Αν A είναι ο άγνωστος πίνακας μειγμάτων, τότε το μοντέλο ανάμειξης γράφεται ως εξής: $x = As$. Υποτίθεται ότι τα σήματα της πηγής είναι ανεξάρτητα από κάθε άλλα και η μήτρα ανάμειξης A είναι αντιστρέψιμη. Με βάση αυτές τις παραδοχές και τα παρατηρούμενες μείγματα, οι ICA αλγόριθμοι προσπαθούν να βρουν την μήτρα ανάμειξης A ή τη μήτρα διαχωρισμού W , έτσι ώστε $u = Wx = WA s$ να είναι μια εκτίμηση των ανεξάρτητων σημάτων της πηγής.

Τεχνικά, η ανεξαρτησία μπορεί να ορίζεται από τις πυκνότητες πιθανότητας. Τα σήματα είναι στατιστικώς ανεξάρτητα όταν:

$$f_u(u) = \prod_i f_{u_i}(u_i)$$

όπου f_u είναι η συνάρτηση πυκνότητας πιθανότητας του u . Οπότε, είναι ισοδύναμο να πούμε ότι το διάνυσμα u είναι ομοιόμορφα κατανομημένο. Δυστυχώς, μπορεί να μην υπάρχει καμία μήτρα W που ικανοποιεί πλήρως την συνθήκη ανεξαρτησίας και να μην υπάρχει κάποια προκαθορισμένη διαδικασία για να αναζητηθεί το W . Αντ' αυτού, υπάρχουν αρκετοί αλγόριθμοι που επαναληπτικά προσεγγίζουν το W , έτσι ώστε να μεγιστοποιηθεί έμμεσα ανεξαρτησία.

Δεδομένου ότι είναι δύσκολο να μεγιστοποιηθεί άμεσα η συνθήκη ανεξαρτησίας που ορίστηκε προηγουμένως, όλοι οι κοινοί ICA αλγόριθμοι αναδιατυπώνουν το πρόβλημα, προκειμένου να βελτιστοποιηθεί επαναληπτικά μια ομαλή συνάρτησή της οποίας το ολικό βέλτιστο (είτε ολικό ελάχιστο είτε ολικό μέγιστο), συμβαίνει όταν τα διανύσματα εξόδου u είναι ανεξάρτητα. Για παράδειγμα, ο αλγόριθμος Infomax [40] βασίζεται στην παρατήρηση ότι η ανεξαρτησία μεγιστοποιείται όταν η εντροπία $H(u)$ έχει μεγιστοποιηθεί, όπου:

$$H(u) = - \int f_u(u) \log f_u(u) du \quad (3)$$

Ο αλγόριθμος του Infomax εκτελεί κλίση ανάβασης στα στοιχεία έτσι ώστε να μεγιστοποιηθεί το $H(\mathbf{u})$ [38][41]. Παίρνει το όνομά του από την παρατήρηση ότι η μεγιστοποίηση του $H(\mathbf{u})$ μεγιστοποιεί επίσης και την κοινή πληροφορία μεταξύ των διανυσμάτων εισόδου και εξόδου.

Ο αλγόριθμος του FastICA [42] είναι αναμφισβήτητα ο πιο γενικός, μεγιστοποίησης:

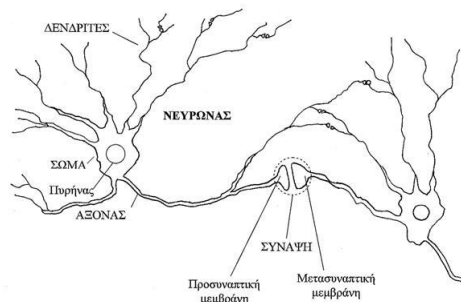
$$J(\mathbf{y}) \approx c [E\{G(\mathbf{y})\} - E\{G(\mathbf{v})\}]^2$$

όπου G είναι μια μη-τετραγωνική συνάρτηση, \mathbf{v} είναι μια Gaussian τυχαία μεταβλητή, και c είναι οποιαδήποτε θετική σταθερά. Έχει αποδειχθεί ότι η μεγιστοποίηση οποιασδήποτε συνάρτησης αυτής της μορφής, θα μεγιστοποιήσει επίσης και την ανεξαρτησία.

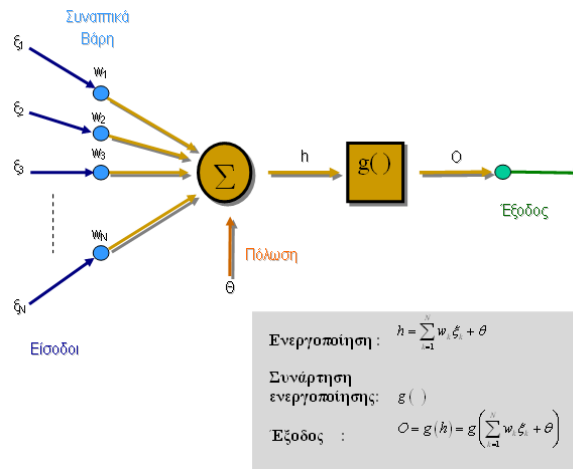
Οι αλγόριθμοι Infomax και FastICA μεγιστοποιούν όλες τις συναρτήσεις με το ίδιο ολικό βέλτιστο. Ως αποτέλεσμα, οι δύο αλγόριθμοι θα πρέπει να συγκλίνουν προς την ίδια λύση για κάθε συγκεκριμένο σύνολο δεδομένων. Στην πράξη, οι διαφορετικές διατυπώσεις της συνθήκης της ανεξαρτησίας έχει σχεδιαστεί για να επιτρέπει διαφορετικές τεχνικές προσέγγισης, και οι αλγόριθμοι βρίσκουν διαφορετικές λύσεις, λόγω των διαφορών μεταξύ αυτών των τεχνικών. Περιορισμένες εμπειρικές μελέτες δείχνουν ότι οι διαφορές απόδοσης μεταξύ των αλγορίθμων είναι ήσσονος σημασίας και εξαρτώνται από το σύνολο των δεδομένων [43].

6.3 Τεχνητά νευρωνικά δίκτυα

Τα τεχνητά νευρωνικά δίκτυα (Τ.Ν.Δ. [44][45]) είναι συστήματα εμπνευσμένα από την λειτουργία του νευρικού συστήματος και του εγκεφάλου. Στην πραγματικότητα είναι ένα μαθηματικό μοντέλο για την επεξεργασία πληροφορίας που προσεγγίζει την υπολογιστική και αναπαραστατική δυνατότητα μέσω συνάψεων. Στη φύση, τα βιολογικά νευρωνικά δίκτυα έχουν ως βασική μονάδα το νευρώνα (neuron) έτσι και στα Τ.Ν.Δ. υπάρχει ο τεχνητός νευρώνας. Όπως βλέπουμε και στα παρακάτω σχήματα ο πυρήνας του νευρώνα δέχεται σήματα από άλλους νευρώνες, μέσω καναλιών εισόδου που ονομάζονται δενδρίτες ($\xi_1, \xi_2, \dots, \xi_N$), και τα επεξεργάζεται (μέσω ενός αθροιστή Σ) για τη δημιουργία ενός καινούριου σήματος (h). Αν το σήμα αυτό είναι αρκετά ισχυρό, ενεργοποιείται (με βάση μια συνάρτηση ενεργοποίησης g) η έξοδος του νευρώνα και παράγεται ένα σήμα εξόδου (O) που μεταδίδεται μέσω ενός καναλιού εξόδου. Το κανάλι εξόδου ονομάζεται άξονας και η σύνδεσή του με τους δενδρίτες των άλλων νευρώνων γίνεται μέσω συνάψεων. Το μεταδιδόμενο σήμα μεταβάλλεται ανάλογα με την ισχύ της αντίστοιχης σύναψης.

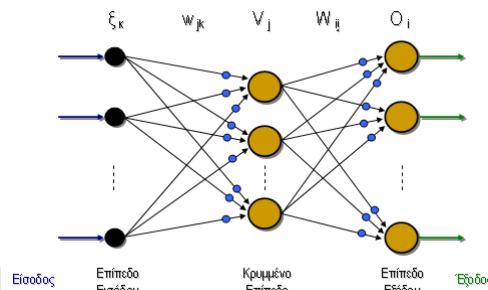


Εικόνα 6.10: Βιολογικό νευρωνικό δίκτυο



Εικόνα 6.11: Τεχνητό νευρωνικό δίκτυο

Τα βάρη των συνάψεων των νευρώνων καθορίζονται κατά την εκπαίδευση ενός δικτύου και στη συνέχεια ευρίσκεται η συνάρτηση ενεργοποίησης. Τα Τ.Ν.Δ. μπορεί να αποτελούνται από έναν ή οι περισσότερους νευρώνες. Στα Τ.Ν.Δ. με πολλούς νευρώνες, οι νευρώνες οργανώνονται σε μια ακολουθία επιπέδων και ονομάζονται πολυστρωματικά. Τα επίπεδα διακρίνονται σε επίπεδα εισόδου, εξόδου και κρυφά.



Εικόνα 6.12: Πολυστρωματικό τεχνητό νευρωνικό δίκτυο

6.4 Αναγνώριση προσώπου σε θερμογραφικές εικόνες

Για την αναγνώριση προσώπων σε θερμογραφικές εικόνες χρησιμοποιήθηκε ένα τεχνητό νευρωνικό δίκτυο οπισθοδιάδοσης τριών επιπέδων. Το επίπεδο εισόδου έχει 17 νευρώνες όσες και οι τιμές των χαρακτηριστικών που χρησιμοποιήθηκαν. Το επίπεδο εξόδου έχει 1 νευρώνα για να υποδείξει με μια τιμή από 0 έως 1, αν αναγνωρίστηκε το πρόσωπο από τις εικόνες ελέγχου και σε τι βαθμό. Οι νευρώνες του κρυφού επιπέδου προκύπτουν από τη σχέση:

$$Ar. \text{ νευρώνων κρυφού επιπέδου} = (ar. \text{ νευρώνων εισόδου} + ar. \text{ νευρώνων εξόδου})/2$$

Επομένως, ο αριθμός νευρώνων του κρυφού επιπέδου στη περίπτωση του παρόντος συστήματος είναι: $(17+1)/2=9$.

Για την εκπαίδευση του συστήματος χρησιμοποιήθηκαν 8 εικόνες του ίδιου προσώπου με διαφορετικές εκφράσεις και υπό διαφορετικές συνθήκες φωτισμού. Στο παρακάτω πίνακα αναφέρεται αναλυτικά τι είναι η κάθε εικόνα:

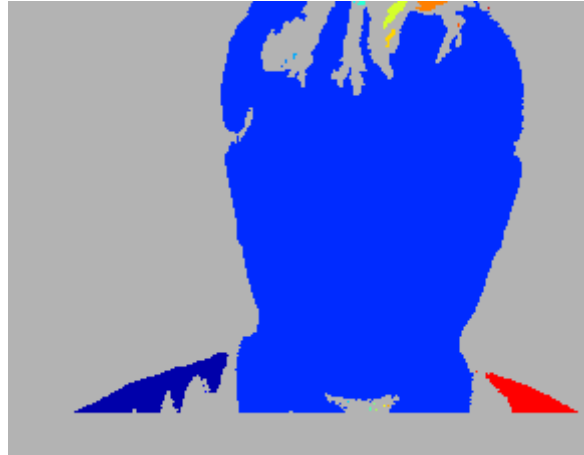
Εικόνα 1	Έκφραση 1 με γυαλιά (κανονικός φωτισμός)
Εικόνα 2	Έκφραση 2 χωρίς γυαλιά (κανονικός φωτισμός)
Εικόνα 3	Έκφραση 3 χωρίς γυαλιά (κανονικός φωτισμός)
Εικόνα 4	Έκφραση 4 χωρίς γυαλιά (κανονικός φωτισμός)
Εικόνα 5	Συνθήκες φωτισμού 1 (κανονικός φωτισμός, με γυαλιά)
Εικόνα 6	Συνθήκες φωτισμού 2 (σκοτάδι, με γυαλιά)
Εικόνα 7	Συνθήκες φωτισμού 3 (αριστερός προβολές, με γυαλιά)
Εικόνα 8	Συνθήκες φωτισμού 4 (προβολέας από μπροστά, χωρίς γυαλιά)

Πίνακας 6.2: Οι εικόνες εκπαίδευσης

Ενδεικτικά για μια (Εικόνα 1) από τις εικόνες εκπαίδευσης του προτεινόμενου συστήματος πήραμε τα εξής αποτελέσματα:



Εικόνα 6.13: Κατωφλίωση Otsu για την εικόνα 1



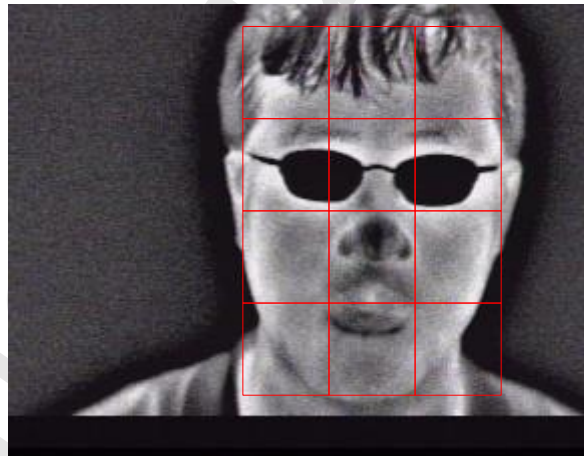
Εικόνα 6.14: Τμηματοποίηση της εικόνας 1 σε περιοχές (με ανοικτό μπλε η μεγαλύτερη περιοχή)



Εικόνα 6.15: Επιλογή της μεγαλύτερης περιοχής έπειτα από την τμηματοποίηση της εικόνας 1



Εικόνα 6.16: Επιλογή της περιοχής του προσώπου για την εικόνα 1



Εικόνα 6.17: Χωρισμός του προσώπου σε περιοχές (πλέγμα προσώπου) για την εικόνα 1

Όσο αφορά τις εικόνες ελέγχου έγιναν δύο πειράματα. Στο πρώτο πείραμα χρησιμοποιήθηκαν οι εξής εικόνες:

Εικόνα 9	Έκφραση 5 με γυαλιά (κανονικός φωτισμός)
Εικόνα 10	Έκφραση 6 με γυαλιά (κανονικός φωτισμός)
Εικόνα 11	Έκφραση 7 χωρίς γυαλιά (κανονικός φωτισμός)

Πίνακας 6.3: Οι εικόνες ελέγχου για το πρώτο πείραμα

Ενδεικτικά για μια (Εικόνα 11) από τις εικόνες ελέγχου του προτεινόμενου συστήματος πήραμε τα εξής αποτελέσματα:



Εικόνα 6.18: Κατωφλίωση Otsu για την εικόνα 11 (ελέγχου)



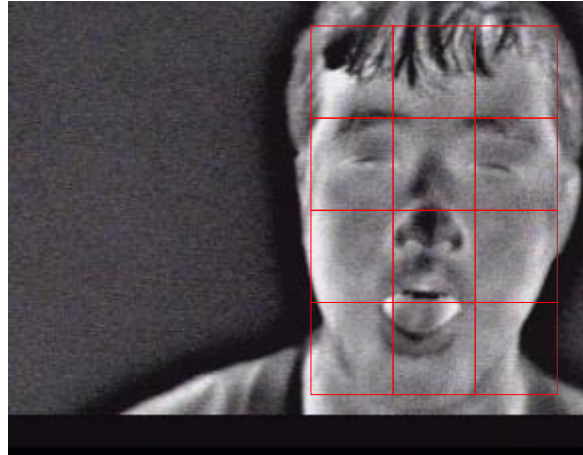
Εικόνα 6.19: Τμηματοποίηση της εικόνας 11 σε περιοχές (με ανοικτό μπλε η μεγαλύτερη περιοχή)



Εικόνα 6.20: Επιλογή της μεγαλύτερης περιοχής έπειτα από την τμηματοποίηση της εικόνας 11



Εικόνα 6.21: Επιλογή της περιοχής του προσώπου για την εικόνα 11



Εικόνα 6.22: Χωρισμός του προσώπου σε περιοχές (πλέγμα προσώπου) για την εικόνα 11

Το αποτέλεσμα της αναγνώρισης στο πρώτο πείραμα για τις τρεις εικόνες ελέγχου είναι:

>> Test

Test =

0.8266 0.7534 0.8754

Στο δεύτερο πείραμα χρησιμοποιήθηκαν εικόνες μεταμφίησης με γυαλιά ηλίου, μάσκα και περούκα:

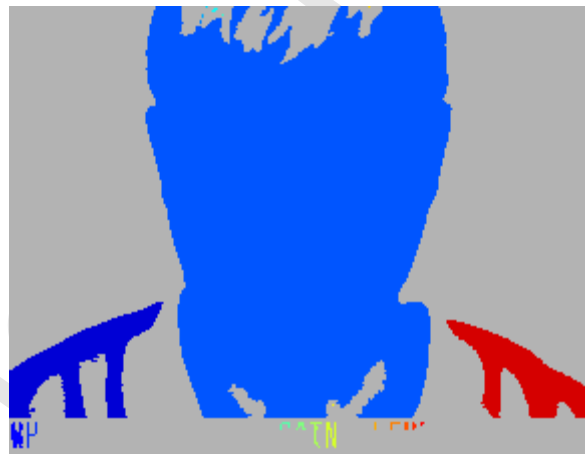
Εικόνα 10	Μεταμφίηση 1 με γυαλιά ηλίου (κανονικός φωτισμός)
Εικόνα 11	Μεταμφίηση 2 με μάσκα (κανονικός φωτισμός)
Εικόνα 12	Μεταμφίηση 3 με περούκα (Κανονικός φωτισμός)

Πίνακας 6.4: Οι εικόνες ελέγχου για το δεύτερο πίνακα

Ενδεικτικά για μια (Εικόνα 10) από τις εικόνες ελέγχου του προτεινόμενου συστήματος πήραμε τα εξής αποτελέσματα:



Εικόνα 6.23: Κατωφλίωση Otsu για την εικόνα 10 (ελέγχου)



Εικόνα 6.24: Τμηματοποίηση της εικόνας 10 σε περιοχές (με ανοικτό μπλε η μεγαλύτερη περιοχή)



Εικόνα 6.25: Επιλογή της μεγαλύτερης περιοχής έπειτα από την τμηματοποίηση της εικόνας 10



Εικόνα 6.26: Επιλογή της περιοχής του προσώπου για την εικόνα 10



Εικόνα 6.27: Χωρισμός του προσώπου σε περιοχές (πλέγμα προσώπου) για την εικόνα 10

Το αποτέλεσμα της αναγνώρισης για τις τρεις εικόνες ελέγχου στο δεύτερο πείραμα είναι:

>> Test

Test =

0.8490 0.6247 0.8767

ΚΕΦΑΛΑΙΟ 7 ΒΕΛΤΙΩΣΕΙΣ - ΕΠΕΚΤΑΣΕΙΣ - ΣΥΜΠΕΡΑΣΜΑΤΑ

Κατά την διάρκεια εκπόνησης της παρούσης εργασίας έγινε μια βιβλιογραφική έρευνα για τις τάσεις που υπάρχουν μέχρι σήμερα στον τομέα της αναγνώρισης προσώπου. Ο στόχος της αναγνώρισης προϋποθέτει μια σειρά από βήματα και λειτουργίες που από μόνα τους είναι ολόκληρα κεφάλαια στην υπολογιστική νοημοσύνη. Σημαντικοί αλγόριθμοι και τεχνικές από κάθε στάδιο μέχρι την αναγνώριση, μελετήθηκαν σε μεγαλύτερο βάθος και υλοποιήθηκαν σε ένα προτεινόμενο σύστημα αναγνώρισης προσώπου για θερμογραφικές εικόνες.

Σε πρώτη φάση η προεπεξεργασία μιας εικόνας που προορίζεται για την αναγνώριση, αφορά σημαντικές λειτουργίες που βελτιώνουν την ίδια την εικόνα και κάνουν πιο εφικτή την διαδικασία της αναγνώρισης. Η πιο γνωστή λειτουργία είναι η αφαίρεση του παρασκηνίου με την μέθοδο Otsu που χρησιμοποιήθηκε στο προτεινόμενο σύστημα. Βελτιώσεις που μπορούν να εφαρμοστούν αφορούν τρόπους κανονικοποίησης σε ένα ενιαίο μέγεθος εικόνας, ώστε για κάθε εικόνα που εισάγεται στο σύστημα, το κεφάλι και το πρόσωπο να βρίσκονται περίπου στην ίδια περιοχή. Αυτό μπορεί να ισχύσει και για εικόνες που το πρόσωπο έχει φωτογραφηθεί υπό γωνία και με κάποιους μετασχηματισμούς να μετατραπεί η προβολή σε πρόσθια. Σε περιπτώσεις που υπάρχει πρόβλημα με τις συνθήκες φωτισμού, στις θερμογραφικές εικόνες δεν υπάρχει ιδιαίτερο πρόβλημα αφού δεν λαμβάνονται υπόψη τιμές φωτεινότητας αλλά θερμοκρασίας. Αν κριθεί σκόπιμο, όμως μπορεί να γίνει αύξηση ή μείωση της μέσης τιμής φωτεινότητας ανά περιοχές της εικόνας, εφαρμόζοντας φίλτρα ενίσχυσης ή εξασθένησης.

Όσο αφορά την εξαγωγή του χάρτη θερμοκρασιών από μια θερμογραφική εικόνα, αξίζει να σημειωθεί ότι σε κάποιες θερμικές κάμερες η εξαγωγή θερμοκρασιών γίνεται αυτόματα και παράγεται μια αντίστοιχη εικόνα που δείχνει τις θερμοκρασίες σε μια κλίμακα αποχρώσεων του κόκκινου. Όσο πιο μεγάλη είναι η θερμοκρασία, τόσο πιο έντονη είναι η απόχρωση και αντίστοιχα, όσο πιο χαμηλή είναι η θερμοκρασία τόσο πιο αχνή είναι η απόχρωση. Αυτή η λειτουργία είναι εφικτό να υλοποιηθεί σε μελλοντική επέκταση της εργασίας, με την δημιουργία της εικόνας θερμοκρασιών με βάση τον χάρτη θερμοκρασιών.

Ο εντοπισμός προσώπου σε εικόνες είναι ένα μεγάλο κεφάλαιο από μόνο του. Οι επιλογές που υπάρχουν για τον εμπλουτισμό του προτεινόμενου συστήματος είναι πάρα πολλές. Στην παρούσα εργασία ο εντοπισμός αφορά μόνο πρόσωπα σε εργαστηριακές εικόνες προσώπων, δηλαδή εύρεση μόνο της περιοχής του προσώπου. Ένα πιο σύνθετο σύστημα θα μπορούσε να περιλαμβάνει εντοπισμό ανθρώπων (πρόσωπο, σώμα, χέρια κ.α.) σε πιο σύνθετες σκηνές με αντικείμενα ή άλλους ανθρώπους. Επίσης, σημαντικό πεδίο έρευνας είναι ο εντοπισμός ματιών (eye detection) και στις θερμογραφικές εικόνες. Πρώτο στοιχείο εντοπισμού είναι ότι οι θερμοκρασίες στα μάτια και στα φρύδια είναι μικρότερες.

Στο επόμενο στάδιο που αφορά την εξαγωγή χαρακτηριστικών, μπορούν να γίνουν προσθήκες επιπλέον χαρακτηριστικών. Για παράδειγμα αν στο εντοπισμό προσώπου γίνεται πιο λεπτομερής εντοπισμός και βρίσκονται και τα μάτια, ένα χαρακτηριστικό θα μπορούσε να είναι η απόσταση τους. Έτσι με βάση το πλέγμα προσώπου, τα μάτια βρίσκονται στα κέντρα των παραλληλογράμμων ακριβώς πάνω από τα παραλληλόγραμμα των 2 μάγουλων, τα οποία προσδιορίστηκαν στο Κεφάλαιο 5.

Τέλος, για το κομμάτι της αναγνώρισης τα πειράματα που έγιναν έδειξαν ότι στις θερμογραφικές εικόνες η μεταμπίεση με περούκες, μάσκες ή ψεύτικα γένια δεν επηρεάζει ιδιαίτερα τα αποτελέσματα. Τα γυαλιά είτε μυωπίας είτε ηλίου αλλοιώνουν λίγο τις πραγματικές θερμοκρασίες στην περιοχή των ματιών αφού οι φακοί τραβούν περισσότερη θερμοκρασία. Όμως, όπως προείπαμε τα μάτια έχουν στατιστικά τις χαμηλότερες θερμοκρασίες στο πρόσωπο (όπως και τα φρύδια), έτσι συνολικά τα γυαλιά δεν δημιουργούν αρνητική επίδραση. Εκτός από

τα τεχνητά νευρωνικά δίκτυα που χρησιμοποιήθηκαν και στο υλοποιημένο σύστημα, υπάρχουν και άλλοι μέθοδοι. Οπότε, είτε μπορεί να γίνει χρήση κάποιας άλλης μεθόδου, είτε συνδυασμός δύο ή περισσότερων. Γενικότερα, αξίζει να σημειωθεί ότι κάποιοι από τους αλγόριθμους και τις τεχνικές χρησιμοποιούνται σε εικόνες ανεξάρτητα αν είναι έγχρωμες, ασπρόμαυρες (στην κλίμακα του γκρι) ή θερμογραφικές. Έτσι, λοιπόν, το προτεινόμενο σύστημα μπορεί να επεκταθεί και σε έγχρωμες ή ασπρόμαυρες εικόνες.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] Kong, S., Heo, J., Abidi, B., Paik, J., and Abidi, M.: "Recent advances in visual and infrared face recognition - A review", *Computer Vision and Image Understanding*, Vol. 97, No.1, pp. 103-135, January 2005.
- [2] Kriegman, D., Yang, M., Ahuja, N.: "Detecting faces in images: a survey", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 1, pp. 34-58, January 2002.
- [3] Gu, Y., Thomas, T.: "A hybrid score measurement for HMM-based speaker verification", *Proceedings of IEEE International Conference Acoustics, Speech, and Signal Processing*, Vol. 1. pp. 317-320, March 1999.
- [4] Chen, Y., T., Wang, M., S.: "Human Face Recognition Using Thermal Image", *Department of Engineering Science, National Cheng Kung University*, 2002.
- [5] Imaging, Robotics, and Intelligent Systems laboratory, "UTK-IRIS database", University of Tennessee, Knoxville, [Online], Available: <http://www.cse.ohiostate.edu/otcbvs-bench>
- [6] Planck, M., "Planck's law", [Online], Available: http://en.wikipedia.org/wiki/Planck's_law
- [7] Incropera, F., P., DeWitt, D., P.: "Introduction to heat transfer", *John Wiley & sons*, 2nd ed., ch. 12, 1990.
- [8] Moghaddam, B., Pentland, A.: "Probabilistic visual learning for object representation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, July 1997.
- [9] Rowley, H., A., Baluja, S., Kanade, T.: "Neural Network-Based Face Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 23, no. 1, pp. 23-38, 1998.
- [10] Darrell, T., Essa, I., Pentland, A., "Correlation and Interpolation Networks for Real-time Expression Analysis/Synthesis", *MIT Press*, 1994.
- [11] Schneiderman, H., Kanade, T.: "A Statistical Method for 3D Object Detection Applied to Faces and Cars", *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 746 – 751, 2000.
- [12] Viola, P., Jones, M.: "Robust real-time face detection", *J. Computer Vision*, vol. 57, no. 2, pp. 137-154, 2004.
- [13] Franklin, C.: "Summed-Area Tables for Texture Mapping", *Crow Computer Sciences Laboratory-Xerox Palo Alto Research Center*, 1984.
- [14] Freund, Y., Schapire, R., E.: "A Short Introduction to Boosting", *Journal of Japanese Society for Artificial Intelligence*, vol. 14(5), pp. 771-780, September 1999.

- [15] Freund, Y., Schapire, R., E.: “A decision-theoretic generalization of on-line learning and an application to boosting”, In *Computational Learning Theory: Eurocolt '95*, pp. 23–37, Springer-Verlag, 1995.
- [16] Haar, A.: “Haar Wavelet – Haar matrix”, [Online], Available: http://en.wikipedia.org/wiki/Haar_wavelet#Haar_matrix
- [17] Fasela I., Fortenberry, B., Movellan, J.: “A generative framework for real time object detection and classification”, *Elsevier - Computer Vision and Image Understanding* 98, pp. 182–210, 2005.
- [18] Friedman J., Hastie, T., Tibshirani, R.: “Additive Logistic Regression: A statistical view of boosting”, *Stanford University, The Annals of Statistics*, Vol.28, No.2, 337-407, 2000.
- [19] Kearns, M.: “Thoughts on Hypothesis Boosting”, *Machine Learning class project*, December 1988.
- [20] Oliveira, V., A., Conci, A.: “Skin Detection using HSV color space”, *Computation Institute – Universidade Federal Fluminense – UFF*, 2009.
- [21] Pantic, M., Rothkrantz, Leon, J., M.: “Affect-Sensitive Multi-Modal Monitoring in Ubiquitous Computing: Advances and Challenges”, *Delft University of Technology, MediaMatica Department*, 2000.
- [22] Pantic, M., Rothkrantz, Leon, J., M.: “Toward an Affect-Sensitive Multimodal Human–Computer Interaction”, *Proceedings of the IEEE*, vol. 91, pp.1370-1390, September 2003.
- [23] Bassili, J., N.: “Facial motion in the perception of faces and of emotional expression”, *Journal of Experimental Psychology: Human Perception and Performance*, vol. 4(3), pp. 373-379, Aug 1978.
- [24] Gokturk, S., B.: “Model-Based Face Tracking for View-Independent Facial Expression Recognition”, In *Proceedings of the Fifth IEEE international Conference on Automatic Face and Gesture Recognition*, 2002
- [25] Chang, E., Chadwicke, O.: “Sketching Articulation and Pose for Facial Animation”, *M.-P. Cani, J. O'Brien (Editors)*, Brown University Eurographics/ACM SIGGRAPH Symposium on Computer Animation, 2006
- [26] Tian, Y.: “Evaluation of Face Resolution for Expression Analysis”, *IBM T. J. Watson Research Center*, 2001
- [27] Zhang, Y., Ji, Q.: “Active and Dynamic Information Fusion for Facial Expression Understanding”, *IEEE*, 2005
- [28] Pantic, M., Patras, I.: “Dynamics of Facial Expression: Recognition of Facial Actions and Their Temporal Segments From Face Profile Image Sequences”, *IEEE Transactions on Systems, Man, and Cybernetics — PART B: Cybernetics*, vol. 36, no. 2, April 2006.

- [29] **Farye, M.:** “Wavelet Transforms and their Applications to Turbulence, *LMD-CNRESc ,Annu. Rev. Fluid Mech*, vol. 24, pp. 395-457, 1992.
- [30] **Turk, M., Pentland, A.:** "Eigenfaces for Recognition", *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71-86, 1991.
- [31] **Pantic, M., Patras, I.:** “Dynamics of facial expression: Recognition of facial actions and their temporal segments from face profile image sequences.”, *IEEE Transactions On Systems, Man and Cybernetics - Part B*, Vol. 36, No. 2, pp. 433-449, 2006.
- [32] **Gokturk, S.B., Bouguet, J.Y., Tomasi, C., Girod, B.:** “Model-based face tracking for view independent facial expression recognition”, *Proc. IEEE Int’l Conf. Automatic Face and Gesture Recognition*, pp. 272-278, 2002.
- [33] **Cohen, I., Sebe, N., Garg, A. Chen, L., Huang, T., S.:** “Facial expression recognition from video sequences: temporal and static modeling”, *Journal Computer Vision and Image Understanding - Special issue on Face recognition*, vol. 91, July 2003.
- [34] **Tao, H., Huang, T., S.:** “Connected vibrations – a model analysis approach to nonrigid motion tracking”, *Proc. IEEE Int’l Conf. Computer Vision and Pattern Recognition*, pp. 735-740, 1998
- [35] **Tao, H., Huang T.,S.:** “Bézier Volume Deformation Model for Facial Animation and Video Tracking” , *Spiengel: Modelling and Motion Capture Techniques for Virtual Environments Lecture Notes in Computer Science*, vol. 1537, pp. 242-253, 1998.
- [36] **Cohn J.,F., Reed L., I., Ambadar, Z, Xiao, J.:** “Automatic Analysis and Recognition of Brow Actions and Head Motion in Spontaneous Facial Behavior, *Proc. IEEE Int’l Conf. Systems, Man & Cybernetics*, 2004.
- [37] **Coan J., A., Allen J.,J., B.:** “Handbook of Emotion Elicitation And Assessment”, *Oxford university press*, 2007
- [38] **Kirby, M., Sirovich, L.:** “Application of the Karhunen-Lokve Procedure for the Characterization of Human Faces”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.12, No1, 1990.
- [39] **Turk, M.,A., Pentland, A.,P.:** “Face recognition using eigenfaces”, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1991.
- [40] **Bell, A.,J., Sejnowski, T.,J.:** “An information maximisation approach to blind separation and blind deconvolution”, *Neural Computation*, November 1995.
- [41] **Sirovich, L., Kirby, M.:** “Low-dimensional procedure for characterization of human faces”, *J. Opt. Soc. Amer.*, vol. 4, no. 3, pp. 519-524, March 1987.
- [42] **Hyvärinen, A.:** “Fast ICA for Noisy Data using Gaussian Moments”, Helsinki University of Technology, *Laboratory of Computer and Information Science*, 1999

- [43] Draper, B., A., Baek, K., Bartlett Stewart, M., Beveridgea Ross, J.: “Recognizing faces with PCA and ICA”, Elsevier - Computer Vision and Image Understanding, vol. 91, issues 1–2, pp. 115–137, July–August 2003.
- [44] Stergiou, C., Siganos, D., : “Neural Networks”, *School of Computer Science & Engineering - The University of New South Wales* [Online], Available: http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html.
- [45] McCulloch, W., S., Pitts, W.: “A logical calculus of the ideas immanent in nervous activity”, *Springer - Bulletin of Mathematical Biology*, vol. 5, nr. 4, pp. 115-133, 1943.