



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Ανίχνευση Προσώπων και
Σημείων Ενδιαφέροντος σε Πρόσωπα**

Παναγιώτης Β. Περάκης

**Επιβλέποντες: Θεοχάρης Θεοχάρης, Αναπληρωτής Καθηγητής ΕΚΠΑ
Γεώργιος Πασσαλής, Διδάκτωρ ΕΚΠΑ**

ΑΘΗΝΑ

ΜΑΙΟΣ 2008

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ανίχνευση Προσώπων και
Σημείων Ενδιαφέροντος σε Πρόσωπα

Παναγιώτης Β. Περάκης

A.M.: M863

Επιβλέποντες: **Θεοχάρης Θεοχάρης**, Αναπληρωτής Καθηγητής ΕΚΠΑ
Γεώργιος Πασσαλής, Διδάκτωρ ΕΚΠΑ

Εξεταστική

Επιτροπή: **Θεοχάρης Θεοχάρης**, Αναπληρωτής Καθηγητής ΕΚΠΑ
Εμμανουήλ Σαγκριώτης, Αναπληρωτής Καθηγητής ΕΚΠΑ

Μάιος 2008

ΠΕΡΙΛΗΨΗ

Στην παρούσα διπλωματική εργασία παρουσιάζονται μέθοδοι για την ανίχνευση προσώπων και την ανίχνευση σημείων ενδιαφέροντος πάνω σε πρόσωπα σε 2D εικόνες. Η ανίχνευση προσώπου είναι το πρώτο βήμα σε ένα αυτοματοποιημένο σύστημα αναγνώρισης προσώπων. Η ευθυγράμμιση προσώπων με τη χρήση σημείων ενδιαφέροντος ή οροσήμεων καθιστά το ζητούμενο πρόσωπο και το υπόδειγμα συγκρίσιμα.

Σε αυτό το πλαίσιο εφαρμόστηκε και αξιολογήθηκε η μέθοδος ανίχνευσης αντικειμένων που αρχικά προτάθηκε από τους Viola & Jones και βελτιώθηκε από τον Lienhart. Επίσης εφαρμόστηκαν και αξιολογήθηκαν μέθοδοι για την ανίχνευση σημείων ενδιαφέροντος στηριγμένες στον ανιχνευτή γωνιών των Harris & Stephens και στον ανιχνευτή συμμετριών του Reissfeld.

Οι αλγόριθμοι που εφαρμόστηκαν πρόκειται να ενσωματωθούν σε ένα ευρύτερο σύστημα αναγνώρισης προσώπων που αναπτύσσεται από το Εργαστήριο Γραφικών του Τμήματος Πληροφορικής του ΕΚΠΑ σε συνεργασία με το Computational Biomedicine Lab του Πανεπιστημίου του Houston.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Γραφικά, Επεξεργασία Εικόνας, Τεχνητή Όραση,
Αναγνώριση Προτύπων, Βιομετρία
ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Αναγνώριση προσώπου, ανίχνευση προσώπου, σημεία
ενδιαφέροντος, ανίχνευση γωνιών, ανίχνευση συμμετριών

ABSTRACT

In this thesis we present methods for human face detection and landmark detection on human faces in 2D images. Face detection is the first step in automated face recognition systems. Facial landmarks aid face alignment.

In this framework the method for object detection initially proposed by Viola & Jones and improved by Lienhart is implemented and evaluated. Also methods for landmark detection based on the corner detector proposed by Harris & Stephens and on the symmetry detector proposed by Reisfeld are implemented and evaluated.

The algorithms implemented will be incorporated in a broader system for face recognition that is developed by the Graphics Lab of the Informatics Department of the University of Athens in collaboration with the Computational Biomedicine Lab of the University of Houston.

SUBJECT AREA: Graphics, Image Processing, Computer Vision, Pattern Recognition, Biometrics
KEY WORDS: Face recognition, face detection, landmark detection, corner detection, symmetry detection

Αφιερώνεται στους Γονείς μου

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΡΟΛΟΓΟΣ	15
Κεφάλαιο 1 :	
ΕΙΣΑΓΩΓΗ	17
1.1 Γενικά	17
1.2 Περιγραφή του Προβλήματος της Αναγνώρισης Προσώπων	18
1.3 Αντικείμενο της Διπλωματικής Εργασίας	20
1.4 Διάρθρωση της Διπλωματικής Εργασίας	20
Κεφάλαιο 2 :	
ΑΝΙΧΝΕΥΣΗ ΠΡΟΣΩΠΩΝ	21
2.1 Γενικά	21
2.2 Περιγραφή του Προβλήματος της Ανίχνευσης Προσώπων	21
2.3 Μέθοδοι Ανίχνευσης/Εντοπισμού Προσώπου	22
2.4 Μικρό Ιστορικό των κυριότερων Μεθόδων Ανίχνευσης	24
2.5 Μέθοδοι Βασισμένες στην Εμφάνιση	25
2.5.1 Δείγματα προσώπων και μη-προσώπων	26
2.5.2 Χρησιμοποιούμενοι Ταξινομητές	28
2.5.3 Αναπαράσταση Προσώπων	28
2.5.4 Προεπεξεργασία	29
2.5.5 Αναζήτηση στον Χώρο και την Κλίμακα	30
2.5.6 Μετεπεξεργασία	30
2.6 Ανίχνευση Προσώπου κατά Viola & Jones	32
2.6.1 Υπολογισμός Χαρακτηριστικών	32
2.6.2 Επιλογή Χαρακτηριστικών με την AdaBoost	34
2.6.3 Κατασκευή του Αδύναμου Ταξινομητή	36
2.6.4 Ταξινόμηση με έναν Καταράκτη Ταξινομητών	37
2.7 Ανίχνευση Προσώπου κατά Lienhart	40
2.8 Σύνοψη του Συστήματος Ανίχνευσης Προσώπου	42

2.9 Μέθοδος Haar+AdaBoost : Συμπεράσματα	42
---	-----------

Κεφάλαιο 3 :

ΥΛΟΠΟΙΗΣΗ ΣΥΣΤΗΜΑΤΟΣ ΑΝΙΧΝΕΥΣΗΣ ΠΡΟΣΩΠΟΥ	45
3.1 Η Βιβλιοθήκη OpenCV της Intel	45
3.2 Ανίχνευση Προσώπων με την OpenCV	45
3.2.1 Η Εκπαίδευση του Ταξινομητή	47
3.2.2 Η Συνάρτηση Ανίχνευσης	47
3.2.3 Οι Μέθοδοι Ανίχνευσης	48
3.3 Πειραματικές Δοκιμές	49
3.3.1 Οι Αρχικές Δοκιμές	50
3.3.2 Περαιτέρω Ανάλυση	51
3.3.3 Προεπεξεργασία με Εξισορρόπηση Ιστογράμματος	53
3.3.4 Κατάργηση της Συγχώνευσης των Ανιχνεύσεων	54
3.3.5 Προσαρμογή του Παραθύρου Ενδιαφέροντος	54
3.3.6 Συνδυασμός των Μεθόδων Discrete και Gentle AdaBoost	55
3.3.7 Ο Αλγόριθμος Μετεπεξεργασίας	57
3.4 Συμπεράσματα : Η Τελική Διαδικασία Ανίχνευσης	58

Κεφάλαιο 4 :

ΑΝΙΧΝΕΥΣΗ ΣΗΜΕΙΩΝ ΕΝΔΙΑΦΕΡΟΝΤΟΣ ΣΕ ΠΡΟΣΩΠΑ	59
4.1 Γενικά	59
4.2 Ο Ανιχνευτής Γωνιών του Harris	61
4.3 Ο Ανιχνευτής Συμμετριών του Reisfeld	64
4.4 Ανίχνευση Τοπικών Ανακλαστικών Συμμετριών	68

Κεφάλαιο 5 :

ΥΛΟΠΟΙΗΣΗ ΣΥΣΤΗΜΑΤΟΣ ΑΝΙΧΝΕΥΣΗΣ ΣΗΜΕΙΩΝ ΕΝΔΙΑΦΕΡΟΝΤΟΣ ΣΕ ΠΡΟΣΩΠΑ	73
5.1 Γενικά	73
5.2 Υλοποίηση των Ανιχνευτών	74

5.2.1 Ανιχνευτής Γωνιών του Harris	74
5.2.2 Ανιχνευτής Συμμετριών του Reifeld	75
5.2.3 Ανιχνευτής Τοπικής Ανακλαστικής Συμμετρίας	77
5.3 Πειραματικές Δοκιμές	78
5.3.1 Δοκιμές πάνω σε Γεωμετρικά Σχήματα	79
5.3.2 Δοκιμές του Ανιχνευτή Γωνιών πάνω σε Πρόσωπα	81
5.3.3 Δοκιμές του Ανιχνευτή Συμμετριών πάνω σε Πρόσωπα	82
5.3.4 Δοκιμές του Ανιχνευτή Τοπικών Συμμετριών πάνω σε Πρόσωπα	84
5.4 Συμπεράσματα : Η Τελική Διαδικασία Ανίχνευσης	85
 Κεφάλαιο 6 :	
ΕΠΙΛΟΓΟΣ	87
ΟΡΟΛΟΓΙΑ	89
ΑΚΡΩΝΥΜΙΑ	91
ΑΝΑΦΟΡΕΣ / ΒΙΒΛΙΟΓΡΑΦΙΑ	93

ΠΡΟΛΟΓΟΣ

Η παρούσα διπλωματική εργασία αποτελεί τον επίλογο των μεταπτυχιακών μου σπουδών στα πλαίσια του Μεταπτυχιακού Διπλώματος Ειδίκευσης «Νέες Τεχνολογίες Πληροφορικής και Επικοινωνιών» του Τμήματος Πληροφορικής και Τηλεπικοινωνιών. Το περιεχόμενο της εργασίας αφορά την ανίχνευση προσώπων και την ανίχνευση σημείων ενδιαφέροντος πάνω σε πρόσωπα που απεικονίζονται σε 2Δ εικόνες.

Πριν την παρουσίαση της εργασίας αυτής, αισθάνομαι την ανάγκη να ευχαριστήσω θερμά τον Αναπληρωτή Καθηγητή Θεοχάρη Θεοχάρη για την καθοδήγηση του αλλά και διότι μου έδωσε τη δυνατότητα να ασχοληθώ με ένα τόσο ενδιαφέρον θέμα. Επίσης θα ήθελα να ευχαριστήσω τον Δρ. Γιώργο Πασσαλή για την πολύπλευρη βοήθεια και τον χρόνο που μου προσέφερε κατά την εκπόνηση της εργασίας αυτής. Ακόμη δεν θα μπορούσα να παραλείψω την συμβολή του Καθηγητή Ιωάννη Κακαδιάρη του Πανεπιστημίου του Houston και της ομάδας του, τους οποίους και ευχαριστώ. Τέλος θέλω επίσης να ευχαριστήσω θερμά όλους όσους μου συμπαραστάθηκαν ή και με ανέχθηκαν κατά τη διάρκεια της εκπόνησής της.

Αθήνα 22 Μαΐου 2008
Παναγιώτης Β. Περάκης

Κεφάλαιο 1

ΕΙΣΑΓΩΓΗ

1.1 Γενικά

Η *αναγνώριση προσώπου* (face recognition) είναι μία από τις κύριες κατευθύνσεις της *βιομετρίας* (biometrics). Τα τελευταία χρόνια έχει γίνει όλο και πιο σημαντική λόγω των αυξημένων απαιτήσεων σε συστήματα ασφαλείας, της ανάγκης για εξακρίβωση ταυτότητας στον ψηφιακό κόσμο του διαδικτύου και των κινητών συσκευών, και έγινε εφικτή λόγω των ταχύτατων εξελίξεων στις τεχνολογίες των ψηφιακών καμερών, αλλά και στις τεχνικές επεξεργασίας εικόνας, αναγνώρισης προτύπων και τεχνητής όρασης. Η αναγνώριση προσώπου έχει σημαντικά πλεονεκτήματα σε σχέση με τις άλλες τεχνολογίες της βιομετρίας, όπως η αναγνώριση δακτυλικού αποτυπώματος, η αναγνώριση φωνής, η αναγνώριση της ίριδας του ματιού και η αναγνώριση υπογραφής: είναι διαδικασία πιά φυσική, λιγότερο αδιάκριτη και ενοχλητική και ευκολότερη στη χρήση της [ΗΙΕΤ00]. Έτσι, η έρευνα στην αναγνώριση προσώπου δεν έχει μόνο ως κίνητρο τις προκλήσεις που θέτει το ίδιο το πρόβλημα, αλλά και τις αναρίθμητες εφαρμογές, όπου είναι απαραίτητη η εξακρίβωση της ταυτότητας των ανθρώπων.

Για τον άνθρωπο η αναγνώριση προσώπων είναι μία καθημερινή, φυσική και εύκολη σχετικά διαδικασία. Αντίθετα, η αναγνώριση προσώπων με τη βοήθεια ενός ευφυούς συστήματος έχει αποδειχθεί μία δύσκολη προσπάθεια, ειδικά σε μη ελεγχόμενα περιβάλλοντα, όπου η θέση παρατήρησης, ο φωτισμός, οι μορφασμοί και τα αντικείμενα ή αξεσουάρ που αποκρύπτουν το πρόσωπο, ποικίλουν χαρακτηριστικά. Παρολαυτά, η απόδοση των συστημάτων αναγνώρισης προσώπου έχει βελτιωθεί σημαντικά από την εποχή που παρουσιάστηκε το πρώτο αυτόματο σύστημα αναγνώρισης προσώπου από τον Kanade [KAN73], ώστε η αναγνώριση να μπορεί πλέον να διεξαχθεί επιτυχώς σε πραγματικό χρόνο κάτω από συγκεκριμένες ελεγχόμενες συνθήκες.

Ένα σύστημα αναγνώρισης προσώπου (face recognition) αναμένεται να αναγνωρίζει αυτόματα πρόσωπα που βρίσκονται σε φωτογραφίες ή βίντεο. Το σύστημα μπορεί να λειτουργεί με δύο τρόπους [LI05, YANG02]:

1. Διακρίβωση/επαλήθευση προσώπου (face verification/authentication), και
2. Ταυτοποίηση προσώπου (face identification).

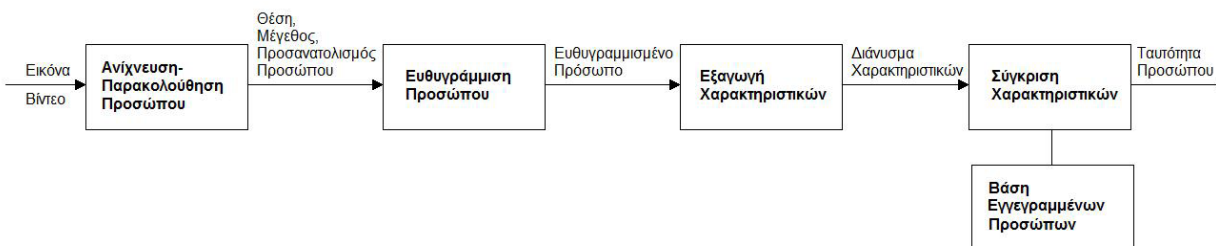
Η *διακρίβωση* συνίσταται στην σύγκριση ενός ζητούμενου προσώπου με το συγκεκριμένο δείγμα του προσώπου του οποίου ισχυρίζεται ότι έχει την ταυτότητα και στην *επαλήθευση* ή όχι της ταυτότητάς του.

Η *ταυτοποίηση* συνίσταται στην σύγκριση ενός ζητούμενου προσώπου με ένα σύνολο από δείγματα προσώπων μιας βάσης για να προσδιοριστεί σε ποιό ταιριάζει περισσότερο ώστε να καθοριστεί η ταυτότητά του.

1.2 Περιγραφή του προβλήματος της Αναγνώρισης Προσώπων

Η αναγνώριση προσώπου είναι ένα πρόβλημα *αναγνώρισης προτύπων* (pattern recognition). Ένα *δοκιμαστικό πρόσωπο* (face probe), ως ένα τρισδιάστατο αντικείμενο που υπόκειται σε μεταβολές ανάλογα με τον φωτισμό, τον προσανατολισμό, τους μορφασμούς κλπ., πρέπει να αναγνωρισθεί συγκρινόμενο με 2-διάστατες ή 3-διάστατες εικόνες μιάς *συλλογής προσώπων* (face gallery), που είναι διαθέσιμες σε μία βάση δεδομένων.

Ένα σύστημα αναγνώρισης προσώπου αποτελείται γενικά από τέσσερα υποσυστήματα [LI05], όπως φαίνεται στην Εικ. 1.1: την ανίχνευση, την ευθυγράμμιση, την εξαγωγή χαρακτηριστικών και την σύγκριση, όπου ο εντοπισμός και η κανονικοποίηση των προσώπων, μέσω της ανίχνευσης και της ευθυγράμμισης, προηγούνται της αναγνώρισης, μέσω της εξαγωγής των χαρακτηριστικών και της σύγκρισης με τη βάση των εγγεγραμμένων προσώπων.



Εικ. 1.1. Διάγραμμα διαδικασίας αναγνώρισης προσώπου [LI05]

Η *ανίχνευση προσώπου* (face detection) αναφέρεται στον έλεγχο αν υπάρχουν πρόσωπα σε μία εικόνα και στον προσδιορισμό της θέσης και των διαστάσεών τους. Ο *εντοπισμός προσώπου* (face localization) αναφέρεται ειδικότερα στον προσδιορισμό της

θέσης και των διαστάσεων του προσώπου, στην περίπτωση που σε μία εικόνα υπάρχει σίγουρα ένα μόνο πρόσωπο. Στην περίπτωση του βίντεο τα πρόσωπα που ανιχνεύονται πρέπει να ανιχνεύονται διαρκώς σε κάθε πλαίσιο, με ένα σύστημα *παρακολούθησης προσώπου (face tracking)*.

Η *ευθυγράμμιση προσώπου (face alignment/registration)* στοχεύει στον εντοπισμό του προσώπου με μεγαλύτερη ακρίβεια, στον προσδιορισμό του μεγέθους και του προσανατολισμού του (πόζα), ενώ η ανίχνευση προσώπου πετυχαίνει απλά ένα χοντρικό προσδιορισμό της θέσης και του μεγέθους του προσώπου. Για την ευθυγράμμιση του προσώπου επιλέγονται χαρακτηριστικά *σημεία αναφοράς ή ορόσημα* του προσώπου (*fiducial points/landmarks*), από συγκεκριμένα *σημεία ενδιαφέροντος (interest points)* που εντοπίζονται στην εικόνα του προσώπου, όπως τα μάτια, τα φρύδια, η μύτη και το στόμα. Τα πρόσωπα κανονικοποιούνται σε σχέση με τις γεωμετρικές τους ιδιότητες, όπως το μέγεθος και ο προσανατολισμός, χρησιμοποιώντας γεωμετρικούς μετασχηματισμούς. Τέλος, μια περαιτέρω κανονικοποίηση μπορεί να εφαρμοστεί, σε σχέση με τον φωτισμό και τις χρωματικές συνιστώσες της εικόνας του προσώπου, ώστε η εικόνα να περιέχει συγκεκριμένα επίπεδα διαβαθμίσεων του γκρι.

Η *εξαγωγή χαρακτηριστικών (feature extraction)* έρχεται μετά τη γεωμετρική και φωτομετρική κανονικοποίηση για να προσφέρει την απαραίτητη πληροφορία, ώστε να μπορεί να γίνει διάκριση μεταξύ των προσώπων διαφορετικών ατόμων, ανεξάρτητα από τις γεωμετρικές και φωτομετρικές διαφοροποιήσεις που υπάρχουν στις εικόνες.

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

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

1.3 Αντικείμενο της Διπλωματικής Εργασίας

Οι στόχοι της εργασίας αυτής είναι ο σχεδιασμός, η ανάπτυξη, η εφαρμογή και η αξιολόγηση αλγοριθμικών μεθόδων που αφορούν τον εντοπισμό και την ευθυγράμμιση προσώπου και θα χρησιμοποιηθούν σε ένα ευρύτερο σύστημα αναγνώρισης προσώπου [ΠΑΣ07, ΚΑΚ07], που έχει αναπτυχθεί από το Εργαστήριο Γραφικών του ΕΚΠΑ σε συνεργασία με Computational Biomedicine Lab [CBL] του Πανεπιστημίου του Houston.

Πιο συγκεκριμένα, στο κομμάτι που αφορά στον εντοπισμό προσώπου σε 2Δ εικόνες θα ασχοληθούμε με τη μέθοδο των Viola & Jones [VIOL01a, VIOL01b], όπως αυτή βελτιώθηκε από τον Lienhart κ.ά. [LIEN02a, LIEN02b]. Όσον αφορά δε στο κομμάτι του εντοπισμού σημείων ενδιαφέροντος σε 2Δ εικόνες προσώπων θα ασχοληθούμε με τη μέθοδο ανίχνευσης γωνιών του Harris [HAR88] και την μέθοδο ανίχνευσης συμμετριών του Reisfeld [REIS92, REIS95].

1.4 Διάρθρωση της Διπλωματικής Εργασίας

Η παρούσα εργασία οργανώνεται ως εξής: Στο Κεφάλαιο 2 παρουσιάζεται η μέθοδος των Viola & Jones και η βελτιωμένη μέθοδος του Lienhart για τον εντοπισμό προσώπου σε 2Δ εικόνες. Στο Κεφάλαιο 3 γίνεται η προσαρμογή και η εφαρμογή της σε πειραματικά δεδομένα για την αξιολόγηση της αποτελεσματικότητάς της. Στο Κεφάλαιο 4 παρουσιάζονται η μέθοδος εντοπισμού γωνιών του Harris και η μέθοδος εντοπισμού συμμετριών του Reisfeld. Στο Κεφάλαιο 5 οι παραπάνω μέθοδοι εφαρμόζονται σε πειραματικά δεδομένα προσώπων για την ανίχνευση σημείων ενδιαφέροντος και αξιολογείται η αποτελεσματικότητά τους. Τέλος στο Κεφάλαιο 6 συγκεντρώνονται τα συμπεράσματα της εργασίας και τίθενται οι κατευθύνσεις περαιτέρω έρευνας.

Κεφάλαιο 2

ΑΝΙΧΝΕΥΣΗ ΠΡΟΣΩΠΩΝ

2.1 Γενικά

Η *ανίχνευση προσώπου* (face detection) είναι το πρώτο βήμα στην διαδικασία *αναγνώρισης προσώπου* (face recognition). Η αξιοπιστία του συστήματος ανίχνευσης παίζει κυρίαρχο ρόλο στην απόδοση και χρησιμότητα ολόκληρου του συστήματος αναγνώρισης προσώπου.

Αν δοθεί μία φωτογραφία ή ένα βίντεο, 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.

Η μέθοδος των Viola & Jones [VIOL01a, VIOL01b] με τις βελτιώσεις του Lienhart [LIEN02a, LIEN02b] που στηρίζεται στην εκμάθηση με την τεχνική Adaboost και την χρήση χαρακτηριστικών γνωρισμάτων της μορφής Haar, εμφανίζεται μέχρι σήμερα να έχει την καλύτερη απόδοση ως προς την αποτελεσματικότητα και την ταχύτητα [LI05].

Το ενδιαφέρον της μεθόδου έγκειται επίσης στο ότι είναι υλοποιημένη από τον Lienhart κ.ά. ως τμήμα της βιβλιοθήκης ανοικτού κώδικα OpenCV της Intel [OCVL].

2.2 Περιγραφή του προβλήματος της Ανίχνευσης Προσώπου

Η ανίχνευση προσώπου μπορεί να θεωρηθεί σαν ένα πρόβλημα ταξινόμησης (classification) δύο κλάσεων, κατά το οποίο μία περιοχή μίας εικόνας ταξινομείται ως "πρόσωπο" ή "μη-πρόσωπο" [YANG02]. Η ανίχνευση προσώπου έγκειται στην ανακάλυψη και εντοπισμό των ανθρώπινων προσώπων σε μια εικόνα ανεξαρτήτως

- της θέσης τους
- του μεγέθους τους (κλίμακας)

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

Όλα τα παραπάνω αποτελούν τις κύριες αιτίες της μεγάλης δυσκολίας που εμφανίζει το πρόβλημα της ανίχνευσης ενός προσώπου [YANG02].

Αυτές οι μεγάλες παραλλαγές στην εμφάνιση του προσώπου, στην στάση, στην έκφραση και τον φωτισμό κάνουν πολύ πολύπλοκο το πολύπτυχο (manifold) του χώρου των προσώπων και δυσδιάκριτα τα όρια προσώπων / μη-προσώπων. Ένας μη γραμμικός ταξινομητής απαιτείται για να είναι εφικτή η διαχείριση αυτής της πολύπλοκης κατάστασης, καθώς επίσης και ένα μεγάλο σύνολο δειγμάτων εκπαίδευσης, τα οποία θα εμπεριέχουν αυτήν την ποικιλότητα στην εμφάνιση των προσώπων. Η ταχύτητα είναι επίσης ένα σημαντικό θέμα για την απόδοση του συστήματος σε πραγματικό χρόνο. Έχει γίνει σημαντική ερευνητική προσπάθεια για την δημιουργία σύνθετων και γρήγορων ταξινομητών και από το 1990 έχει γίνει σημαντική πρόοδος στο θέμα αυτό [L105].

2.3 Μέθοδοι Ανίχνευσης/Εντοπισμού Προσώπου

Οι υπάρχουσες τεχνικές για την ανίχνευση (face detection) ή τον εντοπισμό προσώπων (face localization) σε μονόχρωμες ή έγχρωμες εικόνες κατηγοριοποιούνται κατά τον Yang [YANG02], ως εξής:

- Μέθοδοι βασισμένες στη γνώση (knowledge-based methods):

Αυτές οι μέθοδοι κωδικοποιούν την ανθρώπινη γνώση σε κανόνες αναφορικά με το τί συνιστά ένα τυπικό πρόσωπο. Συνήθως κωδικοποιούνται από τον ερευνητή οι σχέσεις (θέσεις, αποστάσεις) ανάμεσα στα χαρακτηριστικά του προσώπου. Οι δυσκολίες της μεθόδου οφείλονται στην 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 παρατήρησης ή οι συνθήκες φωτισμού ποικίλουν. Με βάση τα δομικά αυτά χαρακτηριστικά (φρύδια, μάτια, στόμα κλπ.) χτίζονται *στατιστικά μοντέλα μορφών* (statistical shape models) [DRYD98], 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):

Εδώ χρησιμοποιούνται διάφορα προκαθορισμένα τυπικά υποδείγματα (templates) για να περιγράψουν το πρόσωπο σαν σύνολο ή τα χαρακτηριστικά του προσώπου 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, την 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 την 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 επιτυχημένες τεχνικές για ανίχνευση προσώπου είναι οι μέθοδοι που βασίζονται στην εμφάνιση (appearance-based). 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 [LI05].

2.4 Μικρό Ιστορικό των κυριότερων Μεθόδων Ανίχνευσης

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) [CTB92]

- 1995: Lanitis et al.

Χρήση Μοντέλων Ενεργών Μορφών (Active Shape Models) [LTC95]

- 1997: Osuna et al.

Χρήση μη γραμμικών Μηχανών Ανυσματικής Στήριξης (SVM) [OSUN97a]

- 1998: Papageorgiou and Poggio

Χρήση κυματιδίων Haar (Haar wavelets) με SVM ταξινομητή [PAP98]

- 1998: Sung and Poggio

Χρήση Γκαουσιανών κατανομών και πολυστρωματικών αντιλήπτρων (multilayer perceptron) [SUNG98]

- 1998: Rowley et al.

Χρήση συνόλου νευρωνικών δικτύων (neural networks) [ROWL98b]

- 2000: Schneiderman and Kanade

Χρήση στατιστικής για την τοπική εμφάνιση και εκμάθηση με AdaBoost [SCHN00b]

- 2001: Viola and Jones.

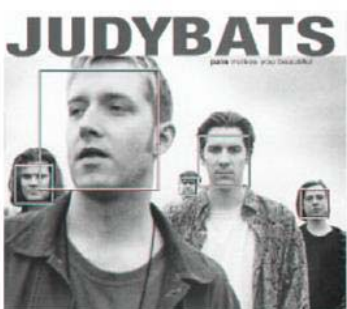
Χρήση των κυματιδίων τύπου Haar σε ταξινομητές ενός-μόνο χαρακτηριστικού, που εκπαιδεύονται με τον AdaBoost αλγόριθμο και οργανώνονται σε μια διαδοχή αποφάσεων [VIOL01b]

- 2002: Lienhart et al.

Χρήση ενός εκτεταμένου συνόλου περιστραμμένων χαρακτηριστικών Haar για την αντιμετώπιση των στροφών στο επίπεδο εικόνας [LIEN02b]

- 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 [LI02]



- Viola & Jones



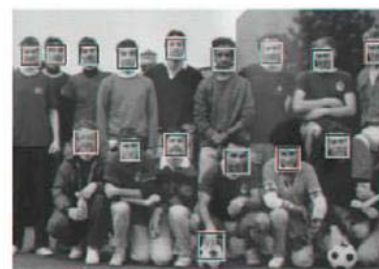
- Li et al.



- Schneiderman et al.



- Rowley et al.



Εικ. 2.1: 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 [YANG02, LI05].

2.5 Μέθοδοι Βασισμένες στην Εμφάνιση

Οι μέθοδοι που είναι βασισμένες στην εμφάνιση (appearance-based methods), εμφανίζουν ορισμένα κοινά χαρακτηριστικά κατά την διαδικασία ανίχνευσης των προσώπων:

- Σύνολο εκπαίδευσης θετικών και συνήθως 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 κοινά χαρακτηριστικά θα εξετάσουμε αναλυτικότερα στις παρακάτω παραγράφους.

2.5.1 Δείγματα προσώπων και μη-προσώπων

Για την εκπαίδευση του ταξινομητή χρησιμοποιούνται θετικά όσο και αρνητικά δείγματα προσώπων:

- Θετικά δείγματα

Λήψη όσο το δυνατό μεγαλύτερης ποικιλίας δειγμάτων, ώστε να περιλαμβάνονται όσο το δυνατόν περισσότερες εκδοχές του προσώπου που πρέπει να ανιχνεύεται.

Κόψιμο και κανονικοποίηση κάθε εικόνας προσώπου σε ένα συγκεκριμένο μέγεθος, π.χ. 20×20 pixels (Εικ. 2.2).

Αύξηση του πλήθους των θετικών δειγμάτων με τη δημιουργία εικονικών δειγμάτων (Εικ. 2.4) [SUNG98].



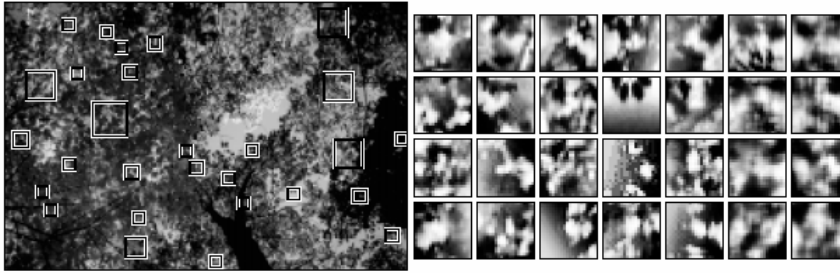
Εικ. 2.2: Πρόσωπα που χρησιμοποιούνται στο σύνολο των θετικών δειγμάτων

- Αρνητικά Δείγματα

Τα αρνητικά δείγματα λαμβάνονται από οποιαδήποτε εικόνα που δεν περιέχει πρόσωπα (Εικ. 2.3).

Οι εικόνες δείγματα των μη-προσώπων λαμβάνονται στο ίδιο μέγεθος με τις εικόνες των προσώπων.

Εφαρμόζεται η τεχνική Bootstrapping [SUNG98]

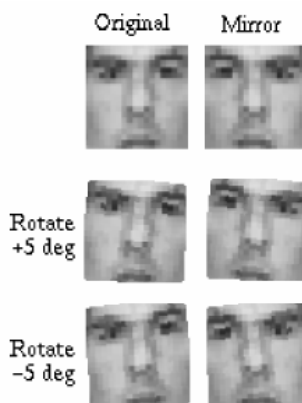


Εικ. 2.3: Μη-πρόσωπα που χρησιμοποιούνται στο σύνολο των αρνητικών δειγμάτων [ROWL98b]

Κατά τη διάρκεια της εκπαίδευσης, το μερικώς εκπαιδευμένο σύστημα εφαρμόζεται σε εικόνες οι οποίες δεν περιέχουν πρόσωπα (όπως στην Εικ. 2.3 αριστερά). Κάθε περιοχή στην εικόνα η οποία ανιχνεύεται εσφαλμένα ως πρόσωπο (τα οποία παρατίθενται στην Εικ. 2.3 δεξιά) είναι δείγματα μη-προσώπων, τα οποία προστίθενται στο σύνολο των αρνητικών δειγμάτων εκπαίδευσης [ROWL98b].

- Εικονικά Θετικά Δείγματα

Για τον πολλαπλασιασμό του συνόλου των δειγμάτων εκπαίδευσης, δημιουργούνται με τεχνητό τρόπο νέες εικόνες-δείγματα από τις υπάρχουσες.



Εικ. 2.4: Τεχνητώς δημιουργούμενα εικονικά δείγματα προσώπων [SUNG98]

Έτσι με κατοπτρισμό, μικρές περιστροφές, μετατοπίσεις και κλιμακώσεις των δειγμάτων των προσώπων (Εικ. 2.4), μπορούμε να πετύχουμε αύξηση του αριθμού των δειγμάτων εκπαίδευσης και μικρότερη ευαισθησία στα σφάλματα λήψης των δειγμάτων [ROWL98b, SUNG98].

2.5.2 Χρησιμοποιούμενοι Ταξινομητές

Διάφοροι ταξινομητές/αλγόριθμοι έχουν χρησιμοποιηθεί κατά καιρούς σε εργασίες ανίχνευσης προσώπων, οι κυριότεροι από τους οποίους είναι:

- Αλγόριθμοι Προσαρμοστικής Ενίσχυσης (Adaptive Boosting – AdaBoost)
- Νευρωνικά Δίκτυα - Πολυστρωματικό Αντίληπτρο (Neural Networks - Multilayer Perceptron)
- Ανάλυση Θεμελειωδών Συνιστωσών (Principal Component Analysis - PCA),
- Μηχανές Ανυσματικής Στήριξης (Support Vector Machines - SVM)
- Μέθοδοι βασισμένες σε Κατανομές (Distribution-based methods)
- Απλοϊκός Ταξινομητής Bayes (Naïve Bayes classifier)
- Μοντέλα Κρυφών Μεταβλητών Markov (Hidden Markov Models)
- Επαγωγική Μάθηση (Inductive Learning)

2.5.3 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 επικαλυμένων ή μη επικαλυμένων περιοχών (μπλοκ) όπως στην Εικ. 2.5.
 - Απεικόνιση σε πολλαπλές κλίμακες

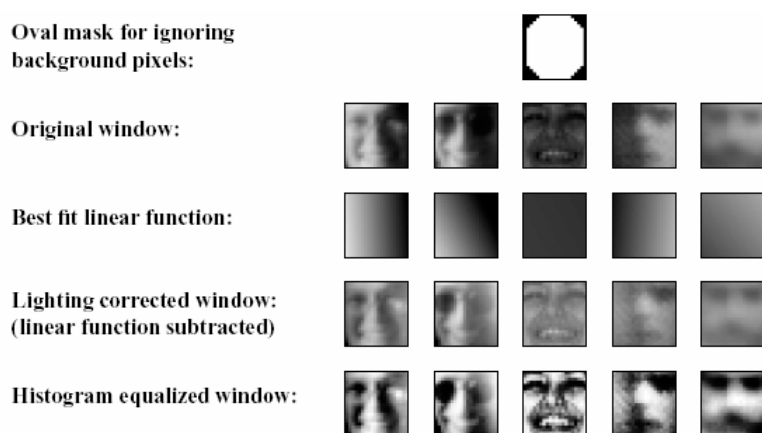
- Επιπλέον επεξεργασία με Διανυσματική Κβάντιση (Vector Quantization), Ανάλυση Θεμελειωδών Συνιστωσών (Principal Component Analysis), Κυματίδια (Wavelets) κλπ.
- Χρήση χαρακτηριστικών γνωρισμάτων (features).



Εικ. 2.5: Διάφορες αναπαραστάσεις στηριγμένες στον τεμαχισμό σε μπλοκ

2.5.4 Προεπεξεργασία

Η προεπεξεργασία με κανονικοποίηση της έντασης των εικονοστοιχείων (pixels) βοηθάει στην διόρθωση των αποκλίσεων των παραμέτρων αποτύπωσης των εικόνων μέσω των καμερών, όπως επίσης και των παραλλαγών στις συνθήκες φωτισμού.



Εικ 2.6: Διαδικασίες προ-επεξεργασίας εικόνας προσώπου [ROWL98b]

Τα βήματα κατά την διαδικασία προ-επεξεργασίας στο υπο-παράθυρο της εικόνας και κατά την διαμόρφωση των δειγμάτων των προσώπων [ROWL98b, SUNG98]:

- Τροποποίηση μεγέθους εικόνας: Τροποποίηση όλων των προτύπων προσώπων σε ένα συγκεκριμένο μέγεθος (π.χ. σε 20×20 pixels).
- Μασκάρισμα: περιορισμός των μη επιθυμητών δομών του υποβάθρου (θόρυβος) κοντά στα όρια του πλαισίου ενός προτύπου προσώπου

- Επιδιόρθωση διαβαθμισμένου φωτισμού: ανεύρεση μιας γραμμικής συνάρτησης η οποία να ταιριάζει στην διαβάθμιση των τιμών έντασης του παραθύρου του προσώπου και της οποίας οι τιμές αφού αφαιρεθούν, επιδιορθώνουν ακραίες διαβαθμίσεις φωτισμού (σκιές που δημιουργούνται από ακραίες γωνίες φωτισμού)
- Εξισορρόπηση Ιστογράμματος: αντισταθμίζει τις επιδράσεις φωτισμού εξαιτίας διαφορετικών συνθηκών φωτισμού και διαφορετικών καμπύλων απόκρισης της κάθε κάμερας λήψης

2.5.5 Αναζήτηση στον Χώρο και την Κλίμακα

Το παράθυρο ανίχνευσης σαρώνει την υπό ανίχνευση εικόνα σε διαφορετικές κλίμακες και θέσεις:

- Σάρωση μιας εικόνας εισόδου με βήμα Δ -pixels (π.χ. 2-pixels) οριζόντια και κάθετα
- Υποδειγματοληψία της εικόνας εισόδου κατά έναν παράγοντα (π.χ. 1.2) και συνέχιση της αναζήτησης
- Συνέχιση της υποδειγματοληψίας της εικόνας εισόδου και αναζήτηση μέχρι η εικόνα που σαρώνεται να γίνει πολύ μικρή, ώστε να είναι αδύνατο να περιέχει πρόσωπο

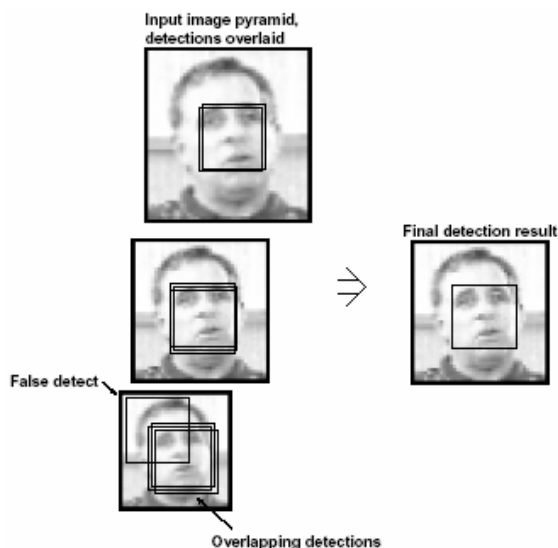
Η επιλογή του βήματος σάρωσης και του παράγοντα υποδειγματοληψίας επηρεάζει την ταχύτητα, όπως επίσης και την ακρίβεια του ανιχνευτή.

Αντί για σμίκρυνση της εικόνας εισόδου μπορούμε να έχουμε αντίστοιχη μεγέθυνση του παραθύρου σάρωσης [VIOL01a, VIOL01b].

2.5.6 Μετεπεξεργασία

Από τη στιγμή που ο ανιχνευτής δεν επηρεάζεται από μικρές μεταβολές στην μετατόπιση και την κλίμακα, πολλαπλές ανιχνεύσεις μπορεί να συμβούν σε κοντινές θέσεις ή σε πολλαπλές κλίμακες κατά την σάρωση μιας εικόνας. Εσφαλμένες ανιχνεύσεις μπορεί επίσης να συμβούν, αλλά συνήθως με λιγότερη συχνότητα από ότι

πολλαπλές ανιχνεύσεις προσώπων. Είναι ωστόσο 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. 2.7: Συγχώνευση πολλαπλών ανιχνεύσεων και απαλοιφή των λανθασμένων [ROWL98b]

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 [LI05, ROWL98b, VIOL01a, VIOL01b]:

- Ανέυρεση 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)

2.6 Ανίχνευση Προσώπου κατά Viola & Jones

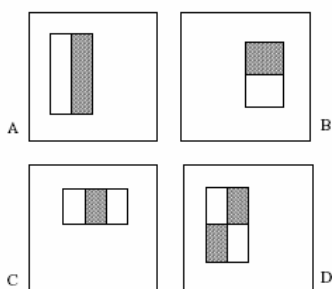
Οι ιδέες-κλειδιά της μεθόδου αυτής είναι οι εξής [VIOL01a, VIOL01b]:

- 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.

2.6.1 Υπολογισμός Χαρακτηριστικών

- Η διαδικασία ανίχνευσης ταξινομεί εικόνες στηριζόμενη στην τιμή απλών βαθμωτών χαρακτηριστικών (features).
- Η χρήση των χαρακτηριστικών αντί της εικόνας έντασης (intensity image) έχει ως σκόπο την μείωση των διαφοροποιήσεων μέσα στην κλάση (intra-class variability) και την αύξηση των διαφοροποιήσεων μεταξύ των κλάσεων (inter-class variability), ώστε η ταξινόμηση να καταστεί ευκολότερη.
- Τα χαρακτηριστικά αυτά περιέχουν γνώση από συγκεκριμένες περιοχές της εικόνας, που είναι δύσκολο να κωδικοποιηθεί χρησιμοποιώντας πεπερασμένα δεδομένα μάθησης.

- Τα χαρακτηριστικά είναι όμοια με τις συναρτήσεις βάσης τύπου Haar (Haar basis) που έχουν χρησιμοποιηθεί και από τον Papageorgiou κ.ά. για τον ίδιο σκοπό [PAP98].
- Εφαρμόζονται σε ασπρόμαυρες εικόνες 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, θεωρώντας τις γκριζες περιοχές θετικές και τις λευκές αρνητικές, όπως φαίνεται στην Εικ. 2.8.
- Τα χαρακτηριστικά καθορίζονται από την θέση, τις διαστάσεις και την τιμή τους.
- Το πλήθος των χαρακτηριστικών που δημιουργούνται για δείγματα προσώπων 24×24 pixels είναι ~ 45.000, που είναι ένα υπερπλήρες σύνολο σε σχέση με τις 576 τιμές έντασης του δείγματος. Για αυτόν τον λόγο απαιτείται μια διαδικασία επιλογής των κυριότερων χαρακτηριστικών από αυτά. Σύμφωνα με τους Viola & Jones ακόμα και ανιχνευτές με 2 χαρακτηριστικά είναι αρκετά αποτελεσματικοί.

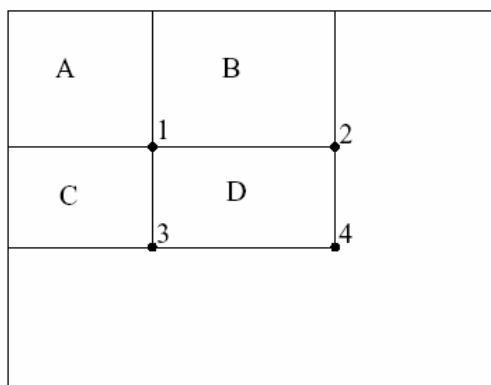


Εικ. 2.8: Μορφή των χαρακτηριστικών Viola & Jones [VIOL01b]

Τα ορθογώνια χαρακτηριστικά των Viola & Jones μπορούν να υπολογιστούν πολύ αποτελεσματικά με τη χρήση μιας βοηθητικής εικόνας, που αναφέρεται ως "εικόνα ολοκλήρωμα" (integral image). Η εικόνα ολοκλήρωμα, I , έχει τιμή στη θέση (x, y) που καθορίζεται ως άθροισμα των εντάσεων των pixels του ορθογωνίου που ορίζεται από την πάνω αριστερή κορυφή $(0, 0)$ και την κάτω δεξιά κορυφή (x, y) :

$$I(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y')$$

όπου i είναι η αρχική εικόνα εισόδου.



Εικ. 2.9: Αναπαράσταση της "Εικόνας Ολοκλήρωμα" [VIOL01a]

Χρησιμοποιώντας την εικόνα ολοκλήρωμα κάθε ορθογώνιο άθροισμα μπορεί να υπολογιστεί σε σταθερό χρόνο με τέσσερις αναφορές στις τιμές ενός πίνακα. Έτσι το άθροισμα εντός του D (Εικ. 2.9) μπορεί να υπολογιστεί σαν $I_4 + I_1 - (I_2 + I_3)$ [VIOL01b].

2.6.2 Επιλογή Χαρακτηριστικών με την 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 τυχειότητα. Ο πιο δημοφιλής αλγόριθμος ενίσχυσης είναι ο αλγόριθμος *προσαρμοστικής ενίσχυσης* (adaptive boosting) AdaBoost, ο οποίος ονομάζεται και *διακριτός* (discrete) AdaBoost, μιάς και αποδίδει διακριτές τιμές εξόδου [THEOD06].

Όπως ήδη έχει αναφερθεί, 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 [FRE95] 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 [VIOL01a, VIOL01b].

Έτσι, η μέθοδος AdaBoost στοχεύει στην επίλυση των παρακάτω 3 θεμελιωδών προβλημάτων [LI05]:

- (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 δημιουργηθέντα χαρακτηριστικά, και
- (3) συνδυασμός των αδύναμων ταξινομητών για την κατασκευή ενός ισχυρού ταξινομητή.

Algorithm: AdaBoost Algorithm for Learning a Strong Classifier [VIOL01b]

Input

Given example images $(x_1, y_1), \dots, (x_n, y_n)$ where $y_i = 0, 1$ for negative and positive examples respectively.

Initialization

Initialize weights $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$ for $y_i = 0, 1$ respectively, where m and l are the number of negatives and positives respectively.

Inclusion

For $t = 1, \dots, T$:

1. Normalize the weights,

$$w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}},$$

so that w_t is a probability distribution.

2. For each feature j , train a classifier h_j , which is restricted to using a single feature.

The error is evaluated with respect to w_t , $\epsilon_j = \sum_i w_i |h_j(x_i) - y_i|$.

3. Choose the classifier h_t , with the lowest error ϵ_t .
4. Update the weights:

$$w_{t+1,i} = w_{t,i} \beta_t^{1-e_i},$$

where $e_i = 0$ if example x_i is classified correctly, otherwise $e_i = 1$, i.e.:

$$w_{t+1,i} = w_{t,i} \beta_t \quad // \text{ if } x_i \text{ classified correctly}$$

$$w_{t+1,i} = w_{t,i} \quad // \text{ if } x_i \text{ classified incorrectly}$$

$$\text{where } \beta_t = \frac{\epsilon_t}{1 - \epsilon_t}.$$

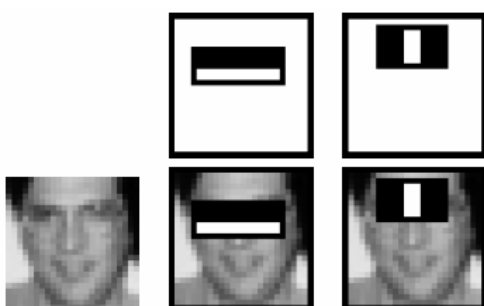
Output

The final strong classifier is:

$$H(x) = \begin{cases} 1, & \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0, & \text{otherwise} \end{cases}, \text{ where } \alpha_t = \ln \frac{1}{\beta_t}$$

Σύμφωνα με τους Viola & Jones [VIOL01a, VIOL01b] κατά την εφαρμογή του AdaBoost:

- Σε κάθε πέρασμα t εκπαιδεύεται ένας νέος αδύναμος ταξινομητής h_t που προστίθεται στο σύνολο με μεγαλύτερο συντελεστή α_t όσο μικρότερο σφάλμα ταξινόμησης ϵ δίνει.
- Το βάρος κάθε δείγματος 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.
- Όπως αποδεικνύεται από τους Freund & Schapire [FRE95] 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.

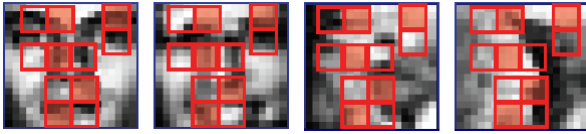


Εικ. 2.10: Τα 2 κυριότερα χαρακτηριστικά εφαρμοσμένα σε ένα τυπικό πρόσωπο [VIOL01b]

2.6.3 Κατασκευή του Αδύναμου Ταξινομητή

Ο απλούστερος τύπος ενός αδύναμου ταξινομητή είναι η "ρίζα" ("stump") ενός δένδρου απόφασης (decision tree). Όταν το χαρακτηριστικό παίρνει πραγματικές τιμές, μπορεί να κατασκευαστεί μία ρίζα απόφασης, συγκρίνοντας απλά την τιμή του επιλεγμένου χαρακτηριστικού με μια συγκεκριμένη τιμή κατωφλίου. Έτσι ο αδύναμος αλγόριθμος μάθησης σχεδιάζεται ώστε να μπορεί να επιλέγει εκείνο το μοναδικό χαρακτηριστικό που διαχωρίζει καλύτερα τα θετικά από τα αρνητικά δείγματα. Για κάθε χαρακτηριστικό, ο αδύναμος ταξινομητής καθορίζει το ιδανικό κατώφλι λειτουργίας της

ταξινόμησης, έτσι ώστε να ελαχιστοποιείται ο αριθμός δειγμάτων που ταξινομείται εσφαλμένα [LI05, VIOL01a, VIOL01b].



Εικ. 2.11: Υπολογισμός της τιμής χαρακτηριστικών πάνω σε πρόσωπα και μη-πρόσωπα

Έτσι ο αδύναμος ταξινομητής $h_j(x)$ αποτελείται από ένα χαρακτηριστικό j και ένα κατώφλι θ_j :

- Για κάθε χαρακτηριστικό j , υπολογίζεται η $f_j(x)$, μια βαθμωτή τιμή του χαρακτηριστικού (εδώ οι διαφορές αθροισμάτων), όπου x είναι ένα θετικό ή αρνητικό δείγμα
- Κάθε χαρακτηριστικό χρησιμοποιείται σαν ένας αδύναμος ταξινομητής
- Καθορισμός τιμής κατωφλίου θ_j για κάθε χαρακτηριστικό έτσι ώστε τα περισσότερα δείγματα να ταξινομούνται σωστά:

$$h_j(x) = \begin{cases} 1, & \text{if } f_j(x) < \theta_j \text{ (or } f_j(x) > \theta_j), \text{ } x : \text{positive} \\ 0, & \text{otherwise, } x : \text{negative} \end{cases}$$

- Επιλογή χαρακτηριστικού και κατωφλίου με το χαμηλότερο σταθμισμένο σφάλμα ταξινόμησης
- Διαδοχική εκτίμηση όλων των χαρακτηριστικών

2.6.4 Ταξινόμηση με έναν Καταρράκτη Ταξινομητών

Για να αποφύγουν την σύγκλιση του AdaBoost που απαιτεί μεγάλο αριθμό δειγμάτων και καταλήγει σε μεγάλο πλήθος αδύναμων ταξινομητών που απαιτούν μεγάλο υπολογιστικό κόστος, οι Viola & Jones εισήγαγαν την έννοια του *καταρράκτη ταξινομητών* (cascade of classifiers) [VIOL01a, VIOL01b].

Με τη μέθοδο αυτή κατασκευάζεται μία διαδοχή ταξινομητών στη μορφή καταρράκτη (cascade) που επιτυγχάνει αυξημένη απόδοση στην ανίχνευση και μειώνει

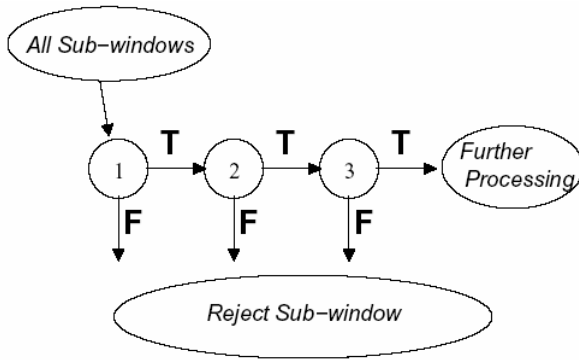
ριζικά τον χρόνο υπολογισμού. Η ιδέα είναι ότι μπορούν να κατασκευαστούν μικροί και ωστόσο αποτελεσματικοί, συνδυασμένοι ταξινομητές που απορρίπτουν πολλά από τα αρνητικά, ενώ ανιχνεύουν σχεδόν όλα τα θετικά περιστατικά.

Η μέθοδος του καταρράκτη ταξινομητών στηρίζεται στο γεγονός ότι σε μία οποιαδήποτε εικόνα η πλειονότητα των παραθύρων ανίχνευσης δεν περιλαμβάνει πρόσωπα. Έτσι πιο απλοποιημένοι και λιγότερο χρονοβόροι ταξινομητές χρησιμοποιούνται για να απορρίψουν την πλειονότητα των παραθύρων ανίχνευσης ως αρνητικά, προτού χρησιμοποιηθούν οι πιο σύνθετοι και περισσότερο χρονοβόροι ταξινομητές που θα επεξεργαστούν τις πιο πολύπλοκες περιπτώσεις και θα επιτύχουν χαμηλά επίπεδα εσφαλμένων θετικών ανιχνεύσεων.

Παράδειγμα: Καταρράκτης Ταξινομητής 32 επιπέδων [VIOL01b]

- 2-χαρακτηριστικών ταξινομητής στο πρώτο επίπεδο
απορρίπτει το 60% των μη-προσώπων ενώ ανιχνεύει 100% τα πρόσωπα
- 5-χαρακτηριστικών ταξινομητής στο δεύτερο επίπεδο
απορρίπτει το 80% των μη-προσώπων ενώ ανιχνεύει 100% τα πρόσωπα
- 20-χαρακτηριστικών ταξινομητής στα επίπεδα 3,4 και 5
- 50-χαρακτηριστικών ταξινομητής στα επίπεδα 6 και 7
- 100-χαρακτηριστικών ταξινομητής στα επίπεδα 8 έως και 12
- 200-χαρακτηριστικών ταξινομητής στα επίπεδα 13 έως και 32

Η συνολική διαδικασία ανίχνευσης είναι παρόμοια με ένα δένδρο απόφασης (decision tree). Ένα θετικό αποτέλεσμα από τον ταξινομητή πρώτου επιπέδου οδηγείται στον ταξινομητή δεύτερου επιπέδου, του οποίου το θετικό αποτέλεσμα οδηγείται στον ταξινομητή τρίτου επιπέδου κ.ο.κ. όπως στην Εικ. 2.12. Τα αρνητικά αποτελέσματα σε κάθε επίπεδο απορρίπτονται χωρίς να επανελέγονται. Έτσι οι ταξινομητές των αρχικών επιπέδων ασχολούνται με τα εύκολα περιστατικά, ενώ οι επόμενοι αντιμετωπίζουν πιο δύσκολες περιπτώσεις



Εικ. 2.12: Σχηματική παράσταση μιας ανίχνευσης με καταρράκτη ταξινομητών [VIOL01b]

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} .

Algorithm: Training Algorithm for Building a Cascade Detector [VIOL01b]

Given:

P : a training set of Positive examples, N : a training set of Negative examples.

f : maximum acceptable false positive rate per layer .

d : minimum acceptable detection rate per layer.

F_{target} : target overall false positive rate.

Initialize:

$F_0 = 1.0$; $D_0 = 1.0$

$i=0$

Iterate:

While $F_i > F_{\text{target}}$ do:

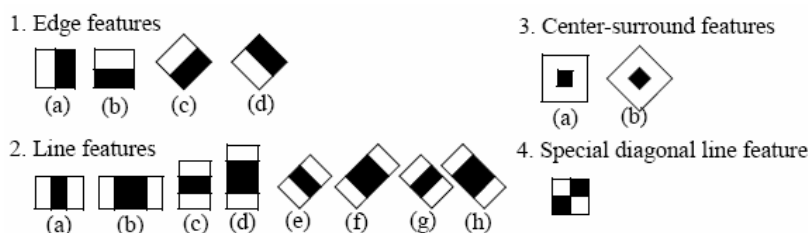
- $i \leftarrow i + 1$
- $n_i = 0$; $F_i = F_{i-1}$
- While $F_i > f \times F_{i-1}$ do:
 - $n_i \leftarrow n_i + 1$
 - Use P and N to train a classifier with n_i features using AdaBoost
 - Evaluate current cascaded classifier on validation test to determine F_i and D_i
 - Decrease threshold for the i th classifier until the current cascaded classifier has a detection rate of at least $d \times D_{i-1}$ (this also affects F_i)
- $N \leftarrow 0$
- If $F_i > F_{\text{target}}$ then evaluate the current cascaded detector on the set of non-face images and put any false detection into the set N

2.7 Ανίχνευση Προσώπου κατά Lienhart

Ο Lienhart κ.ά [LIEN02a, LIEN02b] εισήγαγαν δύο κύριες βελτιώσεις στην μέθοδο ανίχνευσης των Viola & Jones [VIOL01a, VIOL01b], κατά πρώτον ένα εκτεταμένο σύνολο από ορθογώνια χαρακτηριστικά τύπου Haar, και κατά δεύτερον την χρήση της gentle AdaBoost για την κατασκευή του ισχυρού ταξινομητή σε συνδυασμό με τη χρήση CART δένδρων στη θέση των αδύναμων ταξινομητών.

- Το εκτεταμένο σύνολο από Haar χαρακτηριστικά:

Ο Lienhart εμπλούτισε το βασικό σύνολο των Haar χαρακτηριστικών με ένα πιο αποτελεσματικό σύνολο από περιστραμμένα κατά 45° χαρακτηριστικά, προσθέτοντας επιπλέον γνώση στο σύστημα εκπαίδευσης και βελτιώνοντας την απόδοση του συστήματος [LIEN02a, LIEN02b].



Εικ. 2.13: Εκτεταμένο σύνολο στραμμένων Haar χαρακτηριστικών [LIEN02a]

Τα χαρακτηριστικά αυτά φαίνονται στην Εικ. 2.13 και αποτελούνται από τέσσερα χαρακτηριστικά ακμών (Εικ. 2.13.1), οκτώ χαρακτηριστικά γραμμών (Εικ. 2.13.2) και δύο χαρακτηριστικά κέντρων (Εικ. 2.13.3), ενώ το χαρακτηριστικό διαγώνιας γραμμής (Εικ. 2.13.4) δεν χρησιμοποιείται σε αντίθεση με τους Viola & Jones.

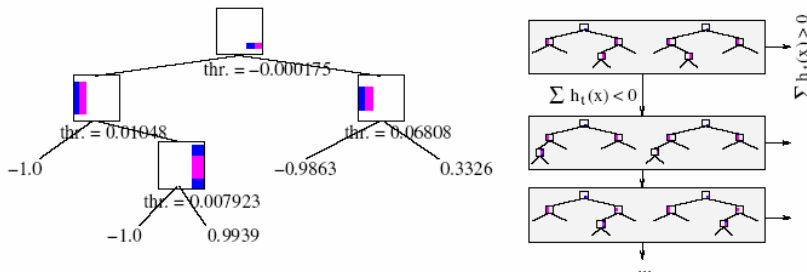
Τα χαρακτηριστικά καθορίζονται από τα ορθογώνια $r = (x, y, w, h, \alpha)$, $\alpha = \{0^\circ, 45^\circ\}$, και παίρνουν τιμές την σταθμισμένη διαφορά των αθροισμάτων των εντάσεων στις λευκές και μαύρες περιοχές:

$$\text{feat}_i = \sum \omega_i \cdot \text{RecSum}(r_i)$$

Το πλήθος των ορθογώνιων χαρακτηριστικών του Lienhart για δείγματα 24×24 pixels ανέρχεται σε ~ 118.000.

- CART δένδρα ως αδύναμοι ταξινομητές:

Επίσης ο Lienhart χρησιμοποίησε ως αδύναμους ταξινομητές αντί για απλές ρίζες απόφασης (stumps), μικρά CART δένδρα με έως 4 χαρακτηριστικά. Τα *Δένδρα Ταξινόμησης και Παλινδρόμησης* (Classification and Regression Trees – CART) είναι μικροί ταξινομητές δένδρων απόφασης, καλοί στην εκμάθηση εξαρτήσεων μεταξύ των χαρακτηριστικών. Η συνάρτηση $h_t(x)$ καθορίζεται σύμφωνα με την διαδρομή μέσα στο δένδρο απόφασης.



Εικ. 2.14: Ένα δένδρο ταξινόμησης CART και ένας καταρράκτης από CART

Η χρήση μικρών δένδρων απόφασης εμφανίζει καλύτερη απόδοση από τις απλές ρίζες απόφασης.

- Χρήση παραλλαγών της AdaBoost:

Από τον Lienhart μελετήθηκαν τρεις παραλλαγές της AdaBoost: η Discrete AdaBoost, που χρησιμοποιήθηκε κι από τους Viola & Jones, και επι πλέον, η Real AdaBoost και η Gentle AdaBoost. Απ'αυτές η Gentle AdaBoost εμφάνισε τα καλύτερα αποτελέσματα και σε μικρότερους χρόνους.

Σε κάθε γύρο της ενίσχυσης (boosting) προστίθεται ο ταξινομητής που ταξινομεί ορθότερα τα δείγματα εκπαίδευσης με τη χρήση του αποτελεσματικότερου CART δένδρου. Ο κάθε ισχυρός ταξινομητής F είναι ένα σταθμισμένο άθροισμα αδύναμων ταξινομητών f_i από CART δένδρα ενός ή περισσότερων κόμβων:

$$F = \text{sign}(c_1 f_1 + c_2 f_2 + \dots + c_n f_n).$$

Τέλος, οι ισχυροί ταξινομητές συνδιάζονται σε διάταξη καταρράκτη (cascade). Καθώς αυξάνεται το επίπεδο του καταρράκτη ανιχνευτή, ο αριθμός των αδύναμων ταξινομητών που αποτελούν το ισχυρό ταξινομητή αυξάνεται, ώστε να επιτυγχάνεται ένας συγκεκριμένος ρυθμός εσφαλμένων ανιχνεύσεων για δεδομένο αριθμό επιτυχών ανιχνεύσεων [LIEN02a, LIEN02b].

2.8 Σύνοψη του Συστήματος Ανίχνευσης Προσώπου

- Υπολογίζονται απλά βαθμωτά χαρακτηριστικά. Υπάρχει ένας μεγάλος αριθμός από υποψήφια χαρακτηριστικά τύπου Haar.
- Επιλέγεται ένα μικρό υποσύνολο από αυτά και οι αντίστοιχοι αδύναμοι ταξινομητές από δένδρα απόφασης εκπαιδεύονται χρησιμοποιώντας AdaBoost.
- Ο ισχυρός ταξινομητής κατασκευάζεται σαν ένας γραμμικός συνδυασμός των αδύναμων και είναι το αποτέλεσμα του αλγόριθμου AdaBoost .
- Ο τελικός ανιχνευτής δημιουργείται από έναν ή από μία διαδοχή, σε διάταξη καταρράκτη, ισχυρών ταξινομητών που χρησιμοποιούν διαρκώς περισσότερα χαρακτηριστικά.
- Η προεπεξεργασία προετοιμάζει την εικόνα που θα ανιχνευτεί.
- Η μετεπεξεργασία συγχωνεύει ή απορρίπτει τις ανιχνεύσεις.

2.9 Μέθοδος Haar+AdaBoost : Συμπεράσματα

Αναφορικά με την προσέγγιση Haar+AdaBoost [LIEN02a, LIEN02b, VIOL01a, VIOL01b], μπορούμε να εξάγουμε τα παρακάτω συμπεράσματα [LI05] 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. [ROWL98b], 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

- Ανάλυση πολύ χρόνου στη φάση της εκπαίδευσης (μπορεί να απαιτεί μέρες εκπαίδευσης)

- Περιορισμένη προσέγγιση των περιπτώσεων διαφορετικού προσανατολισμού των προσώπων

- Απαιτεί αρκετή εργασία υλοποίησης.

Κεφάλαιο 3

ΥΛΟΠΟΙΗΣΗ ΣΥΣΤΗΜΑΤΟΣ ΑΝΙΧΝΕΥΣΗΣ ΠΡΟΣΩΠΟΥ

3.1 Η Βιβλιοθήκη OpenCV της Intel

Η βιβλιοθήκη "Open Source Computer Vision Library" (OpenCV) [OCVL] είναι μία ελεύθερα διαθέσιμη, 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.

3.2 Ανίχνευση Προσώπων με την OpenCV

Η OpenCV [BKP05] παρέχει χαμηλού και υψηλού επιπέδου API για την Ανίχνευση Προσώπων ή άλλων Αντικειμένων, ανάλογα με την εκπαίδευση που έχει προηγηθεί στον ταξινομητή αντικειμένων που προσφέρει.

Ο ταξινομητής αντικειμένων που προσφέρει είναι αυτός που έχει προταθεί από τους Viola & Jones [VIOL01a, VIOL01b] και έχει βελτιωθεί από τον Lienhart κ.ά. [LIEN02a, LIEN02b]. Ο ταξινομητής για τα πρόσωπα έχει εκπαιδευτεί με εκατοντάδες θετικά δείγματα προσώπων μεγέθους 20x20 και αυθαίρετα αρνητικά δείγματα ίδιου μεγέθους.

Η OpenCV περιέχει την εφαρμογή `Haartraining` για την εκπαίδευση ενός ταξινομητή με ένα δεδομένο σύνολο θετικών και αρνητικών δειγμάτων, που αφορούν ένα συγκεκριμένο αντικείμενο. Η εφαρμογή αυτή δημιουργεί ένα `xml` αρχείο που

περιέχει τα δεδομένα που αφορούν τον εν λόγω ταξινομητή και μπορεί να χρησιμοποιηθεί στη συνέχεια για την ανίχνευση του συγκεκριμένου αντικειμένου.

Αφού εκπαιδευτεί ο ταξινομητής, μπορεί να εφαρμοστεί σε μία περιοχή ενδιαφέροντος μιάς εικόνας εισόδου. Ο ταξινομητής αποκρίνεται με "1" αν η περιοχή περιέχει το αναζητούμενο αντικείμενο και με "0" σε αντίθετη περίπτωση. Για να αναζητηθεί το αντικείμενο σε ολόκληρη την εικόνα, το παράθυρο αναζήτησης μπορεί να μετακινηθεί σε διαφορετικές θέσεις και να ελεγχθεί η απόκριση του ταξινομητή. Ο ταξινομητής έχει σχεδιαστεί έτσι ώστε το παράθυρο αναζήτησης να μπορεί να τροποποιηθεί ως προς τις διαστάσεις του εύκολα, για να ανευρεθούν αντικείμενα διαφορετικών διαστάσεων, αντί να απαιτείται αντίστοιχη μεταβολή των διαστάσεων της ίδιας της εικόνας. Έτσι, για να βρεθεί ένα αντικείμενο άγνωστων διαστάσεων σε μία εικόνα, η διαδικασία αναζήτησης μπορεί να γίνει όσες φορές χρειάζεται σε διαφορετικές κλίμακες.

Ο τελικός ταξινομητής αποτελείται από επίπεδα απλούστερων ταξινομητών σε διάταξη καταρράκτη (cascade). Κάθε επίπεδο του καταρράκτη αποτελείται από ένα ισχυρό ταξινομητή (strong classifier) που κατασκευάζεται από απλούστερους αδύναμους ταξινομητές (weak classifiers) με την τεχνική της ενίσχυσης (boosting). Οι τεχνικές ενίσχυσης που υποστηρίζονται είναι η Discrete AdaBoost, η Real AdaBoost, η Gentle AdaBoost και η LogitBoost. Οι αδύναμοι ταξινομητές που χρησιμοποιούνται είναι δένδρα απόφασης (decision trees) με 1 έως 4 κόμβους. Οι αδύναμοι ταξινομητές παίρνουν ως είσοδο ορθογώνια χαρακτηριστικά (features) τύπου Haar, που καθορίζονται από την θέση, τις διαστάσεις, τον προσανατολισμό και την μορφή τους και παίρνουν τιμές μία σταθμισμένη διαφορά των εντάσεων των εικονοστοιχείων στις περιοχές που ορίζουν. Τα αθροίσματα των εντάσεων υπολογίζονται με τη χρήση της εικόνας ολοκλήρωμα (integral image) που έχει προϋπολογιστεί.

Έτσι μια χαμηλού επιπέδου ρουτίνα, για παράδειγμα, περιέχει τη δυνατότητα στον χρήστη να ελέγξει μία συγκεκριμένη περιοχή της εικόνας για το αν περιέχει ή όχι πρόσωπο. Βοηθητικές ρουτίνες υπολογίζουν τις εικόνες ολοκληρώματα, και ρυθμίζουν την κλίμακα αναζήτησης του ταξινομητή για πρόσωπα διαφορετικών μεγεθών κλπ.

Αντίθετα, μία υψηλού επιπέδου συνάρτηση, όπως η `cvHaarDetectObjects` κάνει όλα τα προηγούμενα, χρησιμοποιώντας τον ταξινομητή που είναι αποθηκευμένος σε `xml` αρχείο, παρέχοντας αυτόματα τα αποτελέσματα ανίχνευσης στο χρήστη, πράγμα που είναι τις περισσότερες φορές αρκετό.

3.2.1 Η Εκπαίδευση του Ταξινομητή

Η κατασκευή του ταξινομητή και η εκπαίδευσή του γίνεται με την εφαρμογή Haartraining. Η κλήση της εφαρμογής γίνεται ως εξής:

```
Haartraining
-data eyes_classifier
-vec eyes.vec -w 20 -h 15
-bg backgrounds.idx
-nstages 15
-nsplits 1
-minhitrate 0.995
-maxfalsealarm 0.5
```

Στο παραπάνω παράδειγμα ο ταξινομητής θα αποθηκευτεί στο αρχείο `eyes_classifier.xml`. Τα θετικά δείγματα εκπαίδευσης θα ληφθούν από το αρχείο `eyes.vec` και θα έχουν διαστάσεις 20x15. Τα αρνητικά δείγματα θα ληφθούν από το αρχείο `backgrounds.idx`. Ο καταρράκτης ταξινομητής θα αποτελείται από `nstages` επίπεδα. Κάθε επίπεδο εκπαιδεύεται ώστε ο ταξινομητής που περιέχει να εμφανίζει ρυθμό θετικών ανιχνεύσεων τουλάχιστον `minhitrate` και ρυθμό εσφαλμένων ανιχνεύσεων το πολύ `maxfalsealarm`. Κάθε αδύναμος ταξινομητής θα έχει ακριβώς `nsplits` κόμβους (το 1 αντιστοιχεί σε ρίζα δένδρου απόφασης).

3.2.2 Η Συνάρτηση Ανίχνευσης

Η συνάρτηση `cvHaarDetectObjects` βρίσκει ορθογώνιες περιοχές σε μια δεδομένη εικόνα που είναι πιθανό να περιέχουν αντικείμενα που ο ταξινομητής έχει εκπαιδευτεί να διακρίνει και επιστρέφει αυτές τις περιοχές σαν μία σειρά από ορθογώνια. Η συνάρτηση σαρώνει την εικόνα αρκετές φορές σε διαφορετικές κλίμακες, και κάθε φορά, για κάθε περιοχή, εφαρμόζει τον ταξινομητή για να ελέγξει την ύπαρξη του προς ανίχνευση αντικειμένου.

Αφού προχωρήσει η διαδικασία και συλλεγούν τα υποψήφια ορθογώνια (περιοχές που επαληθεύονται από τον ταξινομητή), ομαδοποιούνται σε ομάδες επικαλυπτόμενων ορθογωνίων και επιστρέφονται σαν μία σειρά από μέσα ορθογώνια πλαίσια, ένα για κάθε ομάδα.

Οι προκαθορισμένες παράμετροι (`scale_factor=1.1`, `min_neighbors=3`, `flags=0`) είναι ρυθμισμένες για μεγαλύτερης ακρίβειας αν και χρονοβόρα ανίχνευση.

Για ταχύτερη λειτουργία σε εικόνες βίντεο οι ρυθμίσεις είναι: `scale_factor=1.2`, `min_neighbors=2`, `flags=CV_HAAR_DO_CANNY_PRUNING`, `min_size=<minimum possible face size>`.

Η συνάρτηση `cvHaarDetectObjects` δέχεται τις εξής παραμέτρους:

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

`image`

Η εικόνα στην οποία ανιχνεύονται τα αντικείμενα.

`cascade`

Η αναφορά στον χρησιμοποιούμενο ταξινομητή.

`storage`

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

`scale_factor`

Ο συντελεστής κατά τον οποίο μεγθύνεται το παράθυρο ανίχνευσης μεταξύ των διαδοχικών σαρώσεων της εικόνας, π.χ., 1.1 σημαίνει αύξηση του παραθύρου κατά 10%.

`min_neighbors`

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

`flags`

Τρόπος λειτουργίας. Ο μόνος διαθέσιμος αυτή τη στιγμή τρόπος λειτουργίας είναι ο `CV_HAAR_DO_CANNY_PRUNING`. Σ' αυτή την περίπτωση, η συνάρτηση χρησιμοποιεί τη μέθοδο ανίχνευσης ακμών του Canny για να απορρίψει κάποιες περιοχές που περιέχουν πολύ λίγες ή πάρα πολλές ακμές σε σχέση με το αντικείμενο που ανιχνεύεται. Οι συγκεκριμένες τιμές έχουν ρυθμιστεί έτσι ώστε να επιταχύνεται η διαδικασία στην ανίχνευση προσώπων.

`min_size`

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

3.2.3 Οι Μέθοδοι Ανίχνευσης

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

- Stump-based 24x24 Discrete Adaboost with cascade

(haarcascade_frontalface_default.xml)

- Stump-based 20x20 Gentle Adaboost with cascade

(haarcascade_frontalface_alt.xml)

- CART-based (2-nodes) 20x20 Gentle Adaboost with cascade

(haarcascade_frontalface_alt2.xml)

- Stump-based 20x20 Gentle Adaboost with tree of stage classifiers

(haarcascade_frontalface_alt_tree.xml)

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

3.3 Πειραματικές Δοκιμές

Οι πειραματικές δοκιμές είχαν ως στόχο την αξιολόγηση των μεθόδων ανίχνευσης προσώπων που προσφέρονται στην OpenCV, και τον προσδιορισμό των κατάλληλων παραμέτρων για την υλοποίηση ενός όσον το δυνατόν πιο αξιόπιστου εντοπιστή προσώπων.

Για τις δοκιμές χρησιμοποιήθηκαν δύο βάσεις δεδομένων, οι οποίες μάς παρασχέθηκαν από το εργαστήριο CBL του Πανεπιστημίου του Houston.

Η 1^η Βάση Δεδομένων αποτελείται από 1578 έγχρωμες εικόνες χαμηλής ανάλυσης, που οι διαστάσεις τους ποικίλουν και κυμαίνονται από 620x650 έως 780x1030 pixels. Κάθε εικόνα απ'αυτές περιέχει ένα μόνο πρόσωπο που καλύπτει το 20%~30% της εικόνας. Τα άτομα αυτά έχουν φωτογραφηθεί σε διαφορετικές συνθήκες φωτισμού, σε διαφορετικά περιβάλλοντα με ή χωρίς εμπόδια και φοράνε γυαλιά ή έχουν μούσια. Κάθε άτομο εμφανίζεται σε 2~4 διαφορετικές στάσεις (πόζες) με ή χωρίς εκφράσεις.

Η 2^η Βάση Δεδομένων αποτελείται από 845 έγχρωμες εικόνες υψηλής ανάλυσης, που οι διαστάσεις τους είναι συγκεκριμένα 2888x2592 pixels. Κάθε εικόνα απ'αυτές περιέχει ένα μόνο πρόσωπο που καλύπτει το 10%~20% της εικόνας. Τα άτομα αυτά έχουν φωτογραφηθεί σε διαφορετικές συνθήκες φωτισμού, σε ελεγχόμενο περιβάλλον

εργαστηρίου χωρίς εμπόδια και φοράνε γυαλιά ή έχουν μούσια. Κάθε άτομο εμφανίζεται σε 3~6 διαφορετικές στάσεις (πόζες) με ή χωρίς εκφράσεις. Οι κύριες διαφορές των δύο βάσεων δεδομένων είναι ότι η πρώτη περιέχει εικόνες με γενικά χειρότερο φωτισμό (υποφωτισμένες), με πιο ακραίες στάσεις των προσώπων, όσον αφορά τις στροφές στο επίπεδο αλλά και εκτός επιπέδου εικόνας, με πιο έντονους ακραίους μορφασμούς και με περισσότερα εμπόδια.

3.3.1 Οι Αρχικές Δοκιμές

Για την αξιολόγηση του OpenCV ως εντοπιστή προσώπων (face locator) χρησιμοποιήθηκαν αρχικά 400 εικόνες από την 1^η Βάση Δεδομένων και μετρήθηκαν τα ποσοστά επιτυχούς ανίχνευσης, μη ανίχνευσης και εσφαλμένης ανίχνευσης.

Τα πρόσωπα που χρησιμοποιήθηκαν έχουν φωτογραφηθεί σε διαφορετικές συνθήκες φωτισμού, με ή χωρίς εκφράσεις, σε διάφορες πόζες και με ή χωρίς εμπόδια, κυρίως γυαλιά.

Στην δοκιμή χρησιμοποιήθηκαν δύο από τις μεθόδους ανίχνευσης, οι Alt (Stump-based Gentle AdaBoost) και Alt2 (CART-based Gentle AdaBoost), διότι με παρατηρήσεις σε μικρό δείγμα (40 εικόνων) παρατηρήθηκε ότι η Default (Stump-based Discrete AdaBoost) εμφανίζει πολλές εσφαλμένες ανιχνεύσεις (false detections) και πολλαπλές ανιχνεύσεις (multiple detections), ενώ η Tree (Stump-based Gentle AdaBoost with Tree of Classifiers) κάνει πολλαπλάσιο χρόνο για να ολοκληρώσει την ανίχνευση χωρίς να βελτιώνει την απόδοση.

Οι χρόνοι για 40 εικόνες είναι : $\frac{t_{alt}}{t_{tree}} = \frac{2 \text{ min}}{13 \text{ min}}$.

Οι εικόνες ταξινομήθηκαν και ως προς τις συνθήκες φωτογράφισης, σε φυσιολογικές (norm), κακού φωτισμού (light), ύπαρξης μορφασμών (expr), στροφών (rot) και εμποδίων (occl), ώστε να προσδιοριστεί η ικανότητα αντιμετώπισης από τον ανιχνευτή τέτοιων καταστάσεων.

Τα αποτελέσματα είναι τα εξής:

OpenCV Alt.xml	Gentle AdaBoost											
			norm		light		expr		rot		occl	
det	390	98%	176	100%	61	95%	118	96%	44	81%	34	94%
no det	10	3%	0	0%	3	5%	5	4%	10	19%	2	6%
all	400	100%	176		64		123		54		36	
false det	37	9%										

Πιν. 3.1: Αποτελέσματα ανίχνευσης προσώπου με τη μέθοδο Alt (Stump-based Gentle AdaBoost)

OpenCV Alt2.xml	CART+ Gentle AdaBoost											
			norm		light		expr		rot		occl	
det	389	97%	176	100%	58	91%	119	97%	45	83%	34	94%
no det	11	3%	0	0%	6	9%	4	3%	9	17%	2	6%
all	400	100%	176		64		123		54		36	
false det	65	16%										

Πιν. 3.2: Αποτελέσματα ανίχνευσης προσώπου με τη μέθοδο Alt2 (CART-based Gentle AdaBoost)

ΠΑΡΑΤΗΡΗΣΕΙΣ:

Οι εσφαλμένες ανιχνεύσεις (false detections) στην Alt (Stump-based Gentle AdaBoost) εμφανίστηκαν πάντα μαζί με θετικές/σωστές ανιχνεύσεις, δηλ., είχαμε πολλαπλές ανιχνεύσεις (multiple detections).

Το μεγαλύτερο ποσοστό των μη ανιχνεύσεων (19% στην Alt και 17% στην Alt2) έχουν τα πρόσωπα με στροφή (rot).

Η απόδοση σε θετικές ανιχνεύσεις των δύο μεθόδων είναι σχεδόν ίδια (97%), όμως η Alt2 (CART-based Gentle AdaBoost) παρουσιάζει σχεδόν διπλάσιες (16% αντί για 9% της Alt) εσφαλμένες ανιχνεύσεις (false detections) και μάλιστα πολλαπλές (multiple detections) στην ίδια εικόνα (εσφαλμένες ανιχνεύσεις μαζί με σωστή ή και διπλές επικαλυπτόμενες σωστές ανιχνεύσεις ή και μόνο εσφαλμένες χωρίς σωστή).

3.3.2 Περαιτέρω Ανάλυση

Για μία βαθύτερη αξιολόγηση της OpenCV χρησιμοποιήθηκε η μέθοδος Alt (Stump-based Gentle AdaBoost), που εμφανίζει καλύτερες επιδόσεις σύμφωνα με την αρχική μελέτη, και επιπλέον χρησιμοποιήθηκαν και οι 1578 διαθέσιμες εικόνες προσώπων της 1^{ης} Βάσης Δεδομένων.

Τα πρόσωπα που χρησιμοποιήθηκαν έχουν φωτογραφηθεί σε διαφορετικές συνθήκες φωτισμού, με ή χωρίς εκφράσεις, σε διάφορες πόζες και με ή χωρίς γυαλιά

και μετρήθηκαν τα ποσοστά επιτυχούς ανίχνευσης, μη ανίχνευσης και εσφαλμένης ανίχνευσης.

Τέλος έγινε μία καταμέτρηση των εσφαλμένων ανιχνεύσεων (false detections) σε συνδυασμό με το αν είχαμε συγχρόνως αποτυχία ή επιτυχία ανίχνευσης και μία καταμέτρηση των περιοχών που ο αλγόριθμος "μπερδεύει" ως πρόσωπα.

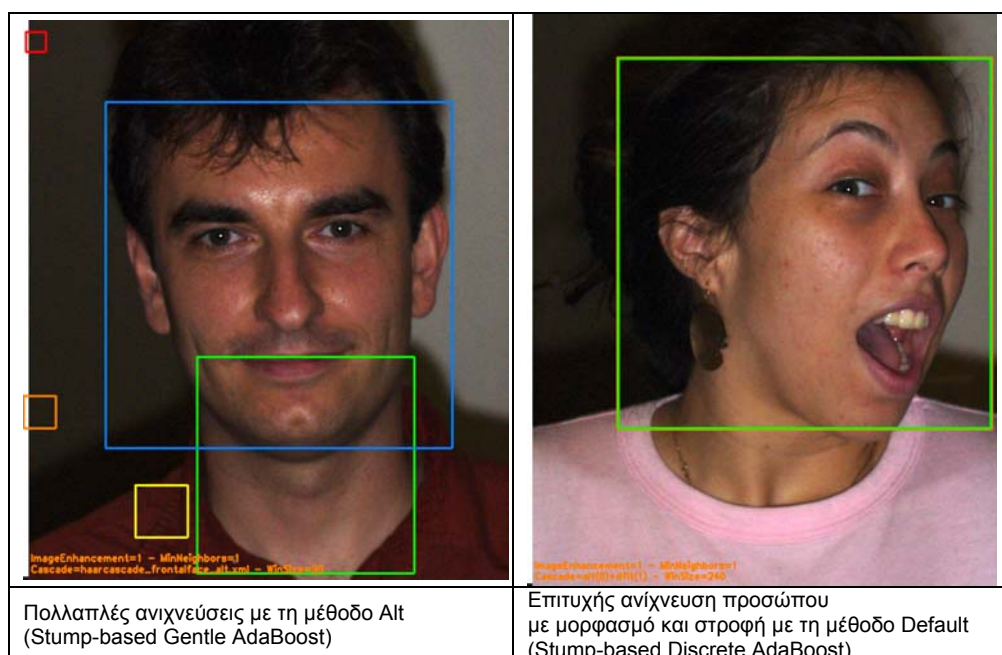
OpenCV Alt.xml	Gentle AdaBoost														
false detections															
		other		face		eye		throat		skin		cloth			
all	1578														
false det	122 8%	42 34%	4 3%	9 7%	16 13%	15 12%	39 32%								
no det	11 9%														
det	111 91%														

Πιν. 3.3: Αποτελέσματα ανίχνευσης προσώπου με τη μέθοδο Alt και κατανομή εσφαλμένων ανιχνεύσεων

Στα 1578 πρόσωπα είχαμε 122 εσφαλμένες ανιχνεύσεις (7.8%).

Στις 122 εσφαλμένες ανιχνεύσεις είχαμε 111 παράλληλες επιτυχείς ανιχνεύσεις (91%), ενώ οι 11 ήταν μόνο εσφαλμένες ανιχνεύσεις (9%).

Τέλος ο αλγόριθμος, εξέλαβε εσφαλμένα ως πρόσωπα, 9 μάτια (7%), 16 λαρύγγια (13%), 15 περιοχές δέρματος (12%) και 39 περιοχές υφασμάτων, ενώ είχαμε και 4 διπλές ανιχνεύσεις προσώπων (3%), στις 122 εσφαλμένες ανιχνεύσεις.



Εικ. 3.1: Παραδείγματα επιτυχούς ανίχνευσης και πολλαπλών ανιχνεύσεων

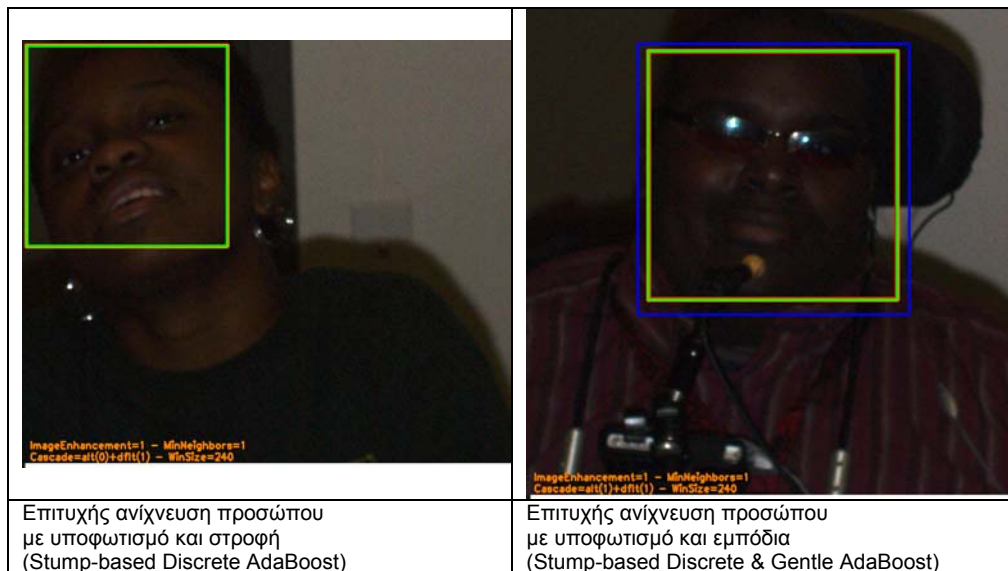
Η ρουτίνα που χρησιμοποιήθηκε για όλους τους παραπάνω ελέγχους είναι η:

```
faces = cvHaarDetectObjects( img, cascade, storage,
                            1.1, 2, CV_HAAR_DO_CANNY_PRUNING,
                            cvSize(30, 30) );
```

με τιμές παραμέτρων τις τιμές όπως αυτές φαίνονται στον παραπάνω τμήμα κώδικα.

3.3.3 Προεπεξεργασία με Εξισορρόπηση Ιστογράμματος

Μία από της αιτίες της μη ανίχνευσης ή της εσφαλμένης ανίχνευσης είναι και ο κακός φωτισμός των εικόνων της 1^{ης} Βάσης Δεδομένων.



Εικ. 3.2: Παραδείγματα επιτυχιών ανιχνεύσεων σε υποφωτισμένες εικόνες

Λόγω της ανάγκης βελτίωσης της απόδοσης του συστήματος, ενσωματώθηκε στον κώδικα προεπεξεργασία της εικόνας με Εξισορρόπηση Ιστογράμματος (Histogram Equalization), με χρήση της διαδικασίας `cvEqualizeHist`. Επειδή όμως η διαδικασία της ισοστάθμισης "αγριεύει" την εικόνα, εφαρμόστηκε στη συνέχεια εξομάλυνση με την εφαρμογή ενός Γκαουσιανού φίλτρου, με την βοήθεια της `cvSmooth` που προσφέρεται από την OpenCV.

Η τροποποίηση αυτή προσέφερε ουσιαστική βελτίωση στο σύστημα, παρέχοντας ένα αξιοσημείωτο θετικό αποτέλεσμα:

OpenCV Alt.xml	Gentle AdaBoost			
	without Equal.		after Equal.	
det	1535	97%	1557	99%
no det	43	3%	21	1%
all	1578		1578	

Πιν. 3.4: Αποτελέσματα ανίχνευσης προσώπου με και χωρίς Εξισορρόπηση Ιστογράμματος

Δηλ., ενώ χωρίς την εξισορρόπηση ιστογράμματος είχαμε 43 (2.7%) αποτυχίες ανίχνευσης, με την εξισορρόπηση ιστογράμματος ανιχνεύθηκαν σωστά τα 1557 πρόσωπα από τα 1578, φθάνοντας τις αποτυχίες ανίχνευσης μόλις στις 21 πρόσωπα (1.3%).

3.3.4 Κατάργηση της συγχώνευσης των ανιχνεύσεων

Στη συνέχεια θέσαμε `min_neighbors=1` για να ελεγχθεί η επίδραση της συγχώνευσης των πλαισίων που ανιχνεύονται ως πρόσωπα που είναι ενσωματωμένη στην ρουτίνα ανίχνευσης `cvHaarDetectObjects`. Καταργώντας λοιπόν την ενσωματωμένη συγχώνευση των ανιχνεύσεων, παρατηρήθηκε βελτίωση στις θετικές ανιχνεύσεις, με τελικό αποτέλεσμα αποτυχίας, δηλαδή καθόλου ανιχνεύσεις μόνο 15/1578 (1.0%).

Η κατάργηση βέβαια της συγχώνευσης είχε ως αποτέλεσμα την κατακόρυφη αύξηση των πολλαπλών εσφαλμένων ανιχνεύσεων στην ίδια εικόνα, φαινόμενο που αντιμετωπίστηκε στη συνέχεια με κατάλληλη μετεπεξεργασία και αύξηση του μεγέθους του παραθύρου ενδιαφέροντος.

3.3.5 Προσαρμογή του Παραθύρου Ενδιαφέροντος

Για να αντιμετωπισθεί το φαινόμενο των πολλαπλών ανιχνεύσεων και επειδή παρατηρήθηκε ότι οι πολλαπλές ανιχνεύσεις εμφανίζονται μαζί με σωστές ανιχνεύσεις και μάλιστα σε μικρότερα πλαίσια απ' αυτό του πραγματικού προσώπου που ανιχνεύεται, επιχειρήθηκε ανίχνευση με παράθυρο ενδιαφέροντος 90×90 pixels ώστε να αποκλειστούν αυτές οι δευτερεύουσες ανιχνεύσεις σε μικρότερες περιοχές υφής.

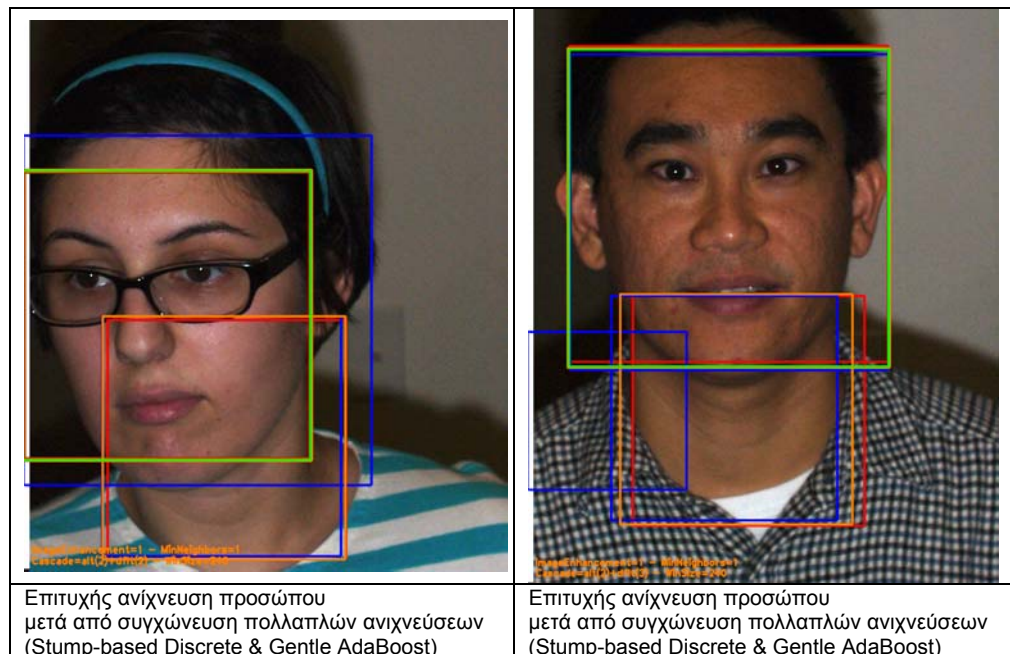
Οι δύο παραπάνω προσαρμογές, της κατάργησης της συγχώνευσης και της αύξησης του παραθύρου ανίχνευσης, επέφεραν δραστική μείωση στο ποσοστό των εσφαλμένων ανιχνεύσεων και στο ποσοστό των δευτερευόντων ανιχνεύσεων, που έφθασε τα 53 πρόσωπα στα 1578 (3.4%).

OpenCV Alt.xml	Gentle AdaBoost	
Equalization & No Grouping & 90x90 min Win		
All	1578	
No det	15	1.0%
false/multi det	53	3.4%

Πιν. 3.5: Αποτελέσματα ανίχνευσης προσώπου χωρίς συγχώνευση και παράθυρο ανίχνευσης 90x90pix

3.3.6 Συνδυασμός των Μεθόδων Discrete και Gentle AdaBoost

Σε τελευταίο στάδιο έγινε ένας πειραματισμός για τη δυνατότητα της μεθόδου Default (Stump-based Discrete AdaBoost) να ανιχνεύσει πρόσωπα στις 15 εικόνες που η alt μέθοδος είχε αποτύχει. Το αποτέλεσμα ήταν να ανιχνευτούν σωστά επιπλέον 8 απ' αυτές, ρίχνοντας τις αποτυχίες στις 7.



Εικ. 3.3: Παραδείγματα πολλαπλών ανιχνεύσεων και διαδικασίας μετεπεξεργασίας

Οι ενδείξεις πως η μέθοδος Default (Stump-based Discrete AdaBoost) έχει δυνατότητα βελτιωμένης επίδοσης στην επιτυχημένη ανίχνευση σε περιπτώσεις που η Alt (Stump-based Gentle AdaBoost) αποτυγχάνει μας οδήγησε στην χρησιμοποίηση και των δύο μεθόδων με μία διαδικασία μετεπεξεργασίας για την συγχώνευση των πολλαπλών ανιχνεύσεων που προέρχονται από τις δύο μεθόδους.

Η λογική του αλγορίθμου μετεπεξεργασίας για την αποδοχή και τη συγχώνευση των υποψηφίων ανιχνευμένων περιοχών είναι είναι η εξής:

Αν default = 0 και alt >= 1
 Τότε επιστρέψε το μεγαλύτερο alt παράθυρο.

Αν default >= 1 και alt = 0
 τότε επιστρέψε το μεγαλύτερο default παράθυρο.

Αν default >= 1 και alt >= 1
 τότε πάρε όλους τους συνδυασμούς default και alt ανά δύο
 αν έχουν κοινή κάλυψη >= 60%
 τότε αν το ένα μέσα στο άλλο επιστρέψε την τομή (εσφαλμένη εστίαση)
 αλλιώς το μέσο παράθυρο (λάθος σκόπευση).

Από όλα τα παράθυρα που συγχωνεύτηκαν σε ένα επέστρεψε το μεγαλύτερο.

Η παραπάνω διαδικασία με συνδυασμό των μεθόδων Default (Stump-based Discrete AdaBoost) και Alt (Stump-based Gentle AdaBoost) έχει τα καλύτερα αποτελέσματα σε θετικές ανιχνεύσεις (1569 στις 1578), με αποτυχία ανίχνευσης στις 7 και με μόνο 2 λανθασμένες ανιχνεύσεις και συγχρόνως μηδενικές δευτερεύουσες ανιχνεύσεις.

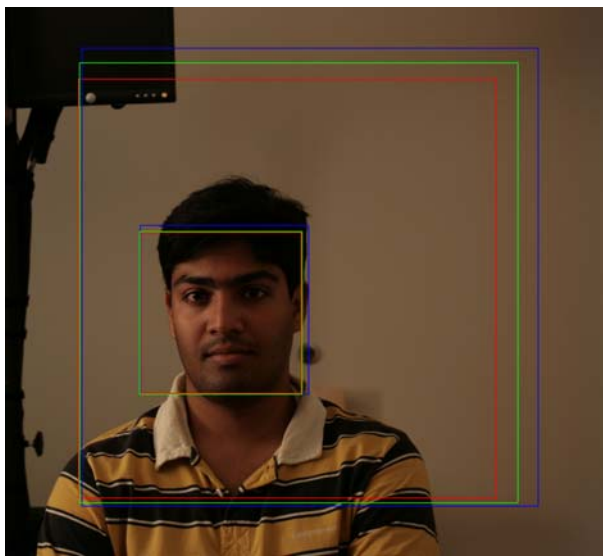
OpenCV Alt.xml + Default.xml	Gentle AdaBoost + Discrete AdaBoost	
Equalization & No Grouping & Post-processing		
All	1578	
no det	7	0.4%
false det	2	0.1%
total failure	9	0.6%

Πιν. 3.6: Αποτελέσματα ανίχνευσης προσώπου χωρίς συγχώνευση και ειδική μετεπεξεργασία

3.3.7 Ο Αλγόριθμος Μετεπεξεργασίας

Για την επιβεβαίωση της αποτελεσματικότητας της παραπάνω διαδικασίας, εκτελέστηκαν νέες δοκιμές στην 2^η Βάση Δεδομένων, υψηλής ευκρίνειας, που αποτελείται από 845 πρόσωπα.

Παρατηρήθηκε πως η αρχικά εφαρμοσμένη διαδικασία μετεπεξεργασίας για την συγχώνευση των πολλαπλών ανιχνεύσεων δεν ήταν επαρκής. Το αδύναμο σημείο του αλγορίθμου είναι η επιλογή του μεγαλύτερου παραθύρου σαν το τελικό παράθυρο του προσώπου που ανιχνεύτηκε.



Εικ. 3.4: Αποτελέσματα ανίχνευσης προσώπου πριν την μετεπεξεργασία

Αυτό έχει ως αποτέλεσμα την επιλογή του εξωτερικού μεγαλύτερου πράσινου πλαισίου αντί του εσωτερικού μικρότερου σαν το τελικό παράθυρο ανίχνευσης, όπως φαίνεται στην Εικ. 3.1.

Ετσι ο αλγόριθμος τροποποιήθηκε, ως προς την επιλογή του μεγαλύτερου παραθύρου, σε έναν πιο έξυπνο αλγόριθμο συγχώνευσης των παραθύρων σε ένα δεύτερο επίπεδο, ως εξής:

--2^ο επίπεδο συγχώνευσης των υποψηφίων παραθύρων--

Για όλα τα συγχωνευμένα παράθυρα λαμβανόμενα κατά ζεύγη

Εάν το ένα εμπεριέχεται στο άλλο

τότε επέστρεψε το μικρότερο

Αλλιώς εάν η επικαλυπτόμενη περιοχή είναι $\geq 80\%$

τότε επέστρεψε το μέσο παράθυρο

Εάν τίποτα από τα παραπάνω επέστρεψε το αρχικό παράθυρο αμετάβλητο

--Τέλος 2^{ου} επιπέδου συγχώνευσης των υποψηφίων παραθύρων--

Για όλα τα συγχωνευμένα παράθυρα από το 2^ο επίπεδο επέστρεψε το μεγαλύτερο ως το τελικό παράθυρο ανίχνευσης.

Η παραπάνω τροποποίηση είχε ως αποτέλεσμα την κατά 100% επιτυχή ανίχνευση προσώπων στην δεδομένη βάση προσώπων υψηλής ανάλυσης με ακριβή θέση του πλαισίου του προσώπου.

Επιπλέον, πρέπει να αναφερθεί ότι η παραπάνω τροποποίηση δεν επέφερε μεταβολή στα αποτελέσματα ανίχνευσης στην πρώτη χαμηλής ανάλυσης βάση δεδομένων (0.6% αποτυχία ανίχνευσης σε 1578 εικόνες προσώπων).

Η συνολική επιτυχής ανίχνευση στο σύνολο των εικόνων και των δύο βάσεων ανέρχεται σε 2414 σε σύνολο 2423, δηλ. 99,63%.

3.4 Συμπεράσματα: Η Τελική Διαδικασία Ανίχνευσης

Σύμφωνα με τους προηγούμενους πειραματισμούς καταλήξαμε:

1) να χρησιμοποιηθεί μία προ-επεξεργασία που αφορά την Ισοστάθμιση Ιστογράμματος (Histogram Equalization) και την εξομάλυνση με την βοήθεια ενός Γκαουσιανού φίλτρου (Gaussian smoothing).

2) Το παράθυρο ανίχνευσης να είναι πάνω από 90 pixels για την γρηγορότερη ανίχνευση και την αποφυγή εσφαλμένων ανιχνεύσεων μικρών περιοχών υφής.

3) να χρησιμοποιηθούν συνδυαστικά δύο μέθοδοι ανίχνευσης η Default και η Alt, που αντιστοιχούν στους ταξινομητές:

- Stump-based 24x24 Discrete Adaboost with cascade
- Stump-based 20x20 Gentle Adaboost with cascade.

4) να καταργηθεί η συγχώνευση των γειτονικών ανιχνεύσεων που κάνει το σύστημα.

5) να εφαρμοστεί εξειδικευμένη μετεπεξεργασία για την αποδοχή και την συγχώνευση των περιοχών που ανιχνεύονται ως πρόσωπα.

Κεφάλαιο 4

ΑΝΙΧΝΕΥΣΗ ΣΗΜΕΙΩΝ ΕΝΔΙΑΦΕΡΟΝΤΟΣ ΣΕ ΠΡΟΣΩΠΑ

4.1 Γενικά

Ο εντοπισμός συγκεκριμένων σημείων ενδιαφέροντος (interest points) σε ένα πρόσωπο είναι πολύ σημαντικός στις εφαρμογές αναγνώρισης προσώπου. Στα συστήματα αναγνώρισης προσώπου η ευθυγράμμιση (alignment-registration) είναι απαραίτητη ώστε το δείγμα και το πρωτότυπο να είναι συγκρίσιμα. Γενικά τα συστήματα αναγνώρισης προσώπων περιλαμβάνουν την ανίχνευση (detection), την ευθυγράμμιση (alignment) και την αναγνώριση (recognition) του προσώπου. Για να είναι η αναγνώριση προσώπου πλήρως αυτοματοποιημένη η ανάδειξη συγκεκριμένων *χαρακτηριστικών του προσώπου* (facial features) είναι ένα κρίσιμο πρώτο βήμα στην διαδικασία αναγνώρισης προσώπου (face recognition). Τα χαρακτηριστικά που χρησιμοποιούνται γενικότερα είναι: το κέντρο του αριστερού και δεξιού ματιού, οι δύο άκρες του αριστερού και δεξιού ματιού, η κορυφή της μύτης, το κέντρο του στόματος, οι άκρες του στόματος, η κορυφή του πηγουνιού, κλπ.

Σ' αυτό το σημείο είναι ενδιαφέρον να δούμε κάποιους ορισμούς που θα μας βοηθήσουν στη συνέχεια. Σύμφωνα με τους Dryden & Mardia [DRYD98]:

Μορφή (shape) είναι η γεωμετρική πληροφορία που απομένει όταν απαλοιφούν η θέση, η κλίμακα, και η στροφή ενός αντικειμένου, δηλ., η μορφή παραμένει αναλλοίωτη κάτω από μετασχηματισμούς ομοιότητας.

Ορόσημο (landmark) είναι ένα σημείο αντιστοιχίας σε κάθε αντικείμενο που παραμένει αναλλοίωτο μεταξύ των αντικειμένων ενός πληθυσμού.

Μία μορφή μπορεί να περιγραφεί με τον καθορισμό ενός πεπερασμένου αριθμού οροσήμων. Στην βιβλιογραφία εμφανίζονται πολλοί συνώνυμοι όροι για τα ορόσημα, όπως *σημεία ενδιαφέροντος* (interest points), *σημεία αναφοράς* (fiducial points), *σημεία κλειδιά* (key points), *σημεία αγκύρωσης* (anchor points), *χαρακτηριστικά σημεία* (feature points), *σημεία ελέγχου* (control points) κ.ά.

Τα ορόσημα διακρίνονται σε τρεις κατηγορίες:

Τα *ανατομικά ορόσημα* (anatomical landmarks) που προσδιορίζονται από ειδικούς και αντιστοιχούν σε οργανισμούς και έχουν βιολογική σημασία, όπως οι γωνίες των ματιών ή η ένωση δύο ραφών μεταξύ οστών του κρανίου.

Τα *μαθηματικά ή γεωμετρικά ορόσημα* (mathematical/geometrical landmarks) είναι σημεία που προσδιορίζονται σε ένα αντικείμενο σύμφωνα με κάποια μαθηματική ή γεωμετρική ιδιότητα της μορφής του αντικειμένου, όπως σημεία με υψηλή καμπυλότητα, ή ένα ακρότατο σημείο. Τα γεωμετρικά ορόσημα είναι ιδιαίτερα χρήσιμα σε συστήματα αυτόματης αναγνώρισης και ανάλυσης.

Τα *ψευδο-ορόσημα* (pseudo-landmarks) είναι κατασκευασμένα σημεία σύμφωνα με συγκεκριμένους κανόνες από τα ανατομικά ή/και τα γεωμετρικά ορόσημα, όπως σημεία που βρίσκονται στο μέσον του τμήματος δύο γεωμετρικών οροσήμων.

Καθορισμένα ορόσημα (labelled landmarks) είναι τα ορόσημα στα οποία έχει καθοριστεί ένας αριθμός ή μία ονομασία (label) ο οποίος και τα ταυτοποιεί, ώστε όταν συγκρίνονται δύο αντικείμενα να χρησιμοποιούνται συγκεκριμένα ζεύγη οροσήμων.

Επομένως στο πρόβλημα που έχουμε να αντιμετωπίσουμε δεν αρκεί μόνο ο εντοπισμός γενικών σημείων ή περιοχών ενδιαφέροντος στα πρόσωπα, αλλά είναι απαραίτητη και η ταυτοποίησή τους, ώστε να αντιστοιχηθούν στα συγκεκριμένα ανατομικά ορόσημα που αναζητούμε.

Σύμφωνα με τους Lu & Jain [LJ05], τα χαρακτηριστικά ενός προσώπου (facial features) ανάλογα με τη μέθοδο αναζήτησής τους μπορούν να είναι τριών τύπων:

Χαρακτηριστικές περιοχές (regions), *σημεία ή ορόσημα* (key points / landmarks), ή *καμπύλες* (contours). Γενικά ο εντοπισμός χαρακτηριστικών σημείων, παρέχει μία μεγαλύτερης ακρίβειας και πιο συνεπή μέθοδο για την εφαρμογή της ευθυγράμμισης και μάλιστα με μικρότερο κόστος στην πολυπλοκότητα και την υλοποίηση από τις άλλες μεθόδους [LJ05].

Οι μέθοδοι εντοπισμού χαρακτηριστικών περιοχών στηρίζονται σε τεχνικές στατιστικής μάθησης με την μέθοδο Adaboost και τη χρήση χαρακτηριστικών τύπου Haar των Viola & Jones [VIOL01] και Lienhart [LIEN02]. Ο εντοπισμός των χαρακτηριστικών γίνεται μετά με σύγκριση της περιοχής ενδιαφέροντος με τα υποδείγματα που έχει "μάθει" ο ταξινομητής. Για παράδειγμα ο ταξινομητής εκπαιδεύεται με δείγματα γωνιών ματιού που μετά αναζητούνται σε εικόνες προσώπων. Η μέθοδος αυτή χρησιμοποιείται από τους Cristinacce & Cootes [CC03], από τους Chen et al. [CHEN04] καθώς και από τους Zuo & With [ZUO04].

Ενδιαφέρουσες μέθοδοι για την εξαγωγή σημείων ενδιαφέροντος είναι η μέθοδος εντοπισμού γωνιών του Harris [HAR88] και η μέθοδος εντοπισμού συμμετριών του Reissfeld [REIS95]. Οι μέθοδοι αυτές είναι γενικές γεωμετρικές μέθοδοι που εφαρμόζονται στην εικόνα έντασης (intensity image) ή στην εικόνα βάρους (depth image).

Σύμφωνα με τους Schmid et al. [SMB00], η μέθοδος εντοπισμού γωνιών του Harris παρουσιάζει τα καλύτερα αποτελέσματα από άλλες μεθόδους εντοπισμού σημείων ή καμπύλων ενδιαφέροντος, εμφανίζοντας ευρωστία σε μεταβολές στροφής, κλίμακας, φωτισμού, προοπτικής και ύπαρξης θορύβου. Επίσης περιέχει την περισσότερη πληροφορία σε σχέση με τις άλλες μεθόδους. Η μέθοδος εντοπισμού συμμετριών του Reissfeld δεν συμπεριλαμβάνεται στην παραπάνω αξιολόγηση.

Η μέθοδος εντοπισμού γωνιών του Harris χρησιμοποιείται για τον εντοπισμό των σημείων ενδιαφέροντος για την ευθυγράμμιση των 3-Δ και 2-Δ μοντέλων στην εργασία αναγνώρισης προσώπων των Lu & Jain [LJ05]. Στη συνέχεια για την ταυτοποίηση των σημείων ενδιαφέροντος χρησιμοποιούνται τεχνικές ικανοποίησης περιορισμών, όπως σε ένα Πρόβλημα Ικανοποίησης Περιορισμών (Constraint Satisfaction Problem – CSP) [LJC06, CSJ05].

Μία παραλλαγή της μεθόδου εντοπισμού συμμετριών του Reissfeld χρησιμοποιείται για τον εντοπισμό και την ευθυγράμμιση των προσώπων στην εργασία εξακρίβωσης προσώπων των Tsalakanidou et al. [TMS05].

4.2 Ο Ανιχνευτής Γωνιών του Harris

Ο ανιχνευτής γωνιών του Harris [HAR88] βασίζεται στην τοπική *συνάρτηση αυτοσυσχετισμού* (autocorrelation function) ενός σήματος, όπου η συνάρτηση αυτοσυσχετισμού μετρά τις τοπικές μεταβολές της 2Δ εικόνας έντασης I , σε μικρές μετατοπίσεις σε διαφορετικές κατευθύνσεις από ένα συγκεκριμένο σημείο της εικόνας.

Δεδομένης μίας μετατόπισης $(\Delta x, \Delta y)$ και ενός σημείου (x, y) , η συνάρτηση αυτοσυσχετισμού ορίζεται ως:

$$c(x, y) = \sum_{(x_i, y_i) \in W} [I(x_i, y_i) - I(x_i + \Delta x, y_i + \Delta y)]^2$$

όπου $I(x, y)$ υποδηλώνει την ένταση της εικόνας και (x_i, y_i) τα σημεία σε ένα παράθυρο W με κέντρο το (x, y) .

Η ένταση των σημείων του παραθύρου W εξομαλύνεται με την χρήση ενός Γκαουσιανού φίλτρου με κέντρο το (x,y) :

$$G(x_i, y_i) = e^{-[(x_i-x)^2+(y_i-y)^2]/2\sigma^2}$$

Το μέγεθος του Γκαουσιανού φίλτρου καθορίζει το ισοζύγιο μεταξύ της ακρίβειας εντοπισμού του σημείου ενδιαφέροντος και της αποφυγής του θορύβου της εικόνας.

Αποδεικνύεται [DERP04] ότι η συνάρτηση αυτοσυσχετισμού προσεγγίζεται, μέσω του πρώτου όρου του αναπτύγματος Taylor, από τη σχέση:

$$c(x, y) = [\Delta x \quad \Delta y] C(x, y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}$$

όπου $C(x,y)$ ο πίνακας συμμεταβολής, που εκφράζει την τοπική μεταβολή της έντασης της εικόνας:

$$C(x, y) = \begin{bmatrix} \sum_w (I_x(x_i, y_i))^2 & \sum_w (I_x(x_i, y_i) \cdot I_y(x_i, y_i)) \\ \sum_w (I_x(x_i, y_i) \cdot I_y(x_i, y_i)) & \sum_w (I_y(x_i, y_i))^2 \end{bmatrix}$$

όπου I_x και I_y εκφράζουν τις μερικές παραγώγους της έντασης εικόνας ως προς x και y αντίστοιχα.

Εάν α και β είναι οι ιδιοτιμές του πίνακα $C(x,y)$ τότε αυτές εκφράζουν την τοπική μεταβολή της έντασης και είναι αναλλοίωτες σε στροφές.

Ανάλογα με τις τιμές των ιδιοτιμών α και β διακρίνουμε τις εξής περιπτώσεις [HAR88, DERP04]:

1. Αν και οι δύο ιδιοτιμές είναι μικρές, τότε η συνάρτηση αυτοσυσχετισμού είναι σχεδόν ομαλή, δηλ., με μικρές μεταβολές σε οποιαδήποτε κατεύθυνση, οπότε η αντίστοιχη περιοχή της εικόνας έχει προσεγγιστικά σταθερή ένταση.
2. Αν η μία ιδιοτιμή είναι μεγάλη και η άλλη μικρή, τότε η συνάρτηση αυτοσυσχετισμού εμφανίζει τη μορφή ράχης, δηλ., μετατοπίσεις κατά μήκος της ράχης έχουν μικρές μεταβολές έντασης, ενώ μετατοπίσεις κάθετα στην ράχη έχουν μεγάλες μεταβολές έντασης. Η περίπτωση αυτή αντιστοιχεί σε ακμή.
3. Αν και οι δύο ιδιοτιμές είναι μεγάλες, τότε η συνάρτηση αυτοσυσχετισμού εμφανίζει τη μορφή αιχμής, δηλ., μετατοπίσεις κατά οποιαδήποτε κατεύθυνση έχουν μεγάλες μεταβολές έντασης. Η περίπτωση αυτή αντιστοιχεί σε γωνία.

Για να αποφευχθεί ο αναλυτικός υπολογισμός των ιδιοτιμών α και β , οι Harris & Stephens [HAR88] επινόησαν ένα μέτρο που εκφράζει την απόκριση των παραπάνω περιπτώσεων.

Αν $\det(C)$ και $\text{tr}(C)$ είναι η ορίζουσα (determinant) και το ίχνος (trace) του πίνακα $C(x,y)$, τότε η απόκριση της γωνιότητας (cornerness response) ορίζεται ως εξής:

$$R(x,y) = \det[C(x,y)] - k \cdot \text{tr}^2[C(x,y)],$$

όπου $\det(C) = \alpha \cdot \beta$ και $\text{tr}(C) = \alpha + \beta$, ενώ η παράμετρος k παίρνει παραδοσιακά τις τιμές $0.04 \sim 0.06$.

Η απόκριση R παίρνει μεγάλες θετικές τιμές στα σημεία που υπάρχουν γωνίες, αρνητικές τιμές στα σημεία που υπάρχουν ακμές και μικρές τιμές σε ομαλές περιοχές.

Η παραπάνω προσέγγιση παρουσιάζει τα εξής προβλήματα:

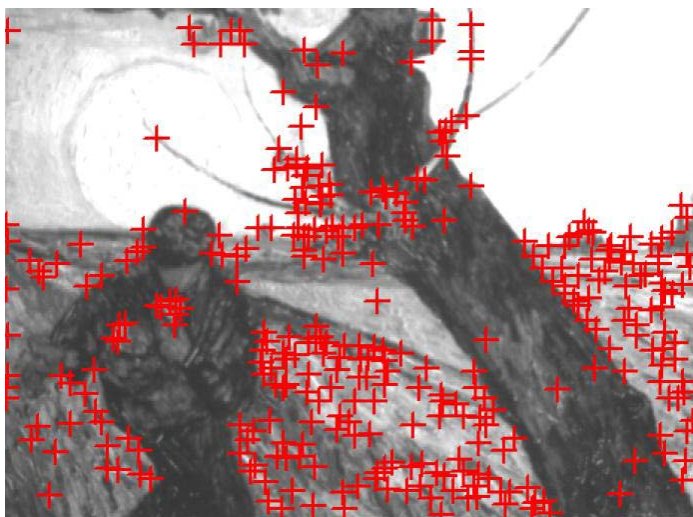
- Επειδή ο πίνακας συμμεταβολής περιέχει σαν στοιχεία τα τετράγωνα της βαθμίδας της έντασης, η απόκριση R εξαρτάται από τις βαθμίδες της έντασης στην τετάρτη δύναμη, πράγμα που συνεπάγεται υψηλή εξάρτηση της R από τις αποκλίσεις της αντίθεσης (contrast) της εικόνας.
- Η εξάρτηση της R από την αντίθεση της εικόνας κάνει δύσκολο τον προσδιορισμό του κατωφλίου, πάνω από το οποίο η απόκριση R θεωρείται ότι παίρνει μεγάλες τιμές, καθώς επίσης και της παραμέτρου k .

Ένα άλλο μέτρο της απόκρισης R που εκφράζει τη γωνιότητα και αποφεύγει την παράμετρο k προτάθηκε από τη Noble [NOB89]:

$$R(x,y) = \frac{\det[C(x,y)]}{\text{tr}[C(x,y)]}$$

Αν μία από τις ιδιοτιμές τείνει στο μηδέν τότε η $\det(C)$ και επομένως η R τείνει στο μηδέν. Αν οι ιδιοτιμές είναι παρόμοιες αλλά μικρές, τότε η R πάλι τείνει στο μηδέν. Όμως, αν οι ιδιοτιμές είναι παρόμοιες αλλά μεγάλες, τότε η $\det(C)$ είναι πολύ μεγαλύτερη του $\text{tr}(C)$ και επομένως η R είναι μεγάλη.

Τέλος ένα υποψήφιο σημείο επιλέγεται σαν σημείο που αντιπροσωπεύει σημείο ενδιαφέροντος αν η απόκρισή του R , είναι ένα τοπικό μέγιστο 8-γειτόνων με τη μέθοδο απαλοιφής των μη-μεγίστων (non-maxima suppression) [HAR88].



Εικ. 4.1: Ανιχνευτής γωνιών του Harris [SMB00]

Ο υπολογισμός των μερικών παραγώγων στην εργασία του Harris γίνεται με τη συνέλιξη της εικόνας έντασης με τους τελεστές του Prewitt $[-1 \ 0 \ +1]$ [GONZ02] σε κάθε κατεύθυνση.

Μία βελτιωμένη εκδοχή του αρχικού ανιχνευτή γωνιών του Harris υλοποιείται αντικαθιστώντας τον αρχικό υπολογισμό των παραγώγων μέσω των τελεστών $[-2 \ -1 \ 0 \ 1 \ 2]$ που προκύπτουν από τις παραγώγους της Γκαουσιανής (derivatives of Gaussian) με $\sigma = 1$ [SMB00].

Στην δική μας εκδοχή για τον υπολογισμό των παραγώγων και την δημιουργία του χάρτη βαθμίδας (gradient map) της εικόνας χρησιμοποιήθηκαν οι τελεστές Sobel

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix},$$

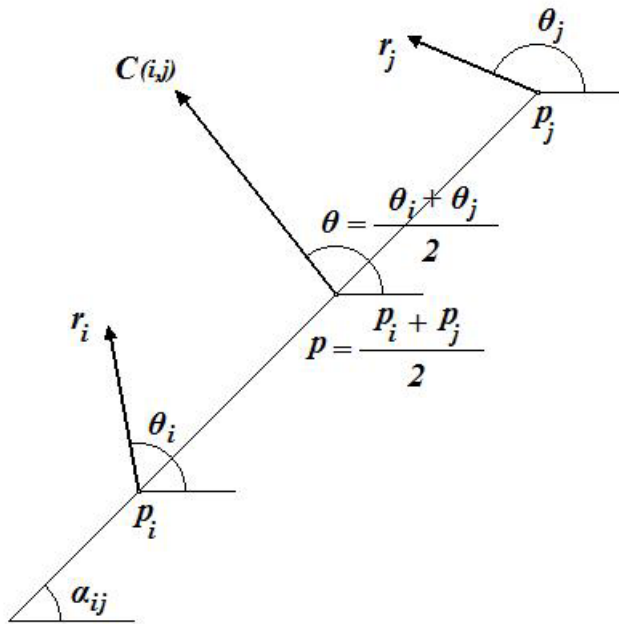
που παρουσιάζουν βελτιωμένη απόδοση σε σχέση με τους τελεστές Prewitt [GONZ02].

4.3 Ο Ανιχνευτής Συμμετριών του Reisfeld

Ο ανιχνευτής συμμετριών του Reisfeld συσχετίζει κάθε εικονοστοιχείο μιας εικόνας με ένα μέτρο συμμετρίας, υπολογίζοντας ένα *χάρτη συμμετρίας* (symmetry map) από τον *χάρτη ακμών* (edge map) ή *χάρτη βαθμίδας* (gradient map) της εικόνας. Ο υπολογιζόμενος χάρτης συμμετρίας δεν απαιτεί την πρότερη γνώση του σχήματος του

αντικειμένου στο οποίο εφαρμόζεται, είναι αναλλοίωτος σε μετασχηματισμούς ομοιότητας και μπορεί να χρησιμοποιηθεί για την εξαγωγή σημείων ενδιαφέροντος, δηλ., των σημείων που παρουσιάζουν υψηλή συμμετρία.

Αν $p_k = (x_k, y_k)$ ένα σημείο της εικόνας με ένταση I_k και $\nabla I_k = \left(\frac{\partial}{\partial x} I_k, \frac{\partial}{\partial y} I_k \right)$ η βαθμίδα (gradient) της έντασης, για κάθε σημείο p_k ορίζουμε το διάνυσμα $v_k = (r_k, \theta_k)$, που εκφράζει το μέτρο $r_k = \log(1 + \|\nabla I_k\|)$ και την κλίση $\theta_k = \arctan\left(\frac{\partial}{\partial y} I_k / \frac{\partial}{\partial x} I_k\right)$ της βαθμίδας.



Εικ. 4.2: Συνεισφορά στο μέτρο συμμετρίας στο p των βαθμίδων έντασης στα p_i και p_j

Για κάθε σημείο p ορίζουμε όλα τα ζεύγη σημείων p_i και p_j που ανήκουν στο σύνολο:

$$\Gamma(p) = \left\{ (i, j) \mid \frac{p_i + p_j}{2} = p \right\},$$

την συνάρτηση βάρους της απόστασης:

$$D_\sigma(i, j) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\|p_i - p_j\| / 2\sigma^2}$$

και την συνάρτηση βάρους της φάσης:

$$P(i, j) = (1 - \cos(\theta_i + \theta_j - 2\alpha_{ij}))(1 - \cos(\theta_i - \theta_j)),$$

όπου α_{ij} η γωνία που σχηματίζει η ευθεία (p_i, p_j) με την οριζόντια.

Έτσι η συνεισφορά των σημείων p_i και p_j στο μέτρο της συμμετρίας στο p (Εικ. 4.1), ορίζεται ως εξής:

$$C(i, j) = D_{\sigma}(i, j) \cdot P(i, j) \cdot r_i \cdot r_j$$

και η κατεύθυνση αυτής της συνεισφοράς:

$$\theta(i, j) = \frac{\theta_i + \theta_j}{2}.$$

Το μέτρο *ισοτροπικής συμμετρίας* (isotropic symmetry magnitude) για κάθε σημείο p μίας εικόνας, ορίζεται ως εξής:

$$M_{\sigma}(p) = \sum_{(i,j) \in \Gamma(p)} C(i, j), \text{ η οποία προσμετρά την συμμετρία σε κάθε κατεύθυνση.}$$

Η *κατεύθυνση συμμετρίας* (symmetry direction) ορίζεται ως η γωνία:

$$\phi(p) = \theta(i, j), \text{ η οποία μεγιστοποιεί την } C(i, j) \text{ για κάποια } (i, j) \in \Gamma(p).$$

Τελικά η συμμετρία στο σημείο p , ορίζεται από το ζεύγος:

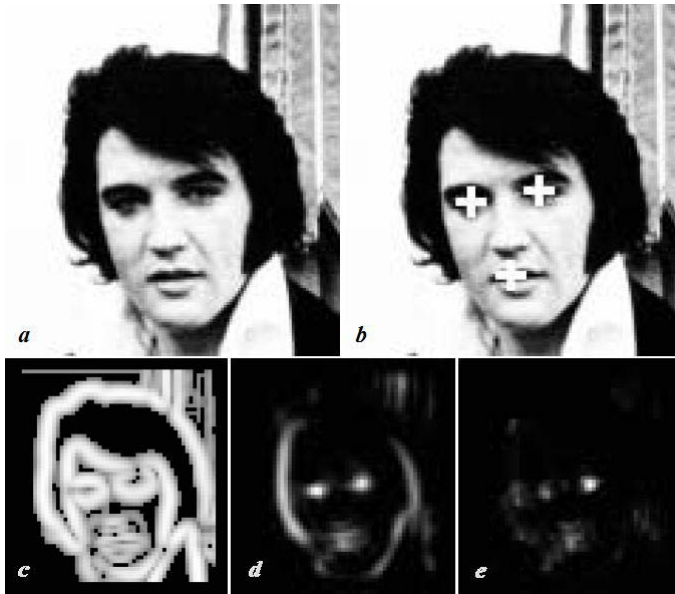
$$S_{\sigma}(p) = (M_{\sigma}(p), \phi(p))$$

Ο παραπάνω ορισμός της συμμετρίας προσδιορίζει την *ανακλαστική* ή *αξονική συμμετρία* και είναι αναλλοίωτος σε μετασχηματισμούς στροφής και μετατόπισης.

Σε αρκετές περιπτώσεις είναι αναγκαίο να προσδιοριστούν σημεία που εμφανίζουν συμμετρία σε πολλαπλές διευθύνσεις και όχι σε μόνο μία κύρια. Μια τέτοια συμμετρία αναφέρεται ως *ακτινική συμμετρία* (radial symmetry) και ορίζεται ως εξής:

$$RS_{\sigma}(p) = \sum_{(i,j) \in \Gamma(p)} C(i, j) \cdot \sin^2(\theta(i, j) - \phi(p))$$

Η παραπάνω συνάρτηση μεγιστοποιείται σε σημεία που περιβάλλονται από ακμές. Ο παραπάνω ορισμός δείχνει πως η ακτινική συμμετρία εκφράζεται με ένα συνεχή τελεστή που δεν είναι ευαίσθητος στις ασυνέχειες των καμπύλων που περιβάλλουν το σημείο p .



Εικ. 4.3: Ανιχνευτής συμμετριών Reisfeld [REIS95]. (a) Αρχική εικόνα (b) Σημεία μέγιστης ακτινικής συμμετρίας (c) Χάρτης ακμών (d) Χάρτης ιστροπικής συμμετρίας (e) Χάρτης ακτινικής συμμετρίας

ΠΑΡΑΤΗΡΗΣΕΙΣ:

1. Η τοπικότητα των παραπάνω τελεστών συμμετρίας εκφράζεται με τη χρήση του Γκαουσιανού φίλτρου $D_{\sigma}(i,j)$. Οι διαφορετικές τιμές της σ επιτρέπουν τον υπολογισμό της συμμετρίας σε παράθυρα διαφορετικής κλίμακας. Η παραπάνω Γκαουσιανή εμφανίζει κυκλικές ισοϋψείς, αν και θα μπορούσαν κατά περίπτωση να χρησιμοποιηθούν και Γκαουσιανές με ελλειπτικές ισοϋψείς π.χ. για τα μάτια σε πρόσωπα.
2. Η συνάρτηση φάσης $P(i,j)$ αποτελείται από δύο όρους. Ο πρώτος όρος $1 - \cos(\theta_i + \theta_j - 2\alpha_{ij})$ καθορίζει τη μέγιστη συμμετρία στην περίπτωση που οι βαθμίδες των εντάσεων στα p_i και p_j έχουν την ίδια κατεύθυνση, δηλ., όταν $(\theta_i - \alpha_{ij}) + (\theta_j - \alpha_{ij}) = \pi$, και μειώνεται σε αντίθετη περίπτωση. Ο δεύτερος όρος $1 - \cos(\theta_i - \theta_j)$ χρησιμοποιείται για να αντισταθμίσει την περίπτωση που $\theta_i - \alpha_{ij} = \theta_j - \alpha_{ij} = \pi/2$, οπότε ο πρώτος όρος γίνεται μέγιστος, πράγμα που συμβαίνει σε μία ευθεία ακμή, που δεν θεωρείται ως ενδιαφέρουσα.
3. Ο όρος $r_i \cdot r_j$ εκφράζει τη συνεισφορά των βαθμίδων της έντασης στη συμμετρία. Η χρήση των βαθμίδων αντί των εντάσεων αποτυπώνει τη σημασία των ακμών των αντικειμένων στον προσδιορισμό της συμμετρίας. Επίσης η χρήση του λογαρίθμου της βαθμίδας $r_k = \log(1 + \|\nabla I_k\|)$, αντί του μέτρου $\|\nabla I_k\|$ της

βαθμίδας, θεωρείται καταλληλότερη γιατί μειώνει τις διαφορές μεταξύ των ισχυρών βαθμίδων και επομένως μειώνει την ευαισθησία της συμμετρίας σε πολύ ισχυρές ακμές.

4. Στην δική μας εκδοχή για τον υπολογισμό των παραγώγων και την δημιουργία του χάρτη βαθμίδας (gradient map) της εικόνας χρησιμοποιήθηκαν οι τελεστές

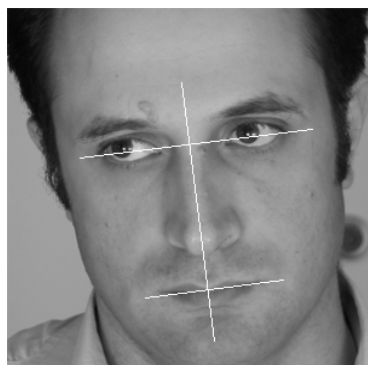
$$\text{Sobel} \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \text{ όπως και στον ανιχνευτή γωνιών του Harris.}$$

4.4 Ανίχνευση Τοπικών Ανακλαστικών Συμμετριών

Οι δύο παραπάνω μέθοδοι ανίχνευσης προσδιορίζουν γενικά γεωμετρικά σημεία ενδιαφέροντος, οπότε είναι φανερό πως απαιτείται να ληφθούν υπόψη τοπολογικές ιδιότητες του προσώπου για να είναι εφικτή η αναγνώριση των ανατομικών σημείων που μας ενδιαφέρουν (anatomical landmarks) και να γίνει απόδοση ετικετών σε αυτά (labelled landmarks).

Έτσι, ένα επόμενο βήμα είναι η υλοποίηση ενός αλγορίθμου που θα προσδιορίζει συγκεκριμένους άξονες συμμετρίας σε συγκεκριμένες περιοχές των εικόνων των προσώπων, ώστε να χρησιμοποιηθούν ως οδηγοί για την αναγνώριση των ανατομικών σημείων ενδιαφέροντος.

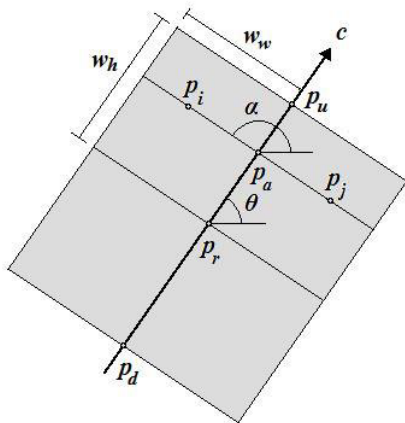
Οι άξονες συμμετρίας που αναζητούμε είναι ο κεντρικός άξονας συμμετρίας του προσώπου, οι άξονες συμμετρίας στην περιοχή των ματιών ή σε κάθε μάτι χωριστά, και στην περιοχή του στόματος. Στόχος μας είναι να προκύψουν άξονες συμμετρίας όπως στην Εικ. 4.4.



Εικ. 4.4: Ανίχνευση τοπικών αξόνων ανακλαστικής συμμετρίας

Ο αλγόριθμος αυτός μπορεί να στηριχθεί στον γενικό αλγόριθμο ανίχνευσης συμμετριών του Reisfeld, και να προσδιορίζει την κατεύθυνση του άξονα συμμετρίας σε σημεία που έχουν ισχυρό μέτρο συμμετρίας μέσα σε συγκεκριμένη περιοχή ενδιαφέροντος.

Για τον προσδιορισμό των τοπικών αυτών συμμετριών εφαρμόστηκε μία δική μας παραλλαγή του γενικού μετασχηματισμού συμμετριών του Reisfeld, εμπνευσμένη από την εργασία εξακρίβωσης προσώπων των Tsalakanidou et al. [TMS05].



Εικ. 4.5: Υπολογισμός μέτρου ανακλαστικής συμμετρίας σε συγκεκριμένο παράθυρο ενδιαφέροντος

Η ιδέα είναι η εξής:

Για ένα σημείο της εικόνας p_r και για συγκεκριμένο άξονα συμμετρίας c ορίζουμε ένα παράθυρο ενδιαφέροντος με πλάτος $2w_w$ και ύψος $2w_h$, με κέντρο το p_r (Εικ. 4.5). Αν p_i και p_j είναι σημεία που είναι συμμετρικά ως προς τον άξονα συμμετρίας, δηλ., $p_a = (p_i + p_j)/2$, αυτά συνεισφέρουν στο μέτρο συμμετρίας που ορίζει το παράθυρο ενδιαφέροντος και αποδίδεται στο σημείο p_r .

Για κάθε σημείο p_r , λοιπόν, ορίζουμε όλα τα ζεύγη σημείων p_i και p_j που ανήκουν στο σύνολο:

$$\Gamma(p_r) = \left\{ (i, j) \mid \frac{p_i + p_j}{2} = p_a \right\},$$

όπου p_a τα σημεία του άξονα συμμετρίας, $p_a \in \text{int line}(p_d, p_u)$,

Αν για κάθε σημείο p_k έχουμε προσδιορίσει το διάνυσμα $v_k = (r_k, \theta_k)$, που εκφράζει το μέτρο $r_k = \log(1 + \|\nabla I_k\|)$ και την κλίση $\theta_k = \arctan\left(\frac{\partial I_k}{\partial y} / \frac{\partial I_k}{\partial x}\right)$ της βαθμίδας της εικόνας έντασης I , τότε η συνάρτηση βάρους της φάσης εκφράζεται ως:

$$P(i, j) = 1 - \cos(\theta_i + \theta_j - 2\alpha_{ij}),$$

όπου α η γωνία που σχηματίζει η ευθεία (p_i, p_j) με την οριζόντια,

και η συνεισφορά των συμμετρικών σημείων p_i και p_j στο μέτρο της συμμετρίας στο p_a , ορίζεται ως εξής:

$$C(i, j) = P(i, j) \cdot r_i \cdot r_j$$

Το ολικό μέτρο ιστροπικής συμμετρίας (isotropic symmetry magnitude) στο p_r , ορίζεται ως εξής:

$$M_\sigma(p_r) = \sum_{(i,j) \in \Gamma(p_r)} C(i, j), \text{ και οποία προσμετρά την συμμετρία από όλα τα } p_i \text{ και } p_j \text{ και}$$

για κάθε p_a .

Μεταβάλλοντας την κατεύθυνση θ του άξονα συμμετρίας c (Εικ. 4.5), βρίσκουμε τον βέλτιστο άξονα συμμετρίας με κατεύθυνση:

$$\phi(p_r) = \theta, \text{ η οποία μεγιστοποιεί την } M_\sigma(p_r) = \sum_{(i,j) \in \Gamma(p_r)} C(i, j).$$

Τελικά η συμμετρία στο σημείο p_r , ορίζεται από το ζεύγος:

$$S_\sigma(p_r) = (M_\sigma(p_r), \phi(p_r))$$

ΠΑΡΑΤΗΡΗΣΕΙΣ:

1. Η τοπικότητα των παραπάνω τελεστών συμμετρίας εκφράζεται με τη χρήση του παραθύρου ενδιαφέροντος ως προς το οποίο υπολογίζεται το μέτρο της ιστροπικής συμμετρίας. Το Γκαουσιανό φίλτρο $D_\sigma(i, j)$ δεν θεωρήθηκε απαραίτητο να χρησιμοποιηθεί, για να μειωθεί το υπολογιστικό κόστος.
2. Στην συνάρτηση φάσης $P(i, j)$ χρησιμοποιήθηκε μόνον ο πρώτος όρος $1 - \cos(\theta_i + \theta_j - 2\alpha_{ij})$ που καθορίζει τη μέγιστη συμμετρία στην περίπτωση που οι βαθμίδες των εντάσεων στα p_i και p_j έχουν την ίδια κατεύθυνση. Η χρήση μόνο του πρώτου όρου αντί των δύο όρων

$P(i, j) = (1 - \cos(\theta_i + \theta_j - 2\alpha_{ij}))(1 - \cos(\theta_i - \theta_j))$ στην αρχική εργασία του Reisfeld, έδωσε γενικά καλύτερα αποτελέσματα.

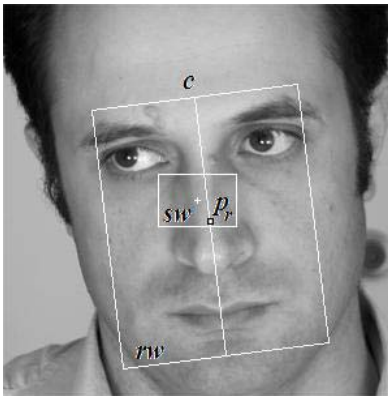
3. Ο όρος $r_i \cdot r_j$ εκφράζει τη συνεισφορά των βαθμίδων της έντασης στη συμμετρία.

4. Για τον υπολογισμό των παραγώγων και την δημιουργία του χάρτη βαθμίδας

(gradient map) της εικόνας χρησιμοποιήθηκαν οι τελεστές Sobel $\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$,

όπως και στον ανιχνευτή γωνιών του Harris.

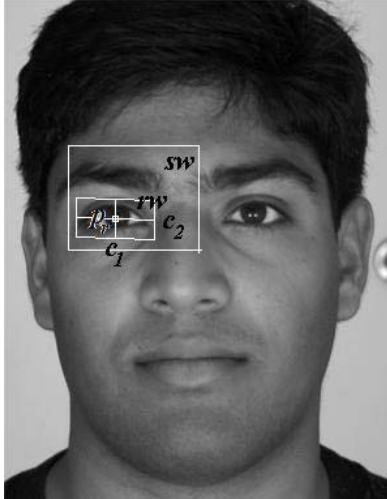
Σύμφωνα με τα παραπάνω, υπολογίζοντας το μέτρο συμμετρίας για διάφορα p_r που ανήκουν σε ένα παράθυρο αναζήτησης sw , βρίσκουμε την θέση p_r και την κατεύθυνση $\phi(p_r)$ του άξονα συμμετρίας c που εμφανίζει το μέγιστο μέτρο συμμετρίας $M_\sigma(p_r)$ στο παράθυρο ενδιαφέροντος rw όπως φαίνεται στην Εικ.4.6.



Εικ. 4.6: Παράθυρο αναζήτησης και ανιχνευμένο παράθυρο μέγιστου μέτρου συμμετρίας

Η μέθοδος αυτή μπορεί να χρησιμοποιηθεί για τον εντοπισμό σε ένα πρόσωπο περιοχών με ισχυρή ανακλαστική συμμετρία. Τέτοιες περιοχές είναι ολόκληρη η περιοχή του προσώπου με κατακόρυφο άξονα συμμετρίας, οι περιοχές των ματιών με κατακόρυφο ή οριζόντιο άξονα και η περιοχή του στόματος με κατακόρυφο ή οριζόντιο άξονα. Οι πειραματικές δοκιμές θα δείξουν ποιές από αυτές τις περιοχές προσδιορίζονται εύρωστα, ώστε να μπορούν να χρησιμοποιηθούν αποτελεσματικά οι συμμετρίες που προκύπτουν.

Μία επέκταση της παραπάνω μεθόδου είναι ο υπολογισμός της βέλτιστης θέσης p_r και προσανατολισμού $\phi(p_r)$ του παράθυρου ενδιαφέροντος r_w με το μέγιστο μέτρο συμμετρίας $M_\sigma(p_r)$ υπολογισμένο ως προς δύο κάθετους μεταξύ τους άξονες c_1 και c_2 , όπως φαίνεται στην Εικ. 4.7.



Εικ. 4.7: Παράθυρο αναζήτησης και παράθυρο μέγιστου μέτρου συμμετρίας ως προς 2 άξονες

Η μέθοδος αυτή προσφέρεται για αντικείμενα που εμφανίζουν συμμετρία ως προς δύο κάθετους μεταξύ τους άξονες, όπως τα μάτια ή και το στόμα.

Κεφάλαιο 5

ΥΛΟΠΟΙΗΣΗ ΣΥΣΤΗΜΑΤΟΣ ΑΝΙΧΝΕΥΣΗΣ ΣΗΜΕΙΩΝ ΕΝΔΙΑΦΕΡΟΝΤΟΣ ΣΕ ΠΡΟΣΩΠΑ

5.1 Γενικά

Σε ένα σύστημα αναγνώρισης προσώπων, η ευθυγράμμιση (alignment / registration) μεταξύ του αναζητούμενου προσώπου δοκιμής (probe) και του εγγεγραμμένου προσώπου υποδείγματος (template) καθιστά τα πρόσωπα συγκρίσιμα. Η ευθυγράμμιση που στηρίζεται στην αντιστοίχιση χαρακτηριστικών σημείων αναφοράς (fiducial points) ή οροσήμων (landmarks) είναι ένα από τα πλέον κρίσιμα βήματα ώστε το σύστημα αναγνώρισης να καταστεί πλήρως αυτοματοποιημένο.

Ο στόχος μας είναι η αυτοματοποιημένη ανίχνευση σημείων ενδιαφέροντος (interest points) σε 2Δ εικόνες προσώπων, που θα οδηγήσει στον προσδιορισμό καθορισμένων οροσήμων (labelled landmarks), που αντιστοιχούν σε ανατομικά χαρακτηριστικά του προσώπου (facial features), όπως το κέντρο του αριστερού ή του δεξιού ματιού, η αριστερή και η δεξιά γωνία κάθε ματιού, η κορυφή της μύτης, το κέντρο του στόματος, οι γωνίες του στόματος κλπ.

Για τον σκοπό αυτό θα υλοποιηθούν οι μέθοδοι εντοπισμού γεωμετρικών χαρακτηριστικών, όπως η μέθοδος ανίχνευσης γωνιών του Harris και η μέθοδος ανίχνευσης συμμετριών του Reisfeld και θα αξιολογηθούν ως προς την ακρίβεια εντοπισμού (localization accuracy), την ευρωστία (robustness) και την επαναληψιμότητα (repeatability).

Επίσης θα υλοποιηθεί η μέθοδος προσδιορισμού τοπικής ανακλαστικής συμμετρίας που αναφέρθηκε στο προηγούμενο κεφάλαιο, στην προσπάθειά μας να ληφθούν υπόψη οι τοπολογικές ιδιότητες του προσώπου, ώστε να είναι εφικτή η αναγνώριση των ανατομικών σημείων που μας ενδιαφέρουν (anatomical landmarks) και να γίνει απόδοση ετικετών σε αυτά (labelled landmarks).

5.2 Υλοποίηση των ανιχνευτών

Παρακάτω παρατίθενται τα βασικά βήματα της υλοποίησης του Ανιχνευτή Γωνιών του Harris, του Ανιχνευτή Συμμετριών του Reissfeld και του Ανιχνευτή Τοπικών Ανακλαστικών Συμμετριών.

5.2.1 Ανιχνευτής Γωνιών του Harris

1. Μετατροπή της εικόνας I σε ασπρόμαυρη.

2. Εφαρμογή των τελεστών Sobel $\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$,

για την εξαγωγή των παραγώγων I_x και I_y της εικόνας.

3. Υπολογισμός των:

$$I_{x2}(x_i, y_i) = (I_x(x_i, y_i))^2$$

$$I_{y2}(x_i, y_i) = (I_y(x_i, y_i))^2$$

$$I_{xy}(x_i, y_i) = (I_x(x_i, y_i) \cdot I_y(x_i, y_i))$$

$$I_{xy2}(x_i, y_i) = (I_{xy}(x_i, y_i))^2$$

4. Καθορισμός του Γκαουσιανού φίλτρου

$$gmask = gauss_mask(conv_win, sigma),$$

με παραμέτρους το εύρος του φίλτρου $conv_win$,

και την τυπική απόκλιση σ ($sigma$).

5. Υπολογισμός των συνελίξεων:

$$GI_{x2} = conv(I_{x2}, gmask)$$

$$GI_y2 = \text{conv}(I_y2, \text{gmask})$$

$$GI_{xy}2 = \text{conv}(I_{xy}2, \text{gmask})$$

ώστε να καθορισθεί ο πίνακας συμμεταβολής:

$$C = \begin{bmatrix} GI_x2 & GI_{xy}2 \\ GI_{xy}2 & GI_y2 \end{bmatrix}$$

6. Υπολογισμός της ορίζουσας Det και του ίχνους Tr του πίνακα C

$$\text{Det} = GI_x2 \cdot GI_y2 - GI_{xy}2^2,$$

$$\text{Tr} = GI_x2 + GI_y2,$$

$$\text{Tr}2 = \text{Tr}^2,$$

ώστε να υπολογιστεί η απόκριση γωνιότητας (cornerness response):

$$R = \text{Det} - k \cdot \text{Tr}2, \text{ με } k = 0.04.$$

7. Επιλογή των τοπικών μεγίστων σε περιοχές εύρους max_win του χάρτη γωνιότητας (cornerness map) R, με τη μέθοδο απαλοιφής των μη-μεγίστων (non-maxima suppression)

$$P_{\max}(x, y) = \text{non_max_suppression}(R, \text{max_win})$$

8. Διάταξη κατά φθίνουσα σειρά των τοπικών μεγίστων P_max

9. Επιλογή των max_points ισχυρότερων μεγίστων

5.2.2 Ανιχνευτής Συμμετριών του Reisfeld

1. Μετατροπή της εικόνας I σε ασπρόμαυρη.

2. Εφαρμογή των τελεστών Sobel $\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$,

για την εξαγωγή των παραγώγων I_x και I_y της εικόνας.

3. Υπολογισμός του μέτρου της βαθμίδας I_r και της γωνίας κλίσης της βαθμίδας I_a :

$$I_r = \log(1 + \sqrt{I_x^2 + I_y^2})$$

$$I_a = \arctan(I_y / I_x)$$

4. Υπολογισμός σε συγκεκριμένο παράθυρο W με κέντρο p για κάθε ζεύγος συμμετρικών σημείων p_i και p_j των βοηθητικών συναρτήσεων:

$$d = \sqrt{(x(i) - x(j))^2 + (y(j) - y(i))^2}$$

$$D(i, j) = \frac{1}{\sqrt{2\pi}\sigma} e^{-d/2\sigma^2}$$

$$\alpha = \arctan(y(j) - y(i) / x(i) - x(j))$$

$$P(i, j) = 1 - \cos(I_a(i) + I_a(j) - 2\alpha)(1 - \cos(I_a(i) - I_a(j)))$$

$$C(i, j) = D(i, j) \cdot P(i, j) \cdot I_r(i) \cdot I_r(j)$$

με παραμέτρους το εύρος του παραθύρου $conv_win$, και την τυπική απόκλιση σ (sigma) του Γκαουσιανού φίλτρου.

5. Υπολογισμός του μέτρου ισοτροπικής συμμετρίας:

$$IS(p) = \sum_{(i,j) \in W} C(i, j),$$

της κατεύθυνσης συμμετρίας:

$$IP(p) = \frac{I_a(i) + I_a(j)}{2}, \text{ η οποία μεγιστοποιεί την } C(i, j) \text{ για κάποια } (i, j) \in W,$$

και της ακτινικής συμμετρίας:

$$IR(p) = \sum_{(i,j) \in W} C(i, j) \cdot \sin^2\left(\frac{I_a(i) + I_a(j)}{2} - IP(p)\right)$$

6. Επιλογή των τοπικών μεγίστων σε περιοχές εύρους max_win του χάρτη ακτινικής συμμετρίας (radial symmetry map) IR , με τη μέθοδο απαλοιφής των μη-μεγίστων (non-maxima suppression)

$$P_{\max}(x, y) = \text{non_max_supression}(\text{IR}, \text{max_win})$$

7. Διάταξη κατά φθίνουσα σειρά των τοπικών μεγίστων P_{\max}

8. Επιλογή των max_points ισχυρότερων μεγίστων

5.2.3 Ανιχνευτής Τοπικής Ανακλαστικής Συμμετρίας

1. Μετατροπή της εικόνας I σε ασπρόμαυρη.

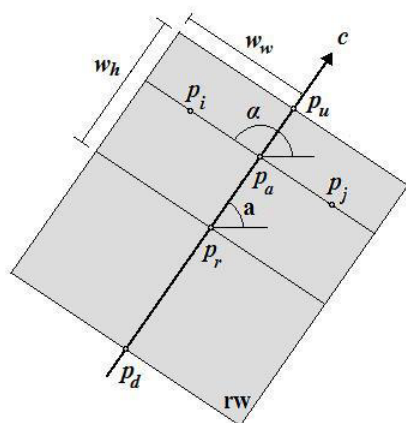
2. Εφαρμογή των τελεστών Sobel
$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix},$$

για την εξαγωγή των παραγώγων I_x και I_y της εικόνας.

3. Υπολογισμός του μέτρου της βαθμίδας I_r και της γωνίας κλίσης της βαθμίδας I_a :

$$I_r = \log(1 + \sqrt{I_x^2 + I_y^2})$$

$$I_a = \arctan(I_y / I_x)$$



Εικ. 5.1: Υπολογισμός μέτρου ανακλαστικής συμμετρίας σε συγκεκριμένο παράθυρο ενδιαφέροντος

4. Για όλα τα σημεία $p_r = (x_r, y_r)$ σε ένα παράθυρο αναζήτησης $sw = [(x_{\min}, y_{\min}), (x_{\max}, y_{\max})]$ θεωρούμε ένα παράθυρο ενδιαφέροντος rw με κέντρο το p_r , διαστάσεις w_h και w_w , και κλίση a ,

και για γωνίες κλίσης a του άξονα συμμετρίας του παραθύρου ενδιαφέροντος από a_{\min} έως a_{\max} με βήμα $asterp$,

υπολόγισε τα ακραία σημεία (x_u, y_u) και (x_d, y_d) του άξονα:

$$x_u = x_r + wh \cdot \cos(a), \quad y_u = y_r + wh \cdot \sin(a)$$

$$x_d = x_r - wh \cdot \cos(a), \quad y_d = y_r - wh \cdot \sin(a)$$

υπολόγισε τα σημεία του ευθύγραμμου τμήματος $[(x_u, y_u), (x_d, y_d)]$:

$$pa[x_a, y_a] = \text{int line}((x_d, y_d), (x_u, y_u))$$

Για κάθε pa και για όλα τα συμμετρικά σημεία p_i και p_j κάθετα στον άξονα c , υπολόγισε τις συναρτήσεις:

$$\alpha = \arctan((y(j) - y(i)) / (x(i) - x(j)))$$

$$P(i, j) = 1 - \cos(|a(i) + a(j) - 2\alpha|)$$

$$C(i, j) = P(i, j) \cdot Ir(i) \cdot Ir(j)$$

6. Υπολογισμός του μέτρου ισοτροπικής συμμετρίας:

$$IS(p_r) = \sum_{(i,j) \in rw} C(i, j)$$

7. Για το σημείο p_r επιλέγουμε σαν κατεύθυνση συμμετρίας την κλίση $IP(p_r) = a$ για την οποία εμφανίζεται το μεγαλύτερο μέτρο συμμετρίας $IS(p_r)$.

8. Τέλος από όλα τα p_r επιλέγεται αυτό με το που εμφανίζει το μεγαλύτερο μέτρο συμμετρίας $IS_{\max}(p_r)$ και κατεύθυνση $IP_{\max}(p_r)$, ώστε να καθοριστεί η θέση και ο προσανατολισμός του παραθύρου ανίχνευσης της ισχυρότερης Τοπικής Ανακλαστικής Συμμετρίας.

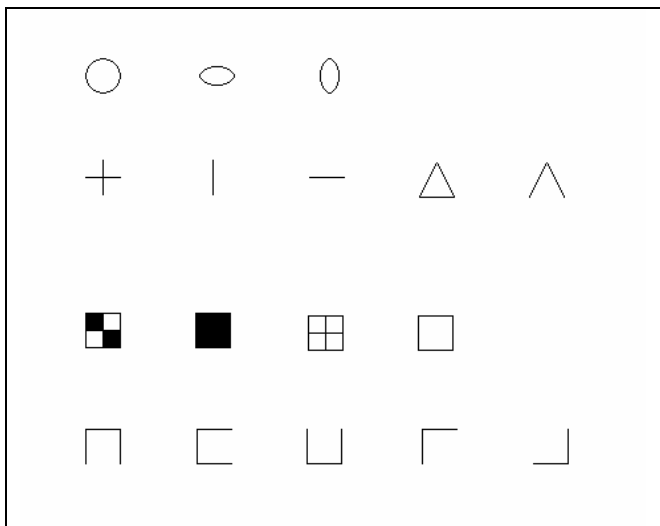
5.3 Πειραματικές Δοκιμές

Για τις πειραματικές δοκιμές χρησιμοποιήθηκε μία μικρή συλλογή από 20 εικόνες προσώπων. Αυτές οι εικόνες περιέχουν πρόσωπα στις διαστάσεις που ανιχνεύτηκαν με

την διαδικασία ανίχνευσης προσώπων. Οι εικόνες έχουν κανονικοποιηθεί σε ίδιες διαστάσεις, ώστε η οριζόντια διάσταση να είναι 300 pixels. Έχουν επιλεγεί κατά τέτοιο τρόπο ώστε να αντιπροσωπεύουν χαρακτηριστικές περιπτώσεις, με διαφορετικές στάσεις, φωτισμό και εκφράσεις και να είναι δυνατή μια πρώτη εκτίμηση της αποτελεσματικότητας των αλγορίθμων που θα δοκιμαστούν. Οι πρώτες 10 έχουν επιλεγεί από την 1^η Βάση Δεδομένων και οι άλλες 10 από την 2^η Βάση Δεδομένων.

5.3.1 Δοκιμές πάνω σε Γεωμετρικά Σχήματα

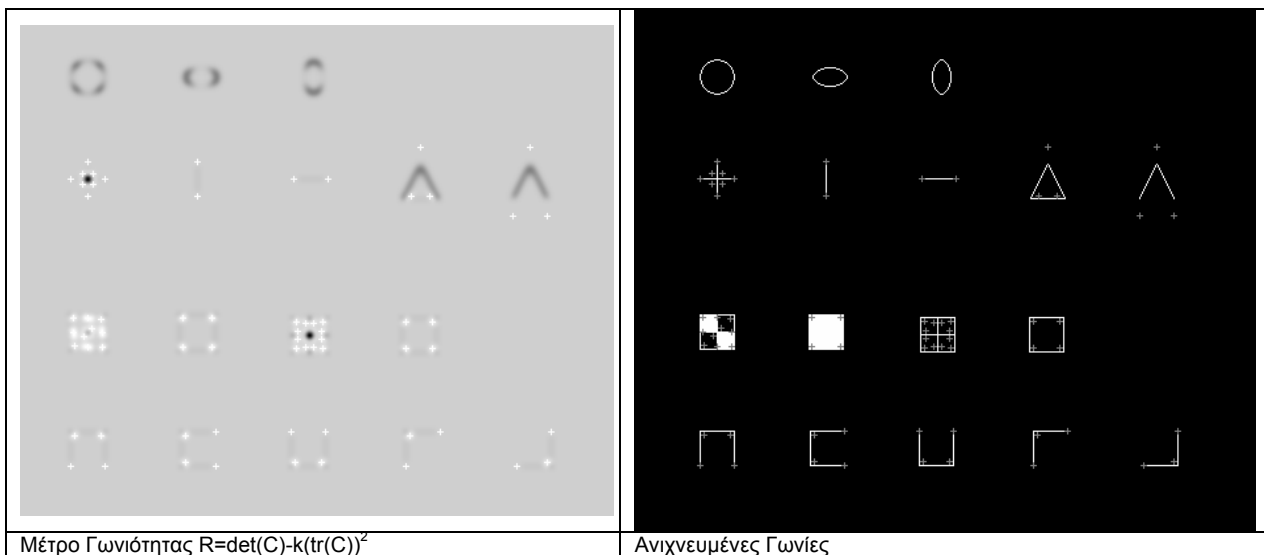
Οι πρώτες δοκιμές είχαν σαν σκοπό την εκτίμηση της ορθότητας της υλοποίησης και της συμπεριφοράς των συγκεκριμένων ανιχνευτών σε απλά γεωμετρικά σχήματα:



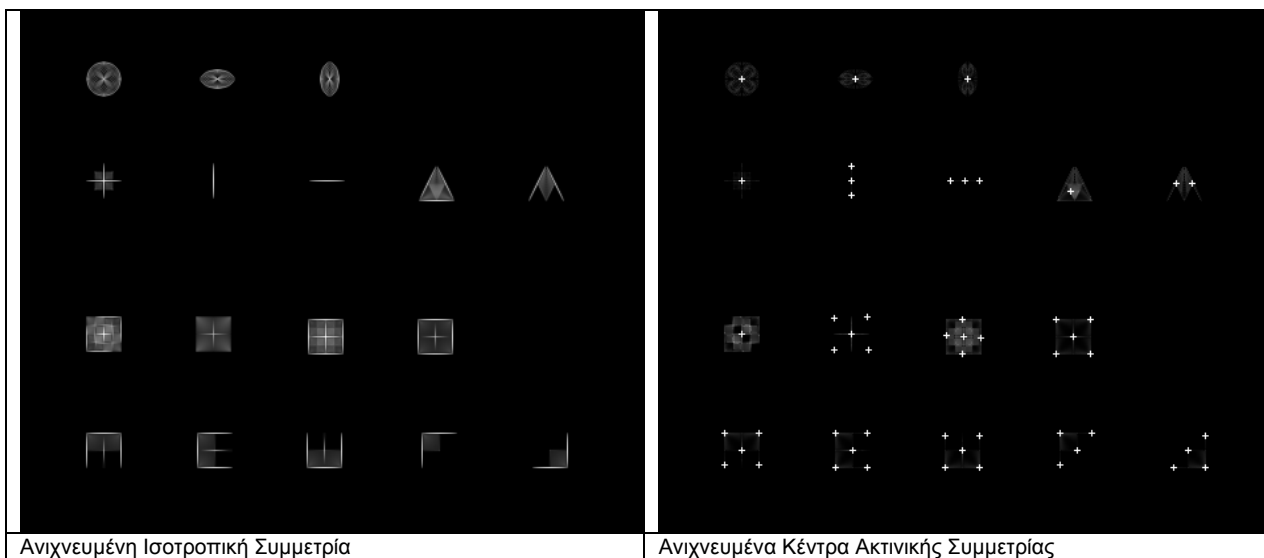
Εικ. 5.2: Γεωμετρικά σχήματα για την ανίχνευση σημείων ενδιαφέροντος

Χρησιμοποιήθηκαν τα σχήματα που εμφανίζονται στην Εικ. 5.2.

Δοκιμάστηκε αρχικά ο ανιχνευτής γωνιών του Harris με διαφορετικές τιμές στις παραμέτρους του ανιχνευτή όπως το μέγεθος του παραθύρου συνέλιξης (convolution window) $conv-win = [4 \sim 16]$, και η τυπική απόκλιση $\sigma = [1 \sim 4]$ του Γκαουσιανού φίλτρου. Τα καλύτερα αποτελέσματα φαίνονται στην Εικ. 5.3 με τιμές παραμέτρων $conv-win = 12$, και $\sigma = 3$. Οι παράμετροι αυτές εξαρτώνται από το μέγεθος της προς ανίχνευση γωνίας και την ακρίβεια στην θέση εντοπισμού της.



Εικ. 5.3: Εφαρμογή του ανιχνευτή γωνιών του Harris σε γεωμετρικά σχήματα



Εικ. 5.4: Εφαρμογή του ανιχνευτή συμμετριών του Reissfeld σε γεωμετρικά σχήματα

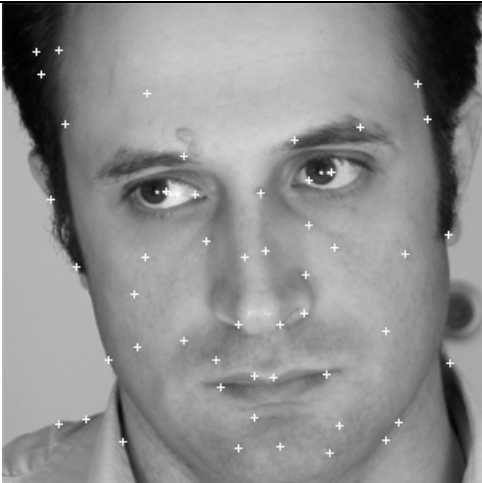
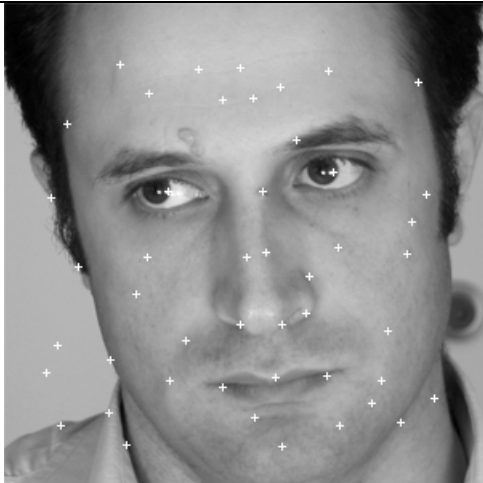
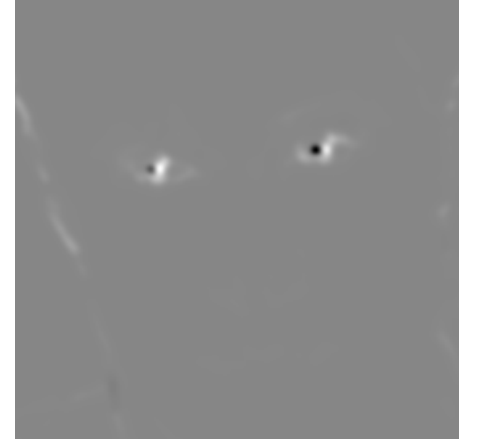
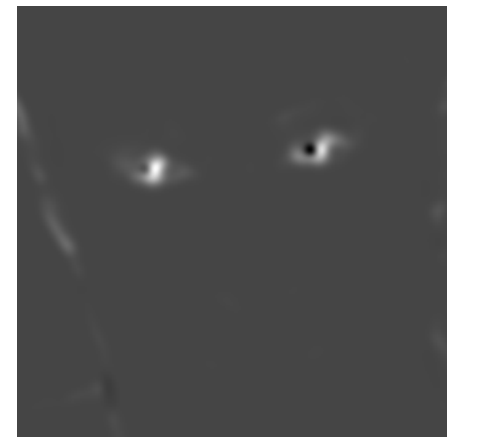
Στην συνέχεια δοκιμάστηκαν ο ανιχνευτής ισοτροπικής συμμετρίας και ο ανιχνευτής ακτινικής συμμετρίας του Reissfeld, με διαφορετικές τιμές στις παραμέτρους του ανιχνευτή, όπως το μέγεθος του παραθύρου συνέλιξης (convolution window) $conv-win = [4 \sim 32]$, και η τυπική απόκλιση $\sigma = [1 \sim 8]$ του Γκαουσιανού φίλτρου. Τα καλύτερα αποτελέσματα φαίνονται στην Εικ. 5.4 με τιμές παραμέτρων $conv-win = 16$, και $\sigma = 4$ και για τους δύο ανιχνευτές. Οι παράμετροι αυτές εξαρτώνται από το μέγεθος της περιοχής στην οποία ανιχνεύεται η συμμετρία και την ακρίβεια στην θέση εντοπισμού της.

Από τα αποτελέσματα κατέστη φανερό πως οι ανιχνευτές δουλεύουν όπως γενικά αναμενόταν, σύμφωνα με τις αναφορές που υπάρχουν στις εργασίες των Harris και Reissfeld.

5.3.2 Δοκιμές του Ανιχνευτή Γωνιών πάνω σε Πρόσωπα

Ο επόμενος κύκλος πειρ

ματισμού αφορά στον προσδιορισμό των κατάλληλων παραμέτρων για την εφαρμογή του ανιχνευτή γωνιών του Harris σε εικόνες προσώπων. Οι παράμετροι που είναι απαραίτητο να προσδιοριστούν είναι το μέγεθος του παραθύρου συνέλιξης (convolution window) conv-win, η τυπική απόκλιση σ του Γκαουσιανού φίλτρου και το μέγεθος του παραθύρου απαλοιφής των μη-μεγίστων (non-maxima suppression window) max-win. Το πλήθος των σημείων ενδιαφέροντος που ανιχνεύονται καθορίζεται ουσιαστικά από τις παραπάνω παραμέτρους.

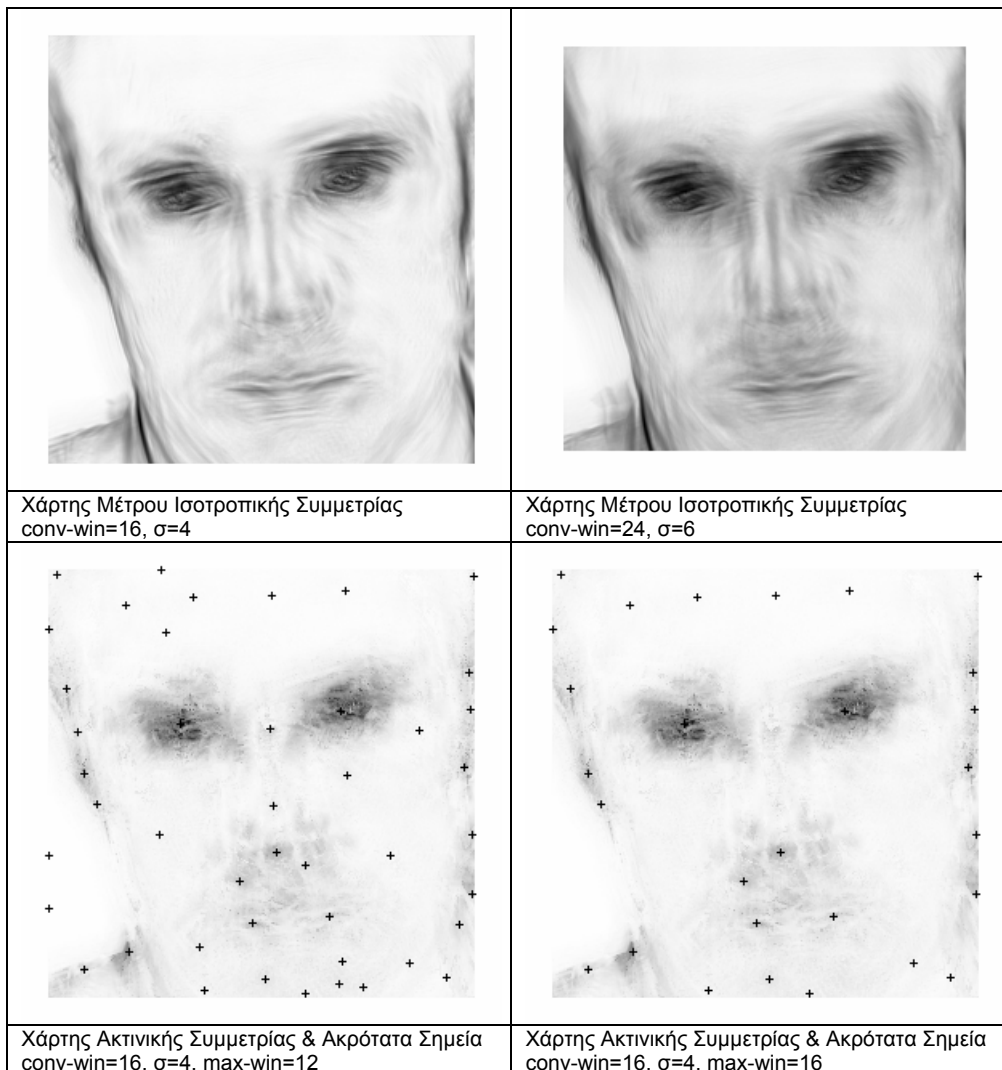
	
Ανιχνευμένες Γωνίες στην αρχική εικόνα conv-win=12, $\sigma=3$, max-win=8	Ανιχνευμένες Γωνίες στην αρχική εικόνα conv-win=16, $\sigma=4$, max-win=8
	
Χάρτης Μέτρου Γωνιότητας $R=\det(C)-k(\text{tr}(C))^2$ conv-win=12, $\sigma=3$	Χάρτης Μέτρου Γωνιότητας $R=\det(C)-k(\text{tr}(C))^2$ conv-win=16, $\sigma=4$

Εικ. 5.5: Εφαρμογή του ανιχνευτή γωνιών του Harris σε πρόσωπο

Δείγματα ανιχνευμένων γωνιών φαίνονται στην Εικ. 5.5. Καλά ανιχνευμένες γωνίες με τιμές παραμέτρων $\text{conv-win} = 12$, $\sigma = 3$ και $\text{max-win} = 8$, είναι οι γωνίες των ματιών προς την μύτη, οι γωνίες των χειλιών, οι κορυφές της μύτης και η κορυφή του πηγουνιού σε εικόνα 300 pix.

5.3.3 Δοκιμές του Ανιχνευτή Συμμετριών πάνω σε Πρόσωπα

Ο επόμενος κύκλος πειραματισμού αφορά στον προσδιορισμό των κατάλληλων παραμέτρων για την εφαρμογή του ανιχνευτή συμμετριών του Reisfeld σε εικόνες προσώπων.





Εικ. 5.6: Εφαρμογή του ανιχνευτή συμμετριών του Reissfeld σε πρόσωπο

Στον μεν ανιχνευτή ιστροπικής συμμετρίας οι παράμετροι που είναι απαραίτητο να προσδιοριστούν είναι το μέγεθος του παραθύρου συνέλιξης (convolution window) conv-win και η τυπική απόκλιση σ του Γκαουσιανού φίλτρου, ώστε να εξαχθεί ο χάρτης του μέτρου της ιστροπικής συμμετρίας.

Ο ανιχνευτής ιστροπικής συμμετρίας εμφάνισε ενδείξεις για την ύπαρξη ανακλαστικών συμμετριών, οι οποίες όμως είναι δύσκολο έως αδύνατο να ταυτοποιηθούν από τον χάρτη ιστροπικής συμμετρίας που προκύπτει (Εικ. 5.6).

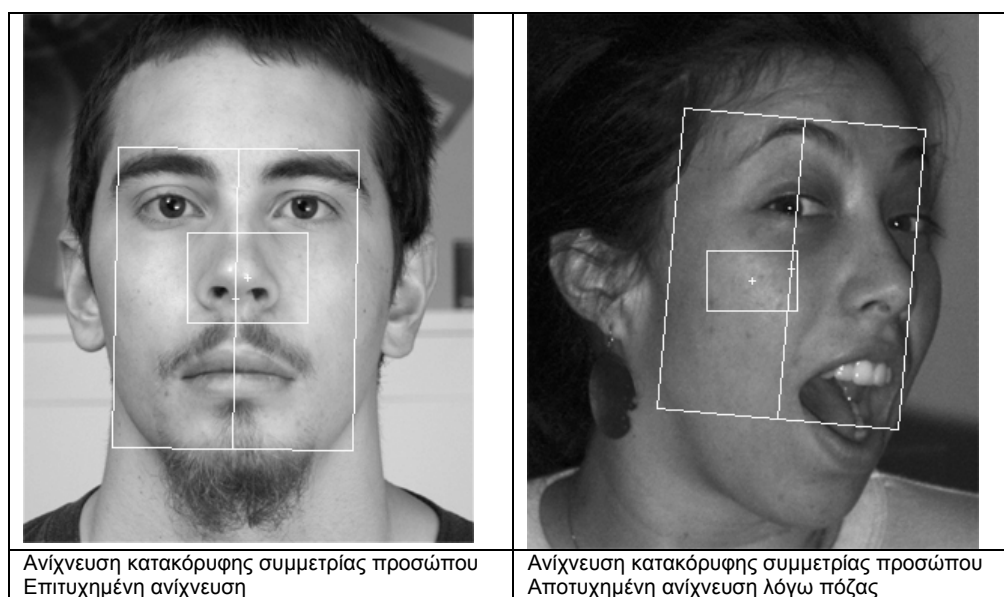
Στον δε ανιχνευτή ακτινικής συμμετρίας οι παράμετροι που είναι απαραίτητο να προσδιοριστούν είναι το μέγεθος του παραθύρου συνέλιξης (convolution window) conv-win, η τυπική απόκλιση σ του Γκαουσιανού φίλτρου και το μέγεθος του παραθύρου απαλοιφής των μη-μεγίστων (non-maxima suppression window) max-win, ώστε να εντοπιστούν τα κέντρα με ισχυρή ακτινική συμμετρία.

Δείγματα ανιχνευμένων κέντρων ακτινικής συμμετρίας φαίνονται στην Εικ. 5.6. Καλά ανιχνευμένα κέντρα ακτινικής συμμετρίας με παραμέτρους conv-win = 16, $\sigma = 4$ και max-win = 12, είναι τα κέντρα των ματιών, οι γωνίες των χειλιών, η κορυφή της μύτης και η κορυφή του πηγουνιού σε εικόνα 300 pix.

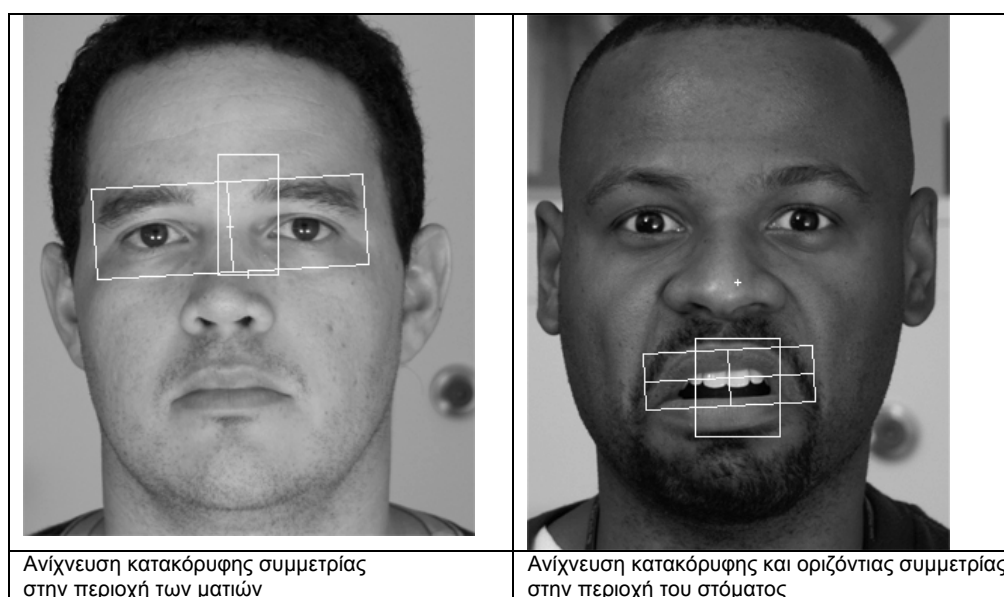
Το πρόβλημα που εμφανίζεται μετά την εφαρμογή των παραπάνω γενικών ανιχνευτών γεωμετρικών σημείων ενδιαφέροντος και που πρέπει να αντιμετωπιστεί είναι η ταυτοποίηση των παραπάνω σημείων, ώστε να αντιστοιχηθούν στα συγκεκριμένα ανατομικά σημεία στα οποία αναφέρονται.

5.3.4 Δοκιμές του Ανιχνευτή Τοπικών Συμμετριών πάνω σε Πρόσωπα

Ο τελευταίος κύκλος πειραματισμού αφορά στην εφαρμογή του αλγορίθμου προσδιορισμού των ισχυρότερων τοπικών συμμετριών πάνω σε πρόσωπα, ώστε να προσδιοριστεί ποιές συμμετρίες είναι πιο εύρωστες για να χρησιμοποιηθούν ως οδηγοί ταυτοποίησης των γεωμετρικών σημείων ενδιαφέροντος που ανιχνεύονται με τις παραπάνω γενικές μεθόδους ανίχνευσης γωνιών και κέντρων ακτινικής συμμετρίας.



Εικ. 5.7: Ανίχνευση τοπικών ανακλαστικών συμμετριών σε πρόσωπο



Εικ. 5.8: Ανίχνευση τοπικών ανακλαστικών συμμετριών σε πρόσωπο

Δοκιμάστηκαν ανιχνεύσεις διαφόρων τοπικών συμμετριών που εμφανίζει το πρόσωπο.

- Ανίχνευση της κεντρικής συμμετρίας όλου του προσώπου με κατακόρυφο άξονα (Εικ. 5.7). Επιτυχείς ανιχνεύσεις 18/20.
- Ανίχνευση της συμμετρίας στα μάτια ως προς οριζόντιο άξονα.
Επιτυχείς ανιχνεύσεις 13/20.
- Ανίχνευση της συμμετρίας στα μάτια ως προς κατακόρυφο άξονα (Εικ. 5.8).
Επιτυχείς ανιχνεύσεις 18/20.
- Ανίχνευση της συμμετρίας στο στόμα ως προς οριζόντιο άξονα.
Επιτυχείς ανιχνεύσεις 4/20.
- Ανίχνευση της συμμετρίας στο στόμα ως προς κατακόρυφο άξονα.
Επιτυχείς ανιχνεύσεις 13/20.
- Ανίχνευση της συμμετρίας στο στόμα ως προς οριζόντιο & κατακόρυφο άξονα (Εικ. 5.8). Επιτυχείς ανιχνεύσεις 12/20.
- Ανίχνευση της συμμετρίας στο αριστερό μάτι ως προς οριζόντιο & κατακόρυφο άξονα. Επιτυχείς ανιχνεύσεις 13/20.
- Ανίχνευση της συμμετρίας στο δεξί μάτι ως προς οριζόντιο & κατακόρυφο άξονα. Επιτυχείς ανιχνεύσεις 11/20.

Οι αποτυχίες οφείλονται σε μεγάλη στροφή του προσώπου (σχεδόν προφίλ) οπότε η συμμετρία εξαφανίζεται και στα εμπόδια στα μάτια, όπως τα γυαλιά.

5.4 Συμπεράσματα : Η Τελική Διαδικασία Ανίχνευσης

Από τις παραπάνω δοκιμές γίνεται φανερό πως με τις γενικές μεθόδους ανίχνευσης σημείων ενδιαφέροντος, είναι αδύνατο να αναγνωρισθούν τα ορόσημα (landmarks) που ανιχνεύονται, ώστε να αντιστοιχηθούν στα ανατομικά ορόσημα που μας ενδιαφέρουν (anatomical landmarks) και να γίνει απόδοση ετικετών σε αυτά (labelled landmarks).

Με την εφαρμογή του αλγορίθμου ανίχνευσης τοπικών ανακλαστικών συμμετριών μπορούν να προσδιοριστούν άξονες συμμετρίας σε συγκεκριμένες περιοχές των εικόνων των προσώπων, και κυρίως της κεντρικής συμμετρίας του προσώπου, της συμμετρίας της περιοχής των ματιών και της συμμετρίας της περιοχής του στόματος, οι οποίοι να χρησιμοποιηθούν ως οδηγοί για την αναγνώριση των σημείων ενδιαφέροντος που ανιχνεύονται από τις γενικές γεωμετρικές μεθόδους.

Κεφάλαιο 6

ΕΠΙΛΟΓΟΣ

Στην παρούσα εργασία εξετάστηκαν μέθοδοι για την ανίχνευση προσώπων καθώς και μέθοδοι για την ανίχνευση σημείων ενδιαφέροντος σε πρόσωπα.

Η μέθοδος ανίχνευσης προσώπων που υλοποιήθηκε βασίστηκε στις εργασίες των Viola & Jones [VIOL01a, VIOL01b] και Lienhart [LIEN02a, LIEN02b]. Οι ανιχνευτές που χρησιμοποιήθηκαν αφορούν την ανίχνευση προσώπων προσανατολισμένων σε στάση ανφάς (frontal). Περαιτέρω μελέτη μπορεί να γίνει για την ενσωμάτωση στο σύστημα και ανιχνευτών για την ανίχνευση προσώπων σε στάση προφίλ (profile). Αυτό μπορεί κάλλιστα να γίνει με την χρήση εκπαιδευμένου ταξινομητή για πρόσωπα σε στάση προφίλ και εφαρμογή της κατάλληλης μετεπεξεργασίας.

Οι μέθοδοι ανίχνευσης σημείων ενδιαφέροντος που υλοποιήθηκαν βασίστηκαν στις εργασίες των Harris & Stephens [HAR88] και Reisfeld [REIS92, REIS95] για την ανίχνευση γωνιών και συμμετριών αντίστοιχα. Επίσης υλοποιήθηκε μία παραλλαγή του ανιχνευτή συμμετριών του Reisfeld για την ανίχνευση ισχυρών τοπικών ανακλαστικών συμμετριών και του προσανατολισμού τους, ώστε να είναι εφικτή η ταυτοποίηση των γεωμετρικών σημείων ενδιαφέροντος που ανιχνεύονται με τις μεθόδους ανίχνευσης γωνιών και συμμετριών. Περαιτέρω μελέτη μπορεί να γίνει για την συγχώνευση των παραπάνω μεθόδων σε ένα ενιαίο εύρωστο σύστημα προσδιορισμού συγκεκριμένων ανατομικών χαρακτηριστικών.

Επίσης αρκετά ενδιαφέρουσα είναι και η μελέτη της χρήσης κατάλληλα εκπαιδευμένου ανιχνευτή βασισμένου στη μέθοδο Viola & Jones που έχει υλοποιηθεί, για την ανίχνευση ανατομικών χαρακτηριστικών σε συνδυασμό με τις μεθόδους ανίχνευσης γενικών γεωμετρικών χαρακτηριστικών. Η μελέτη αυτή μπορεί να στηριχθεί κάλλιστα στην εργασία των Cristinacce & Cootes [CC03].

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

ΟΡΟΛΟΓΙΑ

active shape model (ASM)	μοντέλο ενεργού μορφής
adaptive boosting (AdaBoost)	προσαρμοστική ενίσχυση
anatomical landmarks	ανατομικά ορόσημα
anchor points	σημεία αγκύρωσης
autocorrelation function	συνάρτηση αυτοσυσχετισμού
biometrics	βιομετρία
boosting	ενίσχυση
cascade	διάταξη καταρράκτη
cascade of classifiers	καταρράκτης ταξινομητών
classification	ταξινόμηση
classification and regression trees (CART)	δένδρα ταξινόμησης και παλινδρόμησης
constraint satisfaction problem (CSP)	πρόβλημα ικανοποίησης περιορισμών
contrast	αντίθεση
control points	σημεία ελέγχου
cornerness response	απόκριση γωνιότητας
decision tree	δένδρο απόφασης
deformable model	παραμορφώσιμο μοντέλο
depth image	εικόνα βάθους
edge map	χάρτης ακμών
enrollment	εγγραφή
face alignment/registration	ευθυγράμμιση προσώπου
face detection	ανίχνευση προσώπου
face gallery	πινακοθήκη/συλλογή προσώπων
face identification	ταυτοποίηση προσώπου
face localization	εντοπισμός προσώπου
face locator	εντοπιστής προσώπου
face matching	συνταίριασμα προσώπων
face probe	δοκιμαστικό πρόσωπο
face recognition	αναγνώριση προσώπου
face tracking	παρακολούθηση προσώπου
face verification/authentication	διακρίβωση/επαλήθευση προσώπου
facial features	χαρακτηριστικά προσώπου
feature	χαρακτηριστικό
feature extraction	εξαγωγή χαρακτηριστικών
feature points	χαρακτηριστικά σημεία
fiducial point	σημείο αναφοράς
frontal	κατά μέτωπο, ανφάς
geometrical landmarks	γεωμετρικά ορόσημα
gradient	βαθμίδα
gradient map	χάρτης βαθμίδας
hidden Markov model	μοντελο κρυφών μεταβλητών Markov
histogram equalization	εξισορρόπηση ιστογράμματος
inductive learning	επαγωγική μάθηση
instance	εκδοχή
integral image	εικόνα ολοκλήρωμα
intensity image	εικόνα έντασης
inter-class variability	διαφοροποιήσεις μεταξύ των κλάσεων
interest point	σημείο ενδιαφέροντος
intra-class variability	διαφοροποιήσεις μέσα στην κλάση
isotropic symmetry magnitude	μέτρο ιστροπικής συμμετρίας

key points	σημεία κλειδιά
label	ονομασία/επιγραφή/ετικέτα
labelled landmarks	καθορισμένα ορόσημα (με ετικέτα)
landmark	ορόσημο
localization accuracy	ακρίβεια εντοπισμού
manifold	πολύπτυχο
mathematical landmarks	μαθηματικά ορόσημα
multilayer perceptron	πολυστρωματικό αντίληπτρο
naïve Bayes classifier	απλοϊκός ταξινομητής Bayes
neural network	νευρωνικό δίκτυο
non-maxima suppression	απαλοιφή μη-μεγίστων
orientation	προσανατολισμός
overfitting	υπερπροσαρμογή
pattern recognition	αναγνώριση προτύπων
pixel	εικονοστοιχείο
pose	πόζα/στάση
principal component analysis	αναλυση θεμελειωδών συνιστωσών
profile	κατατομή, προφίλ
pseudo-landmarks	ψευδο-ορόσημα
radial symmetry	ακτινική συμμετρία
repeatability	επαναληψιμότητα
robustness	ευρωστία
shape	μορφή
statistical shape models	στατιστικά μοντέλα μορφών
strong classifier	ισχυρός ταξινομητής
stump	ρίζα δένδρου απόφασης
support vector machines (SVM)	μηχανές ανυσματικής στήριξης (ΜΑΣ)
symmetry direction	κατεύθυνση συμμετρίας
symmetry map	χάρτης συμμετρίας
template	υπόδειγμα
wavelet	κυματίδιο
weak classifier	αδύναμος ταξινομητής

ΑΚΡΩΝΥΜΙΑ

ASM	Active Shape Model
AdaBoost	Adaptive Boosting
BSD	Berkeley Software Distribution
CART	Classification And Regression Trees
CSP	Constraint Satisfaction Problem
xml	Extensible Markup Language
OpenCV	Open Source Computer Vision Library
SVM	Support Vector Machines

ΑΝΑΦΟΡΕΣ / ΒΙΒΛΙΟΓΡΑΦΙΑ

- [ΠΑΣ07]** Γ. Πασσαλής. *Μέθοδοι Γραφικών και Τεχνητής Όρασης για την Ανακατασκευή, Αναπαράσταση και Ανάκτηση Τρισδιάστατων Αντικειμένων με Εφαρμογή στην Βιομετρία*. Διδακτορική Διατριβή, ΕΚΠΑ 2007.
- [BKP05]** G. Bradski, A. Kaehler and V. Pisarevsky. Learning-Based Computer Vision with Intel's Open Source Computer Vision Library. In *Intel Technology Journal*, Vol. 09-Iss.02, May 2005.
- [CBL]** Computational Biomedicine Lab (CBL), University of Houston. www.cbl.uh.edu.
- [CC03]** D. Cristinacce and T. Cootes. Facial Feature Detection using Adaboost with Shape Constraints. In *Proc. 14th British Machine Vision Conference*, 2003.
- [CHEN04]** L. Chen, L. Zhang, L. Zhu, M. Li and H. Zhang. A Novel Facial Feature Localization Method Using Probabilistic-Like Output. In *Proc. Of Asian Conf. On Computer Vision*, 2004.
- [CSJ05]** D. Colbry, G. Stockman and A. Jain. Detection of Anchor Points for 3D Face Verification. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005.
- [CTB92]** I. Craw, D. Tock and A. Benett. Finding Face Features. In *Proc. Second European Conf. Computer Vision*, 1992.
- [DERP04]** K. Derpanis. *The Harris Corner Detector*. Technical Report, York University, 2004.
- [DRYD98]** I. Dryden and K. Mardia. *Statistical Shape Analysis*, Wiley 1998.
- [FORS03]** D. Forsyth and J. Ponce. *Computer Vision, A Modern Approach*, Pearson - Prentice Hall 2003.
- [FRE95]** Y. Freund and R. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting, AT&T Labs, 1995.
- [GONZ02]** R. Gonzalez and R. Woods. *Digital Image Processing*, 2nd Ed. Prentice Hall 2002.
- [HAR88]** C. Harris and M. Stephens. A Combined Corner and Edge Detector. In *Proc. 4th Alvey Vision Conf.*, 1988.
- [HIET00]** R. Hietmeyer. Biometric identification promises fast and secure processing of airline passengers. *The International Civil Aviation Organization Journal*, 55(9):10-11, 2000.
- [KAK07]** I. Kakadiaris, G. Passalis, G. Toderici, N. Murtuza, Y. Lu, N. Karampatziakis and T. Theoharis. 3D face recognition in the presence of facial expressions: an annotated deformable model approach. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007.
- [KAN73]** T. Kanade. *Picture Processing by Computer Complex and Recognition of Human Faces*. PhD Thesis, Kyoto University, 1973.
- [LI02]** 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.
- [LI05]** S. Z. Li and A. K. Jain (ed.). *Handbook of Face Recognition*, Springer 2005
- [LIEN02a]** R. Lienhart and J. Maydt. An Extended Set of Haar-like Features for Rapid Object Detection. IEEE ICIP 2002.
- [LIEN02b]** 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.
- [LJ05]** X. Lu and A. Jain. *Multimodal Facial Feature Extraction for Automatic 3D Face Recognition*. Technical Report MSU-CSE-05-22. Michigan State University, 2005.

- [LJ06] X. Lu and A. Jain. Automatic Feature Extraction for Multiview 3D Face Recognition. In *Proc. Int. Conf. on Automatic Face and Gesture Recognition*, 2006.
- [LJC06] X. Lu, A. Jain and D.Colbry. Matching 2.5D Face Scans to 3D Models. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006.
- [LTC95] A. Lanitis, C. J. Taylor and T. Cootes. An Automatic Face Identification System Using Flexible Appearance Models. In *Image and Vision Computing*, 1995.
- [NOB89] A. Noble. *Descriptions of Image Surfaces*. PhD Thesis, Oxford University, 1989.
- [OCVL] OpenCV Library. www.intel.com/research/mrl/research/opencv.
- [PAP98] C. P. Papageorgiou, M. Oren, and T. Poggio. A General Framework for Object Detection. In *Proceedings of IEEE International Conference on Computer Vision*, 1998.
- [REIS92] D. Reisfeld and Y. Yeshurun. Robust Detection Of Facial Features By Generalized Symmetry. In *Proc. of the 11th IAPR Int. Conf. on Pattern Recognition*, 1992.
- [REIS95] D. Reisfeld, H.Wolfson and Y. Yeshurun. Context Free Attentional Operators: the Generalized Symmetry Transform. In *Int. Journal of Computer Vision*, 1995.
- [ROWL98b] H. A. Rowley, S. Baluja, and T. Kanade. Neural Network-Based Face Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998.
- [RUSS03] S. Russell, P. Norvig. *Artificial Intelligence, A Modern Approach*, Pearson - Prentice Hall 2003.
- [SCHN00b] 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.
- [SMB00] C. Schmid, R.Mohr and C. Bauckhage. Evaluation of Interest Point Detectors. In *International Journal on Computer Vision*, 2000.
- [SUNG98] K-K. Sung and T. Poggio. Example-Based Learning for View-Based Human Face Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998.
- [THEOD06] S. Theodoridis, K. Koutroumbas. *Pattern Recognition*, 3d Ed. Academic Press 2006.
- [TMS05] F. Tsalakanidou, S.Malassiotis and M. Strintzis. Face Authentication Using Color and Depth Images. In *IEEE Transactions on Image Processing*, 2005.
- [VIOL01a] 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.
- [VIOL01b] P. Viola and M. Jones. Robust Real-time Object Detection. In *IEEE ICCV Workshop on Statistical and Computational Theories of Vision*, 2001.
- [YANG02] M.-H. Yang, D. Kriegman, and N. Ahuja. Detecting Faces in Images: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002.
- [ZUO04] F. Zuo and P. With. Fast Facial Feature Extraction Using a Deformable Shape Model with Haar-Wavelet Based Local Texture Attributes. In *Int. Proc. on Image Processing*, 2004.