

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

Τμήμα Διδακτικής της Τεχνολογίας και Ψηφιακών Συστημάτων

Μεταπτυχιακό Τμήμα

Κατεύθυνση Ψηφιακές Επικοινωνίες και Δίκτυα



Image - based Perceptual Services for Context-Awareness in Intelligent Environments

Λογοθέτης Μάριος

Νοέμβριος 2006

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Αφιερώνεται στους γονείς μου

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Περίληψη

Σκοπός της παρούσας εργασίας είναι η μελέτη και η υλοποίηση ενός συστήματος που θα συνειδητοποιεί το περιεχόμενο μιας πληροφορίας σε ένα έξυπνο περιβάλλον διαχείρισης. Ειδικότερα το περιεχόμενο που ουσιαστικά διαχειρίζεται το σύστημα είναι τα πλαίσια των φωτογραφιών που αποτυπώνονται σε έναν μηχανισμό κάμερας. Το σύστημα είναι σε θέση να ελέγξει, εάν υπάρχει κίνηση στο χώρο και να συνεχίσει στην ταυτοποίηση ή μη της συγκεκριμένης εικόνας. Έπειτα μέσα από μια πλατφόρμα μεταφέρει ένα μήνυμα, που στη περίπτωση μας είναι το όνομα της εικόνας, σε ένα οποιοδήποτε απομακρυσμένο σημείο.

Λέξεις Κλειδιά

ανίχνευση κίνησης, σύστημα διαχείρισης ανιχνεύσιμων εικόνων, αναγνώριση προσώπου, αλγόριθμοι αναγνώρισης προσώπου, βάσεις δεδομένων αναγνώρισης

Ευχαριστίες

Της θερμές μου ευχαριστίες εκφράζω στον επίκουρο καθηγητή κ Παναγιώτη Δεμέστιχα για τη συνεχή επίβλεψη και τη σημαντική βοήθεια που μου παρείχε από τη αρχή μέχρι το πέρας της διπλωματικής μου εργασίας.

Ειδικά θα ήθελα να ευχαριστήσω τον Κώστα Τσαγκάρη και τον Νίκο Κουτσουρή, για την υπομονή τους και τη χρήσιμη βοήθεια της σε όλη τη πορεία της εργασίας. Η συνεργασία ήταν άψογη και αυτό φαίνεται στο αποτέλεσμα.

Τέλος εκφράζω την ιδιαίτερη ευγνωμοσύνη μου της γονείς μου και σε της που ψυχολογικά βρίσκονταν δίπλα μου, για την κατανόηση της και την υποστήριξή της κατά την διάρκεια των μεταπτυχιακών σπουδών μου.

Πίνακας Περιεχομένων

Περίληψη.....	iv
Λέξεις Κλειδιά.....	iv
Ευχαριστίες.....	v
Πίνακας Περιεχομένων.....	vi
Κατάλογος Εικόνων.....	ix
Κατάλογος σχημάτων.....	x
Συνοτομογραφίες.....	xi
Εισαγωγή.....	1
1.1 Εισαγωγή.....	1
1.2 Κατανομή Εργασίας.....	2
Πρόγραμμα ανίχνευσης Motion.....	3
2.1 Γενικά.....	3
2.2 Χαρακτηριστικά του προγράμματος motion.....	4
2.3 Το project Motion.....	5
2.4 Περιβάλλον εργασίας.....	6
2.6 Τεχνολογία του Motion.....	7
2.6.1 Ανίχνευση εικόνας.....	7
2.6.2 Ανίχνευση κίνησης.....	8
2.6.3 Motion με μάσκα.....	9
2.7 Τα threads στο Motion.....	10
2.8 Παράδειγμα με τα threads.....	13
2.9 Το Config Αρχείο.....	14
2.10 Συσκευές και εγκατάσταση οδηγών.....	17
2.10.1 Video4linux (V4L) συσκευές.....	17

2.10.2 Network camera συσκευές.....	18
2.11 Μελλοντικές Βελτιώσεις.....	19
2.12 Συμπεράσματα για την εφαρμογή Motion	21
Αναγνώριση Προσώπου.....	22
3.1 Εισαγωγή.....	22
3.2 Γενικά.....	22
3.3 Τα θετικά της αναγνώρισης προσώπου.....	24
3.4 Ιστορία της αναγνώρισης προσώπου.....	26
Ιστορικά γεγονότα στην επιστήμη του Computer Vision	30
3.4.1 Το παρόν στην αναγνώριση προσώπου	31
Αναγνώριση προσώπου και ανίχνευση	32
3.5 Γνωστές πατέντες αναγνώρισης προσώπου.....	34
3.6 Εμπορικά συστήματα και εφαρμογές.....	35
3.7 Αλγόριθμοι Αναγνώρισης.....	37
3.8 Αλγόριθμοι αναγνώρισης από εικόνα.....	39
3.9 Βάσεις Δεδομένων για έλεγχο του αλγορίθμου.....	39
3.10 Εύκολα και δύσκολα σενάρια για αναγνώριση εικόνας.....	41
3.11 Λύσεις για δύσκολα σενάρια αναγνώρισης	44
3.12 Προβλήματα κατά τη διαδικασία της αναγνώρισης.....	47
3.13 Το μέλλον στην αναγνώριση προσώπου.....	53
Υλοποίηση Εφαρμογής - Αποτελέσματα.....	54
4.1 Εισαγωγή.....	54
4.2 Αρχιτεκτονική.....	55
4.3 Περιβάλλον – Βοηθητικά προγράμματα	56
4.3.1 Βιβλιοθήκη Colt.....	56
4.3.2 Jadex.....	57
4.3.3 Κάμερα.....	59

4.3.4 Classpaths.....	59
4.4 Περιγραφή δομής Demo.....	60
4.5 Ανάλυση Εφαρμογής.....	63
Εγκατάσταση του Motion	65
Συμπεράσματα	72
Εγχειρίδιο χρήσης.....	73
Βιβλιογραφικές Αναφορές.....	81

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Κατάλογος Εικόνων

2.1 Εικόνα με κίνηση	8
2.2 Εικόνα προγράμματος.....	8
2.3 Μάσκα εικόνας.....	9
2.5 Εικόνα με μάσκα.....	10
2.6 αισθητήρας κίνησης.....	19
3.1 Αναλύσεις γεωμετρικών προσεγγίσεων.....	28
3.2 Κατηγοριοποίηση αλγορίθμων.....	38
3.3 Σωστά σενάρια αναγνώρισης.....	41
3.3 Δύσκολα σενάρια αναγνώρισης.....	43
3.4 Συνεχόμενες εικόνες από βίντεο.....	45
3.5 Διαφορετικές γωνίες θέασης.....	47
3.6 Πρόσωπα με διαφορετικό φωτισμό.....	49
4.1 Το πρότυπο FIPA.....	58

Κατάλογος σχημάτων

2.1 Τοποθέτηση 3 συσκευών καταγραφής.....	13
3.1 Γνωστές Πατέντες.....	35
4.1 Αρχιτεκτονική Εφαρμογής.....	55
4.2 Λειτουργία σωστής αναγνώρισης.....	69
4.3 Λειτουργία λανθασμένης αναγνώρισης.....	70

Συντομογραφίες

Λατινικές

ASCII	<i>American Standard Code for Information Interchange</i>
JPEG	<i>Joint Photographic Experts Group</i>
PPM	<i>Portable Pixel Map</i>
MPEG	<i>Moving Picture Experts Group</i>
SMS	<i>Short Message Service</i>
XML	<i>Extensible Markup Language</i>
GPL	<i>General Public License</i>
PHP	<i>Hypertext Preprocessor</i>
HTTP	<i>Hypertext Transfer Protocol</i>
USB	<i>Universal Serial Bus</i>
PIN	<i>Personal identification number</i>
CCTV	<i>Closed-circuit television</i>

Ελληνικές

KME	Κεντρική μονάδα Επεξεργασίας
-----	------------------------------

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΚΕΦΑΛΑΙΟ 1

Εισαγωγή

1.1 Εισαγωγή

Στις μέρες μας, όλο και περισσότερο κρίνεται απαραίτητη η αναγνώριση ανθρώπινων προσώπων από μηχανές για την επίτευξη ενός συγκεκριμένου σκοπού. Αυτό για να επιτευχθεί με σωστό τρόπο, χρειάζονται δύο κύριες συνιστώσες. Ένα έξυπνο περιβάλλον καταγραφής της εικόνας του προσώπου και ένας αποδοτικός αλγόριθμος σωστής αναγνώρισης των χαρακτηριστικών του προσώπου.

Η παρούσα εργασία ασχολείται και με τις δύο αυτές συνιστώσες, προσπαθώντας να τις συνδέσει κάτω από μια ομοιογενή πλατφόρμα. Το αποτέλεσμα είναι ένα σύστημα που ουσιαστικά ενοποιεί αυτές τις διαφορετικές πτυχές, κάνοντάς το ένα ενιαίο σύστημα.

Βέβαια σε αυτό το σημείο μπαίνει ο τρόπος που έξυπνα το ίδιο το σύστημα καταγράφει τις εικόνες. Το μεγαλύτερο πρόβλημα που συναντάμε σε τέτοιου είδους υλοποιήσεις, είναι της αποθήκευσης. Τις περισσότερες φορές η καταγραφή ενός χώρου είναι χρονικά συνεχόμενη συνήθως σε εικόνες ή βίντεο, με αποτέλεσμα ο αλγόριθμος αναγνώρισης να αντιμετωπίζει σοβαρό πρόβλημα λόγω του μεγάλου όγκου δεδομένων.

Αυτό το κενό έρχεται επιστημονικά να καλύψει η συγκεκριμένη εργασία, δημιουργώντας ένα συγκεκριμένο και απλό έλεγχο. Αυτό της κίνησης. Έτσι μόνο όταν έχουμε σημαντική κίνηση με έναν μηχανισμό καταγραφής βίντεο ή εικόνας, αποθηκεύεται η συγκεκριμένη εικόνα, η οποία με τη σειρά της αναγνωρίζεται από τον αλγόριθμο.

Έπειτα το ίδιο το σύστημα ταυτοποιώντας την αναγνώριση όπως θα δούμε και στη συνέχεια τρεις φορές στέλνει σε ένα άλλο σύστημα το όνομα της εικόνας.

1.2 Κατανομή Εργασίας

Η εργασία έχει συγκεκριμένη και σαφή δομή, ως προς το περιεχόμενό της. Αρχικά αναλύονται οι δύο τεχνολογίες που χρησιμοποιούνται κυρίως στην υλοποίηση. Έτσι το κεφάλαιο 2 και 3 τις αναφέρουν αναλυτικά .

Έπειτα ακολουθεί το κεφάλαιο της εξήγησης της εφαρμογής και τέλος το παράρτημα που φαίνονται με εικόνες οι τρόποι καθοδήγησης για την υλοποίηση.

ΚΕΦΑΛΑΙΟ 2

Πρόγραμμα ανίχνευσης Motion

2.1 Γενικά

Το motion είναι ένα πρόγραμμα που παρακολουθεί μία ή περισσότερες κάμερες, που είναι συνδεδεμένες στον υπολογιστή και είναι ικανό να ανιχνεύσει εάν ένα συγκεκριμένο κομμάτι της εικόνας που απεικονίζεται έχει αλλάξει. Με άλλα λόγια μπορεί να ανιχνεύσει την κίνηση.

Η κάμερα μπορεί να συνδέεται στον υπολογιστή είτε μέσω της θύρας USB (Universal Serial Bus) είτε μέσω καλωδίου Ethernet στην περίπτωση που είναι δικτυακή κάμερα. Και στις δύο περιπτώσεις, ρυθμίζοντας τις κατάλληλες ιδιότητες το motion λειτουργεί χωρίς κανένα πρόβλημα.

Το πρόγραμμα είναι γραμμένο με γλώσσα C και δημιουργήθηκε για τα λειτουργικά συστήματα αρχιτεκτονικής Linux που κυκλοφορούν στην αγορά όπως Fedora, Suse, Mandrake και άλλα. Το συγκεκριμένο πρόγραμμα δεν μπορεί να λειτουργήσει σε περιβάλλοντα windows.

Το πρόγραμμα ελέγχεται μόνο από την γραμμή εντολών, δεν έχει καθόλου γραφικό περιβάλλον και εκτελείται σαν διεργασία (daemon), τραβώντας μικρή υπολογιστική ισχύ. Όλες οι ιδιότητες παραμετροποιούνται είτε από γραμμή εντολών είτε μέσω ειδικών αρχείων ASCII (config files) που πολύ απλά αλλάζουν με έναν ASCII editor. Πρέπει να επισημάνουμε ότι για να ενεργοποιηθούν οι ιδιότητες θα πρέπει να έχουν γίνει οι αλλαγές πριν την εκτέλεση του προγράμματος.

Το motion είναι ικανό κατά την διαδικασία του να εξάγει τα εξής αρχεία

- εικόνες τύπου jpeg
- εικόνες τύπου ppm και
- ακολουθίες video τύπου mpeg

Αυτά τα αρχεία αποθηκεύονται σε φάκελο τοπικό ή απομακρυσμένο ανάλογα με τις ιδιότητες που έχουμε δημιουργήσει στο αρχείο ASCII.

Μέσα στο συγκεκριμένο αρχείο ιδιοτήτων μπορούμε να παραμετροποιήσουμε αρκετά πράγματα που έχουν να κάνουν με την κάμερα, τον τρόπο αποθήκευσης και άλλα σημαντικά στοιχεία που θα δούμε παρακάτω.

2.2 Χαρακτηριστικά του προγράμματος motion

Μερικά από τα κυριότερα χαρακτηριστικά του προγράμματος motion είναι τα εξής

- Μπορεί να πάρει στιγμιαίες εικόνες (snapshots) κατά την διάρκεια της κίνησης
- Συνδέονται πολλές συσκευές κάμερας την ίδια στιγμή
- Έχει δυνατότητα να μεταδίδει live streaming με ειδικό λογισμικό
- Μπορεί να υποθηκεύσει βίντεο σε format mpeg με τις βιβλιοθήκες από το ffmpeg
- Έχει την ικανότητα να εκτελέσει συγκεκριμένες ενέργειες, όταν ανιχνεύεται κίνηση, όπως να στείλει sms ή ακόμα και email.

- Μπορεί να εκτελέσει συγκεκριμένα προγράμματα – εφαρμογές του λογισμικού σε συγκεκριμένες καταστάσεις όπως όταν παρουσιαστεί κίνηση καθώς και όταν αποθηκευτεί μια εικόνα ή πλήθος αυτών.
- Μπορεί να ακολουθήσει την πορεία της κίνησης εάν η camera το υποστηρίζει (motion tracking)
- Δυνατότητα εικόνας real-time με ειδικούς οδηγούς.
- Καταγράφει τα γεγονότα που του ορίζουμε σε βάσεις δεδομένων MySQL και PostgreSQL
- Ελέγχεται από εφαρμογές web
- Έλεγχος από μέσω φυλλομετρητή (browser) με χρήση xml-rpc

2.3 Το project Motion

Το motion είναι ένα project τύπου open source και έχει άδεια GPL (GNU GENERAL PUBLIC LICENSE) version 2. Αυτό σημαίνει ότι κάποιος μπορεί να εγκαταστήσει και να χρησιμοποιήσει το πρόγραμμα εντελώς δωρεάν, χωρίς την άδεια του δημιουργού. Στην περίπτωση όμως που κάποιος το χρησιμοποιήσει για εμπορικούς λόγους ή αλλάξει τον κώδικα του προγράμματος, τότε υποχρεούται να διαβάσει τους όρους της άδειας.[1]

Υπεύθυνος του όλου project είναι ο Kenneth Lavrsen και αρχικά το πρόγραμμα συντάχτηκε από τον Jeroen Vreeken που εξακολουθεί να συμμετέχει στην ομάδα του project. Μέχρι και σήμερα (Νοέμβριος 2006) το πρόγραμμα έχει φτάσει στην έκδοση 3.2.x

2.4 Περιβάλλον εργασίας

Όπως αναφέραμε και πριν το motion είναι μια εφαρμογή που δημιουργήθηκε αρχικά για περιβάλλοντα αρχιτεκτονικής linux και μόνο. Από την έκδοση όμως 3.1.15 και μετά με την βοήθεια του Angel Carpintero στο project του Motion μπορεί να γίνει εγκατάσταση σε περιβάλλον FreeBSD, χωρίς ακόμα να υποστηρίζονται όλες οι δυνατότητες του προγράμματος. [2]

Επίσης μετά την έκδοση 3.2.4 είναι δυνατή η εγκατάσταση σε MacOSX με τις ίδιες δυνατότητες όπως στο FreeBSD.

2.5 Επεκτασιμότητα

Το Motion αρχικά δημιουργήθηκε με μικρή δυνατότητα επέκτασης, είτε σε θέματα διαχείρισης των αρχείων που έβγαζε, είτε σε θέματα γραφικού περιβάλλοντος. Και αυτό για να είναι ελαφρύ σαν πρόγραμμα μιας και δεν έπρεπε να καταναλώνει πόρους από το σύστημα. Όμως πάνω στη συγκεκριμένη εφαρμογή πολλοί παρουσίασαν επεκτάσεις ή ακόμα και μικρά προγράμματα που εκμεταλλευόντουσαν την κυρίως εφαρμογή για να γίνει πιο εύχρηστη και πιο λειτουργική.

Μερικές επεκτάσεις είναι να μπορεί να συνδεθεί και μικρόφωνο για να έχουμε και καταγεγραμμένο ήχο ή ακόμα και ολόκληρο interface με γλώσσα php για καλύτερη διαχείριση δεδομένων και επιλογών κάμερας.

Πολλές από αυτές τις επεκτάσεις δημιουργήθηκαν και από το project του Motion, αλλά και από άτομα που απλά ήθελαν να επεκτείνουν τις δυνατότητες του.[4]

2.6 Τεχνολογία του Motion

2.6.1 Ανίχνευση εικόνας

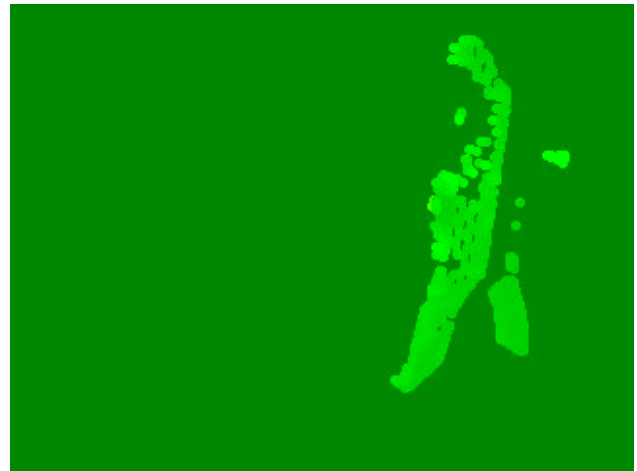
Το motion ανιχνεύει την κίνηση συγκρίνοντας την εικόνα που καταγράφει εκείνη την στιγμή με μία άλλη εικόνα που έχει αρχικά αποθηκευτεί, που λέγεται εικόνα αναφοράς. Προφανώς κάθε φορά που γίνεται και μια αλλαγή στην εικόνα και έχουμε ουσιαστικά κίνηση η εικόνα αναφορά αλλάζει για να είναι μέτρο σύγκρισης με τις υπόλοιπες εικόνες. Ουσιαστικά όταν δεν υπάρχει κίνηση και θόρυβος, η διαφορά μεταξύ της καινούργιας εικόνας και της εικόνας αναφοράς θα πρέπει να είναι μηδενική.

Εάν υπάρχει διαφορά στην εικόνα το αποτέλεσμα προφανώς θα είναι διαφορετικό. Το πρόγραμμα για να προστατεύσει το ενδεχόμενο ότι ο θόρυβος απλά της εικόνας θα την κάνει διαφορετική, χρειάζεται συγκεκριμένο αριθμό διαφορών και μεγεθών στην ένταση της διαφοράς πριν την δήλωση της κίνησης.

Στην παρακάτω εικόνα 2.1 βλέπουμε ουσιαστικά μια κίνηση από έναν άνθρωπο. Πριν μπει ο άνθρωπος μέσα στο πλαίσιο της εικόνας το motion είχε δημιουργήσει την όλη εικόνα σαν ένα πράσινο φόντο. Με το που μπήκε ο άνθρωπος όπως βλέπουμε και από την εικόνα 2.2 άλλαξε το πράσινο χρώμα στα σημεία που είχαμε κίνηση.



2.1 Εικόνα με κίνηση



2.2 Εικόνα προγράμματος

Μετά την διαδικασία της πρώτης ανίχνευσης, εικόνα αναφοράς γίνεται το τελευταίο πλαίσιο (εικόνα 2.2). Έτσι με την αλλαγή του πλαισίου το πρόγραμμα ξανακαταλαβαίνει την διαφορά και βγάζει και άλλες εικόνες. Αυτό σταματάει όταν η εικόνα αναφοράς είναι ίδια με την εικόνα που καταγράφει εκείνη την ώρα η κάμερα.

2.6.2 Ανίχνευση κίνησης

Από όλα τα pixels που έχουν αλλάξει μεταξύ της καινούργιας εικόνας και της εικόνας αναφοράς υπολογίζονται οι μέσοι όροι των τιμών των αξόνων x και y . Αυτό το σημείο θα χρησιμοποιηθεί σαν το κέντρο της κίνηση για την συγκεκριμένη εικόνα. Έτσι μπορούμε κάθε φορά να ξέρουμε σε πιο σημείο της εικόνας παρατηρείται η κίνηση. Σε αυτό το πλαίσιο που δημιουργείται η διαφορά στα pixels, μπορούμε κάθε φορά να είμαστε σίγουροι ότι είναι πραγματική κίνηση και όχι θόρυβος, εάν οι διαφορές είναι πάνω από ένα συγκεκριμένο κατώφλι.

2.6.3 Motion με μάσκα

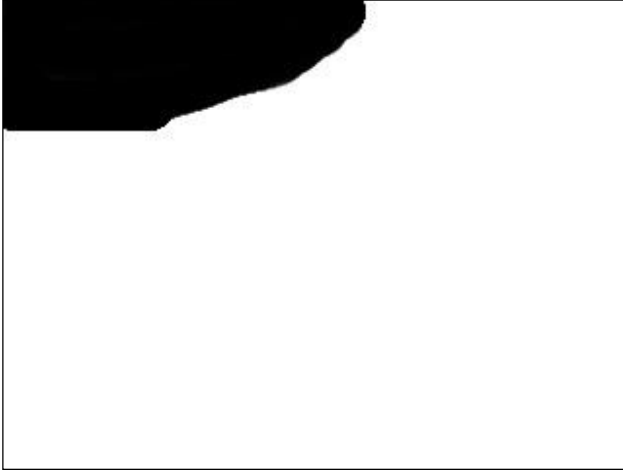
Εάν έχουμε ένα συγκεκριμένο σημείο της εικόνας όπου δεν θέλουμε το motion να ανιχνεύσει την κίνηση, τότε μπορούμε να χρησιμοποιήσουμε μια εικόνα σαν μάσκα για να την φιλτράρουμε. Αν η μάσκα είναι σε άσπρο χρώμα τότε χρησιμοποιούνται για αναγνώριση τα pixel που την αντιστοιχούν. Στο μαύρο χρώμα συμβαίνει το αντίθετο και δεν επιτυγχάνεται η διαδικασία της αναγνώρισης.

Παραδείγματος χάριν η εικόνα 2.3 είναι μια μάσκα που μπορούμε να χρησιμοποιήσουμε για να μην ελέγξουμε συγκεκριμένο τομέα πάνω στην εικόνα και αυτός είναι πάνω αριστερά όπου και είναι μαύρου χρώματος.



2.3 Μάσκα εικόνας

Οπότε στην περίπτωση που έχουμε την εικόνα 2.4 με την παραπάνω μάσκα και ο συγκεκριμένος άνθρωπος διανύσει το μονοπάτι είναι σίγουρο ότι το motion θα δείξει ότι έχουμε κίνηση μιας και η κίνηση συμβαίνει σε διαφορετικό κομμάτι του μαύρου φόντου.



2.5 Εικόνα με μάσκα

2.7 Τα threads στο Motion

Το motion είναι μια περίπλοκη εφαρμογή με πάρα πολλά χαρακτηριστικά. Ένα από τα δύσκολα ζητήματα του είναι τα threads που αρχίζουν όταν ξεκινάει το πρόγραμμα. Ακολουθεί μια περιγραφή με τα threads.

- **Κύριο Thread.** Όταν το πρόγραμμα Motion ξεκινάει το κύριο thread διαβάζει το config αρχείο και ξεκινάει όλα τα threads παιδιά του (child threads). Μόλις αυτό γίνει το κύριο thread περιμένει μέχρι τα άλλα threads που ξεκίνησε να σταματήσουν.
- **Thread Ελέγχου.** Αυτό το thread ξεκινάει την http επαφή ελέγχου. Όταν συνδεόμαστε με την πόρτα ελέγχου του Motion, αυτό είναι το thread με το οποίο επικοινωνούμε. When you read and set option value the control thread simply reads and writes to the data structures that the camera threads use.

- Thread κάμερας. Για κάθε κάμερα που είναι συνδεδεμένη στον υπολογιστή, ξεκινάει ένα thread. Αυτό το thread ελέγχει ότι έχει να κάνει με την σύλληψη (capture) των βίντεο και των εικόνων από την κάμερα, τον εντοπισμό της κίνησης, την αποθήκευση των εικόνων, την δημιουργία και αποθήκευση των βίντεο, τον έλεγχο των συνδέσεων της κάμερας και την διαδικασία αποστολής εικόνων jpeg σε πελάτες που θέλουν τις εικόνες. Όπως φαίνεται αυτό το thread αποτελεί το κυρίως και εμπεριέχει πάρα πολλές λειτουργίες. Στην περίπτωση που η κάμερα είναι δικτυακή το συγκεκριμένο thread έχει και ένα βοηθητικό. Αυτό χρειάζεται για να ρωτάει το ένα thread το άλλο για το επόμενο διαθέσιμο πλαίσιο.

- Thread δικτυακής κάμερας. Όταν η κάμερα είναι τύπου δικτύου, τότε αυτό το ειδικό thread δημιουργείται. Ο βασικός τρόπος της διαδικασίας του Motion είναι ότι το κεντρικό thread της κάμερας τραβάει μια εικόνα από το οδηγό (video4linux) όταν ο τελευταίος αποφασίσει να το κάνει κάτω από τις συνθήκες που του έχουμε ορίσει. Αυτό με μια δικτυακή κάμερα δεν λειτουργεί, επειδή, αυτό το είδος της κάμερας είναι εφοδιασμένη με ένα mjpeg stream. Mjpeg είναι ένα είδος server που παίρνει την εικόνα ή βίντεο και με μια επικοινωνία που δημιουργεί στέλνει τα δεδομένα. Σε αυτή τη περίπτωση η δικτυακή κάμερα ελέγχει ουσιαστικά όλη την ροή. Έτσι το thread το συγκεκριμένο συνέχεια συλλέγει εικόνες από το mjpeg stream και τα αποθηκεύει τοπικά σε ένα είδος αποθηκευτικού χώρου (buffer). Σε αυτή τη φάση δύο πράγματα μπορούν να συμβούν .

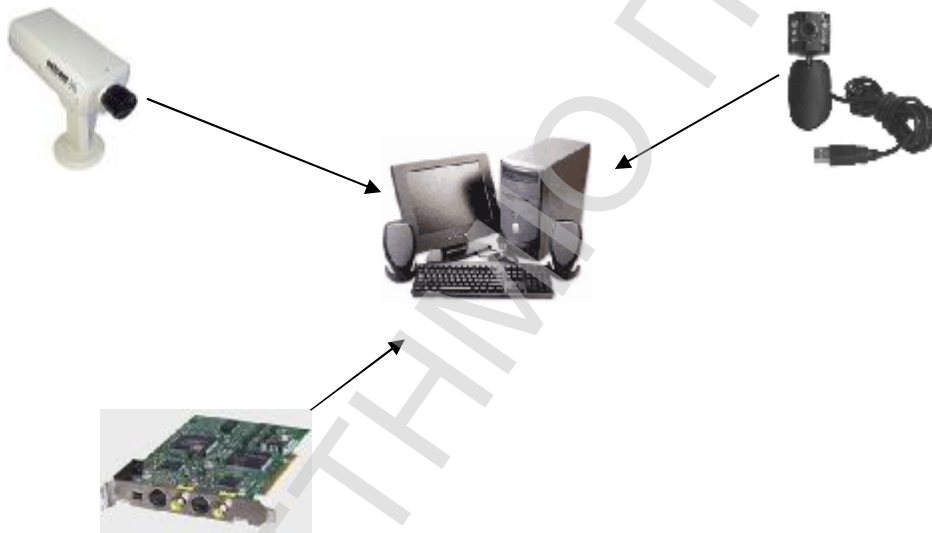
1. Το thread του Motion να είναι πιο γρήγορο από την δικτυακή κάμερα. Τότε το thread του Motion θα ζητήσει από την κάμερα για μια φωτογραφία. Εάν μια καινούργια εικόνα δεν έχει ακόμα φτάσει από την τελευταία φορά που το thread της κάμερας ζήτησε για εικόνα, τότε θα περιμένει το τελευταίο για το thread της δικτυακής κάμερας, που θα πάρει το πλαίσιο από την ίδια την κάμερα.

2. Η δικτυακή κάμερα είναι πιο γρήγορη από ότι το Motion χρειάζεται. Το thread της δικτυακής κάμερας λαμβάνει πλαίσια εικόνων και τα θέτει σε σειρά. Όταν μια καινούρια εικόνα φτάνει, η παλιά που υπάρχει φεύγει. Έτσι όταν το thread της κάμερας ζητήσει για μια εικόνα, τότε του έρχεται η τελευταία εικόνα.

Στην περίπτωση που το Motion ξεκινάει ένα εξωτερικό πρόγραμμα το κύριο thread της κάμερας δημιουργεί μια επιπλέον διαδικασία για την δημιουργία ενός thread παιδιού. Αυτή η ξεχωριστή διαδικασία κληρονομεί όλη την δομή που υπάρχει μέχρι εκείνο το σημείο και διαχωρίζεται εντελώς από το κυρίως πρόγραμμα του Motion και δημιουργεί ένα φλοιό (shell) με τον οποίο εκτελεί το εξωτερικό πρόγραμμα.

2.8 Παράδειγμα με τα threads

Αρχικά το motion ξεκινάει. Υπάρχει ένα motion αρχείο (motion.conf) και 3 αρχεία thread (thread config files) ένα για μια USB κάμερα, για μια κάρτα σύλληψης (capture card) και ένα για μια δικτυακή κάμερα, όπως φαίνεται και στο παρακάτω σχήμα. Ακολουθούν τα threads που θα ξεκινήσουν με χρονολογική σειρά



2.1 Τοποθέτηση 3 συσκευών καταγραφής

- Main thread
- Κάμερα 1 thread (USB cam)
- Κάμερα 2 thread (bttv driver cam)
- Κάμερα 3 thread (netcam)
- Κάμερα 3 network helper thread (netcam)
- Control thread (http control)

Όπως βλέπουμε ξεκινάει το κύριο thread και μετά ανάλογα με τις συσκευές που διαθέτουμε, δημιουργούνται και τα αντίστοιχα thread. Βλέπουμε ότι μόνο στη δικτυακή κάμερα έχει δημιουργηθεί ένα βοηθητικό για το σκοπό που είπαμε και πιο πάνω.

2.9 Το Config Αρχείο

Όπως αναφέραμε και πιο πριν το Motion ελέγχεται μόνο από την γραμμή εντολών, δεν έχει καθόλου γραφικό περιβάλλον και εκτελείται σαν διεργασία (daemon).

Όλες οι ιδιότητες παραμετροποιούνται είτε από γραμμή εντολών είτε μέσω ειδικών αρχείων ASCII (config files) που πολύ απλά αλλάζουν με έναν ASCII editor. Πρέπει να επισημάνουμε ότι για να ενεργοποιηθούν οι ιδιότητες θα πρέπει να έχουν γίνει οι αλλαγές πριν την αρχική εκτέλεση του προγράμματος.

Σε αυτό το σημείο πρέπει να αναφέρουμε τι εννοούμε λέγοντας ότι το Motion εκτελείται σαν διεργασία (daemon). Daemon για τα λειτουργικά συστήματα αρχιτεκτονικής linux είναι μια διεργασία που έχει φτιαχτεί για να τρέχει αυτόνομα, με μικρή έως καθόλου επέμβαση του χρήστη.

Συνήθως περιμένει στο παρασκήνιο (background), εκτελώντας προκαθορισμένες λειτουργίες για να επιτύχει μια εργασία. Παραδείγματος χάριν ο apache web server αποτελεί μία διεργασία (daemon) που περιμένει και ακούει σε συγκεκριμένες πόρτες (ports) για να εξυπηρετήσει ανάλογα με την αίτηση.

Το όνομα του αρχείου πρέπει να είναι *motion.conf*. Όταν ξεκινάει για πρώτη φορά το πρόγραμμα Motion ψάχνει για το ειδικό αρχείο με το όνομα *motion.conf* με την ακόλουθη σειρά.

1. Στο φάκελο που υπάρχει το Motion και καλείται
2. Όταν υπάρχει ένας φάκελος που ονομάζεται '.motion' στους φακέλους που χρησιμοποιούν οι χρήστες όταν δεν είναι διαχειριστές. πχ `/home/marios/.motion/motion.conf`
3. Στο φάκελο που ορίζεται στο `--sysconfdir=DIR` όταν τρέχουμε το `.configure` κατά τη διάρκεια της εγκατάστασης του προγράμματος, όπως θα δούμε και παρακάτω. Εάν αυτή η επιλογή δεν αλλάξει η εξ ορισμού διαδρομή θα είναι η `/usr/local/etc/`.

Το Motion έχει ένα configuration αρχείο μέσα στο πακέτο εγκατάστασης που ονομάζεται *motion-dist.conf*. Όταν κάνουμε εγκατάσταση αυτό το αρχείο αντιγράφεται στη διαδρομή `/usr/local/etc`

Αυτό το αρχείο χρειάζεται να μετονομαστεί από *motion-dist.conf* σε *motion.conf*. Το αρχικό αρχείο ονομάζεται *motion-dist.conf* έτσι ώστε το αρχείο που δουλεύει με το πρόγραμμα *motion.conf* να μην αλλάξει σε περίπτωση που το Motion ξαναεγκατασταθεί ή αναβαθμιστεί σε νεότερη έκδοση.

Σε περίπτωση που έχουμε παραπάνω από μια κάμερες, δεν χρειάζεται να ξεκινήσουμε το Motion επιπλέον φορές. Το πρόγραμμα μπορεί να χειριστεί τις επιπλέον κάμερες χωρίς κανένα εντελώς πρόβλημα και το επιδιώκει δημιουργώντας έναν αριθμό από `thread config` αρχεία, όπως είδαμε και στην προηγούμενη παράγραφο.

Στην περίπτωση που έχουμε μια κάμερα θέλουμε μόνο το ένα αρχείο *motion.conf*. Όταν έχουμε δύο ή παραπάνω κάμερες τότε δημιουργείται ένα `thread config` αρχείο για κάθε κάμερα ανεξάρτητα από το αρχικό *motion.conf* αρχείο.

Κάθε ρύθμιση που τροποποιείται στο αρχικό motion.conf αρχείο ισχύει και για όλα τα άλλα αρχεία που αντιστοιχούν στις κάμερες, εκτός και αν για την ίδια ρύθμιση υπάρχει στο αντίστοιχο config αρχείο μια διαφορετική παράμετρος.

Το πρόγραμμα Motion διαβάζει τις παραμέτρους με την ακόλουθη σειρά. Πρέπει να τονίσουμε ότι εάν η ίδια παράμετρος υπάρχει σε δύο διαφορετικά μέρη τότε ισχύει αυτή που διαβάζεται τελευταία.

1. Το Motion διαβάζει το configuration αρχείο motion.conf από την αρχή μέχρι το τέλος κάθε γραμμής.
2. Εάν η επιλογή "thread" καθορίζεται στο αρχείο motion.conf, τότε τα αρχεία thread config διαβάζονται.
3. Το Motion συνεχίζει και διαβάζει τη συνέχεια του αρχείου motion.conf. Οποιαδήποτε επιλογή από δω και πέρα θα απορρίψει αντίστοιχη ίδια επιλογή που ήταν γραμμένη στο αρχείο thread config .
4. Το Motion τέλος διαβάζει τη γραμμή εντολών και πάλι απορρίπτει οποιαδήποτε ιδιότητα ορίζεται και έχει ήδη γραφτεί προηγουμένως στα αρχεία config.

2.10 Συσκευές και εγκατάσταση οδηγών

Προτού να μπορέσουμε να ξεκινήσουμε την διαδικασία πρέπει να ξέρουμε μερικά βασικά για τη φωτογραφική μας μηχανή και ποιοι είναι οι σωστοί οδηγοί εγκατάστασης. Υπάρχουν ουσιαστικά δύο βασικοί τύποι για τη σύνδεση σε περιβάλλοντα linux.

- Video4linux συσκευές
- Network camera συσκευές

Στην πρώτη περίπτωση έχουμε όλες τις συσκευές που συνδέονται άμεσα στον υπολογιστή, ενώ στη άλλη έχουμε αυτές που χρησιμοποιούν μια web διεπαφή.

2.10.1 Video4linux (V4L) συσκευές

Πρέπει να εγκαταστήσουμε τη φωτογραφική μηχανή μας με το σωστό οδηγό. Με το που γίνει αυτή η διαδικασία τότε οι συσκευές θα πάρουν συγκεκριμένα ονόματα από το λειτουργικό, ανάλογα με το πόσες είναι, δηλαδή /dev/video0, /dev/video1, /dev/video2

Στο λειτουργικό FreeBSD, η ονομασία που δίνει το ίδιο το λειτουργικό είναι διαφορετική, και έτσι έχουμε /dev/bktr0.

Οι φωτογραφικές μηχανές ή κάρτες τηλεόρασης συνήθως δεν υποστηρίζουν οποιοδήποτε μέγεθος εικόνων. Πρέπει να γνωρίζουμε ποιο πλαίσιο υποστηρίζει, τόσο σε πλάτος όσο και σε ύψος. Εάν δεν γνωρίζουμε αυτά τα χαρακτηριστικά, τότε ξεκινάμε με πλάτος 320 και ύψος 240 που οι περισσότερες φωτογραφικές μηχανές και κάρτες τηλεόρασης το υποστηρίζουν.

Πρέπει επίσης να επιλέξουμε και σωστό ρυθμό απόκτησης πλαισίου, δηλαδή σωστό framerate. Όσο πιο γρηγορότερα δεχόμαστε εικόνες από τη φωτογραφική μηχανή, τόσο περισσότερο χρησιμοποιούμε τη κεντρική μονάδα επεξεργασία (ΚΜΕ). Η μικρότερη τιμή είναι δύο. Όπως θα δούμε και παρακάτω εμείς για τη δικιά μας εφαρμογή χρησιμοποιήσαμε τιμή τρία.

Επίσης καλό είναι να γνωρίζουμε εάν η φωτογραφική μηχανή υποστηρίζει την αυτόματη φωτεινότητα. Οι περισσότερες φωτογραφικές μηχανές το υποστηρίζουν αυτό το χαρακτηριστικό. Εάν δεν γνωρίζουμε, δεχόμαστε ότι έχει και δεν χρησιμοποιούμε το χαρακτηριστικό γνώρισμα autobrightness που υπάρχει στο config αρχείο. Τουλάχιστον δεν χρειάζεται να αρχίσει η εφαρμογή με αυτό..

Οι φωτογραφικές μηχανές που συνδέονται στον υπολογιστή μέσω θύρας USB, συνήθως καταναλώνουν μεγάλο εύρος δεδομένων. Μεγαλύτερο πρόβλημα έχουμε στην έκδοση 1.1 όπου ήταν ανέφικτο να υπάρχουν πολλές κάμερες σε USB hub. Καλό είναι σε περιπτώσεις πολλών καμερών να χρησιμοποιηθούν κάρτες επέκτασης USB, για να μην υπάρχει πρόβλημα στη διαδικασία λήψης των εικόνων.

2.10.2 Network camera συσκευές

Στην περίπτωση που έχουμε κάμερα με διαδικτυακή διεπαφή, τότε όλες οι ρυθμίσεις ορίζονται από την κύρια μικρή εφαρμογή που συνήθως έχει από μόνη της η συσκευή. Χρειάζεται μεγάλη προσοχή γιατί οι ρυθμίσεις του config αρχείου σε τέτοιου είδους κάμερα δεν εφαρμόζονται.

2.11 Μελλοντικές Βελτιώσεις

Μια μελλοντική βελτίωση που θα μπορούσε να έχει το πρόγραμμα Motion είναι ένας αισθητήρας κίνησης, όπως φαίνεται παρακάτω που θα μπορούσε να ενεργοποιήσει μια πηγή φωτός σε έναν χώρο. Με αυτό τον τρόπο λύνεται το πρόβλημα που υπάρχει, όταν γίνεται σύλληψη βίντεο ή εικόνων σε περιπτώσεις χαμηλού φωτισμού. Έτσι με το που κινείται κάτι σε έναν χώρο, ενεργοποιείται η πηγή φωτός μέσω του αισθητήρα και έχουμε την λειτουργία του Motion.



2.6 αισθητήρας κίνησης

Φυσικά, μπορεί να γίνει ακόμα πιο αυτοματοποιημένη η διαδικασία στην περίπτωση που συνδεθεί ο αισθητήρας με το Motion. Έτσι με το που ενεργοποιείται το φως αμέσως στέλνεται σήμα για να καταγραφεί εκείνη τη στιγμή ο χώρος, χωρίς να περιμένουμε από μόνο του το πρόγραμμα να καταλάβει την κίνηση.

Μια άλλη σημαντική βελτίωση, ως προς την λειτουργικότητα της εφαρμογής θα μπορούσε να είναι η καλύτερη διαχείριση των αποθηκευτικών αρχείων, είτε είναι αυτό φωτογραφίες είτε είναι βίντεο. Θα μπορούσε να υπήρχε η δυνατότητα τα αρχεία να σβήνονται από μόνα τους μετά από κάποιο χρονικό διάστημα που θα όριζε ο χρήστης. Έτσι για περιπτώσεις όπου ο υπολογιστής δεν χρησιμοποιείται για αρκετό διάστημα να μην υπάρχει πρόβλημα στον αποθηκευτικό χώρο. Επίσης, καλό θα ήταν να κρατάει τις φωτογραφίες που υπάρχει η μεγαλύτερη διαφορά στα pixels περισσότερο χρόνο από τις αντίστοιχες που υπήρχε μικρή διαφορά αλλά είχαμε καταγραφή.

Αυτό μας είναι χρήσιμο στις περιπτώσεις που η κίνηση προέρχεται από πηγές που ουσιαστικά δεν μας ενδιαφέρουν, όπως φύλλα και άλλα μικρά αντικείμενα.

Θα μπορούσαμε ακόμα να προσθέσουμε την λειτουργία της καταγραφής ήχου. Αυτό εφόσον η κάμερα μας φυσικά το επιτρέπει ή ακόμα και η σύνδεση της με ένα εξωτερικό μικρόφωνο. Μάλιστα θα μπορούμε ο ήχος να αποτελέσει κίνητρο για ενεργοποίηση της κάμερας και έτσι να έχουμε αποθηκευμένα βίντεο με ήχο.

2.12 Συμπεράσματα για την εφαρμογή Motion

Το συγκεκριμένο πρόγραμμα Motion, παρέχει μια πολύ καλή λύση για το τεράστιο πρόβλημα που απασχολεί αυτή τη στιγμή τη βιομηχανία των υπολογιστών και δεν είναι άλλο από τη συλλογή των τεράστιων σε όγκο δεδομένων. Με σχετικά μικρή διαδικασία το πρόγραμμα διώχνει τα περιττά για αυτό στοιχεία και καταγράφει ουσιαστικά μόνο τα γεγονότα που όντως χρειαζόμαστε. Και δεν είναι άλλα από αυτά που περιέχουν την κίνηση. Τα αποτελέσματα είναι εκπληκτικά εάν φανταστούμε τη διαφορά με μια κάμερα που καταγράφει συνεχώς έναν χώρο. Οπότε κερδίζουμε και σε όγκο δεδομένων, αλλά το σημαντικότερο είναι ότι κερδίζουμε σε χρόνο, μιας και δεν καθόμαστε να επιλέξουμε κομμάτια που μας ενδιαφέρουν, αλλά έχουμε κατευθείαν αυτά που είχαν κίνηση.

Μελλοντικά είναι σίγουρο ότι θα αποτελέσει βασικό στοιχείο παρακολούθησης χώρου και ότι κάποια στιγμή θα ενσωματωθεί σε συστήματα κάμερας, χωρίς την παρουσία ενός υπολογιστή.

ΚΕΦΑΛΑΙΟ 3

Αναγνώριση Προσώπου

3.1 Εισαγωγή

Την τελευταία δεκαετία, η αναγνώριση προσώπου έγινε δημοφιλής τομέας έρευνας στο πεδίο του computer vision και μια από τις πιο επιτυχημένες εφαρμογές στον τομέα της ανάλυσης και της κατανόησης προσώπου. Εξαιτίας όμως του προβλήματος που προκύπτει λόγω της φυσικής αταξίας του προσώπου, δεν ασχολήθηκε μόνο ο τομέας της επιστήμης των υπολογιστών, αλλά και οι τομείς που έχουν ιατρική σχέση με το πρόσωπο, όπως νευροεπιστήμονες ακόμα και ψυχολόγοι. Πολλές από τις έρευνες που πραγματοποιήσαν και οι δύο τομείς, βοήθησε αντίστοιχα τον άλλο στην διαδικασία της κατανόησης του ανθρώπινου προσώπου.

3.2 Γενικά

Η αναγνώριση προσώπου είναι ένα μέρος μια μεγάλης ομάδας έρευνας, της τεχνολογίας αναγνώρισης. Η αναγνώριση και ιδιαίτερα η αναγνώριση προσώπου καλύπτει μια μεγάλη γκάμα δραστηριοτήτων που συναντάμε καθημερινά. Επιπλέον η αναγνώριση προσώπου είναι κάτι που σε γενικές γραμμές ο άνθρωπος είναι καλός στο να τη κάνει και αυτό μας βοήθησε στο τομέα της επιστήμης των υπολογιστών.

Βασικά η αναγνώριση ενός προσώπου και η αναγνώριση της κίνηση του σε φυσικές δραστηριότητες χρειάζεται βήματα για να επιτευχθεί. Και αυτά θα πρέπει να ακολουθηθούν με μεγάλη προσοχή για την επίτευξη του σκοπού μας.

Φάσεις Αναγνώρισης

- *Απόκτηση εικόνας (Acquisition)*
- *Ομαλοποίησης (Normalization)*
- *Αναγνώρισης (Recognition)*

Επιγραμματικά αναφέρουμε ότι αρχικά χρειαζόμαστε την εικόνα του προσώπου από τη φωτογραφία, είτε η τελευταία πρόκειται για φωτογραφία που εστιάζει μόνο στο πρόσωπο είτε σε ολόκληρο το σώμα. Οπότε καταλαβαίνουμε ότι η απόκτηση από μια συσκευή του πλαισίου του προσώπου σε ένα δυναμικό περιβάλλον είναι βασική διαδικασία και χωρίς αυτή δεν μπορεί να προχωρήσει καμία λειτουργία. Αυτό το στάδιο λέγεται απόκτηση (acquisition)

Το επόμενο στάδιο είναι της ομαλοποίησης (normalization) και εξαρτάται κατά βάση από το πρώτο στάδιο. Στην περίπτωση που το πλαίσιο έχει καταγράψει ένα πρόσωπο σε πλάγια θέση τότε αυτό το στάδιο επιδιορθώνει με αλγόριθμους και δημιουργεί την εικόνα σαν να τραβήχτηκε η φωτογραφία από μπροστά. Ενώ στη περίπτωση που το πλαίσιο τραβήχτηκε έτσι ώστε να φαίνεται το πρόσωπο ολόκληρο τότε αυτό δεν δημιουργεί πρόβλημα στον αλγόριθμο και μπορεί να γίνει εύκολα η αναγνώριση. Έτσι κατανοούμε ότι η δυσκολία αυτού του σταδίου εξαρτάται από την σωστή διαδικασία απόκτησης του πλαισίου της εικόνας. Θα πρέπει να τονίσουμε ότι η επιτυχία του αλγορίθμου αναγνώρισης εξαρτάται βέβαια και από τον ίδιο αλλά και από τον αλγόριθμο ομαλοποίησης της φωτογραφίας. Όσο πιο μεγάλη είναι η επιτυχία στη σωστή απεικόνιση του προσώπου, τόσο πιο πολλές πιθανότητες έχει το στάδιο της αναγνώρισης.

Τέλος συναντάμε το στάδιο της αναγνώρισης, που έχουμε την απεικόνιση και τη δημιουργία του μοντέλου από κοινές εικόνες και τη συσχέτιση των νέων εικόνων με γνωστά μοντέλα.

3.3 Τα θετικά της αναγνώρισης προσώπου

Λαμβάνοντας υπόψη την απαίτηση της κοινωνίας μας για την ταυτότητα των ανθρώπων και την εξακρίβωση των στοιχείων τους, η προφανής ερώτηση είναι ποια τεχνολογία είναι καταλληλότερη για να παρέχει αυτές τις πληροφορίες. Είναι πολλοί οι τρόποι που μπορούν οι άνθρωποι να προσδιορίσουν ο ένας τον άλλον, και έτσι συμβαίνει και για τους υπολογιστές. Στις μέρες μας υπάρχουν πολλές διαφορετικές τεχνολογίες προσδιορισμού της ταυτότητας, πολλές από τις οποίες ήταν σε εμπορική χρήση για χρόνια.

Οι πιο κοινές μέθοδοι επαλήθευσης και προσδιορισμού προσώπων είναι σήμερα τα συστήματα που χρησιμοποιούν κωδικό πρόσβασης / PIN γνωστός ως προσωπικός αριθμός αναγνώρισης (Personal Identification Number). Το πρόβλημα με αυτό ή άλλες παρόμοιες τεχνικές είναι ότι δεν είναι μοναδικοί οι αριθμοί, και είναι δυνατό για κάποιον να το ξεχάσει εύκολα ή ακόμα και να κλέψει κάποιος άλλος. Προκειμένου να υπερνικηθούν αυτά τα προβλήματα, έχει αναπτυχθεί ένα ιδιαίτερο ενδιαφέρον για τα "βιομετρικά" συστήματα προσδιορισμού, τα οποία χρησιμοποιούν τις τεχνικές αναγνώρισης σχεδίων ή μοντέλων για να προσδιοριστούν κατάλληλα οι άνθρωποι χρησιμοποιώντας τα χαρακτηριστικά γνωρίσματά τους.

Μερικές από αυτές τις μεθόδους είναι τα δακτυλικά αποτυπώματα και η αναγνώριση του αμφιβληστροειδούς και των ιριδών. Όμως αυτές οι συγκεκριμένες τεχνικές δεν είναι εύρηστες. Παραδείγματος χάριν στις συναλλαγές στις τράπεζες και κατά την είσοδο μέσα σε ασφαλείς περιοχές, τέτοιες τεχνολογίες έχουν το μειονέκτημα ότι είναι αδιάκριτες και φυσικά και κοινωνικά. Ο χρήστης πρέπει να τοποθετήσει το σώμα του

σχετικά στην ίδια ευθεία με τον αισθητήρα, και έπειτα πρέπει να σταματήσει για ένα δευτερόλεπτο για να δηλωθεί ο ίδιος. Αυτό δεν σημαίνει ότι η αναγνώριση προσώπου δεν χρειάζεται το συγκεκριμένο προσδιορισμό θέσης. Όπως θα δούμε και μετά η θέση του προσώπου και η εμφάνιση παίζει σημαντικό ρόλο στην διαδικασία της αναγνώρισης.

Ενώ η μικρή χρονική διακοπή και η παρούσα αλληλεπίδραση είναι χρήσιμες στα συστήματα υψηλής-ασφάλειας, είναι ακριβώς το αντίθετο από αυτό που απαιτείται κατά την δημιουργία ενός καταστήματος που αναγνωρίζει τους καλύτερους πελάτες του, ή ένα περίπτερο πληροφοριών που θυμάται αυτούς που το επισκέπτονται, ή ένα σπίτι που γνωρίζει τους ανθρώπους που κατοικούν εκεί.

Η αναγνώριση προσώπου από την καταγραφή βίντεο και φωνής έχει μια φυσική θέση σε όλα αυτά τα έξυπνα περιβάλλοντα της επόμενης γενιάς, καθώς μπορούν να λειτουργούν στο περιθώριο, είναι συνήθως παθητικοί, δεν περιορίζουν τη μετακίνηση των χρηστών και τέλος είναι και φτηνοί με χαμηλή κατανάλωση ισχύος. Όμως το σημαντικότερο γεγονός σε όλα αυτά είναι ότι οι άνθρωποι προσδιορίζουν τους άλλους ανθρώπους από το πρόσωπο και τη φωνή τους, επομένως είναι πιθανό να είναι άνετοι με αυτά συστήματα που χρησιμοποιούν τέτοιου είδους αναγνώριση. [5]

Τα θετικά της αναγνώρισης σε σχέση με άλλα βιομετρικά μέσα :

- Ακρίβεια
- Οικονομικώς αποδοτικό μέσο
- Εξοικειωμένη χρήση
- Μη καταπατητικό ως προς τη χρήση
- Χρησιμοποιεί νόμιμα δεδομένα
- Συχνά είναι μόνος κατάλληλος βιομετρικός τρόπος

- Δεν απαιτεί τη συμμετοχή χρηστών
- Χτισμένο στον ανθρώπινο μηχανισμό αναγνώρισης
- Δεν απαιτείται κανένα ειδικό ή δαπανηρό υλικό
- Υψηλότερη κοινωνική αποδοχή
- Επιβολή νόμου
- Ασφάλεια

3.4 Ιστορία της αναγνώρισης προσώπου

Η αναγνώριση προσώπου αναδεικνύεται από τη στιγμή που η μηχανή άρχισε να γίνεται όλο και περισσότερο "ευφυής" και είχε την μεγάλη δυνατότητα να συμπληρώσει, να διορθώσει ή ακόμα και να βοηθήσει στην έλλειψη ανθρώπινων δυνατοτήτων και αισθήσεων.

Το θέμα της αναγνώρισης προσώπου είναι τόσο παλιό όσο και η επιστήμη του computer vision λόγω της πρακτικής σημασίας του θέματος και του θεωρητικού ενδιαφέροντος από τη γνωστική επιστήμη. Η αναγνώριση προσώπου δεν είναι η μόνη μέθοδος αναγνώρισης ανθρώπων.

Ακόμη και οι άνθρωποι μεταξύ τους χρησιμοποιούν τις αισθήσεις τους προκειμένου να αναγνωριστούν. Οι μηχανές έχουν ένα ευρύτερο φάσμα για αναγνωρίσουν ανθρώπους, όπως η χρήση με τα δακτυλικά αποτυπώματα, ή ακόμα πιο πολύπλοκα με τις ανιχνεύσεις ιριδών. Παρά το γεγονός ότι αυτές οι μέθοδοι προσδιορισμού μπορούν να είναι ακριβέστερες με επιτυχία που αγγίζει το 99%, η αναγνώριση προσώπου έχει παραμένει πάντα μια σημαντική εστίαση της έρευνας λόγω της μη καταπατητικής φύσης της και επειδή είναι η αρχική μέθοδος των ανθρώπων για τον προσδιορισμό προσώπων.

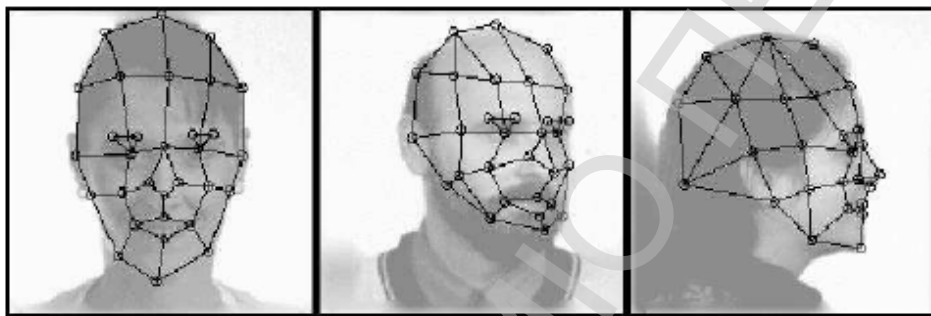
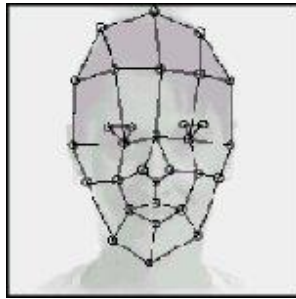
Από την έναρξη εκείνου του τομέα της τεχνολογίας που ασχοληθεί με αυτό το γεγονός υπήρξαν δύο κύριες προσεγγίσεις. Οι δύο κύριες προσεγγίσεις στην αναγνώριση προσώπου είναι

- γεωμετρική προσέγγιση
- εικονογραφική προσέγγιση

Η γεωμετρική προσέγγιση, όπως διακρίνεται και στις παρακάτω εικόνες χρησιμοποιεί τη χωρική διαμόρφωση των χαρακτηριστικών γνωρισμάτων του προσώπου. Αυτό σημαίνει ότι τα κύρια γεωμετρικά χαρακτηριστικά γνωρίσματα του προσώπου όπως τα μάτια, η μύτη και το στόμα καταγράφονται αρχικά και έπειτα τα πρόσωπα ταξινομούνται βάσει των διάφορων γεωμετρικών αποστάσεων και των γωνιών μεταξύ των χαρακτηριστικών γνωρισμάτων.

Αφ' ετέρου, η εικονογραφική προσέγγιση χρησιμοποιεί τα πρότυπα (templates) των χαρακτηριστικών γνωρισμάτων του προσώπου. Αυτή η μέθοδος χρησιμοποιεί τα πρότυπα των σημαντικότερων χαρακτηριστικών γνωρισμάτων του προσώπου καθώς και ολόκληρου προσώπου για να εκτελέσει την αναγνώριση στις μετωπικές απόψεις των προσώπων. Δηλαδή και τα χαρακτηριστικά όπως μέγεθος ματιών ή χειλιών αλλά και ολόκληρου του προσώπου όπως σχήμα ή και μέγεθος. Πολλά από τα προγράμματα που βασίζονται σε αυτές τις δύο προσεγγίσεις έχουν μερικές κοινές επεκτάσεις που χειρίζονται διαφορετικό τρόπο το τεχνικό υπόβαθρο.

Εκτός από αυτές τις δύο τεχνικές έχουμε και άλλες πιο μεταγενέστερες προσεγγίσεις που βασίζονται στα πρότυπα-μοντέλα, και οι οποίες διαμορφώνουν τα πρότυπα από την κλίση της εικόνας, και την προσεγγιστική ανάλυση των κύριων τμημάτων του προσώπου, οι οποίες μπορούν να χαρακτηριστούν ως sub-optimal προσέγγιση προτύπων. Τελικά έχουμε την προσέγγιση της παραμόρφωσης προτύπων (deformable) που συνδυάζει τα στοιχεία και της εικονογραφικής αλλά και τα χαρακτηριστικά γνωρίσματα της γεωμετρικής και έχει εφαρμοστεί στα πρόσωπα σε διαφορετικές καταστάσεις θέσης και έκφρασης.



3.1 Αναλύσεις γεωμετρικών προσεγγίσεων

Από την πρώτη έναρξη της έρευνας στην αναγνώριση προσώπου υπάρχει μια ισχυρή σχέση και μια σύνδεση με την επιστήμη των νευρικών δικτύων (Neural networks). Το πιο πρόωρο διάσημο παράδειγμα ενός συστήματος αναγνώρισης προσώπου, που χρησιμοποιεί τα νευρικά δίκτυα είναι το πρότυπο Kohonen.

Αυτό το σύστημα ήταν ένα απλό νευρικό δίκτυο που ήταν σε θέση να εκτελέσει την αναγνώριση προσώπου για τις ευθυγραμμισμένες και ομαλοποιημένες εικόνες προσώπου. Ο τύπος δικτύου που υιοθέτησε, υπολόγισε μια περιγραφή προσώπου, προσεγγίζοντας τους eigenvectors της μήτρας αυτοσυσχέτισης της εικόνας προσώπου. Αυτά τα eigenvectors είναι τώρα γνωστά και ως "eigenfaces".[6]

Μετά από αυτό το αρχικό γεγονός υπήρξαν άλλες περισσότερες μέθοδοι όπου αναπτύχθηκαν, βασισμένες στις παλαιότερες τεχνικές. Εάν θέλουμε να συνοψίσουμε τις

μεθόδους όπου η "ιδέα" της αναγνώρισης του προσώπου βασίστηκε, έχουμε την γεωμετρική προσέγγιση, την εικονογραφική προσέγγιση, και μετά από αυτήν έχουμε τις μεθόδους όπως τα eigenfaces, την ανάλυση κύριων τμημάτων (Principal Component Analysis-PCA), ή άλλες μεθόδους που επεξεργάζονται τις εικόνες σε συνδυασμό με τα νευρικά δίκτυα ή άλλα έμπειρα συστήματα. Όλες τις μεθόδους και τους αλγόριθμους θα τα δούμε αναλυτικά στο παρακάτω κεφάλαιο.

Στο στάδιο της αναγνώρισης του προσώπου, η είσοδος στοιχείο συγκρίνεται με όλα τα επιλεγμένα μοντέλα-πρότυπα του κάθε προσώπου. Για να συγκρίνει την αρχική εικόνα εισόδου με μια συγκεκριμένη άποψη του πρότυπου-μοντέλου, το πρόσωπο πρώτα γεωμετρικά ευθυγραμμίζεται με την πρότυπη άποψη. Μια σχετική μετατροπή εφαρμόζεται στην εικόνα είσοδος για να φέρει τα χαρακτηριστικά γνωρίσματα του προσώπου σε αντιστοιχία με αυτά του πρότυπου-μοντέλου.

Μια τεχνική η οποία βασίζεται στην οπτική ροή μεταξύ της μετασχηματισμένης εικόνας και του προτύπου χρησιμοποιείται για να αντισταθμίσει οποιοδήποτε μικρό μετασχηματισμό έχει μείνει μεταξύ των δύο προτύπων. Τα πρότυπα από το μοντέλο συγκρίνονται έπειτα με την εικόνα χρησιμοποιώντας τον ομαλοποιημένο συσχετισμό (normalised correlation). Και οι πρότυπες και οι εισαγμένες εικόνες προεπεξεργάζονται με έναν διαφορετικό αλγόριθμο όπως είπαμε και πιο πάνω.

Στο μέλλον, οι ερευνητές προγραμματίζουν το σύστημα έτσι ώστε αυτό να εξετάζει το πρόβλημα μόνο όταν η άποψη του προσώπου είναι διαθέσιμη. Το κλειδί για την παραγωγή αυτής της εργασίας θα είναι ένα σύστημα που θα μαθαίνει από μόνο του να χρησιμοποιεί τις πολλαπλές εικόνες των προσώπων προτύπων που υποβάλλονται στις αλλαγές στην θέση, με τη βοήθεια των νευρικών δικτύων. Το σύστημα θα εφαρμόσει αυτήν την γνώση για να συνθέσει τις νέες εικονικές απόψεις του προσώπου του μοντέλου.

Ας μείνουμε όμως λίγο στο πρώτο διάσημο σύστημα. Το σύστημα του Kohonen δεν ήταν μια πρακτική επιτυχία, λόγω της ανάγκης μετά για ακριβής ευθυγράμμιση και την κανονικοποίηση. Περνώντας τα χρόνια πολλοί ερευνητές δοκίμασαν τα σχέδια αναγνώρισης προσώπου βασισμένα στις άκρες, τις αποστάσεις των χαρακτηριστικών γνωρισμάτων, και άλλες νευρικές καθαρές προσεγγίσεις του προσώπου. Ενώ αρκετές μέθοδοι ήταν επιτυχείς στις μικρές βάσεις δεδομένων των ευθυγραμμισμένων εικόνων, κανένας δεν εξέτασε επιτυχώς το ρεαλιστικότερο πρόβλημα των μεγάλων βάσεων δεδομένων όπου η θέση και η κλίμακα του προσώπου είναι άγνωστες. Μερικές μεγάλες βάσεις δεδομένων θα δούμε σε επόμενο κεφάλαιο.

Έπειτα το 1989 ο Kirby και ο Sirovich εισήγαγαν έναν αλγεβρικό χειρισμό που κατέστησε πιο εύκολο να υπολογιστούν άμεσα τα eigenfaces, και έδειξαν ότι λιγότεροι από 100 eigenfaces χρειαζόντουσαν για να κωδικοποιήσουν ακριβώς τις προσεκτικά ευθυγραμμισμένες και ομαλοποιημένες εικόνες ενός προσώπου.

Ο Turk και ο Pentland το 1991 κατέδειξαν ότι το υπόλοιπο λάθος όταν γινόταν η κωδικοποίηση χρησιμοποιώντας τα eigenfaces μπορούσε να χρησιμοποιηθεί τόσο για να ανιχνεύσει τα πρόσωπα σε ακατάστατες εικόνες, και για να καθορίσει την ακριβή θέση και την κλίμακα των προσώπων σε μια εικόνα. Έπειτα κατέδειξαν ότι με την ένωση αυτής της μεθόδου για να διακρίνουν και να καθορίσουν τα πρόσωπα με τη μέθοδο της αναγνώρισης eigenface, θα μπορούσε να επιτύχει πιο αξιόπιστη, σε πραγματικό χρόνο αναγνώριση των προσώπων σε ένα ελάχιστο περιορισμένο περιβάλλον. Αυτή η επίδειξη με τις δύο τεχνικές, πολύ απλά θα μπορούσε να συνδυαστεί σε πραγματικό χρόνο για να δημιουργήσει ένα χρήσιμο σύστημα αναγνώρισης προσώπου.

Ιστορικά γεγονότα στην επιστήμη του Computer Vision

- **50's** - Η επιστήμη του Computer Vision γεννήθηκε: Πρωτόγονες δοκιμές στα εικονοστοιχεία (pixel) που χρησιμοποιούν τις τεχνικές λείανσης και ανίχνευσης ακρών

- **60's/70's** – Άρχισαν οι άλλες επιστήμες να βοηθούν όπως η στατιστική, η γραμμική άλγεβρα και οι αλγόριθμοι
- **70's** - Οι μελέτες αναγνώρισης προσώπου άρχισαν
- **80's** - Σημαντική ανακάλυψη του αλγορίθμου Eigenface. Ανακαλύπτονται χαρακτηριστικά διανύσματα και τα οποία δημιουργούν συνδυασμούς. Η θεωρία είναι ότι κάθε πρόσωπο αποτελείται από τις παραλλαγές των χαρακτηριστικών γνωρισμάτων.
- **90's** - Οπτική αναγνώριση χαρακτήρα (Optical Character Recognition (OCR)) που προωθείται για τις καταναλωτικές εφαρμογές όπως την σάρωση και την αποστολή φαξ.
- **00's** - Δημιουργείται μια μεγαλύτερη ώθηση για τον τομέα του Computer Vision που αναπτύσσεται στις περιοχές ασφάλειας

3.4.1 Το παρόν στην αναγνώριση προσώπου

Με τη γρήγορη εξέλιξη της τεχνολογίας και της εμπορευματοποίησης των τεχνολογικών επιτευγμάτων, η αναγνώριση προσώπου έγινε όλο και πιο περισσότερο δημοφιλής, όχι μόνο για την έρευνα αλλά και για τη χρήση στα συστήματα ασφάλειας.

Αυτό το γεγονός έδωσε το κινητήριο έναυσμα σε πολλούς ερευνητές, και επίσης στις επιχειρήσεις προκειμένου να αναπτυχθούν οι τεχνικές για αυτοματοποιημένη αναγνώριση προσώπων που θα έβρισκε πολλές εφαρμογές, συμπεριλαμβανομένης της ασφάλειας και της ανθρώπου-υπολογιστή αλληλεπίδρασης.

Παραδείγματος χάριν, μια μηχανή που αναγνωρίζει πρόσωπα θα μπορούσε να επιτρέψει τον αυτοματοποιημένο έλεγχο πρόσβασης για τα κτήρια ή να ενεργοποιήσει έναν υπολογιστή για να αναγνωρίσει το πρόσωπο που βρίσκεται μπροστά στην κονσόλα. Τα περισσότερα υπάρχοντα συστήματα αναγνώρισης προσώπου, εντούτοις, λειτουργούν μόνο για τις μετωπικές ή σχεδόν μετωπικές εικόνες των προσώπων. Με την αναγνώριση

των προσώπων κάτω από διαφορετικές γωνίες θέασης, το σύστημα δεν είναι τόσο αξιόπιστο όπως είχαμε δει και πιο πριν.

Αναγνώριση προσώπου και ανίχνευση

Όπως αναφέρεται και παραπάνω η αναγνώριση προσώπου είναι μια τεχνική αναγνώρισης αλλά δεν είναι απαραίτητο να παραμείνει ο χρήστης σταθερός προκειμένου να ληφθεί μια εικόνα. Αν και υπάρχει ένα πρόβλημα με την αναγνώριση όταν η θέση του προσώπου είναι διαφορετική, ειδικότερα, υπάρχει ένα όριο στις περιστροφές του προσώπου, οι οποίες περιλαμβάνουν κινήσεις προς τα αριστερά ή και δεξιά ή ακόμα και πάνω ή κάτω.

Η ίδια η αναγνώριση προσώπου είναι δύσκολη επειδή είναι μια λεπτή διαδικασία διάκρισης μεταξύ πολλές φορές των παρόμοιων αντικειμένων. Η προσθήκη του προβλήματος της θέασης καθιστά φυσικά το πρόβλημα δυσκολότερο. Αυτό συμβαίνει επειδή η εμφάνιση ενός προσώπου αλλάζει κάτω από την περιστροφή καθώς αυτό έχει μια σύνθετη τρισδιάστατη δομή.

Σε αυτό το σημείο πρέπει να διακρίνουμε την αναγνώριση προσώπου από την ανίχνευση προσώπου. Πράγματι πολλοί άνθρωποι πιστεύουν ότι αυτοί οι δύο όροι είναι οι ίδιοι. Σκεπτόμενοι ακόμη ότι έχουν πολλές παρόμοιες τεχνικές, και ότι είναι βασισμένοι στην ίδια ιδέα και ίδιους αλγόριθμους, είναι δύο διαφορετικά συστήματα. Η κύρια διαφορά είναι το γεγονός ότι η αναγνώριση προσώπου ανιχνεύει τα πρόσωπα και κάνει την αναζήτηση μέσω ενός συνόλου δεδομένων (data base) προκειμένου να βρεθεί μια ακριβής αντιστοιχία. Στην άλλη περίπτωση της ανίχνευσης, ο αλγόριθμος ψάχνει οποιαδήποτε αντιστοιχία και μόλις την βρει η διαδικασία σταματάει.

Ο στόχος της αναγνώρισης προσώπου φαίνεται να είναι διαδοχικός και πολύ συχνά έχει να αντιμετωπίσει καταστάσεις πολλών δεδομένων. Εντούτοις, είναι και υπολογιστικά και φυσικά πιο κατάλληλο να θεωρεί ως σύνολο οπτικών ενοτήτων με την ανατροφοδότηση κλειστών βρόγχων.

Προκειμένου να πραγματοποιηθεί ένα τέτοιο σύστημα, μια ολοκληρωμένη προσέγγιση έχει υιοθετηθεί που θα εκτελέσει την απόκτηση (acquisition), την κανονικοποίηση (normalization) και την αναγνώριση (recognition) με έναν λογικό τρόπο. Οι εικόνες μιας δυναμικής σκηνης υποβάλλονται σε επεξεργασία σε πραγματικό χρόνο για να ομαλοποιηθούν και ευθυγραμμισθούν οι ακολουθίες των προσώπου.

Στην ουσία, αυτή η διαδικασία είναι μια ενότητα κλειστών βρόγχων που περιλαμβάνει τον υπολογισμό και την συγχώνευση τριών διαφορετικών οπτικών παραδειγμάτων: κίνηση, χρώμα και εμφάνιση προσώπου.

Το οπτικό παράδειγμα του χρώματος είναι σημαντικό αλλά οι περισσότεροι από τους αλγόριθμους αναγνώρισης προσώπου δεν ανησυχεί πραγματικά για αυτό. Επίσης αναπτύχθηκαν μηχανισμοί πραγματικού χρόνου για εκτίμηση της γωνίας θέασης του προσώπου και έτσι μπορούμε να πούμε ότι βελτιώθηκε η ευρωστία της ανίχνευσης και της ευθυγράμμισης σε μεγάλο βαθμό.

Γενικά πολλή ερευνητική προσπάθεια έχει συγκεντρωθεί στους στόχους της αναγνώρισης προσώπου στους οποίους μόνο μια ενιαία εικόνα ή το πολύ μερικές εικόνες κάθε προσώπου είναι διαθέσιμες. Μια σημαντική ανησυχία είναι η εξελιξιμότητα στις μεγάλες βάσεις δεδομένων που περιέχουν χιλιάδες ανθρώπους, όπως θα δούμε και παρακάτω.

Εντούτοις, η μεγάλη συλλογή δεδομένων μπορεί να δημιουργήσει και πρόβλημα εάν μερικά από τα πρόσωπα μοιάζουν. Όσο πιο μεγάλη είναι μια βάση δεδομένων προσώπων τόσο πιο σίγουρα ελέγχεται για την σωστή ολοκλήρωσή του ένας αλγόριθμος, αλλά και τόσο πιο πολύπλοκο είναι στη περίπτωση που έχουν πολλά ίδια χαρακτηριστικά.

Οι στόχοι της αναγνώρισης προσώπου απαιτούν συνήθως να γίνεται η διαδικασία της αναγνώρισης σε περιορισμένες δυναμικές σκηνές. Αυτές οι σκηνές χαρακτηρίζονται από τη χαμηλή ευκρίνεια, τις μεγάλης κλίμακας αλλαγές, το μεταβλητό φωτισμό και περιστασιακά την ανακριβή μεταβολή και την ευθυγράμμιση στο πρόσωπο.

Η αναγνώριση που βασίζεται στις εικόνες αυτού του είδους είναι ιδιαίτερα ασυμβίβαστη και αναξιόπιστη. Εντούτοις, τα αποτελέσματα αναγνώρισης αυτών των στοιχείων κατά τη διάρκεια του χρόνου μπορούν να αντισταθμίσουν την κακή ποιότητα των αυτών.

Η αναγνώριση προσώπου είναι ένας ενεργός ερευνητικός τομέας που περιλαμβάνει διαφορετικούς τομείς όπως η φυσική, η ψυχολογία, η βιολογία, τα μαθηματικά, η πληροφορική και διάφορες άλλες. Ένα ευρύ φάσμα των προβλημάτων έχουν προσεγγιστεί, με συνέπεια πολλές ενδιαφέρουσες εφαρμογές.

3.5 Γνωστές πατέντες αναγνώρισης προσώπου

Όλο και περισσότερο, τα διπλώματα ευρεσιτεχνίας είναι πιο σημαντικά όταν κάποιος προσπαθεί να καθιερώσει μια νέα εφεύρεση. Επομένως, η καθιερωμένη επιστημονική έρευνα που γίνεται πρέπει να περιλάβει μια έρευνα διπλωμάτων ευρεσιτεχνίας, για να μπορεί να καθιερωθεί. Παρακάτω δίνεται ένας κατάλογος γνωστών διπλωμάτων ευρεσιτεχνίας που αναφέρονται λίγο πολύ στην ανίχνευση προσώπου από γνωστά στο χώρο τους πανεπιστήμια και εταιρίες.

<i>Χρονολογία</i>	<i>Εταιρεία</i>	<i>Τίτλος</i>
1998	University Central Florida	Αναγνώριση με templates
1997	MIT	Δικτυακά συστήματα και μέθοδοι για αναγνώριση
1996	Lucent Technologies	Μέθοδοι για αναγνώριση από συνεχόμενες εικόνες
1995	Neurmetics Ltd.	Σύστημα προσδιορισμού και επαλήθευσης

3.1 Γνωστές Πατέντες

3.6 Εμπορικά συστήματα και εφαρμογές

Κατά καιρούς, διάφορα προϊόντα αναγνώρισης προσώπου είναι διαθέσιμα στην εκάστοτε αγορά. Οι αλγόριθμοι που αναπτύσσονται, εξετάζονται κάτω από τη βάση δεδομένων FERET, που όπως θα δούμε και παρακάτω θεωρείται πολύ πλούσια σε δεδομένα για να γίνει μια σωστή αξιολόγηση.

Είναι εξαιρετικά δύσκολο να κρίνει κάποιος, ένα από τα τρία συστήματα

- visionics,

- Viisage, και
- Miros

που φαίνεται να είναι στην αγορά οι σύγχρονοι πρωτοπόροι στην αναγνώριση προσώπου.

Το πρόγραμμα αναγνώρισης προσώπου της Visionics, FaceIt βασίζεται στον τοπικό αλγόριθμο ανάλυσης χαρακτηριστικών γνωρισμάτων (Local Feature Analysis algorithm) που αναπτύσσεται στο πανεπιστήμιο Rockefeller. Το FaceIt ενσωματώνεται τώρα σε κλειστό κύκλωμα τηλεόρασης (Close Circuit Television CCTV) για το αποκαλούμενο αντι-έγκλημα σύστημα “Mandrake” στο Ηνωμένο Βασίλειο. Αυτό το σύστημα ψάχνει για τους γνωστούς εγκληματίες στο βίντεο που αποτελείται από 144 θέσεις φωτογραφικών μηχανών τύπου CCTV. Όταν γίνεται μια αντιστοιχία προσώπου με τη βάση δεδομένων εγκληματιών, τότε ειδοποιείται ο ανώτερος υπάλληλος ασφάλειας στο δωμάτιο ελέγχου.[7]

Το Viisage, είναι μια άλλη κορυφαία επιχείρηση στο τομέα της αναγνώρισης προσώπου, η οποία χρησιμοποιεί τον eigenface αλγόριθμο αναγνώρισης που αναπτύσσεται στο εργαστήριο MIT Media Laboratory. Το σύστημά τους χρησιμοποιείται από κοινού με τις κάρτες ταυτοποίησης (identification cards) (π.χ., άδειες του οδηγού και παρόμοιες κάρτες κυβερνητικής ταυτότητας) σε πολλά αμερικανικά κράτη και διάφορα αναπτυσσόμενα έθνη. [8]

Το Miros χρησιμοποιεί τη νευρική τεχνολογία δικτύων για το λογισμικό αναγνώρισης προσώπου TrueFace. Το TrueFace χρησιμοποιείται ουσιαστικά από για τον έλεγχο κατά την εξαργυρώση χρημάτων από ένα σύστημα, και έχει επεκταθεί επί των χαρτοπαικτικών λεσχών, όπως καζίνο και παρόμοιων τόπων σε πολλά αμερικανικά κράτη.[9]

3.7 Αλγόριθμοι Αναγνώρισης

Παρακάτω θα γνωρίσουμε τους περισσότερους αλγόριθμους αναγνώρισης προσώπου που χρησιμοποιούνται είτε σε συστήματα που βρίσκονται αυτή τη στιγμή στην αγορά είτε είναι στο στάδιο της μελέτης και εξέλιξης.

Αρχικά θα δούμε τους αλγόριθμους που κάνουν την διαδικασία, βασισμένη σε φωτογραφίες που είναι τραβηγμένες από ψηφιακές συσκευές. Έπειτα θα παρουσιάσουμε τους αλγόριθμους και την τεχνική στην περίπτωση που έχουμε εικόνα από βίντεο.

Παρακάτω παρουσιάζουμε τους σημαντικότερους αλγόριθμους για αναγνώριση προσώπου από εικόνα, και κάνουμε μια μικρή περιγραφή στους σημαντικότερους.

- **PCA** Principal Component Analysis

Έχει δημιουργηθεί από τον μετασχηματισμό Karhunen-Loeve's. Ουσιαστικά ο PCA δημιουργεί μια μαθηματική διαδικασία που μετασχηματίζει διάφορες (ενδεχομένως) συσχετισμένες μεταβλητές σε έναν (μικρότερο) αριθμό ασύνδετων μεταβλητών αποκαλούμενων κύρια συστατικά. [10]

- **ICA** Independent Component Analysis

- **LDA** Linear Discriminant Analysis

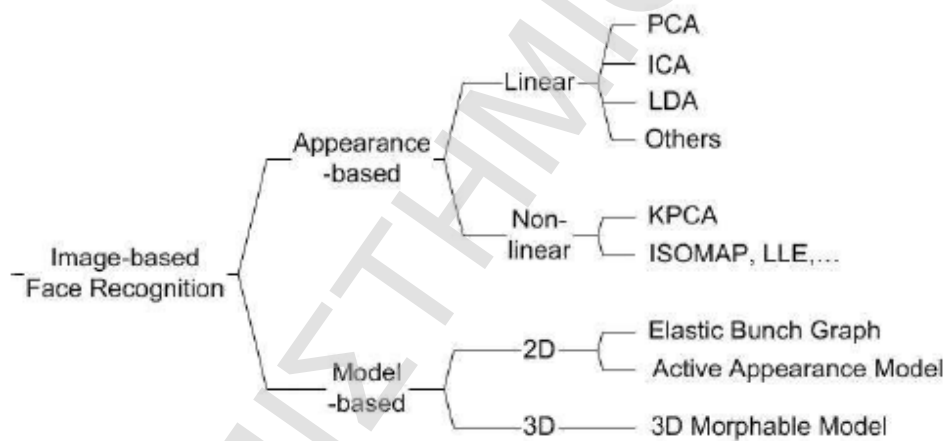
- **EBGM** Elastic Bunch Graph Matching

- **Kernel Methods**

Ο Αλγόριθμος που στηρίζεται σε Kernel Methods ουσιαστικά είναι μια γενίκευση των γραμμικών μεθόδων.

- **Trace Transform**

Είναι ένας αλγόριθμος που μπορεί να χρησιμοποιηθεί για την αναγνώριση των αντικειμένων κάτω από τους μετασχηματισμούς, π.χ. περιστροφή.



3.2 Κατηγοριοποίηση αλγορίθμων

Όπως βλέπουμε και στο παραπάνω σχήμα, γίνεται μια κατηγοριοποίηση των αλγορίθμων αναγνώρισης από στατική εικόνα ανάλογα με το μοντέλο που χρησιμοποιείται ή τον μετασχηματισμό.

Το σημαντικότερο όμως και για τους ερευνητές είναι η συχνή σύγκριση όλων των τύπων αλγορίθμου, σύμφωνα με συγκεκριμένες βάσεις δεδομένων που θα δούμε, ώστε να επιτευχθεί μεγαλύτερη μελλοντική απόδοση.

3.8 Αλγόριθμοι αναγνώρισης από εικόνα

Κατά τη διάρκεια της τελευταίας εικοσαετίας όλο και περισσότερη έρευνα έχει γίνει στον τομέα της αναγνώρισης προσώπου από συνεχόμενες εικόνες που προέρχονται από κάποια συσκευή καταγραφής βίντεο. Η αναγνώριση προσώπου των ανθρώπων από το πραγματικού χρόνου βίντεο είναι δύσκολη λόγω της χαμηλής ποιότητας των εικόνων και επειδή οι εικόνες προσώπου είναι μικρές. Σε αυτό το τομέα χρειάζεται ακόμη πολύ έρευνα και προσπάθεια για να φτάσει η ποιότητα σαν αυτή της αναγνώρισης από εικόνα.

Πάντως οι ψυχολόγοι και οι φυσιολόγοι πιστεύουν πως η μοντελοποίηση και η αναγνώριση που προέρχεται από τη διαδικασία μέσω της καταγραφής από βίντεο, μπορεί να πετύχει πιο μεγάλη απόδοση σε σχέση με τις στατικές εικόνες. [11]

Παρόλα αυτά παρακάτω θα δούμε μια σωστή προσέγγιση στην περίπτωση που η αναγνώριση από εικόνα έχει δύσκολα σενάρια.

3.9 Βάσεις Δεδομένων για έλεγχο του αλγορίθμου

Κατά την αξιολόγηση ενός αλγορίθμου αναγνώρισης, συνιστάται πολλές φορές να χρησιμοποιηθεί ένα τυποποιημένο σύνολο στοιχείων από τους ερευνητές για να είναι σε θέση να συγκρίνουν άμεσα τα αποτελέσματα. Ενώ υπάρχουν πολλές βάσεις δεδομένων

που χρησιμοποιούνται αυτήν τη στιγμή, η επιλογή μιας τέτοιας κατάλληλης πρέπει να γίνει βασισμένη στο στόχο που δίνεται αρχικά από τον ίδιο τον αλγόριθμο (γήρανση, εκφράσεις προσώπου, φωτισμός κ.λπ.).

Ένας άλλος τρόπος είναι να επιλεγεί το συγκεκριμένο σύνολο στοιχείων για συγκεκριμένο σκοπό. Δηλαδή πως συμπεριφέρεται ο αλγόριθμος σε περιπτώσεις που οι εικόνες που του δίνονται μεταβάλλονται ως προς το φωτισμό ή τη γωνία θέασης.

Στη πρώτη περίπτωση συνήθως έχουμε σαν αποτέλεσμα ένα ποσοστό επιτυχίας. Δηλαδή ο x αλγόριθμος στη βάση δεδομένων Feret έχει 85% επιτυχία ενώ ο y 73%. Έτσι είμαστε σε θέση κάθε φορά να μπορούμε να κρίνουμε τη απόδοση και αξιοπιστία ενός αλγόριθμου. Στην δεύτερη περίπτωση, απλά βλέπουμε τι αποτελέσματα μας δίνει ο αλγόριθμος σε διάφορες περιπτώσεις μεταβολών είτε στο χώρο είτε στην έκφραση του προσώπου.[12]

Παρακάτω ακολουθούν μερικές σημαντικές βάσεις δεδομένων που συχνά χρησιμοποιούνται από ερευνητές.

- The Color FERET Database, USA
- The Yale Face Database
- PIE Database, CMU
- Project - Face In Action (FIA) Face Video Database, AMP, CMU
- AT&T "The Database of Faces" (formerly "The ORL Database of Faces")
- MIT-CBCL Face Recognition Database
- The Extended M2VTS Database, University of Surrey, UK
- The AR Face Database, Purdue University, USA
- Cohn-Kanade AU Coded Facial Expression Database

3.10 Εύκολα και δύσκολα σενάρια για αναγνώριση εικόνας

Σε γενικές γραμμές οι προσεγγίσεις που προτείνονταν τα τελευταία χρόνια ήταν σε θέση να λύσουν συγκεκριμένες εφαρμογές αναγνώρισης εικόνων προσώπου. Παραδείγματα των σεναρίων όπου η αναγνώριση προσώπου επιτυγχάνει πολύ καλά αποτελέσματα, φαίνονται στις παρακάτω φωτογραφίες. Αν και πολλές λεπτομέρειες μπαίνουν στην άκρη όπως ποιότητα, μέγεθος της βάσης δεδομένων και γωνία θέασης, μπορεί να θεωρηθεί ότι σε τέτοια σενάρια το πρόβλημα αναγνώρισης προσώπου λύνεται κάλλιστα.



3.3 Σωστά σενάρια αναγνώρισης

Όταν όμως το σενάριο ξεφεύγει από την εύκολη προσέγγιση, τότε οι οποιοσδήποτε προσεγγίσεις αναγνώρισης προσώπου αντιμετωπίζουν σοβαρά προβλήματα. Μεταξύ των προκλήσεων που καλείται να αντιμετωπίσει ο αλγόριθμος είναι οι εξής :

- Γωνία θέασης
- Συνθήκες φωτισμού
- Γυαλιά, μουστάκια, γενειάδες,
- Χαμηλής ποιότητας εικόνα
- Κακές συνθήκες σε εξωτερικό χώρο
- Εικόνες παλιότερης χρονολογίας

Στις παρακάτω εικόνες φαίνονται μερικά από τα προβλήματα που καλείται να αντιμετωπίσει ο αλγόριθμος. Ένα άλλο σημαντικό πρόβλημα, που είναι πάνω από τις εικόνες είναι πώς τα διαφορετικά συστήματα αναγνώρισης προσώπου συγκρίνονται μεταξύ τους, ώστε να βγουν ξεκάθαρα συμπεράσματα.[13]





3.3 Δύσκολα σενάρια αναγνώρισης

Στις παραπάνω εικόνες βλέπουμε τα προβλήματα της αναγνώρισης. Στην πρώτη εικόνα λόγω της κακής ποιότητας αλλά και της μη εστίασης σε συγκεκριμένο πρόσωπο και στη δεύτερη λόγω της διαφορετικότητας σε χαρακτηριστικό προσώπου, όπως είναι το μούσι.

3.11 Λύσεις για δύσκολα σενάρια αναγνώρισης

Σε τέτοιες καταστάσεις, όπως είδαμε στην προηγούμενη ενότητα, η βοήθεια προέρχεται από δύο περιοχές έρευνας.

- Αναγνώριση μέσω βίντεο (Video-based face recognition)
- Προσεγγίσεις multimodal

Αναγνώριση μέσω βίντεο

Αρχικά θα πρέπει να τονίσουμε ότι ακόμα και τα συστήματα αναγνώρισης μέσω βίντεο, χρησιμοποιούν την εικόνα ως μέτρο σύγκρισης. Απλά όπως θα δούμε παρακάτω, υπερτερεί σε πολλούς τομείς σε σχέση με την απλή διαδικασία.

Παρόλο που τις τελευταίες δεκαετίες αυτή η τεχνική από βίντεο έχει αναπτυχθεί αρκετά, τα συστήματα που βρίσκονται στην αγορά εξακολουθούν να χρησιμοποιούν τις μεμονωμένες εικόνες για την αναγνώριση προσώπου.

Αν και η αναγνώριση μέσω βίντεο, δεν υπάρχει σε αρκετές εφαρμογές, έχει πάρα πολλά θετικά στοιχεία έναντι της άλλης τεχνικής, όπως [13],[14]

- Μπορούν να επιλεγτούν τα καλύτερα πλαίσια από μια ακολουθία βίντεο, για να πετύχουμε πιο αποδοτική αναγνώριση.
- Το βίντεο παρέχει τη χρονική συνοχή, που επιτρέπει την επαναχρησιμοποίηση των πληροφοριών αναγνώρισης που λαμβάνεται από τις υψηλής ποιότητας εικόνες στην επεξεργασία των χαμηλής ποιότητας πλαισίων.

- Το βίντεο επιτρέπει την ανίχνευση των εικόνων έτσι ώστε οι εκφράσεις του προσώπου και οι παραλλαγές της πόζας να μπορούν να μεταβληθούν, με συνέπεια τη βελτίωση της αναγνώρισης.
- Η κίνηση, ο βηματισμός και άλλα χαρακτηριστικά γνωρίσματα μπορούν να βοηθήσουν ένα βασισμένο στο βίντεο σύστημα αναγνώρισης προσώπου.

Αφ' ετέρου, προτού να μπορέσουν να χρησιμοποιηθούν οι βασισμένες στο βίντεο τεχνικές αναγνώρισης προσώπου, μερικά άλλα δύσκολα θέματα πρέπει να αντιμετωπιστούν:

- Κατάτμηση του βίντεο
- Ανίχνευση προσώπου
- Εξαγωγή χαρακτηριστικών γνωρισμάτων από το βίντεο
- Χαμηλής ποιότητας συμπίεσμένες εικόνες ανάλογα με την εκάστοτε εφαρμογή

Παρά αυτά τα δύσκολα θέματα, είναι πεποίθηση πολλών ότι η βασισμένη στο βίντεο αναγνώριση προσώπου έχει μια τεράστια δυνατότητα και θα μετακινήσει την έρευνα σε αυτό το τομέα αρκετά μπροστά. Τα αποτελέσματα μέχρι στιγμής φαίνονται πολύ ελπιδοφόρα.



3.4 Συνεχόμενες εικόνες από βίντεο

Προσεγγίσεις multimodal

Τα βασισμένα στο βίντεο συστήματα αναγνώρισης προσώπου φέρνουν και μερικά άλλα χαρακτηριστικά γνωρίσματα που μπορούν να βοηθήσουν ώστε να ελαττωθεί αισθητά το πρόβλημα:

- ομιλία και
- ήχος

Εάν η πληροφορία που χρειαζόμαστε για να κάνουμε την αναγνώριση είναι μέσα σε αυτό το περιεχόμενο, τότε γιατί να μην το χρησιμοποιήσουμε. Η μίξη πολλών διαφορετικών πληροφοριών πάντα δημιουργεί προβλήματα, αλλά τις περισσότερες φορές δίνει και λύσεις. Έχει υπάρξει ήδη μια μεγάλη προσπάθεια σε αυτήν την κατεύθυνση που είναι επίσης πολύ ελπιδοφόρα.[13],[15], [16].

3.12 Προβλήματα κατά τη διαδικασία της αναγνώρισης

Παρακάτω κάνουμε μια πιο λεπτομερή παρουσίαση των προβλημάτων κατά την διαδικασία της αναγνώρισης προσώπου, όπως είδαμε και προηγουμένως. Έτσι εστιάζομαστε σε έξι κύρια προβλήματα, τα οποία είναι τα εξής :

1. Γωνία εξέτασης (Viewing angle)

Το πρόσωπο από τη φύση του έχει μια τρισδιάστατη μορφή. Δεδομένου ότι αλλάζει η θέαση της φωτογραφικής μηχανής κατά τη κίνηση, η εμφάνιση του προσώπου μπορεί να αλλάξει λόγω

α) της παραμόρφωσης, η οποία δημιουργείται κατά το τέντωμα και το τράβηγμα διαφορετικών μερών του προσώπου, και

β) του κλεισίματος και ανοίγματος των μερών του προσώπου. Εάν έχουμε δει τα πρόσωπα μόνο από μια γωνία θέασης, τότε γενικά είναι δύσκολο να αναγνωριστούν.



3.5 Διαφορετικές γωνίες θέασης

2. Φωτισμός

Ακριβώς όπως και με τη μεταβολή στη γωνία θέασης, η μεταβολή φωτισμού είναι αναπόφευκτη. Ο φυσικός φωτισμός αλλάζει πολύ κατά τη διάρκεια των ωρών της ημέρας και μεταξύ των εσωτερικών και υπαίθριων περιβαλλόντων. Φυσικά δεν θα πρέπει να αγνοούμε εκτός από το φυσικό φωτισμό, και τον τεχνικό ο οποίος με τη παρουσία του δημιουργεί μεταβολή φωτισμού.

Λόγω της τρισδιάστατης μορφής του ανθρώπινου προσώπου, η πηγή άμεσου φωτισμού μπορεί δημιουργήσει ισχυρές σκιές και σκιάσεις που τονίζουν ή μειώνουν ορισμένα χαρακτηριστικά γνωρίσματα του προσώπου.

Προηγούμενα συμπεράσματα από ερευνητικές δοκιμές αναγνώρισης προσώπου δείχνουν ότι οι αλλαγές φωτισμού υποβιβάζουν κατά πολύ την διαδικασία της αναγνώρισης. Εντούτοις αυτό το συμπέρασμα είναι δύσκολο να ερμηνευθεί. Η επίδραση της αλλαγής του φωτισμού στις εικόνες μπορεί να οφείλεται σε δύο παράγοντες

- στο έμφυτο που φως, που ανακλάται από το δέρμα
- στη μη γραμμική ρύθμιση της κάμερας, όπως η αντίθεση και η διόρθωση gamma (gamma correction)

Και τα δύο μπορούν να έχουν σημαντικές αλλαγές στην εμφάνιση του προσώπου. Ενώ το τελευταίο είναι λιγότερο αξιοπρόσεχτο από τους ανθρώπους, μπορεί να προκαλέσει σημαντικά προβλήματα για στην επεξεργασία από τους υπολογιστές.



3.6 Πρόσωπα με διαφορετικό φωτισμό

3. Έκφραση προσώπου

Το πρόσωπο είναι ένα αντικείμενο το οποίο από τη φύση του είναι εύκαμπτο. Η έκφραση του προσώπου κατά την συγκίνηση και τη γλωσσολογική επικοινωνία μαζί με τις λεκτικές πράξεις μπορούν να δημιουργήσουν μια μεγάλη παραλλαγή στην εμφάνιση του προσώπου.

Ο αριθμός των πιθανών αλλαγών, σύμφωνα με τις υπάρχουσες πληροφορίες, στην έκφραση του προσώπου είναι πολύ μεγάλος. Ακόμα και σήμερα η επιρροή της έκφρασης του προσώπου στην αναγνώριση δεν είναι απόλυτα κατανοητή.

Προηγούμενες έρευνες πάνω σε αυτό το θέμα, είχαν περιοριστεί σε ουδέτερες εκφράσεις και σε μικρά χαμόγελα του προσώπου. Επειδή η έκφραση του προσώπου έχει προφανείς επιπτώσεις στην γεωμετρική μορφή και τη θέση των χαρακτηριστικών γνωρισμάτων, η επιρροή στην αναγνώριση μπορεί να είναι μεγαλύτερη για αλγορίθμους που βασίζονται στην γεωμετρία παρά από τους άλλους.

Συνήθως χρησιμοποιούμε τη βάση δεδομένων Cohn-Kanade για να ερευνήσουμε αυτά τα ζητήματα. Τα δείγματα αυτών των βάσεων δεδομένων χαρακτήρισαν καλά τις συγκινήσεις και τις εκφράσεις του προσώπου.

4. Φράξιμο

Το πρόσωπο μπορεί να κλειστεί από άλλα αντικείμενα που βρίσκονται εκείνη την ώρα στο χώρο ή από τα γυαλιά ηλίου ή άλλα πράγματα, που μπορεί να έχει ή να φοράει. Το φράξιμο μπορεί να είναι ακούσιο ή ακόμα και σκόπιμο σε περιπτώσεις αποφυγής της αναγνώρισης.

Σε πολλές περιπτώσεις τα άτομα μπορούν να παρακινηθούν για να ανατρέψουν τις προσπάθειες αναγνώρισης, καλύπτοντας σημαντικά μέρη του προσώπου τους. Δεδομένου ότι σε πολλές τέτοιες καταστάσεις, ο στόχος του συστήματος είναι να αναγνωρίσει τα άτομα είτε συνεργάζονται είτε όχι, είναι σημαντικό να ξέρουμε πόσο δύσκολο είναι να αναγνωριστούν οι άνθρωποι, δεδομένου ορισμένων ποσοτικών και ποιοτικών αλλαγών κατά το φράξιμο.

Συνήθως εξετάζεται υπό ποιους όρους τέτοιες προσπάθειες είτε μπορούν, είτε όχι να είναι επιτυχείς. Για να ερευνηθεί καλύτερα αυτό το φαινόμενο χρησιμοποιείται η βάση δεδομένων του AR από το πανεπιστήμιο του Purdue της Αμερικής, η οποία έχει δύο διαφορετικούς τύπους κλεισίματος προσώπου, ένα για τα μάτια, και ένα για τα πρόσωπα που χαμηλώνουν.

5. Προβλήματα λόγω χρόνου

Τα πρόσωπα αλλάζουν με το πέρασμα του χρόνου. Έτσι πολλές φορές πρέπει ο αλγόριθμος να αντιμετωπίσει αλλαγές στα μαλλιά, στο makeup και σε άλλα στοιχεία που μεταβάλλουν τα χαρακτηριστικά γνωρίσματα του προσώπου.

Επίσης έχουμε προβλήματα από το χάλάρωμα των μυών και την απουσία τρίχας του κεφαλιού, και από πιο μακροχρόνια αποτελέσματα που έχουν σχέση με τη γήρανση.

Συνήθως χρησιμοποιείται η βάση δεδομένων του AR για να ερευνηθούν τα αποτελέσματα της χρονικής καθυστέρησης και των αλληλεπιδράσεων μεταξύ της χρονικής καθυστέρησης και της έκφρασης του προσώπου, της μεταβολής του φωτισμού, και του φραξίματος από αντικείμενα.

6. Μεμονωμένοι παράγοντες

Οι αλγόριθμοι της αναγνώρισης προσώπου μπορούν να είναι λίγο έως πολύ ευαίσθητοι για τους άνδρες ή τις γυναίκες ή τα άτομα των διαφορετικών εθνικοτήτων.

Συνήθως εστιάζουμε στις διαφορές μεταξύ των ανδρών και των γυναικών όσον αφορά την απόδοση αλγορίθμου. Διαισθητικά, οι γυναίκες είναι πιο δύσκολα να αναγνωρίσουν λόγω της μεγαλύτερης χρήσης και της καθημερινής παραλλαγής στο makeup ή στα δομικά χαρακτηριστικά γνωρίσματα του προσώπου.

Πάντως τα αρσενικά και τα θηλυκά πρόσωπα διαφέρουν και στα τοπικά χαρακτηριστικά γνωρίσματα και στη μορφή. Τα πρόσωπα των αντρών έχουν κατά μέσον όρο παχύτερα φρύδια.. Στα πρόσωπα των γυναικών, η απόσταση μεταξύ των ματιών και των φρυδιών

είναι μεγαλύτερη, ενώ η προεξοχή της μύτης είναι μικρότερη, και το πηγούνι στενότερο απ'ότι στους άνδρες.

Οι άνθρωποι συνήθως διακρίνουν εύκολα το αρσενικό από το θηλυκό πρόσωπο χρησιμοποιώντας αυτές και άλλες διαφορές (π.χ., στυλ κόμμωσης), και από αυτό το γεγονός προσπαθήσανε οι αλγόριθμοι να πετύχουν την δικιά τους διαδικασία.

Εντούτοις λίγα είναι γνωστά, για την ευαισθησία των αλγορίθμων αναγνώρισης προσώπου στις διαφορές μεταξύ των ανδρών και των γυναικών. Τα σχετικά ποσοστά των ανδρών και των γυναικών στα δείγματα αναφέρονται σπάνια, και τα αποτελέσματα από τη αναγνώριση αποτυγχάνουν να αναφέρουν εάν οι αλγόριθμοι είναι λιγότερο ακριβής για ένα φύλο ή άλλο. Με λίγα λόγια, δεν έχει γίνει ακόμα συστηματική έρευνα πάνω σε αυτό το πεδίο.

Άλλοι παράγοντες που μπορούν να επηρεάσουν την αναγνώριση, είναι οι διαφορές στη μορφή του προσώπου μεταξύ των ατόμων της ευρωπαϊκής, ασιατικής, και αφρικανικής καταγωγής, τα οποία έχουν και αυτά αγνοηθεί.

Συνήθως αυτό το ζήτημα αντιμετωπίζεται, από τη βάση δεδομένων του AR, που έχει ισορροπήσει καλά τα ποσοστά των ανδρών και των γυναικών και από τη FERET αντίστοιχα.

3.13 Το μέλλον στην αναγνώριση προσώπου

Τα συστήματα αναγνώρισης προσώπου σε γενικές γραμμές επεξεργάζονται πολύ καλά τις εικόνες υπό τους περιορισμένους όρους, αν και όλα τα συστήματα λειτουργούν πολύ καλύτερα με τις μετωπικές εικόνες και το σταθερό φωτισμό. Όλοι οι τωρινοί αλγόριθμοι αναγνώρισης προσώπου αποτυγχάνουν κάτω από τους απέραντα ποικίλους όρους από τους οποίους οι άνθρωποι χρειάζονται και είναι σε θέση να προσδιορίσουν άλλους ανθρώπους. Τα συστήματα αναγνώρισης προσώπων επόμενης γενιάς θα πρέπει να αναγνωρίσουν τους ανθρώπους σε πραγματικό χρόνο και στις πολύ λιγότερο περιορισμένες καταστάσεις.

Θεωρείται ότι τα συστήματα προσδιορισμού και ανίχνευσης που αντέχουν στα φυσικά περιβάλλοντα, παρουσία των αλλαγών θορύβου και φωτισμού, δεν μπορούν να στηριχθούν σε μια ενιαία μορφή, και έτσι είναι ουσιαστική η τεχνολογία που χρησιμοποιείται στα έξυπνα περιβάλλοντα, η οποία επιτρέπει στους χρήστες να ενεργήσουν ελεύθερα.

Οι φωτογραφικές μηχανές και τα μικρόφωνα είναι σήμερα πολύ μικρές, ελαφριές και έχουν ενσωματωθεί επιτυχώς σε έξυπνα συστήματα. Τα ακουστικά και τηλεοπτικά συστήματα αναγνώρισης έχουν μεγάλο πλεονέκτημα στο ότι κάνουν χρήση ανθρώπινων μοντέλων για την αναγνώριση. Τέλος, οι ερευνητές αρχίζουν να καταδεικνύουν ότι τα έξυπνα συστήματα προσδιορισμού προσώπων μπορούν να επιτύχουν υψηλά ποσοστά αναγνώρισης χωρίς την ενέργεια του χρήστη.

ΚΕΦΑΛΑΙΟ 4

Υλοποίηση Εφαρμογής - Αποτελέσματα

4.1 Εισαγωγή

Σκοπός της εργασίας αυτής όπως είπαμε και στην εισαγωγή, είναι η δημιουργία μιας αυτόνομης διαδικασίας καταγραφής εικόνας, αναγνώρισης της και αποστολής των αναγνωριστικών σε ένα απομακρυσμένο σύστημα αποφάσεων. Το οποίο με τη σειρά του ανάλογα με τα δεδομένα, πράττει και διαφορετικά.

Χρησιμοποιήσαμε τρεις διαφορετικές τεχνολογίες στην παρούσα εργασία.

1. Πρόγραμμα ανίχνευσης και καταγραφής Motion
2. Αλγόριθμο αναγνώρισης προσώπου του πανεπιστημίου του Toronto
3. Jadex, java πλατφόρμα δημιουργίας agent

Έτσι χρησιμοποιήσαμε το πρόγραμμα Motion για να έχουμε τις φωτογραφίες των προσώπων που περνάνε εκείνη την ώρα από τη κάμερα, ύστερα τον αλγόριθμο για να διαπιστώσουμε ποιος είναι ο χρήστης και τέλος την τεχνολογία jadex για να στείλουμε από agent σε agent τα δεδομένα που πρέπει. Παρακάτω φαίνεται η αρχιτεκτονική της διαδικασίας και αναλύεται περισσότερο.

4.2 Αρχιτεκτονική

Στο παρακάτω σχήμα βλέπουμε την αρχιτεκτονική της εφαρμογής μας.

1. Αρχικά έχουμε την κάμερα που είναι συνδεδεμένη στον υπολογιστή, ο οποίος έχει το πρόγραμμα motion για την ανίχνευση της κίνησης και την καταγραφή των εικόνων.
2. Το πρόγραμμα αποθηκεύει τις φωτογραφίες σε έναν συγκεκριμένο φάκελο, τον οποίο τον παρακολουθεί ο αλγόριθμος αναγνώρισης που ήδη υπάρχει στον υπολογιστή.
3. Έχουμε την διαδικασία αναγνώρισης των εικόνων. Ο αλγόριθμος θα μας βγάλει το όνομα του προσώπου που αντιστοιχεί από μια βάση φωτογραφιών που υπάρχει στον υπολογιστή.



4.1 Αρχιτεκτονική Εφαρμογής

4. Αφού ολοκληρωθεί η διαδικασία της αναγνώρισης, ακολουθεί διαδικασία αποστολής των δεδομένων σε έναν απομακρυσμένο υπολογιστή αποφάσεων. Στον υπολογιστή στέλνονται το όνομα του χρήστη καθώς και ο αριθμός της κάμερας που έκανε την συγκεκριμένη καταγραφή.

Με μεγαλύτερη λεπτομέρεια για την διαδικασία όλη θα ασχοληθούμε πιο μετά.

4.3 Περιβάλλον – Βοηθητικά προγράμματα

- Λειτουργικό σύστημα : FEDORA LINUX VERSION core 3
- Προγραμματιστική γλώσσα : JAVA, έκδοση JDK 1.5
- Βιβλιοθήκες : COLT Matrix Library[17]
- Jadex 0.941
- Java και xml editors

4.3.1 Βιβλιοθήκη Colt

Αυτή η διανομή παρέχει μια υποδομή για επιστημονικούς αλλά και τεχνικούς υπολογισμούς στην java. Αποτελείται από διάφορες δωρεάν βιβλιοθήκες της java, και για την ευκολία των χρηστών έχει διαμορφωθεί κάτω από μια ενιαία ομοιόμορφη διάταξη. Έτσι περιέχει της εξής βιβλιοθήκες

- *Colt* library
- *Jet* library
- *JAL* library
- *RngPack* library
- *CoreJava* library
- *Concurrent* library
- *VNI* library

Η Colt βιβλιοθήκη παρέχει τις θεμελιώδεις γενικής χρήσης δομές δεδομένων που βελτιστοποιούνται για τα αριθμητικά στοιχεία, όπως για multi-dimensional πίνακες, τη γραμμική άλγεβρα και άλλα πολλά.

Η Jet βιβλιοθήκη περιέχει τα μαθηματικά και στατιστικά εργαλεία για την ανάλυση στοιχείων, τις τυχαίες γεννήτριες αριθμών και είναι χρήσιμη για προσομοιώσεις γεγονότων.

Η JAL βιβλιοθήκη, είναι ένα μέρος της C++ Standard Template Library που αναπτύσσεται από τη Silicon Graphics, και περιέχει ένα ευρύ φάσμα κωδικοποιημένων αλγορίθμων για στους πίνακες.

Η βιβλιοθήκη RngPack περιέχει τυχαίες γεννήτριες αριθμών.

Η Concurrent βιβλιοθήκη περιέχει όλα αυτά που χρειάζονται για τον παράλληλο και ταυτόχρονο προγραμματισμό, και τέλος η βιβλιοθήκη VNI περιέχει μαθηματικές λειτουργίες και σύνθετους αριθμούς.

4.3.2 Jadex

Το Jadex όπως αναφέρθηκε και πιο πριν είναι μια java πλατφόρμα δημιουργίας agent-πρακτόρων, η οποία ακολουθεί το BDI Belief Desire Intention μοντέλο και διευκολύνει την ευφή δημιουργία πρακτόρων. Το σημαντικότερο όμως είναι ότι επιτρέπει τη δημιουργία πρακτόρων σε γλώσσα xml και java και φυσικά μπορεί να επεκταθεί σε διαφορετικά middleware όπως το JADE.

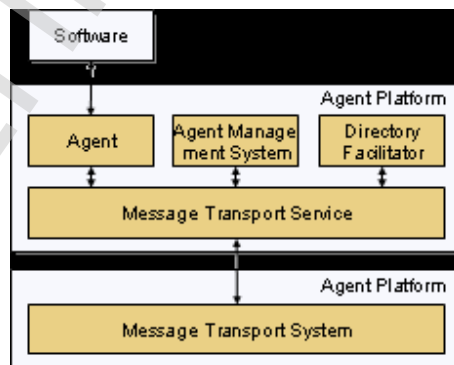
- **Χαρακτηριστικά Jadex**
 - Στηρίζεται στη Java
 - Ακολουθεί το μοντέλο FIPA

Ο στόχος του Jadex, είναι ουσιαστικά φτιάξει ένα περιβάλλον δημιουργίας πρακτόρων-agents, με τον πιο εύκολο τρόπο χωρίς την αλλαγή των βασικών συστατικών, όπως αρχιτεκτονικής ή γλώσσας προγραμματισμού. Γι' αυτό το λόγο έχει σχεδιαστεί με σκοπό να διευκολύνει τη χρησιμοποίηση των πρακτόρων στη διαδομένη γλώσσα προγραμματισμού της Java, επομένως επιτρέποντας να επαναχρησιμοποιήσει ένα μεγάλο αριθμό εργαλείων και βιβλιοθηκών της γλώσσας που ήδη υπάρχουν.

Ένας από τους αρχικούς παράγοντες επιτυχίας αυτής της νέας τεχνολογίας είναι η έγκαιρη διαθεσιμότητα των προτύπων για να εγγυηθεί τη διαλειτουργικότητα μεταξύ των εξελισσόμενων προϊόντων.

Προκειμένου να διευκολυνθεί η διαλειτουργικότητα των ανεξάρτητα αναπτυγμένων συστημάτων πρακτόρων, το Foundation for Intelligent Physical Agents (FIPA) δημιούργησε ένα σύνολο προδιαγραφών, οι οποίες αναφέρονται συνήθως ως "πρότυπα FIPA".

Όπως φαίνεται και στο παρακάτω σχήμα, το πρότυπο FIPA διευκρινίζει μια αρχιτεκτονική πλατφορμών πρακτόρων, η οποία καθορίζει τις υπηρεσίες όπως η διαχείριση πρακτόρων και έναν βοηθό καταλόγου. Αυτή η αρχιτεκτονική επιτρέπει στους πράκτορες να επικοινωνήσουν, χρησιμοποιώντας μιας κοινή γλώσσας επικοινωνίας πρακτόρων.



4.1 Το πρότυπο FIPA

Επίσης το Jadex είναι βασισμένο στο JADE Agent Framework, μια open source ανάπτυξη από το εργαστήριο της Telecom Italia. Το JADE παρέχει την αρχιτεκτονική της πλατφόρμας και τους κύριους μηχανισμούς υπηρεσιών και μεταφορών μηνυμάτων όπως απαιτείται από τις προδιαγραφές FIPA.

4.3.3 Κάμερα

Η κάμερα που χρησιμοποιήθηκε για τις ανάγκες της εργασίας είναι μια web cam της εταιρεία Logitech, quickcam express με id 046d:0870.

4.3.4 Classpaths

Απαραίτητη προϋπόθεση στην εκτέλεση της εφαρμογής μας είναι η δημιουργία μιας μεταβλητής classpath στο περιβάλλον του linux. Φυσικά το ίδιο ισχύει και για να εκτελέσουμε τον κώδικα της java, οπότε ορίζουμε και την path.

Έτσι και για να είναι μόνιμη η δημιουργία των μεταβλητών πληκτρολογούμε στο τερματικό

```
gedit .bashrc
```

και ορίζουμε τις διαδρομές των αρχείων που θέλουμε να χρησιμοποιήσουμε.

4.4 Περιγραφή δομής Demo

Μέσα στο cd που συνοδεύεται η εργασία, εμπεριέχονται όλα τα αρχεία που χρειάζονται και ο κώδικας της εφαρμογής. Παρακάτω περιγράφονται με λεπτομέρεια οι φάκελοι και τα αρχεία που συνθέτουν αυτή την εργασία. Ο φάκελος FaceRecognition έχει την εξής δομή φακέλων.



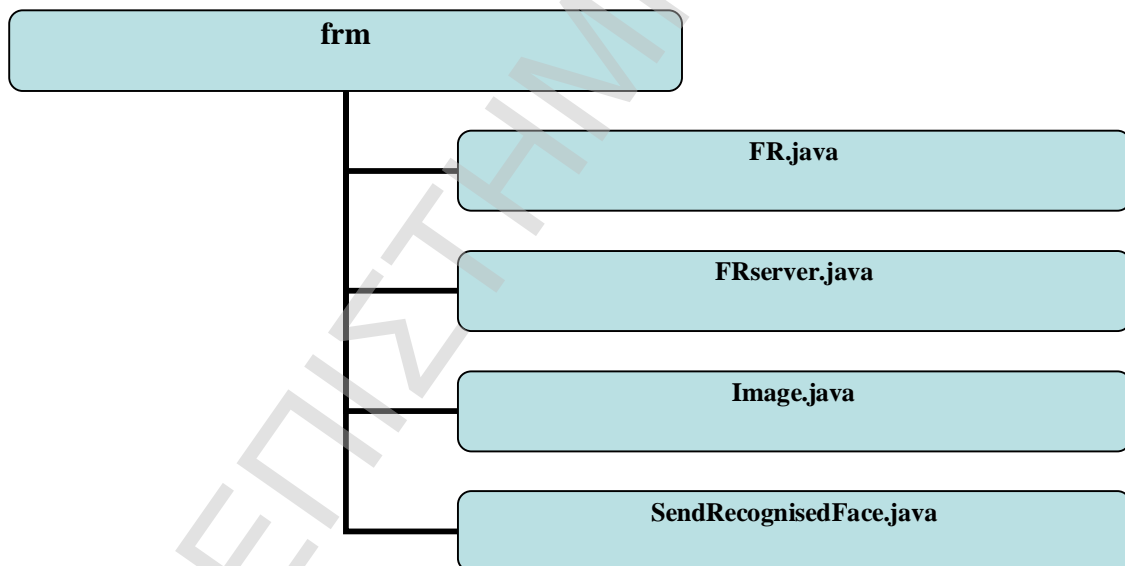
Μέσα στο φάκελο TrainingPics, υπάρχει η βάση δεδομένων των φωτογραφιών. Για την ανάγκη της εργασίας, δεν χρησιμοποιήσαμε κάποια γνωστή βάση προσώπων, όπως είχαμε δει, αλλά δικές μας φωτογραφίες που τραβήχτηκαν από τη κάμερα. Εξάλλου σκοπός της εργασίας δεν ήταν η απόδοση του συγκεκριμένου αλγορίθμου.

Στον φάκελο testImages καταλήγουν όλες οι φωτογραφίες που καταγράφονται από την κάμερα.

Επίσης στους φακέλους lib, onto και jadex εμπεριέχονται όλες οι απαραίτητες βιβλιοθήκες ή προγράμματα που χρησιμοποιεί η κύρια μας εφαρμογή. Μέσα στο lib βρίσκεται και η βασική βιβλιοθήκη για την εφαρμογή μας colt.

Τέλος στον φάκελο src υπάρχει ο κώδικας της εφαρμογής μας και στον classes οι αντίστοιχες κλάσεις του κώδικα java.

Στο φάκελο src και στον υποφάκελο frm βρίσκονται τέσσερα αρχεία κώδικα java που χρειαζόμαστε στην εφαρμογή μας.



Το αρχείο `image.java` είναι ουσιαστικά ο αλγόριθμος που χρησιμοποιούμε για να γίνει η αναγνώριση του προσώπου.

Το κύριο μας αρχείο είναι το `FRserver`, το οποίο ελέγχει συνέχεια τον φάκελο `testImages` και ενεργοποιεί ουσιαστικά την διαδικασία του αλγορίθμου, την διαδικασία της προβολής ενός μικρού παραθύρου και τέλος την διαδικασία αποστολής μηνύματος στον απομακρυσμένο υπολογιστή.

Την δουλειά της προβολής του παραθύρου την κάνει το αρχείο `FR.java` και της αποστολής το `SendRecognisedFace.java`.

Επίσης, εμπεριέχεται και ο κώδικας του άλλου απομακρυσμένου υπολογιστή, στον φάκελο `reas` του `src`, ο οποίος λέγεται `ReceiveRecognizedFacePlan`.

Μέσα στον φάκελο `classes`, υπάρχουν και τα δύο `xml` αρχεία, τα οποία ουσιαστικά δημιουργούν τις πλατφόρμες πάνω στις οποίες η εφαρμογή μας εξελίσσεται. Παρακάτω στο εγχειρίδιο χρήσης θα δούμε ακριβώς πως ουσιαστικά ενεργοποιείται η εφαρμογή μας μέσω της `java` πλατφόρμας `jadex`.

4.5 Ανάλυση Εφαρμογής

Σε αυτό το κομμάτι της εργασίας θα αναλύσουμε με λεπτομέρειες την εφαρμογή, με όπου κρίνεται απαραίτητο τη προσθήκη κομματιών κώδικα και σχημάτων για την καλύτερη κατανόηση.

1. Motion

Στο πρώτο μεγάλο κομμάτι της εργασίας είδαμε όλη την λειτουργία του συγκεκριμένου προγράμματος.

Θα πρέπει να τονίσουμε ότι για την έναρξη του motion, δεν χρειάζεται να πληκτρολογήσουμε στο terminal τη λέξη motion μιας και έγινε προσθήκη στον κώδικα της java, ώστε να γίνεται αυτόματα με το που ξεκινάει η κύρια εφαρμογή μας. Αυτό γίνεται με τον εξής κώδικα

```
Runtime rt= Runtime.getRuntime();  
Process process = rt.exec("motion");
```

Έτσι με αυτό τον τρόπο η εφαρμογή μας γίνεται πιο αυτοματοποιημένη και δεν χρειάζεται πρώτα να ενεργοποιήσουμε το motion για να στείλει τις εικόνες στον φάκελο και μετά να ξεκινήσουμε την πλατφόρμα μας.

Σε αυτό το σημείο πρέπει να τονίσουμε ότι η ανίχνευση και η καταγραφή των εικόνων ξεκινούν μετά από περίπου πέντε δευτερόλεπτα, από την αρχική ενεργοποίηση της εφαρμογής μας.

Σημαντικό γεγονός αποτελεί η παραμετροποίηση του config αρχείου που αρχικά διαβάζει η κάμερα. Μέσα στο αρχείο, όπως φαίνεται και παρακάτω μπορούμε να αλλάξουμε σημαντικές ιδιότητες που έχουν να κάνουν με την κάμερα ή και την διαχείριση των φωτογραφιών, όπως που να αποθηκευτούν και με ποιο όνομα.

Στην δικιά μας τη περίπτωση διορθώνουμε την ιδιότητα target dir στην επιθυμητή μας διαδρομή. Στην συγκεκριμένη περίπτωση επιλέγουμε να βάλουμε τις φωτογραφίες στο φάκελο testImages του αρχικού μας demo. Πρέπει να βάλουμε όλη την διαδρομή του φακέλου.

Επίσης αλλάζουμε τον χρόνο που το πρόγραμμα καταγράφει τις φωτογραφίες και τον κάνουμε τρία δευτερόλεπτα. Αυτός είναι ο χρόνος μεταξύ των διαδοχικών φωτογραφιών που αποθηκεύονται στον φάκελο. Την κίνηση μπορεί να την ανιχνεύει, αλλά αν δεν περάσει ο συγκεκριμένος χρόνος δεν μπορεί να κάνει καμία αποθήκευση.

2. Αποθήκευση και διαχείριση εικόνων

Όπως είδαμε και πιο πριν οι εικόνες αποθηκεύονται στον φάκελο testImages με συγκεκριμένη ανάλυση λόγω της κάμερας. Δεν μας ενδιαφέρει τι όνομα έχουν οι εικόνες μιας και ο αλγόριθμος, όπως θα δούμε παρακάτω, μετά την διαδικασία επιστρέφει το όνομα της φωτογραφίας που ταυτοποίησε.

Στον φάκελο TrainingPics βάζουμε εικόνες με ονόματα αρχίζοντας από το 1. Αυτό μας ενδιαφέρει να στείλουμε στην άλλη πλατφόρμα.

Φυσικά δεν πρέπει να ξεχνάμε ότι οι φωτογραφίες στην βάση πρέπει να έχουν την ίδια ανάλυση με τις φωτογραφίες που δεχόμαστε από την κάμερα. Σε αντίθετη περίπτωση, ο αλγόριθμος απλά δε θα κάνει την διαδικασία της αναγνώρισης. Η ανάλυση που χρησιμοποιούμε στη συγκεκριμένη εφαρμογή είναι 320x240.

Εγκατάσταση του Motion

Αρχικά θα πρέπει να τονίσουμε ότι χρησιμοποιήσαμε το λειτουργικό αρχιτεκτονικής linux, Fedora core 3 και εγκαταστήσαμε όλα τα επιπλέον προγράμματα που εμπεριέχονται στα cd της εγκατάστασής του.

Ακολουθεί με βήματα ή όλη διαδικασία για την ολοκλήρωση της εγκατάστασης.

- **Προετοιμασία**

Το Motion χρησιμοποιεί αρκετές βιβλιοθήκες που αναγκαστικά θα πρέπει να εγκατασταθούν στον υπολογιστή προτού γίνει η αντίστοιχη διαδικασία του Motion. Ανάλογα με το τι θέλουμε να πετύχουμε με την συγκεκριμένη εφαρμογή, υπάρχουν και οι αντίστοιχες βιβλιοθήκες για να χρησιμοποιήσουμε. Όπως αν θέλουμε να χρησιμοποιήσουμε δικτυακή κάμερα (net cam), δυνατότητα καταγραφής βίντεο, δυνατότητα αποθήκευσης γεγονότων σε βάσεις MySQL και PostgreSQL και άλλα πολλά.

Στην δικιά μας περίπτωση, που χρειαζόμαστε το Motion μόνο για την αποθήκευση εικόνων στο τοπικό δίσκο, δε χρειάστηκε η εγκατάσταση κανενός άλλου πακέτου για επιπλέον χαρακτηριστικά.

- **Κατέβασμα αρχείων**

Χρησιμοποιήθηκε η έκδοση 3.1.20 του Motion σε μορφή tar (είδος συμπιεσμένου αρχείου). Το αρχείο βρίσκεται στο cd-Demo της εργασίας.

- **Αποσυμπίεση των αρχείων**

Αποσυμπίεσαμε το αρχείο στο μέρος που θέλαμε να γίνει η εγκατάσταση του προγράμματος. Έτσι το βάλουμε στον φάκελο `usr/local`. Πρέπει να τονίσουμε ότι σε περίπτωση που δεν έχουμε δικαιώματα να κάνουμε αυτή την ενέργεια στον συγκεκριμένο φάκελο, μπορούμε να το κάνουμε στον αντίστοιχο που μας επιτρέπεται. Έτσι πηγαίνουμε στον φάκελο `usr/local` και αποσυμπιέζουμε το αρχείο μας.

```
cd /usr/local
tar -xvzf /path/to/motion-3.1.20.tar.gz
```

- Δημιουργήσαμε έτσι έναν φάκελο με το όνομα `motion-3.1.20`. Μπορούμε να μεταομάσουμε τον φάκελο σε `motion`. Έτσι ακολούθως πληκτρολογούμε

```
mv motion-3.1.X motion
```

- Έτσι μπαίνουμε μέσα στον φάκελο που υπάρχουν τα αρχεία εγκατάστασης του προγράμματος πληκτρολογώντας

```
cd motion
```

- Έτσι τώρα είμαστε έτοιμοι να κάνουμε την εγκατάσταση. Πρώτα τρέχουμε το `configure`

```
./configure
```

- Έπειτα δημιουργούμε τον κώδικα πληκτρολογώντας

```
make
```

- και τέλος κάνουμε την εγκατάσταση

```
make install
```

- Κάνοντας την εγκατάσταση επιτυχώς, στην διαδρομή /etc/motion/etc δημιουργείται το αρχείο ASCII που έχουμε πει και προηγουμένως με την ονομασία motion-dist.conf. Εμείς θα πρέπει να μετονομάσουμε αυτό το αρχείο σε motion.conf για να μπορέσει το πρόγραμμα να ξεκινήσει. Φυσικά καλό είναι να σώσουμε το αρχικό αρχείο ASCII σε άλλο μέρος προτού ξεκινήσουμε το motion, μιας και αν διαγραφούν σημαντικά κομμάτια του κώδικα είναι αδύνατη η ολοκλήρωση της διαδικασίας του προγράμματος.

- Για να ξεκινήσει το πρόγραμμα, ανοίγουμε ένα terminal και πληκτρολογούμε

```
motion
```

- Για να σταματήσουμε την διαδικασία απλά γράφουμε στο terminal

```
pkill motion
```

Δυστυχώς δεν υπάρχει γραφικό περιβάλλον για να καταλάβουμε ότι το Motion τρέχει σαν λειτουργία. Οπότε μόνο βλέποντας στο φάκελο που έχουμε ορίσει στο αρχείο ASCII, τις φωτογραφίες από την κάμερα βλέπουμε ότι όλα δουλεύουν κανονικά.

Για να γίνει απεγκατάσταση του προγράμματος από το λειτουργικό περιβάλλον, πηγαίνουμε στο φάκελο που έχουμε κάνει την εγκατάσταση και πληκτρολογούμε

```
make uninstall
```

Έπειτα χρειάζεται να διαγράψουμε οτιδήποτε εμπεριέχεται στον φάκελο /usr/local και φυσικά οποιοδήποτε link εξωτερικού προγράμματος.

3. Αναγνώριση προσώπου

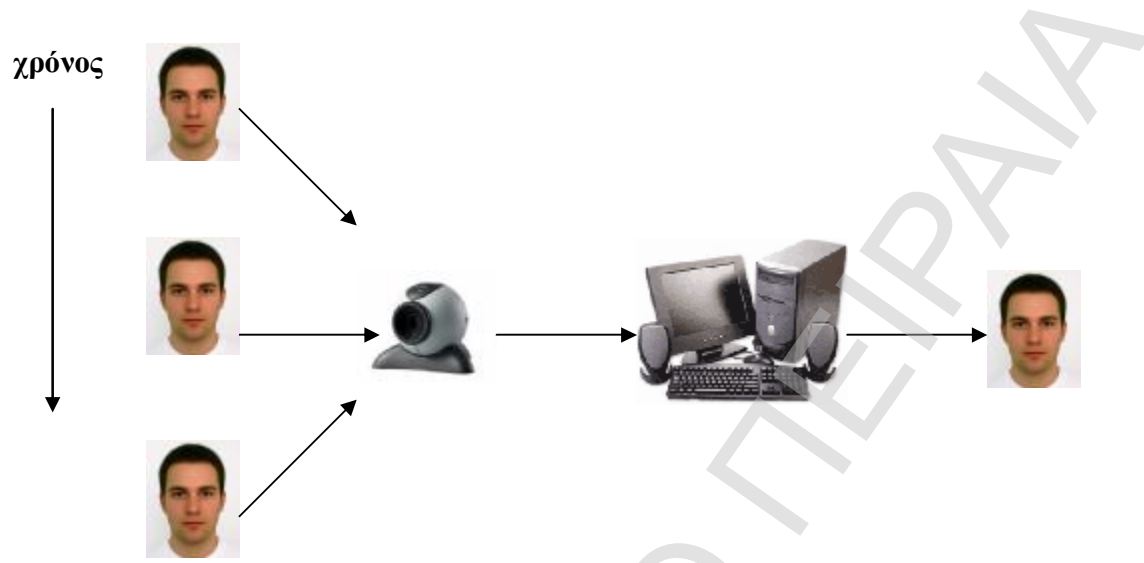
Ας δούμε τώρα ποια είναι η διαδικασία αναγνώρισης της φωτογραφίας. Αρχικά θα πρέπει να τονίσουμε ότι το βασικό αρχείο μας που είναι το FRserver είναι thread και ουσιαστικά ελέγχει συνέχεια τον φάκελο testImages.

Με το που ξεκινάει όμως πρώτα ο αλγόριθμος, ελέγχει τον φάκελο TrainingPics, βλέπει πόσες φωτογραφίες υπάρχουν και ουσιαστικά για κάθε μια δημιουργεί έναν πίνακα με συγκεκριμένα στοιχεία ανάλογα την εικόνα.

Έπειτα για κάθε μια εικόνα που καταγράφεται από το πρόγραμμα motion κάνει την ίδια διαδικασία και έτσι διαπιστώνει ποια εικόνα είναι πιο κοντά στις αρχικές. Μεγαλύτερη έμφαση, δίνεται στην παράγραφο που εξηγούμε τον αλγόριθμο.

Κάθε φορά που έχουμε μια καινούργια φωτογραφία ο αλγόριθμος ελέγχει και συγκρίνει τις εικόνες. Για δικιά μας ασφάλεια για να βγάλει ο αλγόριθμος το όνομα του προσώπου πρέπει να έχει καταγραφεί και αναγνωριστεί τρεις ταυτόχρονες φορές.

Ας δούμε ένα παράδειγμα. Όπως φαίνεται στο παρακάτω σχήμα αρχικά ένα πρόσωπο καταγράφεται από τη κάμερα. Συνεχίζοντας να κινείται μετά από τρία δευτερόλεπτα που ενεργοποιείται το motion επαναλαμβάνεται η διαδικασία. Εάν αυτό γίνει άλλη μια φορά και συνολικά τρεις ταυτόχρονες φορές, τότε και μόνο τότε ο αλγόριθμος μας υποδεικνύει το όνομα της φωτογραφίας.

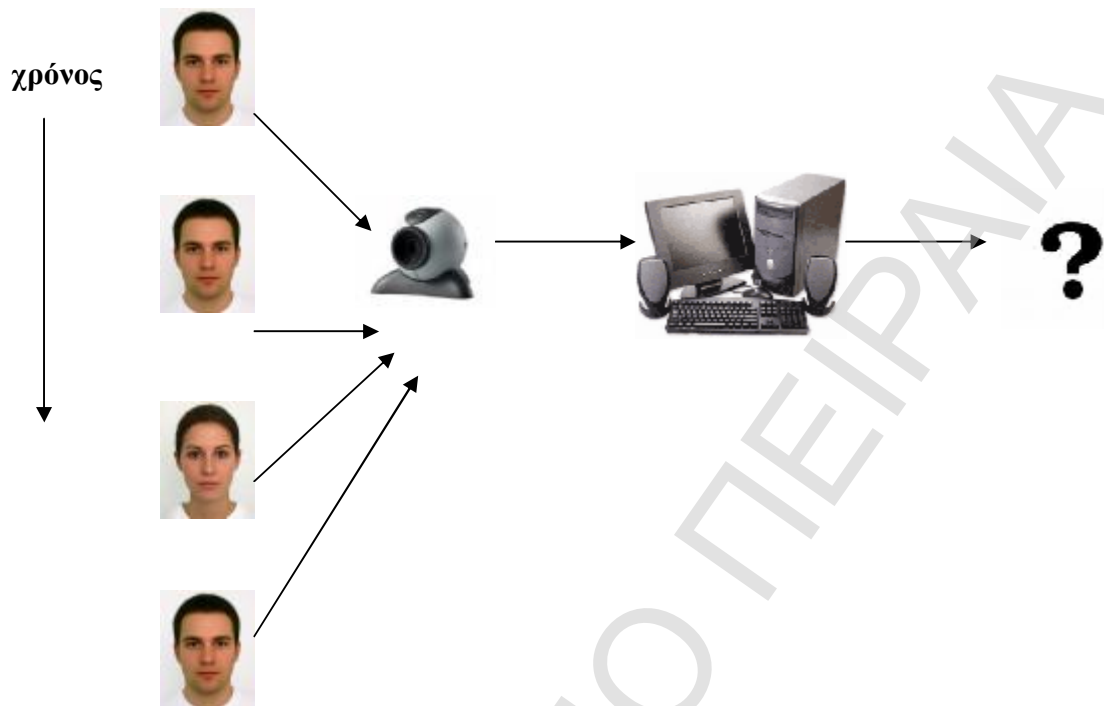


4.2 Λειτουργία σωστής αναγνώρισης

Αυτό γίνεται για να σιγουρέψουμε την αναγνώριση μιας και η επιτυχία του συγκεκριμένου αλγορίθμου αγγίζει το 75%.

Έτσι με αυτό τον τρόπο προσπαθούμε να ανεβάσουμε το ποσοστό επιτυχίας. Φυσικά όσοες περισσότερες φορές αναλύεται το πρόσωπο, τόσο μεγαλύτερη πιθανότητα να είναι σωστή η αναγνώριση έχουμε.

Σε οποιαδήποτε άλλη περίπτωση, ακόμα και τρεις φορές να έχει γίνει η συσχέτιση αλλά να μην είναι ταυτόχρονες, τότε ο αλγόριθμος απλά δεν έχει αποτέλεσμα.



4.3 Λειτουργία λανθασμένης αναγνώρισης

Την ώρα που γίνεται τρεις φορές και έχουμε τη συσχέτιση, στο παραθυράκι που έχει δημιουργηθεί κατά το ξεκίνημα γράφεται ποια εικόνα είναι αυτή. Παρακάτω φαίνεται το σημείο του κώδικα που ελέγχει τρεις φορές αν έχει συμβεί αυτή η διαδικασία.

```

if(sameFaceTimes >= 3){
    System.out.println("IMAGE INDEX = "+lastFaceIndex);
    System.out.println("IMAGE NAME = "+trainingSet.imageFiles[lastFaceIndex].getName());
    String name=trainingSet.imageFiles[lastFaceIndex].getName();
    gui.setMessage("IMAGE NAME = "+name);
    if(name.indexOf(".")!=-1){
        name = name.substring(0,name.indexOf("."));
        faceRecognised(name);
    }
}

```

Βλέπουμε ότι με το που επιτυγχάνεται αυτός ο έλεγχος αμέσως, ενεργοποιείται το επόμενο στάδιο που είναι η αποστολή του ονόματος από μια μέθοδο `faceRecognised` που παίρνει σαν είσοδο τη συγκεκριμένη μεταβλητή `name`.

4. Αποστολή στοιχείων

Η τελευταία διαδικασία είναι της αποστολής δύο στοιχείων. Το πρώτο στοιχείο είναι φυσικά το όνομα του προσώπου, όπως έχει αποθηκευτεί στον φάκελο `TrainingPics` και τον αριθμό της κάμερας. Στην προκειμένη περίπτωση δεν χρησιμοποιήσαμε δεύτερη κάμερα, οπότε δεν υπήρχε και μεταβλητή για τον αριθμό της κάμερας.

Σε περίπτωση που χρησιμοποιήσουμε δεύτερη συσκευή καταγραφής, το μόνο που πρέπει να κάνουμε είναι να πάρουμε τον αριθμό της κάμερας από το `config` αρχείο. Είχαμε όταν αναλύαμε το πρόγραμμα `motion` ότι όταν έχουμε παραπάνω από δύο συσκευές τότε δημιουργούνται και άλλα αρχεία που μέσα στο περιεχόμενό τους έχουν τον αριθμό της κάμερας σαν συσκευή στο `linux`. Δηλαδή `dev0` ή `dev1`.

Η αποστολή των στοιχείων γίνεται μέσω της πλατφόρμας `jadex` που ήδη έχουμε ανεβάσει το κυρίως πρόγραμμα μας.

ΚΕΦΑΛΑΙΟ 5

Συμπεράσματα

Η συγκεκριμένη υλοποίηση όπως είδαμε έρχεται να καλύψει ένα κενό και να οδηγήσει σε ένα επιπλέον καλύτερο σύστημα έξυπνης διαχείρισης ανίχνευσης εικόνων και αναγνώρισης.

Καλό είναι να συνεχιστεί αυτή η αυτοματοποιημένη σκέψη διαχείρισης του συστήματος και σίγουρα έχει ακόμα πολύ δρόμο για την δημιουργία ενός περιβάλλοντος που ο άνθρωπος θα έχει μηδενικό έλεγχο κατά τη διαδικασία του μηχανισμού.

ΠΑΡΑΡΤΗΜΑ Α

Εγχειρίδιο χρήσης

Τα βήματα που πρέπει να ακολουθήσουμε για να εκτελέσουμε με επιτυχία την εφαρμογή είναι τα εξής.

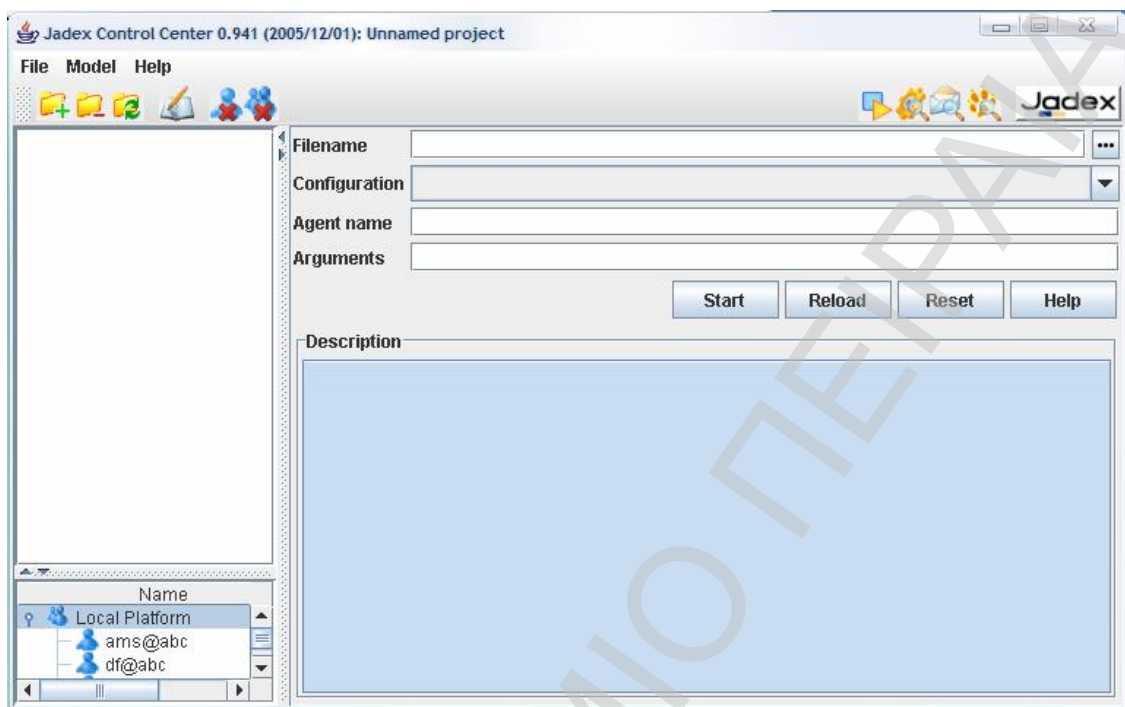
- Εκτέλεση της πλατφόρμας jadex
- Ανεύρεση των xml αρχείων και φόρτωση στην πλατφόρμα jadex
- Ξεκίνημα του agent του απομακρυσμένου υπολογιστή
- Ξεκίνημα του agent του τοπικού υπολογιστή

Ας τα δούμε αναλυτικά παρακάτω με τη σειρά με τη βοήθεια των εικόνων screenshots.

Εκτέλεση της πλατφόρμας jadex

Για να ξεκινήσει η πλατφόρμα jadex, αρκεί να πληκτρολογήσουμε στο terminal το εξής

```
java jadex.adapter.standalone.Platform
```



Περιβάλλον Jadex

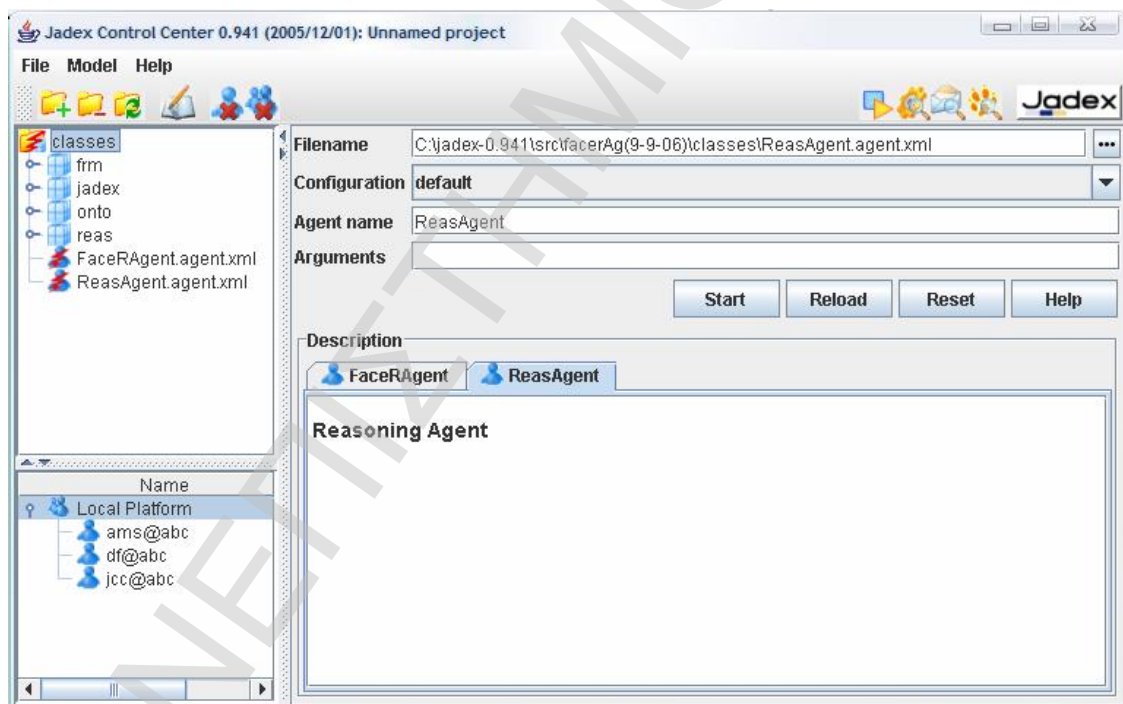
Τότε εμφανίζεται το παραπάνω περιβάλλον. Όπως βλέπουμε κάτω αριστερά υπάρχει μια λίστα με τους agents που ήδη υπάρχουν στο σύστημα. Όταν θα προσθέσουμε τους δικούς μας, τότε θα εμφανιστούν σε αυτό το μέρος.

Πάνω αριστερά μπορούμε να προσθέσουμε τη διαδρομή που βρίσκεται το δικό μας xml αρχείο. Ενώ στο δεξιό κομμάτι έχουμε το σημείο όπου προσθέτουμε τα αρχεία μας και από κάτω μια μικρή περιγραφή, ανάλογα τι έχουμε γράψει στο κομμάτι του xml αρχείου.

Ανεύρεση των xml αρχείων και φόρτωση στην πλατφόρμα jadex

Μόλις γίνει η ανεύρεση των xml αρχείων, είτε από το filename είτε βρίσκοντας τα από τη διαδρομή αριστερά, βλέπουμε ότι έχουν φορτωθεί σωστά στην πλατφόρμα μας και είμαστε έτοιμοι να τα ξεκινήσουμε. Σε περίπτωση που υπήρχε κάποιο πρόβλημα τότε στο πεδίο της περιγραφής, θα έδειχνε το συγκεκριμένο πρόβλημα και όχι την περιγραφή του agent.

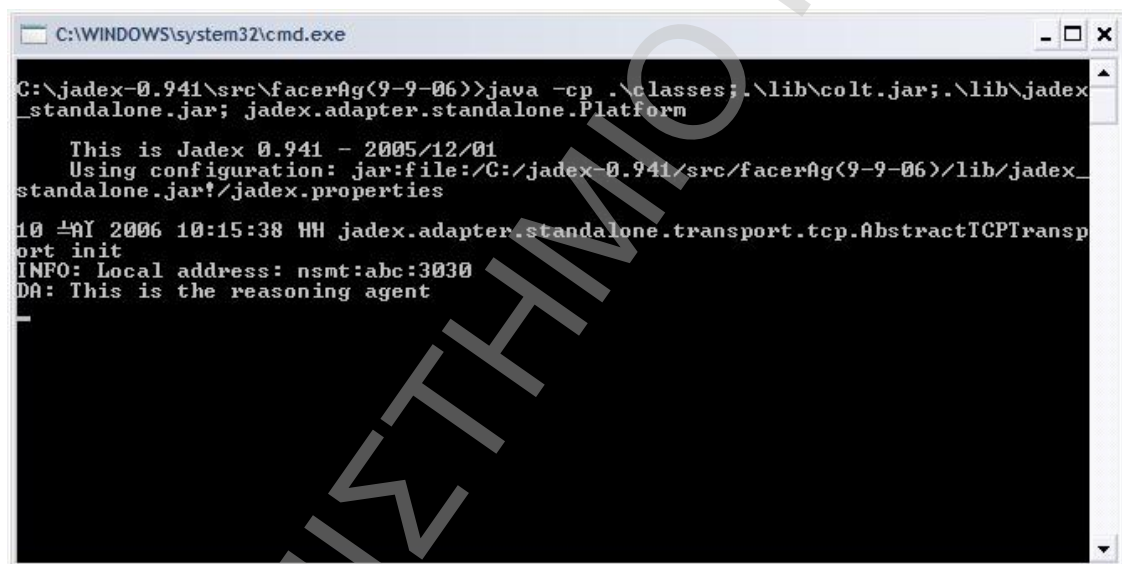
Στην εικόνα που ακολουθεί βλέπουμε ένα στιγμιότυπο της πλατφόρμας λίγο πριν ξεκινήσουν οι agents.



Ξεκίνηση του agent του απομακρυσμένου υπολογιστή

Πρώτα πρέπει να ξεκινήσουμε τον agent του απομακρυσμένου υπολογιστή, έτσι ώστε να είναι έτοιμο να δεχτεί αναπάσα στιγμή τα δεδομένα. Έτσι πατάμε το κουμπί Start για να ξεκινήσει η διαδικασία. Με το που το κάνουμε αυτό αμέσως στο terminal του απομακρυσμένου αναγράφεται

RA : This is the reasoning agent, όπως βλέπουμε και παρακάτω



```
C:\WINDOWS\system32\cmd.exe
C:\jadex-0.941\src\facerAg(9-9-06)>>java -cp .\classes;.\lib\colt.jar;.\lib\jadex_
_standalone.jar; jadex.adapter.standalone.Platform
    This is Jadex 0.941 - 2005/12/01
    Using configuration: jar:file:/C:/jadex-0.941/src/facerAg(9-9-06)/lib/jadex_
_standalone.jar!/jadex.properties
10 ±ΠΙ 2006 10:15:38 HH jadex.adapter.standalone.transport.tcp.AbstractTCPTransp
ort init
INFO: Local address: nsmt:abc:3030
DA: This is the reasoning agent
```

Διαδικασία αναμονής

Έτσι ο απομακρυσμένος agent είναι έτοιμος για να δεχτεί τα δεδομένα. Απλά είναι σε κατάσταση αναμονής από τον άλλο agent.

Ξεκίνηση του agent του τοπικού υπολογιστή

Οπότε η τελευταία διαδικασία, είναι να ξεκινήσουμε το agent για να αρχίσει η κύρια μας εφαρμογή.

Με το που πατάμε το Start, βλέπουμε από την παρακάτω εικόνα ότι ξεκινάει η διαδικασία του αλγορίθμου, διαβάζοντας τις φωτογραφίες που υπάρχουν ήδη στον φάκελο TrainingPics. Θα παρατηρήσουμε ότι όσες περισσότερες βρίσκονται, τόσο πιο πολύ ώρα παίρνει για την εφαρμογή να έρθει σε κατάσταση αναμονής φωτογραφιών από τη κάμερα.

```
C:\WINDOWS\system32\cmd.exe
Total number of training images : 5
C:\jadex-0.941\src\face\ag(9-9-06)\TrainingPics\1.jpg
w=320h=0
w=320h=240
1
6
C:\jadex-0.941\src\face\ag(9-9-06)\TrainingPics\2.jpg
w=320h=240
w=320h=240
2
6
C:\jadex-0.941\src\face\ag(9-9-06)\TrainingPics\3.jpg
w=320h=240
w=320h=240
3
6
C:\jadex-0.941\src\face\ag(9-9-06)\TrainingPics\4.jpg
w=320h=240
w=320h=240
4
6
C:\jadex-0.941\src\face\ag(9-9-06)\TrainingPics\5.jpg
w=320h=240
w=320h=240
5
```

Διαδικασία ελέγχου της βάσης

Μετά από αυτή τη διαδικασία, θα δούμε ότι εμφανίζεται ένα παραθυράκι “Face Recognition”, στο οποίο ουσιαστικά θα φανεί η αναγνώριση του προσώπου



Αναμονή αναγνώρισης

Ενώ στο terminal, όπως φαίνεται και παρακάτω εμφανίζονται οι λειτουργίες του αλγόριθμου, όπως η παρακολούθηση του φακέλου και διαδικασία τις ταυτοποίησης της εικόνας

```
C:\WINDOWS\system32\cmd.exe
Rows = 6 Columns = 6
0.0
1.5533408987488093E9
0.0
1.4100573433953056E9
1.671472908550286E9
1.403286049261468E9
Image -> 0
sameFaceTimes = 1
Ignoring file Thumbs.db
2 files in folder
File 1.jpg deleted
1 files in folder
1 files in folder
1 files in folder
File Thumbs.db deleted
0 files in folder
0 files in folder
0 files in folder
0 files in folder
0 files in folder
0 files in folder
0 files in folder
0 files in folder
0 files in folder
0 files in folder
```

Διαδικασία αναγνώρισης

Με το που γίνει η διαδικασία και η εικόνα αναγνωριστεί τρεις φορές, τότε στο αρχικό περιβάλλον αναγράφεται το όνομα της εικόνας που έχει ουσιαστικά συσχετιστεί.

Αυτό το όνομα είναι βασικά η ονομασία της αρχικής εικόνας που βρίσκεται στον φάκελο της βάσης μας και όχι της εικόνας που αποθηκεύεται.



Συσχέτιση εικόνων

Αμέσως μετά την αναγνώριση γίνεται η αποστολή του στοιχείου, όποτε στον άλλο υπολογιστή ουσιαστικά θα φανεί το όνομα του χρήστη, δηλαδή στο παράδειγμά μας το 1.jpg, όπως φαίνεται παρακάτω.

A screenshot of a Windows command prompt window titled 'C:\WINDOWS\system32\cmd.exe'. The window has a black background with white text. The text shows various numerical values, status messages, and file names related to a face recognition process. The text is as follows:

```
-6.361833E+008
Rows = 5 Columns = 5
0.0
1.5290564816682694E9
1.1419808812572343E9
1.4121616209815464E9
1.6735195202666106E9
Image -> 0
=====
sameFaceTimes = 3
IMAGE INDEX = 0
IMAGE NAME = 1.jpg
FRA: PlanSendRecognisedFace Started
SEND FACE PLAN
FRA: Ready to send message
DA: USER ID = 1 from camera ID = 1
FRA: Face id Recognised = 1
File 1.jpg deleted
0 files in folder
Face belief = 0
0 files in folder
Face belief = 0
0 files in folder
Face belief = 0
```

Ένδειξη αποστολής στοιχείων

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Βιβλιογραφικές Αναφορές

- [1] <http://www.gnu.org/licenses/gpl.html>
- [2] <http://www.lavrsen.dk/twiki/bin/view/Motion/FreeBSD>
- [3] <http://www.lavrsen.dk/twiki/bin/view/Motion/MotionTracking>
- [4] <http://www.lavrsen.dk/twiki/bin/view/Motion/RelatedProjects>
- [5] <http://www.mic.atr.co.jp/events/fg98/>
- [6] History of neural networks, Gregory Tambasis, 1999, University of Wales, Swansea.
- [7] http://www.identix.com/company/comp_history.html
- [8] <http://www.viisage.com/ww/en/pub/home.cfm>
- [9] http://findarticles.com/p/articles/mi_m0EKF/is_1998_Nov_23/ai_53260935
- [10] A tutorial on Principal Components Analysis, Lindsay I Smith, February 26, 2002
- [11] Y. Li, S. Gong, and H. Liddell. Video-based online face recognition using identity surfaces. Technical report, Queen Mary, University of London, 2001.
- [12] R. Gross, Face Databases, Handbook of Face Recognition, Stan Z. Li and Anil K. Jain, ed., Springer-Verlag, February 2005
- [13] W. Zhao, R. Chellappa, A. Rosenfeld and P.J. Phillips. Face Recognition: A literature survey. Technical Report CART-TR-948. University of Maryland, Aug. 2002.

[14] L. Torres, J. Vilà, “Automatic Face Recognition for Video Indexing Applications”, Pattern Recognition, Vol 35/3, pp. 615-625, December 2001.

[15] T. Choudhury, B. Clarkson, T. Jebara and A. Pentland, “Multimodal Person Recognition Using Unconstrained Audio and Video”, International Conference on Audio and Video-based Biometric Authentication, pp, 176-181, Washington D.C., 1999.

[16] A. Albiol, L. Torres, E Delp, “The Indexing of Persons in News Sequences using Audio-visual Data”, International Conference on Acoustics, Speech and Signal Processing, Hong-Kong, China, April 6-10, 2003.

[17] <http://dsd.lbl.gov/~hoschek/colt/>