



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

Διπλωματική Εργασία
της
Λοΐζου Κωνσταντίνας

**Σχεδιασμός και κατασκευή εφαρμογής
ταξινόμησης αντικειμένων σε γραμμή
μεταφοράς προϊόντων με χρήση όρασης
μηχανής**

Επιβλέπων Καθηγητής: Δρ. Μηνάς Δασυγένης

Κοζάνη 2014

© Λοΐζου Κωνσταντίνα
© Πανεπιστήμιο Δυτικής Μακεδονίας

**ΣΧΕΔΙΑΣΜΟΣ ΚΑΙ ΚΑΤΑΣΚΕΥΗ ΕΦΑΡΜΟΓΗΣ ΤΑΞΙΝΟΜΗΣΗΣ
ΑΝΤΙΚΕΙΜΕΝΩΝ ΣΕ ΓΡΑΜΜΗ ΜΕΤΑΦΟΡΑΣ ΠΡΟΪΟΝΤΩΝ ΜΕ ΧΡΗΣΗ
ΟΡΑΣΗΣ ΜΗΧΑΝΗΣ**

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Δασυγένης Μηνάς

Λέκτορας

Φαχαντίδης Νικόλαος

Αναπληρωτής Καθηγητής

*Στην
οικογένεια μου*

Πίνακας Περιεχομένων

Πρόλογος.....	11
Abstract.....	13
Εισαγωγή.....	15
1.1 Κίνητρο Διεξαγωγής Εργασίας	15
1.2 Σκοπός και στόχοι της εργασίας.....	16
1.3 Συνεισφορά στην επιστημονική κοινότητα.....	16
1.4 Θεωρητικό Υπόβαθρο.....	17
1.4.1 Όραση Μηχανής	17
1.4.2 Όραση μηχανής σε συστήματα παραγωγής	23
1.5 Οργάνωση και δομή της εργασίας.....	25
Κατάτμηση Εικόνας και Παρακολούθηση Αντικειμένων.....	27
2.1 Κατάτμηση Εικόνας.....	27
2.1.1 Μετασχηματισμός Κλιμακωτά - Αναλλοίωτων Χαρακτηριστικών (SIFT)	28
2.1.2 Επιταχυνθέντα ισχυρά χαρακτηριστικά (SURF).....	30
2.1.3 Αλγόριθμος Εξαγωγής Χαρακτηριστικών από γρήγορο τμήμα δοκιμής για Ανίχνευση Ακμών (FAST).....	32
2.1.4 Αλγόριθμος Κατάτμησης εικόνας με βάση το χρώμα.....	33
2.2 Ανίχνευση και Παρακολούθηση αντικειμένων.....	34
2.2.1 Ανίχνευση Κινούμενου Αντικειμένου.....	34
2.2.2 Αφαίρεση Φόντου.....	34
2.2.3 Στατιστικές μέθοδοι.....	35
2.2.4 Χρονική Διαφορά.....	35
2.2.5 Οπτική Ροή.....	36
Χρώμα και κατάτμηση.....	37
3.1 Χρωματικοί Χώροι.....	37
3.1.1 Το χρωματικό μοντέλο RGB.....	39
3.1.2 Ο χρωματικός χώρος nRGB ή normalized RGB.....	40
3.1.3 Ο Χρωματικός χώρος YUV.....	41
3.1.4 Ο Χρωματικός χώρος YCbCr.....	43
3.1.5 Περιληπτική σύγκριση χρωματικών χώρων	43
3.2 Κατάτμηση Εικόνας με βάση το Χρώμα.....	44
3.2.1 Τεχνικές κατάτμησης εικόνας με βάση το χρώμα.....	44
3.2.2 Μέθοδοι Κατάτμησης με βάση το χρώμα.....	45
Υλικό και Λογισμικό	51
4.1 Υλικό Μέρος Ανάπτυξης εφαρμογής.....	51
4.1.1 Κάμερα.....	51
4.1.2 Ηλεκτρονικός Υπολογιστής.....	52
4.2 Λογισμικό Μέρος Ανάπτυξης εφαρμογής	53
4.2.1 Λειτουργικό Σύστημα Ubuntu.....	53
4.2.2 Πλατφόρμα ανάπτυξης λογισμικού Eclipse CDT.....	53
4.2.3 Γλώσσα Προγραμματισμού C++.....	56
4.2.4 Η Βιβλιοθήκη όρασης μηχανής OpenCV	56
4.2.5 Χρησιμοποιώντας την OpenCV στην πλατφόρμα Eclipse	59
Μεθοδολογία Υλοποίησης της Εφαρμογής.....	63
5.1 Πρόγραμμα Εργασιών υλοποίησης.....	63
5.2 Αντικείμενο Ταξινόμησης στην γραμμή μεταφοράς.....	64

5.3 Επιλογή Μεθόδου Κατάτμησης Εικόνας	64
5.4 Σχεδιασμός και Υλοποίηση της Εφαρμογής.....	65
5.4.1 Έλεγχοι ταξινόμησης αντικειμένων	70
5.4.2 Προβλήματα που προέκυψαν και η αντιμετώπιση τους.....	73
5.5 Εκτέλεση εφαρμογής από γραμμή εντολών (terminal).....	75
Αποτελέσματα Εφαρμογής	77
6.1 Σενάρια Ελέγχου Δοκιμών	77
6.1.1 Αποτελέσματα Δοκιμών Συστατικών Στοιχείων	78
6.1.2 Αποτελέσματα Δοκιμών Τελικής Έκδοσης.....	84
6.2 Λειτουργίες Qt * GUI	88
6.3 Ορίσματα γραμμής εντολών	89
Συμπεράσματα και Μελλοντικές Προεκτάσεις.....	91
7.1 Συμπεράσματα.....	91
7.2 Μελλοντικές Προεκτάσεις.....	92
ΑΠΟΔΟΣΗ ΑΓΓΛΙΚΩΝ ΟΡΩΝ ΣΤΗΝ ΕΛΛΗΝΙΚΗ ΓΛΩΣΣΑ.....	95
Βιβλιογραφία.....	97

Κατάλογος Πινάκων

Πίνακας 1 Τα χρώματα και τα μήκη κύματος που αντιστοιχούν σε κάθε τύπο κωνίου.	38
Πίνακας 2 Χρώματα του ορατού φάσματος.....	38
Πίνακας 3 Χρωματικός Πίνακας RGB [15].	66

Κατάλογος Σχημάτων

Σχήμα 1 Σχέση ανάμεσα στην Υπολογιστική Όραση και άλλα επιστημονικά πεδία.....	20
Σχήμα 2 Αρχιτεκτονική κατάτμησης εικόνας με βάση το χρώμα.	45
Σχήμα 3 Βασική δομή της OpenCV.....	57
Σχήμα 4 Σχεδιασμός συστήματος εφαρμογής ταξινόμησης αντικειμένων σε γραμμή μεταφοράς.	65
Σχήμα 5 (α) Αντικείμενο βάση των προδιαγραφών (β) Μη αποδεκτό αντικείμενο.....	72
Σχήμα 6 (α) Αντικείμενα που ελέγχονται στο σημείο με διακεκομμένη γραμμή (β) Αντικείμενο μη αποδεκτό που κατά την είσοδο του φαίνεται αποδεκτό.	74
Σχήμα 7 Σενάριο ύπαρξης τριών αντικειμένων στη γραμμή μεταφοράς.....	85
Σχήμα 8 Σενάριο ύπαρξης πέντε αντικειμένων στη γραμμή μεταφοράς.	86

Κατάλογος Εξισώσεων

Εξίσωση 1 Οι νέες τιμές που θα πάρει η συνιστώσα (r,g,b).	40
Εξίσωση 2 Σχέση μετατροπής από RGB σε YUV.	42
Εξίσωση 3 Σχέση για μετατροπή RGB σε YCbCr.....	43
Εξίσωση 4 Τύπος για την εύρεση της διακύμανσης του κέντρου βάρους.....	46
Εξίσωση 5 Εξίσωση εκπαίδευσης λαθών σε τεχνητά νευρωνικά δίκτυα.....	49

Κατάλογος Εικόνων

Εικόνα 1 Χρωματικές ομοιότητες αντικειμένων και φόντου [23].	18
Εικόνα 2 Ανίχνευση και καταμέτρηση αιμοφόρων αγγείων σε εικόνες [8].	22
Εικόνα 3 Υπολογισμός του μεγίστου και ελαχίστου της Γκαουσιανής διαφοράς σε ένα εικονοστοιχείο.	29
Εικόνα 4 Προσέγγιση Λαπλασιανής της Γκαουσιανής με φίλτρο κουτιών.	31
Εικόνα 5 Το σημείο ενδιαφέροντος μιας εικόνας κάτω από δοκιμή και τα 16 εικονοστοιχεία στον κύκλο [1].	33
Εικόνα 6 Το χρωματικό μοντέλο RGB.	40
Εικόνα 7 Μετασχηματισμός εικόνας από RGB σε nRGB [4].	41
Εικόνα 8 (a) Η RGB εικόνα (b) Το στοιχείο Y (c) Το στοιχείο U (d) Το στοιχείο V [9].	42
Εικόνα 9 (a) Η εικόνα χωρίς επεξεργασία (b) Το αποτέλεσμα αν στην εικόνα (a) χρησιμοποιήσουμε την τεχνική ανίχνευσης ακμών.	48
Εικόνα 10 Logitech QuickCam.	51
Εικόνα 11 Το περιβάλλον του Terminal των Ubuntu.	54
Εικόνα 12 Πλατφόρμα ανάπτυξης λογισμικού Eclipse.	55
Εικόνα 13 Προσθήκη φακέλου OpenCV στο Eclipse.	60
Εικόνα 14 Εισαγωγή βιβλιοθηκών και διαδρομής στο Eclipse.	60
Εικόνα 15 Το αντικείμενο που ελέγχεται στη γραμμή μεταφοράς.	64
Εικόνα 16 Λειτουργία της cvFindContours() [14].	68
Εικόνα 17 Μήνυμα λάθους.	78
Εικόνα 18 (α) Εικόνα εισόδου αντικειμένου με βάση τις προδιαγραφές (β) συνιστώσα Y (γ) συνιστώσα Cb (δ) συνιστώσα Cr.	79
Εικόνα 19 Εικόνα εισόδου ελαττωματικού αντικειμένου (β) συνιστώσα Y (γ) συνιστώσα Cb (δ) συνιστώσα Cr.	80
Εικόνα 20 Ανίχνευση αντικειμένου βάση των προδιαγραφών και σχεδιασμός περιγράμματος γύρω από αυτό.	81
Εικόνα 21 Μήνυμα εξόδου.	81
Εικόνα 22 Ανίχνευση ελαττωματικού αντικειμένου και σχεδιασμός περιγράμματος γύρω από αυτό.	82
Εικόνα 23 Μήνυμα Εξόδου.	82
Εικόνα 24 Ανίχνευση δυο αντικειμένων και εξαγωγή περιγραμμάτων γύρω από αυτά.	83
Εικόνα 25 Μήνυμα Εξόδου.	83
Εικόνα 26 Μήνυμα Εξόδου.	85
Εικόνα 27 Μήνυμα για ταξινόμηση πρώτου αντικειμένου.	86
Εικόνα 28 Μήνυμα για ταξινόμηση και καταμέτρηση δύο αντικειμένων.	87
Εικόνα 29 Μήνυμα για ταξινόμηση και καταμέτρηση τριών αντικειμένων.	87
Εικόνα 30 Μήνυμα για ταξινόμηση και καταμέτρηση όλων των αντικειμένων της γραμμής μεταφοράς.	87
Εικόνα 31 Παράθυρο εικόνας εισόδου με λειτουργίες QT.	88

Πρόλογος

Το αντικείμενο της διπλωματικής εργασίας είναι η όραση μηχανής. Με την έλευση υπολογιστών υψηλών ταχυτήτων και καμερών υψηλής ανάλυσης και ευκρίνειας, επεκτάθηκε το ερευνητικό ενδιαφέρον για ανάπτυξη και υλοποίηση αλγορίθμων στον τομέα της μηχανικής όρασης. Οι εφαρμογές αλγορίθμων ποικίλουν ανάλογα με τον τομέα ζήτησής τους. Εφαρμογές μπορούν να αναπτυχθούν με όραση μηχανής για θέματα ασφαλείας, στρατιωτικές και κατασκοπευτικές εφαρμογές, για θέματα υγείας και έρευνας αλλά και εφαρμογές σε θέματα παραγωγής. Ανάμεσα στις εφαρμογές αυτές, μια από τις πιο συχνές εργασίες που εκτελούνται, είναι ο καθορισμός του αριθμού της θέσης αλλά και η ανίχνευση της κίνησης για το κάθε αντικείμενο, στοιχεία τα οποία μας απασχόλησαν στο κομμάτι της εφαρμογής μας.

Στην εποχή μας, η εξωτερική αισθητική και ακεραιότητα των προϊόντων λαμβάνεται σε ίσο βαθμό υπόψη με το ίδιο το προϊόν από το καταναλωτικό κοινό, στοιχείο που απασχολεί ιδιαίτερα τους κατασκευαστές. Για να διασφαλιστεί η μακροχρόνια προώθηση ενός αγαθού, πρέπει η εμφάνιση και η ποιοτική ακεραιότητα του προϊόντος να ελεγχθεί οπτικά πριν την αποστολή του. Το επιθυμητό αυτό αποτέλεσμα μπορεί να διασφαλιστεί αν η διαδικασία ελέγχου γίνει αυτοματοποιημένη και επιτελείται χωρίς ανθρώπινη παρέμβαση.

Με τη χρήση των τεχνικών που προσφέρονται σε θέματα μηχανικής όρασης, σχεδιάστηκε και υλοποιήθηκε μια εφαρμογή ταξινόμησης προϊόντων σε γραμμή μεταφοράς με τη χρήση κάμερας, δηλαδή τη χρήση μιας αλληλουχίας εικόνων. Η εφαρμογή αυτή δίνει τη δυνατότητα στον κάτοχο της να έχει μια πλήρη εικόνα καταγραφής της ποιότητας και της ποσότητας των προϊόντων του, αποφεύγοντας έτσι τυχόν ανθρώπινα λάθη.

Στο σημείο αυτό θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή Δρ. Μηνά Δασυγένη για την πολύτιμη βοήθειά του κατά τη διάρκεια υλοποίησης της διπλωματικής εργασίας και για τις γνώσεις που απέκτησα χάρη σε αυτόν. Το θέμα της διπλωματικής εργασίας προτάθηκε από τον Δρ. Μηνά Δασυγένη και η εφαρμογή αναπτύχθηκε σε γλώσσα προγραμματισμού C++ με τη χρήση των βιβλιοθηκών της OpenCV.

Λέξεις Κλειδιά: OpenCV , C++, Όραση Μηχανής, Γραμμή παραγωγής.

Κοζάνη, Ιούνιος 2014

Λοΐζου Κωνσταντίνα

Abstract

The object of this thesis is computer vision. With the advent of high-speed computers and cameras with high resolution and definition, the research interest was expanded for the development and implementation of algorithms in the field of computer vision. Algorithms for applications vary depending on their field of interest. Applications can be developed with computer vision for security, military and intelligence, for health issues, research and production. Among these applications, one of the most frequently performed task is to determine the position and the motion detection for each subject, all of which occupied us to develop our application.

Nowadays, the external aesthetics and integrity of the products is obtained in an equal degree with regard to the product itself from the general public, a fact that preoccupies manufacturers. To ensure long-term promotion of a product, the appearance and quality product integrity must be visually inspected before dispatch. This desired result can be achieved if the process inspection become automated and is performed without human intervention.

By using the techniques offered on machine vision tasks, we designed and implemented an application for product classification in a transmission line using camera, i.e. the use of a sequence of images. This application enables the holder to have a complete picture regarding quality and quantity of the products thus avoiding any human errors.

At this point I would like to thank my supervisor Dr. Mina Dasygeni for his valuable help during the implementation of this thesis and for the knowledge I gained thanks to him. The topic of the thesis was proposed by Dr. Minas Dasygenis and the application was developed with programming language C++ using the libraries of OpenCV.

Keywords: OpenCV, C + +, Computer Vision, Production line.

Kozani, June 2014

Loizou Konstantina

1

Εισαγωγή

Η διπλωματική αυτή εργασία, αφορά τον σχεδιασμό και την κατασκευή μιας εφαρμογής ταξινόμησης αντικειμένων σε μια γραμμή μεταφοράς, με τη χρήση τεχνικών όρασης μηχανής και επεξεργασίας εικόνας από βίντεο.

Στο κεφάλαιο αυτό, θα εισάγουμε την έννοια της όρασης μηχανής, τις εφαρμογές της αλλά και το λόγο που επιλέχθηκε το συγκεκριμένο γνωστικό αντικείμενο σαν θέμα της διπλωματικής εργασίας.

1.1 Κίνητρο Διεξαγωγής Εργασίας

Η όραση μηχανής είναι ένας από τους τομείς της τεχνολογίας, ο οποίος τα τελευταία χρόνια έχει μεγάλη ανάπτυξη. Οι ψηφιακές εικόνες και τα βίντεο είναι πλέον μέρος της καθημερινότητάς μας. Η ανάπτυξη χρήσιμων εφαρμογών ρομποτικής, ιατρικής, ασφάλειας αλλά και βιομηχανικού αυτοματισμού είναι πλέον αναγκαία.

Η περιέργειά μου για το πως λειτουργεί η αναγνώριση αντικειμένων αλλά και το "πάντρεμα" της όραση μηχανής με τα μαθηματικά, τη φυσική, τη βιολογία, τη μηχανική αλλά και την επιστήμη των υπολογιστών, είναι μερικοί από τους λόγους που με ώθησαν να ασχοληθώ με το συγκεκριμένο γνωστικό αντικείμενο και τη διεκπεραίωση της εργασίας αυτής. Τέλος, ένα από τα σημαντικότερα κίνητρά μου ήταν η κατάρτιση που θα λάμβανα σε ένα αντικείμενο άγνωστο μέχρι τότε σε εμένα, που θα αποτελέσει ένα επιπλέον εφόδιο για την πορεία μου μετά το πανεπιστήμιο.

1.2 Σκοπός και στόχοι της εργασίας

Ο σκοπός του σχεδιασμού και της υλοποίησης της εφαρμογής αυτής έχει να κάνει αφενός με την προσπάθεια συμβολής στην έρευνα πάνω στο συγκεκριμένο γνωστικό πεδίο και αφετέρου με τον εμπλουτισμό των γνώσεών μου και των εμπειριών μου κατά τη διάρκεια υλοποίησης της.

Στόχος της εργασίας αυτής είναι η κατασκευή μιας εφαρμογής, η οποία θα μπορεί εύκολα να χρησιμοποιηθεί για την αναγνώριση και ταξινόμηση αντικειμένων σε γραμμή μεταφοράς, από ένα ευρύ κοινό που δεν έχει προηγμένες γνώσεις στην επιστήμη των υπολογιστών. Σχεδιαστικός στόχος είναι το σύστημα να βασίζεται σε μια ρεαλιστική γραμμή μεταφοράς, ανιχνεύοντας και ταξινομώντας υπαρκτά αντικείμενα και στη συνέχεια να δοκιμαστεί σε συνθήκες πραγματικού χρόνου. Η εφαρμογή αυτή μπορεί να χρησιμοποιηθεί εύκολα για τον ποιοτικό έλεγχο προϊόντων.

1.3 Συνεισφορά στην επιστημονική κοινότητα

Στις μέρες μας, όλο και περισσότεροι επιστήμονες ασχολούνται με τον τομέα της όρασης μηχανής. Υπάρχουν εκατοντάδες παραδείγματα εφαρμογών που αναπτύχθηκαν γύρω από το συγκεκριμένο αντικείμενο. Σε αυτήν τη διπλωματική εργασία ασχοληθήκαμε με δύο προβλήματα, την ανίχνευση αντικειμένων σε διαδοχικές εικόνες και την ταξινόμησή τους. Η μέθοδος ανίχνευσης που χρησιμοποιήσαμε, είναι η κατάτμηση εικόνας με βάση το χρώμα και παρουσιάζεται στο κεφάλαιο 2. Η παρακολούθηση των αντικειμένων γίνεται σε διαδοχικά καρέ και από αυτά εξάγουμε πληροφορίες για το εάν ένα αντικείμενο πληροί ή όχι τις προδιαγραφές που θέσαμε (βλ. Κεφάλαιο 6). Το πρόγραμμα που κατασκευάσαμε χαρακτηρίζεται από την ευχρηστία του, την απλότητα στη χρήση και την απόδοση του.

Πολλά είναι τα παραδείγματα εφαρμογών και έρευνας που βασίζονται στην όραση μηχανής και την κατάτμηση εικόνας. Οι Mohammad F. Hashmi και Avinash G. Keskar στο άρθρο τους [24], παρουσιάζουν τον αλγόριθμο και την εφαρμογή που δημιούργησαν για οπτική επιθεώρηση των σιδηροδρομικών γραμμών. Εκτός από τη μέθοδο τρόπου υλοποίησης, η εφαρμογή αυτή διαφέρει από τη δική μας στο ανιχνεύσιμο σχήμα του αντικείμενου που είναι γνωστό, και από αυτό παίρνουν πληροφορίες για την κατάστασή του.

Ένα άλλο παράδειγμα χρήσης όρασης μηχανής παρουσιάζεται στην εργασία του K.M. Sajjad [25], όπου με τη χρήση όρασης μηχανής κάνει αυτόματη αναγνώριση αριθμών σε πινακίδες αυτοκινήτων. Η κατάτμηση εικόνας γίνεται με βάση το προκαθορισμένο σχήμα των πινακίδων και ακολουθεί η αναγνώριση χαρακτήρων. Όλοι οι χαρακτήρες είναι αποθηκευμένοι σε μια λίστα και από αυτή γίνεται σύγκριση με το κάθε νέο χαρακτήρα που εξάγεται από τις πινακίδες. Η κατάτμηση σε αντίθεση με τη δική μας εργασία γίνεται σε αποχρώσεις του γκρι και η εικόνα εισόδου είναι στατική.

Η Ευαγγελία Α. Κατσαούνου στη διπλωματική της εργασία [26] υλοποίησε ένα σύστημα αυτόματης ανίχνευσης και αναγνώρισης στατικών ανθρώπινων χειρονομιών με την τεχνική σύμπτωσης προτύπων για αναγνώριση των χειρομορφών. Στη διπλωματική αυτή, σε αντίθεση με τη δική μας, το ζητούμενο αντικείμενο ταξινομείται βάση της σύγκρισης του με προκαθορισμένα πρωτότυπα χειρομορφών.

Στη διπλωματική εργασία του Πρωτονοτάριου Ιωάννη [27] έγινε μελέτη τεχνικών αναγνώρισης φωτιάς με χρήση όρασης μηχανής μέσω δικτυακής κάμερας. Η αναγνώριση έγινε με την ανίχνευση του χρώματος της φωτιάς σε κάποιο από τα εικονοστοιχεία (pixel) της οθόνης, το οποίο κινείται με ρυθμό 10Hz. Το αντικείμενο αναζήτησης δεν ταξινομείται με βάση το περίγραμμα του όπως στην δική μας περίπτωση αλλά, ακόμα και η πιο μικρή περιοχή εικονοστοιχείων που πληροί τις παραπάνω προϋποθέσεις αναγνωρίζεται ως φωτιά.

1.4 Θεωρητικό Υπόβαθρο

1.4.1 Όραση Μηχανής

Η Όραση Μηχανής ή Υπολογιστική Όραση είναι ο τομέας της επιστήμης υπολογιστών, ο οποίος προσπαθεί να δώσει στους υπολογιστές ή σε άλλες μηχανές ανάλογες δυνατότητες με αυτές που έχει και ο άνθρωπος, δηλαδή την ενσωμάτωση της λογικής με την όραση. Αυτό το επιτυγχάνει μέσω της εξόρυξης δεδομένων και της ανάλυσης και επεξεργασίας εικόνας, κινούμενης ή στατικής, με τη χρήση ηλεκτρονικών υπολογιστών.

Σκοπός της Όρασης Μηχανής είναι η δημιουργία "έξυπνων συστημάτων" για την ανάκτηση πληροφορίας από εικόνες. Η ανάπτυξη της όρασης μηχανής σε συνδυασμό με τη δημιουργία καμερών υψηλής ευκρίνειας αλλά και των ισχυρών υπολογιστικών συστημάτων, που έχουμε πλέον στη διάθεση μας, βοήθησαν στην άνθιση νέων εφαρμογών που σκοπό έχουν την προσομοίωση της λειτουργίας του ανθρώπινου ματιού.

Στο βιολογικό κόσμο τα όντα που διαθέτουν την αίσθηση της όρασης μπορούν έμφυτα να αναλύσουν το ποσό της πληροφορίας που εισέρχεται μέσω των ματιών. Η είσοδος γίνεται με εικόνες και μέσω αυτών μπορούν να εξαχθούν συμπεράσματα για το χρώμα, το σχήμα, την κίνηση, την υφή αλλά και το πλάτος ενός αντικειμένου. Ο άνθρωπος μπορεί, για παράδειγμα, να αναγνωρίσει έναν άλλο άνθρωπο χωρίς να τον έχει δει ποτέ, μπορεί να αναγνωρίσει ένα αντικείμενο ως τραπέζι χωρίς να έχει συναντήσει ποτέ στο παρελθόν το ίδιο τραπέζι.

Παρατηρούμε πως η ανθρώπινη όραση είναι πολύ ανεπτυγμένη. Στην όραση μηχανής για να γίνει κάτι τέτοιο, σε αντίθεση με τον άνθρωπο, είναι απαραίτητο να προηγηθεί μια επεξεργασία πληροφοριών αλλά και

χαρακτηριστικών στη μορφή μιας ακολουθίας από εικόνες. Η όραση μηχανής παρέχει τις μεθόδους για την επεξεργασία αυτή, αλλά η διαδικασία είναι σαφώς αρκετά πιο περίπλοκη.

Παραθέτουμε συνοπτικά σε αυτό το σημείο, μερικούς από τους παράγοντες που καθιστούν τόσο δύσκολη τη διαδικασία της όρασης μηχανής να αντιγράψει την ανθρώπινη όραση. Για να παραχθεί μια απόφαση απαιτείται ερμηνεία, κάτι που στον άνθρωπο είναι έμφυτο, αλλά στον υπολογιστή όχι. Τα δεδομένα που εξάγονται από τις εικόνες για κατανόησή τους στον υπολογιστή, απαιτούν μεγάλο αποθηκευτικό χώρο, ειδικά αν έχουμε εικόνες με μεγάλη ανάλυση. Στη μηχανική όραση υπάρχει μεγάλη πιθανότητα απώλειας πληροφοριών κατά τη μετατροπή από δυο σε τρεις διαστάσεις και αυτό οφείλεται στο μετασχηματισμό της προοπτικής. Η φωτεινότητα δίνεται από ένα φυσικό σχηματισμό της εικόνας, που είναι αρκετά περίπλοκος, και καθιστά την αντίστροφη διαδικασία αρκετά σύνθετη. Η μεταβολή της οπτικής γωνίας, η διαφορετική κλίμακα των αντικειμένων, η παραμόρφωση, η απόφραξη και η ενδοκατηγοριακή διακύμανση αντικειμένων είναι προβλήματα αρκετά δυσνόητα, που συναντάμε στη μηχανική όραση. Στη δημιουργία εφαρμογών όρασης, ο προγραμματιστής πρέπει πάντα να λαμβάνει υπ' όψιν την πιθανότητα θορύβου. Μια εικόνα με διαφορετικές σκηνές και γεωμετρικές ιδιότητες αλλά σε διαφορετικές συνθήκες φωτισμού ή με όμοια χρώματα θα μπορούσε να οδηγήσει σε όμοιες εικόνες (βλ. Εικόνα 1). Παρατηρούμε λοιπόν, γιατί η διαδικασία μίμησης της ανθρώπινης όρασης είναι τόσο περίπλοκη. Οι εικόνες είναι πολύπλοκες και μπορούν να προκαλέσουν σύγχυση, αλλά μπορούν επίσης να αποκαλύψουν τη δομή του κόσμου γύρω μας μέσω πολυάριθμων ενδείξεων. Η ενασχόληση ενός ερευνητή με τον τομέα αυτό έχει να κάνει με την ερμηνεία αυτών των ενδείξεων.



Εικόνα 1 Χρωματικές ομοιότητες αντικειμένων και φόντου [23].

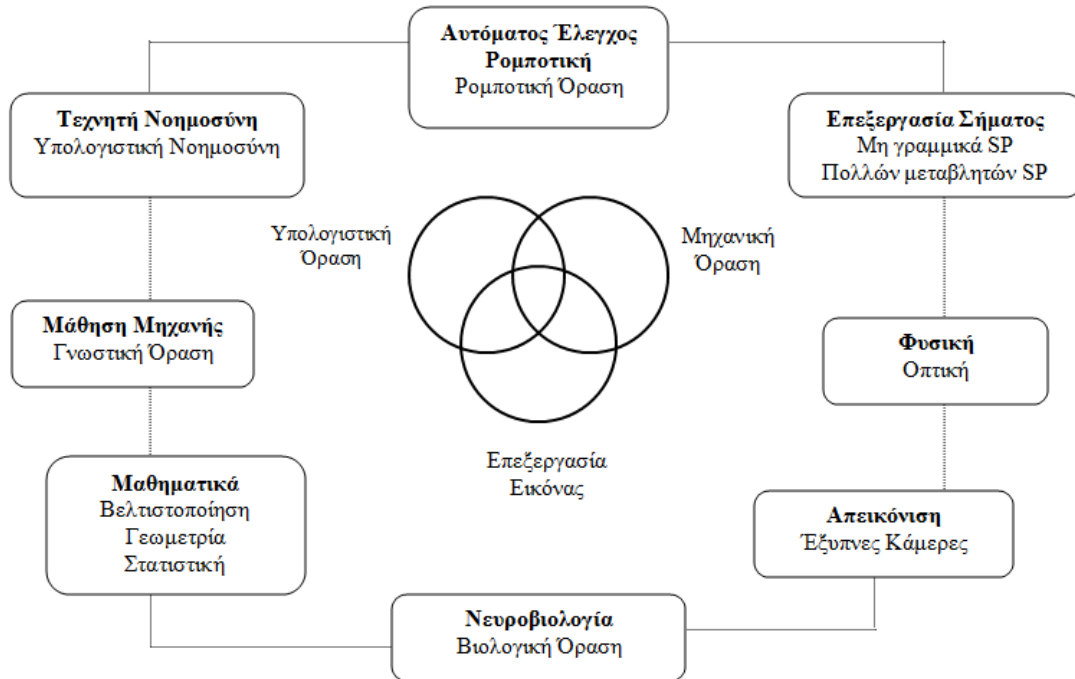
Τα αποτελέσματα που προκύπτουν από την ανάλυση της εικόνας, έχουν να κάνουν με τη διακύμανση στην εμφάνιση σε πολλαπλές προβολές, όπως η κίνηση. Οι ενδείξεις αυτές ποικίλουν ανάλογα με την χρησιμότητα της έρευνας και της εφαρμογής που θέλει να αναπτύξει ο προγραμματιστής.

Επικεντρωνόμαστε σε κάποιες από τις πολύ βασικές όπως είναι η σκίαση και ο χρωματικός τόνος, η ένταση του φωτός, τα περιγράμματα που περικλείουν τα αντικείμενα μέσα στην εικόνα, η υφή, οι γεωμετρικοί περιορισμοί αλλά και η προηγούμενη γνώση του αντικειμένου.

Ιεραρχικά η όραση μηχανής χωρίζεται σε τρία επίπεδα. Ξεκινώντας από το πρώτο, την όραση χαμηλού επιπέδου, όπου γίνεται η επεξεργασία εικόνας για την εξαγωγή χαρακτηριστικών, όπως είναι οι γωνίες, οι άκρες και η οπτική ροή. Το δεύτερο ιεραρχικά επίπεδο είναι η όραση μεσαίου επιπέδου, που εξετάζει την αναγνώριση αντικειμένων, την ανάλυση της κίνησης και την τρισδιάστατη ανακατασκευή με τη χρήση χαρακτηριστικών που προέρχονται από το πρώτο επίπεδο. Τέλος, έχουμε την όραση υψηλού επιπέδου, στην οποία ερμηνεύεται η εξελισσόμενη πληροφορία που προέρχεται από το μεσαίο επίπεδο. Το επίπεδο αυτό κατευθύνει τις διεργασίες που γίνονται στα δυο προηγούμενα επίπεδα. Η ερμηνεία που παρέχει μπορεί να περιλαμβάνει και την εννοιολογική περιγραφή ενός τοπίου, όπως τη δραστηριότητα, την πρόθεση αλλά και την συμπεριφορά.

Η τεχνητή όραση μπορεί να θεωρηθεί ο αντίστροφος κλάδος της πληροφορικής αν τον συγκρίνουμε με τον κλάδο των γραφικών υπολογιστών. Στα γραφικά υπολογιστών παράγονται εικόνες από δισδιάστατα και τρισδιάστατα μοντέλα, σε αντίθεση με τη όραση μηχανής που παράγει μοντέλα από τα δεδομένα μιας εικόνας.

Τα πρώτα βήματα για τη δημιουργία του κλάδου αυτού έγιναν στη δεκαετία του '70 από ερευνητές στον τομέα της Τεχνητής Νοημοσύνης, οι οποίοι προσπαθούσαν να δημιουργήσουν υπολογιστές με δυνατότητα όρασης. Η εξέλιξη της οποίας συνεχίζεται ακόμα και σήμερα. Για τη δημιουργία και ανάπτυξη της όρασης μηχανής τα επιστημονικά πεδία που συνέβαλαν είναι η Τεχνητή Νοημοσύνη, η Επεξεργασία Εικόνας και η Αναγνώριση Προτύπων.



Σχήμα 1 Σχέση ανάμεσα στην Υπολογιστική Όραση και άλλα επιστημονικά πεδία.

Η Τεχνητή Νοημοσύνη παρέιχε στην όραση μηχανής έτοιμους αλγόριθμους οι οποίοι υπέστησαν αλλαγές και χρησιμοποιήθηκαν στη δημιουργία εφαρμογών όρασης μηχανής. Η Επεξεργασία Εικόνας επικεντρώνεται στο χειρισμό εικόνας για τη βελτίωση της ποιότητάς της. Παρέχει διάφορους μετασχηματισμούς που αφορούν διδιάστατες εικόνες και χρησιμεύουν στην εύρεση θορύβου, την εύρεση ακμών, στη μετατόπιση και περιστροφή της εικόνας. Ασχολείται με την απόκτηση ακριβών και αξιόπιστων μετρήσεων όπως επίσης και με τη βαθμονόμηση της κάμερας και την τρισδιάστατη ανακατασκευή. Οι περισσότεροι αλγόριθμοι όρασης μηχανής συνήθως “υποθέτουν” πως ένα σημαντικό ποσοστό επεξεργασίας εικόνας έχει γίνει για τη βελτίωση της ποιότητας της εικόνας. Τέλος, η Αναγνώριση Προτύπων με διάφορες τεχνικές όπως στατιστικές προσεγγίσεις, νευρωνικά δίκτυα και υποστήριξη νευρωνικών μηχανών εξάγει πληροφορίες από κάθε εικόνα μέσω σημάτων και παρέχει τεχνικές εκμάθησης, για αναγνώριση και ταξινόμηση διαφορετικών προτύπων. Οι τεχνικές αναγνώρισης προτύπων χρησιμοποιούνται ευρέως στην όραση μηχανής.

Στο Σχήμα 1 παρουσιάζεται οπτικά η σχέση ανάμεσα στην όραση μηχανής και σε άλλα επιστημονικά πεδία.

Την τελευταία δεκαετία η όραση μηχανής επικεντρώνει το ενδιαφέρον της στην αλληλεπίδραση των γραφικών υπολογιστών με την όραση, έχοντας ως κεντρικό άξονα την ανάλυση και επεξεργασία εικόνας. Εμφανίζονται τεχνικές που έχουν ως βάση την ανίχνευση χαρακτηριστικών των αντικειμένων, στοχεύοντας στην αναγνώρισή τους. Αυτού του είδους τις τεχνικές, επικεντρωνόμαστε στην παρούσα διπλωματική εργασία με βάση το χρώμα του

αντικειμένου. Άλλες ομάδες ερευνητών στοχεύουν στην αναγνώριση με βάση την κατάτμηση περιοχών αλλά και την εύρεση καμπυλών.

Εφαρμογές όρασης μηχανής

Η οργάνωση και η υλοποίηση μιας εφαρμογής ενός συστήματος που βασίζεται στην όραση μηχανής έχει άμεση σχέση με την ίδια την εφαρμογή. Κάποιες από τις εφαρμογές χρησιμοποιούν αυτόνομα συστήματα για την επίλυση ενός προβλήματος ενώ άλλες αποτελούν υπό-σύστημα ενός μεγαλύτερου συστήματος.

Οι βασικές λειτουργίες των εφαρμογών όρασης μηχανής είναι κοινές σε όλους τους τομείς που χρησιμοποιούνται. Το πρώτο στάδιο στην υλοποίηση είναι η ανάκτηση εικόνας, η οποία αποτελεί και την κύρια πληροφορία του συστήματος. Η εικόνα αυτή μπορεί να είναι είτε στατική, είτε κινούμενη. Στη συνέχεια γίνεται μια επεξεργασία της εικόνας και λαμβάνονται κάποια δεδομένα τα οποία και αυτά με τη σειρά τους επεξεργάζονται για να υπακούν στις προϋποθέσεις που θέτει η εφαρμογή. Με βάση τα δεδομένα αυτά γίνεται η εξαγωγή των χαρακτηριστικών της εικόνας όπως είναι το περίγραμμα, που θα αναλύσουμε και στη συνέχεια της διπλωματικής εργασίας. Το επόμενο στάδιο είναι η κατάτμηση της εικόνας και στη συνέχεια έχουμε το τελικό στάδιο στο οποίο παράγονται οι αποφάσεις και η αντίδραση που έχει ορίσει ο προγραμματιστής. Ο σχεδιασμός του αλγορίθμου για τη δημιουργία της εφαρμογής μας, εκτελεί όλα τα παραπάνω βήματα.

Τα πεδία και οι διεργασίες που συναντάμε στις εφαρμογές όρασης μηχανής είναι πάρα πολλές. Παρακάτω αναφέρουμε κάποια ενδεικτικά παραδείγματα:

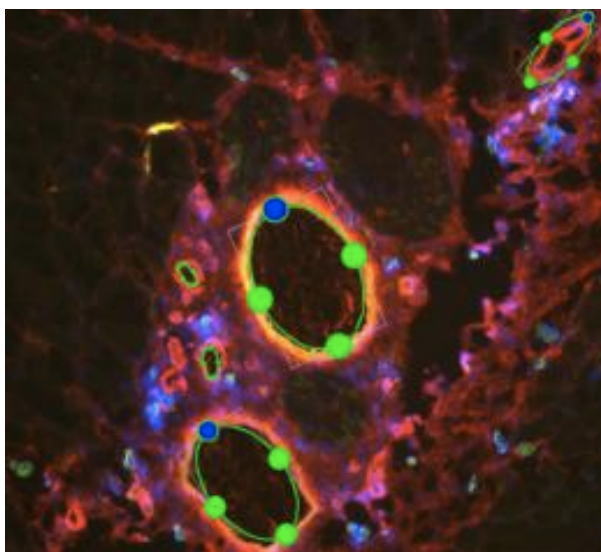
- Ο έλεγχος διεργασιών, ο οποίος συναντάται κυρίως σε ρομποτικές εφαρμογές και αυτόνομα οχήματα.
- Η ανίχνευση γεγονότων, που μπορούμε να την συναντήσουμε σε συστήματα παρακολούθησης και ασφάλειας για κατάτμηση αντικειμένων.
- Η οργάνωση πληροφοριών με τη μέθοδο της δεικτοδότησης βάσεων εικόνων.
- Η κατασκευή μοντέλων ή περιβάλλοντος, που χρησιμοποιείται στον βιομηχανικό έλεγχο, στις ιατρικές εικόνες αλλά και στην τοπογραφία.
- Η διαδραστικότητα μπορεί να χρησιμοποιηθεί για εφαρμογές αλληλεπιδραστικότητας με τον άνθρωπο.

Οι εφαρμογές για στρατιωτικούς σκοπούς είναι ένας από τους μεγαλύτερους τομείς της όρασης μηχανής. Οι εφαρμογές αυτές είναι πολυάριθμες και σκοπό έχουν την τεχνολογική ανωτερότητα των κατόχων τους. Κάποια από τα πιο γνωστά παραδείγματα είναι η καθοδήγηση πυραύλων και η ανίχνευση αντίπαλων στρατευμάτων. Για την καθοδήγηση πυραύλων τα

νέα προηγμένα συστήματα έχουν την επιλογή να στείλουν το βλήμα σε συγκεκριμένη περιοχή και όχι στόχο, για να αποφευχθούν τυχόν λάθη. Ο στόχος επιλέγεται τοπικά μέσω ζωντανής εικόνας.

Για την αύξηση της αξιοπιστίας και με στόχο τη μείωση της πολυπλοκότητας χρησιμοποιείται η αυτόματη επεξεργασία δεδομένων και η χρήση πολλαπλών αισθητήρων εικόνας, που μπορούν σε μια στρατιωτική εφαρμογή να παρέχουν πληροφορίες σχετικές με το τοπίο μιας μάχης έτσι ώστε να εξαχθούν συμπεράσματα και να υποστηριχθούν κάποιες στρατιωτικές αποφάσεις.

Η επεξεργασία ιατρικών δεδομένων από εικόνες είναι ένας άλλος τομέας στον οποίο είναι διαδεδομένες οι εφαρμογές όρασης μηχανής. Τα δεδομένα εικόνας παρέχουν πληροφορίες για τη διάγνωση ασθενειών. Η μορφή των εικόνων αλλάζει ανάλογα με την χρήση των εφαρμογών. Οι εικόνες αυτές μπορεί να προέρχονται από μικροσκόπιο, να πηγάζουν από ακτίνες X, από υπερήχους, τομογράφους αλλά και από άλλα ιατρικά μέσα εξαγωγής εικόνας. Κάποια παραδείγματα χρήσης είναι η εύρεση όγκων, η ανίχνευση και κατηγοριοποίηση κυττάρων (βλ. Εικόνα 2). Προηγμένες ρομποτικές εφαρμογές στον τομέα αυτό, εστιάζουν στη χρήση ρομπότ υποκινούμενων με όραση μηχανής για τη διεξαγωγή χειρουργικών επεμβάσεων. Εφαρμογές στον τομέα αυτό υπάρχουν για την έρευνα ιατρικών θεραπειών και μελέτης, για την πρόληψη αλλά και για τη θεραπεία των ασθενών.



Εικόνα 2 Ανίχνευση και καταμέτρηση αιμοφόρων αγγείων σε εικόνες [8].

Ο τομέας εφαρμογής της όρασης μηχανής τον οποίο υποστηρίζει το πρόγραμμα που συνοδεύει την εργασία αυτή, είναι ο τομέας της βιομηχανίας. Στον τομέα αυτό οι πληροφορίες που εξάγονται από τις εικόνες έχουν σκοπό την υποστήριξη της ομαλής λειτουργίας μιας γραμμής παραγωγής κατά τη διαδικασία παραγωγής των αντικειμένων. Τα προϊόντα ελέγχονται αυτόματα και λεπτομερώς για εύρεση τυχόν ελαττωμάτων κατά τον ποιοτικό έλεγχο, με αποτέλεσμα να απομακρύνονται τα ελαττωματικά και ανεπιθύμητα προϊόντα.

Στον τομέα της ασφάλειας οι εφαρμογές που δημιουργούνται εστιάζουν στα βιομετρικά χαρακτηριστικά των ανθρώπων, όπως είναι η ίριδα του ματιού, τα δακτυλικά αποτυπώματα και η αναγνώριση προσώπων. Με τη χρήση των καμερών ασφαλείας και της αναγνώρισης προσώπων μπορούν οι αρχές να παρακολουθούν ή να συλλαμβάνουν ευκολότερα υπόπτους. Μια άλλη εφαρμογή στον τομέα αυτό είναι η παρακολούθηση και η ανίχνευση ύποπτων δραστηριοτήτων ή συμπεριφοράς για άμεση αντίδραση από τις αρχές.

Προβλήματα με τα οποία ασχολείται η όραση μηχανής

Η όραση μηχανής ασχολείται με πολλά και διαφορετικά προβλήματα τα οποία ποικίλουν ανάλογα με τη χρησιμότητά της. Παρακάτω παραθέτουμε κάποια από τα βασικά και πιο συχνά προβλήματα που συναντά κάποιος στη δημιουργία εφαρμογών.

Ένα από τα πρώτα προβλήματα που καλείται να λύσει κάποιος στη μηχανική όραση είναι η αναγνώριση αντικειμένων (object identification), δηλαδή να γίνει ο απαραίτητος έλεγχος αν υπάρχει ή όχι κάποιο συγκεκριμένο αντικείμενο σε μια εικόνα. Το αντικείμενο αυτό μπορεί να είναι γνωστό, όπως σε μια γραμμή παραγωγής, ή άγνωστο και να χρειαστεί να γίνει η ταυτοποίησή του, όπως στην αναγνώριση προσώπου.

Άλλο πρόβλημα που συναντάμε συχνά στην όραση μηχανής, είναι η κίνηση αντικειμένων (tracking). Αν η εφαρμογή μας βασίζεται σε μια αλληλουχία εικόνων τότε πρέπει να γίνει η απαραίτητη επεξεργασία για να εξαχθεί κάποιο συμπέρασμα για το αντικείμενο αυτό. Στην περίπτωση του βιομηχανικού έλεγχου έχουμε πολλαπλά κινούμενα αντικείμενα για τα οποία γίνονται οι έλεγχοι. Άλλη περίπτωση είναι ο προσδιορισμός της ταχύτητας κίνησης των αντικειμένων.

Σημαντικό πρόβλημα σε πολλές εφαρμογές είναι η ανακατασκευή σκηνής με δεδομένο ότι υπάρχουν περισσότερες από μια εικόνες. Η ανακατασκευή σκηνής έχει σκοπό τον υπολογισμό ενός τρισδιάστατου μοντέλου.

Τέλος, έχουμε το πρόβλημα της ανάκτησης εικόνας που στόχο έχει την αποθρομβοποίηση της εφαρμογής. Ο θόρυβος μπορεί να προκληθεί από σφάλμα των αισθητήρων ή από την κίνηση των αντικειμένων μέσω πολλαπλών εικόνων. Το πρόβλημα αυτό μπορεί να επιλυθεί με την εφαρμογή διαφόρων φίλτρων ή τη χρήση προηγμένων αλγορίθμων για να ξεχωρίζουν τα χρήσιμα αντικείμενα από το θόρυβο.

1.4.2 Όραση μηχανής σε συστήματα παραγωγής

Τα συστήματα βιομηχανικής παραγωγής δεν είναι ιδανικά και χωρίς λάθη για αυτό το λόγο, χρησιμοποιείται ευρέως η όραση μηχανής. Σκοπός των εφαρμογών όρασης είναι να πληροφορούν το χειριστή τους όταν

παρουσιαστούν λάθος αντικείμενα στην γραμμή παραγωγής, έτσι ώστε να παρθούν έγκαιρα διορθωτικές αποφάσεις και να διασφαλιστεί η ποιότητα και η παραγωγικότητα των προϊόντων.

Τα βιομηχανικά συστήματα εξοπλίζονται με χρήσιμα προσόντα όρασης μηχανής τα οποία είναι προσαρμοστικά και μπορούν να αντιμετωπίσουν την αβεβαιότητα του περιβάλλοντος στο οποίο εφαρμόζονται. Με την ενσωμάτωση της όρασης στη διαδικασία παραγωγής μπορούμε να ανιχνεύσουμε μέσω οπτικού ελέγχου όταν υπάρχει κάποιο κατασκευαστικό πρόβλημα. Αυτό κάνει την όραση μηχανής ένα πανίσχυρο εργαλείο χρήσης. Παρακάτω, παραθέτουμε κάποια από τα κίνητρα που έχουν οι βιομηχανίες για τη χρήση τέτοιων εργαλείων.

Ασφάλεια και αξιοπιστία:

Η ασφάλεια και η αξιοπιστία είναι θέματα τα οποία μπορούν να προκαλέσουν προβληματισμό σε ένα περιβάλλον εργασίας επικίνδυνο για τον άνθρωπο, για αυτό το λόγο η όραση μηχανής είναι η προφανής λύση. Η χρήση των συγκεκριμένων εφαρμογών είναι πολύ σημαντική στην παραγωγή προϊόντων όπου ο έλεγχος δεν γίνεται για αισθητικούς σκοπούς αλλά, γιατί είναι καίριας σημασίας η πλήρης διασφάλιση της ποιότητας των αντικειμένων, τα οποία σε περίπτωση κακής κατασκευής μπορούν να στοιχίσουν ανθρώπινες ζωές.

Η όραση μηχανής διευκολύνει την συνοχή των προτύπων επιθεώρησης. Συστήματα όπως αυτά, έχουν σταθερή απόδοση και δεν επηρεάζονται από την ποσοτική επιβάρυνση που θα προκύψει, σε αντίθεση με τον άνθρωπο.

Ποιότητα Προϊόντων:

Βιομηχανικές γραμμές παραγωγής με μεγάλη ποσότητα προϊόντων όταν χρησιμοποιούν αποκλειστικά ανθρώπινο δυναμικό για τον έλεγχο ποιότητας, σπάνια επιτυγχάνουν μια γρήγορη και σωστή επιθεώρηση για όλα τα διαφορετικά εξαρτήματα που απαρτίζουν ένα προϊόν. Η χρήση όρασης μηχανής κάνει αυτή τη διεργασία εφικτή. Για να έχουμε τα επιθυμητά αποτελέσματα, δηλαδή τα προϊόντα που βγαίνουν στην αγορά να πληρούν τις προϋποθέσεις και να έχουν αποτελεσματική απόδοση, πρέπει η εφαρμογή να είναι ανάλογη της πολυπλοκότητας του προϊόντος. Ο προβληματισμός για τη δημιουργία της εφαρμογής είναι ιδιαίτερα σημαντικός ειδικά αν το τελικό προϊόν απαρτίζεται από πολλά μέρη. Το σύστημα όρασης πρέπει να έχει απευθείας σύνδεση με την παραγωγή έτσι ώστε να κάνει επιθεώρηση στο κάθε μέρος του τελικού προϊόντος καθώς παρασκευάζεται ξεχωριστά.

Ευέλικτη Αυτοματοποίηση:

Σε ένα βιομηχανικό περιβάλλον όπου η εξασφάλιση της ποιότητας εξαρτάται από μηχανές, είναι εφικτό να ενσωματωθεί η διεργασία ελέγχου στον πλήρη κύκλο παραγωγής και να έχει την επιλογή να παρέχει πληροφορίες που να διευκολύνουν τον έλεγχο παραγωγής. Αυτό για να γίνει προϋποθέτει μεγάλες

προσαρμοστικές απαιτήσεις από την τεχνολογία παραγωγής αλλά θα έχει σαν αποτέλεσμα τη διευκόλυνση του γενικού ελέγχου παραγωγής από υπολογιστές που τον συναντάμε κυρίως σε μεγάλα προηγμένα και ολοκληρωμένα περιβάλλοντα παραγωγής.

1.5 Οργάνωση και δομή της εργασίας

Η εργασία αυτή χωρίστηκε σε τμήματα με τέτοιο τρόπο ώστε να γίνει κατανοητή από τον αναγνώστη η διαδικασία υλοποίησης της εφαρμογής που αναπτύχθηκε, με θέμα την ταξινόμηση αντικειμένων σε γραμμή μεταφοράς με τη χρήση όρασης μηχανής.

Στην εισαγωγή, εισάγαμε την έννοια της όρασης μηχανής, παρουσιάσαμε τα διάφορα πεδία εφαρμογών της και αναλύσαμε τα προβλήματα με τα οποία ασχολείται η όραση μηχανής. Κλείνοντας λοιπόν την εισαγωγή της εργασίας ακολουθεί μια σύντομη περίληψη των κεφαλαίων που θα ακολουθήσουν και των θεμάτων τους.

Το δεύτερο κεφάλαιο είναι αφιερωμένο στην κατάτμηση εικόνας και την παρακολούθηση κινούμενων αντικειμένων σε μια ακολουθία από εικόνες. Στο κεφάλαιο αυτό εισάγεται η έννοια της κατάτμησης εικόνας όπως την αντιλαμβάνεται ο υπολογιστής και αναλύονται κάποιες από τις πιο δημοφιλείς τεχνικές κατάτμησης, οι οποίες μας βοήθησαν στη δημιουργία του δικού μας αλγορίθμου. Στη συνέχεια αυτού του κεφαλαίου παρουσιάζονται διάφορες τεχνικές ανίχνευσης και παρακολούθησης κινούμενων αντικειμένων.

Η ανίχνευση αντικειμένου αποφασίστηκε να υλοποιηθεί με τη χρήση αλγορίθμου με βάση το χρώμα, για το λόγο αυτό το τρίτο κεφάλαιο είναι αφιερωμένο στο χρώμα και την κατάτμηση με τη χρήση του. Αρχικά παρουσιάζονται οι δημοφιλέστεροι χρωματικοί χώροι, κάποιους από τους οποίους χρησιμοποιήσαμε στην εφαρμογή μας. Η επόμενη θεματική ενότητα του κεφαλαίου αυτού αναλύει τεχνικές κατάτμησης με βάση τους χρωματικούς χώρους.

Στο τέταρτο κεφάλαιο παρουσιάζουμε συνοπτικά κάποια από τα τεχνικά χαρακτηριστικά του υλικού (Hardware) και του λογισμικού (Software) με τη χρήση των οποίων δημιουργήθηκε η εφαρμογή μας. Αναλύονται επίσης, τα εργαλεία που χρησιμοποιήσαμε για την συγγραφή του κώδικα που συνοδεύει την εργασία.

Το πέμπτο κεφάλαιο είναι αφιερωμένο στο κομμάτι του σχεδιασμού και της μεθοδολογίας υλοποίησης της εφαρμογής. Παρουσιάζεται ο τρόπος ανάπτυξης του κώδικα, οι έλεγχοι που απαιτούνται και τα προβλήματα που αντιμετωπίσαμε κατά τη διάρκεια συγγραφής.

Στο έκτο κεφάλαιο παρουσιάζουμε τα αποτελέσματα της εφαρμογής και τα διάφορα σενάρια δοκιμών που εφαρμόσαμε, για να αποδείξουμε ότι το λογισμικό μας, ικανοποιεί τις απαιτήσεις που ορίσαμε.

Στο έβδομο και τελευταίο κεφάλαιο αναλύουμε τα συμπεράσματα στα οποία καταλήξαμε με το πέρας των εργασιών για τη δημιουργία της εφαρμογής και προτείνουμε κάποιες μελλοντικές επεκτάσεις για τη βελτιστοποίηση της.

2

Κατάτμηση Εικόνας και Παρακολούθηση Αντικειμένων

Στο κεφάλαιο αυτό παρουσιάζονται και αναλύονται διάφοροι αλγόριθμοι που εξετάστηκαν για το πρόβλημα της ανίχνευσης (detect) ενός αντικειμένου ή μιας περιοχής ενδιαφέροντος από μία εικόνα και της παρακολούθησης αντικειμένων (tracking) τα οποία εξάγονται από ακολουθίες εικόνων (video). Η διαδικασία της κατάτμησης εικόνας είναι μια διαδικασία που εξαρτάται άμεσα από την εφαρμογή στην οποία θα εφαρμοστεί αλλά και από το είδος των χαρακτηριστικών που θα είναι επιθυμητό να εξαχθούν από την εικόνα εισόδου.

2.1 Κατάτμηση Εικόνας

Στην Όραση Μηχανής ο όρος κατάτμηση εικόνας αναφέρεται στη διαδικασία του διαχωρισμού μιας ψηφιακής εικόνας σε πολλαπλά τμήματα. Χρησιμοποιείται συνήθως για τον εντοπισμό αντικειμένων και των περιγραμμάτων τους. Συγκεκριμένα η κατάτμηση εικόνας είναι η διαδικασία της ανάθεσης μιας ετικέτας σε κάθε εικονοστοιχείο (pixel) της εικόνας αυτής, έτσι ώστε τα εικονοστοιχεία με την ίδια ετικέτα να μοιράζονται κάποια οπτικά χαρακτηριστικά. Το αποτέλεσμα που προκύπτει από την κατάτμηση εικόνας είναι ένα σύνολο τομέων που καλύπτουν συλλογικά ολόκληρη την εικόνα ή ένα σύνολο περιγραμμάτων (contours) τα οποία εξάγονται από την εικόνα. Σε μια περιοχή τα εικονοστοιχεία αναλογούν σε κάποιο χαρακτηριστικό ή ιδιότητα, όπως είναι το χρώμα, η ένταση και η υφή.

Λόγω της μεγάλης σημασίας της κατάτμησης εικόνας έχουν προταθεί και δημιουργηθεί διάφοροι αλγόριθμοι. Ο αλγόριθμος που θα επιλεγεί να χρησιμοποιηθεί, πρέπει να είναι ανάλογος της εφαρμογής και της εικόνας που θα εισαχθεί για να έχουμε τα επιθυμητά αποτελέσματα.

Στο κεφάλαιο αυτό παρουσιάζουμε συνοπτικά κάποιους από τους πιο σημαντικούς αλγόριθμους που εξετάσαμε για την κατάτμηση εικόνας στην εφαρμογή μας.

2.1.1 Μετασχηματισμός Κλιμακωτά - Αναλλοίωτων Χαρακτηριστικών (SIFT)

Ο πρώτος αλγόριθμος που εξετάζουμε είναι ο μετασχηματισμός κλιμακωτά αναλλοίωτων χαρακτηριστικών, ο οποίος ανήκει στο επιστημονικό πεδίο της όρασης μηχανής. Σκοπός του είναι ο εντοπισμός και η περιγραφή τοπικών χαρακτηριστικών. Δημοσιεύθηκε για πρώτη φορά από τον David Lowe το 1999 [18] ο οποίος κατέχει και κατοχυρωμένο δίπλωμα ευρεσιτεχνίας στις Ηνωμένες Πολιτείες Αμερικής κάτω από την ιδιοκτησία του Πανεπιστημίου της Βρετανικής Κολούμπια και παρέχεται για ερευνητική χρήση. Ο αλγόριθμος αυτός έχει πολλές εφαρμογές, κάποιες από αυτές είναι η αναγνώριση αντικειμένων, η παραγωγή χαρτών και η ανίχνευση αντικειμένων σε βίντεο.

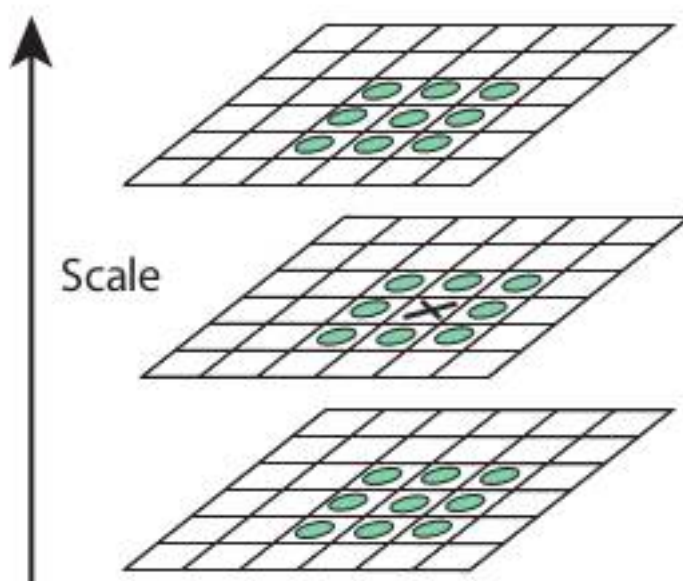
Έστω μια εικόνα η οποία περιέχει αντικείμενα που μας ενδιαφέρουν, με τον αλγόριθμο αυτό για το κάθε ένα από αυτά τα αντικείμενα, εξάγονται σημεία ενδιαφέροντος προκειμένου να δώσουν μια περιγραφή των χαρακτηριστικών του (feature description). Με τον όρο σημείο ενδιαφέροντος εννοούμε το εικονοστοιχείο το οποίο έχει μια καλά καθορισμένη θέση, μπορεί να ανιχνευθεί εύκολα και περιέχει τοπικές πληροφορίες. Η περιοχή χαρακτηριστικού προέρχεται από μία εικόνα που περιέχει μόνο το αντικείμενο που ψάχνουμε, τη λεγόμενη εικόνα εκπαίδευσης. Η εικόνα αυτή χρησιμεύει στον εντοπισμό του αντικειμένου που περιέχει, μέσα σε μια άλλη εικόνα με πολλά και διαφορετικά αντικείμενα. Από τις εικόνες εκπαίδευσης εξάγονται τα κύρια σημεία των αντικειμένων, οι λεγόμενες ακμές, τα οποία αποθηκεύονται σε μια βάση δεδομένων. Για την ανίχνευση του αντικειμένου αυτού γίνεται σύγκριση της νέας εικόνας με το περιεχόμενο της βάσης δεδομένων εξάγοντας έτσι πιθανά χαρακτηριστικά που αντιστοιχούν στο αντικείμενο. Στη συνέχεια γίνονται περαιτέρω έλεγχοι φιλτραρίσματος απορρίπτοντας τις ακραίες τιμές και κρατάμε τα ταιριάσματα, τα οποία αναγνωρίστηκαν ως αξιόπιστα.

Ο αλγόριθμος του μετασχηματισμού κλιμακωτών αναλλοίωτων χαρακτηριστικών (Scale Invariant Feature Transform - SIFT) ακολουθεί τέσσερα βασικά στάδια τα οποία θα αναλύσουμε παρακάτω ένα προς ένα.

Κλιμακωτή ανίχνευση ακρότατων στοιχείων στο χώρο

Το πρώτο στάδιο επεξεργασίας είναι η κλιμακωτή ανίχνευση ακρότατων στοιχείων στο χώρο όπου σε αυτό ερευνώνται όλες οι κλίμακες και οι τοποθεσίες σε μια εικόνα.

Το πρώτο στάδιο της επεξεργασίας είναι χρήσιμο γιατί δεν μπορούμε να χρησιμοποιήσουμε το ίδιο παράθυρο για να ανιχνεύσουμε σημεία - κλειδιά σε διαφορετική κλίμακα. Το στάδιο αυτό έχει υλοποιηθεί με τη χρήση Λαπλασιανών των Γκαουσιανών (Laplacian of Gaussian) φίλτρων για την ανίχνευση πιθανών σημείων ενδιαφέροντος με διάφορες «σίγμα» τιμές. Οι τιμές αυτές λειτουργούν σαν παράγοντας κλίμακας με τη διαφορά των Γκαουσιανών σε μια σειρά από εξομαλύνσεις και επαναδειγματοληφθέντες εικόνες. Στη Γκαουσιανή διαφορά το κάθε εικονοστοιχείο συγκρίνεται με τα γειτονικά του. Το εικονοστοιχείο που επιλέγεται πρέπει να έχει διαφορά φωτεινότητας σε σχέση με τα γειτονικά του εικονοστοιχεία.



Εικόνα 3 Υπολογισμός του μεγίστου και ελαχίστου της Γκαουσιανής διαφοράς σε ένα εικονοστοιχείο.

Στην εικόνα 3 παρουσιάζεται η διαδικασία υπολογισμού του μεγίστου και ελαχίστου της Γκαουσιανής διαφοράς σε ένα εικονοστοιχείο.

Εντοπισμός σημείων κλειδιών:

Το δεύτερο στάδιο είναι ο εντοπισμός σημείων κλειδιών. Όταν εντοπιστούν αρχικά οι τοποθεσίες των σημείων αυτών, θα πρέπει στην συνέχεια να τις προσδιορίσουμε για να μπορούμε να πάρουμε ακριβή αποτελέσματα. Με τη χρήση των σειρών Taylor για την επέκταση της κλίμακας του χώρου, μπορούμε να πάρουμε την ακριβή θέση των ακρότατων και αν η ένταση τους είναι μικρότερη από την τιμή κατωφλίου τότε τα σημεία αυτά απορρίπτονται. Με απλά λόγια, στο στάδιο αυτό τα σημεία κλειδιά επιλέγονται με κριτήριο τη μέτρηση της σταθερότητάς τους.

Ανάθεση προσανατολισμού:

Φτάνοντας στο σημείο αυτό, στο κάθε σημείο κλειδί έχει ανατεθεί ο δικός του προσανατολισμός, ο οποίος βασίζεται στις τοπικές κατευθύνσεις και τις κλίσεις της εικόνας. Το κάθε σημείο μένει αναλλοίωτο στην περιστροφή της εικόνας. Γύρω από το σημείο κλειδί λαμβάνεται η γειτονική περιοχή του, που είναι ανάλογη της κλίμακας, και υπολογίζεται στην περιοχή αυτή το βαθμωτό μέγεθος και η κατεύθυνσή του.

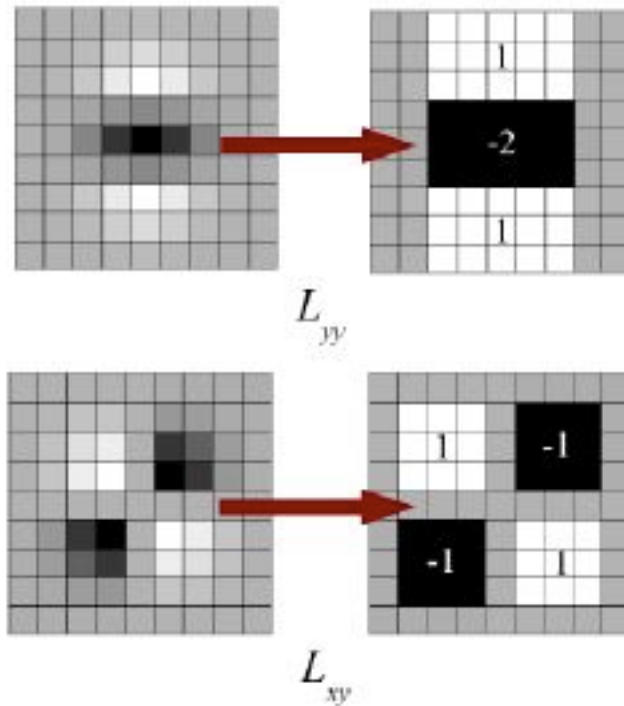
Περιγραφέας σημείων κλειδιών:

Στο στάδιο αυτό έχει δημιουργηθεί μια περιγραφή που αντικατοπτρίζει τα σημεία κλειδιά. Οι βαθμίδες των τοπικών εικόνων γύρω από το κάθε σημείο σε συγκεκριμένη κλίμακα μετατρέπονται σε αναπαραστάσεις που επιτρέπουν αλλαγές στην φωτεινότητα και στη διαστρέβλωση της τοπικής μορφής.

2.1.2 Επιταχυνθέντα ισχυρά χαρακτηριστικά (SURF)

Στην προηγούμενη ενότητα μελετήσαμε συνοπτικά τον αλγόριθμο μετασχηματισμού κλιμακωτών αναλλοίωτων χαρακτηριστικών για την ανίχνευση και περιγραφή σημείων κλειδιών. Η μέθοδος αυτή αποδείχθηκε σχετικά αργή και οι προγραμματιστές χρειάστηκαν μια έκδοση του αλγορίθμου, η οποία θα έκανε τη διαδικασία ταχύτερα. Το 2006 δημοσιεύθηκε από τους Bay, H., Tuytelaars, T. και Van Gool, L [19] ένας άλλος αλγόριθμος, ο αλγόριθμος επιταχυνθέντων ισχυρών χαρακτηριστικών (SURF: Speeded Up Robust Features). Ο αλγόριθμος αυτός είναι μια γρηγορότερη εκδοχή του αλγορίθμου SIFT και θεωρείται ότι είναι πιο έμπιστος στις διάφορες μεταμορφώσεις της εικόνας.

Στον αλγόριθμο αυτό προσεγγίζεται η Λαπλασιανή της Γκαουσιανής με ένα φίλτρο κουτιών όπως βλέπουμε στην εικόνα 4. Η συνέλιξη με φίλτρο κουτιών μπορεί εύκολα να υπολογιστεί με τη βοήθεια των ενσωματωμένων εικόνων και αυτό μπορεί να γίνει παράλληλα για διάφορες κλίμακες. Αυτό είναι ένα από τα μεγάλα πλεονεκτήματα του αλγορίθμου επιταχυνθέντων ισχυρών χαρακτηριστικών. Για τον καθορισμό της κλίμακας και της τοποθεσίας ενός αντικειμένου ο αλγόριθμος αυτός βασίζεται στην διακρίνουσα του πίνακα Έσσιαν (Hessian matrix).



Εικόνα 4 Προσέγγιση Λαπλασιανής της Γκαουσιανής με φίλτρο κουτιών.

Για την αναζήτηση αντιστοιχιών ανάμεσα στο αντικείμενο που ψάχνουμε και στην εικόνα που έχουμε πρέπει να ακολουθήσουμε τρία βασικά βήματα. Στο πρώτο βήμα τα σημεία ενδιαφέροντος εξετάζονται στις διακριτές τοποθεσίες της εικόνας, όπως είναι για παράδειγμα οι γωνίες. Με επανάληψη, τα σημεία ανιχνεύονται αξιόπιστα ακόμα και αν αλλάξουν οι συνθήκες στις οποίες βρίσκεται το ίδιο σημείο που ανιχνεύουμε. Στο επόμενο βήμα τα γειτονικά σημεία του σημείου ενδιαφέροντος περιγράφονται από ένα διάνυσμα χαρακτηριστικών, το οποίο δεν αλλοιώνεται σε περιπτώσεις παραμόρφωσης και θορύβου. Τέλος, τα διανύσματα χαρακτηριστικών εξετάζονται για ταύτιση μεταξύ των διαφόρων εικόνων και αντιστοιχίζονται.

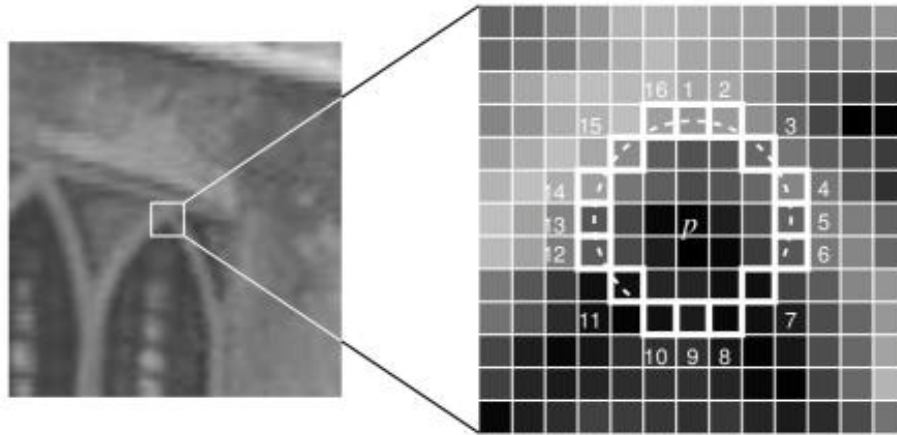
Σε όλα τα βήματα που ακολουθούνται, ο αλγόριθμος επιταχυνθέντων ισχυρών χαρακτηριστικών προσθέτει πολλά χαρακτηριστικά που σκοπό έχουν τη βελτίωση της ταχύτητας. Η απόδοση του στην εξαγωγή χαρακτηριστικών είναι όμοια με αυτή του αλγορίθμου SIFT αλλά σε θέμα ταχύτητας είναι τρεις φορές γρηγορότερος ο SURF. Ο αλγόριθμος επιταχυνθέντων ισχυρών χαρακτηριστικών είναι πολύ καλός στο χειρισμό εικόνων με περιστροφή και θόλωση αλλά δεν είναι τόσο καλός στο χειρισμό της αλλαγής οπτικής γωνίας και φωτισμού.

2.1.3 Αλγόριθμος Εξαγωγής Χαρακτηριστικών από γρήγορο τμήμα δοκιμής για Ανίχνευση Ακμών (FAST)

Οι προαναφερθέντες αλγόριθμοι είναι πολύ καλοί στην εύρεση χαρακτηριστικών σε εικόνες αλλά αν εξετάσουμε τη λειτουργικότητά τους σε μια εφαρμογή πραγματικού χρόνου η ταχύτητα απόδοσης τους είναι σχετικά μικρή. Για τον λόγο αυτό προτάθηκε το 2006 ένας άλλος αλγόριθμος [1] ο αλγόριθμος εξαγωγής χαρακτηριστικών από γρήγορο τμήμα (Features from Accelerated Segment Test – FAST) από τους Edward Rosten και Tom Drummond, ο οποίος αναθεωρήθηκε από τους ίδιους το 2010. Ο αλγόριθμος αυτός έχει σκοπό τον προσδιορισμό σημείων ενδιαφέροντος σε μια εικόνα σε πραγματικό χρόνο.

Για την ανίχνευση χαρακτηριστικών με τη χρήση του αλγορίθμου FAST ακολουθούμε τα εξής βήματα:

- Επιλέγεται ένα σημείο p στην εικόνα το οποίο καθορίζεται αν είναι ή όχι το σημείο ενδιαφέροντος. Την ένταση φωτεινότητας την θέτουμε ως I_p .
- Θέτουμε την ένταση κατωφλίου ως t .
- Θεωρούμε ένα κύκλο αποτελούμενο από 16 εικονοστοιχεία γύρω από το σημείο p (βλ. Εικόνα 5).
- Στο σημείο αυτό το εικονοστοιχείο p αντιστοιχεί σε μια γωνία, με την προϋπόθεση ότι υπάρχει ένα τμήμα με n συνεχόμενα εικονοστοιχεία στον κύκλο, τα οποία είναι πιο φωτεινά από $I_p + t$, ή σκοτεινότερα από $I_p - t$.
- Για να γίνει γρήγορος ο αλγόριθμος πρέπει να συγκρίνουμε πρώτα την ένταση φωτεινότητας των εικονοστοιχείων 1,5,9 και 13 του κύκλου με το I_p . Το σημείο p για να είναι γωνία πρέπει το λιγότερο τρία από τα παραπάνω εικονοστοιχεία να είναι φωτεινότερα του $I_p + t$ ή να είναι σκοτεινότερα από $I_p - t$. Αν το σημείο αυτό αποδειχθεί ότι είναι σημείο ενδιαφέροντος – γωνία τότε εξετάζονται και τα 16 εικονοστοιχεία του κύκλου.
- Η διαδικασία επαναλαμβάνεται για όλα τα εικονοστοιχεία της εικόνας.



Εικόνα 5 Το σημείο ενδιαφέροντος μιας εικόνας κάτω από δοκιμή και τα 16 εικονοστοιχεία στον κύκλο [1].

Στον αλγόριθμο FAST υπάρχουν μερικοί περιορισμοί όπως αν έχουμε μικρότερο από 12, τότε ο αλγόριθμος δεν λειτουργεί όπως θα έπρεπε σε όλες τις περιπτώσεις επειδή ο αριθμός των σημείων ενδιαφέροντος που εντοπίζεται είναι πολύ μεγάλος. Άλλος περιορισμός είναι η σειρά που πρέπει να μπουν τα 16 εικονοστοιχεία του κύκλου, η οποία καθορίζει και την ταχύτητα του αλγορίθμου. Για την αντιμετώπιση αυτών των περιορισμών του αλγορίθμου προστέθηκε μια προσέγγιση μηχανικής μάθησης.

2.1.4 Αλγόριθμος Κατάτμησης εικόνας με βάση το χρώμα

Ο τελευταίος αλγόριθμος που παρουσιάζουμε για κατάτμηση εικόνας είναι ο αλγόριθμος με βάση το χρώμα. Σε αυτόν τον αλγόριθμο η διαδικασία κατάτμησης αναλύει τα χρώματα των εικονοστοιχείων με σκοπό να διαχωρίσει διαφορετικά αντικείμενα που υπάρχουν σε μια εικόνα. Είναι η διαδικασία κατακερματισμού της εικόνας σε υποσύνολα των συνδεδεμένων εικονοστοιχείων, που έχουν τις ίδιες ιδιότητες χρώματος.

Η προσεγγιστική μέθοδος που ακολουθείται στον αλγόριθμο αυτό, είναι να συλλεχθούν δείγματα από τα εικονοστοιχεία και να ελεγχθούν αν ανήκουν στη χρωματική περιοχή ενδιαφέροντος, έτσι ώστε να γίνει ομαδοποίηση των εικονοστοιχείων αυτών με αποτέλεσμα να εξαχθεί το περίγραμμα του αντικειμένου που ψάχνουμε.

Η μεγάλη βελτίωση των δυνατοτήτων επεξεργασίας του υπολογιστή έκανε την κατάτμηση εικόνας με βάση τη πληροφορία του χρώματος να αντιμετωπίζεται με όλο και περισσότερο ενδιαφέρον από την ερευνητική κοινότητα και να χρησιμοποιείται ως βάση πολλών εφαρμογών.

Τα προβλήματα που παρουσιάζει η μέθοδος αυτή έχουν να κάνουν με την ευαισθησία του αλγορίθμου σε αλλαγές συνθηκών φωτισμού και η

πολυπλοκότητα της σε κατάτμηση εικόνας με διάφορα αντικείμενα όμοιου ή και διαφορετικού χρωματισμού.

Η κατάτμηση εικόνας με βάση το χρώμα αναλύεται εις βάθος στο κεφάλαιο 3.

2.2 Ανίχνευση και Παρακολούθηση αντικειμένων

Η διαδικασία της ανίχνευσης ενός αντικειμένου σε μια ή περισσότερες εικόνες είναι ένα από τα πρώτα βήματα που καλείται να αντιμετωπίσει κάποιος για να επιτύχει την ανάπτυξη εφαρμογής στον τομέα της όρασης μηχανής. Η ανίχνευση και παρακολούθηση αντικειμένων, τα οποία κινούνται, αποτελεί υπό-ενότητα του γενικού προβλήματος της κατάτμησης εικόνας και απαρτίζεται από δύο υπό-προβλήματα. Το πρώτο υπό-πρόβλημα είναι ο εντοπισμός του αντικειμένου στην εικόνα και το δεύτερο είναι η παρακολούθηση του, κατά τη διάρκεια πολλαπλών εικόνων.

2.2.1 Ανίχνευση Κινούμενου Αντικειμένου

Κάθε εφαρμογή που επωφελείται από την έξυπνη επεξεργασία βίντεο έχει διαφορετικές ανάγκες, ως εκ τούτου απαιτεί διαφορετική μεταχείριση. Ωστόσο, όλες οι εφαρμογές έχουν κάτι κοινό, τα αντικείμενα που έχουν κίνηση. Έτσι, η ανίχνευση των περιφερειών που αντιστοιχούν σε κινούμενα αντικείμενα στο βίντεο, όπως τα οχήματα, και στη δική μας περίπτωση τα αντικείμενα σε γραμμή παραγωγής είναι το πρώτο βασικό βήμα σχεδόν σε κάθε σύστημα όρασης.

Αλλαγές σε μια φυσική σκηνή, όπως είναι η διακύμανση του φωτισμού και η μεταβολή του καιρού, καθιστούν την ανίχνευση κίνησης μια δύσκολη διαδικασία για αξιόπιστη επεξεργασία. Οι πιο συχνά χρησιμοποιούμενες τεχνικές για την ανίχνευση κινούμενων αντικειμένων είναι η αφαίρεση φόντου, η χρήση στατικών μεθόδων, η χρονική διαφορά και η οπτική ροή. Παρακάτω δίνουμε αναλυτικά τις περιγραφές των τεχνικών αυτών.

2.2.2 Αφαίρεση Φόντου

Η αφαίρεση του φόντου (Background subtraction) είναι μια ευρέως χρησιμοποιούμενη τεχνική για την κατάτμηση εικόνας σε στατικές σκηνές. Επιχειρεί να ανιχνεύσει τις κινούμενες περιοχές με το να θεωρεί την στατική εικόνα, που δεν περιέχει το κινούμενο αντικείμενο, ως το μοντέλο του φόντου. Η κινούμενη εικόνα βρίσκεται από τη διαφορά των καρτέ και του φόντου εξετάζοντας κάθε εικονοστοιχείο με το αντίστοιχο εικονοστοιχείο του προηγούμενου καρτέ. Τα εικονοστοιχεία όπου η διαφορά τους είναι πάνω από το κατώφλι κατηγοριοποιούνται ως παρασκήνιο.

Οι τεχνικές αφαίρεσης φόντου αποδίδουν γενικά καλά στην ανίχνευση κίνησης ακόμα και αν το κινούμενο αντικείμενο παραμένει σταθερό για κάποιο χρονικό διάστημα. Οι τεχνικές αυτές παρουσιάζουν προβλήματα στην περίπτωση που παρατηρηθούν δυναμικές αλλαγές στον φωτισμό ή αν κάποιο μέρος που θεωρήσαμε σαν παρασκήνιο αρχίσει να παρουσιάζει κίνηση, όπως για παράδειγμα η κίνηση από ένα αρχικά σταθμευμένο αυτοκίνητο.

Οι τεχνικές αυτές στην δημιουργία της εφαρμογής μας δεν χρησιμοποιήθηκαν, γιατί το παρασκήνιο της γραμμής παραγωγής δεν ήταν σταθερό με αποτέλεσμα να μην υπάρχουν κάποια σταθερά αρχικά εικονοστοιχεία για τον εντοπισμό κίνησης.

2.2.3 Στατιστικές μέθοδοι

Οι στατιστικές μέθοδοι (Statistical Methods) έχουν αναπτυχθεί για να ξεπεραστούν κάποιες ανεπάρκειες των μεθόδων αφαίρεσης φόντου. Θεωρούνται πιο προηγμένες μέθοδοι από τις μεθόδους αφαίρεσης φόντου, αν και από αυτές πήραν την έμπνευση για τη δημιουργία τους, από την άποψη ότι κρατούν και ενημερώνουν δυναμικά τα στατιστικά των εικονοστοιχείων, τα οποία ανήκουν στο παρασκήνιο της εικόνας. Τα εικονοστοιχεία του προσκηνίου προσδιορίζονται με τη σύγκριση των στατιστικών του κάθε εικονοστοιχείου με αυτά του μοντέλου του παρασκηνίου. Για τη μοντελοποίηση κάθε εικονοστοιχείου συχνά χρησιμοποιούνται γκαουσιανές εξισώσεις.

Άλλη τεχνική των στατιστικών μεθόδων είναι η χρήση της μέγιστης και ελάχιστης τιμής της έντασης και η απόκλιση των τιμών αυτών που προκύπτει από τις διάφορες σκηνές ως στατιστικοί παράμετροι για την μοντελοποίηση του φόντου.

Η προσέγγιση με στατιστικές μεθόδους γίνεται όλο και πιο δημοφιλής λόγω της αξιοπιστίας της σε σκηνές που περιέχουν θόρυβο, σκιές και αλλαγές φωτισμού. Στην εφαρμογή μας η μέθοδος αυτή απορρίφθηκε λόγω μη στατικότητας του φόντου.

2.2.4 Χρονική Διαφορά

Στη μέθοδο χρονικής διαφοράς (Temporal differencing) στόχος είναι η ανίχνευση των κινούμενων περιοχών χρησιμοποιώντας τη διαφορά εικονοστοιχείου προς εικονοστοιχείου διαδοχικών πλαισίων – καρέ σε μία ακολουθία εικόνων. Η μέθοδος αυτή είναι εξαιρετικά προσαρμοστική σε δυναμικές αλλαγές σκηνής, ωστόσο αποτυγχάνει να ανιχνεύσει όλα τα εικονοστοιχεία που απαρτίζουν το κινούμενο αντικείμενο. Η μέθοδος αυτή αποτυγχάνει επίσης να ανιχνεύσει αν είναι στάσιμα τα αντικείμενα. Για την επιτυχία της μεθόδου αυτής σε επεξεργασία υψηλού επιπέδου, στην ανίχνευση αντικειμένων πρέπει να υιοθετηθούν παράλληλα με αυτή και άλλες μέθοδοι προκειμένου να ανιχνευθούν χρήσιμα αντικείμενα τα οποία δεν κινούνται για κάποια χρονική στιγμή.

Η μέθοδος της χρονικής διαφοράς δεν χρησιμοποιήθηκε στην εφαρμογή μας γιατί σκοπός μας είναι η ταξινόμηση των προϊόντων, κάτι που δεν είναι εφικτό αν τα εικονοστοιχεία που ανιχνεύονται είναι ελλιπή.

2.2.5 Οπτική Ροή

Οι μέθοδοι Οπτικής ροής (Optical flow) κάνουν χρήση των διανυσμάτων ροής των κινούμενων αντικειμένων κατά τη χρονική μεταβολή για την ανίχνευση κινούμενων εικόνων. Βασίζονται στην υπόθεση ότι η ένταση των εικονοστοιχείων δεν αλλάζει. Μπορούν να ανιχνεύσουν κινούμενες περιοχές ακόμα και αν το βίντεο προέρχεται από κινούμενη κάμερα. Με την οπτική ροή, όμως, είναι αδύνατον να προσδιορίσουμε την ταχύτητα της εικόνας στην κάθετη διεύθυνση μέσω της βάρθρωσης στην ένταση της εικόνας.

Οι περισσότερες από τις μεθόδους οπτικής ροής είναι υπολογιστικά σύνθετες και δεν μπορούν να χρησιμοποιηθούν σε πραγματικό χρόνο χωρίς εξειδικευμένο υλικό.

3

Χρώμα και κατάτμηση

Η κατάτμηση εικόνας με βάση το χρώμα διαφέρει από τις άλλες τεχνικές κατάτμησης επειδή στοχεύει στην αναγνώριση αντικειμένων λαμβάνοντας υπόψη τις χρωματικές πληροφορίες, σε αντίθεση με τις άλλες τεχνικές που βασίζονται κυρίως σε μονόχρωμη επεξεργασία. Στο παρόν κεφάλαιο θα εισάγουμε την έννοια του χρώματος, όπως την αντιλαμβάνεται η όραση μηχανής, και θα αναλύσουμε την κατάτμηση εικόνας με βάση το χρώμα. Η κατάτμηση εικόνας με βάση τα χρωματικά μοντέλα είναι η μέθοδος που χρησιμοποιήθηκε για την υλοποίηση της εφαρμογής ανίχνευσης και ταξινόμησης αντικειμένων σε γραμμή μεταφοράς που σχεδιάσαμε στην παρούσα διπλωματική εργασία.

3.1 Χρωματικοί Χώροι

Με τον όρο χρωματικοί χώροι εννοούμε τη γεωμετρική αναπαράσταση των χρωμάτων σε ένα χώρο, ο οποίος μας βοηθά να καθορίσουμε τα χρώματα με τη βοήθεια τριών συνιστωσών, των οποίων οι αριθμητικές τους τιμές καθορίζουν ένα συγκεκριμένο χρώμα.

Το χρώμα ορίζεται ως μια κωδικοποίηση του ανθρώπινου νευρικού συστήματος για να διακρίνει τα διάφορα μήκη κύματος του φωτός στο ηλεκτρομαγνητικό φάσμα. Το ανθρώπινο μάτι περιέχει κύτταρα τα οποία λειτουργούν ως φωτοαισθητήρες (Photoreceptor cell) και χωρίζονται σε δύο κατηγορίες. Η πρώτη είναι τα ραβδία (rods), τα οποία είναι υπεύθυνα για την αντίληψη του αμυδρού φωτός και το δεύτερο είδος κυττάρων είναι τα κωνία

(cones), που ανάλογα με τον τύπο τους είναι ευαίσθητα σε διαφορετικά μήκη κύματος με αποτέλεσμα να ερμηνεύουν τα μήκη κύματος σε χρώματα.

Τα κωνία χωρίζονται στα εξής είδη [2]:

- S-κωνία, τα οποία είναι ευαίσθητα σε φωτόνια μικρού κύματος και στο μπλε χρώμα.
- M-κωνία, τα οποία είναι ευαίσθητα σε μεσαίου μήκους κύματος φωτόνια και στο πράσινο χρώμα.
- L-κωνία, τα οποία είναι ευαίσθητα σε φωτόνια μεγάλου μήκους κύματος και στο κόκκινο χρώμα.

Τα χρώματα που αντιστοιχούν σε κάθε κωνίο παρουσιάζονται στον πίνακα 1 [3].

Πίνακας 1 Τα χρώματα και τα μήκη κύματος που αντιστοιχούν σε κάθε τύπο κωνίου.

Φωτοαισθητήρες	Μήκος κύματος σε νανόμετρα (nm)	Peak response σε νανόμετρα (nm)	Διερμηνεία από τον ανθρώπινο εγκέφαλο
Κωνία (τύπου L)	[400 – 680]	564	Κόκκινο
Κωνία (τύπου M)	[400 – 650]	534	Πράσινο
Κωνία (τύπου S)	[370 – 530]	420	Μπλε
Ραβδία	[400 – 600]	498	Απόχρωση του γκρι

Στον πίνακα 2 [2] παρατηρούμε τα χρώματα που αντιστοιχούν στις περιοχές μηκών κύματος και στις περιοχές συχνοτήτων.

Πίνακας 2 Χρώματα του ορατού φάσματος.

Χρώμα	Περιοχή μηκών κύματος (nm)	Περιοχή συχνοτήτων (Hz)
Ερυθρό	~ 630 – 700 nm	~ 476 – 429 x 10 ¹² Hz
Πορτοκαλί	~ 590 – 630 nm	~ 510 – 429 x 10 ¹² Hz
Κίτρινο	~ 560 – 590 nm	~ 535 – 510 x 10 ¹² Hz
Πράσινο	~ 500 – 560 nm	~ 600 – 535 x 10 ¹² Hz
Κυανό	~ 440 – 500 nm	~ 680 – 600 x 10 ¹² Hz
Ιώδες	~ 400 – 440 nm	~ 750 – 680 x 10 ¹² Hz

Τα χρώματα που αντιλαμβάνεται το ανθρώπινο μάτι είναι άπειρα και αυτό οφείλεται στη πρόσπτωση σε αυτό ακτινοβολιών με διαφορετικά μήκη κύματος, με αποτέλεσμα να δημιουργείται ένα νέο χρώμα. Παρατηρούμε, λοιπόν, ότι τα μήκη κύματος που χρειάζεται να μας απασχολήσουν είναι περιορισμένα και εξαρτώνται από το χρωματικό χώρο που θέλουμε να χρησιμοποιήσουμε.

Το χρώμα στην όραση μηχανής μπορεί να αναπαρασταθεί από ένα γραμμικό συνδυασμό τριών πρωταρχικών χρωμάτων, που με άξονα αυτά τα χρώματα μπορούμε να δημιουργήσουμε σχεδόν όλους τους χρωματικούς συνδυασμούς. Τα πρωταρχικά χρώματα χωρίζονται σε δυο μέρη τα προσθετικά και τα αφαιρετικά χρώματα.

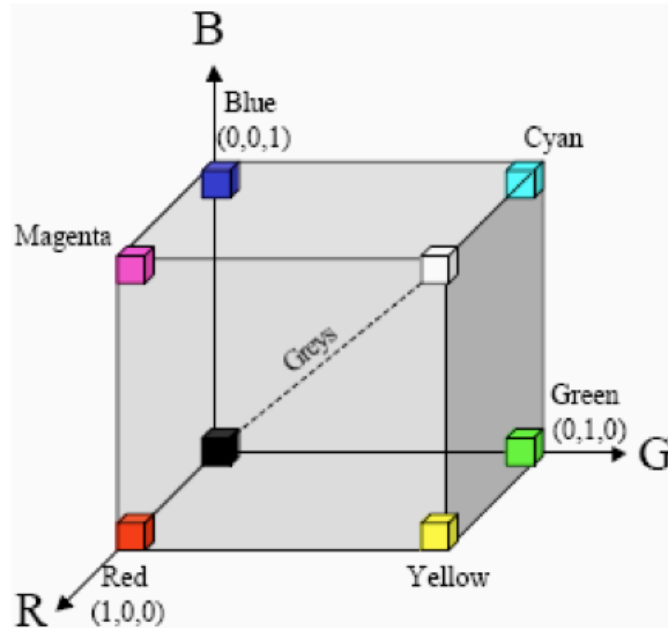
Τα προσθετικά χρώματα απαρτίζονται από το κόκκινο (R – red), το πράσινο (G – green) και το μπλε (B – blue) και αντιστοιχούν στο χρωματικό χώρο RGB. Τα αφαιρετικά χρώματα είναι το κυανό (C – cyan), το πορφυρό (M – magenta) και το κίτρινο (Y – yellow) που αντιστοιχούν στο σύστημα CMY.

3.1.1 Το χρωματικό μοντέλο RGB

Το μοντέλο RGB είναι ο πιο διαδεδομένος χρωματικός χώρος, ο οποίος μπορεί να αναπαρασταθεί με ευκολία σε μια οθόνη. Η αυξημένη πολυπλοκότητα του, τον καθιστά δύσκολο για επεξεργασία σε εφαρμογές όρασης μηχανής και αυτό οφείλεται στο ότι αν σε ένα χρώμα αλλάξει η απόχρωση του, τότε αλλάζουν και οι τιμές του και στους τρεις άξονες.

Το RGB είναι ένα προσθετικό μοντέλο, λόγω του ότι αν προσθέσουμε τα μήκη κύματος της κάθε ακτίνας τότε το άθροισμά τους μας δίνει ως αποτέλεσμα το μήκος κύματος του τελικού χρώματος.

Η τρισδιάστατη αναπαράσταση του μοντέλου RGB παρουσιάζεται γραφικά σαν ένας χρωματικός κύβος, όπως φαίνεται στην εικόνα 6. Οι συντεταγμένες κάθε σημείου μέσα στον κύβο αντιστοιχούν στις τιμές του κόκκινου, πράσινου και του μπλε αντίστοιχα.



Εικόνα 6 Το χρωματικό μοντέλο RGB.

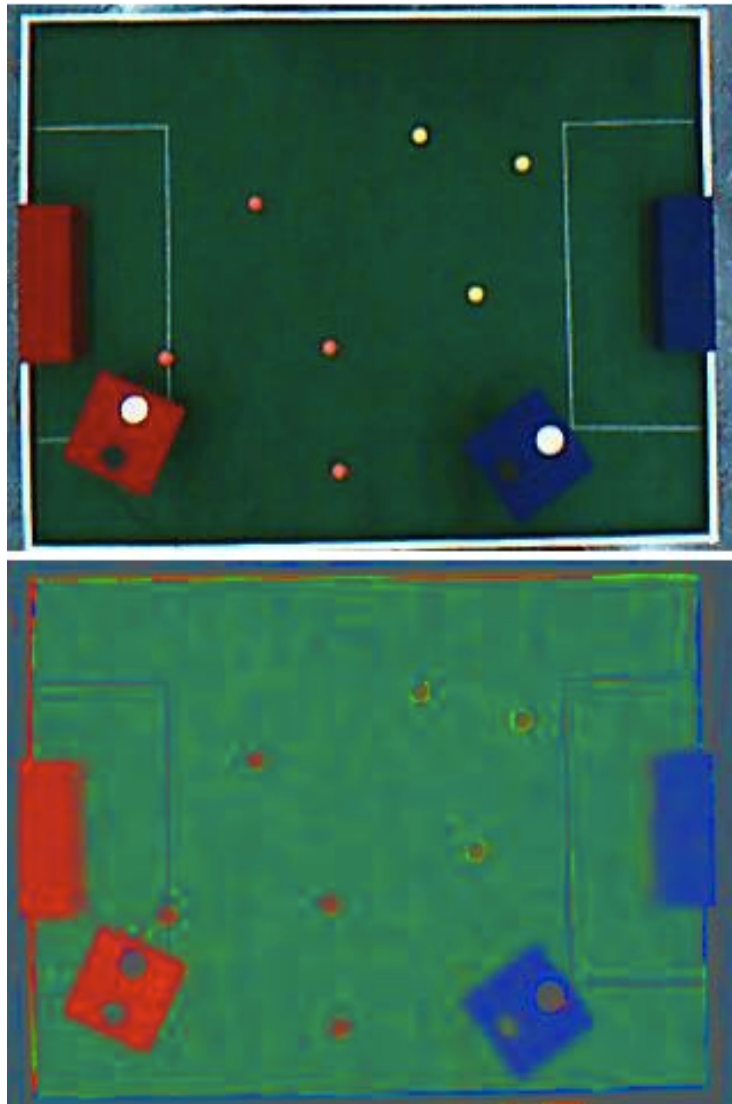
Σύμφωνα με τους νόμους της χρωματομετρίας, οποιοδήποτε χρώμα μπορεί να δημιουργηθεί από τα τρία βασικά χρώματα και ο συνδυασμός των τριών αυτών χρωμάτων είναι μοναδικός. Αν δυο χρώματα είναι ισοδύναμα τότε θα είναι ισοδύναμα και μετά τον πολλαπλασιασμό ή την αφαίρεση, των τριών συντεταγμένων που τα απαρτίζουν, με ένα αριθμό. Η φωτεινότητα ενός μείγματος από χρώματα είναι ίση με το άθροισμα της φωτεινότητας του κάθε χρώματος.

3.1.2 Ο χρωματικός χώρος nRGB ή normalized RGB

Τα φώτα και οι σκιές σε μια εικόνα μπορεί να προκαλέσουν διάφορες αλλοιώσεις σε εικόνα του χρωματικού μοντέλου RGB, για το λόγο αυτό εξετάζουμε το κανονικοποιημένο μοντέλο RGB ή nRGB το οποίο μπορεί να δώσει λύση στο πρόβλημα αυτό. Για να επιτύχουμε την κανονικοποίηση διαιρούμε τις τιμές κάθε εικονοστοιχείου με το άθροισμα των τιμών του εικονοστοιχείου σε όλες τις συνιστώσες. Το κανονικοποιημένο σημείο (r, g, b) δίνεται από τον τύπο στην εξίσωση 1.

Εξίσωση 1 Οι νέες τιμές που θα πάρει η συνιστώσα (r,g,b) .

$$(r, g, b) = \left(\frac{R}{R + G + B}, \frac{G}{R + G + B}, \frac{B}{R + G + B} \right)$$



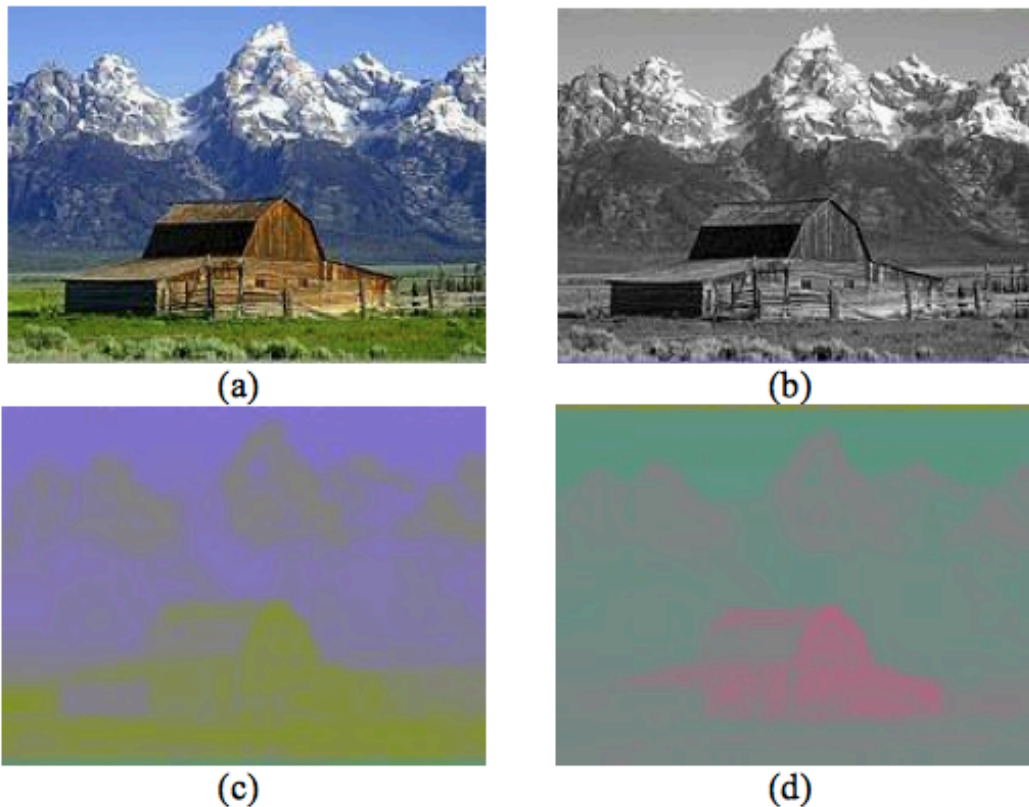
Εικόνα 7 Μετασχηματισμός εικόνας από RGB σε nRGB [4].

Στην εικόνα 7 παρατηρούμε τη διαφορά της πάνω εικόνας, που τραβήχτηκε στο χώρο RGB με την κανονικοποιημένη εκδοχή της. Οι βασικές αλλαγές που εντοπίζουμε στη δεύτερη εικόνα είναι ότι οι σκιές έχουν πλέον εξαφανιστεί, μαύροι και άσπροι κύκλοι έχουν γίνει δυσδιάκριτοι και τα τέρματα δεν έχουν ένα ενιαίο χρώμα.

3.1.3 Ο Χρωματικός χώρος YUV

Ο χρωματικός χώρος YUV αντιστοιχεί στο σύστημα για την κωδικοποίηση χρώματος που χρησιμοποιείται για την αναλογική τηλεόραση. Με το χώρο αυτό μπορούμε να ορίσουμε τη φωτεινότητα και το χρώμα με αριθμούς. Μπορούμε επίσης, να εξάγουμε από την αρχική εικόνα, μια άλλη σε αποχρώσεις του γκρι (Y) και μπορούμε να πάρουμε την πληροφορία χρώματος (U και V). Η πληροφορία χρώματος που εξάγουμε είναι η διαφορά χρώματος των σημάτων,

δηλαδή στο μπλε αφαιρούμε τη φωτεινότητα ($B - Y$) και το ίδιο κάνουμε και στο κόκκινο ($R - Y$). Η μετατροπή μιας RGB εικόνας σε YUV παρουσιάζεται στην εικόνα 8.



Εικόνα 8 (a) Η RGB εικόνα (b) Το στοιχείο Y (c) Το στοιχείο U (d) Το στοιχείο V [9].

Τα δεδομένα που λαμβάνουμε από μία βιντεοκάμερα στο χρωματικό χώρο RGB μέσα από μια διαδικασία που ονομάζεται χρωματική μετατροπή χώρου μετατρέπονται σε σύνθετα αναλογικά σήματα (YUV) ή ψηφιακά YCbCr.

Η μετατροπή του RGB χρωματικού χώρου στον YUV γίνεται βάση της εξίσωσης 2.

Εξίσωση 2 Σχέση μετατροπής από RGB σε YUV.

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad \begin{array}{l} Y \in [0, 255] \\ U \in [-111, 111] \\ V \in [-157, 157] \end{array}$$

3.1.4 Ο Χρωματικός χώρος YCbCr

Ο χρωματικός χώρος YCbCr δημιουργήθηκε για τις ανάγκες του ψηφιακού βίντεο. Είναι ένας χώρος προσαρμοσμένος και αντισταθμισμένος με το χρωματικό χώρο YUV. Το Y αντιπροσωπεύει τη φωτεινότητα και τα Cb, Cr αντιπροσωπεύουν την πληροφορία χρώματος. Το Cb αναπαριστά τη χρωματική διαφορά ως προς το μπλε για παράδειγμα (B - Y) και το Cr εκφράζει τη χρωματική διαφορά ως προς το κόκκινο όπως (R - Y). Αξίζει να σημειωθεί ότι η εικόνα που εξάγουμε από το Y είναι στην ουσία ένα αντίγραφο της αρχικής αλλά σε κλίμακα αποχρώσεων του γκρι.

Ο μετασχηματισμός για τη μετατροπή από το χώρο RGB στο χώρο YCbCr προκύπτει από την εξίσωση 3, στη σχέση αυτή οι τιμές που παίρνουν τα R,G,B βρίσκονται στο διάστημα [0,1] και η τιμή που παίρνει το Y βρίσκεται στο διάστημα [16, 240].

Εξίσωση 3 Σχέση για μετατροπή RGB σε YCbCr.

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112 \\ 112 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

3.1.5 Περιληπτική σύγκριση χρωματικών χώρων

- Ο χώρος RGB είναι ο πιο διαδεδομένος χώρος ο οποίος είναι συμβατός με όλες τις οθόνες και τους αισθητήρες. Μειονέκτημα του είναι η κακή επεξεργασία εικόνας λόγω της υψηλής συσχέτισης που τον χαρακτηρίζει.
- Ο χώρος nRGB παρέχει ανεξαρτησία στο χρώμα από φωτεινότητα και λάμψη αλλά μπορεί να παρουσιάσει προβλήματα θορύβου σε χαμηλή ένταση.
- Οι χώροι YUV και YCbCr μπορούν να υπολογιστούν εύκολα με μεγάλο πλεονέκτημα το Y το οποίο βοηθά στην εύρεση ακμών. Μειονέκτημα τους είναι ότι χρησιμοποιούν γραμμικό μετασχηματισμό και ακατάλληλες αποστάσεις.

3.2 Κατάτμηση Εικόνας με βάση το Χρώμα

Παρουσιάσαμε στην υπό-ενότητα 3.1 τους πιο κλασσικούς χρωματικούς χώρους, σε αυτή την ενότητα της εργασίας θα αναλύσουμε πως οι χρωματικοί χώροι μπορούν να αξιοποιηθούν στην κατάτμηση της εικόνας με βάση το χρώμα. Κανένας από τους χρωματικούς χώρους που έχουμε στη διάθεση μας δεν μπορεί να ξεπεράσει κάποιον άλλο για την κατάτμηση έγχρωμων εικόνων και αυτό οφείλεται στο ότι όλοι οι χώροι έχουν τα μειονεκτήματα και τα πλεονεκτήματά τους. Ο χρωματικός χώρος που θα επιλεγεί πρέπει να είναι ανάλογος της εφαρμογής στην οποία θα εφαρμοστεί.

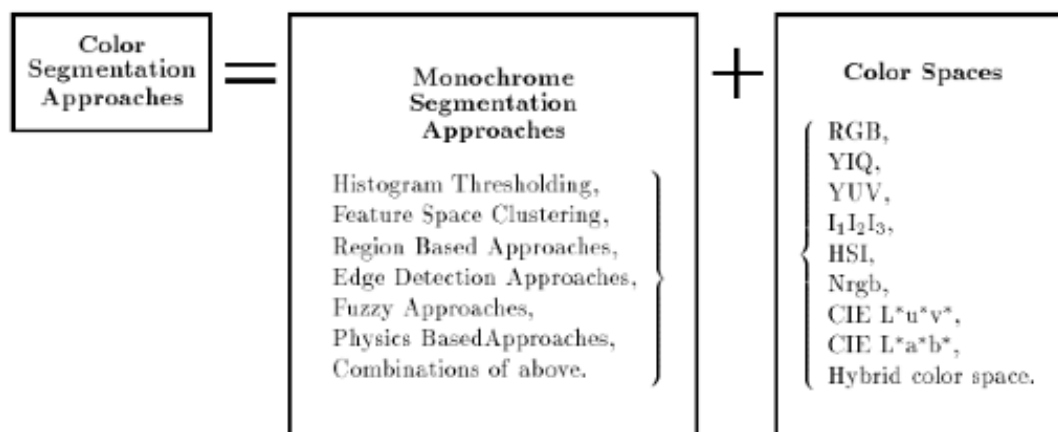
Το ενδιαφέρον για κατάτμηση με βάση το χρώμα ξεκίνησε τη δεκαετία του 90' όταν οι υπολογιστές άρχισαν να έχουν περισσότερες δυνατότητες και η ψηφιακή εικόνα άρχισε να γίνεται όλο και πιο δημοφιλής. Μέχρι τότε η κατάτμηση χρώματος περιοριζόταν στην κατάτμηση αποχρώσεων του γκρι. Η κατάτμηση χρώματος βασίστηκε στην μονοχρωματική κατάτμηση προσθέτοντας σε αυτή μεγάλο ποσοστό αναπαραστάσεων χρώματος.

Σκοπός της κατάτμησης εικόνας με βάση το χρώμα είναι η ομαδοποίηση γειτονικών εικονοστοιχείων, τα οποία μοιράζονται όμοιες χρωματικές ιδιότητες. Τα εικονοστοιχεία της κάθε περιοχής πρέπει να έχουν ομοιογένεια και να υπόκεινται σε συνθήκες συνεκτικότητας.

Μια περιοχή για να θεωρηθεί χρωματική περιοχή πρέπει να πληροί κάποιες προϋποθέσεις. Μια από αυτές είναι ότι το κάθε εικονοστοιχείο απαραίτητα καταχωρείται σε μοναδική και ενιαία περιοχή και όλες οι περιοχές στο σύνολο τους ανταποκρίνονται στα δεδομένα της εικόνας. Άλλη προϋπόθεση είναι ότι κάθε χρωματική περιοχή ορίζεται ως ένα υποσύνολο συνδεδεμένων εικονοστοιχείων, τα οποία έχουν χαρακτηριστικό τους την ομοιομορφία. Οι περιοχές για να ξεχωρίσουν η μια από την άλλη δεν πρέπει να έχουν την ίδια ομοιομορφία. Το αποτέλεσμα που παίρνουμε από την κατάτμηση εικόνας με βάση το χρώμα είναι μια εικόνα στην οποία το κάθε εικονοστοιχείο συνδέεται σε μια ετικέτα που αντιστοιχεί στην περιοχή την οποία υπάγεται.

3.2.1 Τεχνικές κατάτμησης εικόνας με βάση το χρώμα

Για την υλοποίηση εφαρμογής με τη μέθοδο κατάτμησης με άξονα το χρώμα, το πρώτο πρόβλημα που καλούμαστε να αντιμετωπίσουμε είναι η επιλογή της μεθόδου η οποία θα εκτελέσει την κατάτμηση και στη συνέχεια πρέπει να κάνουμε διαλογή της αναπαραστάσης χρώματος ανάμεσα στους χρωματικούς χώρους.



Σχήμα 2 Αρχιτεκτονική κατάτμησης εικόνας με βάση το χρώμα.

Στο σχήμα 2 παρατηρούμε την προσέγγιση που χρειάζεται να γίνει για να κατατμηθεί χρωματικά μια εικόνα. Οι συνδυασμοί μεθόδων κατάτμησης και χρωματικών χώρων εξαρτώνται από το είδος της εικόνας που θέλουμε να επεξεργαστούμε και το σκοπό της κατάτμησης.

3.2.2 Μέθοδοι Κατάτμησης με βάση το χρώμα

Ιστόγραμμα Κατωφλίου - Histogram Thresholding

Ο ευκολότερος τρόπος κατάτμησης είναι με τη χρήση ιστογράμματος κατωφλίου. Με την τεχνική αυτή βασιζόμαστε στο ιστογράμμα της εικόνας το οποίο εξάγεται ως μονοδιάστατο για μονόχρωμες εικόνες και πολυδιάστατο για έγχρωμες. Οι περιοχές ταξινομούνται στο βασικό εύρος τιμών το οποίο εφαρμόζεται στις τιμές έντασης των εικονοστοιχείων μιας εικόνας με τη χρήση του κατωφλίου.

Η τεχνική αυτή είναι μια από τις παλαιότερες μεθόδους κατάτμησης και είναι ευρέως διαδεδομένη ακόμα και σήμερα. Η ανίχνευση κορυφών στο ιστογράμμα οδηγεί στην εύρεση του παρασκηνίου της εικόνας και συνεπώς στην εύρεση αντικειμένων. Το εύρος τιμών του ιστογράμματος οδηγεί στην ταξινόμηση των εικονοστοιχείων με τη χρήση είτε τοπικού είτε ολικού κατωφλίου. Στο τοπικό κατώφλι επιλέγονται διαφορετικές τιμές κατωφλίου για την κάθε περιοχή. Στην ολική τεχνική κατωφλίου επιλέγεται μία μόνο τιμή για όλη την εικόνα. Αν θέλουμε να εφαρμόσουμε την τεχνική κατάτμησης ιστογράμματος κατωφλίου σε σύνθετες εικόνες πρέπει απαραίτητα να χρησιμοποιήσουμε πολυεπίπεδο κατώφλι.

Ένα πρόβλημα που θα χρειαστεί να αντιμετωπίσουμε στη χρήση αυτής της τεχνικής είναι η παρουσία θορύβου στις κορυφές του ιστογράμματος. Για το λόγο αυτό κρίνεται απαραίτητο να γίνει ομαλοποίηση του ιστογράμματος πριν τη χρήση της τεχνικής αυτής.

Τεχνική κατάτμησης με βάση την ομαδοποίηση

Οι τεχνικές κατάτμησης με τη χρήση ιστογράμματος κατωφλίου απαιτούν υπολογισμό του ιστογράμματος, πράγμα που χρειάζεται χρόνο και υπολογιστική μνήμη. Για το λόγο αυτό προτάθηκε η τεχνική ομαδοποίησης σημείων στο χρωματικό χώρο.

Η τεχνική αυτή προϋποθέτει ότι ο αναλυτής προσαρμόζει τον επιθυμητό αριθμό των κλάσεων των εικονοστοιχείων με ένα σημείο βαρύτητας το οποίο είναι και το κέντρο της ομάδας και το οποίο σχετίζεται άμεσα με την κλάση στην οποία υπάγεται. Η τεχνική αυτή ελαχιστοποιεί τον ολικό διασκορπισμό μέσα στην κάθε κλάση.

Αρχικά οι τοποθεσίες των σημείων βαρύτητας καθορίζονται τυχαία από το χρωματικό χώρο. Στη συνέχεια μετά από μια επανάληψη τα σημεία αυτά ενημερώνονται σύμφωνα με τα παρακάτω:

- Το κάθε εικονοστοιχείο ανατίθεται στην κλάση της οποίας το κέντρο βάρους είναι το πλησιέστερο σε χρώμα.
- Το κέντρο βάρους κάθε κλάσης ενημερώνεται με το να λαμβάνει υπόψη τα εικονοστοιχεία που του ανατίθενται.
- Η διακύμανση των κέντρων βάρους προσδιορίζεται ανάμεσα στο προηγούμενο ($t - 1$) και το τρέχον βήμα t από τον τύπο στην εξίσωση 4.

Εξίσωση 4 Τύπος για την εύρεση της διακύμανσης του κέντρου βάρους.

$$\varepsilon = \sum_{j=1 \dots N\omega} || C_j(t) - C_j(t - 1) ||$$

- Αν η επανάληψη γίνει μεγαλύτερη από κάποιο κατώφλι τότε η διαδικασία επιτυγχάνει μια νέα επανάληψη.

Τεχνικές με βάση την περιοχή - Region based

Οι τεχνικές κατάτμησης με βάση την περιοχή είναι τεχνικές βασιζόμενες στη συνέχεια. Οι τεχνικές αυτές διαχωρίζουν την εικόνα σε επιμέρους περιοχές. Βασίζονται σε κοινά πρότυπα στις τιμές έντασης σε μια ομάδα γειτονικών εικονοστοιχείων και έχουν στόχο την ομαδοποίηση των περιοχών ανάλογα με τη λειτουργία ή τον ρόλο τους.

Η επεξεργασία ξεκινά να υφίσταται αφού καθορίσουμε κάποια αρχικά σημεία στην εικόνα. Το κάθε σημείο καθορίζει μια περιοχή την οποία και επεκτείνει προσθέτοντας σε αυτή γειτονικά σημεία τα οποία έχουν το κριτήριο

της ομοιογένειας. Αν κανένα σημείο δεν μπορεί να προστεθεί τότε σταματά η επέκταση. Η επιλογή λοιπόν των αρχικών σημείων είναι καίριας σημασίας γιατί αν κρατήσουμε σταθερό το κριτήριο της ομοιογένειας τότε τα αποτελέσματα θα εξαρτηθούν από αυτά. Για τη βελτιστοποίηση του αποτελέσματος στα τελικά στάδια χρησιμοποιούμε τη μέθοδο της συνένωσης των περιοχών ή τον διαχωρισμό τους σε μικρότερες.

Με τη μέθοδο αυτή μπορούμε να κάνουμε κατάτμηση της εικόνας με βάση το χρώμα γρήγορα και εξασφαλίζουμε την εξαγωγή χωρικά συνδεδεμένων περιοχών.

Τεχνικές ανίχνευσης ακμών - Edge Detection

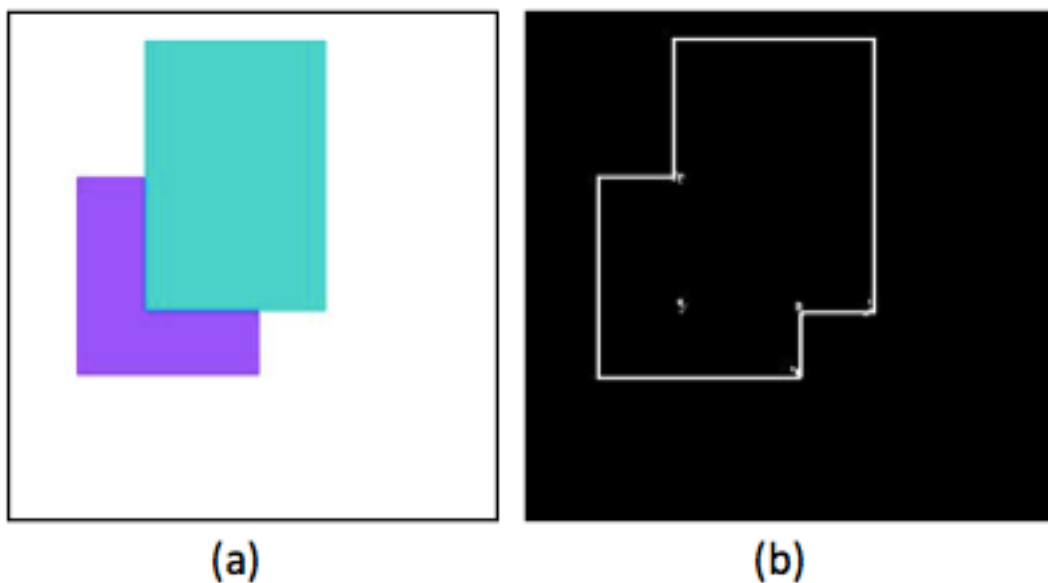
Οι τεχνικές ανίχνευσης ακμών χρησιμοποιούνται εκτενώς στην κατάτμηση εικόνων στην κλίμακα του γκρι και σκοπός τους είναι η ανίχνευση σημείων που παρουσιάζουν απότομες χρωματικές μεταβολές.

Ταξινομούνται κυρίως σε δύο κατηγορίες την παράλληλη ανίχνευση ακμών και τη σειριακή ανίχνευση. Στην παράλληλη τεχνική ανίχνευσης ακμών η απόφαση για το αν ένα σύνολο σημείων αντιπροσωπεύει ή όχι μια ακμή δεν εξαρτάται από το αν άλλα σύνολα σημείων βρίσκονται σε μια άλλη ακμή. Με άλλα λόγια, ο χειριστής ανίχνευσης ακμών μπορεί να εφαρμοστεί παράλληλα σε όλη την εικόνα. Στην σειριακή ανίχνευση ακμών το αποτέλεσμα ενός σημείου εξαρτάται από τα αποτελέσματα των σημείων που εξετάσαμε πριν από αυτό το σημείο. Η απόδοση ενός σειριακού αλγορίθμου ανίχνευσης ακμών εξαρτάται άμεσα από την επιλογή του αρχικού σημείου. Να σημειώσουμε ότι το σημείο τερματισμού του αλγόριθμου είναι δύσκολο να προσδιοριστεί.

Σε μια μονόχρωμη εικόνα οι ακμές καθορίζονται ως μια ασυνέχεια σε αποχρώσεις του γκρι. Σε μια εικόνα με χρώμα η διαδικασία εύρεσης ακμών είναι πολύπλοκότερη και αυτό οφείλεται στο ότι η πληροφορία της εικόνας είναι μεγαλύτερη από αυτήν που παίρνουμε από τις μονόχρωμες.

Οι ακμές σε πολύχρωμες εικόνες καθορίζονται από μια ασυνέχεια στους τρισδιάστατους χρωματικούς χώρους. Οι τεχνικές που ακολουθούνται είναι: η επεξεργασία και στα τρία χρώματα ταυτόχρονα και η επεξεργασία σε κάθε χρώμα χωριστά όπου τα αποτελέσματα συνδυάζονται με διάφορα κριτήρια. Με την παρακολούθηση των χρωματικών αλλαγών στα εικονοστοιχεία βρίσκουμε την κατεύθυνση των αλλαγών αυτών και μπορούμε έτσι να προβλέψουμε τα όρια των ακμών.

Στην εικόνα 9 [6], παρατηρούμε το αποτέλεσμα της χρήσης τεχνικών ανίχνευσης ακμών με βάση το χρώμα.



Εικόνα 9 (a) Η εικόνα χωρίς επεξεργασία (b) Το αποτέλεσμα αν στην εικόνα (a) χρησιμοποιήσουμε την τεχνική ανίχνευσης ακμών.

Τεχνικές Τεχνητών νευρωνικών δικτύων – Neural Networks

Τα τεχνητά νευρωνικά δίκτυα μπορούν να χρησιμοποιηθούν στην κατάτμηση εικόνας με βάση το χρώμα. Διαφέρουν από τις άλλες τεχνικές στο ότι χρησιμοποιούν μια μη γραμμική εκτέλεση λειτουργιών κατάτμησης για να καταλήξουν στην τελική υπόθεση. Οι περιοχές στην εικόνα αναπαριστούνται από τους πολλαπλούς νευρώνες μέσω των διασυνδέσεων τους.

Το πρώτο βήμα που ακολουθείται στην τεχνική αυτή είναι να μετατρέψουμε την εικόνα σε ένα χάρτη εικονοστοιχείων, δηλαδή ένα χάρτη αποτελούμενο από τετράγωνα με διαφορετική ένταση χρώματος γκρι. Ο χάρτης αυτός έχει δισδιάστατη λειτουργία με δυο μεταβλητές εισόδου που αντιστοιχούν στην θέση του κάθε εικονοστοιχείου και μια μεταβλητή εξόδου που αντιστοιχεί στην ένταση του εικονοστοιχείου αυτού.

Η εικόνα μπορεί να αναπαρασταθεί ως ένας δισδιάστατος πίνακας όπου το κάθε στοιχείο της εν λόγω συστοιχίας περιέχει πληροφορίες χρώματος για κάθε ένα εικονοστοιχείο. Οι πληροφορίες αυτές είναι η είσοδος για το τεχνητό νευρωνικό δίκτυο εκπαίδευσης. Η εκπαίδευση ξεκινά με τη χρήση μιας εξίσωσης λάθους (βλ. Εξίσωση 5) όπου τα $n(x, y)$ είναι οι διαστάσεις της εικόνας. Το κάθε εικονοστοιχείο καθορίζεται από τα x, y, z και το z αντιπροσωπεύει την ένταση του εικονοστοιχείου. Το $o(x, y)$ είναι η έξοδος του νευρωνικού δικτύου όταν η είσοδος του είναι η θέση του εικονοστοιχείου i, j .

Εξίσωση 5 Εξίσωση εκπαίδευσης λαθών σε τεχνητά νευρωνικά δίκτυα.

$$E[\omega, \theta] \equiv \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} (o(\mathbf{x}_i, \mathbf{y}_j) - z_{ij})^2$$

Όταν τελειώσει η εκπαίδευση του τεχνητού νευρωνικού δικτύου τότε μπορούμε να το χρησιμοποιήσουμε για ανίχνευση αντικειμένων σε μια εικόνα. Αυτό γίνεται με τη σύγκριση της εξίσωσης λάθους μεταξύ του εκπαιδευμένου τεχνητού δικτύου και μιας διαφορετικής εικόνας.

Τεχνικές ασαφούς λογικής - Fuzzy Techniques

Οι τεχνικές κατάτμησης που προαναφέραμε λαμβάνουν καθαρές αποφάσεις σε ότι έχει να κάνει με τις περιφέρειες (regions) παρόλα αυτά οι περιφέρειες σε μια εικόνα δεν είναι πάντα διακριτές. Αν γίνει μια λανθασμένη υπόθεση στα αρχικά στάδια της επεξεργασίας τότε θα επηρεαστούν και όλα τα επόμενα στάδια που θα ακολουθήσουν αφού οι αποφάσεις σε κάθε επίπεδο βασίζονται στα αποτελέσματα του προηγούμενου επιπέδου.

Ένα σύστημα όρασης μηχανής πρέπει να έχει επαρκή ευελιξία για την επεξεργασία της αβεβαιότητας σε όλα τα επίπεδα έτσι ώστε σε περίπτωση λαθών να μην καταδικάζεται το αποτέλεσμα.

Οι τεχνικές ασαφούς λογικής παρέχουν το μηχανισμό για να αναπαρασταθεί και να διαχειριστεί η αβεβαιότητα και η ασάφεια. Ο μηχανισμός αυτός προσφέρει τη δυνατότητα να μεταφέρονται οι αβέβαιες υποθέσεις και η εκτίμηση τους σε ανώτερα επίπεδα. Σε κάθε επίπεδο συγκρίνονται μεταξύ τους για να αναδείξουν την απόφαση που θα επικρατήσει στο τελευταίο στάδιο επεξεργασίας.

Μια από τις πιο δημοφιλείς τεχνικές ασαφούς λογικής είναι η F.C.M [7], η τεχνική αυτή επιτρέπει ασαφή όρια ανάμεσα στις περιοχές. Τα στάδια λειτουργίας του αλγορίθμου ξεκινούν με τον καθορισμό της αβεβαιότητας ενός εικονοστοιχείου σε σχέση με το σύμπλεγμα (cluster) στο οποίο ανήκει, μέσω της απόστασής του από το κέντρο του συμπλέγματος. Το επόμενο βήμα είναι ο επαναπροσδιορισμός των συμπλεγμάτων και ο υπολογισμός του κέντρου τους. Τέλος, γίνεται επανάληψη των προηγούμενων βημάτων μέχρι η απόσταση που έχουν τα κέντρα των συμπλεγμάτων να φτάσει στο μέγιστο σημείο της.

Τεχνικές βασισμένες στη φυσική - Physics based

Οι τεχνικές κατάτμησης χρώματος με βάση τη φυσική εκμεταλλεύονται τις ιδιότητες των υλικών με τη χρήση των μοντέλων της φυσικής και στοχεύουν

στην ανίχνευση των ορίων των αντικειμένων, εξαλείφοντας παράλληλα τις ψευδείς ακμές που προέρχονται από σκιά ή φωτισμό.

Οι πιο διαδεδομένες τεχνικές κατάτμησης χρώματος με βάση τη φυσική είναι το «διχρωματικό μοντέλο ανάκλασης - dichromatic reflection model» [10] και το «ενδεικτικό μοντέλο ανάκλασης χρώματος - Approximate Color Reflectance Model (ACRM)» [11]. Το A.C.R.M. χωρίζει τα αντικείμενα σε δυο κύριες κλάσεις: οπτικά ομογενή υλικά όπως πλαστικά, υφάσματα και μπογιές, και οπτικά ανομοιογενή υλικά όπως μέταλλα και γυαλιά.

Τα μοντέλα φυσικής, δίνοντας σημασία κυρίως σε αντανάκλασεις, έχουν επιτυχή εφαρμογή σε συγκεκριμένους χώρους, στους οποίους είναι γνωστά από πριν τα φυσικά αντικείμενα της εικόνας και οι ιδιότητές τους.

4

Υλικό και Λογισμικό

4.1 Υλικό Μέρος Ανάπτυξης εφαρμογής

Το σύστημα εφαρμογής που αναπτύξαμε υλοποιήθηκε με τη χρήση ηλεκτρονικού υπολογιστή. Τα δεδομένα εικόνων στο πρόγραμμα προήλθαν από μια κάμερα διεπαφής προσωπικού υπολογιστή με εξωτερική σύνδεση καλωδίου USB.

4.1.1 Κάμερα

Η κάμερα που χρησιμοποιήθηκε ήταν μια Logitech QuickCam Communicate STX Web Cam με ανάλυση 1.3MP. Η συγκεκριμένη κάμερα είναι μια εμπορική συσκευή κατασκευασμένη για χρήση στο σπίτι ή στο γραφείο.



Εικόνα 10 Logitech QuickCam.

Τα τεχνικά χαρακτηριστικά της συσκευής αυτής είναι τα ακόλουθα:

- Αισθητήρας: High quality VGA
- Λήψη Βίντεο: Μέχρι 640 x 480 εικονοστοιχεία
- Λήψη Εικόνας: 1.3 meγαρίxels
- Τύπος Βίντεο: AVI
- Τύποι εικόνας: BMP, JPEG
- Ρυθμός καρτέ: Μέχρι 30 fps - καρτέ ανά δευτερόλεπτο
- Οπτικό πεδίο: 42 μοίρες οριζόντια
- Οπτική: Σταθερή εστίαση

Η εφαρμογή που αναπτύξαμε έχει την ευελιξία να μπορεί να δεχθεί ως είσοδο ακόμα και μια απλή κάμερα. Η συγκεκριμένη κάμερα (βλ. Εικόνα 10) επιλέχθηκε να χρησιμοποιηθεί γιατί πληρούσε τα κριτήρια που απαιτούνται για να έχουμε βέλτιστο αποτέλεσμα. Ένα από τα βασικά κριτήρια που θέσαμε ως απαραίτητα είναι η φορητότητα, δηλαδή να μπορεί η κάμερα να μετακινείται για να σαρώνει τα αντικείμενα που θέλουμε να επεξεργαστεί το πρόγραμμά μας. Άλλα κριτήρια είναι η συνδεσιμότητα και η συμβατότητα με το λειτουργικό μας σύστημα.

Ο πυρήνας του Linux απαιτεί να δημιουργηθεί ένας εικονικός κόμβος συσκευής για την πρόσβαση και τον έλεγχο της εν λόγω συσκευής. Στην περίπτωση που έχουμε περισσότερες από μια κάμερες συνδεδεμένες στον υπολογιστή μας, χρειάζεται να γνωρίζουμε τους κόμβους της κάθε συσκευής για να επιλέξουμε την πιο κατάλληλη για την εφαρμογή μας. Τα βήματα που ακολουθήσαμε για να μάθουμε ποιός κόμβος αντιστοιχεί στη συσκευή που χρησιμοποιήσαμε είναι τα παρακάτω:

- Πριν συνδέσουμε τη κάμερα γράψαμε στο terminal (βλ. Υπό – ενότητα 4.2.2) την εντολή «ls /dev/video *».
- Σημειώσαμε ποιές είναι οι υπάρχουσες συσκευές και συνδέσαμε την κάμερα μας.
- Μετά από ένα μικρό χρονικό διάστημα ξανατρέχουμε την παραπάνω εντολή και το νέο αποτέλεσμα που δημιουργείται ανήκει στην εν λόγω συσκευή. (π.χ. /dev/video0).

4.1.2 Ηλεκτρονικός Υπολογιστής

Η ανάπτυξη του κώδικα έγινε σε φορητό υπολογιστή, ο οποίος έχει τα ακόλουθα τεχνικά χαρακτηριστικά:

- Μοντέλο: Macbook (13- inch, Aluminium, Late 2008)
- Επεξεργαστής Γραφικών: NVIDIA GeFORCE 9400M
- Επεξεργαστής: 2.4GHz Intel Core 2 Duo
- Μνήμη: 4GB of 1066MHz DDR3 SDRAM
- Σκληρός δίσκος: 250GB 5400-rpm Serial ATA

- Λειτουργικό Σύστημα: OS X Mavericks

4.2 Λογισμικό Μέρος Ανάπτυξης εφαρμογής

- Η εφαρμογή αναπτύχθηκε στο ανοιχτού κώδικα λειτουργικό σύστημα Ubuntu 12.04 LTS (Precise Pangolin) 64 – bit.
- Η πλατφόρμα ανάπτυξης λογισμικού που χρησιμοποιήθηκε ήταν το Eclipse CDT (C/C++ Development Tools).
- Η γλώσσα προγραμματισμού με την οποία γράφτηκε ο κώδικας της εφαρμογής είναι η C++ με τη χρήση της βιβλιοθήκης OpenCV που θα αναλύσουμε στη συνέχεια.

4.2.1 Λειτουργικό Σύστημα Ubuntu

Ο φορητός υπολογιστής είχε εγκατεστημένο λειτουργικό Ubuntu 12.04 LTS. Το Ubuntu είναι ένα ελεύθερο λειτουργικό σύστημα που βασίζεται σε πυρήνα Linux. Σαν λειτουργικό είναι εύκολο στη χρήση και μπορεί να χρησιμοποιηθεί από όλους. Η ασφάλεια είναι ένα από τα πιο σημαντικά πλεονεκτήματα του Ubuntu και αυτό γιατί υπάρχουν ελάχιστα κακόβουλα προγράμματα φτιαγμένα για το λειτουργικό αυτό. Η αρχιτεκτονική και οι ρυθμίσεις του αποτρέπουν οποιαδήποτε εξωτερική επίθεση.

Η φιλοσοφία στην οποία στηρίζεται είναι [20]:

- Το λογισμικό πρέπει να είναι ανοιχτού κώδικα και δωρεάν προς όλους.
- Ο χρήστης έχει πρόσβαση στα εργαλεία λογισμικού στη μητρική του γλώσσα.
- Το λογισμικό πρέπει να είναι εύκολο στην χρήση.
- Το λογισμικό μπορεί να τροποποιηθεί από τον κάθε χρήστη ανάλογα με τις ανάγκες του.

Το ότι το Ubuntu είναι ανοιχτού κώδικα και δωρεάν είναι ένας από τους λόγους που το επιλέξαμε ως το λειτουργικό σύστημα της εφαρμογής μας. Άλλος λόγος είναι η ασφάλεια και η ταχύτητα του, χάρη στην έξυπνη διαχείριση πόρων που διαθέτει. Είναι φιλικό προς το χρήστη λόγω του ότι διατίθεται από το διαδίκτυο [20] και έχει ελάχιστες απαιτήσεις από τον υπολογιστή που είναι εγκατεστημένο.

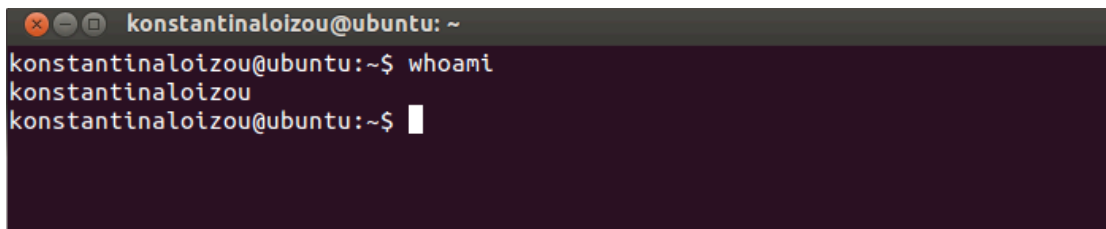
4.2.2 Πλατφόρμα ανάπτυξης λογισμικού Eclipse CDT

Το Eclipse είναι μια πλατφόρμα ανάπτυξης λογισμικού που περιέχει βασικό περιβάλλον εργασίας (workspace) και μπορεί να προσαρμοστεί ανάλογα με τις ανάγκες του προγραμματιστή. Το Eclipse CDT είναι η έκδοση

που χρησιμοποιήσαμε και αντιπροσωπεύει την έκδοση της πλατφόρμας για ανάπτυξη κώδικα σε C/C++. Ο μεταγλωττιστής (compiler) που υποστηρίζει είναι ο gcc.

Εγκατάσταση του Eclipse CDT στο λειτουργικό σύστημα Ubuntu 12.04 μέσω terminal

Το terminal είναι ένα περιβάλλον στο οποίο ο χρήστης μπορεί να πληκτρολογήσει και να εκτελέσει εντολές με τη μορφή διαδοχικών γραμμών κειμένου [16] (βλ. Εικόνα 11).



Εικόνα 11 Το περιβάλλον του Terminal των Ubuntu.

Πριν την εγκατάσταση της πλατφόρμας Eclipse πρέπει να εγκαταστήσουμε στο λειτουργικό μας τη java JRE [21]. Η εγκατάσταση γίνεται γράφοντας την παρακάτω εντολή στο terminal:

```
sudo apt-get install openjdk-7-jre
```

Επόμενο βήμα, είναι να κάνουμε λήψη της τελευταίας έκδοσης του προγράμματος και αφού ολοκληρωθεί η διαδικασία να την μετακινήσουμε στον κατάλογο /opt γράφοντας:

```
sudo mv eclipse-CDT-X.X.X-linux-gtk.tar.gz  
/opt/tk.tar.gz /opt/
```

Αποσυμπιέζουμε το αρχείο στον κατάλογο αυτό:

```
cd /opt  
sudo tar -xvf eclipse-CDT-X.X.X-linux-gtk.tar.gz
```

Δημιουργούμε ένα αρχείο επιφάνειας εργασίας (desktop file) με το όνομα eclipse.desktop στον κατάλογο /usr/share/applications/ γράφοντας:

```
gedit eclipse.desktop
```

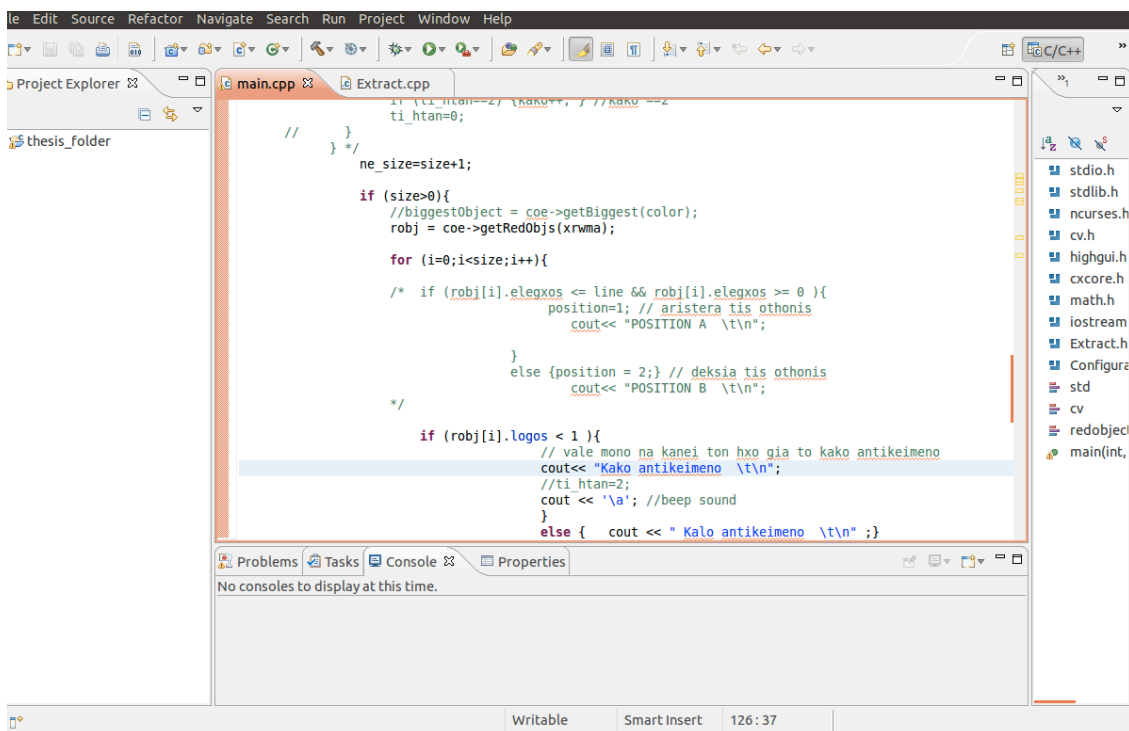
Μέσα στο eclipse.desktop προσθέτουμε τις εξής γραμμές:

```
[Desktop Entry]  
Name=Eclipse
```

```
Type=Application
Exec=eclipse
Terminal=false
Icon=eclipse
Comment=Integrated Development Environment
NoDisplay=false
Categories=Development;IDE;
Name[en]=Eclipse
```

Τέλος, για αυτόματη εγκατάσταση γράφουμε στο terminal την εντολή:

```
desktop-file-install eclipse.desktop
```



Εικόνα 12 Πλατφόρμα ανάπτυξης λογισμικού Eclipse.

Το περιβάλλον ανάπτυξης εφαρμογών Eclipse CDT φαίνεται στην εικόνα 12.

Η εκτέλεση και η μεταγλώττιση ενός προγράμματος στο Eclipse CDT χρειάζεται τα ακόλουθα βήματα:

- Δημιουργία ενός νέου workspace ή χρήση κάποιου από τα υπάρχοντα. Τα workspaces είναι οι κατάλογοι όπου το Eclipse αποθηκεύει τις πληροφορίες και τους φακέλους που σχετίζονται με ένα ή περισσότερα προγράμματα.

- Δημιουργία ενός νέου C/C++ έργου (project). Τα projects περιέχουν τα προγράμματα και είναι υπεύθυνα για το διαχειρισμό τους έτσι ώστε να διαμορφωθεί ένα ενιαίο εκτελέσιμο πρόγραμμα.
- Τέλος, προσθέτουμε αρχεία στο project που δημιουργήσαμε, το κάνουμε Build και το τρέχουμε.

4.2.3 Γλώσσα Προγραμματισμού C++

Η γλώσσα προγραμματισμού C++ είναι μια γλώσσα γενικού σκοπού, μεσαίου επιπέδου και αυτό οφείλεται στο γεγονός ότι περιλαμβάνει χαρακτηριστικά τόσο από γλώσσες χαμηλού επιπέδου όσο και από υψηλού επιπέδου. Υποστηρίζει διάφορα είδη προγραμματισμού όπως για παράδειγμα το δομημένο και τον αντικειμενοστραφή.

Η γλώσσα εμπνεύστηκε και αναπτύχθηκε το 1979 από τον Bjarne Stroustrup καθώς εργαζόταν για τη διδακτορική του διατριβή. Αρχικός σκοπός της ήταν η βελτίωση της ήδη υπάρχουσας γλώσσας προγραμματισμού C και για αυτό το πρώτο της όνομα ήταν «C with Classes» δηλαδή C με κλάσεις. Το 1983 το όνομα της άλλαξε και μετονομάστηκε σε C++. Κάποια από τα νέα χαρακτηριστικά που προστέθηκαν τότε στη γλώσσα, εκτός από τις κλάσεις, είναι οι εικονικές συναρτήσεις (virtual functions), η υπερφόρτωση τελεστών (function overloading), η πολλαπλή κληρονομικότητα και πολλά άλλα.

Το 1990 ο Bjarne Stroustrup δημοσίευσε το βιβλίο με θέμα «The Annotated C++ Reference Manual» [17] το οποίο λειτούργησε ως καταλυτικός παράγοντας το 1998 για τον παγκόσμιο ορισμό της με το πρότυπο ISO/IEC 14882:1998. Η γλώσσα C++ εξελίσσεται ακόμα και σήμερα [13].

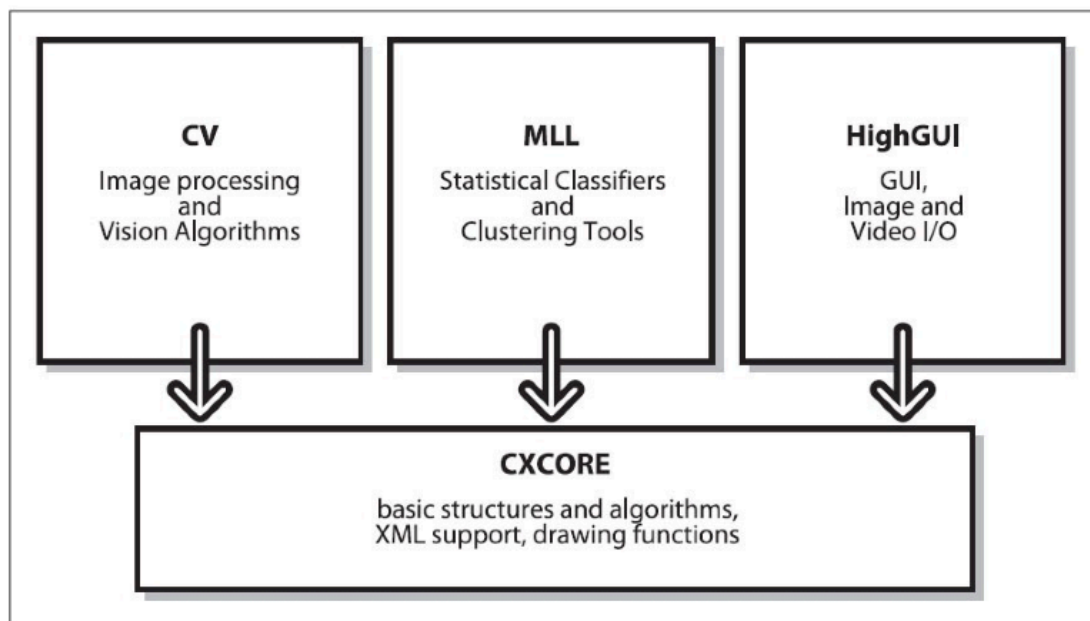
Κάποια από τα χαρακτηριστικά που έχει η C++ ως γλώσσα προγραμματισμού είναι ότι μεταγλωττίζεται άμεσα με τον εγγενή κώδικα μηχανής κάτι που την κάνει πολύ γρήγορη. Υποστηρίζει έλεγχο στατικού αλλά και δυναμικού τύπου προσφέροντας ευελιξία στον προγραμματιστή. Υπάρχουν πάρα πολλές πλατφόρμες οι οποίες την υποστηρίζουν. Ένα μεγάλο πλεονέκτημα που έχει η C++ είναι η συμβατότητα της με τη C δηλαδή μπορεί να χρησιμοποιήσει σχεδόν όλο τον κώδικα της C αλλά και τις βιβλιοθήκες της χωρίς να χρειάζεται να γίνει οποιαδήποτε μετατροπή σε αυτές.

4.2.4 Η Βιβλιοθήκη όρασης μηχανής OpenCV

Η βιβλιοθήκη της OpenCV (Open Computer Vision) είναι μια βιβλιοθήκη όρασης μηχανής η οποία αναπτύχθηκε αρχικά από την Intel και είναι ελεύθερη τόσο για ερευνητική χρήση όσο και για εμπορική, υπό την άδεια ανοιχτού λογισμικού BSD (Berkeley Software Distribution) [12]. Περιέχει μια μεσαίου έως υψηλού επιπέδου διασύνδεση εφαρμογών με πάνω από πεντακόσιες συναρτήσεις γραμμένες σε C και C++. Η OpenCV σχεδιάστηκε με σκοπό να

εκμεταλλευτεί τους υπολογιστικούς πόρους σε εφαρμογές πραγματικού χρόνου.

Στο σχήμα 3 παρουσιάζονται τα δομικά συστατικά της OpenCV. Το CV περιέχει αλγόριθμους για επεξεργασία εικόνας και όρασης μηχανής. Το ML αναπαριστά στατιστικούς ταξινομητές και τα εργαλεία για ομαδοποίηση. Το HighGUI περιέχει τις συναρτήσεις για την προβολή και αποθήκευση βίντεο, τις ρουτίνες εισόδου εξόδου αλλά και όλες τις διεπαφές του χρήστη με την εφαρμογή. Τέλος έχουμε το CxCore και το CxAux. Στο CxCore βρίσκουμε τις βασικές δομές δεδομένων και αλγορίθμων. Το CxAux περιέχει αλγορίθμους οι οποίοι δεν χρησιμοποιούνται και αυτό οφείλεται είτε στη μη λειτουργικότητα τους είτε στο ότι βρίσκονται ακόμα σε πειραματικό στάδιο.



Σχήμα 3 Βασική δομή της OpenCV.

Το βασικότερο πλεονέκτημα που έχει η OpenCV είναι η ταχύτητα επεξεργασίας βίντεο σε πραγματικό χρόνο. Άλλο πλεονέκτημα είναι ότι η OpenCV χρησιμοποιείται από πολλές εταιρίες όπως η Intel, η IBM, η Google η Sony και η Toyota κάτι που βοηθά στην ραγδαία ανάπτυξη της. Μεγάλο πλεονέκτημα της OpenCV είναι ότι υποστηρίζει από τον πιο απλό εξοπλισμό μέχρι και τον πιο πολύπλοκο. Η εφαρμογή μας για παράδειγμα έγινε με τη χρήση μιας απλής webcam.

Η OpenCV συμβάλλει σημαντικά στον τομέα της όρασης μηχανής. Διαθέτει ρουτίνες για την ανάγνωση, την απεικόνιση, την επεξεργασία και την αποθήκευση συμβατικών εικόνων αλλά και διαδοχικών εικόνων από κάμερα. Οι ρουτίνες αυτές υποστηρίζουν, αλγόριθμους ανίχνευσης αντικειμένων, επεξεργασία εικόνας, αναγνώρισης προτύπων και υπολογιστικής γεωμετρίας.

Βασικός στόχος της βιβλιοθήκης είναι η παροχή των βασικών εργαλείων που απαιτούνται για να λυθούν προβλήματα στον τομέα της όρασης μηχανής. Με τη χρήση της μπορούν να λυθούν προβλήματα αρκετά σύνθετα προσφέροντας μια πλήρη λύση σε αυτά χάρη στην εκμετάλλευση των πολλών συναρτήσεων που παρέχει.

Εγκατάσταση της OpenCV μέσω terminal

Η εγκατάσταση της OpenCV είναι μια διαδικασία που αρκετός κόσμος αντιμετωπίζει πρόβλημα, για το λόγο αυτό παρέχεται από την επίσημη ιστοσελίδα του Ubuntu ένα έτοιμο σενάριο (script) [22]. Γράφουμε το script στο gedit και το αποθηκεύουμε ως opencv.sh. Το περιεχόμενο του πρέπει να είναι:

```
version="$(wget -q -O - http://sourceforge.net/projects/opencvlibrary/files/opencv-unix | egrep -m1 -o '\"[0-9](\.[0-9])+' | cut -c2-)"
echo "Installing OpenCV" $version
mkdir OpenCV
cd OpenCV
echo "Removing any pre-installed ffmpeg and x264"
sudo apt-get -qq remove ffmpeg x264 libx264-dev
echo "Installing Dependences"
sudo apt-get -qq install libopencv-dev build-essential checkinstall cmake pkg-config yasm libjpeg-dev libjasper-dev libavcodec-dev libavformat-dev libswscale-dev libdc1394-22-dev libxine-dev libgstreamer0.10-dev libgstreamer-plugins-base0.10-dev libv4l-dev python-dev python-numpy libtbb-dev libqt4-dev libgtk2.0-dev libfaac-dev libmp3lame-dev libopencore-amrnb-dev libopencore-amrwb-dev libtheora-dev libvorbis-dev libxvidcore-dev x264 v4l-utils ffmpeg
echo "Downloading OpenCV" $version
wget -O OpenCV-$version.zip http://sourceforge.net/projects/opencvlibrary/files/opencv-unix/$version/opencv-"$version".zip/download
echo "Installing OpenCV" $version
unzip OpenCV-$version.zip
cd opencv-$version
mkdir build
cd build
cmake -D CMAKE_BUILD_TYPE=RELEASE -D CMAKE_INSTALL_PREFIX=/usr/local -D WITH_TBB=ON -D BUILD_NEW_PYTHON_SUPPORT=ON -D WITH_V4L=ON -D INSTALL_C_EXAMPLES=ON -D INSTALL_PYTHON_EXAMPLES=ON
```

```

-D BUILD_EXAMPLES=ON -D WITH_QT=ON -D WITH_OPENGL=ON
..
make -j2
sudo checkinstall
sudo sh -c 'echo "/usr/local/lib" >
/etc/ld.so.conf.d/opencv.conf'
sudo ldconfig
echo "OpenCV" $version "ready to be used"

```

Στο terminal γράφουμε:

```

chmod +x opencv.sh
./opencv.sh

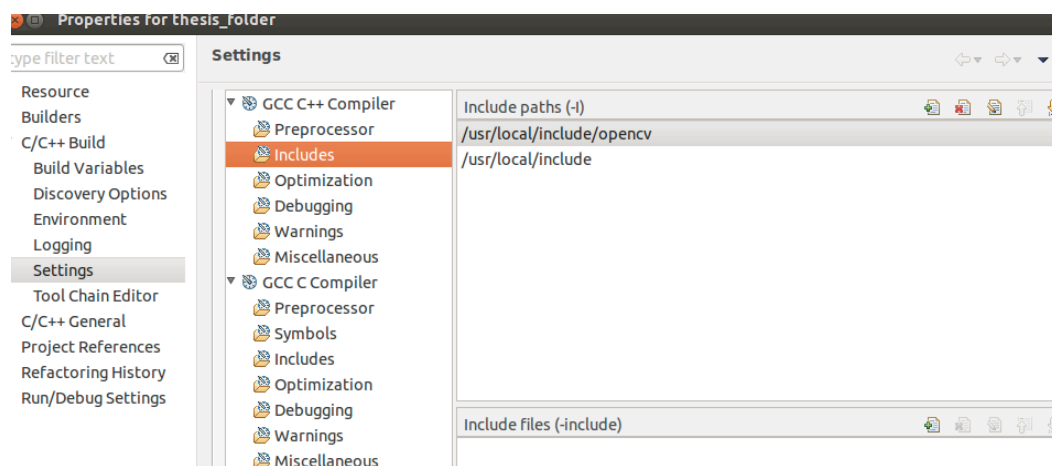
```

4.2.5 Χρησιμοποιώντας την OpenCV στην πλατφόρμα Eclipse

Αφού εγκαταστήσουμε στο λειτουργικό μας σύστημα την πλατφόρμα Eclipse CDT for C/C++ Developers και την OpenCV τότε πρέπει να δημιουργήσουμε ένα νέο πρόγραμμα στο Eclipse όπου σε αυτό θα συνδέσουμε τις βιβλιοθήκες της OpenCV.

Ξεκινάμε με το File -> New -> C/C++ Project και στη συνέχεια δίνουμε ένα όνομα στην εφαρμογή μας. Για να προσθέσουμε κώδικα, δημιουργήσαμε ένα φάκελο μέσα στο Project με το όνομα src. Στον φάκελο αυτό θα δημιουργηθούν όλα τα αρχεία που απαρτίζουν τον κώδικα μας. Για να μπορέσουμε να χρησιμοποιήσουμε συναρτήσεις της OpenCV θα πρέπει να συνδέσουμε χειροκίνητα τα headers και τις βιβλιοθήκες της OpenCV. Τα βήματα που θα πρέπει να ακολουθήσουμε δίνονται πιο κάτω [12].

Πατάμε Project->Properties και στο C/C++ Build, επιλέγουμε τα Settings και στα δεξιά πατάμε στην καρτέλα Tool Settings. Στο GCC C++ Compiler, πηγαίνουμε στο Includes. Μέσα στο Include paths(-I) προσθέτουμε τη διαδρομή του φακέλου που εγκαταστήσαμε την OpenCV (βλ. Εικόνα 13).



Εικόνα 13 Προσθήκη φακέλου OpenCV στο Eclipse.

Για να βρούμε το φάκελο αυτό στη γραμμή εντολών - Terminal γράφουμε:

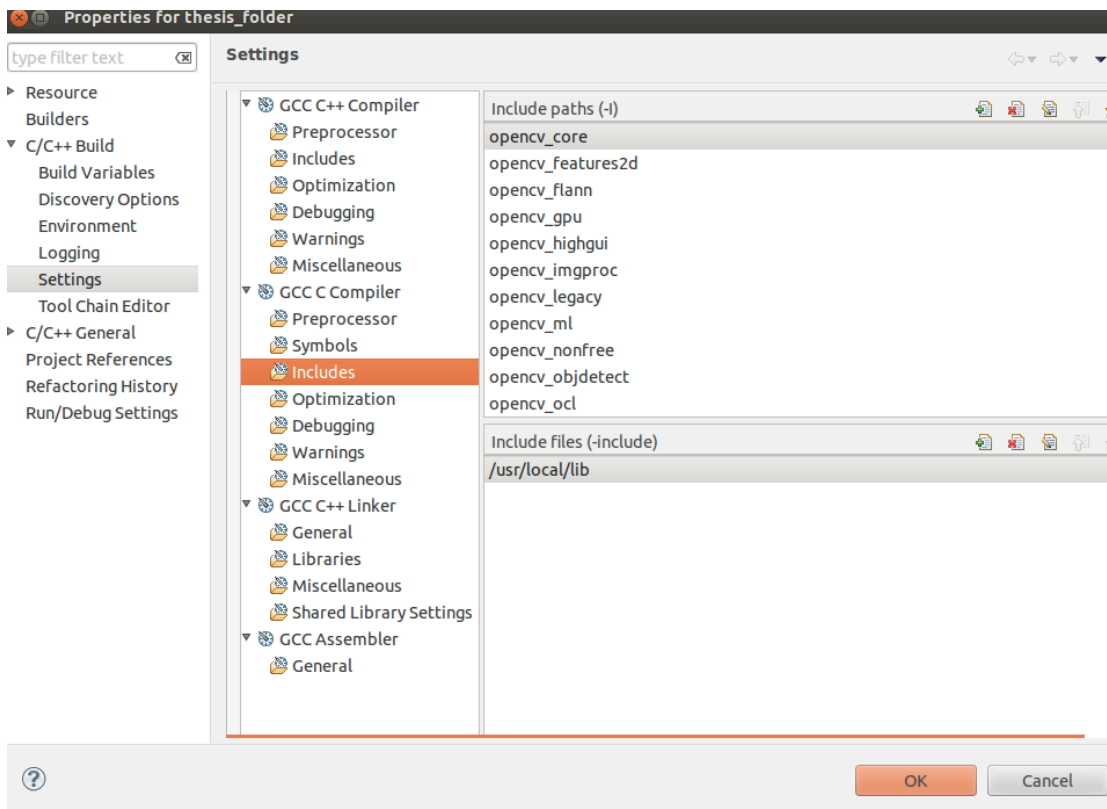
```
pkg-config --cflags opencv
```

Στη συνέχεια πηγαίνουμε στο GCC C++ Linker όπου αρχικά πρέπει να γράψουμε τη διαδρομή που οδηγεί στις βιβλιοθήκες της OpenCV μέσα στο πεδίο Library search path (-L).

Για να βρούμε τη διαδρομή των βιβλιοθηκών στο Terminal γράφουμε:

```
pkg-config --libs opencv
```

Στο πεδίο Libraries(-l) προσθέτουμε μια-μια τις βιβλιοθήκες που θα χρησιμοποιήσουμε στο σύστημα μας (βλ. Εικόνα 14). Οι πιο δημοφιλείς από αυτές είναι οι opencv_core, opencv_imgproc, opencv_highgui, opencv_ml, opencv_video, opencv_features2d, opencv_calib3d, opencv_objdetect, opencv_contrib, opencv_legacy και opencv_flann.



Εικόνα 14 Εισαγωγή βιβλιοθηκών και διαδρομής στο Eclipse.

Αν ακολουθήσουμε τα παραπάνω βήματα καταφέρνουμε να δημιουργήσουμε το αρχείο στο οποίο υπάρχουν συνδεδεμένες οι βιβλιοθήκες της OpenCV ώστε να μπορέσουμε να υλοποιήσουμε τον αλγόριθμό μας.

5

Μεθοδολογία Υλοποίησης της Εφαρμογής

Η κύρια λειτουργία του συστήματος που υλοποιήσαμε, είναι η επεξεργασία διαδοχικών εικόνων που λαμβάνουμε από την κάμερα σε πραγματικό χρόνο με τη βοήθεια της βιβλιοθήκης όρασης μηχανής OpenCV. Στόχος μας είναι η ανίχνευση και αναγνώριση κινούμενων αντικειμένων σε μια γραμμή παραγωγής. Ως αντικείμενο ταξινόμησης ορίσαμε ένα κόκκινο κουτάκι αναψυκτικού το οποίο, εκτός από την αναγνώριση του, το σύστημα πρέπει να το ταξινομεί στην γραμμή παραγωγής σαν αντικείμενο το οποίο πληροί ή όχι τις προδιαγραφές που θέσαμε.

5.1 Πρόγραμμα Εργασιών υλοποίησης

Με τη λήψη του θέματος της διπλωματικής το πρώτο βήμα που ακολουθήσαμε ήταν η ανάλυση των απαιτήσεων της εργασίας. Έγινε μελέτη για την κατανόηση του θέματος και των κριτηρίων που πρέπει να πληροί. Στη συνέχεια ακολούθησε η βιβλιογραφική έρευνα για τη μέθοδο κατάτμησης που θα επιλέγαμε και για τον τρόπο υλοποίησης. Επόμενο βήμα ήταν η μελέτη της γλώσσας προγραμματισμού C++ αλλά και η εκμάθηση της βιβλιοθήκης OpenCV. Μετά την εξοικείωση μας με τις C++ και OpenCV ξεκίνησε ο σχεδιασμός και η ανάπτυξη του αλγορίθμου. Στο στάδιο του σχεδιασμού καθορίσαμε την αναλυτική δομή του λογισμικού μαζί με τις απαιτήσεις της λειτουργία της διεπαφής του προγράμματος με το χρήστη. Στο στάδιο της υλοποίησης έγινε η ανάπτυξη του λογισμικού, η διαχείριση των αλλαγών και η αποσφαλμάτωση. Επόμενο στάδιο, ήταν η βελτιστοποίηση στον αλγόριθμο για να πληροί τις

αρχικές προδιαγραφές που θέσαμε. Τέλος είχαμε το στάδιο ελέγχου του συστήματος, σε αυτό ο έλεγχος βασίστηκε σε πολλαπλά σενάρια και με διάφορες προσεγγίσεις ελέγχθηκε το σύστημα και η απόδοση του, για την εξασφάλιση της υλοποίησης σύμφωνα με τις καθορισμένες προδιαγραφές.

5.2 Αντικείμενο Ταξινόμησης στην γραμμή μεταφοράς

Το αντικείμενο ταξινόμησης το οποίο θα αναγνωρίζει και θα ανιχνεύει η εφαρμογή ως αντικείμενο παραγωγής της γραμμής μεταφοράς προϊόντων είναι ένα κόκκινο σε χρώμα κουτάκι αναψυκτικού όπως φαίνετε στην εικόνα 15, με διαστάσεις:

Ύψος: ~11cm

Πλάτος: ~ 5 cm



Εικόνα 15 Το αντικείμενο που ελέγχεται στη γραμμή μεταφοράς.

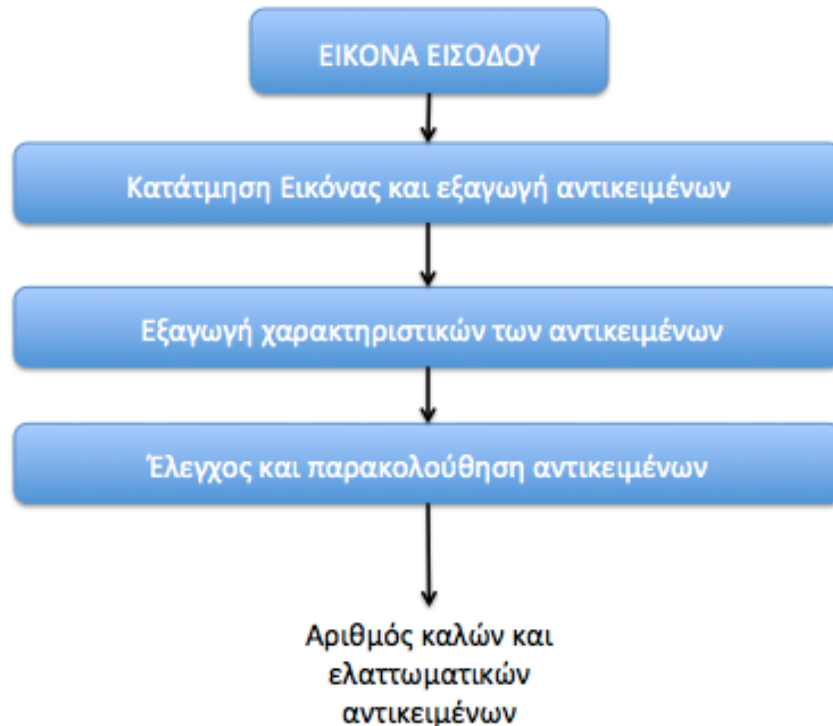
5.3 Επιλογή Μεθόδου Κατάτμησης Εικόνας

Στην υλοποίηση της εφαρμογής μας το πρώτο πρόβλημα που μας απασχόλησε είναι η επιλογή της μεθόδου που θα επιλεγεί για να εκτελεστεί η κατάτμηση εικόνας. Αρχικά μελετήσαμε διάφορους αλγορίθμους κατάτμησης που υποστηρίζει η OpenCV (βλ. Κεφάλαιο 2) και καταλήξαμε στη χρήση του αλγορίθμου που βασίζεται στο χρώμα. Οι αλγόριθμοι SIFT και SURF απορρίφθηκαν επειδή τα αντικείμενα που αναγνωρίζουν είναι συγκεκριμένα και απορρέουν από μια αρχική εικόνα, η οποία περιέχει το αντικείμενο. Η εφαρμογή που θέλαμε να δημιουργήσουμε πρέπει να βρίσκει εκτός από το

αντικείμενο αναφοράς, και άλλα αντικείμενα της ίδιας γραμμής παραγωγής τα οποία θα θεωρηθούν ελαττωματικά. Ο παραπάνω λόγος είναι η βασική αιτία επιλογής αλγόριθμου με προσέγγιση χρώματος.

5.4 Σχεδιασμός και Υλοποίηση της Εφαρμογής

Το σχήμα 4 παρουσιάζει τα στάδια σχεδίασης του συστήματος ανίχνευσης αντικειμένων σε διαδοχικές εικόνες.



Σχήμα 4 Σχεδιασμός συστήματος εφαρμογής ταξινόμησης αντικειμένων σε γραμμή μεταφοράς.

Το πρώτο στάδιο της εφαρμογής είναι η είσοδος εικόνας, που στην δική μας περίπτωση είναι είσοδος διαδοχικών εικόνων με τη χρήση μιας webcam. Για λόγους ευκολίας αρχικά η είσοδος γινόταν μέσω ενός έτοιμου αρχείου .avi και στη συνέχεια έγιναν οι κατάλληλες τροποποιήσεις για είσοδο από κάμερα.



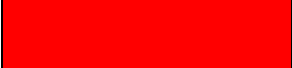













```
CvCapture* capture = cvCaptureFromCAM( -1 );
```

Η παραπάνω συνάρτηση της OpenCV χρησιμοποιήθηκε για είσοδο εικόνας από κάμερα. Το (-1) μας επιτρέπει αν έχουμε περισσότερες από μια κάμερες, να κάνουμε επιλογή αυτής που θέλουμε. Η συνάρτηση cvCaptureFromCAM φορτώνει τα δεδομένα από την κάμερα και επιστρέφει ένα δείκτη στην δομή CvCapture. Η δομή αυτή περιέχει όλα τα δεδομένα από το αρχείο που διαβάζεται.

Με την είσοδο των διαδοχικών εικόνων στο σύστημα αυτές αποθηκεύονται ως δομές IplImage, δηλαδή δομές στον βασικό πίνακα της OpenCV.

Το επόμενο βήμα στην υλοποίηση ήταν η αναγνώριση αντικειμένου και ο σχεδιασμός περιγράμματος γύρω από αυτό. Η ανίχνευση έγινε με χρήση χρώματος κόκκινου. Στον πίνακα 3, βλέπουμε τα βασικά χρώματα στις τιμές του RGB. Τα χρώματα κόκκινο, πράσινο και μπλε χρησιμοποιούν 8 bits το κάθε ένα και έχουν τιμές από 0 μέχρι 255. Αυτό συνεπάγεται ότι έχουμε $256*256*256=16777216$ πιθανά χρώματα.

Πίνακας 3 Χρωματικός Πίνακας RGB [15].

Χρώμα	Όνομα HTML/CSS	16-αδικός Κωδικός #RRGGBB	10-αδικός Κωδικός (R,G,B)
	Μαύρο	#000000	(0,0,0)
	Άσπρο	#FFFFFF	(255,255,255)
	Κόκκινο	#FF0000	(255,0,0)
	Λάιμ	#00FF00	(0,255,0)
	Μπλε	#0000FF	(0,0,255)
	Κίτρινο	#FFFF00	(255,255,0)
	Κυανό	#00FFFF	(0,255,255)
	Πορφυρό / Φούξια	#FF00FF	(255,0,255)
	Ασημί	#C0C0C0	(192,192,192)
	Γκρι	#808080	(128,128,128)
	Βυσσινί	#800000	(128,0,0)
	Λαδί	#808000	(128,128,0)
	Πράσινο	#008000	(0,128,0)
	Μωβ	#800080	(128,0,128)
	Γαλαζοπράσινο	#008080	(0,128,128)
	Σκούρο Μπλε	#000080	(0,0,128)

Οι τιμές που έχει το χρώμα κόκκινο στον πιο πάνω πίνακα είναι πολύ συγκεκριμένες και δεν μπορούμε να τις χρησιμοποιήσουμε ακριβώς έτσι, για το λόγο ότι στο αντικείμενο, αν και φαινομενικά υπάρχει ομοιόμορφο χρώμα, δημιουργούνται διάφορες διακυμάνσεις στις αποχρώσεις που οφείλονται στον φωτισμό, στις σκιές και στην οπτική γωνία. Στην εφαρμογή μας για να αποφύγουμε διάφορα προβλήματα όπως η μη ανίχνευση αντικειμένου, ορίσαμε σαν χρώμα κόκκινο μια εμβέλεια χρωμάτων για να υπάρχει ευελιξία στην ανίχνευση αντικειμένου. Οι ελάχιστες τιμές χρώματος R, G, B, είναι (κόκκινο 200, πράσινο 0, μπλε 0) και οι μέγιστες είναι (κόκκινο 255, πράσινο 150, μπλε 150). Όσα εικονοστοιχεία έχουν χρωματική ιδιότητα σε αυτή την ακτίνα τιμών θεωρούνται εικονοστοιχεία του αντικειμένου.

Το δεύτερο πρόβλημα αφορούσε την αναπαράσταση των χρωματικών χώρων που επιλέξαμε. Οι χρωματικοί χώροι που χρησιμοποιήσαμε για την ανίχνευση κινούμενου αντικειμένου ήταν ο κλασικός χώρος RGB και οι χώροι YUV – YCrCb.

```
cvCvtColor(img,YUV,CV_RGB2YCrCb);
```

Η πιο πάνω συνάρτηση που χρησιμοποιήθηκε μετατρέπει ένα χώρο σε άλλο δηλαδή τον RGB στον YCrCb. Στη συνέχεια της εφαρμογής με τη χρήση του `IplImage` και της `cvCreateImage` αποσυνθέτουμε την εικόνα σε διαφορετικά επίπεδα όπως θα δούμε και πιο κάτω στα αποτελέσματα τα οποία παίρνουμε από την εφαρμογή.

```
c = cvWaitKey(33);  
if( c == 27 ) break;
```

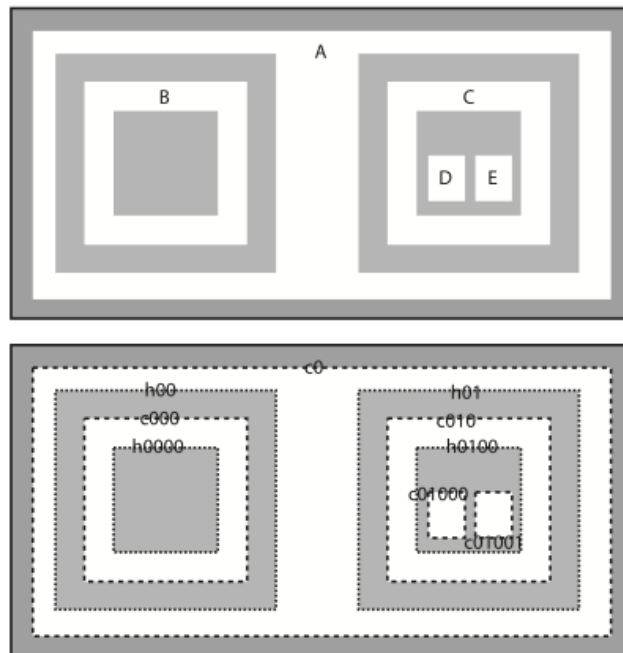
Η ανίχνευση αντικειμένων και οι έλεγχοι που γίνονται σε αυτά μπαίνουν σε έναν ατέρμονο βρόγχο ο οποίος σταματά να εκτελείται μόνο αν ο χρήστης πατήσει το πλήκτρο ESC (ASCII= 27) ή αν περάσει ένα χρονικό διάστημα αδράνειας (33 δευτερόλεπτα). Ο σκοπός αυτού του βρόγχου είναι να γίνονται οι έλεγχοι σε κάθε καρέ του βίντεο.

```
frame = cvQueryFrame( capture );
```

Μέσα στον ατέρμονο βρόγχο ξεκινάμε να διαβάζουμε τα δεδομένα που εισέρχονται από την κάμερα. Η `cvQueryFrame(capture)` παίρνει ως όρισμα ένα δείκτη σε μια δομή της `CvCapture`. Στη συνέχεια παίρνει το επόμενο καρέ στη μνήμη και επιστρέφεται ένας δείκτης στο καρέ εκείνο.

Το αντικείμενο ορίζεται ως η περιοχή στην οποία τα εικονοστοιχεία της έχουν χρώμα μέσα στην κλίμακα που δώσαμε πιο πάνω. Αν ένα εικονοστοιχείο και τα γειτονικά του έχουν την ιδιότητα αυτή, τότε γίνεται ομαδοποίησή τους με τεχνικές κατάτμησης με βάση την περιοχή. Για λόγους διεπαφής με το χρήστη αλλά και για να μπορούν να γίνουν στη συνέχεια οι έλεγχοι στις ομάδες εικονοστοιχείων που εμφανίζονται, προσθέσαμε περιγράμματα (`contours`)

γύρω από τα αντικείμενα. Η συνάρτηση που χρησιμοποιήσαμε για να βρούμε τα περιγράμματα είναι η `cvFindContours()`.



Εικόνα 16 Λειτουργία της `cvFindContours()` [14].

Η εικόνα 16 απεικονίζει τη λειτουργικότητα της συνάρτησης `cvFindContours()`. Στο πάνω μέρος αναπαριστάται θεωρητικά μια εικόνα η οποία θα περαστεί μέσα στη συνάρτηση. Οι περιοχές μας A,B,C,D,E έχουν λευκό χρώμα και το φόντο έχει γκρι. Η εικόνα στο κάτω μέρος αναπαριστά το αποτέλεσμα της συνάρτησης όταν εφαρμοστεί στην εικόνα εισόδου. Παρατηρούμε τα δύο είδη περιγραμμάτων που υποστηρίζονται. Το ένα είναι το εξωτερικό περίγραμμα με διακεκομμένες γραμμές και το άλλο είναι το εσωτερικό περίγραμμα με συνεχόμενες τελείες.

Στον κώδικα της εφαρμογής θέσαμε ως ένα από τα ορίσματα της συνάρτησης το `CV_PETR_EXTERNAL` το οποίο βρίσκει μόνο τα ακραία εξωτερικά περιγράμματα. Στην εικόνα 16 υπάρχει μόνο ένα τέτοιο περίγραμμα. Το όρισμα αυτό το βάλουμε γιατί δεν μας απασχολούν τα εσωτερικά στοιχεία του αντικειμένου, ο έλεγχος της γραμμής παραγωγής αφορά μόνο τη δομή του.

Στο σημείο αυτό θέσαμε περιγράμματα που οριοθετούν τις περιοχές εικονοστοιχείων με κοινό χρώμα βρίσκοντας έτσι το ζητούμενο αντικείμενο. Χάρη στο σχήμα του αντικειμένου, ορίσαμε τα περιγράμματα να έχουν το γεωμετρικό σχήμα ενός ορθογωνίου.

Βιβλιοθήκη HighGUI

Η βιβλιοθήκη HighGUI (High – level graphical user interface) περιέχει τις συναρτήσεις της OpenCV που επιτρέπουν την επικοινωνία του χρήστη με τον υπολογιστή και τις περιφερειακές συσκευές. Στην OpenCV η βιβλιοθήκη HighGUI χωρίζεται σε τρεις δομές. Η πρώτη είναι κυρίως για τις συσκευές και συγκεκριμένα τις κάμερες και την αλληλεπίδραση με αυτές. Η δεύτερη έχει να κάνει με το σύστημα αρχείων για ανάκτηση και αποθήκευση εικόνων. Η Τρίτη και τελευταία αφορά το σύστημα παραθύρων (window system). Η δομή αυτή περιέχει τις συναρτήσεις για άνοιγμα παραθύρων και εγγραφή δεδομένων σε αυτά.

Κατά την υλοποίηση του κώδικα της εφαρμογής μας χρησιμοποιήσαμε ευρέως συναρτήσεις και από τις τρεις δομές της βιβλιοθήκης της OpenCV HighGUI. Παραθέτουμε συνοπτικά κάποιες από αυτές τις συναρτήσεις και τη χρήση τους.

Το `cvCapture` που αναλύσαμε πιο πάνω είναι μία από τις συναρτήσεις της βιβλιοθήκης HighGUI.

Δημιουργούμε παράθυρο με το όνομα `image` και με το όρισμα `CV_WINDOW_AUTOSIZE` αλλάζει το μέγεθος του παραθύρου για να χωράει η εικόνα. Το όρισμα `c_str()` μετατρέπει ένα `string` σε `char`.

```
cvNamedWindow(image.c_str(),CV_WINDOW_AUTOSIZE);
```

Όταν τελειώσουμε την ανίχνευση αντικειμένων καταστρέφουμε το παράθυρο που ανοίξαμε, με τη συνάρτηση:

```
cvDestroyWindow(image,c_str());
```

Για να φορτώσουμε την εικόνα από το βίντεο στο παράθυρο που δημιουργήσαμε χρησιμοποιούμε τη συνάρτηση:

```
cvShowImage(imagename.c_str(), image);
```

Το πρώτο όρισμα καθορίζει ποια εικόνα θα βάλουμε στο παράθυρο και το δεύτερο στο δεύτερο όρισμα βάζουμε το όνομα του παραθύρου που θέλουμε να μπει η εικόνα.

Όταν σταματήσουμε να χρειαζόμαστε την εικόνα χρησιμοποιούμε τη συνάρτηση:

```
cvReleaseImage(&img);
```

Με τον τρόπο αυτό ελευθερώνουμε τη μνήμη που έχει εκχωρηθεί. Μέσα στη συνάρτηση πρέπει να βάλουμε ως όρισμα ένα δείκτη στο `IplImage` έτσι ώστε μετά την κλήση της ο δείκτης αυτός θα πάρει την τιμή `NULL`.

Άλλη συνάρτηση που χρειαστήκαμε για την υλοποίηση είναι η:

```
cvGetCaptureProperty(capture, CV_CAP_PROP_FRAME_WIDTH);
```

Με τη συνάρτηση αυτή μπορούμε να εξάγουμε πληροφορίες για τη πηγή του βίντεο. Οι πληροφορίες αυτές μπορεί να είναι για το ύψος, το πλάτος του και για το πόσα καρέ ανά δευτερόλεπτο παίρνουμε. Το όρισμα που δώσαμε στην παραπάνω συνάρτηση έχει να κάνει με το πόσα εικονοστοιχεία έχει το πλάτος της εικόνας που εξάγουμε από το βίντεο.

5.4.1 Έλεγχοι ταξινόμησης αντικειμένων

Σκοπός της παρούσας διπλωματικής εργασίας εκτός από την ανίχνευση αντικειμένων είναι και η ταξινόμησή τους. Με τον όρο ταξινόμηση αντικειμένων σε μια γραμμή παραγωγής εννοούμε την κατάταξη του προϊόντος που κατασκευάστηκε ως επιθυμητό ή ανεπιθύμητο. Η κατάταξη αυτή απορρέει από την επιθυμία του κατασκευαστή να έχει την επιλογή της διαλογής των προϊόντων καθώς αυτά βρίσκονται στην γραμμή παραγωγής, επιτυγχάνοντας έτσι μαζική παραγωγή γρήγορα και αξιόπιστα.

Έλεγχος Συνδεσιμότητας:

Ο πρώτος έλεγχος που κάνουμε είναι να σιγουρευτούμε πως το μέσο εισόδου εικόνας, που στην περίπτωση της εφαρμογής μας είναι η κάμερα είναι συνδεδεμένο με τον υπολογιστή. Αν δεν ανιχνευτεί κάμερα τότε μας εμφανίζει μήνυμα λάθους.

Αλγόριθμος Ελέγχου Συνδεσιμότητας

- 1 Επιλογή_Κάμερας = Αληθές
- 2 Είσοδος ← Εικόνα_Από_Κάμερα
- 3 Αν Όχι_Κάμερα
- 4 Τότε
- 5 Εμφάνισε Μήνυμα_Λάθους
- 6 Έξοδος_Προγράμματος
- 7 Τέλος_Αν

Έλεγχος Θορύβου:

Για να αποφύγουμε ανεπιθύμητα περιγράμματα από σκιάσεις ή από το γεγονός ότι το εύρος ζώνης χρώματος που βάλαμε είναι μεγάλο, ελέγχουμε το μέγεθος του κάθε περιγράμματος. Η ανίχνευση αντικειμένων με μέγεθος μικρότερο από 3000 εικονοστοιχεία οφείλεται σε θόρυβο και παραβλέπεται.

Αν ένα αντικείμενο είναι μεγαλύτερο από το κατώφλι θορύβου τότε σχηματίζουμε το περίγραμμά του και το υπολογίζουμε ως αντικείμενο που βρίσκεται πάνω στη γραμμή μεταφοράς.

Αλγόριθμος Έλεγχου Θορύβου:

```
1 Ελάχιστο_Εμβαδόν ← 3000
2 ...
3 Για Περίγραμμα=Αληθές και Περίγραμμα ← Επόμενο_Περίγραμμα Επανάλαβε
4     Αρχή
5         Εμβαδόν ← Εμβαδόν_Περιγράμματος
6         Αν Εμβαδόν > Ελάχιστο_Εμβαδόν
6             Τότε
7                 // Πράξεις
8                 Σχεδίασε_Περίγραμμα
9             Τέλος_Αν
10        Τέλος
```

Έλεγχος Ποιότητας:

Ο κυριότερος έλεγχος είναι αυτός της ποιότητας των αντικειμένων. Κάθε ένα τα αντικείμενα πρέπει να ελέγχονται και να κατηγοριοποιούνται σε αποδεκτό ή μη αποδεκτό αντικείμενο. Παίρνουμε τα δεδομένα από το κάθε εξωτερικό περίγραμμα των αντικειμένων που ανιχνεύονται και από αυτά βρίσκουμε το ύψος και το πλάτος των ορθογωνίων. Ως κριτήριο για να είναι ένα αντικείμενο αποδεκτό από τη γραμμή παραγωγής, ορίσαμε το λόγο ύψους προς πλάτος να είναι μεγαλύτερος από 1,5.

Αλγόριθμος Εύρεσης Λόγου:

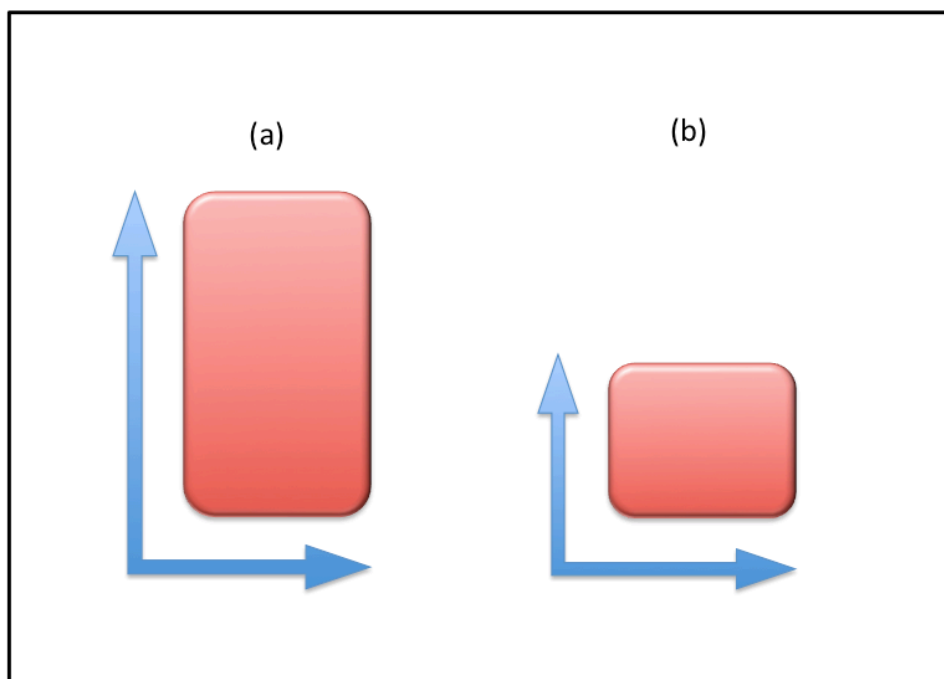
```
1 Για Περίγραμμα και Περίγραμμα ← Επόμενο_Περίγραμμα Επανάλαβε
2     Αρχή
3         Πάρε Υψος_Ορθογωνίου
4         Πάρε Πλάτος_Ορθογωνίου
5         ...
6         Λόγος ← Υψος_Ορθογωνίου / Πλάτος_Ορθογωνίου
7         ...
8     Τέλος
```

Οι πράξεις για την εύρεση του λόγου μπορούν να υλοποιηθούν με βάση τον παραπάνω ψευδοκώδικα. Για τον έλεγχο της ποιότητας παίρνουμε το λόγο των περιγραμμάτων και στην κύρια συνάρτηση κάνουμε τους ελέγχους:

Αλγόριθμος Ελέγχου Ποιότητας:

```
1 Αν Λόγος > 1,5
2   Τότε
3     Εκτύπωσε Αντικείμενο_Αποδεκτό
4     Μετρητής_Αποδεκτών ← Μετρητής_Αποδεκτών + 1
5   Αλλιώς
6     Εκτύπωσε Αντικείμενο_Μη_Αποδεκτό
7     Μετρητής_Μη_Αποδεκτών ← Μετρητής_Μη_Αποδεκτών + 1
8   Εκτύπωσε Κουδούνι
8 Τέλος Αν
```

Στο σχήμα 5 βλέπουμε ένα υποθετικό σενάριο ενός στιγμιότυπου, όπου το (a) είναι το αποδεκτό αντικείμενο με λόγο μεγαλύτερο από το 1,5 και το (b) είναι ένα μη αποδεκτό αντικείμενο με λόγο μικρότερο από το 1,5. Το αντικείμενο (b) πρέπει να απομακρυνθεί από την γραμμή παραγωγής.



Σχήμα 5 (a) Αντικείμενο βάση των προδιαγραφών (b) Μη αποδεκτό αντικείμενο.

Ο έλεγχος για το αν ένα αντικείμενο είναι επιθυμητό ή όχι ξεκινά από τη στιγμή που το αντικείμενο μπαίνει στην οθόνη και συνεχίζεται μέχρι και τη στιγμή που αρχίζει να βγαίνει από την οθόνη. Η σκοπιμότητα που έχει ο συνεχής αυτός έλεγχος είναι η αντικατάσταση των μηχανισμών αφαίρεσης ελαττωματικών προϊόντων σε μια πραγματική γραμμή μεταφοράς με ένα σήμα ήχου (beep). Κάθε φορά που ακούμε συνεχόμενο ήχο, το πρόγραμμά μας ειδοποιεί πως υπάρχει ένα ή περισσότερα ελαττωματικά αντικείμενα στην οθόνη.

Έλεγχος Ποσότητας

Το τελικό αποτέλεσμα που θέλουμε να μας δίνει η εφαρμογή είναι η συνολική ποσότητα αντικειμένων που ελέγχθηκαν, πόσα από αυτά τα αντικείμενα ήταν αποδεκτά και πόσα έχουν απορριφθεί. Υπάρχουν στον κώδικα τρεις μετρητές, ένας για τα προϊόντα που πληρούν τις προδιαγραφές, ένας για τα ελαττωματικά προϊόντα και ένας για το σύνολο των προϊόντων που εξετάστηκαν από το πρόγραμμα.

5.4.2 Προβλήματα που προέκυψαν και η αντιμετώπιση τους

Σε αυτή την υπό-ενότητα παραθέτουμε μερικά από τα πιο σημαντικά προβλήματα που χρειάστηκε να αντιμετωπίσουμε και τον τρόπο αντιμετώπισης τους, κατά τη διάρκεια της υλοποίησης της εφαρμογής.

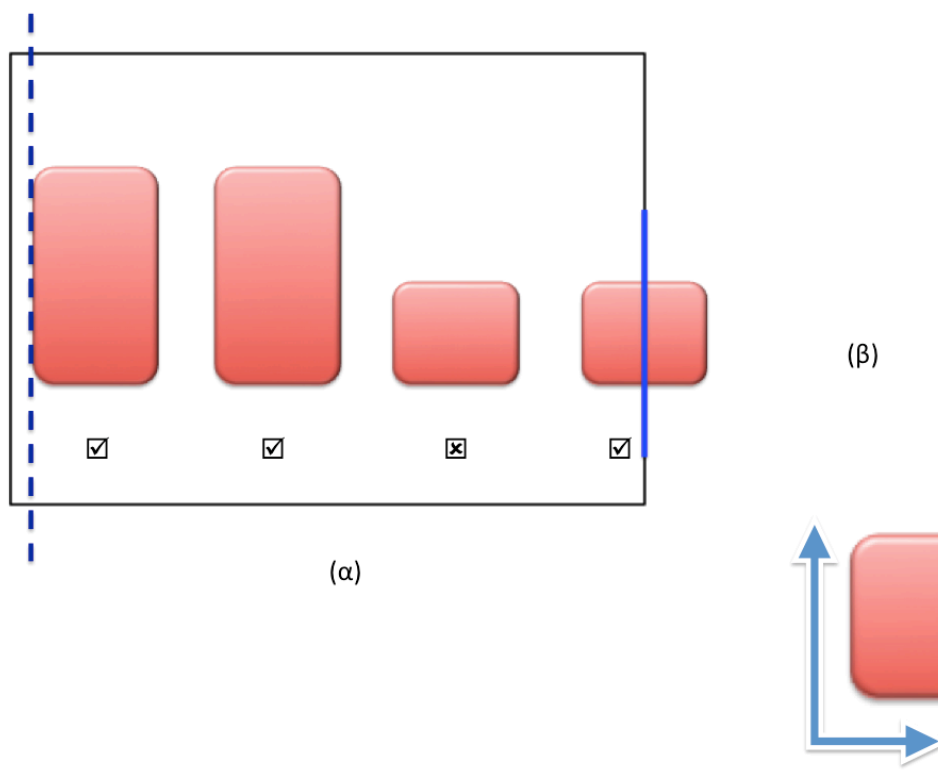
Πριν ξεκινήσουμε την ανάπτυξη του κώδικα, γνωρίζαμε ότι οι έλεγχοι θα πρέπει να γίνουν σε μία δική μας γραμμή μεταφοράς λόγω του ότι δεν είχαμε τη δυνατότητα να δοκιμάσουμε το σύστημα μας σε μία βιομηχανική γραμμή. Για τη λύση αυτού του προβλήματος θέσαμε ως μη στατικά, τόσο το φόντο όσο και τα προϊόντα.

Το πρώτο πρόβλημα που αντιμετωπίσαμε κατά την συγγραφή του κώδικα, ήταν η ανίχνευση περιγραμμάτων που δεν αντιστοιχούσαν σε κάποιο από τα ζητούμενα αντικείμενα. Εκτός από τα περιγράμματα των αντικειμένων ανιχνεύονταν και περιγράμματα που οφείλονταν σε σκιές ή στην διαβάθμιση των χρωμάτων. Το πρόβλημα αυτό αντιμετωπίστηκε ως θόρυβος (βλ. Υπό-ενότητα 5.4.1). Ορίσαμε τα περιγράμματα που είναι κατά πολύ μικρότερα σε εμβαδό από αυτά που ψάχνουμε να μη λαμβάνονται υπόψη με αποτέλεσμα να λυθεί και το πρόβλημα αυτό.

Το επόμενο πρόβλημα που συναντήσαμε ήταν ένα σύνθετο ζήτημα και αφορούσε την ανίχνευση ελαττωματικών αντικειμένων αλλά και την καταμέτρηση όλων των αντικειμένων της γραμμής μεταφοράς. Ο αρχικός μας έλεγχος ταξινόμησης ξεκινούσε από την είσοδο του αντικειμένου στην οθόνη μέχρι και την έξοδο του. Κάθε φορά που δημιουργούνταν ένα περίγραμμα καταγράφαμε τα στοιχεία του και το ταξινομούσαμε ως επιθυμητό ή ανεπιθύμητο. Το πρόβλημα που προέκυψε ήταν ότι με αυτό τον τρόπο μετρούσαμε τα αντικείμενα που υπήρχαν στο κάθε καρέ. Δηλαδή αν σε ένα στιγμιότυπο ανιχνεύονταν δύο αντικείμενα και στο επόμενο αυτά τα αντικείμενα υπήρχαν ακόμα στην οθόνη, ο μετρητής κατέγραφε έλεγχο τεσσάρων αντικειμένων, ενώ στην πραγματικότητα τα αντικείμενα ήταν δύο.

Το δεύτερο σκέλος του προβλήματος σχετίζεται με τον έλεγχο ποιότητας των αντικειμένων. Όταν ένα αντικείμενο εισέρχεται στην οθόνη με βάση τα κριτήρια ελέγχου είναι πάντα αποδεκτό. Η κατανόηση αυτού του ζητήματος μπορεί να γίνει ευκολότερα αν παρατηρήσουμε το υποθετικό παράδειγμα στο σχήμα 6. Στο 6 (α) βλέπουμε να εισέρχεται στην οθόνη ένα ελαττωματικό αντικείμενο αλλά για τα πρώτα στιγμιότυπα της εισόδου του το αντικείμενο

θεωρείται ως αποδεκτό. Όπως παρουσιάζεται και στο σχήμα 6 (β) ο λόγος ύψους προς πλάτος του αντικειμένου αυτού κατά την είσοδο του είναι μεγαλύτερος από το 1,5.



Σχήμα 6 (α) Αντικείμενα που ελέγχονται στο σημείο με διακεκομμένη γραμμή (β) Αντικείμενο μη αποδεκτό που κατά την είσοδο του φαίνεται αποδεκτό.

Η λύση που δώσαμε και στα δύο αυτά υπό-πρόβλημα ήταν κοινή. Θέσαμε τον έλεγχο για την καταμέτρηση των αντικειμένων να γίνεται λίγο πριν τη στιγμή που το αντικείμενο βγει από την οθόνη, όπως βλέπουμε στο σχήμα 6 (α) με διακεκομμένη γραμμή. Στο σημείο εκείνο, το αντικείμενο καταγράφεται ότι πέρασε από τη γραμμή μεταφοράς και αυξάνεται ο μετρητής καλών ή ελαττωματικών αντικειμένων, ανάλογα με τη φύση του. Τα αντικείμενα για να θεωρηθούν ότι πέρασαν από επιθεώρηση χρειάζεται να βγαίνουν από την οθόνη. Με τον τρόπο αυτό λύνεται το σύνθετο πρόβλημα που αντιμετωπίσαμε και τα αποτελέσματα που παίρνουμε είναι τα σωστά.

Τέλος, με τη παραπάνω λύση προέκυψε άλλο ένα πρόβλημα. Ο ήχος που ειδοποιεί ότι υπάρχει ελαττωματικό αντικείμενο, ακουγόταν μόνο σε ένα στιγμιότυπο πράγμα που μπορούσε να μη γίνει αντιληπτό και δεν έδινε περιθώρια αντίδρασης για αφαίρεση του προϊόντος από τη γραμμή μεταφοράς. Βάλαμε έναν άλλο έλεγχο για το κάθε περίγραμμα στην οθόνη, ξεχωριστό από τον έλεγχο του μετρητή έτσι, αν ένα αντικείμενο είναι αποδεκτό δεν γίνεται κάποια ενέργεια, αν δεν είναι τότε ενεργοποιείται το κουδούνι (BELL).

5.5 Εκτέλεση εφαρμογής από γραμμή εντολών (terminal)

Για την εκτέλεση του προγράμματος που υλοποιήσαμε ανοίγουμε τη γραμμή εντολών (terminal) και πληκτρολογούμε την εντολή « ls -l », το αποτέλεσμα που θα πάρουμε είναι μια λίστα με τα αρχεία που περιέχει ο τρέχον κατάλογος. Με την εντολή cd μπορούμε να μετακινηθούμε από ένα κατάλογο σε έναν άλλο. Με αυτό τον τρόπο πηγαίνουμε στον κατάλογο που περιέχει τον πηγαίο κώδικα και μετά στο αρχείο Debug. Στο σημείο αυτό, πληκτρολογούμε την εντολή:

```
ldd Inspection_Control
```

Η εντολή αυτή εκτυπώνει τις κοινόχρηστες βιβλιοθήκες που απαιτεί το πρόγραμμα για να εκτελεστεί.

Οι βιβλιοθήκες που χρειαζόμαστε από την OpenCV και πρέπει να προσθέσουμε στο φάκελο με τα αρχεία του κώδικα μας (βλ. Υπό – ενότητα 4.2.5) είναι οι :

```
libopencv_core.so.2.4 libopencv_highgui.so.2.4 libopencv_imgproc.so.2.4
```

Οι βιβλιοθήκες που χρησιμοποιούμε από το Ubuntu – Linux/GNU είναι οι παρακάτω:

```
linux-vdso.so.1 libstdc++.so.6 libgcc_s.so.1 libc.so.6 libpthread.so.0 libz.so.1  
librt.so.1 libGL.so.1 libtbb.so.2 libm.so.6 libjpeg.so.8 libpng12.so.0 libtiff.so  
libjasper.so.1 libl1m1mf.so.6 libHalf.so.6 libgstreamer-0.10.so.0 libgobject2.0.so.0  
libglib-2.0.so.0 libgstapp-0.10.so.0 libgstvideo-0.10.so.0 libdc1394.so.22  
libv4l1.so.0 libavcodec.so.53 libavformat.so.53 libavutil.so.51 libswscale.so.2  
libQtOpenGL.so.4 libQtGui.so.4 libQtTest.so.4 libQtCore.so.4 libglapi.so.0  
libXext.so.6 libXdamage.so.1 libXfixes.so.3 libX11-xcb.so.1 libX11.so.6 libxcb-  
glx.so.0 libxcb-dri2.so.0 libxcb.so.1 libXxf86vm.so.1 libdrm.so.2 libdl.so.2  
liblmath.so.6 liblex.so.6 libllmThread.so.6 libgmodule-2.0.so.0 libxml2.so.2  
libffi.so.6 libpcre.so.3 libgstbase-0.10.so.0 liborc-0.4.so.0 libraw1394.so.11  
libusb-1.0.so.0 libv4l2.so.0 libvpx.so.1 libvorbisenc.so.2 libvorbis.so.0  
libtheoraenc.so.1 libtheoradec.so.1 libspeex.so.1 libschroedinger-1.0.so.0  
libgsm.so.1 libva.so.1 libbz2.so.1.0 libfreetype.so.6 libXrender.so.1  
libfontconfig.so.1 libaudio.so.2 libSM.so.6 libICE.so.6 libXi.so.6 libXau.so.6  
libXdmpc.so.6 libv4lconvert.so.0 libogg.so.0 libexpat.so.1 libXt.so.6 libuuid.so.1
```

Για την εκτέλεση του κώδικα σε αυτό το σημείο θα πρέπει από την πλατφόρμα Eclipse να έχουμε αποθηκεύσει και μεταγλωττίσει ολόκληρο το αρχείο με τα περιεχόμενα της εφαρμογής. Τέλος πληκτρολογούμε την εντολή:

```
./Inspection_Control
```

Το αποτέλεσμα της παραπάνω εντολής είναι η εκτέλεση του κώδικα μας. Αξίζει να σημειώσουμε ότι στην εφαρμογή χρειαζόμαστε τον ήχο (beep), αλλά τα νέα τερματικά έχουν ως προεπιλογή την οπτικοποίηση του ήχου (visible beep), το

οποίο είναι η εμφάνιση ενός μηνύματος κειμένου BELL. Χρησιμοποιείται κυρίως σε συνδέσεις απομακρυσμένων διακομιστών διαδικτύου, όπου η δημιουργία του beep δεν είναι δυνατό να ακουστεί. Αυτό που έχουμε να κάνουμε πριν την εκτέλεση στο τερματικό απαραίτητα, είναι να βάλουμε την εντολή « xset b on » που θα ενεργοποιήσει τον ήχο.

6

Αποτελέσματα Εφαρμογής

Στο κεφάλαιο αυτό κάνουμε μια σύντομη περιγραφή των αποτελεσμάτων της εφαρμογής που υλοποιήθηκε στα πλαίσια της διπλωματικής εργασίας. Αναλύουμε τα αποτελέσματα και εξάγουμε συμπεράσματα για το ποσοστό επιτυχίας του αρχικού στόχου.

6.1 Σενάρια Ελέγχου Δοκιμών

Τα σενάρια δοκιμών των αποτελεσμάτων της εφαρμογής έχουν ως στόχο να αναδείξουν ότι το λογισμικό που αναπτύξαμε ικανοποιεί τις καθορισμένες απαιτήσεις και να εντοπιστούν τυχόν ατέλειες. Οι προδιαγραφές που θέσαμε για το σύστημα είναι εξής:

- Είσοδος εικόνας από κάμερα.
- Μη στατικό φόντο.
- Μετατροπή στο χώρο YCbCr.
- Το αντικείμενο της γραμμής παραγωγής έχει σχήμα ορθογωνίου με ύψος περίπου 11 εκατοστά, πλάτος περίπου 5 εκατοστά και χρώμα κόκκινο.
- Ανίχνευση του αντικειμένου.
- Σχεδιασμός ορθογώνιου περιγράμματος γύρω από αυτό.
- Διαχωρισμός ελαττωματικών και μη ελαττωματικών αντικειμένων.
- Προειδοποίηση με τη μορφή ήχου στην περίπτωση ανίχνευσης ελαττωματικού αντικειμένου.

- Συνολική καταμέτρηση αντικειμένων που πέρασαν από τη γραμμή μεταφοράς.
- Ξεχωριστή καταμέτρηση μη ελαττωματικών και ελαττωματικών προϊόντων.
- Έξοδος από την εφαρμογή σε περίπτωση που ο χρήστης πατήσει το πλήκτρο ESC.
- Έξοδος από τη εφαρμογή αν περάσει ένα χρονικό διάστημα χωρίς ανίχνευση αντικειμένου.

Οι δοκιμές χωρίστηκαν σε δύο κατηγορίες, τις δοκιμές συστατικών στοιχείων όπου, ελέγχονται μεμονωμένα τα συστατικά στοιχεία του προγράμματος και τις δοκιμές τελικής έκδοσης. Στη δεύτερη κατηγορία δοκιμών γίνεται έλεγχος του συστήματος ως ομάδα συστατικών στοιχείων που πληροί τις προδιαγραφές.

Τα σενάρια ελέγχου δοκιμών έγιναν για να πιστοποιηθεί η ομαλότητα του συστήματος και να αποδειχθεί ότι κάθε απαίτηση έχει ελεγχθεί.

6.1.1 Αποτελέσματα Δοκιμών Συστατικών Στοιχείων

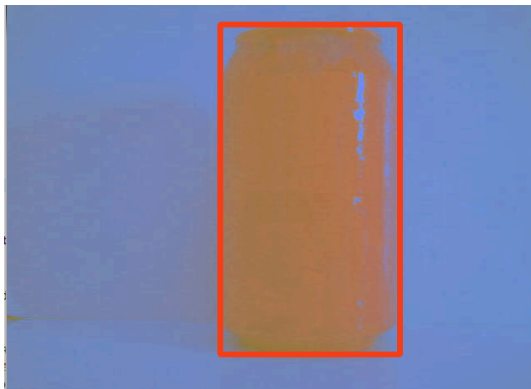
Κατά τη διάρκεια της ανάπτυξης του πηγαίου κώδικα, η αποσφαλμάτωση έγινε με επιτυχία όπως και η επικύρωση της ορθότητας των μονάδων και των βιβλιοθηκών.

Το πρώτο σενάριο δοκιμής αφορούσε τη συνδεσιμότητα της κάμερας. Εκτελέσαμε την εφαρμογή χωρίς να έχουμε συνδέσει την κάμερα και μας εμφανίστηκε το ακόλουθο μήνυμα λάθους (βλ. Εικόνα 17). Στη συνέχεια, συνδέσαμε τη κάμερα στον υπολογιστή, κάναμε επανεκτέλεση και άνοιξε το παράθυρο με τη εικόνα εισόδου από την κάμερα, επικυρώνοντας έτσι την ορθότητα λειτουργίας του ελέγχου συνδεσιμότητας.

ERROR: Please connect your camera!

Εικόνα 17 Μήνυμα λάθους.

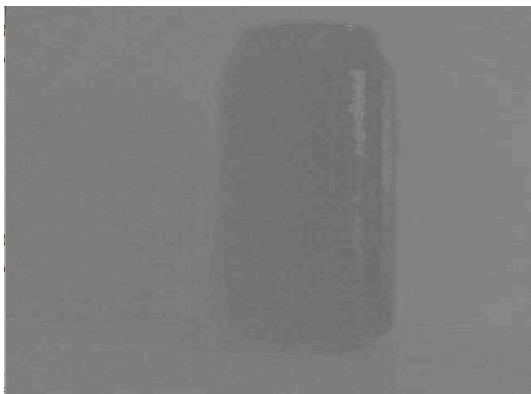
Η επόμενη δοκιμή αφορούσε την ανίχνευση αντικειμένου και τη σωστή μετατροπή της εικόνας εισόδου στο χρωματικό χώρο YCbCr. Η συνιστώσα με τη μεγαλύτερη σημασία για την ανίχνευση αντικειμένων είναι η Y αλλά για λόγους ελέγχου κατάτμησης ορίσαμε στον πηγαίο κώδικα το άνοιγμα παραθύρων για τη κάθε μία από αυτές.



(α)



(β)



(γ)

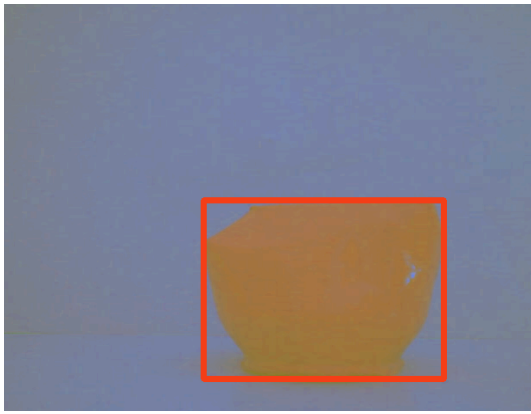


(δ)

Εικόνα 18 (α) Εικόνα εισόδου αντικειμένου με βάση τις προδιαγραφές (β) συνιστώσα Y (γ) συνιστώσα Cb (δ) συνιστώσα Cr.

Η εικόνα 18 (α) είναι η εικόνα εισόδου από τη κάμερα η οποία περιέχει ένα από τα ζητούμενα αντικείμενα. Η εικόνα 18 (β) αναπαριστά τη συνιστώσα Y του χρωματικού χώρου YCbCr. Στις εικόνες 18 (γ) και 18 (δ) αντίστοιχα παρουσιάζονται οι συνιστώσες Cb και Cr.

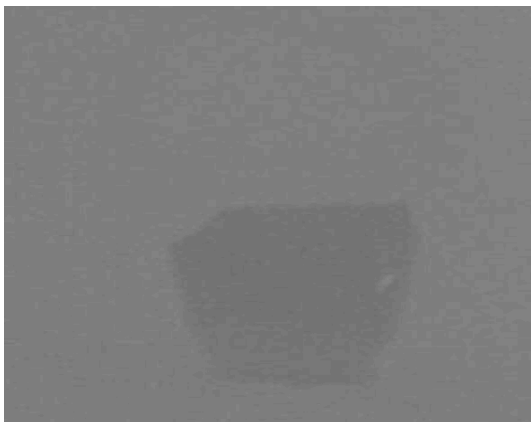
Στο σημείο αυτό έγιναν και οι δοκιμές ανίχνευσης αντικειμένων και εξαγωγής περιγραμμάτων. Στην εικόνα 18 (α) παρατηρούμε το σωστό σχεδιασμό περιγράμματος γύρω από ένα μη ελαττωματικό αντικείμενο. Το περίγραμμα που σχηματίζεται έχει το γεωμετρικό σχήμα ενός ορθογωνίου, όπως ορίζουν οι προδιαγραφές και το ύψος είναι μεγαλύτερο από το πλάτος. Οι ίδιες δοκιμές έγιναν και για την ανίχνευση ελαττωματικού αντικειμένου όπως βλέπουμε στην εικόνα 19.



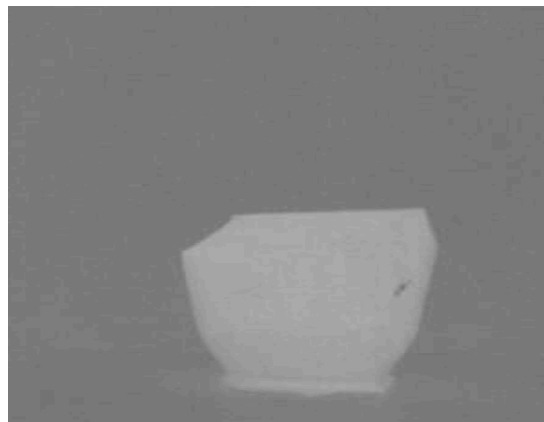
(α)



(β)



(γ)



(δ)

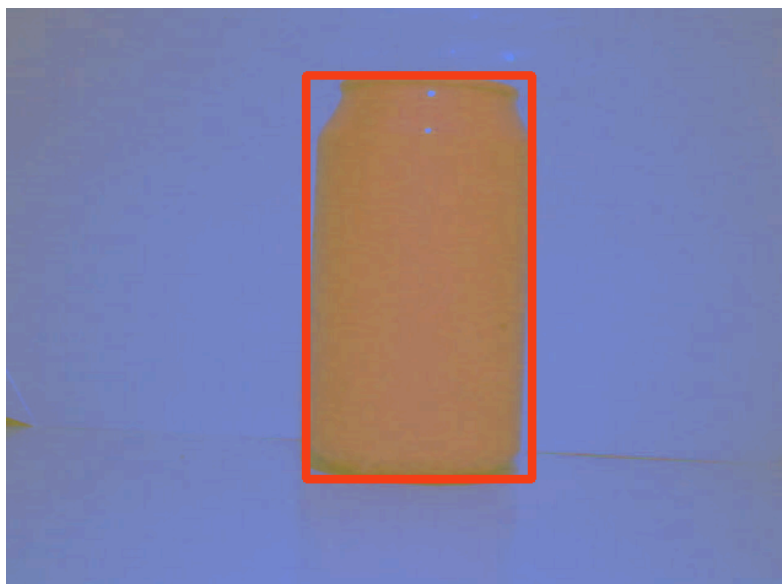
Εικόνα 19 Εικόνα εισόδου ελαττωματικού αντικειμένου (β) συνιστώσα Y (γ) συνιστώσα Cb (δ) συνιστώσα Cr.

Ο σχεδιασμός περιγράμματος κατά την ανίχνευση ελαττωματικού αντικειμένου όπως παρατηρούμε και από την εικόνα 19 (α) είναι ο αναμενόμενος τόσο στο ύψος του ορθογωνίου που σχηματίστηκε, όσο και στο πλάτος.

Η ανίχνευση έγινε με τη χρήση μιας κάμερας που τη μετακινούσαμε, για να κάνουμε τον έλεγχο ποιότητας σε μια ερασιτεχνική γραμμή μεταφοράς. Το φόντο όπως και η κάμερα άλλαζε σε κάθε καρέ για το λόγο αυτό, έγιναν δοκιμές σε θορυβώδες φόντο. Το αποτέλεσμα των δοκιμών ήταν επιτυχημένο μόνο στην περίπτωση που δεν υπήρχαν αντικείμενα σε αποχρώσεις του κόκκινου στο φόντο. Ορίσαμε έτσι, ως μία από τις προδιαγραφές της γραμμής μεταφοράς τη χρήση λευκού φόντου, χωρίς λοιπά αντικείμενα σε αυτό.

Η επόμενη μεμονωμένη δοκιμή ήταν για να επιβεβαιώσουμε την παρακολούθηση των αντικειμένων και την ταξινόμηση τους σε αντικείμενα που πληρούν ή όχι τις προδιαγραφές.

Στην εικόνα 20 έχουμε ένα σωστό αντικείμενο. Το αποτέλεσμα που παίρνουμε ως έξοδο από την εφαρμογή (βλ. Εικόνα 21) είναι η επιβεβαίωση της ορθής λειτουργίας του ελέγχου ποιότητας του αντικειμένου. Ότι δηλαδή, παρακολουθεί ένα αντικείμενο το οποίο περνά τον έλεγχο ποιότητας.



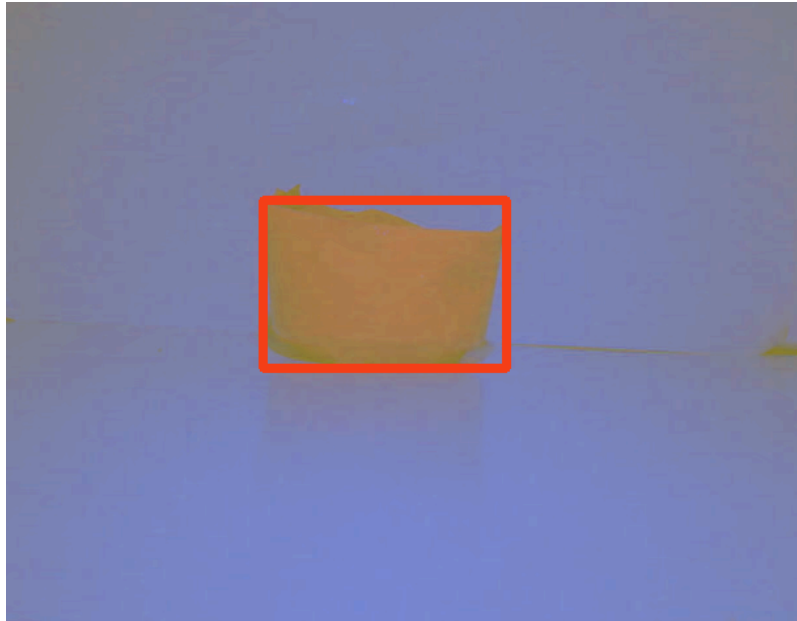
Εικόνα 20 Ανίχνευση αντικειμένου βάση των προδιαγραφών και σχεδιασμός περιγράμματος γύρω από αυτό.

Inspection Passed

Tracking 1 RED product

Εικόνα 21 Μήνυμα εξόδου.

Στην εικόνα 22 κάνουμε ανίχνευση και παρακολούθηση ενός ελαττωματικού προϊόντος που πρέπει να αφαιρεθεί από τη γραμμή μεταφοράς. Όπως έχουμε προαναφέρει για πρακτικούς λόγους οι μηχανισμοί αφαίρεσης αντικαθίστονται από ένα ήχο (beep) τον οποίο και ακούμε κατά τη διάρκεια της παρακολούθησης. Στην εικόνα 23 βλέπουμε τα αποτελέσματα από την ανίχνευση του αντικειμένου. Το πρόγραμμα δηλώνει την παρακολούθηση του αντικειμένου και εμφανίζει μήνυμα ότι το συγκεκριμένο αντικείμενο δεν περνά τον έλεγχο ποιότητας. Κάτω από το μήνυμα ύπαρξης αποτυχίας στην παραγωγή, βλέπουμε ένα σύμβολο που περιέχει τον αριθμό 007. Ο αριθμός αυτός είναι η ASCII τιμή του ήχου.



Εικόνα 22 Ανίχνευση ελαττωματικού αντικειμένου και σχεδιασμός περιγράμματος γύρω από αυτό.

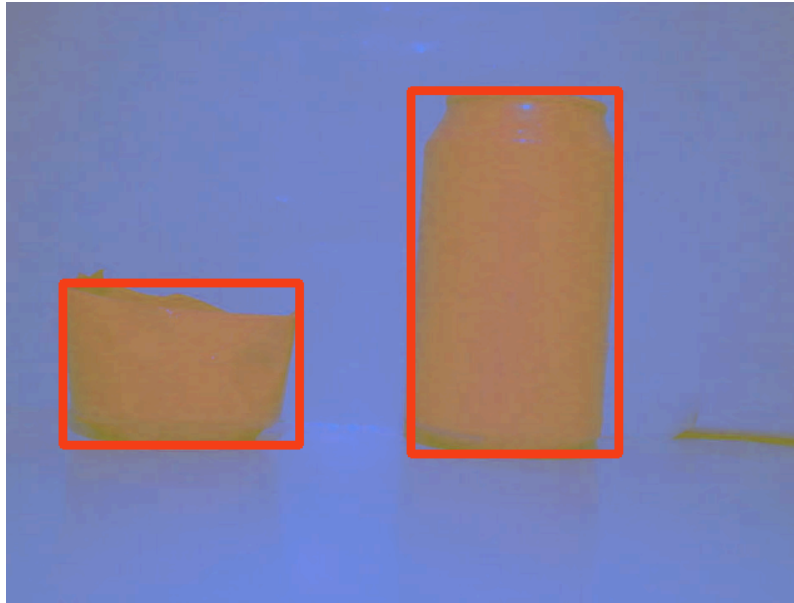
Inspection Failed



Tracking 1 RED product

Εικόνα 23 Μήνυμα Εξόδου.

Σε αυτό το στάδιο έγιναν δοκιμές παρακολούθησης και ανίχνευσης πολλαπλών αντικειμένων. Στην εικόνα παρατηρούμε την ανίχνευση δύο αντικειμένων. Το πρώτο αντικείμενο παρακολούθησης, δεν πληροί τις προδιαγραφές και το δεύτερο τις πληροί.



Εικόνα 24 Ανίχνευση δυο αντικειμένων και εξαγωγή περιγραμμάτων γύρω από αυτά.

Inspection Failed



Inspection Passed

Tracking 2 RED product

Εικόνα 25 Μήνυμα Εξόδου.

Το μήνυμα που εμφανίζεται στην περίπτωση που έχουμε ένα ελαττωματικό και ένα μη ελαττωματικό αντικείμενο μπροστά στην κάμερα, είναι αυτό της εικόνας 25. Επιβεβαιώνουμε έτσι την παρακολούθηση δυο προϊόντων και τον διαχωρισμό τους σε ελαττωματικό και μη ελαττωματικό.

Ο τελευταίος μεμονωμένος έλεγχος είναι της εξόδου από το σύστημα. Ελέγχουμε ότι με το να πατήσουμε το πλήκτρο ESC σταματά το πρόγραμμα και κλείνουν όλα τα ανοιχτά παράθυρα. Επίσης, ελέγχουμε την έξοδο από το σύστημα σε περίπτωση αδράνειας. Αν για 33 δευτερόλεπτα δεν ανιχνευτεί αντικείμενο έχουμε και πάλι έξοδο.

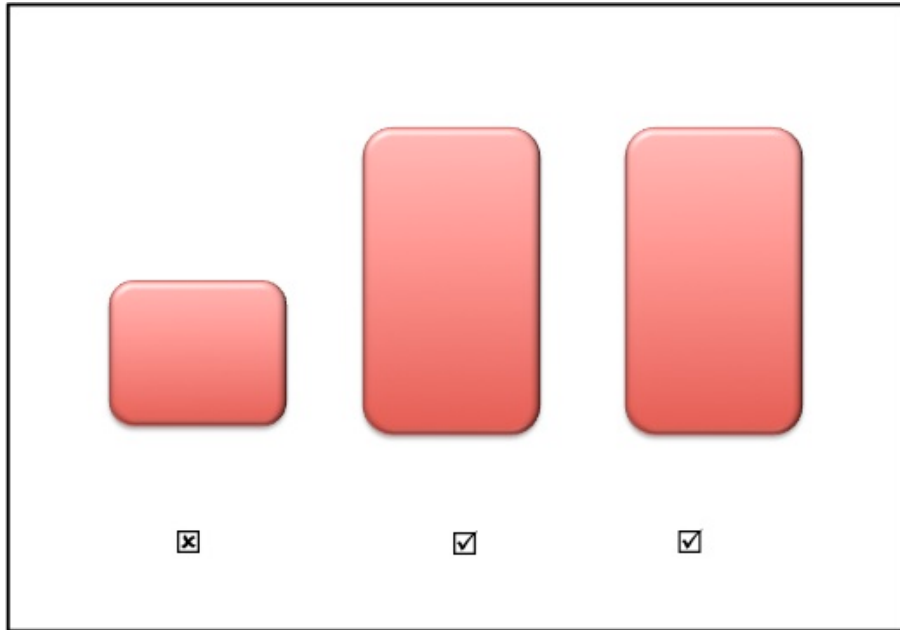
6.1.2 Αποτελέσματα Δοκιμών Τελικής Έκδοσης

Με τη βοήθεια των παραπάνω σεναρίων δοκιμών, ελέγξαμε τη λειτουργία του κάθε συστατικού που απαρτίζει τον κώδικα μας. Σε αυτή την υπό – ενότητα οι δοκιμές γίνονται στην τελική έκδοση της εφαρμογής μας. Θέλουμε να πιστοποιήσουμε ότι τα αποτελέσματα αντικατοπτρίζουν τις προδιαγραφές του συστήματος.

Μετά από πολλαπλά σενάρια δοκιμών στη γραμμή μεταφοράς αποφασίστηκε ότι η βέλτιστη απόσταση που πρέπει να έχει το ένα αντικείμενο από το άλλο είναι 8 εκατοστά και η κάμερα πρέπει να έχει απόσταση περίπου 30 εκατοστά από τη γραμμή. Σε διαφορετική περίπτωση έχουμε πάλι ορθή ανίχνευση αντικειμένων, αλλά σε κάποια από αυτά τα σενάρια παρουσιάστηκε αστάθεια στο σύστημα μας.

Άλλη μια παρατήρηση που πρέπει να σημειωθεί αφορά την κάμερα. Για το λόγο ότι δεν είχαμε στη διάθεση μας κάποια βιομηχανική γραμμή παραγωγής, όπου η κάμερα είναι τοποθετημένη σε ένα σταθερό σημείο, κινούσαμε εμείς την κάμερα. Για να έχουμε βέλτιστη απόδοση η κάμερα πρέπει να κινείται με σταθερό ρυθμό χωρίς να γίνονται οποιεσδήποτε παλινδρομήσεις, πράγμα που ισχύει και για τη γραμμή μεταφοράς ενός βιομηχανικού περιβάλλοντος. Σε διαφορετική περίπτωση τα αποτελέσματα μας θα είναι λανθασμένα και αυτό οφείλεται στο γεγονός ότι η καταμέτρηση γίνεται τη στιγμή που το αντικείμενο βγαίνει από την οθόνη.

Στο σημείο αυτό θα εξετάσουμε και θα παρουσιάσουμε μερικά από τα σενάρια που δημιουργήσαμε στη γραμμή παραγωγής. Στο πρώτο σενάριο έχουμε μια γραμμή μεταφοράς τριών αντικειμένων. Τα αντικείμενα εισέρχονται στην οθόνη ελέγχονται και καταμετρούνται. Το πρώτο αντικείμενο που παράγεται είναι ελαττωματικό και τα άλλα δυο είναι σύμφωνα με τις προδιαγραφές. Το υποθετικό αυτό σενάριο παρουσιάζεται στο σχήμα 7.



Σχήμα 7 Σενάριο ύπαρξης τριών αντικειμένων στη γραμμή μεταφοράς.

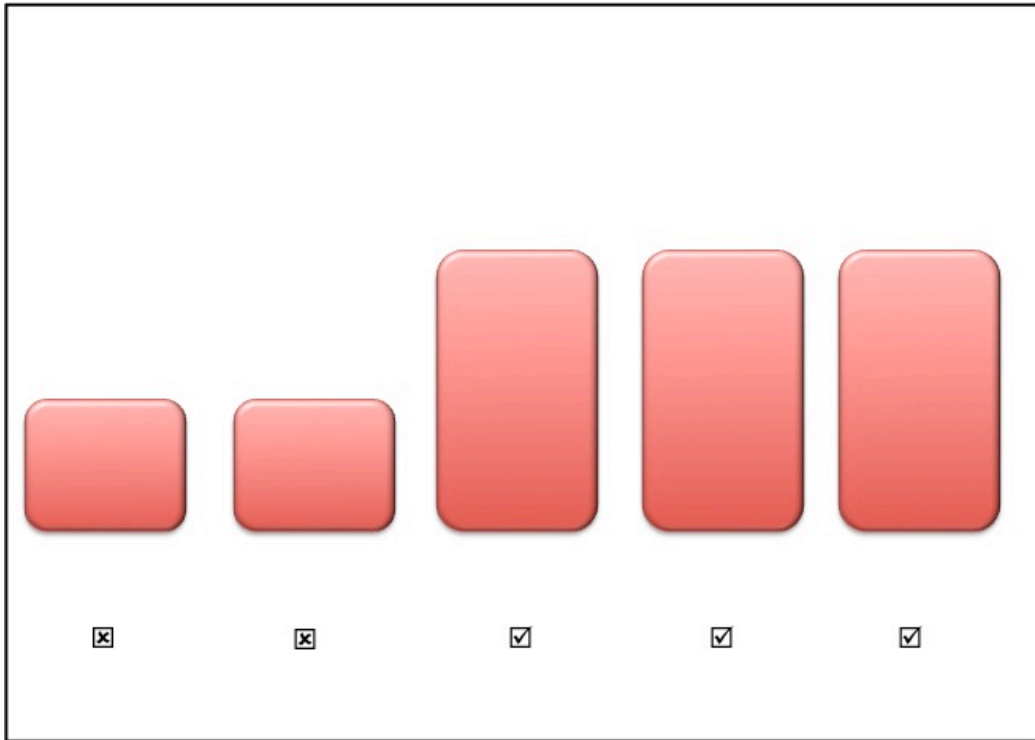
Μετά το πέρας της ταξινόμησης, η γραμμή μεταφοράς κλείνει και εξάγεται το τελικό αποτέλεσμα της καταμέτρησης, όπως φαίνεται στην εικόνα 26. Επιβεβαιώνουμε ότι πλέον όλα τα αντικείμενα έχουν περάσει και έχουμε ως αποτέλεσμα, δυο καλά και ένα ελαττωματικό αντικείμενο, σε μια γραμμή παραγωγής τριών αντικειμένων.

Tracking 0 RED product

Total of 2 good products.
Total of 1 corrupted products.
In a production line of 3 products.

Εικόνα 26 Μήνυμα Εξόδου.

Το τελευταίο σενάριο δοκιμών παρουσιάζεται στο σχήμα 8. Η γραμμή περιέχει πέντε αντικείμενα, τρία βάση των προδιαγραφών και δύο ελαττωματικά. Η σειρά με την οποία τοποθετήθηκαν τα αντικείμενα είναι αυτή του σχήματος 8.



Σχήμα 8 Σενάριο ύπαρξης πέντε αντικειμένων στη γραμμή μεταφοράς.

Το αποτέλεσμα που εμφανίζεται κατά την καταμέτρηση του πρώτου αντικειμένου είναι (βλ. Εικόνα 27):

```

-----
Tracking 2 RED product
-----
Total of 0 good products.
Total of 1 corrupted products.
In a production line of 1 products.

```

Εικόνα 27 Μήνυμα για ταξινόμηση πρώτου αντικειμένου.

Το αποτέλεσμα μετά την καταμέτρηση και του δεύτερου αντικειμένου είναι (βλ. Εικόνα 28) :

Tracking 1 RED product

Total of 0 good products.
Total of 2 corrupted products.
In a production line of 2 products.

Εικόνα 28 Μήνυμα για ταξινόμηση και καταμέτρηση δύο αντικειμένων.

Στη συνέχεια το αντικείμενο που εισέρχεται παύει να είναι ελαττωματικό και παίρνουμε ως έξοδο (βλ. Εικόνα 29):

Tracking 1 RED product

Total of 1 good products.
Total of 2 corrupted products.
In a production line of 3 products.

Εικόνα 29 Μήνυμα για ταξινόμηση και καταμέτρηση τριών αντικειμένων.

Τέλος, καταμετρούνται και τα πέντε αντικείμενα και έτσι το πρόγραμμα εμφανίζει το συνολικό αριθμό ταξινόμησης (βλ. Εικόνα 30).

Tracking 0 RED product

Total of 3 good products.
Total of 2 corrupted products.
In a production line of 5 products.

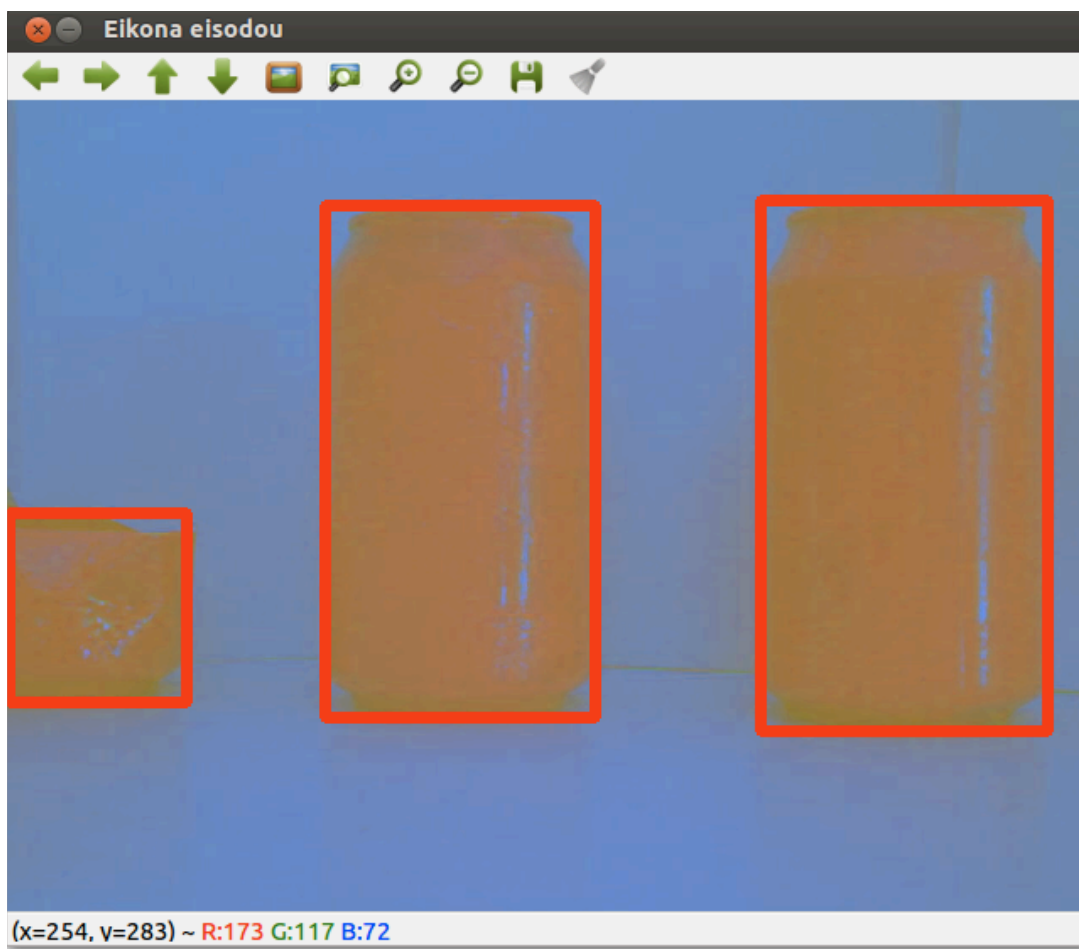
Εικόνα 30 Μήνυμα για ταξινόμηση και καταμέτρηση όλων των αντικειμένων της γραμμής μεταφοράς.

6.2 Λειτουργίες Qt * GUI

Για τις διεπαφές του χρήστη με την εφαρμογή χρησιμοποιήσαμε τις νέες λειτουργίες που παρέχονται μέσω της OpenCV. Οι λειτουργίες αυτές εμφανίζονται στην εικόνα 31. Το παράθυρο που ανοίξαμε με τη χρήση των λειτουργιών του Qt και της βιβλιοθήκης HighGUI έχει το όνομα «Eikona eisodou». Το παράθυρο περιέχει μια γραμμή κατάστασης (statusbar), μια γραμμή εργαλείων (toolbar), καθώς και έναν πίνακα ελέγχου (control panel).

Η γραμμή εργαλείων βρίσκεται κάτω από το όνομα του παραθύρου. Πάνω σε αυτή υπάρχουν κουμπιά με τα οποία ο χρήστης μπορεί να κάνει διάφορες λειτουργίες όπως η αποθήκευση της εικόνας, η μεγέθυνση και η σμίκρυνση της.

Η γραμμή κατάστασης βρίσκεται στην κάτω πλευρά του παραθύρου. Οι παράμετροι x και y μας δίνουν τις συντεταγμένες στην οθόνη στο σημείο που βρίσκεται το ποντίκι του χρήστη. Για το ίδιο σημείο αναγράφονται και οι αριθμοί χρώματος της κάθε συνιστώσας του RGB. Στην εικόνα 31 το ποντίκι μας βρίσκεται στο σημείο $x=254$, $y=283$ και οι τιμές RGB στο σημείο αυτό είναι Red=173, Green=117, Blue=72.



Εικόνα 31 Παράθυρο εικόνας εισόδου με λειτουργίες QT.

Με τη χρήση των τιμών RGB από τη γραμμή κατάστασης, μπορούμε να εξάγουμε με ακρίβεια τις μέγιστες και ελάχιστες τιμές χρωμάτων για το αντικείμενο που ανιχνεύουμε.

6.3 Ορίσματα γραμμής εντολών

Μια σημαντική δυνατότητα που προσφέρει η εφαρμογή μας είναι η είσοδος ορισμάτων από το τερματικό. Με απλά λόγια μπορούμε να ορίσουμε ανίχνευση αντικειμένων διαφορετικού χρώματος από το προκαθορισμένο κόκκινο, μέσω της γραμμής εντολών. Αν ο χρήστης δεν ορίσει χρώμα τότε η ανίχνευση γίνεται με βάση τις προδιαγραφές, διαφορετικά η εφαρμογή ανιχνεύει το ανάλογο χρώμα που ζήτησε ο χρήστης.

Τα ορίσματα που χρειάζεται να δοθούν για να επωφεληθεί κάποιος από αυτή τη δυνατότητα είναι επτά. Πρώτο όρισμα είναι το όνομα της εφαρμογής. Τα επόμενα τρία ορίσματα είναι οι ελάχιστες τιμές RGB αντίστοιχα για το χρώμα που θέλουμε. Οι τελευταίες τρεις είναι οι μέγιστες τιμές RGB του ίδιου χρώματος.

Μπορούμε να χρησιμοποιήσουμε τις Λειτουργίες Qt * GUI που αναλύσαμε στην παραπάνω υπό - ενότητα για να βρούμε με ακρίβεια τις τιμές χρώματος που έχει το αντικείμενο που θέλουμε να ανιχνευτεί. Ένα παράδειγμα είναι να έχουμε το ίδιο αντικείμενο παρακολούθησης αλλά το χρώμα του να είναι μπλε. Αυτό που έχουμε να κάνουμε είναι να ακολουθήσουμε τα βήματα που δόθηκαν στην υπό-ενότητα 5.5 για την εκτέλεση της εφαρμογής από τη γραμμή εντολών και στην τελευταία εντολή να προσθέσουμε τις ελάχιστες και μέγιστες τιμές του μπλε όπως:

```
./Inspection_Control 0 150 50 100 255 200
```

Οι τιμές αυτές αντιστοιχούν σε ελάχιστο χρώμα (R,G,B) → (0, 150, 50) και μέγιστο (R,G,B) → (100, 255, 200).

Το αποτέλεσμα που θα πάρουμε είναι η ανίχνευση αντικειμένων μπλε χρώματος και όχι κόκκινου.

7

Συμπεράσματα και Μελλοντικές Προεκτάσεις

Στο κεφάλαιο αυτό κλείνουμε τη διπλωματική μας εργασία τόσο σε θεωρητικό όσο και στο επίπεδο της υλοποίησης. Παραθέτουμε τα συμπεράσματα στα οποία καταλήξαμε βάση των αποτελεσμάτων της εφαρμογής και προτείνουμε ορισμένες μελλοντικές προεκτάσεις για να αποκτήσει η εφαρμογή μια πιο προηγμένη μορφή.

7.1 Συμπεράσματα

Αντικείμενο της διπλωματικής εργασίας αποτέλεσε η ανάπτυξη και υλοποίηση ενός αλγορίθμου για ανίχνευση και ταξινόμηση αντικειμένων σε μια γραμμή μεταφοράς, η κίνηση των οποίων καταγράφεται σε μια ακολουθία εικόνων πραγματικού χρόνου. Μελετήσαμε διάφορες μεθόδους κατάτμησης εικόνας και καταλήξαμε στην κατάτμηση με βάση το χρώμα. Η επιλογή μας αυτή αποδείχθηκε σωστή, αφού επιτύχαμε την ανίχνευση του προκαθορισμένου αντικειμένου. Η παρακολούθηση αντικειμένων έγινε με επιτυχία σε διαφορετικά καρέ και με τη χρήση των περιγραμμάτων, τα οποία επεξεργαστήκαμε, εξαγάγαμε ταξινομημένα προϊόντα.

Εν συντομία, η εφαρμογή μας όπως μπορούμε να παρατηρήσουμε και από τα αποτελέσματα κατατάσσεται ως αρκετά υποσχόμενη στον τομέα της όρασης μηχανής. Η ταχύτητα επεξεργασίας και εξαγωγής αποτελεσμάτων είναι πολύ

ικανοποιητική. Η εφαρμογή που υλοποιήθηκε έχει ευελιξία και προσαρμοστικότητα σε νέα δεδομένα αφού μπορεί να ανιχνεύσει αντικείμενα διαφορετικών χρωμάτων. Με μικρές αλλαγές στον πηγαίο κώδικα μπορεί να εντοπίσει αντικείμενα διαφορετικού γεωμετρικού σχήματος από το προκαθορισμένο, σχηματίζοντας το αντίστοιχο περίγραμμα.

Η γραμμή μεταφοράς που χρησιμοποιήσαμε ήταν ερασιτεχνική και η απόσταση της κάμερας από αυτή δεν ήταν προκαθορισμένη, με αποτέλεσμα να χρειάζεται χειροκίνητη μεταφορά της συσκευής λήψης διαδοχικών εικόνων, για να γίνεται η παρακολούθηση αντικειμένων. Με γνώμονα το πρόβλημα αυτό, καταφέραμε να υλοποιήσουμε την εφαρμογή μας με τη δυνατότητα της ανίχνευσης και ταξινόμησης των αντικειμένων ανεξάρτητα από την κλίμακα και το εμβαδό που έχουν στην οθόνη.

Η ενασχόληση μας με τη βιβλιοθήκη όρασης μηχανής OpenCV μας οδήγησε στο συμπέρασμα ότι η βιβλιοθήκη αυτή περιέχει ένα τεράστιο πλήθος συναρτήσεων με απεριόριστες δυνατότητας επεξεργασίας εικόνας, στατικής και μη, κάνοντας την ανίχνευση και παρακολούθηση αντικειμένων μια διαδικασία με μεγάλο βαθμό αξιοπιστίας. Χάρη στη βιβλιοθήκη OpenCV καταφέραμε να υλοποιήσουμε ένα σύστημα που μπορεί να εφαρμοστεί ως έχει, για ανίχνευση αντικειμένων σε μια μικρή βιομηχανία για έλεγχο παραγωγής. Ο χρήστης της εφαρμογής χωρίς να έχει εξειδικευμένες γνώσεις, μπορεί να αξιοποιήσει πλήρως τις δυνατότητες του συστήματος που αναπτύξαμε για έλεγχο ποιότητας στην παραγωγή καταλήγοντας σε μια γραμμή που δεν μεταφέρει ελαττωματικά προϊόντα. Συμπεραίνουμε, λοιπόν, ότι η όραση μηχανής με τη βοήθεια της βιβλιοθήκης OpenCV αποτελεί ένα πανίσχυρο εργαλείο στον τομέα της βιομηχανίας για αξιόπιστο έλεγχο ποιότητας και ποσότητας των προϊόντων.

7.2 Μελλοντικές Προεκτάσεις

Το σύστημα που αναπτύξαμε βρίσκεται σε πρωταρχικό στάδιο και κάποιες από τις πιθανές μελλοντικές προεκτάσεις και βελτιώσεις της εφαρμογής που αναπτύξαμε παρουσιάζονται σε αυτή την υπό-ενότητα, με αποτέλεσμα να γίνει πιο φιλικό προς τον χρήστη και να μεγιστοποιηθούν οι δυνατότητες επεξεργασίας.

Η εφαρμογή θα πρέπει να δοκιμαστεί σε πραγματικές βιομηχανικές συνθήκες, για να ελέγξουμε την ταχύτητα αναγνώρισης και ταξινόμησης των αντικειμένων, κάτι που με τα μέσα που είχαμε ήταν αδύνατο. Στην περίπτωση που η ταχύτητα δεν είναι ικανοποιητική, θα πρέπει να αναζητηθεί κάποια προγραμματιστική λύση. Αν εφαρμογή χρησιμοποιηθεί σε μια πραγματική γραμμή μεταφοράς, η κάμερα θα είναι στατική και η απόσταση της από τη γραμμή μεταφοράς θα είναι προκαθορισμένη. Τα αντικείμενα θα έχουν σταθερή απόσταση μεταξύ τους. Το συμπέρασμα που εξάγουμε από τα δεδομένα αυτά είναι ότι μελλοντικά μπορούμε να διαφοροποιήσουμε τον πηγαίο κώδικα στον τρόπο ταξινόμησης, βρίσκοντας έτσι περισσότερα στοιχεία για το κάθε

αντικείμενο, όπως για παράδειγμα το εμβαδόν του. Από τα στοιχεία αυτά τα αντικείμενα μπορούν να καταταχθούν ως ελαττωματικά αν έχουν ακόμα και την πιο μικρή ατέλεια.

Για να αυξήσουμε την αξιοπιστία της ανίχνευσης των προϊόντων, ως μελλοντική προέκταση προτείνουμε την εισαγωγή νέων φίλτρων στην εικόνα εισόδου. Όσον αφορά την αξιοπιστία της ταξινόμησης συνιστούμε τη δυνατότητα ύπαρξης μια ποικιλίας περιγραμμάτων αποτελούμενα από διάφορα γεωμετρικά σχήματα. Τα περιγράμματα αυτά θα μπορούν να σχεδιαστούν εσωτερικά και εξωτερικά του αντικειμένου κάνοντας δυνατή την ανίχνευση πολυπλοκότερων προϊόντων.

Ο ήχος που ειδοποιεί για την παρουσία ελαττωματικού αντικειμένου στη γραμμή μεταφοράς, μπορεί εύκολα να αντικατασταθεί με μια ενέργεια που κάνει ο υπολογιστής. Ένα ενδεικτικό παράδειγμα είναι η σύνδεση του με ένα έμβολο εκτόνωσης, που θα ωθεί βιαίως εκτός της γραμμής παραγωγής, το ελαττωματικό αντικείμενο. Η σύνδεση με το περιφερειακό μπορεί να γίνει σε μία έξοδο του υπολογιστή, όπως είναι η θύρα USB ή μια παράλληλη. Η τροποποίηση στον κώδικα που απαιτείται είναι ασήμαντη.

Αν και ο πηγαίος κώδικας αναπτύχθηκε σε linux x86, εντούτοις μπορεί εύκολα να μεταφερθεί σε μια ενσωματωμένη συσκευή beagleboard με το να γίνει compile για linux ARM, αφού όλες οι βιβλιοθήκες είναι διαθέσιμες και σε αυτή την αρχιτεκτονική. Αυτό το ενσωματωμένο σύστημα θα έχει το όφελος του ότι η αναγνώριση και ταξινόμηση αντικειμένων να μπορεί να γίνει ανεξάρτητα από τον υπολογιστή, προσφέροντας φορητότητα στην εφαρμογή.

ΑΠΟΔΟΣΗ ΑΓΓΛΙΚΩΝ ΟΡΩΝ ΣΤΗΝ ΕΛΛΗΝΙΚΗ ΓΛΩΣΣΑ

Luminance	Φωτεινότητα
Pixel	Εικονοστοιχείο
RGB ή red – green – blue	Κόκκινο – Πράσινο – Μπλε
normalizedRGB	Κανονικοποιημένος RGB
Frame	Καρέ
Computer vision	Όραση μηχανής
Object Detection	Ανίχνευση αντικειμένου
Image segmentation	Κατάτμηση εικόνας
Tracking	Παρακολούθηση
Contours	Περιγράμματα
SIFT	Μετασχηματισμός κλιμακωτά αναλλοίωτων χαρακτηριστικών
SURF	Επιταχυνθέντα ισχυρά χαρακτηριστικά
FAST	Εξαγωγή χαρακτηριστικών από γρήγορο τμήμα δοκιμής
Laplacian of Gaussian	Λαπλασιανή της Γκαουσιανής
Hessian Matrix	Πίνακας Έσσιαν
Background subtraction	Αφαίρεση φόντου
Temporal differencing	Χρονική διαφορά
Optical flow	Οπτική ροή
Photoreceptor cell	Φωτοαισθητήρες
Rods	Ραβδία
Cones	Κωνία
Terminal	Τερματικό / γραμμή εντολών
Histogram Thresholding	Ιστόγραμμα κατωφλίου
Fuzzy Techniques	Τεχνικές ασαφούς λογικής
Script	Σενάριο
Cluster	Σύμπλεγμα
Compiler	Μεταγλωττιστής
Workspace	Περιβάλλον εργασίας
Function overloading	Υπερφόρτωση τελεστών

Βιβλιογραφία

- [1] E. Rosten and T. Drummond, *Machine learning for high speed corner detection*, in 9th European Conference on Computer Vision, vol. 1, 2006, pp. 430–443.
- [2] Χρωματικοί χώροι – Color space, http://en.wikipedia.org/wiki/Color_space, Ιούνιος 2014 .
- [3] T.B. Moeslund, *Introduction to Video and Image Processing*, 1 Undergraduate Topics in Computer Science, DOI 10.1007/978-1-4471-2503-7_1, © Springer-Verlag London Limited 2012
- [4] Normalized RGB, <http://www.aishack.in/2010/01/normalized-rgb/>, Ιούνιος 2014.
- [5] YUV, <http://en.wikipedia.org/wiki/YUV>, Ιούνιος 2014.
- [6] Mohamed D Almadhoun, *Improving and measuring color edge detection algorithm in RGB color space*, The Society of Digital Information and Wireless Communications, 2013 (ISSN: 2225-658X) pp. 20.
- [7] T. L. Huntsberger, C. L. Jacobs, R. L. Cannon, *Iterative Fuzzy Image Segmentation*, Pattern Recognition, 1998.
- [8] *Future Computer Vision Tools to Aid Medical Research and Healthcare*, http://www.jacobsschool.ucsd.edu/news_events/releases/release.sfe?id=1056, Ιούνιος 2014.
- [9] A. Saleh, A.Sagheer, *Automatic Skin Detection in Color Images A Color Space Study*, Center for Artificial Intelligence and Robotics, 2012.
- [10] S. A. Shafer, *Using Color to Separate Reflection Components*, Color Research and Application 10(4):210-218, 1985.
- [11] G. Healey, *Using Color for Geometry – insensitive Segmentation*. Optical Society of America, vol. 22, No. 1, 920 -937, 1989.
- [12] OpenCv, <http://opencv.org/>, Ιούνιος 2014.
- [13] C++ Programming, <http://www.cplusplus.com>, Ιούνιος 2014.
- [14] Gary Bradski, Adrian Kaehle, *Learning OpenCV*, 2008, Εκδόσεις O'Reilly Media, ISBN: 978-0-596-51613-0, pp. 13.

[15] RGB Color Codes Chart, http://www.rapidtables.com/web/RGB_Color.htm, Ιούνιος 2013.

[16] Command – line interface, http://en.wikipedia.org/wiki/Command-line_interface, Ιούνιος 2013.

[17] Margaret A. Ellis, Bjarne Stroustrup, *The Annotated C++ Reference Manual*, Published by Addison – Wesley Professional, ISBN-13: 978-0201514599, 1990.

[18] Lowe G. David, *Object recognition from local scale – invariant features*, Proceedings of the International Conference on Computer Vision 2. pp. 1150–1157, 1999.

[19] Herbert Bay, Tinne Tuytelaars, Luc Van Gool, *SURF: Speeded Up Robust Features*, Computer Vision and Image Understanding, 2006.

[20] Ubuntu, <http://ubuntu-gr.org>, Ιούνιος 2014.

[21] Java, <http://www.oracle.com/technetwork/java/index.html>, Ιούνιος 2014.

[22] OpenCV, <https://help.ubuntu.com/community/OpenCV>, Ιούνιος 2014.

[23] Lin ZHANG, *Introduction to Computer Vision*, Lecture 1, Tongji University, 2011.

[24] Mohammad F. Hashmi, Avinash G. Keskar *Computer – Vision Based Visual Inspection and Crack Detection of Railroad Tracks*, Recent Advances in Electrical and Computer Engineering, Μάρτιος 2014.

[25] K.M. Sajjad, *Automatic License Plate Recognition using Python and OpenCV*, Department of Computer Science and Engineering M.E.S. College of Engineering, Kuttippuram, Kerala.

[26] Ευαγγελία Α. Κατσαούνου, *Παρακολούθηση και Αναγνώριση Ανθρώπινων Χειρονομιών με Τεχνικές Υπολογιστικής Όρασης*, Διπλωματική Εργασία, Εθνικό Μετσόβιο Πολυτεχνείο, Οκτώβριος 2006.

[27] Πρωτονοτάριος Ιωάννης, *Ανάπτυξη Συστήματος Ενσωματωμένων Αισθητήρων για Ασύρματη Μετάδοση Εικόνας και Δεδομένων*, Διπλωματική Εργασία, Πανεπιστήμιο Πατρών, Φεβρουάριος 2011.