



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ

ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

Διπλωματική Εργασία

της

Γούλα – Δημητρίου Μιχαήλας

**Σχεδιασμός και υλοποίηση ενός έξυπνου παιχνιδιού για
τη συναισθηματική παρακολούθηση και το γονικό έλεγχο
ενός παιδιού**

Επιβλέπων Καθηγητής: Δρ. Μηνάς Δασυγένης

Εργαστήριο Ψηφιακών Συστημάτων και Αρχιτεκτονικής Υπολογιστών

Κοζάνη | Μάρτιος 2018



University of Western Macedonia
FACULTY OF ENGINEERING
DEPARTMENT OF INFORMATICS & TELECOMMUNICATIONS
ENGINEERING

Thesis

by

Michaila Goula - Dimitriou

**Design and implementation of an intelligent toy to
provide emotional observation and paternal monitoring
of a child**

Supervisor: Dr. Minas Dasygenis

Laboratory of Digital Systems & Computer Architecture

Kozani | March 2018

Περιεχόμενα

Κατάλογος Πινάκων	9
Κατάλογος Εικόνων	11
Περίληψη	13
Abstract	15
Δήλωση πνευματικών δικαιωμάτων	17
Ευχαριστίες	19
Κεφάλαιο 1: Εισαγωγή	21
1.1 Αντικείμενο διπλωματικής εργασίας	22
1.2 Κίνητρα διεξαγωγής διπλωματικής εργασίας	25
1.3 Συνεισφορά στην επιστημονική κοινότητα και παρόμοιες εργασίες	26
1.4 Σύνοψη διπλωματικής εργασίας	28
Κεφάλαιο 2: Θεωρητικό υπόβαθρο	31
2.1 Μικροϋπολογιστής BeagleBoard	31
2.1.1 Τύποι BeagleBoard	32
2.1.2 Λογισμικό	34
2.2 C++	35
2.3 Matlab	37
2.4 Motion	38
2.5 Νήματα (Threads)	40
2.6 Βιβλιοθήκη Mpg123	42
2.7 cURL	43
POST	44
Κεφάλαιο 3: Μηχανική όραση και περιγραφείς χαρακτηριστικών	47
3.1 Όραση υπολογιστή	47

3.3.1 Εφαρμογές της μηχανικής όρασης	51
3.3.2 Προβλήματα που αντιμετωπίζει η μηχανική όραση	52
3.2 Βιβλιοθήκη OpenCV	54
3.3 Περιγραφείς χαρακτηριστικών (feature descriptors)	58
3.3.1 Histogram of oriented gradient (HOG)	61
3.3.2 Haar features	63
3.3.3 Σύγκριση περιγραφέων	65
Κεφάλαιο 4: Υλοποίηση του συστήματος	67
4.1 Εκπαίδευση ταξινομητών	67
4.2 Λογισμικό Μέρος	76
4.2.1 Αυτοματοποίηση Διαδικασίας Σύνδεσης στο Διαδίκτυο	77
4.2.2 Ανίχνευση συναισθήματος	82
4.2.3 Επικοινωνία υπολογιστή - μικροϋπολογιστή	86
4.2.3.1 Υποδοχές Ρευμάτων (Sockets)	87
4.2.3.2 Επικοινωνία Client/Server	88
4.2.4 Ανταλλαγή διεύθυνσης IP	91
4.3 Ασφάλεια του συστήματος	96
4.4 S.W.O.T. ανάλυση του συστήματος	97
4.5 Μετρικές του έργου	98
Κεφάλαιο 5: Κατασκευή ενσωματωμένου συστήματος	101
5.1 Διάταξη των ηλεκτρονικών μέσων της κατασκευής	102
5.2 Απενεργοποίηση-Ενεργοποίηση Συσκευής	104
5.2.1 Κομβίο Τερματισμού	104
5.2.2 Διακόπτης τροφοδοσίας	106
5.3 Τελική μορφή της κατασκευής	107
Κεφάλαιο 6: Συμπεράσματα και μελλοντικές επεκτάσεις	109
6.1 Συμπεράσματα	109

6.2 Προβλήματα που αντιμετωπίστηκαν κατά την υλοποίηση	111
6.3 Μελλοντικές επεκτάσεις	113
Παράρτημα	115
Βιβλιογραφία	117

Κατάλογος Πινάκων

Πίνακας 1: Μικροϋπολογιστές της οικογένειας BeagleBoard.....	34
Πίνακας 2: Κατάλογος βασικών ενοτήτων του OpenCV με τις λειτουργίες τους.....	57
Πίνακας 3: Δεδομένα που χρησιμοποιήθηκαν κατά την εκπαίδευση καθενός από τους ταξινομητές.	75
Πίνακας 4: Μετρικές κώδικα στον υπολογιστή	98
Πίνακας 5: Μετρικές κώδικα στον μικροϋπολογιστή.....	98
Πίνακας 6: Μετρικές κώδικα στο server του πανεπιστημίου	98

Κατάλογος Εικόνων

Εικόνα 1: Στιγμιότυπο που καταγράφηκε από κάμερα πριν οποιαδήποτε επεξεργασία. .	39
Εικόνα 2: Στιγμιότυπο μετά την εφαρμογή του φίλτρου despeckle από το Motion. Μόνο η μπλε περιοχή αναγνωρίζεται ως κίνηση.	39
Εικόνα 3: Ο πλαϊνός καθρέπτης από τη μεριά του οδηγού, όπως γίνεται αντιληπτός από τη μηχανική όραση. Κάθε αριθμός του πλέγματος αντιστοιχίζεται σε κάποια συνιστώσα του αντικειμένου με έντονο το στοιχείο του θορύβου.	49
Εικόνα 4: Συσχέτιση του τομέα της μηχανικής όρασης με άλλα επιστημονικά πεδία.	51
Εικόνα 5: Βασικά στοιχεία που απαρτίζουν το OpenCV.....	57
Εικόνα 6: Παράδειγμα χρήσης περιγραφικών χαρακτηριστικών (feature descriptors).	58
Εικόνα 7: Διαδικασία περικοπής εικόνων πριν την εφαρμογή του περιγραφέα HOG.	61
Εικόνα 8: Πυρήνες που χρησιμοποιούνται για το φιλτράρισμα των εικόνων.	62
Εικόνα 9: Σειρά διαδικασιών η οποία ακολουθείται προκειμένου να παραχθεί το ιστόγραμμα της εικόνας. Τα βήματα εξαγωγής περιγραφών HOG: (a) πρωτότυπη εικόνα, (b) υπολογισμός διανύσματος κλίσης, (c) εικονοστοιχείο ομαδοποίηση κυττάρων, (d) υπολογισμός ιστόγραμμα.	63
Εικόνα 10: Χαρακτηριστικά Haar που χρησιμοποιούνται από μια εικόνα.....	64
Εικόνα 11: Παραδείγματα φωτογραφιών που χρησιμοποιήθηκαν για την παραγωγή θετικών δειγμάτων που απεικονίζουν χαμογελαστό στόμα.....	70
Εικόνα 12: Παράδειγμα φωτογραφίας με διαστρεβλωμένη μορφή στόματος που χρησιμοποιήθηκε σαν αρνητική εικόνα	71
Εικόνα 13: Παραδείγματα φωτογραφιών στις οποίες έχει οριοθετηθεί η περιοχή ενδιαφέροντος	72
Εικόνα 14: Παράδειγμα της δομής του τελικού αρχείου που περιέχει το σύνολο των οριοθετημένων περιοχών ενδιαφέροντος από τις θετικές εικόνες.....	73
Εικόνα 15: Απόσπασμα από τα στάδια εκπαίδευσης ταξινομητή με τη συνάρτηση trainCascadeObjectDetector.....	75
Εικόνα 16: Παραδείγματα ανίχνευσης α) χαρούμενου, β) λυπημένου και γ) ανοιχτού στόματος με χρήση ταξινομητών Haar. Οι ανιχνεύσεις δεν είναι οι επιθυμητές.....	76
Εικόνα 17: Φόρμα εισαγωγής στοιχείων σύνδεσης	78
Εικόνα 18: Μήνυμα αποτροπής καταχώρισης σύνδεσης με κενό όνομα δικτύου.....	78

Εικόνα 19: Διάγραμμα ροής που περιγράφει τη διαδικασία προετοιμασίας του παιχνιδιού	81
Εικόνα 20: Κανένα συναίσθημα δεν ανιχνεύεται σε περίπτωση που τα μάτια είναι κλειστά.	83
Εικόνα 21: Επιτυχείς αναγνωρίσεις α) χαρούμενου, β) λυπημένου και γ) ανοιχτού στόματος.....	84
Εικόνα 22: Περιεχόμενα κεφαλίδας ενός μηνύματος που μεταδίδεται μέσω πρωτοκόλλου TCP. Τα περιεχόμενα αυτά καθιστούν το πρωτόκολλο μετάδοσης TCP ασφαλές και αξιόπιστο.	88
Εικόνα 23: Ο πελάτης αιτείται σύνδεσης στη θύρα που «ακούει» ο εξυπηρετητής.....	88
Εικόνα 24: Η σύνδεση μεταξύ πελάτη και εξυπηρετητή είναι αμφίδρομη.....	89
Εικόνα 25: Η επικοινωνία μεταξύ του πελάτη και του εξυπηρετητή μέσω υποδοχής ρευμάτων συμπεριλαμβανομένων όλων των βημάτων της διαδικασίας.	89
Εικόνα 26: Στο ενσωματωμένο σύστημα που αναπτύχθηκε ο μικροϋπολογιστής αποφορτίζεται από την επεξεργασία του βίντεο για την ανίχνευση συναισθήματος. Η διαδικασία αυτή μετατάσσεται στον υπολογιστή. Το σύνολο των διεργασιών και της επικοινωνίας μεταξύ των οντοτήτων.	91
Εικόνα 27: Σχηματική περιγραφή της ανταλλαγής διεύθυνσης IP μεταξύ μικροϋπολογιστή και υπολογιστή.	93
Εικόνα 28: Διάγραμμα ροής του συνολικού προγράμματος.....	95
Εικόνα 29: Εμπρόσθια όψη σκεύους αποθήκευσης. Φαίνονται αναλυτικά όλα τα στοιχεία που βρίσκονται στο εξωτερικό της κατασκευής	102
Εικόνα 30: Εσωτερικό του σκεύους αποθήκευσης.....	103
Εικόνα 31: Αναλυτικά η αρίθμηση και τα σήματα των συνδέσμων επέκταση επεξεργαστή	105
Εικόνα 32: Τελική εικόνα της κατασκευής: α) μπροστινό και β) πίσω μέρος	108

Περίληψη

Η φροντίδα και η προστασία των παιδιών είναι πάντα οι πρώτες προτεραιότητες των γονέων. Η ανατροφή τους κατά την βρεφική ηλικία είναι ένας τομέας που πάντα επιδέχεται βελτιώσεις. Τα παιχνίδια που κυκλοφορούν στο εμπόριο είναι κατασκευασμένα για να συντροφεύουν τα παιδιά και να τους παρέχουν τις πρώτες γνώσεις. Οι απαιτήσεις, όμως, των γονέων από ένα παιχνίδι είναι, ιδανικά, πολύ περισσότερες. Λαμβάνοντας υπόψη αυτές τις απαιτήσεις και συνδυάζοντας τη μηχανική όραση με τα ενσωματωμένα συστήματα δημιουργήθηκε ένα έξυπνο παιχνίδι. Στόχος του είναι η καινοτομία στον κόσμο των παιχνιδιών και η παροχή νέων δυνατοτήτων.

Το εν λόγω παιχνίδι καλύπτει τόσο τις ανάγκες των παιδιών όσο και των γονέων. Έχει τη μορφή ενός λούτρινου παιχνιδιού, στο οποίο ενσωματώνεται ένας μικροϋπολογιστής με μία κάμερα και ένα ηχείο συνδεδεμένα πάνω του. Το ολοκληρωμένο σύστημα λαμβάνει εικόνα πραγματικού χρόνου από την κάμερα και την επεξεργάζεται για να ανιχνεύσει σε αυτήν εκφράσεις προσώπου. Στη συνέχεια, ανάλογα με το συναίσθημα που έχει ανιχνευτεί αντιδράει αναπαράγοντας έναν ήχο μέσω του ηχείου. Για να επιτευχθεί το αποτέλεσμα αυτό, προηγήθηκε εκπαίδευση κατάλληλων ταξινομητών. Στην επεξεργασία του βίντεο συντελεί ένας υπολογιστής, υψηλής επεξεργαστικής ισχύος και συνεπώς πραγματοποιείται επικοινωνία μεταξύ αυτού και του μικροϋπολογιστή με σκοπό την ανταλλαγή δεδομένων. Επιπλέον, το παιχνίδι προσφέρει τη δυνατότητα στους γονείς να αποκτούν πρόσβαση στην εικόνα που λαμβάνεται από την κάμερα, για να μπορούν να παρακολουθούν το παιδί ανά πάσα στιγμή.

Το παιχνίδι χαρακτηρίζεται από αυτονομία και είναι φορητό, καθώς η σύνδεση των τμημάτων που το απαρτίζουν γίνεται δυναμικά και δεν απαιτεί παροχή ρεύματος, δεδομένου ότι είναι φορτισμένο. Επιπρόσθετο χαρακτηριστικό του είναι η επεκτασιμότητα, καθώς μπορεί να εμπλουτιστεί ώστε να αναγνωρίζει πλήθος συναισθημάτων. Μπορεί να χρησιμοποιηθεί από πολλές ηλικίες και να προσφέρει συντροφιά και πνευματική ανάπτυξη.

Λέξεις κλειδιά: OpenCV, Beagleboard –xM, Ενσωματωμένο Σύστημα, Αναγνώριση Συναισθήματος, Έξυπνο παιχνίδι

Abstract

The care and protection of children are always the first priorities of parents. Their upbringing during the infantile age is a sector that has always been susceptible to improvements. The commercially distributed toys are manufactured to keep the children company and accommodate them with their very first acknowledgment of things. However, the parents' demands of a toy are, ideally, many more. Having taken these requirements into consideration and by combining the computer vision with the embedded systems, an intelligent toy has been implemented. The aim of this toy is the innovation in the world of toys and the provision of new capabilities.

The toy in question covers the needs both of children and parents. It has the form of a teddy bear, in which a single board computer with a camera and a speaker are embedded. The complete system process the live streaming video of the camera in order to identify facial expressions. Afterwards, it reacts to the feeling recognized by putting a sound on playback. To accomplish this feeling identification, the appropriate cascade classifiers were trained. The video processing is the responsibility of a personal computer with high processing power. Therefore, a connection between the computer and the single board computer is established for data exchange. Furthermore, the toy provides access to the video streaming, so the parents can monitor the child at all times.

The introduced toy is autonomous and portable since the connections between its parts are all dynamic and as long as it is charged there is no need to be plugged in. In addition, scalability is one of its characteristics as it can be enriched to recognize many more feelings. It can be used by children of many ages to provide company and mental growth.

Keywords: OpenCV, Beagleboard –xM, Embedded system, Emotion recognition, Intelligent toy

Δήλωση πνευματικών δικαιωμάτων

Δηλώνω ρητά ότι, σύμφωνα με το άρθρο 8 του Ν. 1599/1986 και τα άρθρα 2,4,6 παρ. 3 του Ν. 1256/1982, η παρούσα Διπλωματική Εργασία με τίτλο *«Σχεδιασμός και υλοποίηση ενός έξυπνου παιχνιδιού για τη συναισθηματική παρακολούθηση και το γονικό έλεγχο ενός παιδιού»*

καθώς και τα ηλεκτρονικά αρχεία και πηγαίοι κώδικες που αναπτύχθηκαν ή τροποποιήθηκαν στα πλαίσια αυτής της εργασίας και αναφέρονται ρητώς μέσα στο κείμενο που συνοδεύουν, και η οποία έχει εκπονηθεί στο Τμήμα Μηχανικών Πληροφορικής και Τηλεπικοινωνιών του Πανεπιστημίου Δυτικής Μακεδονίας, υπό την επίβλεψη του μέλους του Τμήματος κ. Δρ. Μηνά Δασυγένη, αποτελεί αποκλειστικά προϊόν προσωπικής εργασίας και δεν προσβάλλει κάθε μορφής πνευματικά δικαιώματα τρίτων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον. Τα σημεία όπου έχω χρησιμοποιήσει ιδέες, κείμενο, αρχεία ή / και πηγές άλλων συγγραφέων, αναφέρονται ευδιάκριτα στο κείμενο με την κατάλληλη παραπομπή και η σχετική αναφορά περιλαμβάνεται στο τμήμα των βιβλιογραφικών αναφορών με πλήρη περιγραφή.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα. Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και μόνο.

Copyright (C) Μιχαέλα Γούλα - Δημητρίου, Δασυγένης Μηνάς, 2018, Κοζάνη

Υπογραφή φοιτητή:

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον Δρ. Μηνά Δασυγένη για την εμπιστοσύνη και την υπομονή που έδειξε κατά την περάτωση της παρούσας διπλωματικής εργασίας, αλλά και την καθοδήγησή του καθ' όλη την διάρκεια εκπόνησής της.

Επιπλέον, θα ήθελα να ευχαριστήσω την οικογένειά μου και τους φίλους μου για την υποστήριξη και τη βοήθειά τους όλα αυτά τα χρόνια και κυρίως τα άτομα που συνέβαλαν εθελοντικά στην εργασία όταν αυτό τους ζητήθηκε.

Κεφάλαιο 1: Εισαγωγή

Ο ερχομός ενός παιδιού στον κόσμο είναι το πιο σημαντικό γεγονός στη ζωή των γονέων. Από τη γέννηση του και μετά, το κύριο μέλημά τους είναι να είναι αυτό ασφαλές, ειδικά στα πρώτα στάδια της ζωής του, να αναπτυχθεί σωστά σε όλους τους τομείς και να έχει τα απαραίτητα εφόδια για τη συνέχεια. Κάθε γονέας επιθυμεί για το παιδί του τη σωστή ανατροφή, με τον ίδιο να είναι διαρκώς παρών. Η ευθύνη του είναι μεγάλη καθώς πρέπει να φροντίζει ανελλιπώς για τις βασικές ανάγκες του παιδιού του, αλλά ταυτόχρονα είναι υπεύθυνος να το βοηθήσει να αναπτυχθεί τόσο πνευματικά όσο και συναισθηματικά και να διαμορφώσει κριτική σκέψη.

Λόγω του πλήθους των υποχρεώσεων εντός και εκτός σπιτιού, οι γονείς δεν έχουν πάντα τη δυνατότητα να αφιερώνουν όλο τους το χρόνο στα παιδιά. Καθημερινά χρειάζονται βοήθεια και καταφεύγουν σε λύσεις που μπορούν σε ένα βαθμό να τους αντικαταστήσουν, εξακολουθώντας, όμως, να έχουν τον έλεγχο και τη σιγουριά ότι τα παιδιά τους είναι ασφαλή και μεγαλώνουν με τις καλύτερες προϋποθέσεις. Σε αυτή τους την προσπάθεια μπορούν να ωφεληθούν τόσο από άλλα φυσικά πρόσωπα όσο και από τεχνητά μέσα, όπως συσκευές παρακολούθησης και εκπαιδευτικά παιχνίδια.

Η ραγδαία ανάπτυξη της τεχνολογίας έρχεται για να καλύψει κενά και να διευκολύνει τα πράγματα ακόμα και σε τόσο «ευαίσθητους» τομείς, όπως η ανατροφή ενός παιδιού. Οι άπειρες δυνατότητες που προσφέρουν οι μικροϋπολογιστές και κατ' επέκταση οι συσκευές στις οποίες ενσωματώνονται, μπορούν να υποστηρίξουν τους γονείς στην καθημερινότητα τους με τα παιδιά. Μπορούν επίσης να συμβάλλουν σημαντικά στην πνευματική και συναισθηματική σταθερότητα και ανάπτυξη των παιδιών, διεγείροντας τις αισθήσεις τους και προτρέποντάς τα να σκεφτούν, να πράξουν και να αναρωτηθούν σχετικά με πράγματα και καταστάσεις που συναντούν στην καθημερινότητα τους.

Λαμβάνοντας υπόψη όλα τα παραπάνω και έχοντας σαν κύριο γνώμονα τη φροντίδα των παιδιών, σχεδιάστηκε και αναπτύχθηκε ένα έξυπνο παιχνίδι. Το παιχνίδι αυτό εκμεταλλεύεται και συνδυάζει τα προνόμια που προσφέρουν οι μικροϋπολογιστές αλλά και η αναγνώριση προτύπων μέσω τεχνητής όρασης. Σαν στόχο έχει να κρατάει τα παιδιά, κυρίως μικρής ηλικίας, απασχολημένα, ασφαλή και σε εγρήγορση, να τα συντροφεύει και να προλαβαίνει τις ανάγκες τους. Ταυτόχρονα διευκολύνει τους γονείς δίνοντας τους κάποια παραπάνω ευελιξία. Αυτοί πλέον μπορούν να αφήνουν τα παιδιά μόνα για μικρά χρονικά διαστήματα χωρίς ιδιαίτερη ανησυχία, αφού τα έχουν πάντα υπό την επίβλεψή τους απομακρυσμένα.

1.1 Αντικείμενο διπλωματικής εργασίας

Ο τομέας των παιχνιδιών, μέχρι σήμερα, περιοριζόταν στην κατασκευή και διάθεση διαδραστικών παιχνιδιών τα οποία σχετίζονται με την κινησιολογική ανάπτυξη του παιδιού και όχι τόσο με την πνευματική ή τη συναισθηματική. Συνεπώς, τα παιχνίδια αυτά αλληλεπιδρούν με το παιδί ύστερα από κάποιο ερέθισμα, όπως είναι το πάτημα ενός κουμπιού ή μέσω κάποιας φωνητικής εντολής, η οποία χρησιμοποιείται σε πιο προηγμένα συστήματα.

Με την ανάπτυξη της τεχνολογίας και κυρίως την πρόοδο γύρω από τον τομέα της μηχανικής όρασης, δημιουργήθηκε η δυνατότητα αναγνώρισης ανθρώπων, αλλά και πιο συγκεκριμένα, χαρακτηριστικών των προσώπων τους και άλλων γνωρισμάτων όπως η ηλικία, το φύλο κ.α. Έτσι τα προϊόντα τεχνολογίας αρχίζουν να έχουν περισσότερο ανθρωποκεντρικό χαρακτήρα και να προσαρμόζονται ανάλογα με τις εκάστοτε ανάγκες του ατόμου, διαχωρίζοντάς το πλέον από το πλήθος και αντιμετωπίζοντας το σαν ξεχωριστή οντότητα με δικά της γνωρίσματα.

Στην παρούσα διπλωματική εργασία, αξιοποιήθηκαν οι δυνατότητες που προσφέρει η όραση μηχανής και σε συνδυασμό με τα πλεονεκτήματα των ενσωματωμένων συστημάτων αναπτύχθηκε ένα καινοτόμο παιχνίδι, το οποίο μπορούν να αξιοποιήσουν τόσο τα παιδιά, όσο και οι γονείς. Το έξυπνο αυτό παιχνίδι εισάγει την αναγνώριση συναισθημάτων για να ενισχύσει ακόμα περισσότερο τον τομέα της ψυχαγωγίας των

παιδιών και να επικεντρώσει το ενδιαφέρον του στη φροντίδα και την ασφάλειά τους. Ταυτόχρονα, συμβάλλει στην διανοητική ανάπτυξη των νεαρών ατόμων και στη διέγερση της κριτικής τους σκέψης.

Η κατασκευή έχει τη μορφή ενός λούτρινου παιχνιδιού, για να είναι προσιτό στα παιδιά, το οποίο έχει εμπλουτιστεί με έναν μικροϋπολογιστή, ένα μέσο τροφοδοσίας, μία κάμερα και ένα ηχείο. Το ολοκληρωμένο αυτό σύστημα έχει σαν στόχο την ανίχνευση του προσώπου ενός παιδιού και την αντίδρασή του βάσει των μορφασμών που αναγνωρίζει σε αυτό. Οι μορφασμοί οι οποίοι υποστηρίζονται είναι η χαρά, η λύπη και το χασμουρητό. Ανάλογα με τον μορφασμό του παιδιού, το σύστημα αναπαράγει τον αντίστοιχο ήχο με στόχο να το προσελκύσει και να διατηρήσει την καλή του διάθεση, να βελτιώσει τη διάθεσή του με ένα γνωστό τραγούδι που μπορεί να το παρακινήσει να τραγουδήσει, αλλά και να το νανουρίσει, αντίστοιχα.

Το λογισμικό της παρούσας εφαρμογής λαμβάνει δεδομένα εικόνας σαν είσοδο και αλληλεπιδρά επιστρέφοντας ήχο σαν έξοδο. Πιο συγκεκριμένα, το σύστημα δέχεται συνεχώς ροή βίντεο σε πραγματικό χρόνο, στο οποίο προσπαθεί να ανιχνεύσει πρόσωπα. Η ανίχνευση αυτή λαμβάνει χώρα χάρη σε συγκεκριμένους ταξινομητές που έχουν εκπαιδευτεί σχετικά. Η εκπαίδευση των ταξινομητών έχει βασιστεί σε ένα πλήθος συλλεγμένων εικόνων που απεικονίζουν κατά βάση το στοιχείο που τίθεται προς ανίχνευση. Εφόσον αναγνωριστεί κάποιο πρόσωπο, άλλοι ταξινομητές, εξειδικευμένοι σε συγκεκριμένα χαρακτηριστικά του, προσπαθούν τόσο να τα αναγνωρίσουν όσο και να αποφανθούν σε ποια κατάσταση αυτά βρίσκονται. Εάν ο έλεγχος των ταξινομητών καταλήξει στην ύπαρξη κάποιου εμφανούς, για συγκεκριμένο χρόνο, συναισθήματος στο πρόσωπο, τότε πραγματοποιείται η αλληλεπίδραση του συστήματος. Ένας επιλεγμένος ήχος, τέτοιος ώστε να ταιριάζει στο συναίσθημα που αναγνωρίστηκε, φορτώνεται και ξεκινά να αναπαράγεται. Η αναπαραγωγή αυτή εξακολουθεί έως ότου τα χαρακτηριστικά του προσώπου αλλάξουν κατάσταση και παραπέμπουν πια σε κάποιο άλλο συναίσθημα. Τότε η αναπαραγωγή διακόπτεται και τη θέση του προηγούμενου ήχου παίρνει ένας νέος, πιο ταιριαστός στην ανίχνευση που πραγματοποιήθηκε.

Η παρούσα εργασία φέρει και ερευνητικό χαρακτήρα, καθώς πέρα από την υλοποίηση του τελικού συστήματος, στόχος της ήταν η μελέτη και η ανάπτυξη ενός προϊόντος το οποίο θα μπορεί να χρησιμοποιείται ευρέως και να υποστηρίζει τα όσα

αναφέρθηκαν παραπάνω σε συνθήκες πραγματικού χρόνου. Για το λόγο αυτό, έγιναν δοκιμές για δύο εκδόσεις του συστήματος, που η κύρια διαφορά τους έγκειται στο υλικό μέρος, το οποίο είναι διαφορετικό σε υπολογιστική ισχύ και επιδόσεις.

Η πρώτη έκδοση καθιστά τη συσκευή πλήρως αυτόνομη, καθώς αυτή χρησιμοποιεί τον ενσωματωμένο μικροϋπολογιστή για όλες τις διεργασίες και τους υπολογισμούς που απαιτούνται. Αυτό έχει σαν αποτέλεσμα, το παιχνίδι να είναι ανεξάρτητο και ευέλικτο στις μετακινήσεις, δεδομένου ότι είναι φορτισμένο. Η δεύτερη έκδοση μεταθέτει ολόκληρη την υπολογιστική πολυπλοκότητα σε έναν σταθερό υπολογιστή αυξημένων επιδόσεων. Ο μικροϋπολογιστής συνεχίζει να έχει καθοριστικό ρόλο, καθώς είναι υπεύθυνος για την είσοδο και την έξοδο δεδομένων, αλλά έχει πλέον αποφορτιστεί από τις διεργασίες που απαιτούν υψηλή υπολογιστική ισχύ. Τελικά, η δεύτερη έκδοση επιτυγχάνει μεγάλες ταχύτητες, απόκριση σε πραγματικό χρόνο και συνεπώς μεγάλα ποσοστά επιτυχούς αναγνώρισης, έναντι της πρώτης η οποία, λόγω έλλειψης πόρων, εισάγει μεγάλη καθυστέρηση στα δεδομένα εικόνας με αποτέλεσμα η ανίχνευση του συναισθήματος να γίνεται ετεροχρονισμένα.

Ταυτόχρονα με τις δυνατότητες ανίχνευσης και αναγνώρισης, το εν λόγω προηγμένο παιχνίδι προσφέρει και μία λειτουργία ειδικά σχεδιασμένη για τους γονείς. Εφόσον σημαντικό για τους γονείς είναι να ξέρουν ανά πάσα στιγμή την κατάσταση του παιδιού τους, το σύστημα προσφέρει δυνατότητα συνεχούς απομακρυσμένου ελέγχου και παρακολούθησής του από αυτούς. Για την παροχή αυτής της λειτουργίας, γίνεται εκμετάλλευση του βίντεο και της συνεχούς ροής των εικόνων που τίθενται προς αναγνώριση. Στόχος είναι οι γονείς να μπορούν να αποκτούν πρόσβαση σε αυτή την εικόνα, πράγμα που καθίσταται δυνατό από οποιαδήποτε απομακρυσμένη συσκευή, αρκεί αυτή να βρίσκεται συνδεδεμένη στο ίδιο δίκτυο. Με τον τρόπο αυτό επιτυγχάνεται τόσο η ασφάλεια του παιδιού, όσο και των προσωπικών δεδομένων της οικογένειας.

Έχοντας τα παραπάνω χαρακτηριστικά, το εν λόγω προηγμένο παιχνίδι αποτελεί ένα καινοτόμο και ταυτόχρονα ασφαλές μέσο για παρακολούθηση και φροντίδα ενός παιδιού μικρής ηλικίας. Οι δυνατότητές του είναι αυξημένες, καθώς, μέσα από τη χρήση του μπορούν να ικανοποιηθούν τόσο οι ανάγκες των παιδιών, όσο και των γονέων.

1.2 Κίνητρα διεξαγωγής διπλωματικής εργασίας

Οι άνθρωποι έχουν σαν έμφυτο χάρισμα να αναγνωρίζουν τα αντικείμενα και την κατάσταση στην οποία αυτά βρίσκονται, εκμεταλλευόμενοι την αίσθηση της όρασης. Σήμερα οι ανάγκες για αυτοματοποίηση της διαδικασίας αναγνώρισης είναι ιδιαίτερα αυξημένες. Έτσι, ο τομέας της μηχανικής όρασης έχει αρχίσει να γνωρίζει μεγάλη άνθιση, με τους υπολογιστές να διαδραματίζουν τον πιο σημαντικό ρόλο. Ταυτόχρονα, υπάρχει διαρκής ζήτηση για βελτιστοποίηση της διαδικασίας επεξεργασίας εικόνας και την παραγωγή όσο γίνεται περισσότερων αποτελεσμάτων από αυτήν, τα οποία θα μπορούν να συσχετίζονται και να επηρεάζουν ποικίλους τομείς. Οι υπολογιστές, λοιπόν, συνδυάζουν την όραση μηχανής με την επεξεργασία εικόνας και τα αποτελέσματα γίνονται ολοένα και πιο θεαματικά.

Αναλαμβάνοντας την παρούσα διπλωματική εργασία, στόχος ήταν η ένταξη στο πεδίο της τεχνητής όρασης και η απόκτηση γνώσεων πάνω σε αυτό. Εξίσου ενδιαφέρουσα και άξια ενασχόλησης αποτέλεσε η διαρκής διεύρυνση του τομέα των μικροϋπολογιστών. Είναι εντυπωσιακό το πώς ένας τόσο μικρός υπολογιστής προσπαθεί να αντικαταστήσει αυτούς που είναι γνωστοί και χρησιμοποιούνται μέχρι τώρα, να σταθεί αντάξιος και να επιτελέσει πολλές και διαφορετικές λειτουργίες. Το πιο σημαντικό χαρακτηριστικό των μικροϋπολογιστών, που αποδείχτηκε ιδιαίτερα χρήσιμο, είναι η δυνατότητα ενσωμάτωσής τους σε αντικείμενα προσιτά και καθημερινής χρήσης, καταφέρνοντας έτσι να τα μεταλλάξουν. Τα μετατρέπουν πλέον σε προγραμματισμένες μηχανές με αναρίθμητες δυνατότητες.

Έχοντας υπόψη όλα τα παραπάνω, συνδυάστηκαν οι τομείς της τεχνητής όρασης και των ενσωματωμένων συστημάτων για να δημιουργηθεί ένα καινοτόμο αντικείμενο, με δυνατότητα ευρείας χρήσης σε έναν εξίσου σημαντικό τομέα, που αυτή τη φορά δεν ανήκει στα πλαίσια της τεχνολογίας. Ο τομέας που επιλέχτηκε είναι η ψυχαγωγία και η διαπαιδαγώγηση των παιδιών, ένας τομέας μείζονος σημασίας, καθώς επηρεάζει πολύ και επηρεάζεται από την ψυχολογία τους. Σκοπός, λοιπόν, ήταν η απόκτηση γνώσεων και η εξοικείωση πάνω σε δύο πολύ διαδεδομένους τομείς της επιστήμης των υπολογιστών και η χρήση αυτών για την ανάπτυξη ενός προϊόντος με ανθρωποκεντρικό χαρακτήρα.

1.3 Συνεισφορά στην επιστημονική κοινότητα και παρόμοιες εργασίες

Τα τελευταία χρόνια έχουν αναπτυχθεί πολλές εφαρμογές πάνω στην όραση μηχανής και πολλές που χρησιμοποιούν ενσωματωμένα συστήματα. Ωστόσο υπάρχουν και παραδείγματα εφαρμογών που συνδυάζουν τους δύο αυτούς επιστημονικούς τομείς, όπως και η παρούσα διπλωματική εργασία. Αυτή ασχολείται με την αναγνώριση προτύπων και συγκεκριμένα την αναγνώριση προσώπου και επιμέρους χαρακτηριστικών αυτού. Η επίτευξη των πολλαπλών αναγνωρίσεων που απαιτούνται πραγματοποιήθηκε με τη μέθοδο ειδικά εκπαιδευμένων ταξινομητών, μία μέθοδο η οποία εισήχθη το 2001 από τους Paul Viola και Michael Jones [1]. Περισσότερα για τη μέθοδο αυτή, αλλά και λεπτομέρειες που αφορούν την εκπαίδευση των ταξινομητών θα αναφερθούν σε επόμενο κεφάλαιο. Επιπλέον, καθοριστικό ρόλο για την υλοποίηση της εργασίας είχε η έρευνα γύρω από τα ενσωματωμένα συστήματα, η χρήση ενός μικροϋπολογιστή και η ζεύξη ανάμεσα σε αυτόν και σε έναν κανονικό υπολογιστή ώστε να επιτευχθούν τα καλύτερα δυνατά αποτελέσματα και αποδόσεις.

Όλα αυτά συνδυάστηκαν για να αναπτυχθεί μία εφαρμογή αναγνώρισης εκφράσεων προσώπου. Η εφαρμογή αυτή προσανατολίζεται και αφορά μία μείζονος σημασίας περίοδο της ζωής του ατόμου, την περίοδο της βρεφικής και προσχολικής ηλικίας. Στόχο έχει να εκμεταλλευτεί στο έπακρο τις νέες τεχνολογίες, που βρίσκονται στο προσκήνιο και ανθίζουν και να αποτελέσει τον κύριο άξονα ενός καινοτόμου προϊόντος. Το προϊόν αυτό έχει τη μορφή ενός έξυπνου παιχνιδιού, προσιτού και με διαδραστικό χαρακτήρα. Οι ανθρώπινοι μορφασμοί έχουν τη δυνατότητα να επικοινωνούν συναισθήματά και να προσδιορίζουν συμπεριφορές. Είναι, λοιπόν, κλασικά δείγματα που αντανακλούν την εσωτερική ψυχική κατάσταση του ατόμου. Μεγαλύτερο και πιο σημαντικό ρόλο παίζουν όταν προέρχονται από πρόσωπα πολύ μικρής ηλικίας, που τις περισσότερες φορές δεν έχουν άλλον τρόπο να δείξουν τι αισθάνονται.

Με αναγνώριση προτύπων, επίσης, ασχολήθηκαν και σε σχετική τους εργασία οι Qing Chen, Nicolas D. Georganas, Emil M. Petriu [2]. Στο εγχείρημά τους επέλεξαν να αναγνωρίσουν έναν άλλο τρόπο έκφρασης των ανθρώπων, τις χειρονομίες. Θεωρούν ότι οι τύποι εισόδων που δίνονται μέσω πληκτρολογίου ή ποντικιού είναι πλέον πολύ απαρχαιωμένοι σε σχέση με τα όλο και πιο εκσυγχρονισμένα συστήματα εικονικού

περιβάλλοντος (Virtual Environment systems). Έτσι, αποφάσισαν να χρησιμοποιήσουν μία συσκευή που θα μπορεί να αναγνωρίζει κινήσεις σώματος με τη χρήση μιας κάμερας, και χρησιμοποιώντας την AdaBoost μέθοδο, να εκπαιδεύσουν από την αρχή κάποιους ταξινομητές. Οι ταξινομητές αυτοί εκπαιδεύτηκαν με την ίδια διαδικασία που ακολουθήθηκε και στην παρούσα διπλωματική εργασία και αναγνωρίζουν σε πραγματικό χρόνο σήματα ανθρώπινων χεριών, που πιστεύεται ότι είναι ένας φυσικός και αποτελεσματικός τρόπος για να μπορέσει να επικοινωνήσει ένας άνθρωπος με έναν υπολογιστή.

Ο Duc-Minh Pham σε συνεργασία με τους Trong-Nhan Dam-Nguyen, Phuc-Thinh Nguyen-Vo, Minh-Triet Tran [3]. παρουσίασαν το 2013 ένα εγχείρημα που έχει τον ίδιο στόχο με την παρούσα εργασία, να κρατάει τα παιδιά απασχολημένα και να ενισχύει τις διανοητικές τους ικανότητες. Για να επιτύχουν αυτό το σκοπό χρησιμοποίησαν και εκείνοι ένα λούτρινο αρκουδάκι, το οποίο εμπλούτισαν με ένα σύστημα που αναγνωρίζει παιδικά βιβλία μέσα από την εικονογράφηση τους και κατόπιν αναπαράγει το περιεχόμενό τους. Για την αναγνώριση των εικόνων, βασίστηκαν στο μοντέλο Bag-of-Word, το οποίο διαχειρίζεται τα χαρακτηριστικά των εικόνων σαν λέξεις. Από τα δείγματα εικόνων, λοιπόν, η μέθοδος αυτή, αποσπά έντονα γνωρίσματα, όπως γωνίες και άκρες, τα μετατρέπει σε διανύσματα για να μπορεί να προσδιορίζεται η θέση τους στο χώρο και τα κατηγοριοποιεί σε συστάδες. Τέλος, οι συγγραφείς, χρησιμοποίησαν SVM ταξινομητές για να γίνει πιθανολογικά η ταύτιση της εικόνας με κάποιο πρότυπο.

Αναγνώριση έκφρασης προσώπου, αλλά και μικροϋπολογιστή συνδύασαν επίσης οι Srikanth Sriram και Babu Illuri [4]. κατά την εκπόνηση αντίστοιχης διπλωματικής εργασίας. Στη δική τους εργασία επέλεξαν, επίσης, να χρησιμοποιήσουν τον Adaboost αλγόριθμο για να συγκρίνει και να προσπαθήσει να ταυτίσει χαμόγελο σε πρόσωπα με ταξινομητές που οι ίδιοι εκπαίδευσαν. Η γλώσσα προγραμματισμού που επέλεξαν για την εφαρμογή τους είναι η Python και το λογισμικό τους τρέχει σε μικροϋπολογιστή Raspberry Pi. Το λειτουργικό σύστημα που επέλεξαν για το Raspberry είναι το Raspbian και χρησιμοποίησαν τον OVA5647 CMOS αισθητήρα για καταγραφή εικόνας 5MP.

1.4 Σύνοψη διπλωματικής εργασίας

Το κείμενο δομείται σε έξι κεφάλαια, το καθένα από τα οποία πραγματεύεται ένα διαφορετικό ζήτημα που ερευνήθηκε κατά την υλοποίηση του εν λόγω συστήματος. Στα επιμέρους αυτά κεφάλαια γίνεται αναλυτική παρουσίαση, τόσο του υλικού μέρους, όσο και του λογισμικού που αναπτύχθηκαν. Επίσης, παρουσιάζονται τα προβλήματα που ανέκυψαν και οι τρόποι επίλυσης αυτών.

Στο παρόν κεφάλαιο γίνεται παρουσίαση του αντικειμένου της διπλωματικής εργασίας, τα κίνητρα ανάληψης και διεκπεραίωσης της έρευνας καθώς και μία σύνοψη της δομής που θα ακολουθήσει.

Στο δεύτερο κεφάλαιο, αναλύεται λεπτομερώς το θεωρητικό υπόβαθρο και τα εργαλεία τα οποία χρησιμοποιήθηκαν για την ανάπτυξη του συστήματος. Το κεφάλαιο αυτό είναι σημαντικό καθώς κατά την ανάγνωσή του ο αναγνώστης μπορεί να κατανοήσει τα μέσα και το υλικό τα οποία συντέλεσαν στο να επιτευχθεί ο στόχος της εργασίας.

Το τρίτο κεφάλαιο κάνει εκτενή αναφορά στην υπολογιστική όραση, τη βιβλιοθήκη που χρησιμοποιήθηκε και γενικότερα στις βασικές έννοιες τις επεξεργασίας εικόνας οι οποίες είναι απαραίτητες για την υπόλοιπη εργασία. Επιπλέον, στο κεφάλαιο αυτό παρουσιάζονται τα χαρακτηριστικά των περιγραφών οι οποίοι χρησιμοποιήθηκαν για τις δοκιμές και την τελική ανάπτυξη του λογισμικού ανίχνευσης και αναγνώρισης των συναισθημάτων.

Το τέταρτο κεφάλαιο, αποτελεί την παρουσίαση του λογισμικού μέρους της εργασίας. Αναλύει όλο τον κύκλο εργασιών από την εκκίνηση του συστήματος μέχρι και την επίτευξη της αναγνώρισης συναισθήματος. Εστιάζει σε όλες τις λειτουργίες που εκτελούνται τόσο στον υπολογιστή όσο και στον μικροϋπολογιστή, αλλά και στη μεταξύ τους επικοινωνία.

Στο πέμπτο κεφάλαιο γίνεται ανάλυση του υλικού μέρους της εργασίας. Πιο συγκεκριμένα, παρουσιάζεται ο τρόπος ενσωμάτωσης των ηλεκτρονικών μέσων στο παιχνίδι, αλλά και οι προσθήκες που έγιναν για να είναι αυτό εύχρηστο και προσιτό.

Τέλος, το τελευταίο κεφάλαιο συνοψίζει όλα τα στοιχεία που έχουν παρουσιαστεί στην υπόλοιπη εργασία. Επιπροσθέτως, καταγράφει τους περιορισμούς και τα εμπόδια που υπήρξαν κατά την ανάπτυξη του συστήματος και τις μελλοντικές επεκτάσεις οι οποίες μπορεί να υπάρχουν στην εργασία αυτή.

Σύνοψη Κεφαλαίου

Το παρόν κεφάλαιο αποτέλεσε μια εισαγωγή για τον αναγνώστη στο θέμα που πραγματεύεται η διπλωματική εργασία. Αναφέρθηκαν τα κίνητρα που οδήγησαν στη σύλληψη της ιδέας και παρόμοιες εργασίες, τόσο όσον αφορά το στόχο όσο και τις τεχνολογίες που χρησιμοποιήθηκαν. Τέλος, σημειώνεται περιληπτικά το περιεχόμενο των ακόλουθων κεφαλαίων.

Κεφάλαιο 2: Θεωρητικό υπόβαθρο

Στο κεφάλαιο που ακολουθεί αναλύονται όλα τα μέσα που συντέλεσαν στην υλοποίηση της παρούσας διπλωματικής εργασίας. Εκτός από το λογισμικό που χρησιμοποιήθηκε για την ανάπτυξη των διαφορετικών τμημάτων της εφαρμογής, παρουσιάζονται και τα χαρακτηριστικά του υλικού μέρους που την πλαισιώνει.

2.1 Μικροϋπολογιστής BeagleBoard

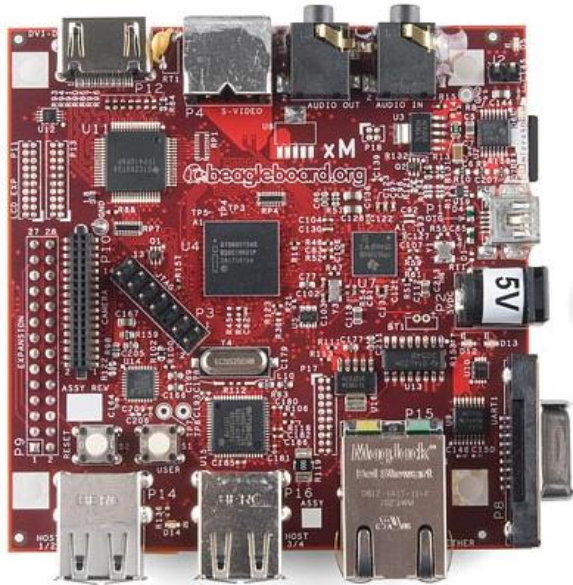
Η οικογένεια BeagleBoard είναι μία οικογένεια ολοκληρωμένων υπολογιστών πολύ μικρού μεγέθους. Η ιδέα για τη δημιουργία τους ξεκίνησε όταν, σε μία ομάδα με έντονο ενδιαφέρον γύρω από τους υπολογιστές προτάθηκε να κάνει ό,τι καλύτερο μπορούσε με έναν μικροϋπολογιστή. Η πρόταση αυτή προήλθε από τον Gerald Coley, ένα μέλος του στρατηγικού μάρκετινγκ της εταιρείας Texas Instruments και ο εν λόγω μικροϋπολογιστής έφερε μεγαλύτερη απόδοση από όλους όσους διέθεταν επεξεργαστή βασισμένο σε αρχιτεκτονική ARM. Η πρώτη σκέψη τους για να ενισχύσουν τον μικροϋπολογιστή ήταν να μπορέσουν να ενσωματώσουν σε αυτήν όσο περισσότερα χαρακτηριστικά ενός υπολογιστή μπορούσαν, όπως LCD οθόνη, Wifi και Ethernet. Σύντομα όμως κατέληξαν στην άποψη ότι για να έχει πολλές δυνατότητες αυτός ο μικροϋπολογιστής, δεν θα έπρεπε να επιβαρυνθεί με on-board προσθήκες.

Καταργώντας, λοιπόν, όλα τα περιφερειακά, τόσο αναλογικά όσο και ψηφιακά, πάνω από τον μικροϋπολογιστή και ταυτόχρονα ενσωματώνοντας διαύλους επέκτασης, όπως μεγάλης ταχύτητας USB 2.0, SDIO (Secure Digital Input Output) και DVI-D, οι προγραμματιστές έχουν πλέον τη δυνατότητα να είναι ευέλικτοι και να χρησιμοποιούν τις δικές τους περιφερειακές συσκευές. Τελικά, δημιουργήθηκε ένας χαμηλού κόστους, χωρίς ανεμιστήρα υπολογιστής, που όμως έχει όλες τις δυνατότητες επέκτασης που



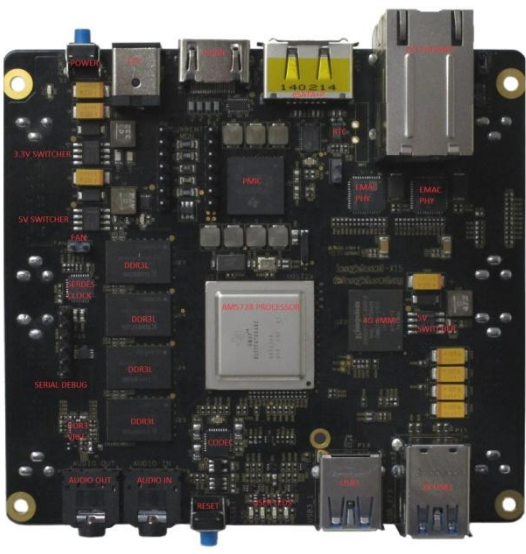
έχουν οι σημερινοί σταθεροί υπολογιστές, μη επιβαρυνόμενος με τον όγκο, τις δαπάνες, αλλά και το θόρυβο. Πλέον, η οικογένεια BeagleBoard μικροϋπολογιστών έχει μεγαλώσει και υποστηρίζεται από μία ολόκληρη κοινότητα ανοιχτού κώδικα (open-source), που ενθαρρύνει συνεχώς νέους προγραμματιστές να συνεισφέρουν σε αυτήν και υποστηρίζει τα έργα τους. [5]

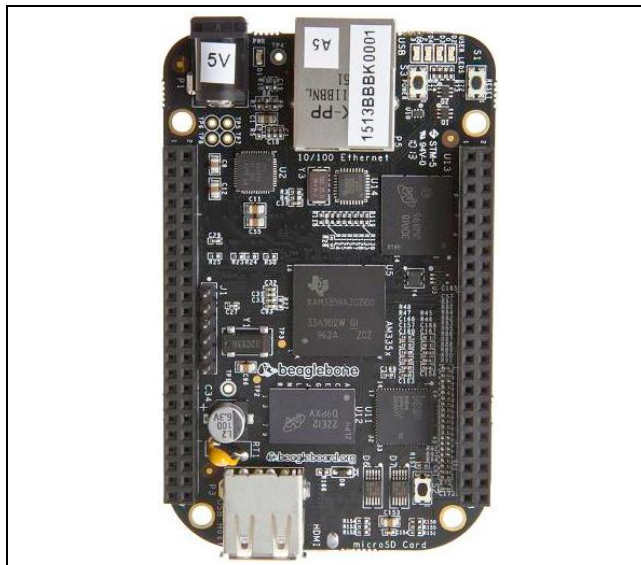
2.1.1 Τύποι BeagleBoard

Ο μικροϋπολογιστής που χρησιμοποιήθηκε για την υλοποίηση της παρούσας διπλωματικής εργασίας ανήκει στην οικογένεια BeagleBoard και συγκεκριμένα είναι ο μικροϋπολογιστής τύπου BeagleBoard –xM. [5]. Στον Πίνακα 1 παρουσιάζονται όλοι οι μικροϋπολογιστές οι οποίοι ανήκουν στον τύπο αυτό.

Υλικό	Χαρακτηριστικά																
	<table border="1"><tr><td>Επεξεργαστής</td><td>AM335x 1GHz ARM® Cortex-A8</td></tr><tr><td>Μνήμη (SDRAM)</td><td>512 MiB DDR2</td></tr><tr><td>DSP</td><td>TMS320C64x+ (800 MHz)</td></tr><tr><td>Αποθήκευση στην Πλακέτα</td><td>microSD card Supported</td></tr><tr><td>Δίκτυο</td><td>Fast Ethernet (via USB hub with Ethernet)</td></tr><tr><td>USB θύρες</td><td>4 x Standard A host port 1x mini AB OTG port (direct)</td></tr><tr><td>Ρεύμα</td><td>350 mA @5 V</td></tr><tr><td>Διαστάσεις</td><td>85.09 mm × 85.6 mm</td></tr></table>	Επεξεργαστής	AM335x 1GHz ARM® Cortex-A8	Μνήμη (SDRAM)	512 MiB DDR2	DSP	TMS320C64x+ (800 MHz)	Αποθήκευση στην Πλακέτα	microSD card Supported	Δίκτυο	Fast Ethernet (via USB hub with Ethernet)	USB θύρες	4 x Standard A host port 1x mini AB OTG port (direct)	Ρεύμα	350 mA @5 V	Διαστάσεις	85.09 mm × 85.6 mm
Επεξεργαστής	AM335x 1GHz ARM® Cortex-A8																
Μνήμη (SDRAM)	512 MiB DDR2																
DSP	TMS320C64x+ (800 MHz)																
Αποθήκευση στην Πλακέτα	microSD card Supported																
Δίκτυο	Fast Ethernet (via USB hub with Ethernet)																
USB θύρες	4 x Standard A host port 1x mini AB OTG port (direct)																
Ρεύμα	350 mA @5 V																
Διαστάσεις	85.09 mm × 85.6 mm																

Σχεδιασμός και υλοποίηση ενός έξυπνου παιχνιδιού για τη συναισθηματική παρακολούθηση και το γονικό έλεγχο ενός παιδιού

	<table border="1"> <tbody> <tr> <td>Επεξεργαστής</td> <td>AM335x 1GHz ARM® Cortex-A8</td> </tr> <tr> <td>Μνήμη (SDRAM)</td> <td>256 MiB DDR2</td> </tr> <tr> <td>DSP</td> <td>-</td> </tr> <tr> <td>Αποθήκευση στην Πλακέτα</td> <td>microSD card 3.3 V Supported</td> </tr> <tr> <td>Δίκτυο</td> <td>Fast Ethernet (MII based)</td> </tr> <tr> <td>USB θύρες</td> <td>1 x Standard A host port 1x mini B device port</td> </tr> <tr> <td>Ρεύμα</td> <td>300–500 mA @5 V</td> </tr> <tr> <td>Διαστάσεις</td> <td>86.40 mm × 53.3 mm</td> </tr> </tbody> </table>	Επεξεργαστής	AM335x 1GHz ARM® Cortex-A8	Μνήμη (SDRAM)	256 MiB DDR2	DSP	-	Αποθήκευση στην Πλακέτα	microSD card 3.3 V Supported	Δίκτυο	Fast Ethernet (MII based)	USB θύρες	1 x Standard A host port 1x mini B device port	Ρεύμα	300–500 mA @5 V	Διαστάσεις	86.40 mm × 53.3 mm	<table border="1"> <tbody> <tr> <td>Επεξεργαστής</td> <td>AM335x 1GHz ARM® Cortex-A8</td> </tr> <tr> <td>Μνήμη (SDRAM)</td> <td>128 MiB (rev B) DDR 256 MiB (rev C+) DDR</td> </tr> <tr> <td>DSP</td> <td>TMS320C64x+ (520 MHz)</td> </tr> <tr> <td>Αποθήκευση στην Πλακέτα</td> <td>256MB NAND Flash, SD/MMC card</td> </tr> <tr> <td>Δίκτυο</td> <td>-</td> </tr> <tr> <td>USB θύρες</td> <td>1 x Standard A host port 1x mini AB OTG port</td> </tr> <tr> <td>Ρεύμα</td> <td>350 mA</td> </tr> <tr> <td>Διαστάσεις</td> <td>78.74 x 76.2mm</td> </tr> </tbody> </table>	Επεξεργαστής	AM335x 1GHz ARM® Cortex-A8	Μνήμη (SDRAM)	128 MiB (rev B) DDR 256 MiB (rev C+) DDR	DSP	TMS320C64x+ (520 MHz)	Αποθήκευση στην Πλακέτα	256MB NAND Flash, SD/MMC card	Δίκτυο	-	USB θύρες	1 x Standard A host port 1x mini AB OTG port	Ρεύμα	350 mA	Διαστάσεις	78.74 x 76.2mm
Επεξεργαστής	AM335x 1GHz ARM® Cortex-A8																																	
Μνήμη (SDRAM)	256 MiB DDR2																																	
DSP	-																																	
Αποθήκευση στην Πλακέτα	microSD card 3.3 V Supported																																	
Δίκτυο	Fast Ethernet (MII based)																																	
USB θύρες	1 x Standard A host port 1x mini B device port																																	
Ρεύμα	300–500 mA @5 V																																	
Διαστάσεις	86.40 mm × 53.3 mm																																	
Επεξεργαστής	AM335x 1GHz ARM® Cortex-A8																																	
Μνήμη (SDRAM)	128 MiB (rev B) DDR 256 MiB (rev C+) DDR																																	
DSP	TMS320C64x+ (520 MHz)																																	
Αποθήκευση στην Πλακέτα	256MB NAND Flash, SD/MMC card																																	
Δίκτυο	-																																	
USB θύρες	1 x Standard A host port 1x mini AB OTG port																																	
Ρεύμα	350 mA																																	
Διαστάσεις	78.74 x 76.2mm																																	
	<table border="1"> <tbody> <tr> <td>Επεξεργαστής</td> <td>AM335x 1GHz ARM® Cortex-A8</td> </tr> <tr> <td>Μνήμη (SDRAM)</td> <td>128 MiB (rev B) DDR 256 MiB (rev C+) DDR</td> </tr> <tr> <td>DSP</td> <td>TMS320C64x+ (520 MHz)</td> </tr> <tr> <td>Αποθήκευση στην Πλακέτα</td> <td>256MB NAND Flash, SD/MMC card</td> </tr> <tr> <td>Δίκτυο</td> <td>-</td> </tr> <tr> <td>USB θύρες</td> <td>1 x Standard A host port 1x mini AB OTG port</td> </tr> <tr> <td>Ρεύμα</td> <td>350 mA</td> </tr> <tr> <td>Διαστάσεις</td> <td>78.74 x 76.2mm</td> </tr> </tbody> </table>	Επεξεργαστής	AM335x 1GHz ARM® Cortex-A8	Μνήμη (SDRAM)	128 MiB (rev B) DDR 256 MiB (rev C+) DDR	DSP	TMS320C64x+ (520 MHz)	Αποθήκευση στην Πλακέτα	256MB NAND Flash, SD/MMC card	Δίκτυο	-	USB θύρες	1 x Standard A host port 1x mini AB OTG port	Ρεύμα	350 mA	Διαστάσεις	78.74 x 76.2mm	<table border="1"> <tbody> <tr> <td>Επεξεργαστής</td> <td>Dual ARM Cortex-A15 + Dual ARM M4 (212 MHz) + Quad PRU (200 MHz)</td> </tr> <tr> <td>Μνήμη (SDRAM)</td> <td>2048 MiB DDR3L</td> </tr> <tr> <td>DSP</td> <td>Dual TMS320C66x (700 MHz)</td> </tr> <tr> <td>Αποθήκευση στην Πλακέτα</td> <td>8-bit eMMC 4 GB, microSD card</td> </tr> <tr> <td>Δίκτυο</td> <td>Dual Gigabit Ethernet</td> </tr> <tr> <td>USB θύρες</td> <td>3 x USB 3.0 Type A Host 4 x USB 2.0 Host 1 x micro USB Type B</td> </tr> <tr> <td>Ρεύμα</td> <td>5A @12 V</td> </tr> <tr> <td>Διαστάσεις</td> <td>107 mm × 102 mm</td> </tr> </tbody> </table>	Επεξεργαστής	Dual ARM Cortex-A15 + Dual ARM M4 (212 MHz) + Quad PRU (200 MHz)	Μνήμη (SDRAM)	2048 MiB DDR3L	DSP	Dual TMS320C66x (700 MHz)	Αποθήκευση στην Πλακέτα	8-bit eMMC 4 GB, microSD card	Δίκτυο	Dual Gigabit Ethernet	USB θύρες	3 x USB 3.0 Type A Host 4 x USB 2.0 Host 1 x micro USB Type B	Ρεύμα	5A @12 V	Διαστάσεις	107 mm × 102 mm
Επεξεργαστής	AM335x 1GHz ARM® Cortex-A8																																	
Μνήμη (SDRAM)	128 MiB (rev B) DDR 256 MiB (rev C+) DDR																																	
DSP	TMS320C64x+ (520 MHz)																																	
Αποθήκευση στην Πλακέτα	256MB NAND Flash, SD/MMC card																																	
Δίκτυο	-																																	
USB θύρες	1 x Standard A host port 1x mini AB OTG port																																	
Ρεύμα	350 mA																																	
Διαστάσεις	78.74 x 76.2mm																																	
Επεξεργαστής	Dual ARM Cortex-A15 + Dual ARM M4 (212 MHz) + Quad PRU (200 MHz)																																	
Μνήμη (SDRAM)	2048 MiB DDR3L																																	
DSP	Dual TMS320C66x (700 MHz)																																	
Αποθήκευση στην Πλακέτα	8-bit eMMC 4 GB, microSD card																																	
Δίκτυο	Dual Gigabit Ethernet																																	
USB θύρες	3 x USB 3.0 Type A Host 4 x USB 2.0 Host 1 x micro USB Type B																																	
Ρεύμα	5A @12 V																																	
Διαστάσεις	107 mm × 102 mm																																	
	<table border="1"> <tbody> <tr> <td>Επεξεργαστής</td> <td>Dual ARM Cortex-A15 + Dual ARM M4 (212 MHz) + Quad PRU (200 MHz)</td> </tr> <tr> <td>Μνήμη (SDRAM)</td> <td>2048 MiB DDR3L</td> </tr> <tr> <td>DSP</td> <td>Dual TMS320C66x (700 MHz)</td> </tr> <tr> <td>Αποθήκευση στην Πλακέτα</td> <td>8-bit eMMC 4 GB, microSD card</td> </tr> <tr> <td>Δίκτυο</td> <td>Dual Gigabit Ethernet</td> </tr> <tr> <td>USB θύρες</td> <td>3 x USB 3.0 Type A Host 4 x USB 2.0 Host 1 x micro USB Type B</td> </tr> <tr> <td>Ρεύμα</td> <td>5A @12 V</td> </tr> <tr> <td>Διαστάσεις</td> <td>107 mm × 102 mm</td> </tr> </tbody> </table>	Επεξεργαστής	Dual ARM Cortex-A15 + Dual ARM M4 (212 MHz) + Quad PRU (200 MHz)	Μνήμη (SDRAM)	2048 MiB DDR3L	DSP	Dual TMS320C66x (700 MHz)	Αποθήκευση στην Πλακέτα	8-bit eMMC 4 GB, microSD card	Δίκτυο	Dual Gigabit Ethernet	USB θύρες	3 x USB 3.0 Type A Host 4 x USB 2.0 Host 1 x micro USB Type B	Ρεύμα	5A @12 V	Διαστάσεις	107 mm × 102 mm	<table border="1"> <tbody> <tr> <td>Επεξεργαστής</td> <td>Dual ARM Cortex-A15 + Dual ARM M4 (212 MHz) + Quad PRU (200 MHz)</td> </tr> <tr> <td>Μνήμη (SDRAM)</td> <td>2048 MiB DDR3L</td> </tr> <tr> <td>DSP</td> <td>Dual TMS320C66x (700 MHz)</td> </tr> <tr> <td>Αποθήκευση στην Πλακέτα</td> <td>8-bit eMMC 4 GB, microSD card</td> </tr> <tr> <td>Δίκτυο</td> <td>Dual Gigabit Ethernet</td> </tr> <tr> <td>USB θύρες</td> <td>3 x USB 3.0 Type A Host 4 x USB 2.0 Host 1 x micro USB Type B</td> </tr> <tr> <td>Ρεύμα</td> <td>5A @12 V</td> </tr> <tr> <td>Διαστάσεις</td> <td>107 mm × 102 mm</td> </tr> </tbody> </table>	Επεξεργαστής	Dual ARM Cortex-A15 + Dual ARM M4 (212 MHz) + Quad PRU (200 MHz)	Μνήμη (SDRAM)	2048 MiB DDR3L	DSP	Dual TMS320C66x (700 MHz)	Αποθήκευση στην Πλακέτα	8-bit eMMC 4 GB, microSD card	Δίκτυο	Dual Gigabit Ethernet	USB θύρες	3 x USB 3.0 Type A Host 4 x USB 2.0 Host 1 x micro USB Type B	Ρεύμα	5A @12 V	Διαστάσεις	107 mm × 102 mm
Επεξεργαστής	Dual ARM Cortex-A15 + Dual ARM M4 (212 MHz) + Quad PRU (200 MHz)																																	
Μνήμη (SDRAM)	2048 MiB DDR3L																																	
DSP	Dual TMS320C66x (700 MHz)																																	
Αποθήκευση στην Πλακέτα	8-bit eMMC 4 GB, microSD card																																	
Δίκτυο	Dual Gigabit Ethernet																																	
USB θύρες	3 x USB 3.0 Type A Host 4 x USB 2.0 Host 1 x micro USB Type B																																	
Ρεύμα	5A @12 V																																	
Διαστάσεις	107 mm × 102 mm																																	
Επεξεργαστής	Dual ARM Cortex-A15 + Dual ARM M4 (212 MHz) + Quad PRU (200 MHz)																																	
Μνήμη (SDRAM)	2048 MiB DDR3L																																	
DSP	Dual TMS320C66x (700 MHz)																																	
Αποθήκευση στην Πλακέτα	8-bit eMMC 4 GB, microSD card																																	
Δίκτυο	Dual Gigabit Ethernet																																	
USB θύρες	3 x USB 3.0 Type A Host 4 x USB 2.0 Host 1 x micro USB Type B																																	
Ρεύμα	5A @12 V																																	
Διαστάσεις	107 mm × 102 mm																																	

	<table border="1"><tr><td>Επεξεργαστής</td><td>AM335x 1GHz ARM® Cortex-A8</td></tr><tr><td>Μνήμη (SDRAM)</td><td>512MB DDR3 RAM</td></tr><tr><td>DSP</td><td>-</td></tr><tr><td>Αποθήκευση στην Πλακέτα</td><td>8-bit eMMC microSD card 3.3 V Supported</td></tr><tr><td>Δίκτυο</td><td>Fast Ethernet (MII based)</td></tr><tr><td>USB θύρες</td><td>1 x Standard A host port 1x mini B device port</td></tr><tr><td>Ρεύμα</td><td>210–460 mA @5 V</td></tr><tr><td>Διαστάσεις</td><td>86.40 mm × 53.3 mm</td></tr></table>	Επεξεργαστής	AM335x 1GHz ARM® Cortex-A8	Μνήμη (SDRAM)	512MB DDR3 RAM	DSP	-	Αποθήκευση στην Πλακέτα	8-bit eMMC microSD card 3.3 V Supported	Δίκτυο	Fast Ethernet (MII based)	USB θύρες	1 x Standard A host port 1x mini B device port	Ρεύμα	210–460 mA @5 V	Διαστάσεις	86.40 mm × 53.3 mm
Επεξεργαστής	AM335x 1GHz ARM® Cortex-A8																
Μνήμη (SDRAM)	512MB DDR3 RAM																
DSP	-																
Αποθήκευση στην Πλακέτα	8-bit eMMC microSD card 3.3 V Supported																
Δίκτυο	Fast Ethernet (MII based)																
USB θύρες	1 x Standard A host port 1x mini B device port																
Ρεύμα	210–460 mA @5 V																
Διαστάσεις	86.40 mm × 53.3 mm																

Πίνακας 1: Μικροϋπολογιστές της οικογένειας BeagleBoard.

2.1.2 Λογισμικό

Όσον αφορά το λογισμικό, δεν υπάρχει ενιαία συλλογή λογισμικού για το μικροϋπολογιστή BeagleBoard καθώς ο ίδιος δε διαθέτει κάποιο λειτουργικό σύστημα ή λογισμικό εφαρμογών. Ο μικροϋπολογιστής υποστηρίζει τη χρήση αρκετών λειτουργικών συστημάτων και εφαρμογών που προέρχονται από την ανοιχτή κοινότητα λογισμικών. Από τα λογισμικά τα οποία μπορούν να χρησιμοποιηθούν είναι:

- **Ångström:** Το λογισμικό αυτό αποτελεί μια διανομή Linux για μια ποικιλία ενσωματωμένων συσκευών. Η διανομή του λογισμικού αυτού είναι μια συνεργασία με το Poky Linux. Το Ångström βασίζεται στο έργο OpenEmbedded, συγκεκριμένα στο στρώμα OpenEmbedded Core (OE-Core) και διαθέτει μια σειρά εργαλείων, η οποία είναι και επίσημα συνδεδεμένη με το Poky Linux.

Το Ångström διαφέρει, κυρίως, από το Poky Linux ως δυαδική διανομή (όπως π.χ. οι διανομές Debian, Fedora, OpenSuse ή Ubuntu GNU / Linux), χρησιμοποιώντας `orkg` για τη διαχείριση πακέτων. Ως εκ τούτου, ένα βασικό μέρος της κατασκευής του Ångström είναι μια δυαδική δέσμη πακέτων, που επιτρέπει την εγκατάσταση του λογισμικού που διανέμει ως πακέτα `orkg`, χωρίς να χρειάζεται να μεταγλωττιστούν πρώτα. [6]

- **Ubuntu:** Οι εκδόσεις που υποστηρίζονται από τον εν λόγω μικροϋπολογιστή είναι αρκετές, ανάλογα με τις ανάγκες της εκάστοτε εφαρμογής που θα αναπτυχθεί. Η κυριότερη έκδοση η οποία χρησιμοποιείται κατά κόρον μέχρι και σήμερα είναι η έκδοση 16.04.
- **Debian:** Και από το λειτουργικό αυτό σύστημα μπορεί να χρησιμοποιηθεί η έκδοση που απαιτείται ανάλογα με την εφαρμογή η οποία θα αναπτυχθεί. Από την προγραμματιστική κοινότητα προτείνεται να χρησιμοποιηθεί η έκδοση 9.

Για τις ανάγκες υλοποίησης της παρούσας εργασίας, χρησιμοποιήθηκε η έκδοση Ubuntu 12.04 η οποία είναι επίσης εγκατεστημένη στον υπολογιστή και δοκιμασμένη με τη λειτουργία του λογισμικού OpenCV. Ωστόσο δοκιμάστηκε και το λειτουργικό σύστημα Ångström το οποίο δεν μπόρεσε να υποστηρίξει την εν λόγω βιβλιοθήκη.

2.2 C++

Η εφαρμογή που αναπτύχθηκε στην παρούσα διπλωματική εργασία βασίστηκε στη βιβλιοθήκη OpenCV για την οποία θα μιλήσουμε αναλυτικά σε επόμενο κεφάλαιο. Ξεκινώντας την ανάπτυξη μιας εφαρμογής που απαιτεί τη χρήση της εν λόγω βιβλιοθήκης, είναι απαραίτητο να γίνει επιλογή της γλώσσας προγραμματισμού που θα χρησιμοποιηθεί. Υπάρχουν δύο πιθανές γλώσσες προγραμματισμού γι' αυτό το σκοπό, η C++ και η Python. Η γλώσσα που τελικά χρησιμοποιήθηκε είναι η C++ για πολλούς λόγους.

Η Python σαν γλώσσα προγραμματισμού είναι βραδύτερη από την C++. Αυτό γιατί δεν χρειάζεται μεταγλώττιση σε αντίθεση με την C++ η οποία μεταγλωττίζεται σε γλώσσα μηχανής και έτσι καταφέρνει να προσφέρει πλεονεκτήματα ως προς την απόδοση. Η Python επιλέγεται γενικά για την απλότητά και τη φορητότητα που προσφέρει αλλά και για την ενίσχυση της δημιουργικότητας των προγραμματιστών, καθώς με αυτήν είναι απαλλαγμένοι από την γνωστική πολυπλοκότητα και επικεντρώνονται αποκλειστικά στους αλγόριθμους που αναπτύσσουν.

Με το OpenCV, όμως, τα πράγματα είναι διαφορετικά. Ο συνδυασμός Python-OpenCV λειτουργεί κάπως σαν περιτύλιγμα γύρω από τον πρωτότυπο κώδικα C/C++. Αυτό συμβαίνει γιατί η βιβλιοθήκη OpenCV είναι γραμμένη σε γλώσσες C και C++. Συνεπώς ο παραπάνω συνδυασμός χρησιμοποιείται για να συνυπάρξουν τα πλεονεκτήματα και των δύο γλωσσών, η υψηλή απόδοση των C/C++ και η απλότητα της Python. Έτσι, όταν καλείται μία συνάρτηση του OpenCV με Python, αυτό που στην ουσία εκτελείται από πίσω υπόκειται σε κώδικα C/C++. Συνεπώς η διαφορά στην απόδοση υπάρχει αλλά δεν είναι μεγάλη, με τη χειρότερη περίπτωση, υπολογισμένη κατά προσέγγιση σύμφωνα με μερικές βασικές λειτουργίες του OpenCV, να έχει ένα μειονέκτημα της τάξης του <4%, όπου το μειονέκτημα αυτό προκύπτει από τον τύπο:

μέγιστος χρόνος που καταναλώνεται σε Python – ελάχιστος χρόνος που καταναλώνεται σε C++ / ελάχιστος χρόνος που καταναλώνεται σε C++

Το μεγαλύτερο πρόβλημα προκύπτει όταν ο κώδικας που αναπτύσσεται για τη εφαρμογή αποτελείται από πολλές εγγενείς Python συναρτήσεις. Αν, δηλαδή, οι λειτουργίες που δεν είναι διαθέσιμες από το OpenCV, επιλεγεί να αναπτυχθούν σε Python, τότε ο κώδικας αυτός θα εκτελεστεί εγγενώς σε Python και η διαφορά στην απόδοση θα είναι πλέον εμφανής.

Αξίζει να σημειωθεί ότι υπάρχει μία τεράστια κοινότητα από προγραμματιστές με χιλιάδες μέλη, περί των 47 χιλιάδων, που χρησιμοποιούν και υποστηρίζουν το OpenCV. Η βιβλιοθήκη έχει κατέβει στους υπολογιστές των χρηστών πάνω από 9 εκατομμύρια φορές. Η κοινότητα αποτελείται από μέλη που ανήκουν σε πολλούς τομείς και βιομηχανίες και η ανάπτυξη της βιβλιοθήκης χρηματοδοτείται από επιχειρήσεις όπως η Intel, η AMD και η Google. Συνεπώς, τα εγχειρίδια χρήσης της βιβλιοθήκης, των συναρτήσεων της και της γλώσσας προγραμματισμού στην οποία είναι γραμμένη είναι άπειρα και πλήρως ενημερωμένα. [7]

2.3 Matlab

Η γλώσσα προγραμματισμού Matlab [8] άρχισε να αναπτύσσεται στα τέλη του 1970 από τον Cleve Moler, διευθυντή του τμήματος πληροφορικής του Πανεπιστημίου του New Mexico. Ο Moler σχεδίασε τη συγκεκριμένη γλώσσα για να δώσει τη δυνατότητα στους φοιτητές να έχουν πρόσβαση και να μπορούν να χρησιμοποιούν τις βιβλιοθήκες LINPACK και EISPACK, βιβλιοθήκες γραμμένες στη γλώσσα προγραμματισμού Fortran, χωρίς οι ίδιοι να χρειάζεται να μάθουν τη συγκεκριμένη γλώσσα. Γρήγορα η φήμη της πλατφόρμας εξαπλώθηκε στα άλλα Πανεπιστήμια και άρχισε να γνωρίζει μεγάλη απήχηση στην κοινότητα των εφαρμοσμένων μαθηματικών. Το 1984 ο Moler σε συνεργασία με τον μηχανικό Jack Little ξαναέγραψαν τη γλώσσα προγραμματισμού Matlab σε C και ίδρυσαν την εταιρία MathWorks, εταιρία που δραστηριοποιείται γύρω από τα λογισμικά μαθηματικής υπολογιστικής. Το 2000, η Matlab ξαναγράφηκε για να μπορεί να χρησιμοποιεί το πλέον σύγχρονο σύνολο βιβλιοθηκών για τον χειρισμό πινάκων.

Το περιβάλλον Matlab (matrix laboratory) αποτελεί ένα περιβάλλον αριθμητικής υπολογιστικής που υποστηρίζει πολλαπλές προγραμματιστικές τεχνολογίες. Επιπλέον, είναι μία τέταρτης γενιάς, κατοχυρωμένη γλώσσα προγραμματισμού που προσφέρει χειρισμό πινάκων, γραφική απεικόνιση συναρτήσεων και δεδομένων, εφαρμογές αλγορίθμων, δημιουργία διαφορετικών περιβαλλόντων για τους χρήστες καθώς και διασύνδεση με άλλες γλώσσες προγραμματισμού όπως τη C, τη C++, τη C#, τη Java, τη Fortran και την Python. Τα ενσωματωμένα γραφικά του περιβάλλοντος Matlab, καθιστούν εύκολη την οπτικοποίηση και εξοικείωση με τα δεδομένα. Όλες οι παραπάνω δυνατότητες έχουν δημιουργηθεί και ελεγχθεί αυστηρά για να λειτουργούν μαζί και αποδοτικά. [9]

Πέρα από τα παραπάνω πλεονεκτήματα του περιβάλλοντος Matlab, αυτό περιλαμβάνει ένα Σύστημα Εργαλείων Όρασης Υπολογιστή (Computer Vision System Toolbox), το οποίο υπήρξε πολύ χρήσιμο για την παρούσα διπλωματική εργασία. Η συγκεκριμένη εργαλειοθήκη προσφέρει συναρτήσεις, αλγόριθμους και εφαρμογές για το σχεδιασμό και την προσομοίωση συστημάτων μηχανικής όρασης και επεξεργασίας βίντεο. Δίνει τη δυνατότητα ανίχνευσης, εξαγωγής και ταύτισης γνωρισμάτων, αναγνώρισης αντικειμένων, προσδιορισμού κίνησης κ.α. Επιπλέον, βρίσκει εφαρμογή

και στον τομέα της τρισδιάστατης όρασης μηχανής. Τέλος, μέσω πλαισίων προσανατολισμένων στην εκπαίδευση μηχανής, μπορεί να γίνει εκπαίδευση συστημάτων για ανίχνευση και αναγνώριση αντικειμένων και ανάκτηση εικόνας.

Όσον αφορά στην εκπαίδευση μηχανής συγκεκριμένα, το παραπάνω εργαλείο προσφέρει εκπαιδευμένους ταξινομητές, έτοιμους να χρησιμοποιηθούν για αναγνώριση προτύπων, αλλά και τη δυνατότητα στους χρήστες να εκπαιδεύσουν κάποιον ταξινομητή δικής τους προτίμησης. Επιπλέον, ο χρήστης έχει τη δυνατότητα να χρησιμοποιήσει αυτούς τους ταξινομητές μέσω της ήδη έτοιμης εφαρμογής για αναγνώριση αντικειμένων, η οποία χρησιμοποιεί τον αλγόριθμο Viola-Jones και ανιχνεύει πρόσωπα, μάτια, μύτη, στόμα ή το επάνω μέρος του ανθρώπινου κορμού. Στην παρούσα διπλωματική εργασία το εργαλείο αυτό χρησιμοποιήθηκε για να εκπαιδευτούν συγκεκριμένοι ταξινομητές για χαμόγελο, λύπη και ανοιχτό στόμα και να χρησιμοποιηθούν στον αλγόριθμο που αναπτύχτηκε. Για να παραχθούν τα τελικά αρχεία που χρησιμοποιήθηκαν για την εκπαίδευση, έγινε εκμετάλλευση, επίσης, του εργαλείου Training Image Labeler για να σημειωθεί η περιοχή ενδιαφέροντος στις εικόνες που συλλέχτηκαν. [8]

2.4 Motion

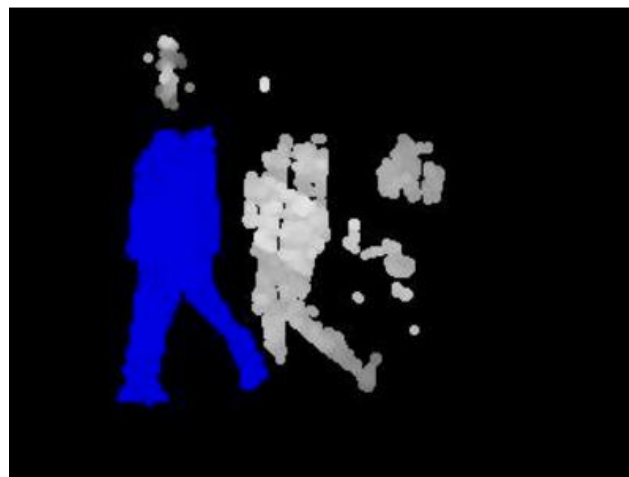
Το Motion είναι μία εφαρμογή χαμηλών σε υπολογιστική ισχύ απαιτήσεων, αλλά παρόλα αυτά μπορεί να χρησιμοποιηθεί από κλειστά συστήματα παρακολούθησης μέσω λογισμικού Linux. Δύναται να λειτουργήσει με οποιαδήποτε κάμερα συμβατή με το λογισμικό αυτό, συμπεριλαμβανομένων όλων των V4L (Video4Linux) Webcam, πολλών IP καμερών, Axis καμερών και υποστηρίζει λειτουργίες περιστροφής της κάμερας δεξιά, αριστερά, πάνω και κάτω. [10]

Η κύρια λειτουργία της εφαρμογής Motion είναι η ανίχνευση κίνησης. Στην Εικόνα 1 φαίνεται ένα στιγμιότυπο κίνησης ενός ανθρώπου που ανιχνεύτηκε από κάμερα. Η Εικόνα 2 είναι η ίδια, με τη διαφορά ότι έχει εφαρμοστεί πάνω της φίλτρο για την απομάκρυνση θορύβου, σκόνης κ.α. (despeckle) από το Motion. Η μπλε περιοχή είναι η μεγαλύτερη και είναι και αυτή που γίνεται αντιληπτή από το Motion ως κίνηση. Στην

εικόνα αυτή φαίνεται πως το Motion διατηρεί ένα πλαίσιο αναφοράς, το οποίο δεν αποτελείται μόνο από το τελευταίο στιγμιότυπο, αλλά από ένα μαθηματικό υπολογισμό προηγούμενων εικόνων. Με τον τρόπο αυτό διευρύνεται η ανίχνευση κίνησης και είναι αρκετά δύσκολο να γίνει παράβλεψη αυτής.



Εικόνα 1: Στιγμιότυπο που καταγράφηκε από κάμερα πριν οποιαδήποτε επεξεργασία.



Εικόνα 2: Στιγμιότυπο μετά την εφαρμογή του φίλτρου despeckle από το Motion. Μόνο η μπλε περιοχή αναγνωρίζεται ως κίνηση.

Για την καλύτερη αξιοποίηση των αποτελεσμάτων, το Motion προσφέρει την επιλογή `ffmpeg` μέσω της οποίας γίνεται παραγωγή MPEG βίντεο πολύ γρήγορα, μόλις ανιχνευτεί οποιαδήποτε κίνηση. Επιπλέον, λειτουργεί σαν μία κανονική κάμερα και μπορεί να τραβάει φωτογραφίες ανά τακτά χρονικά διαστήματα. Υπάρχει η δυνατότητα εμπλουτισμού αυτών των στιγμιότυπων με κάποιο κείμενο επιλογής του χρήστη, που να

επιδεικνύει π.χ. την ώρα λήψης ή και άλλες χρήσιμες πληροφορίες και αποθήκευσής τους με όνομα επίσης επιλογής του χρήστη.

Τέλος, η πιο χρήσιμη λειτουργία του Motion για την παρούσα διπλωματική εργασία αποτέλεσε η συμπεριφορά του σε webcam server. Η μετάδοση του βίντεο γίνεται σε μορφή multipart jpeg (mjpeg), μέσω μίας προκαθορισμένης πόρτας, της 8080. Σε περίπτωση που υπάρχουν πολλαπλά παράλληλα νήματα (thread) ο χρήστης μπορεί να ορίσει τις πόρτες για τη μετάδοση βίντεο του καθενός. Το βίντεο αυτό μπορεί να προβληθεί από Mozilla Firefox περιηγητή ή από μία εφαρμογή της Java, το Combozola. Το βίντεο αυτό, φορτώνεται μέσω της σχετικής ηλεκτρονικής διεύθυνσης στην εφαρμογή που κατασκευάστηκε, ώστε να πραγματοποιείται η οποιαδήποτε αναγνώριση πάνω σε αυτό. [11]

2.5 Νήματα (Threads)

Ως νήμα (thread) χαρακτηρίζεται μία ανεξάρτητη ακολουθία εντολών και ελέγχων εντός προγράμματος. Η έννοια του νήματος πολλές φορές συγχέεται με αυτή της διεργασίας, καθώς τα νήματα αναφέρονται συχνά σαν ελαφριές διεργασίες. Ωστόσο η βασική διαφορά ανάμεσά τους είναι ότι τα νήματα μπορούν και μοιράζονται τον ίδιο χώρο διευθύνσεων, ενώ οι διεργασίες συνήθως όχι. Έτσι, δύο διαφορετικά νήματα του ίδιου προγράμματος έχουν δυνατότητα πρόσβασης σε κάποιο αντικείμενο αυτού, ενώ δύο διαφορετικές διεργασίες συνήθως βλέπουν διαφορετικά αντίγραφα του ίδιου αντικειμένου. Μία διεργασία μπορεί να διαθέτει πολλαπλά νήματα, τα οποία περιέχουν τον κώδικά προς εκτέλεση της διεργασίας, καθώς και το εννοιολογικό της πλαίσιο, δηλαδή τις τιμές των μεταβλητών σε οποιαδήποτε χρονική στιγμή. [12]

Σε έναν απλό επεξεργαστή, η πολυνημάτωση (multithreading) πραγματοποιείται με τη μέθοδο της πολυπλεξίας με διαίρεση χρόνου, δηλαδή, πιο συγκεκριμένα, ο επεξεργαστής μεταπηδά μεταξύ των διάφορων νημάτων. Αυτή η εναλλαγή μεταξύ των διεργασιών ονομάζεται μεταγωγή περιβάλλοντος, και συμβαίνει σε πολύ τακτά χρονικά διαστήματα, τέτοια έτσι ώστε ο χρήστης έχει την εντύπωση ότι τα νήματα εκτελούνται

την ίδια στιγμή. Μόνο σε έναν επεξεργαστή με πολλούς επεξεργαστικούς πυρήνες, τα νήματα εκτελούνται πραγματικά ταυτόχρονα και κάθε πυρήνας εκτελεί ένα συγκεκριμένο νήμα ή εργασία.

Ο πολυνηματικός προγραμματισμός είναι ένας καλός τρόπος να βελτιωθεί η εκτέλεση κάποιων προγραμμάτων. Τα πλεονεκτήματα του πολυνηματισμού περιλαμβάνουν:

- Πιο αποτελεσματική χρήση των διαθέσιμων πόρων.
- Απλούστερα προγράμματα κατά περιπτώσεις.
- Καλύτερη απόκριση προγραμμάτων.

Έστω ότι έχει αναπτυχθεί μία εφαρμογή η οποία διαβάζει δεδομένα από αρχεία που βρίσκονται στο δίσκο και μετά τα επεξεργάζεται. Από σχεδιαστικής άποψης, είναι ευκολότερο να δημιουργηθούν δύο νήματα καθένα από τα οποία θα εκτελεί μία από αυτές τις εργασίες παρά να υπήρχε ένα μοναδικό νήμα το οποίο θα έπρεπε να παρακολουθεί τις διαδικασίες ανάγνωσης όλων των αρχείων. Τα δύο νήματα μπορούν να ξεκινήσουν μαζί και το καθένα να περιμένει να ελευθερωθεί ο δίσκος για να τον χρησιμοποιήσει. Με τον τρόπο αυτό επιτυγχάνεται καλύτερη χρήση τόσο του δίσκου όσο και της CPU.

Η βελτίωση της απόκρισης μιας εφαρμογής είναι ακόμη ένας λόγος για να μετατραπεί σε πολυνηματική. Πιο συγκεκριμένα, στη περίπτωση που μία εφαρμογή περιλαμβάνει μία χρονοβόρα εργασία, το βέλτιστο θα ήταν να ανατεθεί αυτή σε ένα νήμα εργασίας. Αποτέλεσμα αυτής της αλλαγής θα είναι η εφαρμογή να ελευθερώνεται για άλλες εργασίες, καθώς το εν λόγω νήμα εργασίας είναι απασχολημένο.

Εκτός από πλεονεκτήματα, βέβαια, στον πολυνηματισμό υπάρχουν και μειονεκτήματα. Οι εφαρμογές που τον εφαρμόζουν είναι επιρρεπείς στην ύπαρξη λαθών. Αυτά έχουν να κάνουν κυρίως με την πρόσβαση σε διαμοιραζόμενους πόρους όπως μνήμη, αρχεία, βάσεις δεδομένων κ.α. Επιπλέον, πρέπει να λαμβάνονται πάντα υπόψη οι επιπτώσεις στην απόδοση που προκύπτουν από τις εναλλαγές περιβάλλοντος αλλά και ο χώρος μνήμης και οι πόροι του λειτουργικού συστήματος που απαιτεί ένα νήμα για τα δεδομένα και τη διαχείρισή του. [13]

Στην παρούσα διπλωματική εργασία η δημιουργία νημάτων καθίσταται απαραίτητη, καθώς είναι απαραίτητο να εκτελούνται ταυτόχρονα κάποιες λειτουργίες του συστήματος, χωρίς η εφαρμογή να μένει εν αναμονή ολοκλήρωσής τους.

2.6 Βιβλιοθήκη Mpeg123

Η διανομή του mpeg123 περιλαμβάνει έναν πραγματικού χρόνου αναπαραγωγή-αποκωδικοποιητή ήχου για τα επίπεδα 1, 2 και 3, με πιο συνηθισμένο τον MPEG 1.0 επιπέδου 3, δηλαδή MP3. Επιπλέον περιλαμβάνει δύο επαναχρησιμοποιήσιμες βιβλιοθήκες, αποκωδικοποίησης και εξαγωγής, την libmpeg123 και την libout123, αντίστοιχα. Μεταξύ άλλων, λειτουργεί σε GNU/Linux, MacOSX, BSDs, Solaris, AIX, HPUX, SGI Irix, OS/2 και Cygwin ή και σε MS Windows. Είναι ελεύθερο λογισμικό με άδεια LGPL 2.1.

Στην παρούσα διπλωματική χρησιμοποιήθηκαν η βιβλιοθήκη libmpeg123 του λογισμικού και μία αντίστοιχης λειτουργίας με τη libout123, η libao για την ανάπτυξη αλγορίθμου αναπαραγωγής MP3 ήχων.

Η βιβλιοθήκη libmpeg123 μεταξύ των άλλων προσφέρει [14]:

- Αποκωδικοποίηση των ακουστικών ρευμάτων MPEG 1.0/2.0/2.5 επιπέδου 1, 2 και 3 σε PCM
- Αναζήτηση ακριβούς δείγματος και αναπαραγωγή χωρίς κενά μέσω του LAME πλαισίου πληροφοριών
- Επιλογές για την κωδικοποίηση του δείγματος, από οκτώ bit μlaw σε τριάντα δύο bit ακολουθία
- Assembly/SIMD βελτιστοποιήσεις για τις λειτουργίες κωδικοποίησης και αποκωδικοποίησης των δειγμάτων
- Επιλογές μεταξύ ακρίβειας και αποδοτικότητας
- Διακριτικό χειρισμό των λαθών στα ρεύματα
- Συντελεστές προσαύξησης συχνότητας ήχου για ισοσταθμιστές χωρίς κόστος
- Ένταση ήχου μέσω συντελεστή προσαύξησης χωρίς επιβάρυνση της CPU

- Απενεργοποίηση χαρακτηριστικών για τη δημιουργία μικρότερων εκδόσεων για τα ενσωματωμένα συστήματα

Η libao είναι μία διαπλατφορμική βιβλιοθήκη που επιτρέπει σε αλγορίθμους να αναπαράγουν PCM δεδομένα ήχου μέσω των συνδεδεμένων ακουστικών συσκευών. Αυτή τη στιγμή υποστηρίζει [15]:

- OSS (Open Sound System)
- ESD (ESound)
- ALSA (Advanced Linux Sound Architecture)
- Sun audio system (χρησιμοποιείται σε Solaris, OpenBSD και NetBSD)
- aRts (Analog Realtime Synthesizer)
- PulseAudio
- RoarAudio
- OpenBSD sndio
- Windows MMISound
- IRIX
- MacOS X (AU HAL)
- NAS
- AIX

2.7 cURL

Το cURL [16] είναι λογισμικό που προσφέρει μια βιβλιοθήκη και ένα εργαλείο γραμμής εντολών για μεταφορά δεδομένων χρησιμοποιώντας ποικίλα πρωτόκολλα. Το όνομά του προέρχεται από τη φράση «see URL». Η πρώτη του έκδοση του έγινε το 1997 και ο συντάκτης και βασικός του δημιουργός ήταν ο Σουηδός προγραμματιστής Daniel Stenberg.

Οι βασικές λειτουργίες που υποστηρίζει το cURL είναι η αποστολή και λήψη αρχείων χρησιμοποιώντας URL σύνταξη. Δεδομένου ότι το cURL χρησιμοποιεί τη βιβλιοθήκη libcurl, υποστηρίζει μία σειρά από πρωτόκολλα διαδικτύου,

συμπεριλαμβανομένων των HTTP, HTTPS, FTP, FTPS, SCP, SFTP, TFTP, LDAP, DAP, DICT, TELNET, FILE, IMAP, POP3, SMTP και RTSP. Όταν χρησιμοποιείται με κάποιο πρωτόκολλο ασφαλείας, όπως το HTTPS, εφαρμόζει SSL επαλήθευση πιστοποιητικών από προεπιλογή. Όταν συνδέεται σε κάποιον απομακρυσμένο διακομιστή μέσω του HTTPS, αρχικά λαμβάνει το πιστοποιητικό του διακομιστή και στη συνέχεια ελέγχει την εγκυρότητά του σύμφωνα με τις αρχές πιστοποίησης που διαθέτει, για να βεβαιωθεί ότι ο διακομιστής είναι όντως αυτός που υποστηρίζει ότι είναι. Το cURL επιστρέφει μήνυμα λάθους αν ο απομακρυσμένος διακομιστής χρησιμοποιεί πιστοποιητικό υπογεγραμμένο από τον εαυτό του ή αν αυτό δεν συμπεριλαμβάνεται στις αρχές πιστοποίησης που διαθέτει. Σε αντίθετη περίπτωση, δηλαδή αν ο απομακρυσμένος διακομιστής είναι πιστοποιημένος, οι αρχές πιστοποίησης του cURL ανανεώνονται και ο διακομιστής πλέον συμπεριλαμβάνεται σε αυτές. [16]

Στην εφαρμογή που υλοποιήθηκε το cURL χρησιμοποιήθηκε για την αποστολή και λήψη δεδομένων σε συγκεκριμένα αρχεία που βρίσκονται σε απομακρυσμένο εξυπηρετητή. Η αποστολή των δεδομένων αυτών γίνεται με τη μέθοδο POST.

POST

Η μέθοδος POST είναι μία μέθοδος αίτησης που υποστηρίζεται από το HTTP πρωτόκολλο. Ο ρόλος της POST μεθόδου ανάμεσα στο πλήθος των HTTP μεθόδων είναι να αντιπροσωπεύσει μία οντότητα στον εξυπηρετητή με σκοπό αυτή να αποθηκευτεί σαν υφιστάμενη του πόρου που προσδιορίζεται από το URI. Στην μέθοδο POST τα δεδομένα δεν αποστέλλονται σαν τμήμα της σειράς χαρακτήρων του URL, αλλά σαν μέρος του σώματος του μηνύματος.

Η μέθοδος POST προσφέρει ασφάλεια ακριβώς γιατί τα δεδομένα προς αποστολή δεν είναι ορατά στο URL. Ωστόσο, για περισσότερη ασφάλεια οι ακολουθίες χαρακτήρων μπορούν να κρυπτογραφηθούν χρησιμοποιώντας HTTP με SSL πρωτόκολλο. Συνεπώς, όλες οι ευαίσθητες και εμπιστευτικές πληροφορίες που αποστέλλονται σε κάποιον εξυπηρετητή πρέπει να χρησιμοποιούν αιτήσεις POST μέσω HTTPS. Σχεδόν όλα τα αιτήματα επιβεβαίωσης αποστέλλονται με τη μέθοδο POST στον κόσμο του HTTP. Επιπλέον, η POST χρησιμοποιείται για την υποβολή πληροφοριών σε

κεντρικούς υπολογιστές, για οποιεσδήποτε πληροφορίες που μπορούν να αλλάξουν την κατάσταση της εφαρμογής, όπως η προσθήκη προϊόντων στο καλάθι αγορών, η πραγματοποίηση πληρωμών κ.α. [25]

Σύνοψη κεφαλαίου

Στο κεφάλαιο αυτό αναφέρθηκαν και αναλύθηκαν όλες οι τεχνολογίες, υλικού και λογισμικού, που χρησιμοποιήθηκαν για την υλοποίηση της παρούσας διπλωματικής εργασίας, οι λόγοι για τους οποίους επιλέχτηκαν και οι στόχοι που τελικά επιτέλεσαν.

Κεφάλαιο 3: Μηχανική όραση και περιγραφείς χαρακτηριστικών

Στο ακόλουθο κεφάλαιο εισάγονται οι έννοιες της μηχανικής όρασης για να ενταχθεί ο αναγνώστης σε αυτές και να αντιληφθεί το υπόβαθρο της παρούσας διπλωματικής εργασίας. Πιο συγκεκριμένα, αναλύονται ο τρόπος ανάπτυξης, οι χρήσεις της μηχανικής όρασης, οι τομείς που επηρεάζονται από αυτή και οι βιβλιοθήκες που χρησιμοποιούνται για την ανάπτυξη εφαρμογών βασισμένων σε αυτή. Τέλος, γίνεται μία αναλυτική παρουσίαση δύο ειδών ταξινομητών για την ανίχνευση αντικειμένων και των χαρακτηριστικών που οδήγησαν στην επιλογή του τύπου που τελικά χρησιμοποιήθηκε στον αλγόριθμο που αναπτύχτηκε.

3.1 Όραση υπολογιστή

Η αίσθηση της όρασης θεωρείται η πιο σημαντική από τις πέντε αισθήσεις, καθώς δίνει τη δυνατότητα στους ανθρώπους να αντιλαμβάνονται και να κατανοούν τον κόσμο γύρω τους. Τη δυνατότητα αυτή επιδιώκει να αντιγράψει ή ακόμα και να ξεπεράσει η μηχανική ή υπολογιστική ή αλλιώς τεχνητή όραση. Αυτή αποτελεί ένα επιστημονικό πεδίο της τεχνητής νοημοσύνης που προσπαθεί αλγοριθμικά να αναπαραστήσει μία τρισδιάστατη σκηνή μέσω της επεξεργασίας των δισδιάστατων εικόνων που λαμβάνει σαν είσοδο από οπτικούς αισθητήρες, όπως κάμερες, με κόστος την απώλεια μεγάλου όγκου χρήσιμης πληροφορίας. [18]

Πιο συγκεκριμένα, ως όραση υπολογιστή ορίζεται η μετατροπή δεδομένων από μία κάμερα σταθερής εικόνας ή βίντεο είτε σε μία απόφαση είτε σε μία άλλου είδους απεικόνιση. Όλες αυτές οι μετατροπές γίνονται για την επίτευξη ενός συγκεκριμένου στόχου. Τα δεδομένα εισόδου μπορεί να περιέχουν κάποιες συναφείς πληροφορίες όπως

το ότι «η κάμερα είναι τοποθετημένη σε ένα αυτοκίνητο» ή το ότι «το λέιζερ εύρεσης απόστασης υποδεικνύει ότι υπάρχει αντικείμενο σε απόσταση ενός μέτρου». Η απόφαση που τελικά θα ληφθεί μπορεί να είναι του τύπου «υπάρχει ένα άτομο στην εικόνα» ή «υπάρχουν δεκατέσσερα καρκινικά κύτταρα σε αυτή την απεικόνιση». Μια νέα απεικόνιση μπορεί να σημαίνει τη μετατροπή μιας έγχρωμης εικόνας σε μία με χρώματα που ανήκουν στην κλίμακα του γκρι ή την αφαίρεση της κίνησης από μία ροή εικόνων.

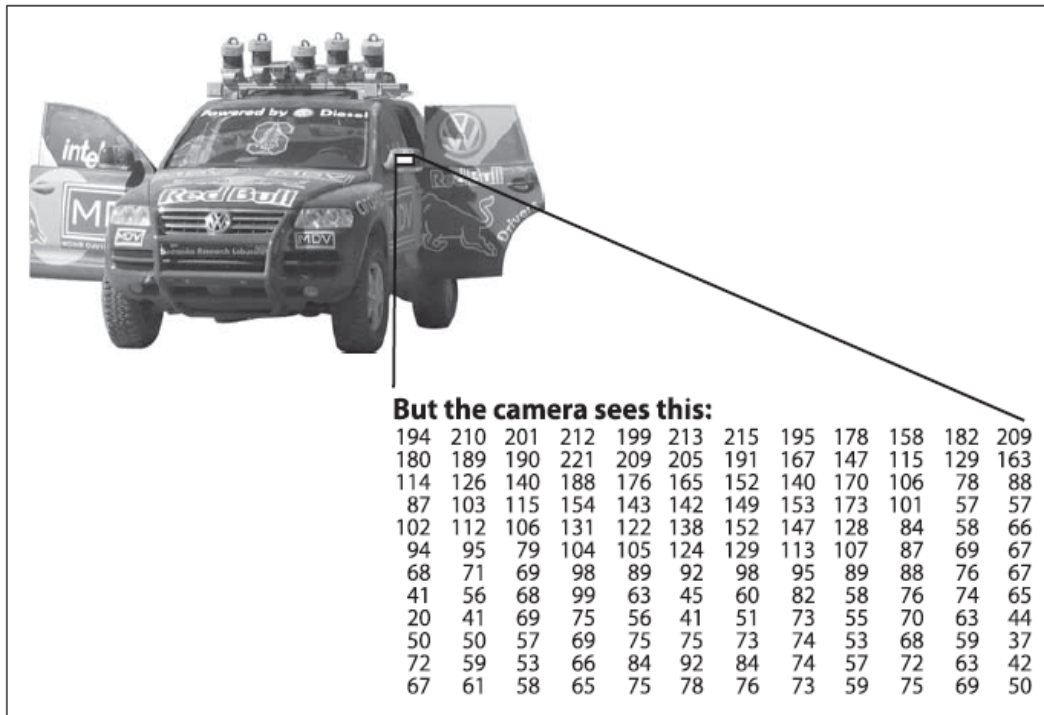
Η μηχανική όραση επιδιώκει να εφαρμόσει θεωρίες και μοντέλα για την κατασκευή μηχανικών συστημάτων με δυνατότητα όρασης. Παραδείγματα στα οποία τέτοια συστήματα βρίσκουν εφαρμογή είναι τα εξής:

1. Έλεγχος διαδικασιών (π.χ. ένα βιομηχανικό ρομπότ ή ένα αυτόνομο όχημα)
2. Ανίχνευση συμβάντων (π.χ. οπτική επιτήρηση)
3. Οργάνωση πληροφοριών (π.χ. ευρετηριοποίηση βάσεων δεδομένων και ακολουθιών εικόνων)
4. Εξομοίωση αντικειμένων και περιβαλλόντων (π.χ. βιομηχανική επιθεώρηση, ιατρική ανάλυση εικόνας ή τοπογραφική εξομοίωση)
5. Αλληλεπίδραση χρηστών με υπολογιστικά συστήματα (π.χ. ως είσοδος σε μια συσκευή επικοινωνίας ανθρώπου / μηχανής).

Οι άνθρωποι, ως οπτικά όντα, μπορεί να θεωρήσουν ότι οι εργασίες που είναι σχετικές με την όραση μηχανής πιθανότατα είναι εύκολες. Αυτό συμβαίνει γιατί εκ πρώτης όψεως είναι εύκολο για το ανθρώπινο μάτι να αντιληφτεί πού βρίσκεται, ένα αντικείμενο μέσα σε μία εικόνα. Όμως πίσω από αυτή τη φαινομενικά εύκολη δουλειά κρύβεται μια πολύπλοκη διαδικασία. Ο ανθρώπινος εγκέφαλος διαιρεί το οπτικό σήμα σε πολλά κανάλια μέσω των οποίων ρέουν διαφορετικά είδη πληροφοριών. Ο εγκέφαλος διαθέτει ένα σύστημα που αναγνωρίζει, με τρόπο τέτοιο που εξαρτάται από τον εκάστοτε στόχο, τα σημαντικά μέρη της εικόνας ενώ παράλληλα καταστέλλει την εξέταση άλλων περιοχών. Κατά τη διάρκεια αυτής της διαδικασίας γίνεται μαζική ανατροφοδότηση του οπτικού ρεύματος, η οποία μέχρι τώρα έχει γίνει ελάχιστα κατανοητή. Ταυτόχρονα υπάρχουν ευρείες συνειρμικές είσοδοι από μυϊκούς αισθητήρες ελέγχου, αλλά και από άλλες αισθήσεις που επιτρέπουν στον εγκέφαλο να χρησιμοποιήσει όλες τις γνώσεις και συσχετίσεις που έχουν προκύψει από τα πολλαπλά χρόνια ζωής του ανθρώπου στον κόσμο. Η ανατροφοδότηση κάνει κύκλους μέσα στον εγκέφαλο, πηγαίνοντας πίσω σε

όλα τα στάδια επεξεργασίας, συμπεριλαμβανομένων των αισθητήρων, στην περίπτωση αυτή των ματιών, τα οποία ελέγχουν μηχανικά το φωτισμό μέσω της ίριδας και συντονίζουν τη λήψη σήματος στην επιφάνεια του αμφιβληστροειδή.

Αντιθέτως, τα συστήματα όρασης, κατά βάση, είναι ακόμα αρκετά αφελή. Ως επί το πλείστον, σε αυτά δεν υπάρχει καμία ενσωματωμένη αναγνώριση προτύπων, κανένας αυτόματος έλεγχος εστίασης και διαφράγματος κάμερας και καμία συσχέτιση λόγω ετών εμπειρίας. Στην Εικόνα 3, για παράδειγμα, το ανθρώπινο μάτι παρατηρώντας το μαρκαρισμένο σημείο, βλέπει τον πλαϊνό καθρέφτη του αυτοκινήτου από τη μεριά του οδηγού.



Εικόνα 3: Ο πλαϊνός καθρέπτης από τη μεριά του οδηγού, όπως γίνεται αντιληπτός από τη μηχανική όραση. Κάθε αριθμός του πλέγματος αντιστοιχίζεται σε κάποια συνιστώσα του αντικειμένου με έντονο το στοιχείο του θορύβου.

Από την άλλη μεριά, το μόνο που «βλέπει» ο υπολογιστής ενός συστήματος όρασης στην ίδια εικόνα είναι ένα πλέγμα αριθμών. Κάθε μεμονωμένος αριθμός από αυτό το πλέγμα περιέχει ένα μεγάλο στοιχείο θορύβου και έτσι από μόνος του δίνει ελάχιστες πληροφορίες για το αντικείμενο. Στόχος της μηχανικής όρασης είναι να μετατραπεί αυτό το θορυβώδες σύνολο αριθμών στην αντίληψη «πλαϊνός καθρέφτης».

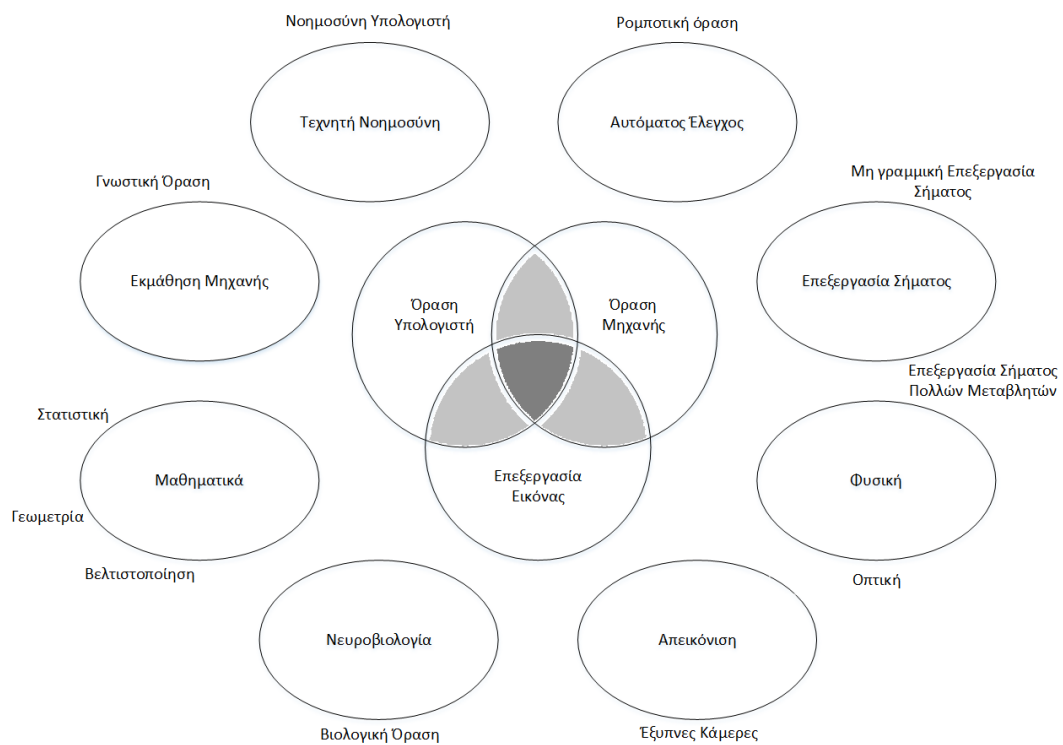
Ο τομέας της μηχανικής όρασης έχει άμεση σχέση με την ψηφιακή επεξεργασία εικόνας και πολλές φορές συγχέεται το πού σταματάει ο ένας τομέας και πού αρχίζει ο άλλος (Εικόνα 4). Για να γίνει πιο κατανοητό αυτό το όριο μεταξύ των τομέων ορίζονται τα παρακάτω τρία επίπεδα επεξεργασίας εικόνας, τα οποία ακολουθούνται και από τα συστήματα τεχνητής όρασης για να επιτύχουν το στόχο τους [19]:

- Χαμηλό επίπεδο: Σε έναν αλγόριθμο χαμηλού επιπέδου, τόσο η είσοδος όσο και η έξοδος είναι εικόνες (π.χ. Αποθορυβοποίηση, απαλοιφή θολώματος κλπ)
- Ενδιάμεσο επίπεδο: Η είσοδος, σε ένα αλγόριθμο επεξεργασίας ενδιάμεσου επιπέδου, είναι εικόνα αλλά η έξοδος είναι κάποια χαρακτηριστικά ανωτέρου επιπέδου (π.χ. Ακμές, περιγράμματα περιοχών κλπ.)
- Υψηλό επίπεδο: Η είσοδος, σε ένα αλγόριθμο επεξεργασίας υψηλού επιπέδου, είναι εικόνα αλλά η έξοδος είναι πληροφορία για το σημαντικό περιεχόμενο της εικόνας (π.χ. Ποια αντικείμενα αυτή περιέχει)

Η μηχανική όραση αποτελεί συμπληρωματικό και συνδεδετικό κρίκο μεταξύ πολλών επιστημών. Πολλές φορές αντιμετωπίζεται σαν τμήμα της τεχνητής νοημοσύνης καθώς οι δύο αυτοί τομείς μοιράζονται θέματα που αφορούν την αναγνώριση προτύπων και τις τεχνικές εκμάθησης. Επιπλέον, όσον αφορά το πεδίο της ρομποτικής, μπορεί να προσφέρει, μέσω ολοκληρωμένων συστημάτων, ακριβείς πληροφορίες περιβάλλοντος για την αυτόνομη πλοήγηση ρομπότ μέσα στο χώρο.

Έναν ακόμα τομέα, στενά συνδεδεμένο με την όραση μηχανής, αποτελεί ο τομέας της φυσικής. Τα περισσότερα συστήματα τεχνητής όρασης χρησιμοποιούν αισθητήρες εικόνας, σχεδιασμένους με τη χρήση κβαντικής φυσικής, για την ανίχνευση ηλεκτρομαγνητικών πεδίων. Σημαντική σύνδεση υπάρχει και με τη νευροβιολογία, καθώς η μελέτη των ματιών, των νευρώνων και της δομής του εγκεφάλου, τα οποία είναι υπεύθυνα για την επεξεργασία οπτικών ερεθισμάτων σε ανθρώπους και ζώα, οδήγησε στη δημιουργία ενός καινούριου τομέα της μηχανικής όρασης, όπου τα τεχνητά συστήματα σχεδιάζονται για να μιμηθούν τις συμπεριφορές των βιολογικών συστημάτων, σε διαφορετικά επίπεδα πολυπλοκότητας. Συνεπώς, η διεπιστημονική ανταλλαγή στη βιολογική και την υπολογιστική όραση έχει αποδειχτεί καρποφόρος και για τους δύο τομείς.

Όσον αφορά την επεξεργασία σήματος, η τεχνητή όραση μπορεί να επεκτείνει τις μεθόδους της και από επεξεργασία μονοδιάστατων σημάτων να γίνεται πλέον επεξεργασία τόσο δισδιάστατων όσο και πολυδιάστατων. Επιπλέον, η τεχνητή όραση με τα γραφικά υπολογιστών έχουν συνάφεια σαν τομείς, καθώς και οι δύο ασχολούνται με τη δημιουργία τρισδιάστατων μοντέλων από δεδομένα εικόνας. Τέλος, πρέπει να σημειωθεί ότι οι μέθοδοι και οι λειτουργίες αυτού του πολυδιάστατου τομέα μελετώνται από την οπτική γωνία των μαθηματικών και βασίζονται σε υποτομείς τους.



Εικόνα 4: Συσχέτιση του τομέα της μηχανικής όρασης με άλλα επιστημονικά πεδία.

3.3.1 Εφαρμογές της μηχανικής όρασης

Ένας από τους πιο εξέχοντες και σημαντικούς τομείς εφαρμογής της μηχανικής όρασης είναι η επεξεργασία εικόνων για ιατρικούς σκοπούς. Κατ' αυτή τη διαδικασία εξάγονται πληροφορίες από δεδομένα εικόνων με σκοπό την ιατρική διάγνωση ενός ασθενή. Τα δεδομένα αυτά μπορεί να προέρχονται από κάποιο μικροσκόπιο, από ακτινογραφίες, αγγειογραφίες, υπερήχους ή τομογραφίες. Αποτελέσματα των εφαρμογών μηχανικής όρασης στην ιατρική μπορεί να είναι η εύρεση όγκων, η ανίχνευση κάποιας κυτταρικής ή χρωμοσωμικής αντίθεσης, η μέτρηση των διαστάσεων οργάνων ή της ροής

του αίματος, πληροφορίες για τη δομή του εγκεφάλου ή την αποτελεσματικότητα κάποιας θεραπείας, αλλά και η αφαίρεση θορύβου από δεδομένα εικόνων ώστε να γίνονται πιο εύκολα οι διαγνώσεις.

Παρόλα αυτά, ο στρατιωτικός τομέας είναι ίσως ο μεγαλύτερος στον οποίο η όραση μηχανής βρίσκει εφαρμογή. Τα πιο προφανή παραδείγματα χρήσης της είναι η ανίχνευση εχθρών ή εχθρικών οχημάτων και η καθοδήγηση βλημάτων. Τα πιο προηγμένα συστήματα καθοδήγησης βλημάτων στέλνουν πλέον τα βλήματα σε περιοχές και όχι σε συγκεκριμένους στόχους, ώστε να γίνει αργότερα η επιλογή του ακριβούς στόχου, βασισμένη στα δεδομένα εικόνας που προέκυψαν από τη σάρωση της περιοχής. Τα πιο μοντέρνα στρατιωτικά συστήματα χρησιμοποιούν πλήθος αισθητήρων, συμπεριλαμβανομένων των αισθητήρων εικόνας, για να συλλέξουν πλούσιο σύνολο πληροφοριών από το πεδίο της μάχης και να τα χρησιμοποιήσουν για τη λήψη στρατηγικών αποφάσεων. Στις περιπτώσεις αυτές γίνεται αυτόματη επεξεργασία των δεδομένων ώστε να μειωθεί η πολυπλοκότητα και να αυξηθεί η αξιοπιστία.

Ένας από τους νεότερους τομείς εφαρμογής της όρασης μηχανής είναι αυτός των αυτόματων οχημάτων που περιλαμβάνει υποβρύχια, χερσαία, εναέρια και μη επανδρωμένα εναέρια οχήματα. Το επίπεδο της αυτονομίας αυτών των οχημάτων κυμαίνεται από οχήματα στα οποία η μηχανική όραση υποστηρίζει τους κυβερνήτες σε διάφορες καταστάσεις ως πλήρως αυτόνομα οχήματα. Τα πλήρως αυτόνομα οχήματα χρησιμοποιούν την όραση υπολογιστή για πλοήγηση και πιο συγκεκριμένα, για να γνωρίζουν πού βρίσκονται, για την παραγωγή χάρτη του περιβάλλοντός τους, για την ανίχνευση εμποδίων αλλά και για αποστολές ανίχνευσης συγκεκριμένων γεγονότων, όπως παραδείγματος χάριν της ύπαρξης δασικής πυρκαγιάς. Από την άλλη μεριά, παραδείγματα συστημάτων υποστήριξης αποτελούν τα συστήματα προειδοποίησης εμποδίων στα αυτοκίνητα, αλλά και τα συστήματα για αυτόνομη προσγείωση των αεροσκαφών. [20]

3.3.2 Προβλήματα που αντιμετωπίζει η μηχανική όραση

Η όραση μηχανής συναντά στο δρόμο της πολλαπλά προβλήματα που καλείται να υπερπηδήσει ή και να λύσει. Το πιο σημαντικό από αυτά τα προβλήματα είναι, κατά

βάση, άλυτο. Έχοντας σαν δεδομένο μια δισδιάστατη οπτική ενός τρισδιάστατου κόσμου, υπάρχουν άπειροι τρόποι να ανακατασκευαστεί το αντίστοιχο τρισδιάστατο σήμα. Ακόμα και αν τα δεδομένα ήταν τέλεια, μία δισδιάστατη εικόνα θα μπορούσε να αναπαριστά οποιονδήποτε από τους άπειρους συνδυασμούς τρισδιάστατων σκηνών. Συνεπώς στο πρόβλημα αυτό δεν υπάρχει μοναδική και οριστική λύση.

Κατά το σχεδιασμό ενός πρακτικού συστήματος μηχανικής όρασης, οι όχι τόσο ρητές πληροφορίες, που απορρέουν από τα συμφραζόμενα μπορούν να χρησιμοποιηθούν για να αποφευχθούν οι περιορισμοί που επιβάλλονται από τους οπτικούς αισθητήρες. Οι πληροφορίες αυτές μπορούν επίσης να διαμορφωθούν ρητά με τις τεχνικές εκμάθησης μηχανών. Έτσι, κρυμμένες πληροφορίες όπως το μέγεθος, ο προσανατολισμός της βαρύτητας κ.α. μπορούν να συσχετίζονται με τις τιμές τους σε ένα σύνολο εκπαίδευσης με ετικέτες. Εναλλακτικά, στο πρακτικό σύστημα μπορούν να προστεθούν επιπρόσθετοι αισθητήρες για να προσδιοριστούν οι κρυμμένες συνιστώσες. Συνεπώς, η χρήση μιας σειράς από λέιζερ ανίχνευσης για προσδιορισμό βάθους επιτρέπει την ακριβή μέτρηση του μεγέθους ενός αντικειμένου στο χώρο.

Τα δεδομένα εικόνας, όμως, με τα οποία τροφοδοτούνται τα συστήματα τεχνητής όρασης κάθε άλλο παρά τέλεια είναι. Ο θόρυβος και οι διάφορες διαστρεβλώσεις τα αλλοιώνουν επιδεινώνοντας ακόμα περισσότερο το έργο που προσπαθεί να επιτύχει η τεχνητή όραση. Τέτοιες αλλοιώσεις συνήθως προέρχονται από αλλαγές συνθηκών στο κόσμο γύρω μας (καιρός, φωτισμός, αντανάκλασεις, μετακινήσεις), ατέλειες του φακού, μηχανικές ρυθμίσεις, πεπερασμένο χρόνο ενσωμάτωσης στον αισθητήρα (θολές κινήσεις), ηλεκτρικό θόρυβο στον αισθητήρα ή σε άλλα ηλεκτρονικά εξαρτήματα αλλά και πιθανές συμπίεσεις μετά την αποτύπωση στιγμιότυπου της εικόνας.

Ο τρόπος με τον οποίο αντιμετωπίζεται ο θόρυβος σε αυτές τις περιπτώσεις είναι η χρήση στατιστικών μεθόδων. Για παράδειγμα, υπάρχει περίπτωση να είναι αδύνατο να ανιχνευτεί μία άκρη σε μία εικόνα μόνο με τη σύγκριση ενός σημείου με τους άμεσους γείτονές του. Εάν εξετάζονται, όμως, οι στατιστικές πέρα από μία τοπική περιοχή, η ανίχνευση ακρών γίνεται αισθητά ευκολότερη. Μια πραγματική άκρη θα πρέπει τελικά να εμφανίζεται ως μία ακολουθία απαντήσεων από τους άμεσους γείτονες μέσα σε μία συγκεκριμένη περιοχή, των οποίων ο προσανατολισμός είναι σύμφωνος με αυτόν των γειτόνων τους. Είναι, επίσης, δυνατό να αντισταθμιστεί ο θόρυβος με τη λήψη

στατιστικών μέσα στο χρόνο. Επιπροσθέτως, άλλες τεχνικές αντιμετωπίζουν το θόρυβο ή άλλες διαστρεβλώσεις φτιάχνοντας ρητά πρότυπα βασισμένα αποκλειστικά στα διαθέσιμα δεδομένα. Παραδείγματος χάριν, λόγω του ότι οι διαστρεβλώσεις φακών γίνονται σε μεγάλο βαθμό κατανοητές, η εκπαίδευση μπορεί να γίνει μόνο με βάση τις παραμέτρους ενός απλού πολυωνυμικού μοντέλου έτσι ώστε να μπορούν να περιγραφούν και κατ' επέκταση να διορθωθούν τέτοιου είδους διαστρεβλώσεις.

3.2 Βιβλιοθήκη OpenCV

Το OpenCV [17] είναι μία ανοιχτού κώδικα βιβλιοθήκη που χρησιμοποιείται σε εφαρμογές σχετικές με την όραση υπολογιστή και την εκπαίδευση μηχανών. Η βιβλιοθήκη αυτή είναι γραμμένη σε γλώσσες προγραμματισμού C και C++ και μπορεί να χρησιμοποιηθεί σε διάφορα λογισμικά όπως Linux, Windows και Mac OS X. Ταυτόχρονα, γίνεται ανάπτυξη της βιβλιοθήκης ώστε να είναι συμβατή με τις γλώσσες προγραμματισμού Python, Ruby, Matlab κ.α. Το OpenCV σχεδιάστηκε για να προσφέρει υπολογιστική επάρκεια, στοχεύοντας κυρίως στις εφαρμογές πραγματικού χρόνου. Λόγω του ότι έχει γραφτεί σε βελτιστοποιημένη, ως προς την απόδοση, γλώσσα προγραμματισμού C μπορεί να εκμεταλλευτεί τα οφέλη που προσφέρουν οι πολυπύρρηνοι επεξεργαστές. Βέβαια, τα πιο χρήσιμα, ίσως, χαρακτηριστικά του OpenCV είναι η αρχιτεκτονική και η διαχείριση μνήμης. Προσφέρει μία δομή στην οποία ο προγραμματιστής μπορεί να δουλέψει πάνω σε εικόνες και βίντεο με όποιον τρόπο θέλει, χρησιμοποιώντας αλγόριθμους του OpenCV ή και δικούς του, χωρίς να απασχολείται με το να δεσμεύει και να αποδεσμεύει μνήμη για τα δεδομένα εικόνας του.

Η βιβλιοθήκη OpenCV ξεκίνησε επίσημα σαν ένα έργο στα πλαίσια του Intel Research που είχε ως στόχο να αναπτύξει τεχνολογίες γύρω από τις εφαρμογές με μεγάλες απαιτήσεις σε CPU. Ανάμεσα στους πιο σημαντικούς συνεισφέροντες του έργου αυτού ήταν μέλη του Intel Research της Ρωσίας και η ομάδα του Intel Performance Library. Τα αντικείμενα αυτού του έργου ήταν τα παρακάτω:

- Εξελιγμένη έρευνα γύρω από την όραση μηχανής προσφέροντας τόσο ανοιχτό όσο και βελτιστοποιημένο κώδικα για τη βασική υποδομή όρασης.
- Διαδεδομένη γνώση γύρω από την όραση, παρέχοντας μία κοινή υποδομή για τους προγραμματιστές, ώστε ο κώδικάς τους πάνω σε αυτήν να είναι εύκολα προσβάσιμος και μεταφέρσιμος.
- Εμπορικές εφαρμογές βασισμένες στην αναπτυγμένη όραση κάνοντας τον μεταφέρσιμο και βελτιστοποιημένο, ως προς τις επιδόσεις, κώδικα, δωρεάν διαθέσιμο –με μία άδεια η οποία δεν απαιτεί να είναι ανοιχτές και δωρεάν οι ίδιες οι εφαρμογές.

Η πρώτη έκδοση του OpenCV παρουσιάστηκε στο κοινό κατά την IEEE διάσκεψη για την όραση μηχανής και την αναγνώριση προτύπων το 2000. Το OpenCV, ανήκει μέχρι και σήμερα, σε ένα μη κερδοσκοπικό οργανισμό, τον λεγόμενο OpenCV.org. [37]

Ένας από τους στόχους του OpenCV είναι να προσφέρει μία εύκολη προς τη χρήση υποδομή όρασης υπολογιστή, η οποία βοηθάει τους χρήστες να αναπτύξουν γρήγορα κάποιες αρκετά εξεζητημένες εφαρμογές όρασης. Η βιβλιοθήκη OpenCV περιλαμβάνει παραπάνω από πεντακόσιες συναρτήσεις, οι οποίες καλύπτουν αρκετές περιοχές του τομέα της όρασης, όπως τον έλεγχο εργοστασιακών προϊόντων, τις ιατρικές απεικονίσεις, την ασφάλεια, τις διεπαφές χρηστών, τη βαθμονόμηση καμερών, τη στερεοσκοπική όραση και τη ρομποτική. Καθώς η όραση υπολογιστή και η νοημοσύνη μηχανής τις περισσότερες φορές σχετίζονται άμεσα, το OpenCV περιέχει επιπλέον μία ολοκληρωμένη, γενικής χρήσης Machine Learning Library (MLL). Αυτή η υποβιβλιοθήκη είναι εστιασμένη στη στατιστική αναγνώριση και ομαδοποίηση μοτίβων. Είναι ζωτικής χρησιμότητας για την ολοκλήρωση των εργασιών που έχουν να κάνουν με την όραση και αποτελούν τον πυρήνα της βιβλιοθήκης OpenCV, αλλά είναι και αρκετά γενική ώστε να μπορέσει να χρησιμοποιηθεί και να λύσει κάθε πρόβλημα τεχνητής νοημοσύνης.

Οι ενσωματωμένες δομές της OpenCV βιβλιοθήκης είναι ισχυρές και αρκετά ευπροσάρμοστες για να μπορούν να επιλύσουν τα περισσότερα ζητήματα σχετικά με τη μηχανική όραση, για τα οποία διατίθενται καλά δομημένες λύσεις. Χρησιμοποιώντας το OpenCV οι χρήστες έχουν πολλές δυνατότητες, οι οποίες, μεταξύ άλλων, είναι να περικόπτουν εικόνες, να τις βελτιώνουν αλλάζοντας τη φωτεινότητα, την οξύτητα και την αντίθεσή τους, να αναγνωρίζουν σχήματα μέσα σε αυτές, να διαιρούν τις εικόνες σε

αισθητά εμφανείς περιοχές, να ανιχνεύουν κινούμενα αντικείμενα σε βίντεο, να αναγνωρίζουν γνωστά αντικείμενα, να υπολογίζουν την κίνηση ενός ρομπότ από δεδομένα κάμερας ή και να έχουν τρισδιάστατη, πλέον, άποψη για τον κόσμο χρησιμοποιώντας στερεοσκοπικές κάμερες.

Παρακάτω (Πίνακας 2) παρουσιάζεται μία σειρά από ενότητες που ανήκουν στο OpenCV καθώς και οι λειτουργίες που μπορούν να επιτελέσουν.

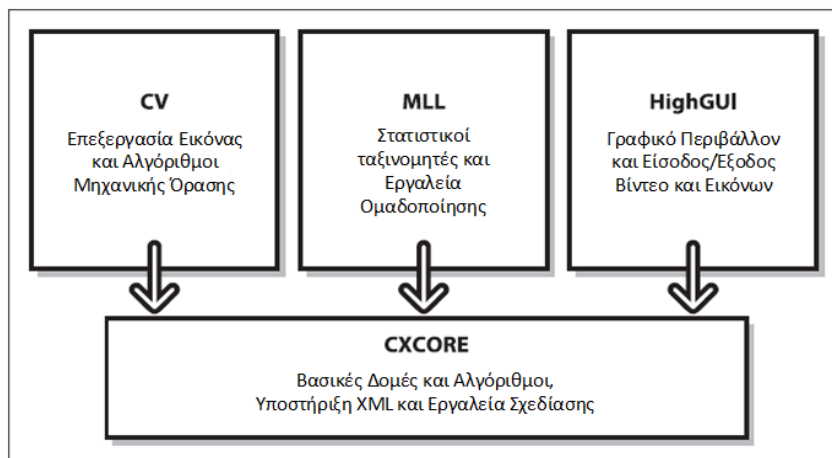
Ενότητα	Λειτουργία
Core	Πηγαίες δομές δεδομένων, τύποι στοιχείων και διαχείριση μνήμης
Imgproc	Φιλτράρισμα εικόνας, γεωμετρικοί μετασχηματισμοί εικόνας, δομή και ανάλυση μορφής
Highgui	Γραφικό περιβάλλον χρήστη, ανάγνωση και καταγραφή εικόνας και βίντεο
Video	Ανάλυση κίνησης και ανίχνευση αντικειμένων σε βίντεο
Calib3d	Βαθμονόμηση κάμερας και τρισδιάστατη ανακατασκευή από πολλαπλές οπτικές
Features2d	Εξαγωγή, περιγραφή, και ταίριασμα χαρακτηριστικών γνωρισμάτων
Objdetect	Αναγνώριση αντικειμένων με τη χρήση ταξινομητών καταρράκτη και ιστογράμματος κλίσης
ML	Στατιστικά πρότυπα και αλγόριθμοι ταξινόμησης για χρήση σε εφαρμογές όρασης υπολογιστών
Flann	Fast Library for Approximate Nearest Neighbors –γρήγορες αναζητήσεις σε χώρους υψηλών διαστάσεων
GPU	Παραλληλοποίηση αλγορίθμων με σκοπό τη γρήγορη εκτέλεση σε GPU

Stitching	Στρέβλωση, συνδυασμός και ρύθμιση συνδέσμων για την ένωση εικόνων
Nonfree	Εφαρμογές αλγορίθμων που είναι κατοχυρωμένοι με δίπλωμα ευρεσιτεχνίας σε κάποιες χώρες

Πίνακας 2: Κατάλογος βασικών ενοτήτων του OpenCV με τις λειτουργίες τους

Η βιβλιοθήκη OpenCV, κατά μία ευρύτερη έννοια, είναι δομημένη σε πέντε δομικά στοιχεία, τέσσερα από τα οποία φαίνονται στο σχήμα που ακολουθεί (Εικόνα 5). Το τμήμα CV περιλαμβάνει τους βασικούς αλγορίθμους για επεξεργασία εικόνας και υψηλότερου επιπέδου όραση υπολογιστή, το τμήμα ML αποτελεί τη βιβλιοθήκη που χρησιμοποιείται για την εκπαίδευση μηχανής και περιλαμβάνει πολλούς στατιστικούς ταξινομητές και εργαλεία ομαδοποίησης, το τμήμα HighGUI περιέχει ρουτίνες εισόδου/εξόδου και μεθόδους για αποθήκευση και φόρτωση βίντεο και εικόνων και το τμήμα CXCore περιέχει τις βασικές δομές δεδομένων και περιεχόμενο.

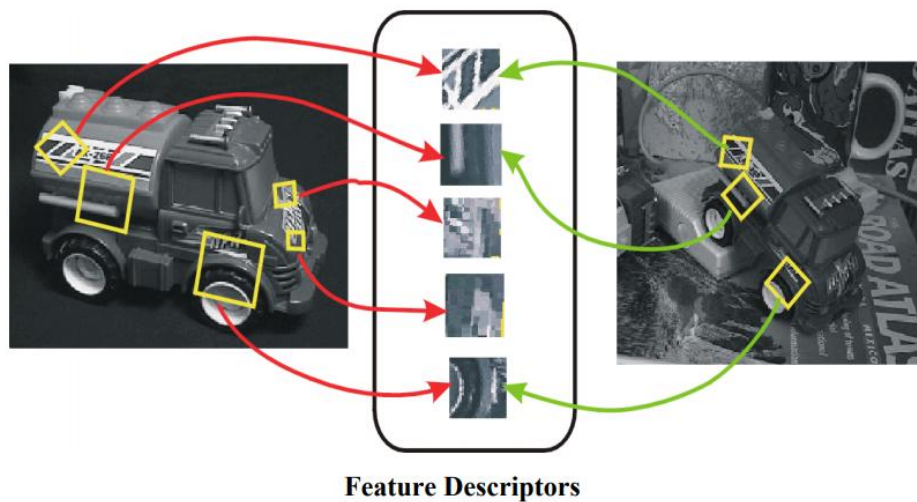
Το πέμπτο δομικό στοιχείο της βιβλιοθήκης OpenCV λέγεται CvAux και περιλαμβάνει τόσο μη χρησιμοποιήσιμα πλέον στοιχεία όσο και πειραματικούς αλγόριθμους. Μεταξύ άλλων καλύπτει πεδία όπως η στερεοσκοπική όραση, ο εντοπισμός ματιών και στόματος, η εύρεση του σκελετού αντικειμένων σε μια ολοκληρωμένη σκηνή και C++ κλάσεις για βαθμονόμηση καμερών. Κάποιες από αυτές τις λειτουργίες είναι πιθανόν να συμπεριληφθούν στο τμήμα CV. [17]



Εικόνα 5: Βασικά στοιχεία που απαρτίζουν το OpenCV.

3.3 Περιγραφείς χαρακτηριστικών (feature descriptors)

Ο περιγραφέας χαρακτηριστικών (feature descriptor) αποτελεί την αναπαράσταση μιας εικόνας ή μιας ενημερωμένης έκδοσης κώδικα που απλοποιεί την εικόνα, κρατώντας χρήσιμες πληροφορίες και απομακρύνοντας ξένες και άχρηστες πληροφορίες (Εικόνα 6). Οι περιγραφείς χαρακτηριστικών έχουν αρκετά πλεονεκτήματα, όπως είναι η παρουσίαση των τοπικών χαρακτηριστικών, η διαφοροποίησή τους και η δημιουργία μεγάλων βάσεων δεδομένων αντικειμένων τα οποία μπορούν να χρησιμοποιηθούν ευρέως, η ευκολία απομύζησης τους από μια εικόνα, και η πληθώρα τους καθώς είναι εκατοντάδες μόνο σε μια εικόνα. Γενικότερα οι περιγραφείς χαρακτηριστικών αυξάνουν την αποδοτικότητα της ανίχνευσης και της αναγνώρισης ενός αντικειμένου σε πραγματικό χρόνο, καθώς παράλληλα εκμεταλλεύονται διαφορετικούς τύπους χαρακτηριστικών σε διαφορετικές καταστάσεις. [26]



Εικόνα 6: Παράδειγμα χρήσης περιγραφέων χαρακτηριστικών (feature descriptors).

Οι ακόλουθες ιδιότητες είναι σημαντικές για τη χρήση ενός ανιχνευτή χαρακτηριστικών σε περιπτώσεις επεξεργασίας εικόνας και υπολογιστικής όρασης [27]:

- Ανθεκτικότητα -ο αλγόριθμος ανίχνευσης χαρακτηριστικών θα πρέπει να είναι σε θέση να εντοπίσει το ίδιο χαρακτηριστικό σε θέσεις ανεξάρτητες από κλιμάκωση, περιστροφή, μετατόπιση, φωτομετρικές παραμορφώσεις, αντικείμενα συμπίεσης και θόρυβο.

- Επαναληψιμότητα -ο αλγόριθμος ανίχνευσης χαρακτηριστικών θα πρέπει να είναι σε θέση να εντοπίσει το ίδιο χαρακτηριστικό της ίδιας σκηνής ή αντικειμένου επανειλημμένα υπό ποικίλες συνθήκες προβολής.
- Ακρίβεια -ο αλγόριθμος ανίχνευσης χαρακτηριστικών θα πρέπει να εντοπίζει με ακρίβεια την εικόνα (ίδιες θέσεις εικονοστοιχείων), ειδικά σε εργασίες αντιστοίχισης εικόνας, όπου οι αντιστοιχίες είναι απαραίτητες για την εκτίμηση της επιπολικής γεωμετρίας.
- Γενικότητα -ο αλγόριθμος ανίχνευσης χαρακτηριστικών θα πρέπει να είναι σε θέση να ανιχνεύει χαρακτηριστικά και να μπορεί να χρησιμοποιηθεί σε διαφορετικές εφαρμογές.
- Απόδοση -ο αλγόριθμος ανίχνευσης χαρακτηριστικών θα πρέπει να είναι σε θέση να ανιχνεύει χαρακτηριστικά καινούργια, με μεγάλη ταχύτητα για να υποστηρίξει εφαρμογές σε πραγματικό χρόνο.
- Ποσότητα -ο αλγόριθμος ανίχνευσης χαρακτηριστικών θα πρέπει να είναι σε θέση να εντοπίσει το σύνολο ή το μεγαλύτερο μέρος των χαρακτηριστικών της εικόνας, όπου, η πυκνότητα των εντοπισμένων χαρακτηριστικών θα πρέπει να αντικατοπτρίζει το περιεχόμενο της εικόνας για την παροχή συμπαγούς αντιπροσώπευσης εικόνας.

Οι ανιχνευτές χαρακτηριστικών μπορούν να ταξινομηθούν ευρέως σε τρεις κατηγορίες [27]:

- ανιχνευτές μίας κλίμακας,
- ανιχνευτές πολλαπλών κλιμάκων και
- συναφείς ανιχνευτές.

Η ενιαία κλίμακα σημαίνει ότι υπάρχει μόνο μία αναπαράσταση για τα χαρακτηριστικά ή τα περιγράμματα αντικειμένων χρησιμοποιώντας εσωτερικές παραμέτρους του ανιχνευτή. Οι ανιχνευτές μίας κλίμακας είναι αμετάβλητοι στην εικόνα, σε μετασχηματισμούς όπως η περιστροφή, οι αλλαγές σε φωτισμούς και η προσθήκη θορύβου. Έχοντας δύο εικόνες της ίδιας σκηνής που σχετίζονται με μια αλλαγή κλίμακας, πρέπει να καθοριστεί αν τα ίδια σημεία ενδιαφέροντος μπορούν να εντοπιστούν ή όχι. Ως εκ τούτου, είναι απαραίτητο να οικοδομηθεί πολυ-κλίμακα ανιχνευτών που μπορούν να εξάγουν αξιόπιστα τα χαρακτηριστικά γνωρίσματα υπό μεταβολές της κλίμακας.

Μόλις εντοπιστεί ένα σύνολο σημείων ενδιαφέροντος από μια εικόνα σε μια τοποθεσία $p(x, y)$, η οποία έχει κλίμακα s και προσανατολισμό θ , τότε το περιεχόμενο ή η δομή της εικόνας τους στη γειτονιά p των εικονοστοιχείων (pixels) κωδικοποιείται στον κατάλληλο περιγραφέα, ώστε να υπάρχει διακριτική αντιστοίχιση και να αποφεύγεται η ευαισθησία σε τοπικές παραμορφώσεις εικόνας. Ο περιγραφέας θα πρέπει να ευθυγραμμίζεται με θ μοίρες και να είναι ανάλογος στην κλίμακα s . Υπάρχει ένας μεγάλος αριθμός περιγραφέων χαρακτηριστικών εικόνας στη βιβλιογραφία. Οι πιο συνηθισμένοι περιγραφείς περιγράφονται στις επόμενες ενότητες [28]:

- Scale Invariant Feature Transform (SIFT)
- Gradient Location-Orientation Histogram (GLOH)
- Speeded-Up Robust Features Descriptor (SURF)
- Histogram of oriented gradient (HOG)
- PCA-SIFT
- Pyramidal HOG (PHOG)
- Pyramidal Histogram Of visual Words (PHOW)

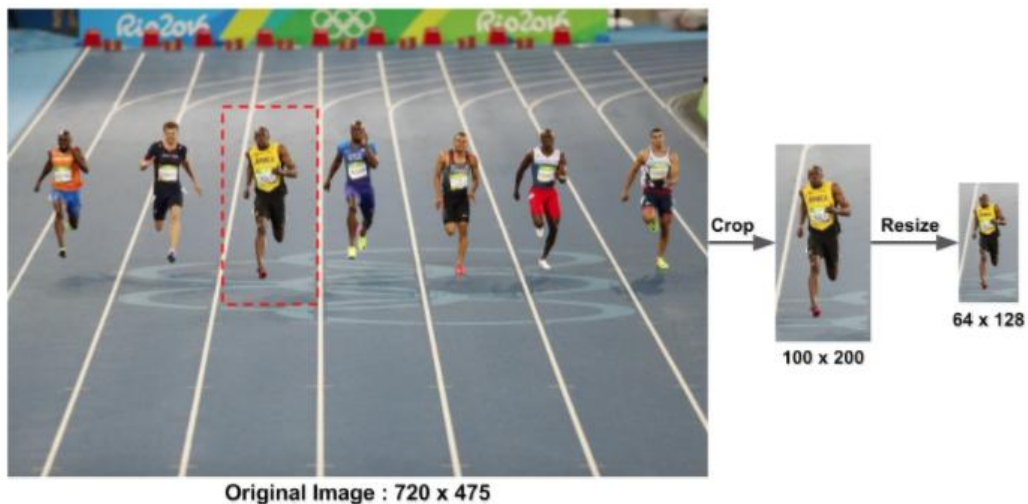
Η επεξεργασία εικόνας και η ανάλυση της με βάση τη συνεχή ή διακριτή μεταμόρφωσή της είναι κλασικές τεχνικές. Οι μετασχηματισμοί εικόνας χρησιμοποιούνται ευρέως σε φιλτράρισμα εικόνας, περιγραφή δεδομένων κ.λπ. Σήμερα τα θεωρήματα wavelet συνθέτουν πολύ δημοφιλείς μεθόδους επεξεργασίας εικόνας, αποθορυβοποίησης και συμπίεσης. Θεωρώντας ότι οι λειτουργίες Haar είναι οι απλούστερες wavelets, πολλές μέθοδοι διακριτών μετασχηματισμών εικόνας και επεξεργασίας χρησιμοποιούν αυτές οι μορφές. Η θεωρία μετασχηματισμού εικόνας είναι ένα πηγάδι γνωστής περιοχής που χαρακτηρίζεται από ακριβές μαθηματικό υπόβαθρο, αλλά σε πολλές περιπτώσεις υπάρχουν μεταμορφώσεις που έχουν ιδιαίτερες ιδιότητες οι οποίες δεν έχουν ακόμη διερευνηθεί. [29]

Στην παρούσα εργασία χρησιμοποιήθηκαν και οι δύο παραπάνω τεχνικές προκειμένου να γίνει σύγκριση και έλεγχος για να παραχθεί το καλύτερο αποτέλεσμα στην ανίχνευση και την αναγνώριση του χαρακτηριστικού. Με βάση τα στοιχεία που έχουν ειπωθεί παραπάνω, και οι δύο μέθοδοι περιγραφέων χρησιμοποιήθηκαν από την παρούσα εργασία για την εξαγωγή των χαρακτηριστικών και την εκπαίδευση του συστήματος. Στις ακόλουθες υποενότητες παρουσιάζονται τα στοιχεία και

χαρακτηριστικά του εκάστοτε περιγραφέα καθώς επίσης και μια μικρή συγκριτική ανάλυση μεταξύ των δύο μεθόδων.

3.3.1 Histogram of oriented gradient (HOG)

Τυπικά, ένας περιγραφέας χαρακτηριστικών μετατρέπει μια εικόνα μεγέθους: πλάτος x ύψος x 3 (κανάλια) σε ένα διάνυσμα χαρακτηριστικών / πίνακα μήκους n . Στην περίπτωση του περιγραφέα χαρακτηριστικών Histogram of Oriented Gradient (HOG), η εικόνα εισόδου έχει μέγεθος $64 \times 128 \times 3$ και το διάνυσμα χαρακτηριστικών εξόδου είναι μήκους 3780. Τα μήκη αυτά χρησιμοποιούνται ως βασικοί παράμετροι και προτείνονται για την καλύτερη εξαγωγή των χαρακτηριστικών, ωστόσο ποικίλουν ανάλογα με το μέγεθος και τα κανάλια τα οποία φέρει ή θα χρησιμοποιηθούν από μια εικόνα. Ως παράδειγμα δίνεται μια μεγάλη εικόνα μεγέθους 720×475 (Εικόνα 7). Από την εικόνα επιλέγεται ένα άτομο και η εικόνα κόβεται σε μέγεθος 100×200 για τον υπολογισμό του περιγραφέα χαρακτηριστικών HOG [30]. Προκειμένου να μπορέσει να χρησιμοποιηθεί καλύτερα η εικόνα, γίνεται σμίκρυνση αυτής σε μέγεθος 64×128 . Μόλις γίνει η σμίκρυνση της εικόνας εφαρμόζετε ο περιγραφέας HOG.



Εικόνα 7: Διαδικασία περικοπής εικόνων πριν την εφαρμογή του περιγραφέα HOG.

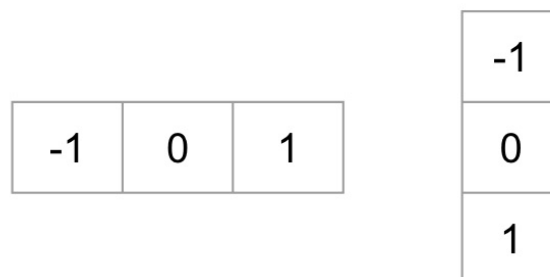
Στον περιγραφέα χαρακτηριστικών HOG, κύριο ρόλο φέρει η παραγωγή των χαρακτηριστικών γνωρισμάτων κατανομής (ιστογράμματα) και διευθύνσεων κλίσεων (προσανατολισμένες κλίσεις). Οι διαβαθμίσεις (παράγωγα x και y) μιας εικόνας είναι

χρήσιμες καθώς το μέγεθος των διαβαθμίσεων είναι μεγάλο γύρω από τις άκρες και τις γωνίες (περιοχές με απότομες αλλαγές έντασης). Σημειώνεται ότι οι άκρες και οι γωνίες μιας εικόνας φέρουν πολύ περισσότερες πληροφορίες σχετικά με το σχήμα του αντικειμένου από ότι οι επίπεδες περιοχές.

Η διαδικασία απόσπασης του περιγραφέα μπορεί να χωριστεί σε τέσσερα βήματα:

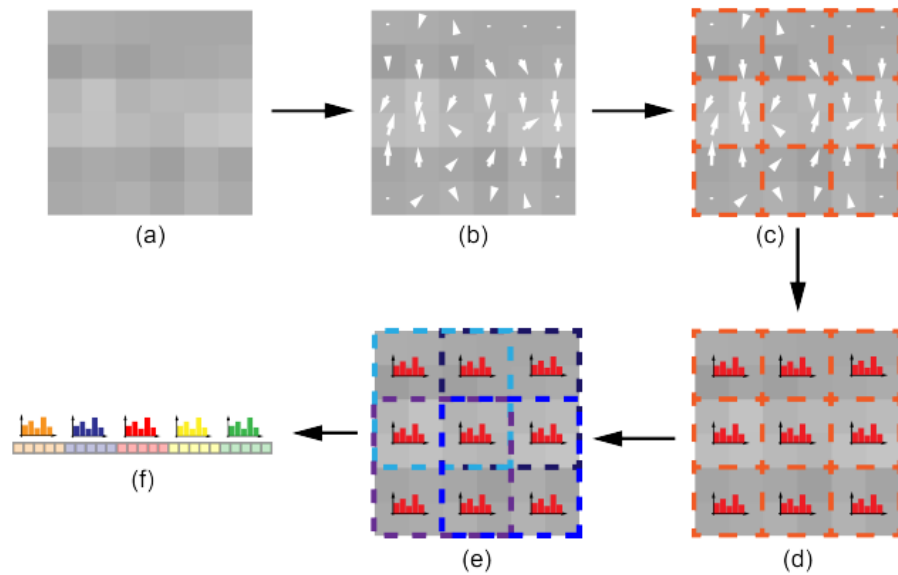
1. Τον υπολογισμό της κλίσης του φορέα για κάθε ένα από τα εικονοστοιχεία
2. Την ομαδοποίηση των εικονοστοιχείων σε κύτταρα
3. Τη δημιουργία μπλοκ από ομάδες κυττάρων σε και
4. Τη δημιουργία του περιγραφέα. [31]

Για να υπολογιστεί ένας περιγραφέας HOG, πρέπει πρώτα να υπολογιστούν οι οριζόντιες και κάθετες κλίσεις σε μια εικόνα. Ο υπολογισμός των κλίσεων γίνεται μέσα από τον καθορισμό των τιμών με βάση τα εικονοστοιχεία της εικόνας, προκειμένου να υπολογιστεί το ιστόγραμμα των κλίσεων. Αυτό επιτυγχάνεται εύκολα φιλτράροντας την εικόνα με τους ακόλουθους πυρήνες (Εικόνα 8).



Εικόνα 8: Πυρήνες που χρησιμοποιούνται για το φιλτράρισμα των εικόνων.

Στη συνέχεια, τα εικονοστοιχεία ομαδοποιούνται σε κύτταρα, όπως φαίνεται στην Εικόνα 9 (c). Στο τρίτο βήμα, δημιουργούνται μπλοκ με την ομαδοποίηση των κυττάρων, όπως φαίνεται στην Εικόνα 9 (e) και Εικόνα 9 (b). Στο τέταρτο βήμα, Εικόνα 9 (d), ο περιγραφέας δημιουργείται. Ο περιγραφέας αποτελείται ουσιαστικά από μια λίστα ιστογραμμάτων κυττάρων όλων των μπλοκ. Η τοπική διακύμανση του φωτός ή η υψηλή αντίθεση μεταξύ του προσκηνίου και του φόντου μπορεί να είναι ένα ζήτημα στην επεξεργασία εικόνας. Ο τρόπος μείωσης του προβλήματος είναι η ομαλοποίηση των ιστογραμμάτων σύμφωνα με τους γείτονές τους.



Εικόνα 9: Σειρά διαδικασιών η οποία ακολουθείται προκειμένου να παραχθεί το ιστόγραμμα της εικόνας. Τα βήματα εξαγωγής περιγραφών HOG: (a) πρωτότυπη εικόνα, (b) υπολογισμός διανύσματος κλίσης, (c) εικονοστοιχείο ομαδοποίηση κυττάρων, (d) υπολογισμός ιστόγραμμα.

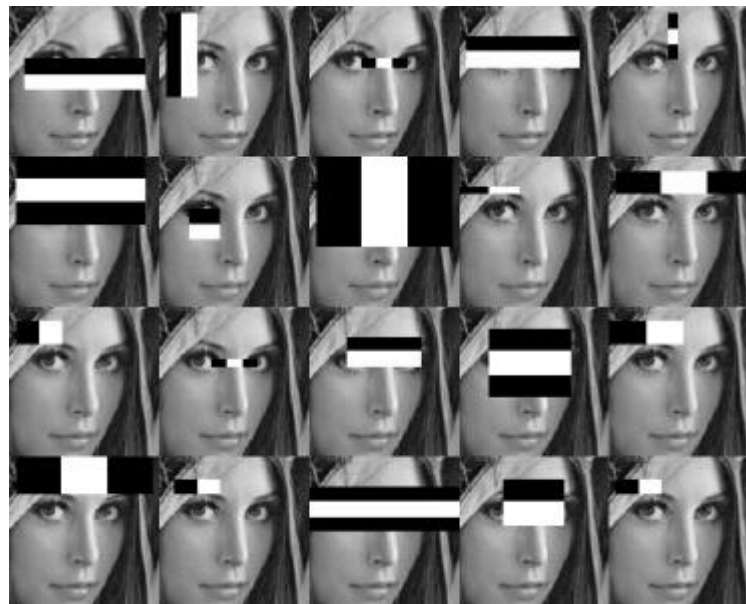
3.3.2 Haar features

Τα χαρακτηριστικά Haar είναι χαρακτηριστικά ψηφιακής εικόνας που χρησιμοποιούνται στην αναγνώριση αντικειμένων. Οφείλουν το όνομά τους στην διαισθητική τους ομοιότητα με τα κύματα Haar και χρησιμοποιήθηκαν στον πρώτο ανιχνευτή προσώπου πραγματικού χρόνου. [33]

Ιστορικά, η επεξεργασία εικόνας με τις τιμές των εικονοστοιχείων RGB σε κάθε εικόνα έκαναν την εργασία υπολογισμού χρονοβόρα και δαπανηρή. Μια δημοσίευση των Παπαγεωργίου και άλλων [34]. παρουσίασε ένα εναλλακτικό σύνολο χαρακτηριστικών βασισμένο σε κύματα Haar αντί για τις συνήθεις εντάσεις εικόνας. Οι Viola και Jones [33] προσάρμοσαν την ιδέα της χρήσης κυματοειδών μετασχηματισμών Haar και ανέπτυξαν τα λεγόμενα χαρακτηριστικά Haar (Haar features).

Ένα χαρακτηριστικό γνώρισμα Haar θεωρεί γειτονικές ορθογώνιες περιοχές σε μια συγκεκριμένη θέση σε ένα παράθυρο ανίχνευσης, συνοψίζει τις εντάσεις των εικονοστοιχείων σε κάθε περιοχή και υπολογίζει τη διαφορά μεταξύ αυτών των αθροισμάτων. Αυτή η διαφορά στη συνέχεια χρησιμοποιείται για την κατηγοριοποίηση

υποτομέων μιας εικόνας. Για να γίνει κατανοητή η χρήση των Haar χαρακτηριστικών, θεωρείται πως υπάρχει μια βάση δεδομένων εικόνων με ανθρώπινα πρόσωπα (Εικόνα 10). Είναι κοινή παρατήρηση ότι μεταξύ όλων των προσώπων η περιοχή των ματιών είναι πιο σκούρα από την περιοχή των μάγουλων. Επομένως, ένα κοινό χαρακτηριστικό γνώρισμα Haar για την ανίχνευση προσώπου είναι ένα σύνολο δύο παρακείμενων ορθογωνίων που βρίσκονται πάνω από το μάτι και την περιοχή των μάγουλων. Η θέση αυτών των ορθογωνίων καθορίζεται σε σχέση με ένα παράθυρο ανίχνευσης που ενεργεί σαν ένα κιβώτιο οριοθέτησης στο αντικείμενο-στόχο (το πρόσωπο σε αυτή την περίπτωση). [34]



Εικόνα 10: Χαρακτηριστικά Haar που χρησιμοποιούνται από μια εικόνα.

Στη φάση ανίχνευσης του πλαισίου των αντικειμένων Viola-Jones, ένα παράθυρο του μεγέθους στόχου μετακινείται πάνω από την εικόνα εισόδου, και για κάθε υποενότητα της εικόνας υπολογίζεται το χαρακτηριστικό γνώρισμα Haar. Αυτή η διαφορά στη συνέχεια συγκρίνεται με ένα μαθημένο όριο που χωρίζει τα μη αντικείμενα από τα αντικείμενα. Επειδή ένα τέτοιο χαρακτηριστικό γνώρισμα Haar είναι μόνο ένας ασθενής ταξινομητής, ένας μεγάλος αριθμός χαρακτηριστικών τύπου Haar είναι απαραίτητος για την περιγραφή ενός αντικειμένου με επαρκή ακρίβεια. Στο πλαίσιο ανίχνευσης αντικειμένων Viola-Jones, επομένως, τα χαρακτηριστικά γνωρίσματα Haar είναι οργανωμένα έτσι ώστε να σχηματίσουν έναν ισχυρό ταξινομητή, ο οποίος θα μπορεί να χρησιμοποιηθεί στην ανίχνευση των αντικειμένων.

Το βασικό πλεονέκτημα ενός χαρακτηριστικού τύπου Haar έναντι των περισσότερων άλλων χαρακτηριστικών είναι η ταχύτητα υπολογισμού του. Λόγω της χρήσης ολοκληρωμένων εικόνων, ένα χαρακτηριστικό γνώρισμα Haar οποιουδήποτε μεγέθους μπορεί να υπολογιστεί σε σταθερό χρόνο.

3.3.3 Σύγκριση περιγραφέων

Στις παραπάνω παραγράφους έχει παρουσιαστεί η λειτουργία κάθε περιγραφέα χωριστά, καθώς και τα χαρακτηριστικά τους, τα οποία είναι απαραίτητα για την κατανόηση των λειτουργιών τους και του τρόπου με τον οποίο χρησιμοποιούνται. Οι συνήθεις λόγοι για τους οποίους επιλέγεται η εξαγωγή των χαρακτηριστικών αντί για τις τιμές των εικονοστοιχείων έγκειται στην ευκολία των πρώτων, να κωδικοποιήσουν πληροφορίες αντικειμένων υψηλού επιπέδου (τμήματα, υφή). Αντίθετα, τα εικονοστοιχεία μπορούν να αποτυπώσουν μόνο τιμές έντασης, κάτι το οποίο κάνει το σύστημα πιο αργό κατά την επεξεργασία των εικόνων σε οποιοδήποτε στάδιο. [32]

Αρκετοί ερευνητές προσεγγίζουν βιβλιογραφικά το θέμα, προσπαθώντας να κατανοήσουν τη διαφορετικότητα των μεθόδων και των στοιχείων που χρησιμοποιούνται. Σε σχετική έρευνα [35] παρουσιάζονται διαφορετικές δοκιμές και με τους δύο περιγραφείς συνοδευόμενες με τα ποσοστά αναγνώρισης. Από τους δύο περιγραφείς, εκείνος με τα υψηλότερα ποσοστά αναγνώρισης είναι ο HOG. Ωστόσο, σημειώνεται ότι με τη χρήση αυτού του περιγραφέα παρουσιάζεται καθυστέρηση στην παραγωγή των αποτελεσμάτων η οποία είναι ανάλογη του μεγέθους και του αριθμού των εικόνων.

Όπως επισημάνθηκε σε παραπάνω ενότητα ο περιγραφέας HOG βασίζεται στην παραγωγή ιστογράμματος, το οποίο αντικατοπτρίζει την ένταση των εικονοστοιχείων που έχει η εικόνα. Επομένως το μέγεθος της εικόνας φέρει σημαντικό ρόλο στο χρόνο της επεξεργασίας. Ακόμη, εξίσου σημαντικό στοιχείο είναι και οι χρωματισμοί των εικόνων. Μια εικόνα έγχρωμη με αρκετά κανάλια απαιτεί μια πιο χρονοβόρα επεξεργασία απ' ό τι μια εικόνα ασπρόμαυρη.

Αντίθετα, τα Haar features είναι χαρακτηριστικά τα οποία εξάγονται από την επεξεργασία της εικόνας, εφαρμόζοντας σε αυτή έτοιμες «φόρμες» για την εξαγωγή τους.

Το στοιχείο αυτό καθιστά πιο εύκολη την επεξεργασία των εικόνων και πιο άμεση τη διαδικασία. Ωστόσο επειδή κάθε εικόνα είναι διαφορετική και φέρει ιδιαίτερα χαρακτηριστικά, θα πρέπει η εφαρμογή των φορμών να επιλέγεται και να γίνεται πιο προσεκτικά, προκειμένου να δομηθεί ένα σωστό σύνολο εικόνων το οποίο θα χρησιμοποιηθεί για την εκπαίδευση κάποιου συστήματος. Προκειμένου να γίνουν σωστά τα δείγματα θα πρέπει να γίνει σωστή προ-επεξεργασία και να χρησιμοποιηθούν στοιχεία με σωστή δομή. [36]

Επομένως, το θέμα της σύγκρισης τίθεται στην χρήση της εφαρμογής και στους σκοπούς της. Από τη μια πλευρά ο περιγραφέας HOG είναι πιο αξιόπιστος, αλλά πιο αργός, από την άλλη τα Haar είναι μια πιο άμεση μορφή εξαγωγής χαρακτηριστικών, τα οποία χρησιμοποιούνται για λόγους αμεσότητας και ταχύτητας.

Στην παρούσα εφαρμογή, όπως θα παρουσιαστεί και στο επόμενο κεφάλαιο, μετά από δοκιμές, χρησιμοποιήθηκαν οι περιγραφείς HOG για την εξαγωγή των στοιχείων και την εκπαίδευση του συστήματος. Στο σύστημα το οποίο αναπτύχθηκε κύριο και σημαντικό ρόλο φέρει η αύξηση των επιτυχημένων αποτελεσμάτων ανίχνευσης και όχι τόσο η ταχύτητα. Άλλωστε η εκπαίδευση του συστήματος λαμβάνει χώρα μια φορά, επομένως το δείγμα είναι ήδη εκπαιδευμένο όταν ξεκινά η ανίχνευση. Τα αποτελέσματα των δοκιμών παρουσιάζονται αναλυτικότερα στο Κεφάλαιο 4.

Σύνοψη κεφαλαίου

Στο παρόν κεφάλαιο αναλύθηκε σε βάθος ο τομέας της μηχανικής όρασης και οι εφαρμογές της, που αποτέλεσαν κινητήριο δύναμη για την υλοποίηση της διπλωματικής εργασίας. Επιπροσθέτως, έγινε εκτενής αναφορά στη βιβλιοθήκη OpenCV, η οποία αποτέλεσε τη βάση για την ανάπτυξη του λογισμικού μέρους και πραγμάτωσε τους σκοπούς της μηχανικής όρασης για την παρούσα υλοποίηση. Τέλος, αναφέρθηκαν και συγκρίθηκαν δύο είδη περιγραφέων, οι Haar και οι HOG, οι οποίοι παρήχθησαν για τους σκοπούς του συστήματος.

Κεφάλαιο 4: Υλοποίηση του συστήματος

Στο παρόν κεφάλαιο θα αναλυθεί διεξοδικά ο τρόπος λειτουργίας της εφαρμογής καθώς και όλα τα τμήματά που αναπτύχθηκαν ξεχωριστά μέχρι να παραχθεί στο τελικό προϊόν. Πιο αναλυτικά, αναφέρονται ο τρόπος εκπαίδευσης των ταξινομητών αναγνώρισης, τα στάδια από τα οποία περνάει ο αλγόριθμος καθώς και όλες οι λεπτομέρειες για τη δημιουργία των συνδέσεων μεταξύ του υπολογιστή και του μικροϋπολογιστή με σκοπό την ανταλλαγή μηνυμάτων.

4.1 Εκπαίδευση ταξινομητών

Εκτενές και πολύ σημαντικό μέρος της διαδικασίας ανάπτυξης της εφαρμογής αποτέλεσε η εκπαίδευση των ταξινομητών που τελικά χρησιμοποιήθηκαν. Η βιβλιοθήκη OpenCV παρέχει κάποιους ήδη εκπαιδευμένους ταξινομητές, έτοιμους προς χρήση σε αλγόριθμους αναγνώρισης. Από αυτούς τους ταξινομητές, δύο φάνηκαν χρήσιμοι και τελικά χρησιμοποιήθηκαν στον αλγόριθμο που αναπτύξαμε. Ο πρώτος είναι ο ταξινομητής που εκπαιδεύτηκε για την ανίχνευση προσώπου και ο δεύτερος, αυτός που εκπαιδεύτηκε για την ανίχνευση ματιών. Ωστόσο, επιτακτικής σημασίας χαρακτηριστικό το οποίο έπρεπε να ανιχνευτεί για την επίτευξη του στόχου που τέθηκε ήταν το ανθρώπινο στόμα. Η βιβλιοθήκη OpenCV παρέχει έτοιμο ταξινομητή και για την ανίχνευση στόματος, κάτι το οποίο για την παρούσα εφαρμογή δεν ήταν αρκετό. Για να επιτύχουμε στην αναγνώριση συναισθήματος ήταν απαραίτητη η χρήση πιο ειδικευμένων ταξινομητών πάνω στην ανίχνευση των τριών εκφράσεων στόματος, του χαμογελαστού, του λυπημένου και του ανοιχτού, που αντιστοιχούν τελικά στα τρία συναισθήματα προς αναγνώριση, της χαράς, της λύπης και της νύστας.

Για να ολοκληρωθεί η εκπαίδευση των ταξινομητών χρειάζονται τα εξής στάδια:

- Συλλογή θετικών και αρνητικών εικόνων
- Δημιουργία δειγμάτων από τις θετικές εικόνες
- Εκπαίδευση με βάση τα δείγματα και τις αρνητικές εικόνες

Ο όρος θετικές εικόνες (positive images) αναφέρεται σε εικόνες που έχουν σαν βασικό αντικείμενό τους το αντικείμενο ενδιαφέροντος, αυτό δηλαδή που στοχεύουν τελικά να αναγνωρίσουν οι ταξινομητές. Οι εικόνες αυτές είτε αποτελούνται εξ ολοκλήρου από το αντικείμενο ενδιαφέροντος είτε το περιέχουν ανάμεσα σε άλλα αντικείμενα και διαφορετικά φόντα. Το αντικείμενο σε αυτές μπορεί να βρίσκεται σε διαφορετικές θέσεις, βάθη και φωτισμό. Το σύνολο αυτών των εικόνων είναι αυτό στο οποίο θα βασιστεί η διαδικασία εκπαίδευσης για να «διδάξει» στους ταξινομητές ποιο θα είναι το αντικείμενο για το οποίο ψάχνουν και στοχεύουν να ανιχνεύσουν.

Εστιάζοντας στην εφαρμογή που αναπτύχθηκε, οι θετικές εικόνες που έπρεπε να συλλεχθούν ήταν τριών ειδών, εικόνες με χαμόγελα, εικόνες με λυπημένα στόματα και εικόνες με ανοιχτά στόματα (Εικόνα 11). Υψίστης σημασίας για την επιτυχία της διαδικασίας και για να παραχθούν ταξινομητές οι οποίοι θα καταφέρουν να αναγνωρίζουν την κατάσταση στόματος για την οποία εκπαιδεύτηκαν, ήταν η συλλογή εικόνων με αληθοφανή και πολλών ειδών στόματα, με καθαρές εκφράσεις και προπαντός εύκολα διακριτές μεταξύ τους. Μεγάλος στόχος, εκτός από την ανίχνευση και κατηγοριοποίηση του συναισθήματος, ήταν και η αποφυγή σύγχυσης του με τα λοιπά συναισθήματα. Είναι σαφώς προτιμότερο να προσπεραστεί κάποια έκφραση χωρίς να κατορθωθεί η αναγνώρισή της, παρά να ταυτιστεί λαθεμένα με κάποιο άλλο είδος έκφρασης.

Για να είναι οι θετικές εικόνες όσο το δυνατό πιο αντιπροσωπευτικές ως προς το συναίσθημα το οποίο απεικονίζουν αλλά και για να συλλεχθούν σε πλεονασμό, επιλέχθηκε να έχουν τη μορφή φωτογραφιών καθημερινών ανθρώπων με αληθοφανείς και αυθόρμητες εκφράσεις. Πιο συγκεκριμένα, ζητήθηκε η βοήθεια από συγγενικά και φιλικά πρόσωπα ώστε να συγκεντρωθεί ποικιλία φωτογραφιών με τη σωστή μορφή χειλιών, αφού από αυτά θα εξαχθεί αργότερα το συμπέρασμα του συναισθήματος. Το δείγμα που συλλέχθηκε βασίστηκε κατά πολύ σε άτομα μικρής ηλικίας που θα ήταν και τα πλέον ιδανικά, δεδομένου ότι στόχος της εφαρμογής τέθηκε να είναι η αναγνώριση εκφράσεων σε παιδικά πρόσωπα. Ωστόσο, για ερευνητικούς λόγους και λόγους

ποικιλομορφίας προτιμήθηκαν και γυναικεία πρόσωπα, στα οποία τα χαρακτηριστικά είναι περισσότερο ομαλά, για να συμπληρώσουν το δείγμα.

Σκοπός της διαδικασίας της δειγματοληψίας ήταν να συλλεχθούν όσο περισσότερες φωτογραφίες που θα απεικόνιζαν το αντικείμενο ενδιαφέροντος. Για να επιτευχθεί ο στόχος αυτός χωρίς ιδιαίτερη προσπάθεια και σπατάλη χρόνου, κυρίως από τα άτομα που συμμετείχαν εθελοντικά στη διαδικασία, αναπτύχθηκε ένα πρόγραμμα που αυτοματοποίησε κατά πολύ τις ενέργειες της φωτογράφισης. Στον κώδικα του προγράμματος αυτού αξιοποιήθηκαν οι βιβλιοθήκες που παρέχει το OpenCV. Πιο συγκεκριμένα, κατά την εκτέλεση αυτού του προγράμματος η κάμερα του υπολογιστή ανοίγει αυτόματα και για όσο παραμένει ανοιχτή, λαμβάνονται και αποθηκεύονται στιγμιότυπα εικόνας σε οποιαδήποτε μορφή και όνομα επιλεγεί. Το εκάστοτε άτομο που συμμετείχε στη δειγματοληψία έπρεπε να βρίσκεται μπροστά στην κάμερα, με θέση στη οποία θα φαίνεται καθαρά το πρόσωπό του, έχοντας κάθε φορά την αντίστοιχη έκφραση για την οποία συλλέγονταν οι φωτογραφίες και να κάνει ελαφριές κινήσεις κεφαλιού, ώστε οι λήψεις να έχουν διαφορετικές γωνίες και βάθος. Η απόφαση αυτή ήταν στρατηγικής σημασίας, καθώς τα δείγματα που συλλέχθηκαν είχαν τη ροή ενός βίντεο και τελικά χρησιμοποιήθηκαν για την εκπαίδευση ταξινομητών που εφαρμόστηκαν για ανίχνευση εκφράσεων σε βίντεο.





Εικόνα 11: Παραδείγματα φωτογραφιών που χρησιμοποιήθηκαν για την παραγωγή θετικών δειγμάτων που απεικονίζουν χαμογελαστό στόμα

Εξίσου σημαντικές, αν όχι περισσότερο, για την επιτυχία της εκπαίδευσης των ταξινομητών, είναι και οι αρνητικές εικόνες. Με τον όρο αρνητικές εικόνες εννοούμε εικόνες οι οποίες είναι εντελώς άσχετες με το αντικείμενο ενδιαφέροντος. Οι εικόνες αυτές δεν πρέπει να περιέχουν το αντικείμενο ή αν το περιέχουν αυτό να βρίσκεται σε μία διαστρεβλωμένη του μορφή. Ιδανικά, για να εκπαιδευτεί ταξινομητής και να έχει αυξημένα αποτελέσματα αναγνώρισης και ανίχνευσης, πρέπει να συλλεχθεί πλήθος αρνητικών φωτογραφιών που φαινομενικά θα είναι πανομοιότυπες με τις θετικές εικόνες, με τη διαφορά ότι δεν θα περιέχουν το αντικείμενο προς αναγνώριση. Για παράδειγμα, αν στόχος είναι να αναγνωριστεί μία πινακίδα υποχρεωτικής διακοπής πορείας, οι αρνητικές εικόνες που θα έπρεπε να χρησιμοποιηθούν κατά τη διαδικασία εκπαίδευσης θα πρέπει να περιέχουν δρόμους ή ακόμα και άλλου τύπου πινακίδες.

Εκτός από το είδος των εικόνων, υψίστης σημασίας δεδομένο είναι και η ποσότητα των αρνητικών και των θετικών εικόνων οι οποίες θα χρησιμοποιηθούν προκειμένου να δημιουργηθεί ένα ολοκληρωμένο δείγμα. Η ποσότητα αρνητικών εικόνων που απαιτείται είναι κατά πολύ μεγαλύτερη από αυτή των θετικών εικόνων. Αυτό γιατί ενώ με τις θετικές εικόνες ο ταξινομητής θα εκπαιδευτεί για το τι να ψάχνει και πού να στοχεύει, με τις αρνητικές εικόνες θα καταλάβει τι πρέπει να αποφεύγει και με τι δεν πρέπει να συγγέει το αντικείμενο ενδιαφέροντος.

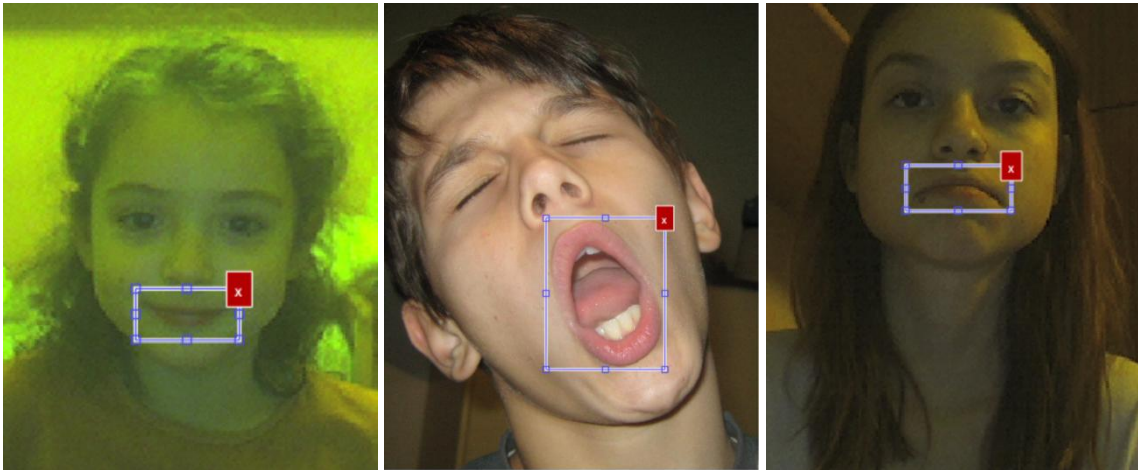
Έχοντας υπόψη όλα τα παραπάνω, κατά τη διαδικασία της φωτογράφισης για τη δειγματοληψία, ζητήθηκε από τα άτομα που συμμετείχαν να κάνουν διάφορες

γκριμάτσες. Με αυτό τον τρόπο εξασφαλίσαμε το κοινό φόντο στα δείγματα, δηλαδή την ύπαρξη προσώπου, και παράλληλα τη διαστρεβλωμένη εικόνα του στόματος (Εικόνα 12) σε σχέση πάντα με την έκφραση στην οποία στοχεύουμε. Το θετικό με τα δείγματα που συλλέχτηκαν και απεικόνιζαν τις εκφράσεις ενδιαφέροντος, ήταν ότι αυτά αποτέλεσαν επίσης ιδανικά δείγματα αρνητικών εικόνων και κατ' επέκταση χρησιμοποιήθηκαν συμπληρωματικά και γι' αυτόν το σκοπό. Πιο συγκεκριμένα, οι εικόνες με χαμογελαστό και αυτές με ανοιχτό στόμα χρησιμοποιήθηκαν σαν αρνητικές για την εκπαίδευση του ταξινομητή που τελικά θα αναγνώριζε λυπημένο στόμα. Αντίστοιχα λειτούργησαν και οι εικόνες των υπόλοιπων εκφράσεων.



Εικόνα 12: Παράδειγμα φωτογραφίας με διαστρεβλωμένη μορφή στόματος που χρησιμοποιήθηκε σαν αρνητική εικόνα

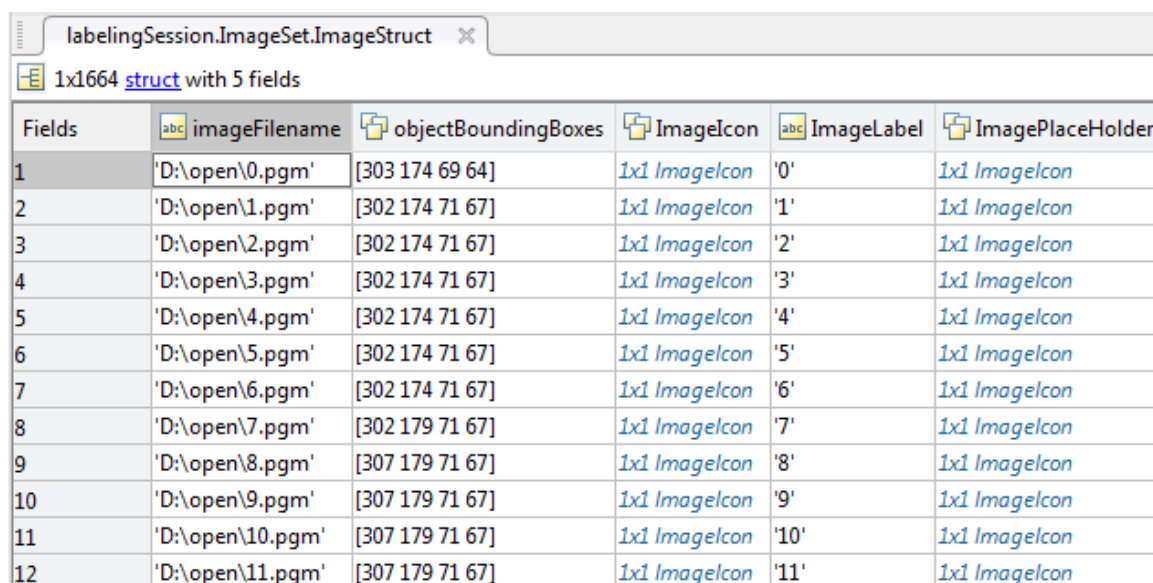
Το πιο χρήσιμο τμήμα των θετικών εικόνων είναι η περιοχή που περιλαμβάνει το αντικείμενο ενδιαφέροντος. Για να χρησιμοποιηθούν ουσιαστικά οι θετικές εικόνες και να μπορέσει να γίνει κατανοητό, κατά τη διαδικασία της εκπαίδευσης, ποιο είναι το αντικείμενο προς αναγνώριση, θα πρέπει να οριοθετηθεί με κάποιον τρόπο η περιοχή ενδιαφέροντος (Region Of Interest, ROI). Η οριοθέτηση αυτή πρέπει να γίνει με προσοχή και συγκεκριμένες προδιαγραφές. Το αντικείμενο ενδιαφέροντος θα πρέπει να απεικονίζεται μέσα σε ένα ορθογώνιο πλαίσιο ή να απομονωθεί εντελώς από την αρχική εικόνα, όντας πάλι σε ορθογώνιο σχήμα (Εικόνα 13). Επιπλέον, οι οριοθετημένες περιοχές πρέπει να έχουν όσο γίνεται παρόμοιο μέγεθος μεταξύ των εικόνων.



Εικόνα 13: Παραδείγματα φωτογραφιών στις οποίες έχει οριοθετηθεί η περιοχή ενδιαφέροντος

Για το σκοπό αυτό χρησιμοποιήθηκε ένα χρήσιμο και πολύ εύχρηστο εργαλείο της Matlab. Το εργαλείο αυτό είναι μία προσθήκη στη βασική έκδοση της Matlab και ονομάζεται Training Image Labeler. Χρησιμοποιώντας το ο χρήστης έχει τη δυνατότητα να προσδιορίσει ορθογώνιες περιοχές ενδιαφέροντος για οποιαδήποτε κατηγορία αντικειμένου. Επιπλέον, μετά τον προσδιορισμό μπορεί να εξάγει τις περιοχές που προσδιόρισε και να χρησιμοποιήσει πλέον το τελικό αρχείο για να εκπαιδεύσει ή να αξιολογήσει έναν ταξινομητή αντικειμένων. Η δυνατότητα που προσφέρει αυτό το εργαλείο και αποδείχτηκε πολύ χρήσιμη κατά τη διαδικασία του προσδιορισμού του στόματος στις φωτογραφίες που συλλέχθηκαν, είναι η αντιγραφή και η επικόλληση της ορθογώνιας περιοχής από τη μία εικόνα στην άλλη ακριβώς στην ίδια θέση. Έτσι, καθώς οι φωτογραφίες ήταν παρόμοιες μεταξύ τους, εξοικονομήθηκε πολύς χρόνος και εξασφαλίστηκε ότι οι οριοθετημένες περιοχές θα έχουν το ίδιο μέγεθος, με λίγες εξαιρέσεις όπου είναι απαραίτητο.

Αφού ολοκληρώθηκε η διαδικασία προσδιορισμού των περιοχών ενδιαφέροντος, τα αποτελέσματα εξήχθησαν στο τελικό mat αρχείο που πλέον περιείχε όλη την πληροφορία για κάθε φωτογραφία (Εικόνα 14). Η συνολική πληροφορία έχει τη μορφή μιας δομής με πέντε πεδία. Στο ένα πεδίο, μεταξύ αυτών, περιέχεται η τοποθεσία της εκάστοτε φωτογραφίας μέσα στον υπολογιστή και σε ένα άλλο οι συντεταγμένες της περιοχής ενδιαφέροντος που προσδιορίστηκε σε αυτή. Το αρχείο ήταν πλέον έτοιμο να εισαχθεί στη διαδικασία της εκπαίδευσης.



Fields	imageFilename	objectBoundingBoxes	ImageIcon	ImageLabel	ImagePlaceholder
1	'D:\open\0.pgm'	[303 174 69 64]	1x1 ImageIcon	'0'	1x1 ImageIcon
2	'D:\open\1.pgm'	[302 174 71 67]	1x1 ImageIcon	'1'	1x1 ImageIcon
3	'D:\open\2.pgm'	[302 174 71 67]	1x1 ImageIcon	'2'	1x1 ImageIcon
4	'D:\open\3.pgm'	[302 174 71 67]	1x1 ImageIcon	'3'	1x1 ImageIcon
5	'D:\open\4.pgm'	[302 174 71 67]	1x1 ImageIcon	'4'	1x1 ImageIcon
6	'D:\open\5.pgm'	[302 174 71 67]	1x1 ImageIcon	'5'	1x1 ImageIcon
7	'D:\open\6.pgm'	[302 174 71 67]	1x1 ImageIcon	'6'	1x1 ImageIcon
8	'D:\open\7.pgm'	[302 179 71 67]	1x1 ImageIcon	'7'	1x1 ImageIcon
9	'D:\open\8.pgm'	[307 179 71 67]	1x1 ImageIcon	'8'	1x1 ImageIcon
10	'D:\open\9.pgm'	[307 179 71 67]	1x1 ImageIcon	'9'	1x1 ImageIcon
11	'D:\open\10.pgm'	[307 179 71 67]	1x1 ImageIcon	'10'	1x1 ImageIcon
12	'D:\open\11.pgm'	[307 179 71 67]	1x1 ImageIcon	'11'	1x1 ImageIcon

Εικόνα 14: Παράδειγμα της δομής του τελικού αρχείου που περιέχει το σύνολο των οριοθετημένων περιοχών ενδιαφέροντος από τις θετικές εικόνες.

Για τις αρνητικές εικόνες η διαδικασία δεν ήταν τόσο πολύπλοκη. Αρκούσε μόνο να υπάρχουν συγκεντρωμένες σε έναν φάκελο ώστε να μπορεί να προσδιοριστεί η τοποθεσία τους στον υπολογιστή και να αποθηκευτούν, για τις ανάγκες της εκπαίδευσης, σε μία συλλογή δεδομένων εικόνας.

Εστιάζοντας στην εκπαίδευση των ταξινομητών, αυτή έγινε χρησιμοποιώντας την `trainCascadeObjectDetector` μέθοδο που προσφέρει η Matlab. Για να επιτευχθεί αποδεκτή ακρίβεια του ταξινομητή θα πρέπει να γίνει προσεκτική επιλογή διάφορων παραμέτρων της μεθόδου όπως ο αριθμός των σταδίων στα οποία θα γίνει η εκπαίδευση, το πλήθος των δειγμάτων που θα χρησιμοποιηθούν, θετικά και αρνητικά, και το μοντέλο που ταιριάζει περισσότερο στο αντικείμενο που στοχεύεται να αναγνωριστεί (Πίνακας 3). Κατά την επιλογή των παραπάνω παραμέτρων λήφθηκαν υπόψη κάποια αντισταθμίσματα που αφορούν και επηρεάζουν το ποσοστό των αληθινά και λανθασμένα θετικών αναγνωρίσεων και την ποσότητα των σταδίων από τα οποία θα περάσει η αναγνώριση. Πιο συγκεκριμένα ισχύουν οι παρακάτω συνεπαγωγές:

Μεγάλο δείγμα από εικόνες (σε κλίμακα των χιλιάδων) -> Αύξηση του αριθμού των σταδίων και του ποσοστού λαθεμένων θετικών αναγνωρίσεων ανά στάδιο

Μικρό δείγμα από εικόνες -> Μείωση του αριθμού των σταδίων και του ποσοστού λαθεμένων θετικών αναγνωρίσεων ανά στάδιο

Μείωση της πιθανότητας παράληψης αναγνώρισης -> Αύξηση του ποσοστού αληθινών θετικών αναγνωρίσεων, με το ρίσκο να μην επιτευχθεί το επιθυμητό ποσοστό λαθεμένων θετικών αναγνωρίσεων ανά στάδιο, κάνοντας τον ταξινομητή τελικά επιρρεπή σε λαθεμένες αναγνωρίσεις

Μείωση του αριθμού λαθεμένων αναγνωρίσεων -> Αύξηση του αριθμού των σταδίων ή μείωση του ποσοστού λαθεμένων αναγνωρίσεων ανά στάδιο

Η ίδια η μέθοδος υπολογίζει τον αριθμό των θετικών δειγμάτων που χρειάζεται και σε κάθε στάδιο εκπαίδευσης του ταξινομητή ακολουθείται διαφορετική διαδικασία. Στο πρώτο στάδιο χρησιμοποιούνται τα θετικά δείγματα που υπολογίστηκαν ότι είναι απαραίτητα σε σχέση πάντα με τις υπόλοιπες παραμέτρους που επιλέχθηκαν. Τα δείγματα αυτά είναι λιγότερα από εκείνα που εξ αρχής παρείχε ο χρήστης για τη διαδικασία. Επίσης χρησιμοποιούνται και τα αρνητικά δείγματα που παρήχθησαν από τον ίδιο τον αλγόριθμο με βάση τις αρνητικές εικόνες με τις οποίες τον τροφοδότησε ο χρήστης. Στα επόμενα στάδια χρησιμοποιούνται τα αποτελέσματα από όλα τα προηγούμενα, ταξινομούνται τα θετικά δείγματα και απορρίπτονται αυτά που ταξινομήθηκαν λάθος σαν αρνητικά. Από τα θετικά δείγματα που έχουν απομείνει χρησιμοποιούνται τόσα όσα είχαν χρησιμοποιηθεί και στο πρώτο στάδιο. Τέλος, παράγονται αρνητικά δείγματα μέσω επεξεργασίας των αρνητικών εικόνων με τη μέθοδο της κύλισης παραθύρου αλλά και μέσω χρήσης των λαθεμένων θετικών ταξινομημένων δειγμάτων. [21]

Έχοντας υπόψη όλα τα παραπάνω, εκπαιδεύτηκαν δύο είδη ταξινομητών για κάθε μία από τις επιθυμητές εκφράσεις, ταξινομητές τύπου Haar και ταξινομητές τύπου Hog. Η μέθοδος που προσφέρει το εργαλείο της Matlab για το σκοπό αυτό παίρνει σαν ορίσματα το όνομα του ταξινομητή που θα παραχθεί με την επέκταση xml, το αρχείο που περιέχει τα θετικά δείγματα εικόνων με την επέκταση mat, τη διαδρομή του φακέλου που περιέχει τις αρνητικές εικόνες, το ποσοστό λάθους και το πλήθος των σταδίων εκπαίδευσης. Χρησιμοποιώντας αυτά τα ορίσματα το αποτέλεσμα της μεθόδου είναι ένας ταξινομητής τύπου HOG. Για να προκύψουν οι Haar ταξινομητές προστέθηκε άλλον ένα

όρισμα στην εντολή, η λέξη haar. Παρακάτω φαίνονται αναλυτικά τα δεδομένα που χρησιμοποιήθηκαν για την εκπαίδευση και των δύο τύπων ταξινομητών (Εικόνα 15).

	Χαρούμενο Στόμα	Λυπημένο Στόμα	Ανοιχτό στόμα
Θετικά Δείγματα	1593	1880	1664
Αρνητικά Δείγματα	11527	12074	11636
Ποσοστό Λάθους	0.2	0.2	0.2
Στάδια	6	6	6

Πίνακας 3: Δεδομένα που χρησιμοποιήθηκαν κατά την εκπαίδευση καθενός από τους ταξινομητές.

```

Automatically setting objectTrainingSize to [ 36, 32 ]
Using at most 1186 of 1258 positive samples per stage
Using at most 2372 negative samples per stage

Training stage 1 of 13
[.....]
Used 1186 positive and 2372 negative samples
Time to train stage 1: 5 seconds

Training stage 2 of 13
[.....]
Used 1186 positive and 2372 negative samples
Time to train stage 2: 13 seconds

Training stage 3 of 13
[.....]
Used 1186 positive and 2372 negative samples
Time to train stage 3: 27 seconds

Training stage 4 of 13
[.....]
Used 1186 positive and 2372 negative samples
Time to train stage 4: 158 seconds
    
```

Εικόνα 15: Απόσπασμα από τα στάδια εκπαίδευσης ταξινομητή με τη συνάρτηση trainCascadeObjectDetector.

Τελικά, με βάση όσα έχουν αναφερθεί παραπάνω για τη σύγκριση μεταξύ των μοντέλων HOG και Haar αλλά και λαμβάνοντας υπόψη τις δοκιμές που έγιναν με την εφαρμογή και των δύο ειδών ταξινομητών πάνω στον αλγόριθμο, αποφασίστηκε να χρησιμοποιηθούν οι ταξινομητές τύπου HOG, αφού επιτυγχάνουν το σκοπό τους και προσφέρουν μεγάλα ποσοστά αναγνώρισης (Εικόνα 16).



(α)

(β)



(γ)

Εικόνα 16: Παραδείγματα ανίχνευσης α) χαρούμενου, β) λυπημένου και γ) ανοιχτού στόματος με χρήση ταξινομητών Haar. Οι ανιχνεύσεις δεν είναι οι επιθυμητές.

4.2 Λογισμικό Μέρος

Σαν πρωταρχικός στόχος, στην παρούσα διπλωματική εργασία, ήταν να δημιουργηθεί μία εφαρμογή η οποία θα υποστηρίζεται και θα λειτουργεί εξ' ολοκλήρου στον μικροϋπολογιστή BeagleBoard-xM. Το γεγονός, όμως, ότι η εν λόγω εφαρμογή έχει μεγάλες απαιτήσεις σε υπολογιστική ισχύ, για να μπορεί να προσφέρει πραγματικού χρόνου αποτελέσματα, οδήγησε στην ανάγκη για χρήση ενός υπολογιστή με αυξημένες υπολογιστικές δυνατότητες. Εάν, όμως, ο μικροϋπολογιστής έχανε εντελώς το ρόλο του στην εφαρμογή, θα χανόταν και η έννοια του ενσωματωμένου συστήματος και κατ' επέκταση του αυτόνομου παιχνιδιού, που εξ' αρχής θέλαμε να κατασκευάσουμε.

Αποφασίστηκε, λοιπόν, ο μικροϋπολογιστής να εκτελεί όλες τις λειτουργίες εισόδου και εξόδου δεδομένων, ο υπολογισμός των αποτελεσμάτων να ανατεθεί στον πιο ισχυρό υπολογιστή και οι δύο αυτές οντότητες να αλληλεπιδρούν μεταξύ τους και να ανταλλάζουν δεδομένα.

4.2.1 Αυτοματοποίηση Διαδικασίας Σύνδεσης στο Διαδίκτυο

Το λούτρινο παιχνίδι που κατασκευάστηκε αποτελεί ένα ενσωματωμένο σύστημα. Συνεπώς, ο χρήστης δεν έχει τη δυνατότητα να χρησιμοποιήσει το μικροϋπολογιστή σαν έναν κανονικό υπολογιστή, να έχει πρόσβαση σε γραφικό περιβάλλον και να κάνει αλλαγές, όπως για παράδειγμα να τον συνδέσει στο διαδίκτυο ή να θέσει σε λειτουργία κάποιον αλγόριθμο. Όλες αυτές οι λειτουργίες έπρεπε να αυτοματοποιηθούν ώστε να καταστήσουν την κατασκευή όσο το δυνατόν καλύτερα διαμορφωμένη, με το χρήστη να πρέπει να κάνει τις λιγότερες δυνατές ενέργειες για να την ενεργοποιήσει.

Για να είναι λειτουργικός ο μικροϋπολογιστής και να επιτελέσει τις ενέργειες που απαιτούνται από το ολοκληρωμένο σύστημα θα πρέπει να είναι συνδεδεμένος στο διαδίκτυο και να εκτελέσει τον αλγόριθμο που είναι υπεύθυνος να δέχεται μηνύματα από τον υπολογιστή και να αναπαράγει τον ήχο που αντιστοιχεί σε κάθε συναίσθημα. Εφόσον η κατασκευή είναι φορητή θα πρέπει η σύνδεσή της στο διαδίκτυο να μπορεί να επιτευχθεί ασύρματα και επιπλέον να υπάρχει δυνατότητα αλλαγής των στοιχείων σύνδεσης. Ακόμα, δεδομένου ότι δεν είναι επιθυμητό ο μικροϋπολογιστής να βρίσκεται μόνιμα σε λειτουργία θα πρέπει κάθε φορά που το παιχνίδι επανεκκινείται να εκκινείται αυτόματα και ο αλγόριθμος, ώστε να μπορεί να γίνει άμεση σύνδεση με τον υπολογιστή.

Προκειμένου να λειτουργήσει, λοιπόν, για πρώτη φορά το παιχνίδι είναι απαραίτητο να δηλωθούν τα στοιχεία σύνδεσης στο ασύρματο δίκτυο. Για να διευκολύνεται ο χρήστης να καταχωρεί τα στοιχεία που θα χρησιμοποιεί για τη σύνδεση του μικροϋπολογιστή στο wifi, τα οποία μπορεί να αλλάζουν είτε λόγω αλλαγής τοποθεσίας, είτε λόγω αλλαγής ρυθμίσεων, δημιουργήθηκε μία φόρμα εισαγωγής δεδομένων. Η φόρμα αυτή βρίσκεται ανεβασμένη στο χώρο που προσφέρει το πανεπιστήμιο και είναι προσβάσιμη από οποιαδήποτε συσκευή, εφόσον είναι

συνδεδεμένη στο διαδίκτυο. Τα πεδία που διαθέτει προς συμπλήρωση είναι το όνομα του ασύρματου δικτύου και ο μυστικός κωδικός πρόσβασης σε αυτό (Εικόνα 17).

Set Your Wi-Fi

Wi-Fi Name

Wi-Fi Password

Εικόνα 17: Φόρμα εισαγωγής στοιχείων σύνδεσης

Η φόρμα επιτρέπει στο χρήστη να μην εισάγει κωδικό πρόσβασης, καθώς μπορεί να συνδέεται σε ανοιχτό δίκτυο, αλλά δεν επιτρέπει την καταχώριση δεδομένων χωρίς τη συμπλήρωση του ονόματος του δικτύου (Εικόνα 18).

Set Your Wi-Fi

Wi-Fi Name

Wi-Fi Password

! Please fill out this field.

Εικόνα 18: Μήνυμα αποτροπής καταχώρισης σύνδεσης με κενό όνομα δικτύου

Με το πάτημα του κουμπιού «Save» και δεδομένου ότι το όνομα του δικτύου έχει συμπληρωθεί, τα στοιχεία αποθηκεύονται σε δύο διαφορετικά αρχεία κειμένου. Αυτό συμβαίνει για να μπορεί να διαχωρίζεται αν ο χρήστης δεν καταχώρησε κωδικό πρόσβασης. Σε περίπτωση, λοιπόν, που το αρχείο στο οποίο αποθηκεύεται ο κωδικός πρόσβασης είναι κενό, συμπεραίνεται ότι επιδιώκεται η σύνδεση σε κάποιο ανοιχτό δίκτυο. Η πληροφορία αυτή έχει σημασία για τη διαμόρφωση της σύνδεσης στη συνέχεια.

Για να επιτευχθεί τελικά η σύνδεση του μικροϋπολογιστή στο διαδίκτυο έχει προγραμματιστεί, κατά την εκκίνησή του, να τρέχει ένας bash αλγόριθμος που θα

διαβάζει τις τιμές από τα αρχεία που αναφέρθηκαν παραπάνω. Προκειμένου ο μικροϋπολογιστής να έχει πρόσβαση στα αρχεία αυτά είναι απαραίτητο να έχει πρόσβαση στο διαδίκτυο. Για τον παραπάνω λόγο, αλλά και για να γνωρίζει ο αλγόριθμος πότε επιδιώκεται διαμόρφωση στοιχείων σύνδεσης, έγινε η παραδοχή ότι, για τις περιπτώσεις αυτές, το παιχνίδι θα πρέπει να είναι συνδεδεμένο στο internet μέσω καλωδίου Ethernet. Κάθε άλλη φορά που το παιχνίδι θα τίθεται σε λειτουργία, θα συνδέεται στο διαδίκτυο ασύρματα. Ο αλγόριθμος έχει προγραμματιστεί να τρέχει με την εκκίνηση του συστήματος, ώστε να μην απαιτείται ενέργεια του χρήστη.

Κατά συνέπεια, ο πρώτος έλεγχος που πραγματοποιείται μέσα στον αλγόριθμο είναι το αν υπάρχει συνδεδεμένο καλώδιο Ethernet ώστε να γνωρίζει αν θα προχωρήσει στην διαμόρφωση δικτύου. Σε περίπτωση που δεν βρεθεί συνδεδεμένο καλώδιο Ethernet, το μόνο που κάνει ο αλγόριθμος είναι να θέτει σε λειτουργία το πρόγραμμα για τη ζεύξη της επικοινωνίας με τον υπολογιστή. Ο έλεγχος αυτός γίνεται επίσης για να βεβαιωθεί ότι υπάρχει πρόσβαση στο διαδίκτυο και συνεπώς μπορεί να προχωρήσει και να εφαρμόσει cURL στις διευθύνσεις URL των αρχείων που περιέχουν τα στοιχεία της νέας ασύρματης σύνδεσης στο διαδίκτυο.

Αφού εφαρμοστεί το εργαλείο cURL επιβεβαιώνεται ότι η τιμή που επιστράφηκε και αντιστοιχεί στο όνομα του δικτύου δεν είναι κενή. Ο έλεγχος για την τιμή αυτή ήδη πραγματοποιήθηκε, όπως αναφέρθηκε παραπάνω, μέσω της ίδιας της φόρμας εισαγωγής δεδομένων. Παρόλα αυτά επιδιώχθηκε να αποφευχθεί τυχόν κενή τιμή λόγω συμπλήρωσης της φόρμας, παραδείγματος χάριν, μέσω φυλλομετρητή που δεν υποστηρίζει τη λειτουργία για τον έλεγχο κενών απαιτούμενων πεδίων κατά την υποβολή της φόρμας. Σε περίπτωση που το όνομα του δικτύου μείνει κενό, η διαδικασία δεν ακολουθείται κανονικά ώστε το παιχνίδι να συνδεθεί ασύρματα στο διαδίκτυο, αλλά ο χρήστης θα πρέπει να διατηρήσει το καλώδιο του Ethernet συνδεδεμένο.

Για να διαμορφωθεί τελικά η σύνδεση πρέπει να δημιουργηθεί ένα αρχείο `wpa_supplicant.conf` μέσα στον `etc` φάκελο του μικροϋπολογιστή. Το αρχείο αυτό περιέχει πληροφορίες για τα χαρακτηριστικά του δικτύου, μεταξύ των οποίων είναι το όνομά του, το αν αυτό είναι ασφαλισμένο ή όχι και ο κωδικός πρόσβασής του. Κάθε φορά που ο αλγόριθμος μπαίνει σε διαδικασία διαμόρφωσης δικτύου και αφού έχει βεβαιωθεί ότι ο χρήστης έχει εισάγει τις απαραίτητες πληροφορίες για το δίκτυό του,

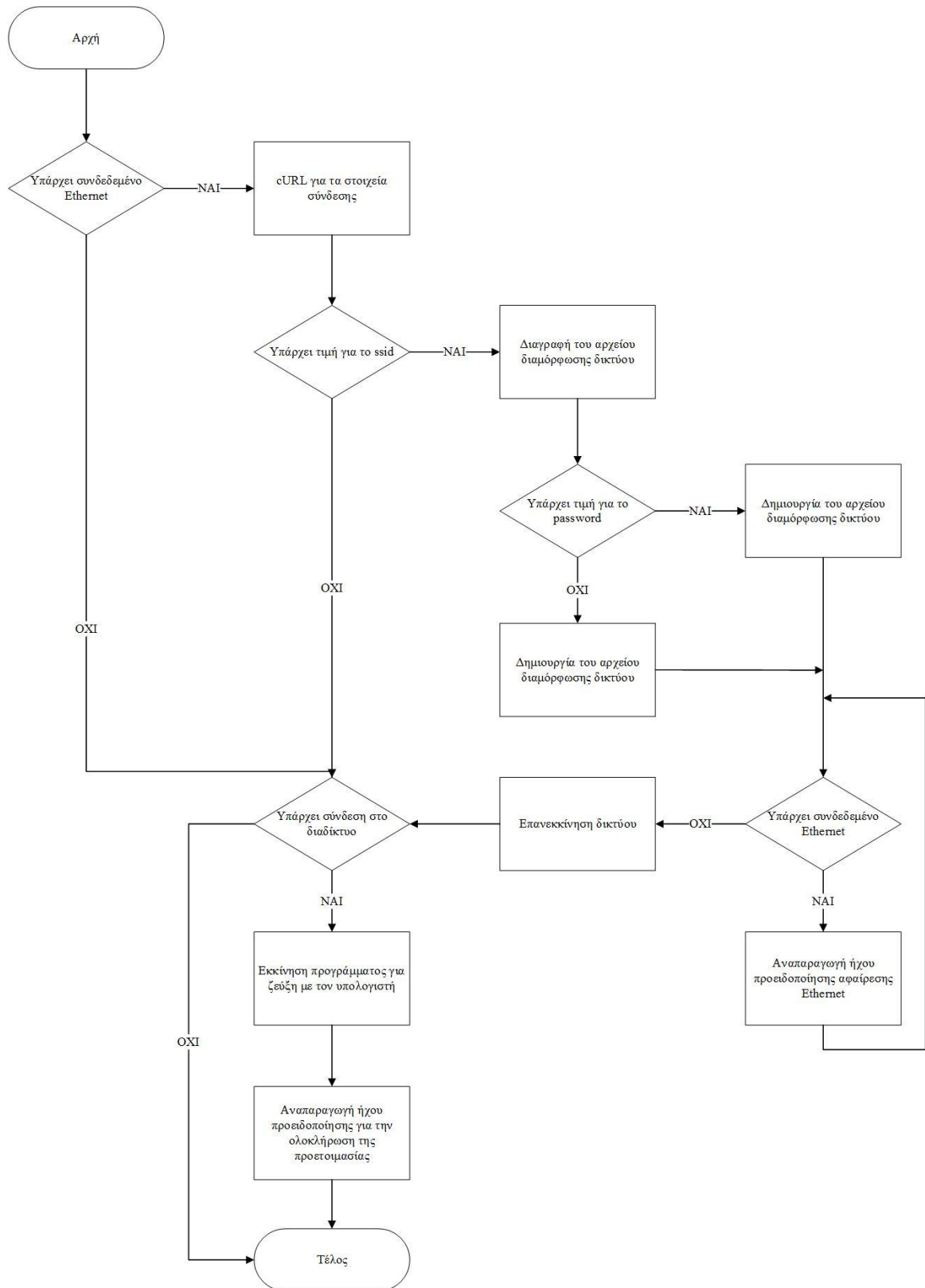
σβήνει το αρχείο αυτό και το δημιουργεί από την αρχή με τα στοιχεία για τη νέα σύνδεση.

Από την ύπαρξη ασφάλειας στο καινούριο δίκτυο εξαρτάται και το αν το αρχείο διαμόρφωσής του θα πρέπει να περιέχει κωδικό πρόσβασης και είδος πρωτοκόλλου. Συνεπώς, ο αλγόριθμος ελέγχει αν η τιμή που ανακτήθηκε από το αρχείο για τον κωδικό πρόσβασης είναι κενή και δημιουργεί το αρχείο με τα ανάλογα περιεχόμενα.

Αφού δημιουργηθεί το αρχείο, ο μικρουπολογιστής είναι έτοιμος να επανεκκινήσει το δίκτυό του για να εγκαθιδρύσει την καινούρια σύνδεση στο διαδίκτυο. Στο σημείο αυτό πρέπει να αφαιρεθεί το καλώδιο Ethernet από τη συσκευή για να μπορεί αυτή πλέον να συνδέεται ασύρματα. Δεδομένου του ότι ο χρήστης δεν μπορεί να ξέρει την ακριβή χρονική στιγμή στην οποία θα έχει ολοκληρωθεί η παραπάνω διαδικασία για να αφαιρέσει το καλώδιο, θα πρέπει με κάποιον τρόπο να ενημερωθεί. Για το σκοπό αυτό, ο αλγόριθμος, στο συγκεκριμένο σημείο, αναπαράγει έναν ήχο ειδοποίησης για το χρήστη, ότι είναι έτοιμος να αποσυνδέσει το καλώδιο. Ο ήχος αναπαράγεται συνεχόμενα, μέχρις ότου ανιχνευτεί από τον αλγόριθμο ότι δεν υπάρχει πια ενσύρματη σύνδεση.

Από εκεί και έπειτα η ροή του αλγορίθμου συνεχίζεται κανονικά, πράγμα που σημαίνει ότι επανεκκινείται το δίκτυο του μικροϋπολογιστή για να τεθεί σε λειτουργία με τις νέες ρυθμίσεις. Τέλος, ξεκινάει το πρόγραμμα που πραγματοποιεί τη ζεύξη του υπολογιστή με το παιχνίδι, ώστε να μπορέσει πλέον το ολοκληρωμένο σύστημα να λειτουργήσει κανονικά και να επιτύχει τους στόχους του. Ο χρήστης ενημερώνεται με έναν προειδοποιητικό ήχο ότι η συσκευή είναι έτοιμη προς χρήση και έτσι γνωρίζει πότε πρέπει να εκκινήσει το πρόγραμμα που θα εκτελείται στον υπολογιστή για την ανίχνευση συναισθήματος. Η όλη παραπάνω διαδικασία παρουσιάζεται στο διάγραμμα στην Εικόνα 19.

Σχεδιασμός και υλοποίηση ενός έξυπνου παιχνιδιού για τη συναισθηματική παρακολούθηση και το γονικό έλεγχο ενός παιδιού

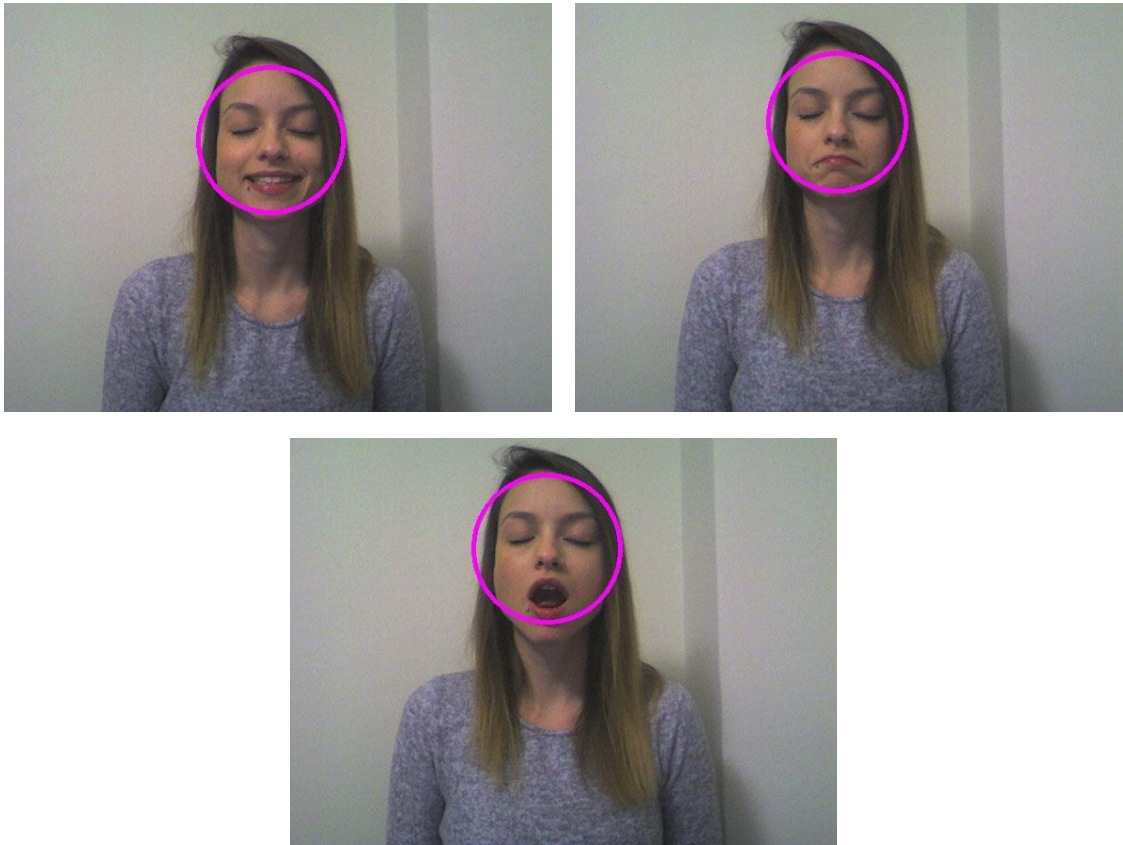


Εικόνα 19: Διάγραμμα ροής που περιγράφει τη διαδικασία προετοιμασίας του παιχνιδιού

4.2.2 Ανίχνευση συναισθήματος

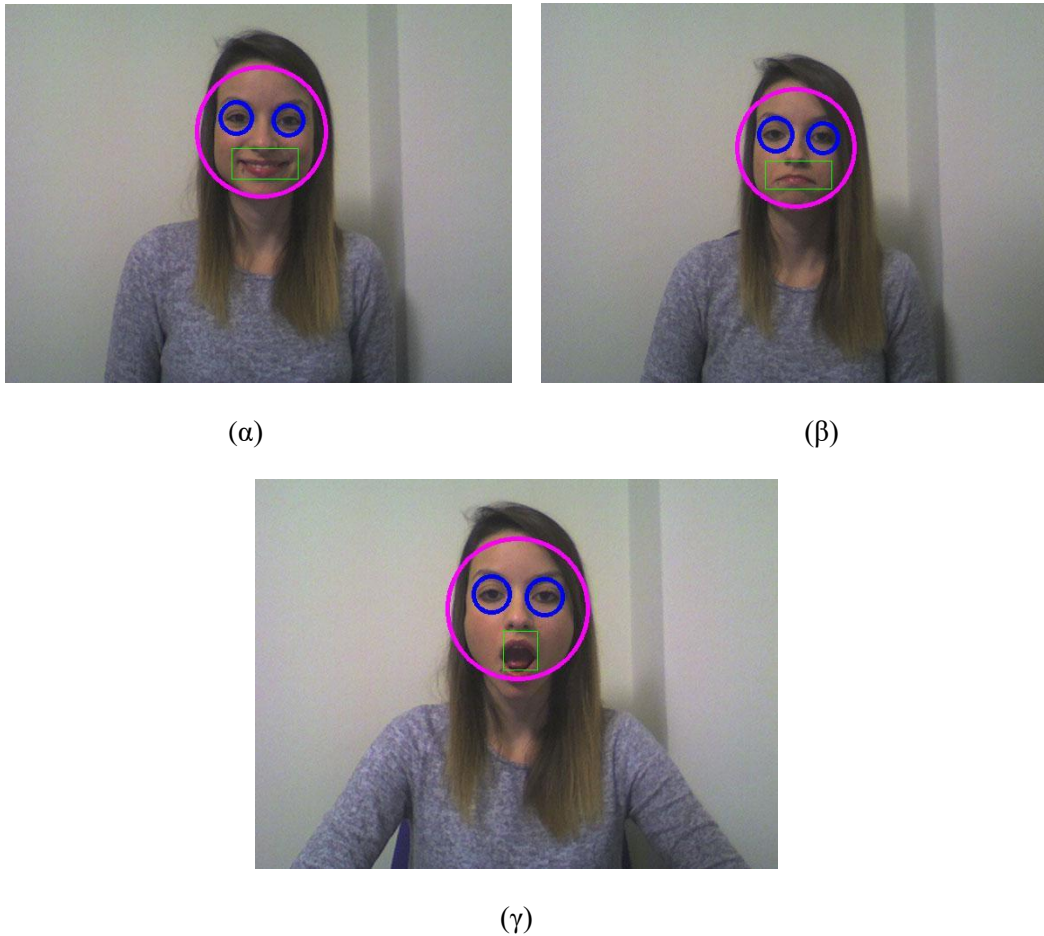
Ξεκινώντας από τη μεριά του υπολογιστή και εφόσον έχουν εκπαιδευτεί οι ταξινομητές και έχουν παραχθεί τα τελικά xml αρχεία, είναι έτοιμα να φορτωθούν στον αλγόριθμο. Στο σύνολό τους, οι ταξινομητές που χρησιμοποιούνται είναι πέντε, ένας για ανίχνευση προσώπου, ένας για ανίχνευση ματιών και τρεις για τις τρεις καταστάσεις του στόματος, χαμογελαστό, λυπημένο και ανοιχτό. Οι ταξινομητές φορτώνονται και ο αλγόριθμος τροφοδοτείται με την εικόνα που λαμβάνεται από την κάμερα που βρίσκεται συνδεδεμένη στον μικροϋπολογιστή. Για να ξεκινήσει η αναγνώριση η εικόνα μετατρέπεται από έγχρωμη σε ασπρόμαυρη και αποθηκεύεται σε έναν πίνακα με τη μορφή αριθμών. Άλλωστε και οι ταξινομητές με τους οποίους θα συγκριθεί στη συνέχεια έχουν τη μορφή αριθμών και με τον τρόπο αυτό τα δύο αυτά στοιχεία γίνονται συμβατά μεταξύ τους.

Η πρώτη ανίχνευση που επιδιώκεται να γίνει στην εικόνα είναι η ανίχνευση προσώπου, αφού χωρίς να υπάρχει πρόσωπο δεν υπάρχει νόημα να προχωρήσει ο αλγόριθμος στην ανίχνευση κάποιου συναισθήματος. Με την προϋπόθεση ότι έχει ανιχνευτεί κάποιο πρόσωπο, ο αλγόριθμος συνεχίζει και επιδιώκει αναγνώριση ματιών. Εδώ πρέπει να σημειωθεί ότι τα μάτια που εντοπίζονται στην εικόνα είναι μόνο τα ανοιχτά. Αυτό έχει ιδιαίτερη σημασία καθώς δεν είναι επιθυμητό να γίνει ανίχνευση στόματος και κατ' επέκταση να τεθεί κάποιος ήχος σε αναπαραγωγή την ώρα που το παιδί κοιμάται, αφού ανάμεσα στους ρόλους του παιχνιδιού είναι το νανούρισμα του παιδιού και η διασφάλιση της ηρεμίας του. Σε κάθε άλλη περίπτωση, η αναγνώριση των ματιών θα γινόταν χωρίς να προσφέρει κάποια βασική λειτουργικότητα στην εφαρμογή. Κάθε φορά που εντοπίζεται κάποιο πρόσωπο στην εικόνα, έχει οριστεί αυτό να οριοθετείται σε ένα σχήμα έλλειψης του οποίου οι διαστάσεις είναι σαφώς ορισμένες. Το ίδιο συμβαίνει για καθένα από τα δύο μάτια, μόνο που γι' αυτά το σχήμα που προτιμήθηκε είναι ο κύκλος (Εικόνα 20).



Εικόνα 20: Κανένα συναίσθημα δεν ανιχνεύεται σε περίπτωση που τα μάτια είναι κλειστά.

Αφού εξασφαλιστούν οι παραπάνω προϋποθέσεις, η εικόνα ελέγχεται προς ανίχνευση στόματος με συγκεκριμένη έκφραση. Οι μορφές του στόματος που οι ταξινομητές εκπαιδεύτηκαν να αναγνωρίζουν υπάρχει περίπτωση να ανιχνευτούν λαθεμένα σε άλλα σημεία της εικόνας. Για παράδειγμα ο ταξινομητής του ανοιχτού στόματος μπορεί να θεωρήσει ότι βρίσκει ταύτιση σε κάποιο μάτι. Συνεπώς, ενδέχεται να καταλήξει να αναπαράγεται ένα τραγούδι χωρίς να έχει επιτευχθεί πραγματική ανίχνευση συναισθήματος. Για να διασφαλιστεί ότι δεν θα προκύψει τέτοια πιθανότητα λάθους, η περιοχή της εικόνας στην οποία οι ταξινομητές στόματος εφαρμόζονται οριοθετήθηκε και ορίστηκε να είναι η περιοχή στο κάτω μισό της περιοχής όπου ανιχνεύτηκε το πρόσωπο. Σε περίπτωση, λοιπόν, που εντοπιστεί στόμα χαρούμενο, λυπημένο ή ανοιχτό στην περιοχή της εικόνας που αναφέρθηκε παραπάνω, αυτό μαρκάρεται με ένα ορθογώνιο συγκεκριμένων διαστάσεων που ορίστηκαν ακριβώς μετά από δοκιμές (Εικόνα 21).



Εικόνα 21: Επιτυχείς αναγνώρισεις α) χαρούμενου, β) λυπημένου και γ) ανοιχτού στόματος

Κάθε φορά που ο αλγόριθμος ταυτίζει ένα στόμα στην εικόνα με έναν συγκεκριμένο ταξινομητή, δηλαδή όταν ένα συγκεκριμένο συναίσθημα ανιχνεύεται, ο αντίστοιχος μετρητής που ορίστηκε εξ' αρχής μηδενικός, αυξάνεται. Εκ των προτέρων έχει οριστεί ένα ανώτατο όριο στο οποίο ο κάθε μετρητής μπορεί να φτάσει για να θεωρηθεί ότι το παιδί έχει μία συγκεκριμένη έκφραση στο πρόσωπό του. Όταν ο μετρητής φτάσει σε αυτό το όριο δημιουργείται ένα καινούριο νήμα στην εφαρμογή στο οποίο πραγματοποιείται η σύνδεση του υπολογιστή με τον μικροϋπολογιστή ώστε να ενημερωθεί η δεύτερη για το ποιο συναίσθημα ανιχνεύτηκε και κατά συνέπεια ποιο τραγούδι θα θέσει σε αναπαραγωγή. Το πώς επιτυγχάνεται η σύνδεση αυτή θα περιγραφεί σε επόμενη ενότητα.

Το καινούριο νήμα είναι σημαντικό να δημιουργηθεί, έτσι ώστε ανεξάρτητα με το τι συμβαίνει μέσα σε αυτό, η εφαρμογή να συνεχίζει κανονικά να λαμβάνει εικόνα και να

εφαρμόζει τον αλγόριθμο ανίχνευσης πάνω της. Κάτι τέτοιο είναι απαραίτητο γιατί, καθώς παίζει το τραγούδι για ένα συγκεκριμένο συναίσθημα, η έκφραση του προσώπου του παιδιού μπορεί να αλλάξει, δηλώνοντας ακόμα και κάποια αποστροφή προς το συγκεκριμένο ήχο. Αυτή η έκφραση, λοιπόν, πρέπει να ανιχνευτεί και κατ' επέκταση να μπει σε αναπαραγωγή ο ήχος που τη συνοδεύει.

Μέσα στο νήμα που δημιουργείται, καλείται η συνάρτηση `posix_spawn()` για να εκτελεστεί το πρόγραμμα που είναι υπεύθυνο για τη ζεύξη του υπολογιστή με τον μικροϋπολογιστή. Μέσω του εν λόγω προγράμματος γίνεται η αποστολή του αναγνωριστικού του τραγουδιού από τον υπολογιστή και έτσι ο μικροϋπολογιστής ειδοποιείται για το ποιο τραγούδι πρέπει να αναπαράγει μετά την ανίχνευση συναισθήματος. Στη θέση αυτής της συνάντησης θα μπορούσε να έχει χρησιμοποιηθεί η συνάρτηση `fork()`. Στόχος και των δύο αυτών συναρτήσεων είναι να δημιουργήσουν μία θυγατρική διεργασία ή οποία θα εκτελείται παράλληλα με την κύρια διεργασία-γονέα και θα αλληλεπιδρά με αυτήν.

Αυτό που κάνει η συνάρτηση `fork()`, ωστόσο, είναι να αναγκάζει τον πυρήνα να αντιγράψει την ήδη υπάρχουσα διεργασία σε ένα καινούριο κομμάτι μνήμης, ορίζοντας έτσι την καινούρια διεργασία και στη συνέχεια να της δώσει πόρους και μνήμη για να μπορέσει να την υποστηρίξει. Η διαδικασία αυτή την καθιστά «βαριά» για το σύστημα με αποτέλεσμα να εισάγει καθυστέρηση στην εφαρμογή και να επιβαρύνει την απόκρισή της. Αντίθετα, η συνάρτηση `posix_spawn()` δεν αντιγράφει όλη τη μνήμη της διεργασίας για να δημιουργήσει μία θυγατρική και για το λόγο αυτό μπορεί να θεωρηθεί μία ελαφριά έκδοση της `fork()` και κατά συνέπεια επιλέχτηκε για την εφαρμογή.

Τα ορίσματα με τα οποία καλείται η `posix_spawn()`, είναι δύο και είναι αυτά που χρειάζεται το πρόγραμμα που πραγματοποιεί τη σύνδεση του υπολογιστή με τον μικροϋπολογιστή. Το πρώτο είναι η διεύθυνση IP του μικροϋπολογιστή και το δεύτερο είναι το αναγνωριστικό του τραγουδιού που θα πρέπει να μπει σε αναπαραγωγή ώστε το παιχνίδι τελικά να αλληλεπιδράσει με το παιδί μετά την ανίχνευση του συναισθήματός του.

4.2.3 Επικοινωνία υπολογιστή - μικροϋπολογιστή

Τυπικά δύο διαδικασίες επικοινωνούν μεταξύ τους σε ένα ενιαίο σύστημα μέσω μιας από τις ακόλουθες τεχνικές:

- Αγωγούς
- Ουρές Μηνυμάτων
- Κοινή μνήμη

Υπάρχουν και άλλοι τρόποι επικοινωνίας, με τους παραπάνω να είναι οι πιο συνήθεις. Βέβαια, δύο διαδικασίες μπορούν να επικοινωνήσουν και μέσω δικτύου. Για παράδειγμα, όταν φορτώνεται μία ιστοσελίδα, στο τοπικό σύστημα του χρήστη η διεργασία που τρέχει είναι η μηχανή αναζήτησης ιστού, ενώ στο απομακρυσμένο σύστημα, η διεργασία που τρέχει είναι ο κεντρικός υπολογιστής δικτύου. Κατά συνέπεια η διαδικασία αυτή αποτελεί, επίσης, παράδειγμα επικοινωνίας μεταξύ διεργασιών, αλλά η τεχνική αυτή τη φορά ονομάζεται Υποδοχή (Socket).

Ως υποδοχή ορίζεται ένα ακροσημείο επικοινωνίας μεταξύ δύο συστημάτων μέσα σε ένα δίκτυο. Πιο συγκεκριμένα, μία υποδοχή είναι ένας συνδυασμός διεύθυνσης IP και θύρας σε ένα σύστημα. Έτσι, σε κάθε σύστημα μια υποδοχή υφίσταται για μια διεργασία που αλληλεπιδρά με την υποδοχή ενός άλλου συστήματος μέσω δικτύου. Ένας συνδυασμός τοπικής υποδοχής και της υποδοχής ενός απομακρυσμένου συστήματος είναι γνωστός σαν «Τετραπλή Πλειάδα (4-tuple)». Λέγεται τετραπλή γιατί περιλαμβάνει τη διεύθυνση IP, τη θύρα του πελάτη υπολογιστή και το αντίστοιχο ζεύγος του διακομιστή. Μέσω αυτής μπορεί να προσδιοριστεί μοναδικά η σύνδεση δύο διεργασιών που τρέχουν σε διαφορετικά συστήματα.

Σε μία σύνδεση υποδοχής οι αριθμοί θύρας κωδικοποιούνται στην κεφαλίδα του πακέτου πρωτοκόλλου μεταφοράς και μπορούν εύκολα να ερμηνευτούν όχι μόνο από τους αποστολείς και παραλήπτες υπολογιστές, αλλά και από άλλα μέρη της υποδομής δικτύωσης. Ιδίως τα τείχη προστασίας είναι συνήθως διαμορφωμένα να διαχωρίζουν τα πακέτα της επικοινωνίας βασισμένα στους αριθμούς θυρών της προέλευσης και του προορισμού τους.

Σε έναν συγκεκριμένο συνδυασμό διεύθυνσης IP και αριθμού θύρας, μόνο μία διεργασία μπορεί να συνδεθεί, χρησιμοποιώντας το ίδιο πρωτόκολλο μεταφοράς. Σε διαφορετική περίπτωση, θα υπάρχουν συγκρούσεις ανάμεσα στα προγράμματα που προσπαθούν να συνδεθούν. Συνεπώς, σκοπός των θυρών είναι να διαφοροποιήσουν τις πολλαπλές υποδοχές που υπάρχουν σε μία διεύθυνση δικτύου.

4.2.3.1 Υποδοχές Ρευμάτων (Sockets)

Υπάρχουν πολλοί διαφορετικοί τύποι υποδοχών που καθορίζονται από τη δομή του στρώματος μεταφοράς. Οι πιο κοινοί τύποι είναι οι υποδοχές ρευμάτων και οι υποδοχές διαγραμμάτων δεδομένων.

Οι υποδοχές ρευμάτων προσφέρουν αξιόπιστη αμφίδρομη επικοινωνία παρόμοια με αυτή που έχουν δύο άτομα μέσω τηλεφώνου. Η μία πλευρά αρχικοποιεί τη σύνδεση με την άλλη και αφού αυτή εγκαθιδρύεται, οποιαδήποτε από τις δύο πλευρές μπορεί να επικοινωνήσει με την άλλη. Επιπλέον, υπάρχει άμεση επιβεβαίωση ότι το μήνυμα που στάλθηκε πράγματι έφτασε στον προορισμό του. Στις υποδοχές ρευμάτων χρησιμοποιείται ένα πρωτόκολλο ελέγχου μετάδοσης (TCP), το οποίο υπάρχει στο στρώμα μεταφοράς του προτύπου Διασύνδεσης Ανοιχτών Συστημάτων (Open Systems Interconnection - OSI). Τα δεδομένα συνήθως μεταδίδονται με τη μορφή πακέτων, τα οποία αποτελούνται από μία κεφαλίδα και από το περιεχόμενο του μηνύματος. Το πρωτόκολλο TCP είναι σχεδιασμένο έτσι ώστε τα πακέτα δεδομένων να φτάσουν στον προορισμό τους χωρίς λάθη και με τη σωστή σειρά (Εικόνα 22). Οι διακομιστές διαδικτύου, οι διακομιστές ταχυδρομείου και οι αντίστοιχες αιτήσεις πελατών τους χρησιμοποιούν πρωτόκολλο TCP και υποδοχές ρευμάτων για να επικοινωνήσουν.

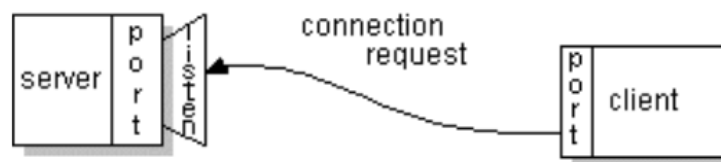
Στις υποδοχές διαγραμμάτων δεδομένων η σύνδεση είναι μονόδρομη και αναξιόπιστη. Για παράδειγμα, αν σταλεί μια σειρά από γράμματα δεν υπάρχει βεβαιότητα ότι αυτά θα παραληφθούν με την ίδια σειρά ή ακόμα περισσότερο αν αυτά θα καταφέρουν να φτάσουν στον προορισμό τους. Τα datagram socket χρησιμοποιούν User Datagram Protocol (UDP). Στην ουσία η σύνδεση που δημιουργείται δεν είναι πραγματική σύνδεση, αλλά μία απλή μέθοδος για αποστολή δεδομένων από ένα σημείο σε ένα άλλο.

TCP Segment Header Format								
Bit #	0	7	8	15	16	23	24	31
0	Source Port				Destination Port			
32	Sequence Number							
64	Acknowledgment Number							
96	Data Offset	Res	Flags			Window Size		
128	Header and Data Checksum				Urgent Pointer			
160...	Options							

Εικόνα 22: Περιεχόμενα κεφαλίδας ενός μηνύματος που μεταδίδεται μέσω πρωτοκόλλου TCP. Τα περιεχόμενα αυτά καθιστούν το πρωτόκολλο μετάδοσης TCP ασφαλές και αξιόπιστο.

4.2.3.2 Επικοινωνία Client/Server

Συνήθως ο εξυπηρετητής (server) τρέχει σε έναν συγκεκριμένο υπολογιστή και έχει μία υποδοχή συνδεδεμένη με μία συγκεκριμένη πόρτα (Εικόνα 23). Αυτό που κάνει είναι να περιμένει, ακούγοντας πάντα στην υποδοχή, περιμένοντας κάποιον πελάτη (client) να αιτηθεί για σύνδεση. Από τη μεριά του πελάτη, αυτός γνωρίζει το όνομα του υπολογιστή στον οποίο τρέχει ο εξυπηρετητής και τη θύρα στην οποία ακούει. Για να υποβάλλει ένα αίτημα σύνδεσης, ο πελάτης προσπαθεί να συναντηθεί με τον εξυπηρετητή στο συγκεκριμένο μηχάνημα και θύρα. Ο πελάτης επίσης είναι απαραίτητο να ταυτοποιήσει τον εαυτό του στον εξυπηρετητή και το σύστημα αναλαμβάνει να δεσμεύσει γι' αυτό το σκοπό μία τοπική θύρα που χρησιμοποιηθεί κατά τη διάρκεια της σύνδεσης.

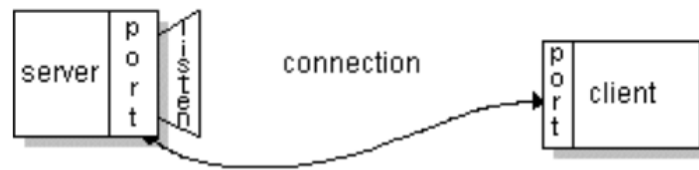


Εικόνα 23: Ο πελάτης αιτείται σύνδεσης στη θύρα που «ακούει» ο εξυπηρετητής

Αν όλα πάνε καλά, ο εξυπηρετητής δέχεται τη σύνδεση. Σε περίπτωση αποδοχής, ο εξυπηρετητής παίρνει μία καινούρια υποδοχή συνδεδεμένη στην ίδια τοπική θύρα και επιπλέον θέτει το απομακρυσμένο ακροσημείο του στη διεύθυνση και τη θύρα του

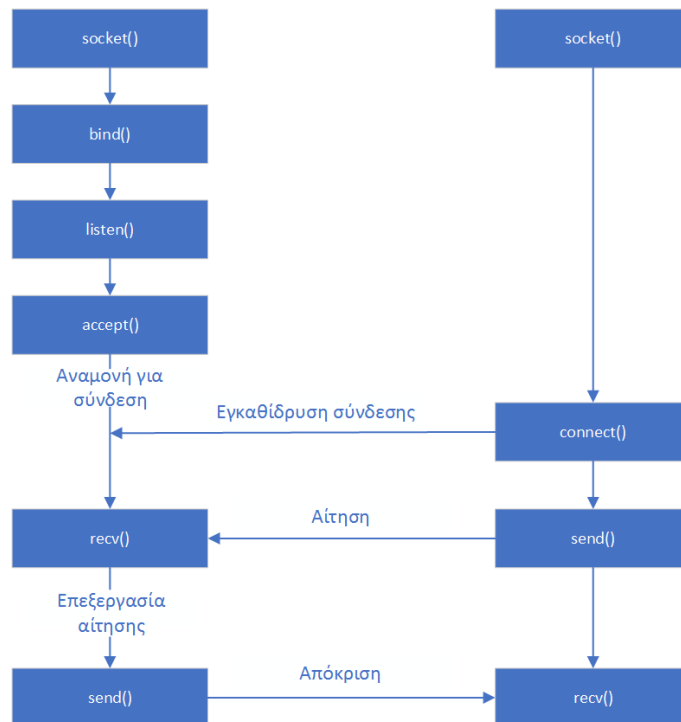
πελάτη. Χρειάζεται την καινούρια υποδοχή ώστε να μπορεί να συνεχίσει να ακούει στην αρχική για τυχόν αιτήσεις σύνδεσης, ενώσω εξυπηρετεί τις ανάγκες του ήδη συνδεδεμένου πελάτη.

Από τη μεριά του πελάτη, αν η σύνδεση γίνει αποδεκτή, μια υποδοχή δημιουργείται έτοιμη να τη χρησιμοποιήσει για να επικοινωνήσει με τον εξυπηρετητή (Εικόνα 24). [22]



Εικόνα 24: Η σύνδεση μεταξύ πελάτη και εξυπηρετητή είναι αμφίδρομη

Τελικά, σ' ένα δικτυακό μοντέλο εξυπηρετητή-πελάτη, οι πελάτες αιτούνται σε έναν εξυπηρετητή στέλνοντάς του μηνύματα και οι εξυπηρετητές απαντούν στους πελάτες τους για κάθε αίτηση και επιστρέφουν αποτελέσματα. Ένα παράδειγμα επικοινωνίας παρουσιάζεται στην Εικόνα 25.

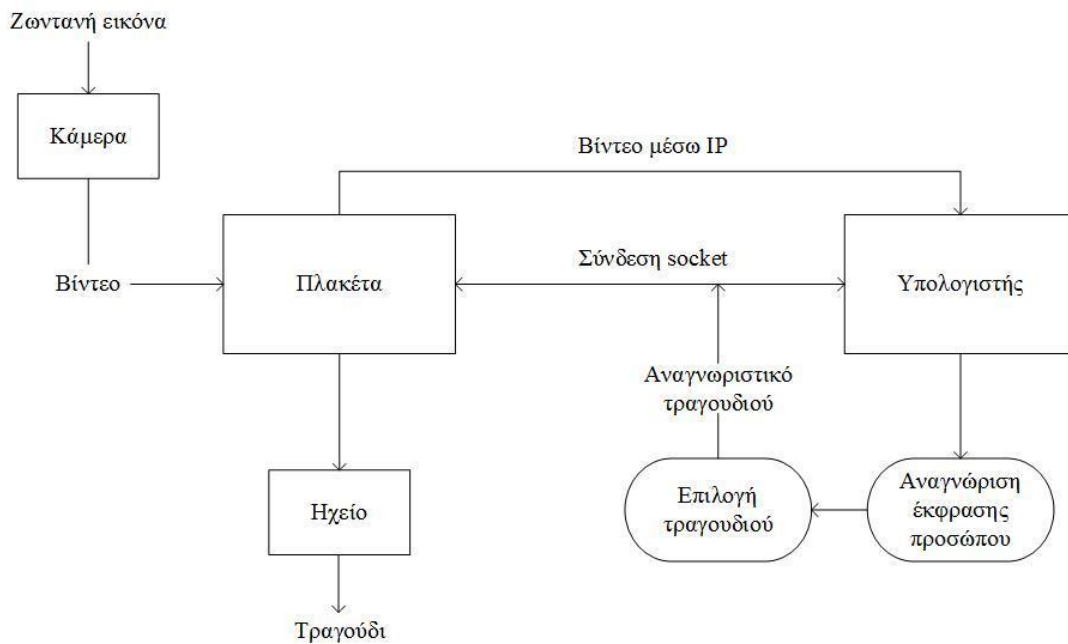


Εικόνα 25: Η επικοινωνία μεταξύ του πελάτη και του εξυπηρετητή μέσω υποδοχής ρευμάτων συμπεριλαμβανομένων όλων των βημάτων της διαδικασίας.

Ένας εξυπηρετητής, κατά γενική ομολογία, υποστηρίζει πολυάριθμους πελάτες και πολλαπλοί εξυπηρετητές μπορούν να συνδεθούν σε ένα δίκτυο για να διαχειριστούν τον αυξανόμενο επεξεργαστικό φόρτο, καθώς οι πελάτες γίνονται όλο και περισσότεροι. Στην περίπτωση πιο απλών εξυπηρετητών, αν μία σύνδεση έχει ήδη εγκαθιδρυθεί, οι περαιτέρω πελάτες περιμένουν σε ουρά αναμονής μέχρι να απελευθερωθεί ο εξυπηρετητής.

Μερικές από τις δημοφιλέστερες διαδικτυακές εφαρμογές ακολουθούν το μοντέλο πελάτη/εξυπηρετητή, συμπεριλαμβανομένων των υπηρεσιών ηλεκτρονικού ταχυδρομείου, FTP και ιστού. Καθένας από αυτούς τους πελάτες προσφέρουν μία διεπαφή χρήστη και μία εφαρμογή πελατών που επιτρέπει στο χρήστη να συνδεθεί με έναν εξυπηρετητή. Στις περιπτώσεις ηλεκτρονικού ταχυδρομείου και FTP ο χρήστης εισάγει στη διεπαφή ένα όνομα υπολογιστή ή μια διεύθυνση IP για να εγκαθιδρύσει συνδέσεις με τον εξυπηρετητή. [23]

Στην εφαρμογή που αναπτύχθηκε χρησιμοποιήθηκε το μοντέλο εξυπηρετητή-πελάτη με τον τύπο της υποδοχής ρευμάτων για να επιτευχθεί η επικοινωνία μεταξύ του υπολογιστή και του μικροϋπολογιστή. Στην επικοινωνία αυτή, τη θέση του εξυπηρετητή την έχει ο μικροϋπολογιστής και τη θέση του πελάτη την έχει ο υπολογιστής (Εικόνα 26). Ο μικροϋπολογιστής ως εξυπηρετητής ανοίγει μία υποδοχή έτοιμη να δεχτεί συνδέσεις από τον πελάτη υπολογιστή. Σκοπός της επικοινωνίας αυτής είναι να ανταλλάσσονται μηνύματα από τον υπολογιστή προς τον μικροϋπολογιστή τα οποία θα υποδηλώνουν κάθε φορά ποιο συναίσθημα έχει ανιχνευτεί από τον αλγόριθμο. Καθένα από αυτά τα μηνύματα είναι ένας αριθμός ο οποίος έχει αντιστοιχηθεί στο αντίστοιχο συναίσθημα, όπως αναφέρθηκε παραπάνω. Κάθε φορά που ο εξυπηρετητής δέχεται ένα τέτοιο μήνυμα, γνωρίζει ότι πρέπει να καλέσει τον αλγόριθμο που είναι υπεύθυνος για την αναπαραγωγή του τραγουδιού με όρισμα την «ταυτότητα» του τραγουδιού, δηλαδή τον αριθμό που έλαβε σαν μήνυμα από τον πελάτη. Με τον τρόπο αυτό αναπαράγεται το σωστό τραγούδι που συνοδεύει κάθε συναίσθημα.



Εικόνα 26: Στο ενσωματωμένο σύστημα που αναπτύχθηκε ο μικροϋπολογιστής αποφορτίζεται από την επεξεργασία του βίντεο για την ανίχνευση συναισθήματος. Η διαδικασία αυτή μετατάσσεται στον υπολογιστή. Το σύνολο των διεργασιών και της επικοινωνίας μεταξύ των οντοτήτων.

4.2.4 Ανταλλαγή διεύθυνσης IP

Στο σημείο αυτό θα πρέπει να σημειωθεί ο τρόπος με τον οποίο ο υπολογιστής-πελάτης γνωρίζει τη διεύθυνση IP του μικροϋπολογιστή-εξυπηρετητή μέσω της οποίας θα αιτηθεί σύνδεση στην υποδοχή του. Καθώς το σύστημα χαρακτηρίζεται από κάποια φορητότητα, ο υπολογιστής και το λούτρινο παιχνίδι μπορούν να μεταφέρονται και να συνδέονται κάθε φορά σε διαφορετικό δίκτυο. Για το λόγο αυτό, ο πελάτης δεν πρέπει να γνωρίζει μία συγκεκριμένη διεύθυνση εξυπηρετητή, αλλά θα πρέπει να μπορεί να τη λαμβάνει δυναμικά, κάθε φορά που αυτή αλλάζει. Βέβαια και από τη μεριά του ο εξυπηρετητής, για να εγκαθιδρύσει την υποδοχή, θα πρέπει να γνωρίζει την διεύθυνση IP που του δόθηκε. Στην περίπτωση του εξυπηρετητή δημιουργήθηκε μία συνάρτηση γραμμένη σε κώδικα C++, που επιστρέφει κάθε φορά την IP που λαμβάνει το μηχάνημα από το δίκτυο στο οποίο βρίσκεται συνδεδεμένο. Αυτό σαν διαδικασία είναι αρκετό, καθώς αφού εγκαθιδρυθεί η υποδοχή, παραμένει ανοιχτή έως ότου το σύστημα τερματιστεί.

Εκτός από την αποστολή του αναγνωριστικού του τραγουδιού από τον υπολογιστή στον μικροϋπολογιστή, ο πρώτος είναι αναγκαίο να γνωρίζει τη διεύθυνση IP του δεύτερου για έναν ακόμη λόγο. Μιας και όπως αναφέρθηκε παραπάνω, ο μικροϋπολογιστής είναι υπεύθυνος για όλες τις ενέργειες εισόδου εξόδου της εφαρμογής, είναι υπεύθυνος και για τη λήψη ροής εικόνας, μέσω κάμερας. Ο υπολογιστής, λοιπόν, αποκτά πρόσβαση στην εικόνα αυτή, μέσω της διεύθυνσης του μικροϋπολογιστή, για να μπορέσει στη συνέχεια να εφαρμόσει σε αυτήν τον αλγόριθμο που αναπτύχθηκε.

Για να λάβει όμως ο πελάτης τη διεύθυνση του εξυπηρετητή η διαδικασία είναι πιο περίπλοκη (Εικόνα 27). Ο μικροϋπολογιστής σαν εξυπηρετητής είναι υπεύθυνος να κάνει, με κάποιον τρόπο, γνωστή την IP του σε συσκευές που αναμένει να επιδιώξουν σύνδεση μαζί του. Επειδή δεν είναι γνωστό εκ των προτέρων κάθε πότε θα απενεργοποιείται/ενεργοποιείται το παιχνίδι και κατ' επέκταση κάθε πότε ο μικροϋπολογιστής θα λαμβάνει καινούρια διεύθυνση, προγραμματίστηκε να λαμβάνεται και να αποστέλλεται αυτή η διεύθυνση, ανά συγκεκριμένο χρονικό διάστημα, σε ένα αρχείο όπου ο εν δυνάμει πελάτης έχει πρόσβαση. Η λειτουργία αυτή πραγματοποιήθηκε με τη χρήση του εργαλείου `crontab`.

Το `cron` είναι ένα daemon του συστήματος που χρησιμοποιείται για να εκτελέσει, στο παρασκήνιο, επιθυμητά καθήκοντα σε ορισμένους χρόνους. Το όνομά του το έχει πάρει από τον Κρόνο, τον Έλληνα θεό του χρόνου.

Ένα αρχείο `crontab`, συντομογραφία του `cron table`, είναι ένα απλό αρχείο κειμένου που περιέχει μία λίστα από εντολές που προορίζονται να τρέξουν σε καθορισμένους χρόνους. Η επεξεργασία αυτού του αρχείου μπορεί να γίνει με την εντολή `crontab` και σ' αυτό μπορούν να προστεθούν, να αφαιρεθούν ή και να τροποποιηθούν προγραμματισμένες εργασίες. Οι εντολές στο αρχείο, αλλά και οι αντίστοιχοι χρόνοι τους ελέγχονται από το `cron daemon`, το οποίο τις εκτελεί στο παρασκήνιο. Κάθε χρήστης, συμπεριλαμβανομένου του `root`, έχει ένα `crontab` αρχείο. Το `cron daemon` ελέγχει το `crontab` αρχείο του χρήστη ανεξάρτητα από το αν ο χρήστης είναι συνδεδεμένος στο σύστημα. [24]

Για τις ανάγκες του παρόντος συστήματος δημιουργήθηκε ένα SH αρχείο μέσα στο οποίο υπάρχει μια εντολή για την εύρεση της διεύθυνσης IP του εξυπηρετητή κάθε ένα λεπτό και άλλη μία για την αποστολή της με τη χρήση του εργαλείου `cURL`.

Η μέθοδος POST που χρησιμοποιεί το εργαλείο αυτό, είναι ιδανική, καθώς η IP διεύθυνση του μικροϋπολογιστή, ως πληροφορία που μεταφέρεται, είναι ευαίσθητη και δεν θα πρέπει να είναι προσβάσιμη από οποιονδήποτε. Περισσότερες λεπτομέρειες για την ασφάλεια που διέπει το σύστημα θα αναφερθούν σε επόμενη ενότητα.

Για τη διαχείριση της πληροφορίας που αποστέλλεται με cURL, έγινε χρήση του περιβάλλοντος που προσφέρεται στους φοιτητές στον server του πανεπιστημίου. Εκεί όντας εγκατεστημένη η PHP, μπόρεσε να επιτευχθεί η αποθήκευση της IP διεύθυνσης του μικροϋπολογιστή. Πιο συγκεκριμένα, δημιουργήθηκαν τρία αρχεία, δύο php και ένα txt. Το set_ip.php αρχείο είναι αυτό που αποθηκεύει την IP στο txt αρχείο. Κάθε φορά που μία καινούρια τιμή λαμβάνεται για τη μεταβλητή, το προηγούμενο περιεχόμενο του αρχείου κειμένου διαγράφεται και τη θέση του παίρνει η καινούρια τιμή. Το δεύτερο php αρχείο αναλαμβάνει να επιστρέψει το περιεχόμενο του αρχείου κειμένου. Επομένως, για να αποκτήσει ο υπολογιστής τη διεύθυνση IP του μικροϋπολογιστή-εξυπηρετητή εφαρμόζει cURL στη διεύθυνση URL του αρχείου αυτού.



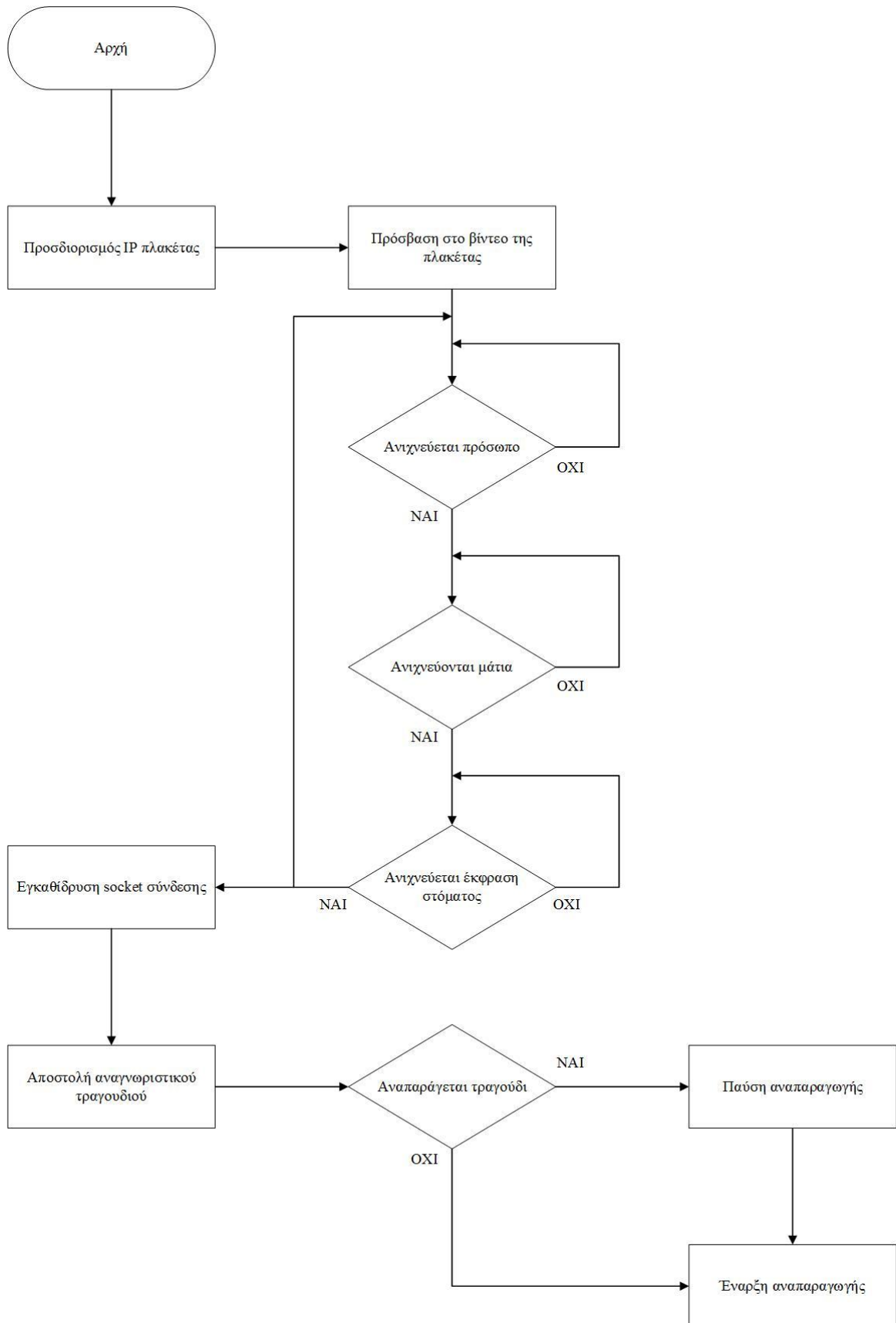
Εικόνα 27: Σχηματική περιγραφή της ανταλλαγής διεύθυνσης IP μεταξύ μικροϋπολογιστή και υπολογιστή.

Αφού αναλύθηκε πλήρως η επικοινωνία του υπολογιστή με τον μικροϋπολογιστή και το πώς αυτή επιτυγχάνεται, θα εστιάσουμε στην μεριά του μικροϋπολογιστή. Εκεί δύο ενέργειες εκτελούνται με την εκκίνηση του συστήματος και μετά τη διαμόρφωσή του. Η μία είναι η εκτέλεση του εργαλείου motion για τη λήψη ροής εικόνας από την κάμερα και η άλλη είναι η εκτέλεση του προγράμματος που μετατρέπει τον μικροϋπολογιστή σε server, έτοιμο να συνάψει επικοινωνία με τον υπολογιστή-client. Οι λειτουργίες αυτές συνεχίζονται μέχρι το σύστημα να τερματιστεί.

Όπως αναφέρθηκε παραπάνω, ο μικροϋπολογιστής σαν εξυπηρετητής δέχεται από τον υπολογιστή μηνύματα που περιέχουν το αναγνωριστικό του τραγουδιού που πρέπει να αναπαράγει. Κάθε φορά που δέχεται ένα τέτοιο μήνυμα εκκινεί, μέσω της συνάρτησης

`posix_spawn()`, το πρόγραμμα που είναι υπεύθυνο για την αναπαραγωγή του τραγουδιού με τη χρήση της βιβλιοθήκης `mpg123`. Όπως και στη μεριά του υπολογιστή, έτσι και εδώ, η `posix_spawn()` εκτελείται σε ένα καινούριο νήμα κάθε φορά. Το σημαντικό σε αυτή την περίπτωση είναι κάθε φορά να γίνεται γνωστή η ταυτότητα του νήματος που εκκινήθηκε.

Σε περίπτωση που ένα μήνυμα καταφτάσει ενώ ένα τραγούδι ήδη αναπαράγεται, που σημαίνει ότι το συναίσθημα που ανιχνεύτηκε έχει αλλάξει πριν προλάβει να τελειώσει το τραγούδι που αντιστοιχούσε στο προηγούμενο, σε φυσιολογική ροή προγράμματος, θα συνέβαινε σύγκρουση. Η σύγκρουση εδώ μεταφράζεται σαν παράλληλη αναπαραγωγή δύο τραγουδιών. Η ταυτότητα του νήματος, έρχεται να δώσει τη λύση για την αποφυγή αυτής της κατάστασης. Κρατώντας την, με το που έρχεται ένα νέο τραγούδι για αναπαραγωγή, το νήμα «σκοτώνεται», με αποτέλεσμα να τερματιστεί το προηγούμενο τραγούδι. Το νέο ξεκινάει να αναπαράγεται σε ένα νέο νήμα, με νέα ταυτότητα και η ίδια διαδικασία συνεχίζεται. Στην Εικόνα 28 παρουσιάζεται το συνολικό διάγραμμα ροής του προγράμματος.



Εικόνα 28: Διάγραμμα ροής του συνολικού προγράμματος.

4.3 Ασφάλεια του συστήματος

Όπως αναφέρθηκε παραπάνω, για την επίτευξη της λειτουργίας του ολοκληρωμένου συστήματος είναι αναγκαία η επικοινωνία του υπολογιστή με τον μικροϋπολογιστή για την ανταλλαγή μηνυμάτων. Επιπλέον, είναι απαραίτητη η γνωστοποίηση κάποιων στοιχείων στον μικροϋπολογιστή, ώστε αυτός να διαμορφωθεί κατάλληλα και να μπορεί να επιτύχει τους στόχους του. Εφόσον όλες οι πληροφορίες ανταλλάσσονται μέσω διαδικτύου, η προστασία τους ήταν επιτακτικής σημασίας.

Αρχικά, ο χρήστης εισάγει τα αναγνωριστικά του ασύρματου δικτύου του μέσω της φόρμας εισαγωγής στοιχείων και αυτά αποθηκεύονται σε δύο αρχεία. Στη συνέχεια, ο μικροϋπολογιστής πρέπει να αποκτήσει πρόσβαση σε αυτά για να μπορέσει να διαμορφώσει την ασύρματη σύνδεσή του. Η ανάγνωση των στοιχείων γίνεται μέσω cURL. Εφόσον η διεύθυνση στην οποία βρίσκονται τα αρχεία είναι γνωστή από το url της φόρμας, οποιοσδήποτε θα μπορούσε με διάφορους τρόπους να αποκτήσει πρόσβαση στα στοιχεία του ασύρματου δικτύου του χρήστη, γνωρίζοντας τα ονόματα των αρχείων στα οποία αυτά αποθηκεύονται. Για να αποτραπεί η διαρροή δεδομένων, στην εντολή cURL που χτυπάει με την μέθοδο POST το rhp αρχείο, αποστέλλεται παράλληλα και ένα μοναδικό αναγνωριστικό. Από τη μεριά του rhp αρχείου ελέγχεται η εγκυρότητα του αναγνωριστικού και στη συνέχεια επιστρέφεται το περιεχόμενο του αρχείου κειμένου που ζητήθηκε από την εντολή. Εάν υπάρξει ανεπιτυχής ταυτοποίηση, επιστρέφεται ανάλογο μήνυμα.

Στην περίπτωση της IP διεύθυνσης του μικροϋπολογιστή γίνεται αποστολή αναγνωριστικού και κατά την εγγραφή της στο αρχείο κειμένου, αλλά και κατά την ανάγνωσή της από τον υπολογιστή. Ο έλεγχος για την ορθότητά του κατά την εγγραφή έγινε από αντίστοιχο, με αυτό που περιγράφηκε παραπάνω, αρχείο. Ο έλεγχος αυτός ήταν εξίσου σημαντικός, καθώς είναι αναγκαίο να αποτραπεί η εγγραφή δεδομένων από οποιονδήποτε άλλον εκτός του μικροϋπολογιστή, αφού λάθος εγγραφή διεύθυνσης IP θα οδηγήσει σε αδυναμία ζεύξης μεταξύ υπολογιστή και παιχνιδιού.

Τέλος, όλα τα αρχεία κειμένου έχουν τοποθετηθεί σε ένα φάκελο στον οποίο έχει απαγορευτεί η πρόσβαση προς όλους. Με τον τρόπο αυτό, κανείς δεν μπορεί να

εφαρμόσει cURL απευθείας στα αρχεία και να ζητήσει το περιεχόμενό τους, προσπαθώντας, έτσι, να αποφύγει τον έλεγχο του μοναδικού αναγνωριστικού.

4.4 S.W.O.T. ανάλυση του συστήματος

Σε αυτό το σημείο, καθώς το ολοκληρωμένο σύστημα που αναπτύχθηκε είναι καινοτόμο, θα πρέπει να αναλυθούν τα δυνατά σημεία, οι αδυναμίες, οι ευκαιρίες και οι απειλές που το χαρακτηρίζουν. Η ανάλυση αυτή συντελεί στον καθορισμό των απαιτήσεων και των τεχνικών χαρακτηριστικών για την περαιτέρω βελτίωση του συστήματος σε μελλοντικές εκδόσεις. Σύμφωνα με το μοντέλο S.W.O.T. υπάρχουν θετικά και αρνητικά χαρακτηριστικά του έργου που προκύπτουν τόσο από εσωτερικούς όσο και από εξωτερικούς παράγοντες.

Ξεκινώντας από τα θετικά χαρακτηριστικά του συστήματος που αναπτύχθηκε, αξίζει να αναφερθεί ότι αυτό είναι καινοτόμο για την κατηγορία που ανήκει. Βασίστηκε σε μία ιδέα που δεν έχει υλοποιηθεί ξανά στον τομέα της παιδικής ψυχαγωγίας και εκπαίδευσης. Είναι ένα παιχνίδι που εστιάζει στη συναισθηματική κατάσταση του παιδιού και προσπαθεί να τη διατηρήσει αν αυτή είναι καλή ή να τη βελτιώσει. Η ανίχνευση συναισθήματος που επιτυγχάνει είναι ακριβής και από αυτή προέρχονται οι μελωδίες που αναπαράγει.

Ακόμα ένα δυνατό χαρακτηριστικό του είναι πως παρά το πολύπλοκο της κατασκευής του, είναι ιδιαίτερα εύχρηστο και φιλικό προς το χρήστη-γονέα. Δεν απαιτεί ιδιαίτερη προετοιμασία και σαν αυτόνομη συσκευή λειτουργεί σχεδόν όπως τα κλασικά παιχνίδια του εμπορείου. Επιπλέον, προσφέρει τη δυνατότητα παρακολούθησης του παιδιού από το γονέα, λειτουργία που είναι ιδιαίτερα χρήσιμη, επιδιώκεται να υπάρχει και μέχρι τώρα προσφέρεται από ξεχωριστές συσκευές.

Βέβαια, όπως κάθε σύστημα στην πρώτη έκδοσή του, έτσι και αυτό χαρακτηρίζεται από κάποια αδύναμα σημεία. Εφόσον η ανίχνευση συναισθήματος βασίζεται στην κατάσταση του στόματος του προσώπου, υπάρχει ο παράγοντας απόσταση που την επηρεάζει. Για να είναι ανιχνεύσιμο το στόμα και να μην γίνεται αντιληπτό από την

κάμερα σαν μία κουκίδα, η απόσταση που παιχνιδιού από το πρόσωπο πρέπει να είναι μικρή, ώστε να μπορεί να αναγνωριστεί η κατάσταση του στόματος.

Επιπλέον, το κόστος του συγκεκριμένου μικροϋπολογιστή που ενσωματώθηκε στο παιχνίδι είναι αρκετά μεγάλο και σε συνδυασμό με τον υπολογιστή που απαιτείται για την ολοκληρωμένη λειτουργία του συστήματος, ενέχει τον κίνδυνο να χαρακτηριστεί μη προσιτό οικονομικά για μεγάλη μερίδα χρηστών.

4.5 Μετρικές του έργου

Πίνακας 4: Μετρικές κώδικα στον υπολογιστή	
Αριθμός Αρχείων	3
Συνολικός Αριθμός Γραμμών Κώδικα (χωρίς σχόλια)	316
Αριθμός Συναρτήσεων	5
Μέγεθος Κώδικα	153003 bytes

Πίνακας 5: Μετρικές κώδικα στον μικροϋπολογιστή	
Αριθμός Αρχείων	6
Συνολικός Αριθμός Γραμμών Κώδικα (χωρίς σχόλια)	253
Αριθμός Συναρτήσεων	3
Μέγεθος Κώδικα	22371 bytes

Πίνακας 6: Μετρικές κώδικα στο server του πανεπιστημίου	
Αριθμός Αρχείων	5
Συνολικός Αριθμός Γραμμών Κώδικα (χωρίς σχόλια)	51
Αριθμός Συναρτήσεων	0
Μέγεθος Κώδικα	-

Σύνοψη κεφαλαίου

Στο κεφάλαιο που μόλις ολοκληρώθηκε αναλύθηκε το λογισμικό μέρος της παρούσας διπλωματικής εργασίας. Πιο συγκεκριμένα, περιγράφηκε όλη η διαδικασία που ακολουθήθηκε για την εκπαίδευση των ταξινομητών αναγνώρισης προσώπου και χαρακτηριστικών, αυτή για τη διαμόρφωση του μικροϋπολογιστή ώστε να επιτελέσει τις λειτουργίες του, αναφέρθηκε η πλήρης λειτουργία του αλγορίθμου αναγνώρισης και δόθηκε ιδιαίτερη βαρύτητα στην επικοινωνία του υπολογιστή με το παιχνίδι για την ολοκλήρωση του συστήματος. Επιπλέον, έγινε η S.W.O.T. ανάλυση, η ανάλυση ασφάλειας του συστήματος και παρατέθηκαν οι μετρικές του κώδικα.

Κεφάλαιο 5: Κατασκευή ενσωματωμένου συστήματος

Στο ακόλουθο κεφάλαιο αναλύονται όλα τα επιμέρους τμήματα της κατασκευής, αλλά και ο τρόπος σύνδεσής τους, ώστε να αποτελέσουν ένα ολοκληρωμένο σύστημα. Επιπροσθέτως, αναφέρονται όλοι οι σχεδιασμοί και υλοποιήσεις, ώστε το σύστημα αυτό να ενσωματωθεί με επιτυχία στο παιχνίδι και ταυτόχρονα να είναι αυτόνομο, εύχρηστο και φιλικό προς το χρήστη.

Σα βάση της κατασκευής επιλέχτηκε ένα λούτρινο παιχνίδι με τη μορφή αρκούδου σε μεσαίο μέγεθος. Η μορφή αυτή είναι κατάλληλη για το λόγο ότι είναι ελκυστική προς τα παιδιά πολλών ηλικιών και διαχρονική. Επιπλέον, το κόστος του είναι χαμηλό –άλλος ένας λόγος που το καθιστά πρακτικό. Ακόμα ένα σημαντικό χαρακτηριστικό είναι το μαλακό της κατασκευής του, από ύφασμα και αφρολέξ, που το κάνει εύκολα διαχειρίσιμο και διευκολύνει την ενσωμάτωση της κατασκευής στο εσωτερικό του. Τέλος, η κάμερα μπορούσε εύκολα να ενσωματωθεί στο ένα από τα δύο μάτια του αρκούδου, αποφεύγοντας έτσι το να είναι εμφανής και αποκρουστική προς τα παιδιά.

Το ενσωματωμένο σύστημα αποτελείται από:

1. Ένα σκεύος αποθήκευσης φαγητού
2. Έναν μικροϋπολογιστή Beagleboard -xM
3. Μια κάμερα webcam
4. Ένα ηχείο χωρίς τροφοδοσία
5. Έναν αντάπτορα ασύρματου δικτύου
6. Ένα καλώδιο προέκτασης θύρας Ethernet
7. Ένα powerbank
8. Ένα ανεμιστηράκι
9. Έναν διακόπτη (switch)
10. Ένα κομβίο (button)

5.1 Διάταξη των ηλεκτρονικών μέσων της κατασκευής

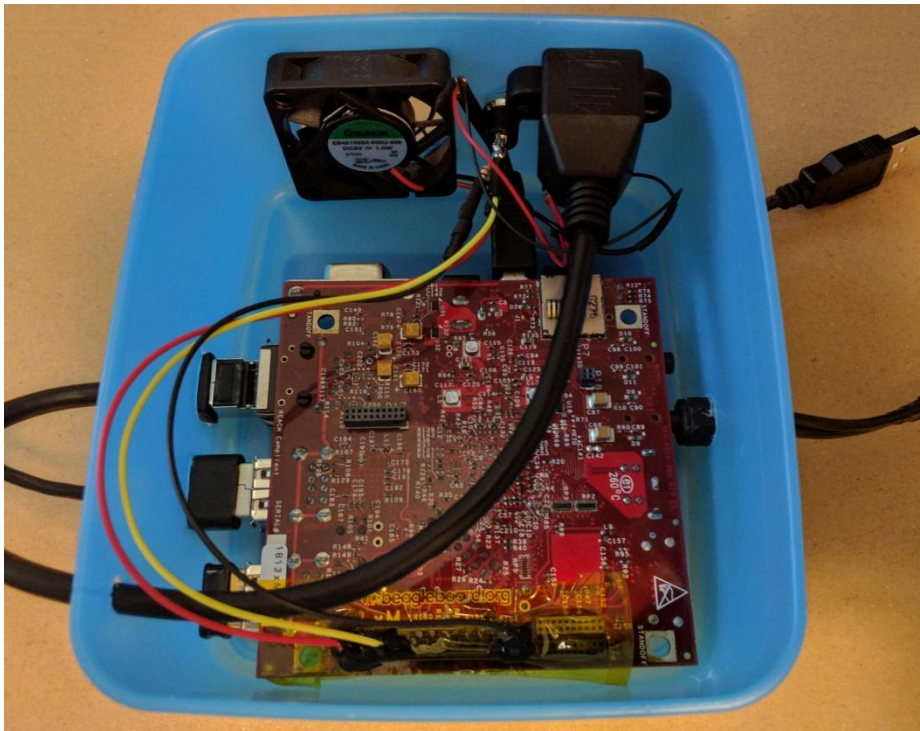
Το σκεύος αποθήκευσης φαγητού θεωρήθηκε μια πρακτική και οικονομική λύση για να αποτελέσει το χώρο αποθήκευσης των ηλεκτρονικών μέσων. Ένας τέτοιος χώρος ήταν απαραίτητος για την προστασία του μικροϋπολογιστή και των καλωδιώσεων, αποφεύγοντας να τοποθετηθούν στο εσωτερικό του παιχνιδιού χωρίς κάποιο περίβλημα. Το υλικό κατασκευής του σκεύους είναι το πλαστικό και οι διαστάσεις του είναι: Μήκος x Πλάτος x Ύψος: 12x14x6 εκατοστά. Για να είναι σταθερή η συσκευή στο εσωτερικό του σκεύους, αλλά και για να υπάρχουν οι κατάλληλοι έξοδοι για τις καλωδιώσεις και τα στοιχεία στα οποία πρέπει να έχει πρόσβαση ο χρήστης, το σκεύος μεταποιήθηκε με τη βοήθεια πολυεργαλείου.

Στη μία από τις μικρές πλευρές του σκεύους αποθήκευσης έχουν εφαρμόσει όλοι οι διακόπτες και οι θύρες στις οποίες χρειάζεται πρόσβαση ο χρήστης για τη λειτουργία του παιχνιδιού. Η συγκέντρωση αυτή εξυπηρετεί ούτως ώστε μόνο η συγκεκριμένη πλευρά να βρίσκεται εξωτερικά του λούτρινου παιχνιδιού. Εφόσον αποφασίστηκε αυτό, στην ίδια πλευρά ενσωματώθηκε και το ανεμιστηράκι για την ψύξη της συσκευής, ώστε να μπορεί να ανακυκλώνεται ο αέρας με επιτυχία. Για χωροταξικούς λόγους, λοιπόν, αλλά και για λόγους απόδοσης επιλέχτηκε ένα ανεμιστηράκι διαστάσεων 4x4 εκατοστών (Εικόνα 29).



Εικόνα 29: Εμπρόσθια όψη σκεύους αποθήκευσης. Φαίνονται αναλυτικά όλα τα στοιχεία που βρίσκονται στο εξωτερικό της κατασκευής

Πιο συγκεκριμένα, οι θύρες που βρίσκονται στην εξωτερική πλευρά της κατασκευής είναι η θύρα Ethernet και η θύρα micro USB για το καλώδιο τροφοδοσίας του powerbank. Επιπροσθέτως, στην εξωτερική πλευρά εφάπτονται ο διακόπτης και το κομβίο, των οποίων η χρήση θα αναλυθεί εκτενώς παρακάτω. Η θηλυκή κατάληξη του καλωδίου προέκτασης Ethernet, που αναφέρθηκε παραπάνω, στερεώθηκε στη συγκεκριμένη πλευρά του σκεύους και η αρσενική προσαρτήθηκε στην αντίστοιχη θύρα του μικροϋπολογιστή. Η προσθήκη αυτή έγινε καθώς, χωροταξικά, δεν ήταν βολικό, για την υπόλοιπη κατασκευή, να εφάπτεται ολόκληρη η πλευρά του μικροϋπολογιστή στην πλευρά του σκεύους που θα είναι εμφανής στο χρήστη (Εικόνα 30). Όσον αφορά στο powerbank, αυτό δεν έχει τοποθετηθεί στο εσωτερικό του σκεύους αποθήκευσης, αλλά προσαρτήθηκε στην κάτω του πλευρά με τη θύρα micro USB να βρίσκεται στην ίδια ευθεία με την εξωτερική πλευρά για να είναι εξίσου εύκολα προσβάσιμη από το χρήστη.



Εικόνα 30: Εσωτερικό του σκεύους αποθήκευσης


5.2 Απενεργοποίηση-Ενεργοποίηση Συσκευής

5.2.1 Κομβίο Τερματισμού

Πολύ σημαντικό μέρος της λειτουργίας του συστήματος αποτελεί και η απενεργοποίησή του. Καθώς ο μικροϋπολογιστής δε διαθέτει κάποιο ειδικό κομβίο για ενεργοποίηση και απενεργοποίηση και δεδομένου ότι ο χρήστης δεν θα έχει καμία πρόσβαση σε αυτόν για να τερματίσει τη λειτουργία του μέσω του λογισμικού, ο μόνος τρόπος για να απενεργοποιηθεί είναι η απομάκρυνσή του από την τροφοδοσία. Για να αποφευχθεί η «βίαιη» απενεργοποίηση του λειτουργικού συστήματος και κατ' επέκταση ολόκληρης της συσκευής, με διακοπή τροφοδοσίας, προσαρτήθηκε και προγραμματίστηκε ένα κομβίο.

Η λειτουργία που προσφέρει το κομβίο είναι αυτή του τερματισμού του λειτουργικού συστήματος του μικροϋπολογιστή, κλείνοντας έτσι το παιχνίδι. Ωστόσο, αυτό δε συνεπάγεται και παύση της τροφοδοσίας της συσκευής, πράγμα που επιτυγχάνει ο διακόπτης, όπως θα αναφερθεί παρακάτω.

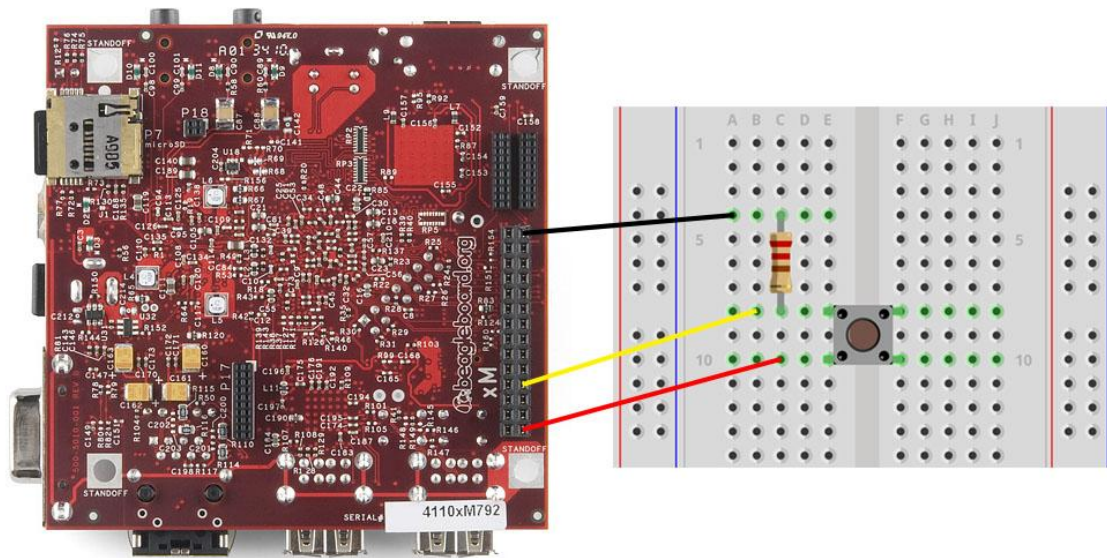
Εισχωρώντας στα πιο τεχνικά, για τη σύνδεση του κομβίου με το μικροϋπολογιστή αξιοποιήθηκαν οι σύνδεσμοι επέκτασης που διαθέτει ο δεύτερος. Για την ακρίβεια, από αυτούς χρησιμοποιήθηκαν ο ακροδέκτης για τροφοδοσία 1,8 Volt, ένας από τους ακροδέκτες για γείωση και ένας από αυτούς που δέχονται σήματα εισόδου/εξόδου. Παρακάτω παρατίθεται ο πίνακας των συνδέσμων επέκτασης επεξεργαστή του μικροϋπολογιστή Beagleboard -xM (Εικόνα 31). [38]



EXP	Processor	0	1	2	3	4	5	6	7
1		VIO 1V8							
2		DC 5V							
3	AE3	MMC2_DAT7	*	*	*	GPIO 139	*	*	Z
4	AB26	UART2_CTS	McBSP3_DX	GPT9_PWMEVT	X	GPIO 144	X	X	Z
5	AF3	MMC2_DAT6	*	*	*	GPIO 138	*	X	Z
6	AA25	UART2_TX	McBSP3_CLKX	GPT11_PWMEVT	X	GPIO 146	X	X	Z
7	AH3	MMC2_DAT5	*	*	*	GPIO 137	*	X	Z
8	AE5	McBSP3_FSX	UART2_RX	X	X	GPIO 143	*	X	Z
9	AE4	MMC2_DAT4	*	X	*	GPIO 136	X	X	Z
10	AB25	UART2_RTS	McBSP3_DR	GPT10_PWMEVT	X	GPIO 145	X	X	Z
11	AF4	MMC2_DAT3	McSPI3_CS0	X	X	GPIO 135	X	X	Z
12	V21	McBSP1_DX	McSPI4_SIMO	McBSP3_DX	X	GPIO 158	X	X	Z
13	AG4	MMC2_DAT2	McSPI3_CS1	X	X	GPIO 134	X	X	Z
14	W21	McBSP1_CLK X	X	McBSP3_CLKX	X	GPIO_162	X	X	Z
15	AH4	MMC2_DAT1	X	X	X	GPIO 133	X	X	Z
16	K26	McBSP1_FSX	McSPI4_CS0	McBSP3_FSX	x	GPIO 161	X	X	Z
17	AH5	MMC2_DAT0	McSPI3_SOMI	X	X	GPIO 132	X	X	Z
18	U21	McBSP1_DR	McSPI4_SOMI	McBSP3_DR	X	GPIO 159	X	X	Z
19	AG5	MMC2_CMD	McSPI3_SIMO	X	X	GPIO 131	X	X	Z
20	Y21	McBSP1_CLK R	McSPI4_CLK	X	X	GPIO_156	X	X	Z
21	AE2	MMC2_CLKO	McSPI3_CLK	X	X	GPIO 130	X	X	Z
22	AA21	McBSP1_FSR	X	*	Z	GPIO 157	X	X	Z
23	AE15	I2C2_SDA	X	X	X	GPIO 183	X	X	Z
24	AF15	I2C2_SCL	X	X	X	GPIO 168	X	X	Z
25	25	REGEN							
26	26	Nreset							
27	27	GND							
28	28	GND							

Εικόνα 31: Αναλυτικά η αρίθμηση και τα σήματα των συνδέσεων επέκταση επεξεργαστή

Το σήμα που χρειάστηκε για τη λειτουργικότητα του κομβίου ήταν το GPIO (General Purpose I/O) που υποδηλώνει ότι ο αντίστοιχος σύνδεσμος μπορεί να χρησιμοποιηθεί ως διακόπτης. Καθαρά για λόγους προτίμησης, ο ακροδέκτης που αξιοποιήθηκε για το σήμα ήταν αυτός με τον αριθμό 7 και σήμα εισόδου/εξόδου 137 (GPIO_137). Η τροφοδοσία δόθηκε μέσω του ακροδέκτη 1 και η γείωση μέσω του ακροδέκτη 27 (Εικόνα 32). Για να υλοποιηθεί το κύκλωμα του κομβίου δε χρησιμοποιήθηκε breadboard, αλλά οι καλωδιώσεις και η αντίσταση συνδέθηκαν απευθείας από τους ακροδέκτες του κομβίου σε αυτούς του μικροϋπολογιστή.



Εικόνα 32: Πραγματική απεικόνιση του κυκλώματος του κομβίου

Για τον προγραμματισμό του κομβίου αναπτύχθηκε ένας bash αλγόριθμος, ο οποίος τρέχει με την εκκίνηση του συστήματος μέσω crontab. Στον αλγόριθμο αυτό, αφού ανοίξει η 137 GPIO θύρα, παρακολουθείται συνεχώς η τιμή του σήματος που λαμβάνει. Τη στιγμή που η τιμή αυτή αλλάζει, ένδειξη ότι το κομβίο έχει πατηθεί, στέλνεται εντολή για τερματισμό του λειτουργικού συστήματος.

5.2.2 Διακόπτης τροφοδοσίας

Όπως ήδη αναφέρθηκε, η τροφοδοσία του συστήματος γίνεται με το powerbank που βρίσκεται στο εσωτερικό του παιχνιδιού και προσαρτάται στη mini USB θύρα του μικροϋπολογιστή. Ο χρήστης δεν έχει πρόσβαση στο καλώδιο αυτό για να μπορέσει να το αφαιρέσει, με αποτέλεσμα η συσκευή να δέχεται ρεύμα ακόμα και όταν δεν βρίσκεται σε λειτουργία. Αυτή η συμπεριφορά οδηγεί στην αποφόρτιση της μπαταρίας χωρίς λόγο. Προς αποφυγήν της συχνής αποφόρτισης της μπαταρίας προσαρτήθηκε στο καλώδιο τροφοδοσίας ένας διακόπτης. Συγκεκριμένα, το καλώδιο κόπηκε στη μέση και μονώθηκαν τα εσωτερικά καλώδια που αφορούν την ανταλλαγή δεδομένων. Χρησιμοποιήθηκαν μόνο τα καλώδια για την τροφοδοσία και τη γείωση, τα οποία συνδέθηκαν κατάλληλα στους ακροδέκτες του διακόπτη. Με τον τρόπο αυτό το παιχνίδι απέκτησε έναν φιλικό προς το χρήστη τρόπο για να διακόπτεται η τροφοδοσία του.

Στους ακροδέκτες του διακόπτη συνδέθηκαν επίσης κατάλληλα και τα καλώδια τροφοδοσίας και γείωσης του ανεμιστήρα. Συνεπώς, κάθε φορά που κλείνει το κύκλωμα μέσω του διακόπτη, εκκινείται η λειτουργία της συσκευής και του ανεμιστήρα. Σε περίπτωση που το λειτουργικό σύστημα τερματιστεί και ο διακόπτης παραμείνει πατημένος στο ON, που σημαίνει ότι η συσκευή συνεχίζει να δέχεται ρεύμα, συνεχίζει να λειτουργεί και το ανεμιστηράκι για διατήρηση καλής θερμοκρασίας στο εσωτερικό του σκεύους.

5.3 Τελική μορφή της κατασκευής

Για να ολοκληρωθεί η κατασκευή και να έχει το παιχνίδι την τελική του μορφή έπρεπε να συνδεθούν η κάμερα για εισαγωγή εικόνας και το ηχείο για εξαγωγή ήχου. Η κάμερα αντικατέστησε το ένα μάτι του αρκούδου-παιχνιδιού, ώστε να μην είναι εύκολα εμφανής και υποδηλώνει κάποιας μορφής διαφορετικότητα για το παιχνίδι. Το ηχείο εισήχθη, επίσης, στο εσωτερικό του παιχνιδιού, σε σημείο τέτοιο ώστε ο ήχος να ακούγεται όσο το δυνατόν καλύτερα (Εικόνα 33).

Συνοψίζοντας, για να ενεργοποιηθεί η λειτουργία της συσκευής ο χρήστης πρέπει να κλείσει το κύκλωμα μέσω του διακόπτη και για να απενεργοποιηθεί εντελώς θα πρέπει να πατηθεί το κουμπί απενεργοποίησης του λειτουργικού συστήματος και στη συνέχεια να ανοίξει το κύκλωμα μέσω του διακόπτη.



(α)

(β)

Εικόνα 33: Τελική εικόνα της κατασκευής: α) μπροστινό και β) πίσω μέρος

Σύνοψη Κεφαλαίου

Στο παρόν κεφάλαιο αναλύθηκε το υλικό μέρος της διπλωματικής εργασίας. Συγκεκριμένα, περιγράφηκε η μορφή του τελικού παιχνιδιού, εσωτερικά και εξωτερικά, οι προσθήκες που έγιναν για τη διευκόλυνση του χρήστη στο χειρισμό του και όλες οι συνδέσεις που υλοποιήθηκαν για να καταλήξει η συσκευή να χαρακτηρίζεται ένα ολοκληρωμένο ενσωματωμένο σύστημα.

Κεφάλαιο 6: Συμπεράσματα και μελλοντικές επεκτάσεις

Η παρούσα διπλωματική εργασία εισήγαγε ένα ευφύες παιχνίδι με δυνατότητες αναγνώρισης συναισθήματος, αντίδρασης στα συναισθήματα με ήχους και διαρκούς παρακολούθησης από το γονέα. Η εκπόνηση του συστήματος χωρίστηκε σε τρία τμήματα. Το πρώτο τμήμα αφορά στην εκπαίδευση των ταξινομητών για τις εκφράσεις του στόματος, το δεύτερο στην ανάπτυξη του αλγορίθμου ανίχνευσης συναισθήματος και αναπαραγωγής ήχου και το τελευταίο στη μετατροπή του σε ενσωματωμένο σύστημα με την τοποθέτηση του μικροϋπολογιστή και των περιφερειακών στο εσωτερικό ενός λούτρινου παιχνιδιού.

6.1 Συμπεράσματα

Το παιχνίδι που υλοποιήθηκε στην παρούσα διπλωματική εργασία έχει σαν στόχο να παρέχει φροντίδα, προστασία και ψυχαγωγία σε παιδιά νεαρής ηλικίας. Μέσω της ανίχνευσης συναισθήματος που πραγματοποιεί, μπορεί να αντιδράσει σωστά και να διαχειριστεί την κατάσταση με τον καλύτερο δυνατό τρόπο, αφήνοντας τελικά το παιδί ευχαριστημένο και σε εγρήγορση. Επιπλέον, προσφέρει στους γονείς τη δυνατότητα να ελέγχουν απομακρυσμένα τα παιδιά τους από οποιαδήποτε συσκευή συνδεδεμένη στο διαδίκτυο κι έτσι να είναι βέβαιοι για την ασφάλεια τους.

Το ολοκληρωμένο έξυπνο παιχνίδι που κατασκευάστηκε υπερτερεί έναντι άλλων παιχνιδιών με παρόμοιες ιδιότητες. Τα παιχνίδια που κυκλοφορούν στο εμπόριο και απευθύνονται σε παιδιά ίδιας ηλικίας με το παρόν, προσφέρουν εκπαίδευση με το να αντιδρούν σε πατήματα κουμπιών και κάποια, περιορισμένα σε πλήθος, αναγνωρίζοντας

μεμονωμένες λεκτικές εκφράσεις. Όσο για την παρακολούθηση, υπάρχουν αυτόνομα συστήματα ενδοεπικοινωνίας, που βασίζονται κυρίως στους ήχους και δεν έχουν τη μορφή παιχνιδιού. Εν αντιθέσει, το παρόν σύστημα έρχεται να εμπλουτίσει τον τομέα των παιχνιδιών με ένα παιχνίδι ανθρωποκεντρικού χαρακτήρα, που βασίζει τις αντιδράσεις του στα συναισθήματα που ανιχνεύει στο εκάστοτε παιδί. Ταυτόχρονα έχει τη μορφή ενός λούτρινου κάνοντάς το προσιτό στα παιδιά, αλλά και διαχρονικό, καθώς μεγαλώνοντας το παιδί κι έχοντας μάθει πια τη λειτουργία του, θα μπορεί κάνοντας «γκριμάτσες» που αντιστοιχούν στα ανιχνεύσιμα συναισθήματα, να ακούει το τραγούδι που επιθυμεί κάθε φορά.

Βασικό πλεονέκτημα του παιχνιδιού που κατασκευάστηκε είναι ότι, παρά την πολύπλοκη λειτουργία του, εξακολουθεί να χαρακτηρίζεται από πλήρη μεταφερσιμότητα. Βασική προϋπόθεση είναι να υπάρχει σύνδεση στο διαδίκτυο και να είναι συνδεδεμένες σε αυτήν και οι δύο συσκευές που το απαρτίζουν. Συνεπώς, έχοντας όλα τα παραπάνω χαρακτηριστικά, συμπεριλαμβανομένης της φορητότητας, ξεχωρίζει από άλλα παιχνίδια της κατηγορίας του.

Επιπλέον, στο σύνολο του το σύστημα χαρακτηρίζεται από επεκτασιμότητα, καθώς καινούριες δυνατότητες μπορούν εύκολα να ενσωματωθούν στην εφαρμογή. Για παράδειγμα, καινούριοι ταξινομητές μπορούν να εκπαιδευτούν ή ακόμα να βελτιωθούν οι ήδη υπάρχοντες και να φορτωθούν στον αλγόριθμο με εύκολο τρόπο, χωρίς να απαιτούνται μεγάλες αλλαγές στο λογισμικό. Το ίδιο ισχύει και με την προσθήκη μουσικών κομματιών, τα οποία θα συνοδεύουν τους νεοεισαχθέντες ταξινομητές και θα αυξηθούν σε ποσότητα ανάλογα με την επέκταση της εφαρμογής. Επιπλέον, υπάρχει δυνατότητα απομάκρυνσης του ενσωματωμένου συστήματος από το παιχνίδι που επιλέχτηκε και τοποθέτησης του, με μικρές παραλλαγές, για να εφαρμόζει, για παράδειγμα, η κάμερα σε κάποιο σημείο, σε κάποιο άλλο περισσότερο επιθυμητό από τους χρήστες (π.χ. άλλου είδους λούτρινου).

Πέρα από τα πλεονεκτήματα που προσφέρει στους χρήστες το παρόν σύστημα, αξίζει να σημειωθεί και το ερευνητικό μέρος πίσω από τον αλγόριθμο ανίχνευσης συναισθήματος. Μέσω της ενασχόλησης με την όραση μηχανής και τη βιβλιοθήκη OpenCV διαπιστώθηκε ότι οι δυνατότητες γύρω από αυτόν τον τομέα είναι αναρίθμητες. Έχοντας οποιαδήποτε δεδομένα εικόνας και με την κατάλληλη επεξεργασία τους

μπορούν να υλοποιηθούν εφαρμογές που αφορούν μεγάλο εύρος τομέων. Σε συνδυασμό με τα ενσωματωμένα συστήματα, οι εφαρμογές αυτές μπορούν να αποτελέσουν ολοκληρωμένα συστήματα και να διευκολύνουν την μετατροπή απλών συσκευών σε ευφυείς, με ανεξάντλητες δυνατότητες. Εκμεταλλευόμενοι, λοιπόν, τα πλεονεκτήματα αυτά καταφέραμε να κατασκευάσουμε ένα ολοκληρωμένο σύστημα με προσανατολισμό στον πολύπλοκο τομέα της ανατροφής των παιδιών και να παράγουμε επαρκή και αξιόπιστα αποτελέσματα.

6.2 Προβλήματα που αντιμετωπίστηκαν κατά την υλοποίηση

Σε όλα τα στάδια της διπλωματικής εργασίας που εκπονήθηκε προέκυψαν ζητήματα των οποίων η επίλυση αποτέλεσε πρόκληση, έτσι ώστε τελικά να παραχθεί το καλύτερο αποτέλεσμα που θα καλύπτει όλες τις ανάγκες του συστήματος.

Ξεκινώντας την υλοποίηση της εργασίας έπρεπε να γίνει η εκπαίδευση των ταξινομητών για τις εκφράσεις του στόματος. Για να γίνει αυτό απαιτήθηκε ένας μεγάλος όγκος εικόνων που να απεικονίζουν τα επιθυμητά συναισθήματα. Ωστόσο, δεν υπάρχει κάποια διαθέσιμη βάση δεδομένων με παρόμοιο περιεχόμενο που να μπορεί να χρησιμοποιηθεί για να εξαχθούν τα δείγματα που χρειάζονται για τη διαδικασία της εκπαίδευσης. Συνεπώς, έπρεπε να βρεθεί τρόπος να συλλεχθούν οι κατάλληλες φωτογραφίες γι' αυτό το σκοπό. Για το λόγο αυτό ζητήθηκε από φιλικά και συγγενικά πρόσωπα να συνεργαστούν και να φωτογραφηθούν με τις κατάλληλες εκφράσεις. Για να επιτευχθεί ομοιομορφία μεταξύ των δειγμάτων και δεδομένου ότι ο όγκος που απαιτείται είναι μεγάλος, ώστε να δώσει ο αλγόριθμος υψηλά ποσοστά αναγνώρισης, αναπτύχθηκε ο αλγόριθμος που ανοίγει την κάμερα του υπολογιστή και τραβάει φωτογραφίες μέχρις ότου το χρήστης πατήσει το κουμπί ESC. Ο εν λόγω αλγόριθμος αναπτύχθηκε με τη χρήση της βιβλιοθήκης OpenCV.

Σε επόμενο στάδιο και αφού οι ταξινομητές εκπαιδεύτηκαν και ενσωματώθηκαν στον αλγόριθμο, αυτός έπρεπε να τρέχει και να παράγει αποτελέσματα μέσω του μικροϋπολογιστή BeagleBoard –xM. Εξ αρχής σκοπός ήταν να αναπτυχθεί ένα αυτόνομο

ενσωματωμένο σύστημα με τη μορφή και τις λειτουργίες ενός έξυπνου παιχνιδιού. Ο μικροϋπολογιστής που χρησιμοποιήθηκε, όμως, για το σκοπό αυτό δεν επαρκούσε, από άποψη υπολογιστικής ισχύος, για να υποστηρίξει τόσο την τροφοδότηση του αλγόριθμου με εικόνα πραγματικού χρόνου, όσο και την επεξεργασία της για την παραγωγή αποτελεσμάτων συναισθήματος. Πιο συγκεκριμένα, οι λειτουργίες της ανίχνευσης και της αναγνώρισης μπορούν να γίνουν κανονικά, αλλά η καθυστέρηση που εισάγεται είναι αυξημένη με αποτέλεσμα η απόκριση της εφαρμογής να μην είναι πραγματικού χρόνου. Για το λόγο αυτό ήταν επιτακτική η ανάγκη για την χρήση και την επικοινωνία με τον υπολογιστή, που διαθέτει αυξημένη επεξεργαστική ισχύ, ώστε να αναλάβει την επεξεργασία των δεδομένων βίντεο και να εξάγει τα συναισθήματα που ανιχνεύονται. Κατά τα λοιπά, ο μικροϋπολογιστής συνεχίζει να παίζει σημαντικό ρόλο στο σύστημα, το οποίο εξακολουθεί να χαρακτηρίζεται ως ενσωματωμένο.

Αφού αποφασίστηκε ότι το σύστημα στην ολοκληρωμένη που μορφή θα αποτελούνταν από δύο συσκευές, τον υπολογιστή και τον μικροϋπολογιστή, το πρόβλημα που προέκυψε ήταν αυτό της επικοινωνίας μεταξύ τους. Ο υπολογιστής έπρεπε να λαμβάνει τα δεδομένα εικόνας από τον μικροϋπολογιστή και ο μικροϋπολογιστής να λαμβάνει το μήνυμα με την επιλογή του τραγουδιού, βασισμένο στο συναίσθημα που ανιχνεύτηκε. Στο πρώτο σκέλος της επικοινωνίας βοήθησε το εργαλείο Motion, που προσφέρει πρόσβαση στην εικόνα της κάμερας μέσω IP διεύθυνσης και το OpenCV που μπορεί να ανοίξει προς επεξεργασία βίντεο από οποιαδήποτε διεύθυνση. Για το δεύτερο σκέλος της επικοινωνίας αναπτύχτηκε η σύνδεση socket για ανταλλαγή μηνυμάτων από τον υπολογιστή προς τον μικροϋπολογιστή. Όλα τα παραπάνω ήταν εφικτά αφού οι IP διευθύνσεις των δύο συσκευών ήταν γνωστές.

Για να γίνει, όμως εφικτή η επικοινωνία τους σε οποιοδήποτε δίκτυο, δεν έπρεπε να ληφθούν ως δεδομένες οι IP διευθύνσεις, αλλά έπρεπε να λαμβάνονται δυναμικά και αυτό είναι και το τελευταίο πρόβλημα που αντιμετωπίστηκε. Η IP διεύθυνση του μικροϋπολογιστή στέλνεται σε έναν ενδιάμεσο server στον οποίο έχει πρόσβαση ο υπολογιστής για να διαβάσει και να γνωρίζει πλέον από πού θα λάβει το βίντεο προς επεξεργασία και πού θα στείλει τα αποτελέσματά του.

6.3 Μελλοντικές επεκτάσεις

Κατά τη διάρκεια εκπόνησης της παρούσας διπλωματικής εργασίας, και λόγω των προκλήσεων που αναφέρθηκαν παραπάνω, προέκυψαν ιδέες για μελλοντικές επεκτάσεις που αφορούν πολλά τμήματα του συστήματος.

Η πιο σημαντική επέκταση που μπορεί να γίνει στο σύστημα είναι να μετατραπεί αυτό σε απόλυτα αυτόνομο. Πιο συγκεκριμένα, η εφαρμογή θα μπορεί να υποστηρίζεται εξ ολοκλήρου από τον μικροϋπολογιστή, έτσι ώστε να υπάρχει πλέον ένα ενσωματωμένο σύστημα, πάντα με τη μορφή παιχνιδιού, πλήρως μεταφέρσιμο και λειτουργικό. Αυτό μπορεί να γίνει με την χρήση ενός άλλου μικροϋπολογιστή, με μεγαλύτερη επεξεργαστική ισχύ ή με την επέκταση της ήδη υπάρχουσας. Λύση αυτού του προβλήματος θα μπορούσε να αποτελέσει και ο διαμοιρασμός των διεργασιών ανάμεσα σε δύο πλακέτες, δηλαδή τη θέση του υπολογιστή να την πάρει ένας μικροϋπολογιστής. Η λύση αυτή θα διατηρούσε το προνόμιο της αυτονομίας αλλά το ολοκληρωμένο σύστημα θα ήταν πλέον πολύ βαρύ για να αποτελεί ένα φορητό παιχνίδι.

Μια ακόμα, εξίσου σημαντική επέκταση είναι ο εμπλουτισμός των ταξινομητών. Με τον τρόπο αυτό η εφαρμογή θα μπορεί να αναγνωρίζει πολλαπλάσια συναισθήματα, όπως πόνος, τρόμος, θυμός κ.α. και έτσι θα μπορεί να ανταποκρίνεται καλύτερα στις ανάγκες γονέων και παιδιών. Επιπλέον, οι ήδη υπάρχοντες ταξινομητές, θα ήταν ενδιαφέρον να εκπαιδευτούν εκ νέου με δείγματα φωτογραφιών αποκλειστικά και μόνο από μικρής ηλικίας άτομα και συγκεκριμένα από αυτής για τα οποία θα απευθύνεται. Κατά συνέπεια, ο αλγόριθμος θα έχει σαν άξονά του, αποκλειστικά, τα παιδιά και η αποτελεσματικότητά του θα είναι αυξημένη.

Τέλος, για να γίνει πιο διαδραστικό το παιχνίδι θα μπορούσε να εμπλουτιστεί με περισσότερους αισθητήρες. Για παράδειγμα, ένα μικρόφωνο θα ήταν ιδανικό για να καταγράφει τη φωνή του παιδιού, στη συνέχεια αυτή να υπόκειται κάποιας επεξεργασίας και τελικά να ανιχνεύεται πιο ολοκληρωμένα κάποιο συναίσθημα. Ακόμα, με την

προσθήκη αυτή, το παιδί θα δύναται να ζητάει ακριβώς ποιο τραγούδι θέλει να ακούσει σε μία δεδομένη στιγμή.

Παράρτημα

Οδηγίες Εγκατάστασης στον Υπολογιστή

Για την εκτέλεση του προγράμματος στον υπολογιστή είναι απαραίτητη η εγκατάσταση τριών βιβλιοθηκών και των εξαρτήσεών τους. Οι βιβλιοθήκες αυτές είναι:

- 1) OpenCV
- 2) cURL
- 3) pthread

Συνίσταται η λήψη της τελευταίας πιστοποιημένης έκδοσης OpenCV για εγκεκριμένα και αξιόπιστα αποτελέσματα. Η έκδοση που χρησιμοποιήθηκε για την ανάπτυξη του συγκεκριμένου λογισμικού είναι η 2.4.11.

Για να τρέξει η εφαρμογή στον υπολογιστή απαιτείται η μεταγλώττιση δύο αρχείων. Για το σκοπό αυτό δημιουργήθηκε ένα αρχείο Makefile που αφορά και τα δύο προγράμματα, μέσα στο φάκελο της εργασίας. Αρκεί, λοιπόν, να εκτελεστεί η εντολή «make all» στην τοποθεσία του φακέλου της εργασίας.

Οδηγίες Χρήσης Παιχνιδιού

Παρακάτω αναφέρονται αναλυτικά τα βήματα που πρέπει να ακολουθηθούν για να ξεκινήσει η λειτουργία του παιχνιδιού:

Εάν είναι η πρώτη φορά λειτουργίας της συσκευής για κάποιο ασύρματο δίκτυο:

- 1) Ο χρήστης πρέπει να πλοηγηθεί στο σύνδεσμο <http://zafora.ict.e.uowm.gr/~ictest00303/> και να εισάγει τα στοιχεία της σύνδεσής του. Μετά το πέρας της φοίτησης ο νέος ιστότοπος για τη φόρμα θα είναι http://108.167.189.65/diplomatiki_michaila/
- 2) Το καλώδιο Ethernet πρέπει να συνδεθεί στο πίσω μέρος του αρκούδου.
- 3) Ο διακόπτης τροφοδοσίας πρέπει να βρεθεί στη θέση ON.
- 4) Ο χρήστης πρέπει να αφαιρέσει το καλώδιο Ethernet με το άκουσμα του χαρακτηριστικού ήχου προειδοποίησης.
- 5) Ο χρήστης πρέπει να εκκινήσει το πρόγραμμα στον υπολογιστή με το άκουσμα του χαρακτηριστικού ήχου προειδοποίησης ότι η συσκευή είναι έτοιμη.

Εάν η συσκευή έχει συνδεθεί έστω και μία φορά στο ασύρματο δίκτυο του χρήστη, αυτός θα πρέπει να παραλείψει τα βήματα 1), 2) και 4).

Για την απενεργοποίηση του συστήματος πρέπει να τερματιστεί το πρόγραμμα που εκτελείται στον υπολογιστή, να πατηθεί το κουμπί που βρίσκεται στο πίσω μέρος του αρκούδου και στη συνέχεια να γυρίσει ο διακόπτης στη θέση OFF.

Βιβλιογραφία

- [1]. Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." *Computer Vision and Pattern Recognition*, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. Vol. 1. IEEE, 2001.
- [2]. Chen, Qing, Nicolas D. Georganas, and Emil M. Petriu. "Real-time vision-based hand gesture recognition using haar-like features." *Instrumentation and Measurement Technology Conference Proceedings, 2007. IMTC 2007. IEEE. IEEE*, 2007.
- [3]. Pham, Duc-Minh, et al. "Smart teddy bear a vision-based story Teller." *Control, Automation and Information Sciences (ICCAIS), 2013 International Conference on. IEEE*, 2013.
- [4]. Sriram, Srikanth, and Babu Illuri. "Real Time Smile Detection using Haar Classifiers on SoC." *International Journal of Computer Applications* 104.10, 2014.
- [5]. BeagleBoard, Ιστότοπος, <https://beagleboard.org/about>, Νοέμβριος 2017
- [6]. Ångström, Ιστότοπος, <http://www.angstrom-software.com/>, Νοέμβριος 2017
- [7]. Σύγκριση C++/Python για το OpenCV, Ιστότοπος, <https://www.learnopencv.com/opencv-c-vs-python-vs-matlab-for-computer-vision/>, Νοέμβριος 2017
- [8]. Matlab, Ιστότοπος, <https://en.wikipedia.org/wiki/MATLAB>, Νοέμβριος 2017
- [9]. Εργαλεία Μηχανικής Όρασης του εργαλείου Matlab, Ιστότοπος, https://www.mathworks.com/help/vision/index.html?s_tid=srchtitle, Νοέμβριος 2017
- [10]. Motion, Ιστότοπος, <https://www.linux.com/learn/how-operate-linux-spycams-motion>, Νοέμβριος 2017
- [11]. Οδηγός Motion, Ιστότοπος, <http://www.lavrsen.dk/foswiki/bin/view/Motion/MotionGuideBasicFeatures>, Νοέμβριος 2017
- [12]. Νήμα, Ιστότοπος, http://arch.ict.e.uowm.gr/courses/parallel/parsys_08_oc.pdf, Νοέμβριος 2017
- [13]. Πολυνηματισμός, Ιστότοπος, https://aetos.it.teithe.gr/~adamidis/OOP/Threads_notes.pdf, Νοέμβριος 2017
- [14]. Mpg123, Ιστότοπος, <https://mpg123.org/features.shtml>, Νοέμβριος 2017

- [15]. Libao, Ιστότοπος, <https://xiph.org/ao/doc/>, Νοέμβριος 2017
- [16]. cURL, Ιστότοπος, <https://curl.haxx.se/>, Νοέμβριος 2017
- [17]. Bradski, Gary, and Adrian Kaehler. "Learning OpenCV: Computer vision with the OpenCV library". O'Reilly Media, Inc., 2008.
- [18]. Forsyth, D. A., & Ponce, J. "Computer vision: a modern approach". Prentice Hall Professional Technical Reference, 2002.
- [19]. Όραση μηχανής-Επεξεργασία εικόνας, Ιστότοπος, <http://www2.cs.ucy.ac.cy/~nicolast/courses/teds150/lectures/IP01.pdf>, Νοέμβριος 2017
- [20]. Umbaugh, S. E. "Computer Vision and Image Processing: A Practical Approach Using Cviptools with Cdrom". Prentice Hall PTR, 1997.
- [21]. Εκπαίδευση ταξινομητή αντίχρευσης, Ιστότοπος, <https://www.mathworks.com/help/vision/ug/train-a-cascade-object-detector.html>, Νοέμβριος 2017
- [22]. Sockets, Ιστότοπος, <https://docs.oracle.com/javase/tutorial/networking/sockets/definition.html>, Νοέμβριος 2017
- [23]. Server-Client, Ιστότοπος, http://www.bogotobogo.com/cplusplus/sockets_server_client.php, Νοέμβριος 2017
- [24]. Cron, Ιστότοπος, <https://help.ubuntu.com/community/CronHowto>, Νοέμβριος 2017
- [25]. R. Fielding, J. C. Mogul, H. Frystyk, P. Leach, T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1.", June 1999
- [26]. Rezaei, M., Nafchi, H. Z., & Morales, S. "Global Haar-like features: A new extension of classic Haar features for efficient face detection in noisy images". In Pacific-Rim Symposium on Image and Video Technology (pp. 302-313). Springer Berlin Heidelberg, October 2013.
- [27]. Hassaballah, M., Abdelmgeid, A. A., & Alshazly, H. A. "Image Features Detection, Description and Matching". In Image Feature Detectors and Descriptors (pp. 11-45). Springer International Publishing, 2016.
- [28]. Winder, S. A., & Brown, M. "Learning local image descriptors. In Computer Vision and Pattern Recognition", 2007. CVPR'07. IEEE Conference on (pp. 1-8). IEEE, June 2007.
- [29]. Gupta, D., & Choubey, S. "Discrete wavelet transform for image processing. International Journal of Emerging Technology and Advanced Engineering", 4(3), (pp.598-602), 2015.

- [30]. Dalal, N., & Triggs, B. "Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recognition", 2005. CVPR 2005. IEEE Computer Society Conference on (Vol. 1, pp. 886-893). IEEE, June 2005.
- [31]. Tsai, Grace. "Histogram of oriented gradients." University of Michigan, 2010.
- [32]. Abrah & atildeo, W., et al. "A comparison of Haar-like, LBP and HOG approaches to concrete and asphalt runway detection in high resolution imagery."
- [33]. Ole Helvig Jensen. "Implementing the Viola-Jones Face Detection Algorithm". Kongens Lyngby, 2008.
- [34]. Papageorgiou, Oren and Poggio, "A general framework for object detection", International Conference on Computer Vision, 1998.
- [35]. Negri, Pablo, Xavier Clady, and Lionel Prevost. "Benchmarking haar and histograms of oriented gradients features applied to vehicle detection." ICINCO-RA (1). 2007.
- [36]. Brehar, Raluca, and Sergiu Nedevschi. "A comparative study of pedestrian detection methods using classical Haar and HoG features versus bag of words model computed from Haar and HoG features." Intelligent Computer Communication and Processing (ICCP), 2011 IEEE International Conference on. IEEE, 2011.
- [37]. Samarth Brahmhatt. "Practical OpenCV". Apress, 2013.
- [38]. beagleboard.org, BeagleBoard-xM Rev C System Reference Manual, Revision 1.0 April 4, 2010.