



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Kinect : Αποθήκευση Κινήσεων

Χρήστος Ν. Διαμαντόγιαννης

Επιβλέποντες: Ιωάννης Ιωαννίδης, Καθηγητής

Κατερίνα Ελ Ράχεμπ, Υποψήφια Διδακτορικός

ΑΘΗΝΑ

ΟΚΤΩΒΡΙΟΣ 2015

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Kinect : Αποθήκευση Κινήσεων

Χρήστος Ν. Διαμαντόγιαννης

A.M.: 1115201000021

Επιβλέποντες: Ιωάννης Ιωαννίδης, Καθηγητής

Κατερίνα Ελ Ράχεμπ, Υποψήφια Διδακτορικός

ΠΕΡΙΛΗΨΗ

Αυτή η πτυχιακή παρουσιάζει την υλοποίηση μιας εφαρμογής, η οποία καταγράφει και αποθηκεύει κινήσεις του χρήστη. Για την καταγραφή των κινήσεων χρησιμοποιείται μία κάμερα Kinect. Κάθε δευτερόλεπτο αποθηκεύονται κάποια frames της κίνησης που εκτελεί ο χρήστης σε μία βάση. Ουσιαστικά αποθηκεύονται τα joints του κάθε frame, τα οποία έχουν 3-Δ συντεταγμένες στο χώρο (X, Y, Z). Συνολικά από το Kinect στέλνει 20 joints.

Ο χρήστης θα μπορεί να πραγματοποιεί μία απλή κίνηση (βήμα, τρέξιμο, ...) ή κάποια κίνηση από ένα συγκεκριμένο είδος χορού ή ακόμα και να εισάγει κάποια δική του κίνηση ο ίδιος (ακόμα και νέο χορό).

Παρουσιάζονται τα χαρακτηριστικά του Kinect, καθώς επίσης και σχετική εργασία επί του ίδιου θέματος. Αναφέρονται οι τεχνολογίες και τα εργαλεία που χρησιμοποιήθηκαν για την ανάπτυξη της εφαρμογής και της βάσης δεδομένων που αυτή χρησιμοποιεί. Αναλύονται τα μειονεκτήματα και τα πλεονεκτήματα, όπως και οι προβληματισμοί που προέκυψαν κατά την δημιουργία της. Τέλος γίνεται αναφορά στα συμπεράσματα, μελλοντική εργασία – ανοιχτά θέματα.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Movement Database

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Αναγνώριση κίνησης, Kinect, εκπαίδευση υπολογιστή, βάση δεδομένων, χορευτικές κινήσεις.

ABSTRACT

This thesis presents the implementation of an application, which records and stores human movements. A Kinect sensor is used in order to record these movements. Several frames of the human movement are stored in a database. Basically the joints of every frame, which have 3d coordinates, are stored. Kinect sensor sends 20 joints.

Every user can make a simple movement (step, run) or a movement which belongs to a specific type of dance even adds his own (even new dance).

Here are presented the features of Kinect, as well as related work on the same issue. Technologies and tools that used to develop the application are indicated and also the database which is being used. The advantages and disadvantages are analyzed, as well as the concerns raised during the creation. Finally reference is made to the conclusions, future work - open issues.

SUBJECT AREA: Movement Database

KEYWORDS: Movement recognition, Kinect, computer training, database, dance movements

Στην οικογένειά μου

ΕΥΧΑΡΙΣΤΙΕΣ

Για τη διεκπεραίωση της παρούσας Πτυχιακής Εργασίας θα ήθελα να ευχαριστήσω τους επιβλέποντες κ. Ιωάννη Ιωαννίδη και Κατερίνα Ελ Ράχεμπ για την καθοδήγησή τους.

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΡΟΛΟΓΟΣ.....	10
1.ΕΙΣΑΓΩΓΗ.....	11
1.1 Θέμα εργασίας.....	11
1.2 Σκοπός εργασίας.....	11
1.3 Διάρθρωση εργασίας.....	11
2.ΚΙΝΕCΤ.....	13
2.1 Κυκλοφορία και τεχνικά χαρακτηριστικά.....	13
2.2 Χρήση Kinect.....	15
3.ΣΧΕΤΙΚΗ ΕΡΓΑΣΙΑ.....	17
4.ΥΛΟΠΟΙΗΣΗ ΚΑΙ ΤΕΧΝΟΛΟΓΙΕΣ.....	19
4.1 Περιγραφή υλοποίησης.....	19
4.2 Τεχνολογίες που χρησιμοποιήθηκαν.....	27
4.3 Σχήμα βάσης.....	28
5.ΔΥΣΚΟΛΙΕΣ ΚΑΤΑ ΤΗΝ ΥΛΟΠΟΙΗΣΗ.....	34
5.1Προβληματισμοί.....	34
5.2Μειονεκτήματα.....	35
5.3 Πλεονεκτήματα.....	35
6.ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΗ ΕΡΓΑΣΙΑ.....	36
6.1 Ανασκόπηση.....	36
6.2 Συμπεράσματα.....	36
6.3 Μελλοντική εργασία - ανοιχτά θέματα.....	36
ΠΙΝΚΑΣ ΟΡΟΛΟΓΙΑΣ.....	38
ΣΥΝΤΜΗΣΕΙΣ – ΑΡΚΤΙΚΟΛΕΞΑ – ΑΚΡΩΝΥΜΙΑ.....	39
ΠΑΡΑΡΤΗΜΑ.....	40
ΑΝΑΦΟΡΕΣ.....	41

ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ

Εικόνα 1: Kinect.....	13
Εικόνα 2: Μέρη του Kinect.....	14
Εικόνα 3: Τα joints που αναγνωρίζει το Kinect.....	15
Εικόνα 4: Παράθυρο Log in.....	19
Εικόνα 5: Παράθυρο Register.....	20
Εικόνα 6: Παράθυρο για Βοήθεια.....	21
Εικόνα 7: Παράθυρο για πληροφορίες του δημιουργού.....	21
Εικόνα 8: Παράθυρο όπου ο χρήστης επιλέγει αν θέλει να κάνει Generic ή Dance Movement.....	22
Εικόνα 9: Παράθυρο όπου ο χρήστης διαλέγει ποια Generic Movement θα εκτελέσει.....	23
Εικόνα 10: Παράθυρο όπου ο χρήστης εισάγει νέο χορό και την αντίστοιχη νέα κίνηση που θέλει να εκτελέσει.....	23
Εικόνα 11: Παράθυρο όπου ο χρήστης διαλέγει αν θα εκτελέσει μία προτεινόμενη κίνηση ή αν θα κάνει μία δική του.....	24
Εικόνα 12: Παράθυρο όπου ο χρήστης ονοματίζει τη νέα Generic κίνηση που θέλει να εκτελέσει.....	24
Εικόνα 13: Σχήμα βάσης.....	28
Εικόνα 14: Οντότητα user.....	28
Εικόνα 15: Οντότητα gesture.....	29
Εικόνα 16: Το Table που ταιριάζει τους users με τις αντίστοιχες κινήσεις που εκτέλεσε ο καθένας.....	29
Εικόνα 17: Οντότητα frame.....	29
Εικόνα 18: Το table που αντιστοιχίζει τις κινήσεις με τα αντίστοιχα frames.....	30
Εικόνα 19: Οντότητα joint.....	30
Εικόνα 20: Το table που αντιστοιχίζει τα frames με τα joints τους.....	30
Εικόνα 21: Η οντότητα generic movements.....	31
Εικόνα 22: Η οντότητα dance movements.....	31
Εικόνα 23: Ψαλιδάκι.....	32
Εικόνα 24: Arabesque.....	33

Εικόνα 25: Attitude.....	33
--------------------------	----

ΠΡΟΛΟΓΟΣ

Η παρούσα εργασία εκπονήθηκε στο Τμήμα Πληροφορικής και Τηλεπικοινωνιών του Εθνικού Καποδιστριακού Πανεπιστημίου Αθηνών στα πλαίσια της Πτυχιακής Εργασίας. Η διάρκεια της διεξαγωγής της μέχρι και την ολοκλήρωσή της ήταν ένας χρόνος.

1. ΕΙΣΑΓΩΓΗ

1.1 Θέμα εργασίας

Το θέμα της παρούσας εργασίας είναι η υλοποίηση μιας εφαρμογής, που έχει σαν στόχο την αποθήκευση απλών (βήμα, τρέξιμο) ή χορευτικών κινήσεων συγκεκριμένων χορών με τη χρήση μιας κάμερας Kinect [1]. Ο χρήστης μπορεί να διαλέγει κάποια από τις κινήσεις (απλές ή χορευτικές) που προτείνει η εφαρμογή και να την πραγματοποιεί. Ακολουθώντας η κίνηση αυτή θα αποθηκεύεται στη βάση δεδομένων του υπολογιστή. Ο χρήστης δεν είναι περιορισμένος να κάνει μόνο τις κινήσεις που του προτείνει η βάση. Αν το επιθυμεί μπορεί να εισάγει μία δική του κίνηση ή ακόμα να εισάγει και δικό του είδος χορού (με τις αντίστοιχες του κινήσεις).

1.2 Σκοπός εργασίας

Ο στόχος της εργασίας είναι η ανάπτυξη μιας εφαρμογής που θα αποθηκεύει απλές ή και χορευτικές κινήσεις που πραγματοποιεί ο χρήστης σε μία βάση δεδομένων η οποία θα χρησιμοποιηθεί για πειραματισμό ο οποίος θα εξηγηθεί παρακάτω.

Κίνηση είναι ένα σύνολο θέσεων του ανθρώπινου σκελετού. Με τον όρο απλή κίνηση εννοούμε την κίνηση που εκτελεί ο μέσος άνθρωπος στην καθημερινή του ζωή (π.χ. περπάτημα, τρέξιμο, άλμα, χαιρετισμός). Με τον όρο χορευτική κίνηση εννοούμε αποκλειστικά κινήσεις που αφορούν χορούς.

Έχοντας κατασκευάσει μια βάση δεδομένων που περιέχει κινήσεις, είναι εφικτό να συμπεράνουμε το κατά πόσο ρεαλιστικά μπορεί ο υπολογιστής να αναπαραστήσει μια φυσική κίνηση με τον εξής τρόπο: Θα φτιάξουμε ένα πρόγραμμα το οποίο θα καταγράφει κινήσεις του χρήστη (με τη βοήθεια της κάμερας Kinect) και μετά θα ελέγχει αν υπάρχει αντιστοιχία με των αποθηκευμένων κινήσεων στη βάση.

1.3 Διάρθρωση εργασίας

Η διάρθρωση των κεφαλαίων έχει ως εξής:

Κεφάλαιο 2: Σε αυτό το κεφάλαιο αναφέρονται η κυκλοφορία όπως και τα τεχνικά χαρακτηριστικά της κάμερας Kinect. Αναφέρονται επίσης οι αρθρώσεις του ανθρώπινου σκελετού που αναγνωρίζει και η χρησιμότητα του Kinect.

Κεφάλαιο 3: Σε αυτό το κεφάλαιο περιγράφονται εργασίες παραπλήσιες στην εφαρμογή που αναπτύχθηκε στα πλαίσια της πτυχιακής εργασίας αυτής.

Κεφάλαιο 4: Σε αυτό το κεφάλαιο γίνεται αναφορά στη λειτουργία του κάθε αρχείου της εφαρμογής και στη γενικότερη λειτουργία της εφαρμογής.

Περιγράφονται οι τεχνολογίες και τα εργαλεία που χρησιμοποιήθηκαν για την ανάπτυξη της εφαρμογής καθώς επίσης και το σχήμα της βάσης.

Κεφάλαιο 5: Σε αυτό το κεφάλαιο γίνεται περιγραφή των δυσκολιών που προέκυψαν κατά την υλοποίηση της εφαρμογής, των προβληματισμών που δημιουργήθηκαν. Επίσης αναφέρονται τα πλεονεκτήματα και τα μειονεκτήματα της εφαρμογής.

Κεφάλαιο 6: Σε αυτό το κεφάλαιο γίνεται ανασκόπηση της εργασίας, καταγράφονται τα συμπεράσματα που προέκυψαν κατά την υλοποίηση της εφαρμογής και γίνεται αναφορά σε μελλοντική εργασία – ανοιχτά θέματα.

Παράρτημα: Εδώ γίνεται αναφορά των ελάχιστων απαιτήσεων του συστήματος για να τρέξει η εφαρμογή.

2. KINECT

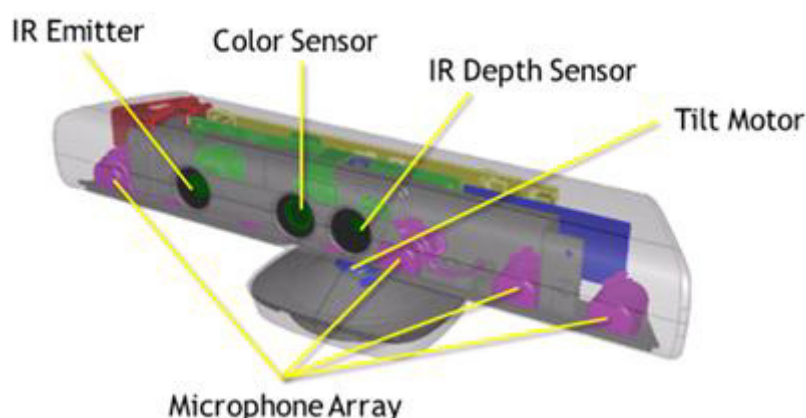
2.1 Κυκλοφορία και τεχνικά χαρακτηριστικά

Το Kinect (Εικόνα 1) είναι μία συσκευή αναγνώρισης κίνησης (gesture recognition) [2], που κυκλοφόρησε από την εταιρεία Microsoft το Νοέμβριο του 2010, και χρησιμοποιείται τόσο στα παιχνίδια του Xbox 360 καθώς επίσης και σε υπολογιστές με λειτουργικό σύστημα Windows.



Εικόνα 1: Kinect

Το Kinect βασίζεται σε τεχνολογία λογισμικού η οποία έχει αναπτυχθεί από την Rare, θυγατρική της Microsoft. Η τεχνολογία της κάμερας αναπτύχθηκε από την ισραηλινή εταιρία PrimeSense, η οποία ανέπτυξε ένα σύστημα που μπορεί να ερμηνεύσει συγκεκριμένες χειρονομίες (gestures), καθιστώντας δυνατόν τον έλεγχο ηλεκτρονικών συσκευών χωρίς κάποια άλλη συσκευή εισόδου, αλλά μόνο χειρονομίες σώματος. Για να γίνει αυτό δυνατόν χρησιμοποιείτε μια υπέρυθρη κάμερα, ένας προβολέας υπέρυθρων και ένα ειδικό μικροτσίπ (όπως φαίνονται και στην εικόνα στην επόμενη σελίδα) για να παρακολουθεί την κίνηση των αντικειμένων και τα άτομα σε τρεις διαστάσεις (3D). Αυτό το σύστημα 3D scanner που ονομάζεται Light Coding χρησιμοποιεί μια παραλλαγή εικόνας (Χάρτης Βάθους) η οποία μπορεί να αναπαρασταθεί σε τρισδιάστατο χώρο. Ο αισθητήρας Kinect έχει σχεδιαστεί για να τοποθετείται κατά μήκος πάνω ή κάτω από την οθόνη. Η συσκευή διαθέτει μία "RGB" κάμερα, αισθητήρα βάθους και multi-array μικρόφωνο και είναι ικανό να παρέχει πλήρη σωματική απεικόνιση 3D, καταγραφή κίνησης, αναγνώριση προσώπου και δυνατότητες αναγνώρισης φωνής. Ο αισθητήρας βάθους αποτελείται από ένα υπέρυθρο προβολέα λέιζερ σε συνδυασμό με ένα μονόχρωμη αισθητήρα CMOS, ο οποίος καταγράφει δεδομένα βίντεο σε 3D

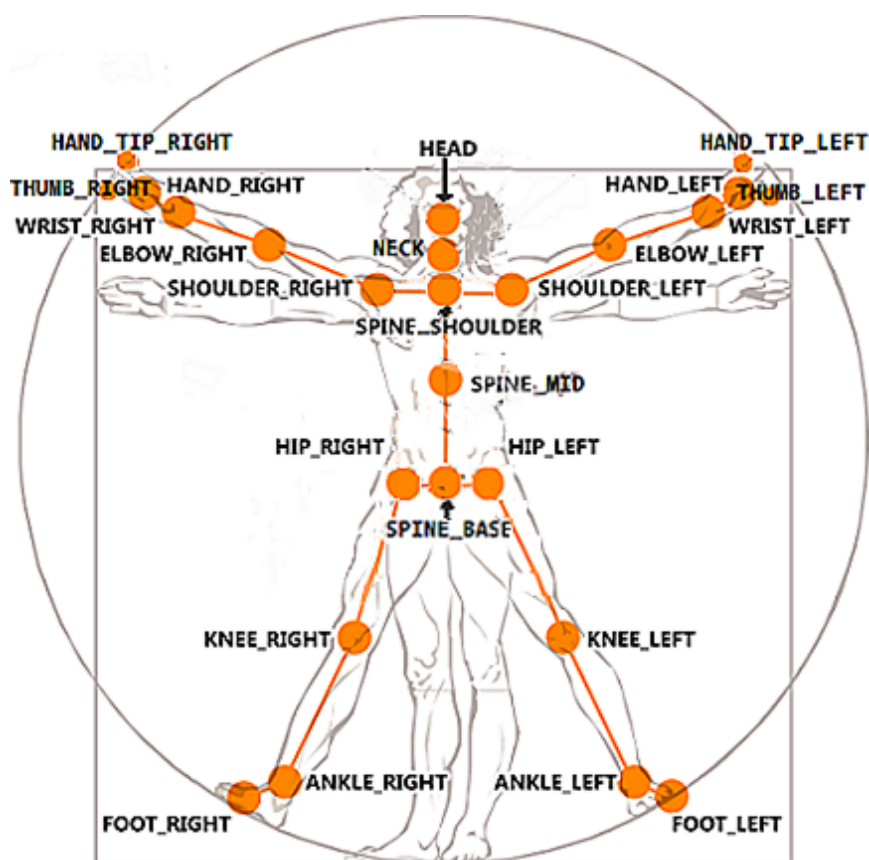


Εικόνα 2: Μέρη του Kinect

κάτω από οποιεσδήποτε συνθήκες φωτισμού. Η τεχνολογία του λογισμικού επιτρέπει την προηγμένη αναγνώριση χειρονομιών (gesture recognition), αναγνώριση προσώπου (facial recognition) και αναγνώριση φωνής (speech recognition) [12]. Σύμφωνα με τις πληροφορίες που παρέχονται στους εμπόρους λιανικής πώλησης, το Kinect είναι σε θέση να εντοπίζει ταυτόχρονα έως έξι άτομα (αλλά να κάνει tracking μόνο 2). Ο χάρτης βάθους (depth map) που δημιουργείτε από τον αισθητήρα βάθους, είναι στην ουσία μία εικόνα, η οποία χρησιμοποιώντας διαφορετικές αποχρώσεις χρώματος από λευκό έως μπλε ανάλογα με την απόσταση των αντικειμένων που "βλέπει". Το Kinect μπορεί να δει από 0.8 έως 4.0 μέτρα βάθος [3]. Αναγνωρίζει μέχρι και 20 σημεία του ανθρώπινου σώματος (joints) [4], τα οποία είναι τα εξής:

1.Base of the spine	11.Right wrist
2.Middle of the spine	12.Right hand
3.Neck	13.Left hip
4.Head	14.Left knee
5.Left shoulder	15.Left ankle
6.Left elbow	16.Left foot
7.Left wrist	17.Right hip
8.Left hand	18.Right knee
9.Right shoulder	19.Right ankle
10.Right elbow	20.Right foot

Τα σημεία του ανθρώπινου σώματος που αναγνωρίζονται φαίνονται και στην εικόνα παρακάτω:



Εικόνα 3: Τα joints που αναγνωρίζει το Kinect.

Το Kinect έχει 2 tracking modes [5]: μία λειτουργία για εντοπισμό καθιστών και μία για εντοπισμό όρθιων ανθρώπων. Η πρώτη λειτουργία έχει σχεδιαστεί για να εντοπίζει που είναι καθισμένοι σε καρέκλα ή καναπέ καθώς επίσης και για τον εντοπισμό ανθρώπων που το χαμηλότερο μέρος του σώματος τους για κάποιον λόγο δεν είναι ορατό από την κάμερα. Η δεύτερη λειτουργία αντιθέτως αναγνωρίζει και εντοπίζει ανθρώπους που στέκονται και είναι απόλυτα ορατοί από την κάμερα.

2.2 Χρήση Kinect

Το Kinect χρησιμοποιείται ευρέως για την αναγνώριση κίνησης, και λόγω του χαμηλού του κόστους και λόγω της σε πραγματικό χρόνο εκτίμησης των 3-Δ συντεταγμένων των θέσεων των αντικειμένων ή και των ανθρώπων στο χώρο.

Γενικά χρησιμοποιείται σε πολλούς τομείς:

1.gait recognition [6]

Gait recognition είναι μία τεχνολογία, η οποία χρησιμοποιείται για την αναγνώριση ανθρώπων από την ανάλυση του τρόπου που περπατάνε.

2.νοηματική γλώσσα (sign language) [7]

Η νοηματική γλώσσα χρησιμοποιείται για την επικοινωνία με ανθρώπους που δεν έχουν τη δυνατότητα να ακούνε. Οι περισσότεροι άνθρωποι (από αυτούς που ακούνε), δεν καταλαβαίνουν τη νοηματική γλώσσα. Με τη χρήση του Kinect μπορούν να αναπτυχθούν εφαρμογές που βοηθούν στην μετάφραση της νοηματικής στην καθομιλουμένη γλώσσα.

3.καθοδήση robot [8]

Με τη χρήση της κάμερας Kinect είναι δυνατή η καθοδήση ενός ρομπότ. Ο χρήστης θα βρίσκεται μπροστά από την κάμερα και θα έχει τη δυνατότητα να το κατευθύνει ανάλογα με τις κινήσεις που θα εκτελεί.

4.παροχή βοήθειας σε ανθρώπους με εγκεφαλικά [9]

Με τη χρήση του Kinect μπορεί να γίνει πιο ευχάριστη η αποθεραπεία αυτών των ανθρώπων. Μπορούν μέσα από διαδραστικά παιχνίδια να εκτελούν τις κινήσεις που τους αναθέτουν οι γιατροί για την αντιμετώπιση των προβλημάτων τους.

3. ΣΧΕΤΙΚΗ ΕΡΓΑΣΙΑ

Σε αυτή την ενότητα αναφέρονται δύο σχετικές εργασίες, που χρησιμοποιούν μία βάση δεδομένων που περιέχει κινήσεις. Η πρώτη αναφέρει τον τρόπο δημιουργίας μιας βάσης δεδομένων κινήσεων (κάτι το οποίο πραγματεύεται και η παρούσα πτυχιακή) ενώ η δεύτερη περιγράφει το χρήση μιας τέτοιας βάσης στα πλαίσια ενός μεγαλύτερου project.

Σχετικά με αυτό λοιπόν οι Alena Korpanicakova και Maria Vircikova έκαναν ένα project [10], το οποίο αναφέρεται σε έναν DTW Gesture Recognizer [16]. Έχουν δημιουργήσει ένα πρόγραμμα το οποίο δημιουργεί μία βάση κινήσεων και μετά αναγνωρίζουν κινήσεις χρησιμοποιώντας τις πληροφορίες από τη βάση. Στη συγκεκριμένη εφαρμογή επικεντρώνονται στην ανθρωποκεντρική ρομποτική, δηλαδή στη διάδραση ανθρώπου - μηχανής(ρομπότ). Πιστεύουν ότι μεταφέροντας την ανθρώπινη επικοινωνία (όπως τις κινήσεις) στα ρομπότ, ανοίγει το δρόμο για τη διάδραση ανθρώπου - ρομπότ. Το πρόγραμμα επιτρέπει στο χρήστη να καταγράψει τις δικές του κινήσεις. Αν ο χρήστης είναι ευχαριστημένος από την κίνηση που εκτέλεσε, μπορεί να την αποθηκεύσει. Διαφορετικά μπορεί να την επαναλάβει μέχρι να ικανοποιηθεί. Έτσι λοιπόν σιγά-σιγά χτίζει τη δική του βιβλιοθήκη κινήσεων. Ο χρήστης μπορεί να δημιουργήσει τους δικούς του τύπους κινήσεων. Οι κινήσεις που έχει εκτελέσει στο παρελθόν, μπορούν εύκολα να να φορτωθούν και να χρησιμοποιηθούν ξανά. Μετά αυτές οι κινήσεις ελέγχονται για το αν ταιριάζουν με άλλες κινήσεις που υπάρχουν στη βάση. Για να γίνει αυτό χρησιμοποιείται ο DTW αλγόριθμος [10], ο οποίος δεν δίνει σημασία στη χρονική διάρκεια ή στην ταχύτητα με την οποία πραγματοποιήθηκε η κίνηση.

Ένα άλλο σχετικό project [11] είναι των Lu Xia, Chia-Chih Chen και K. Aggarwal. Προσπαθούν να κάνουν αναγνώριση κίνησης με τη χρήση ιστογραμμάτων 3-Δ σημείων του ανθρώπινου σκελετού, που αναγνωρίζει το Kinect. Οι βάση τους αποτελείται από 200 3-Δ ακολουθίες 10 κινήσεων που έχουν πραγματοποιηθεί από 10 ανθρώπους. Η μέθοδός τους είναι σε πραγματικό χρόνο και έχει αρκετά καλά αποτελέσματα για βάση δεδομένων με 3-Δ στοιχεία. Το ανθρώπινο σώμα σύστημα από τμήματα που συνδέονται από αρθρώσεις και ανθρώπινη κίνηση είναι το σύνολο της αλλαγής των θέσεων των τμημάτων αυτών σε κάποιο δεδομένο χρονικό διάστημα. Εδώ χρησιμοποιούν τις θέσεις των αρθρώσεων για να δημιουργήσουν μία συμπαγή παρουσίαση κάποιας στάσης του ανθρώπινου σώματος. Εξάγουν τις 3-Δ θέσεις των αρθρώσεων από μία εικόνα βάθους (depth image). Εκτιμούν τις 3-Δ θέσεις τους στο χώρο χρησιμοποιώντας μία προσέγγιση, που βασίζεται στη μέση τιμή με τα βάρη του Gaussian πυρήνα. Το μεγάλο και ποικίλο εκπαιδευτικό σύνολο που έχουν επιτρέπει στον ταξινομητή να εκτιμά τα ανθρώπινα μέρη ανεξαρτήτως στάσης, σχήμα σώματος, ντυσίματος. Εν συνεχεία χρησιμοποιούν τις αρθρώσεις αυτές για να φτιάξουν την παρουσίαση των στάσεων. Μεταξύ των αρθρώσεων hand - wrist και foot - ankle οι αποστάσεις είναι πολύ μικρές και γι αυτό δε χρησιμοποιούνται στην περιγραφή των στάσεων. Επίσης τα σημεία spine, neck, shoulder δεν χρησιμοποιούνται κατά κύριο λόγο σε εσωτερικές δραστηριότητες. Άρα ουσιαστικά υπολογίζουν το ιστογράμμο βασιζόμενοι στα 12 από τα 20 σημεία.

Πρέπει να έχουμε στα υπόψη ότι οι θέσεις της κάθε άρθρωσης που επιστρέφει το Kinect είναι ανάλογη της θέσης ανθρώπου - κάμερας.

Οι Johannes Preis, Moritz Kessel, Martin Werner και Claudia Linnhoff-Popien [20] χρησιμοποίησαν την κάμερα Kinect και μία βάση δεδομένων από κινήσεις για gait recognition. Χρησιμοποιώντας τα 20 σημεία του ανθρώπινου σκελετού που αναγνωρίζει το Kinect, ορίζουν 13 βιομετρικά χαρακτηριστικά για την αναγνώριση ενός ατόμου που είναι τα εξής: ύψος, μήκος ποδιών, τα δύο πόδια, οι δύο μηροί, τα δύο χέρια, οι δύο πήχεις, το μήκος του βήματος και η ταχύτητα. Επίσης χρησιμοποίησαν τρεις ταξινομητές (1R, ένα C4.5 δέντρο απόφασης και τον ταξινομητή Bayes) για να πειραματιστούν με 9 ανθρώπους και να αποφασίσουν ποιος ταξινομητής είναι περισσότερο αποδοτικός για gait recognition. Αυτοί οι τρεις ταξινομητές εκπαιδεύτηκαν από μία βάση δεδομένων που περιέχει κινήσεις του ανθρώπινου σκελετού 9 ανθρώπων.

Ο 1R παράγει έναν κανόνα ταξινόμησης βασιζόμενος σε ένα μόνο χαρακτηριστικό των δεδομένων εκπαίδευσης. Μετά ταξινομεί τα δεδομένα με βάση την τιμή του συγκεκριμένου χαρακτηριστικού.

Ο C4.5 αλγόριθμος παράγει ένα δέντρο απόφασης του οποίου οι εσωτερικοί κόμβοι αποτελούν ουσιαστικά δυαδικούς ελέγχους ενός χαρακτηριστικού των δεδομένων, ενώ τα φύλλα αποτελούν ουσιαστικά τις κλάσεις. Η ταξινόμηση ενός νέου στιγμιότυπου γίνεται ακολουθώντας κάποιο μονοπάτι του δέντρου ελέγχοντας σε κάθε κόμβο του δέντρου την τιμή του χαρακτηριστικού που μας ενδιαφέρει.

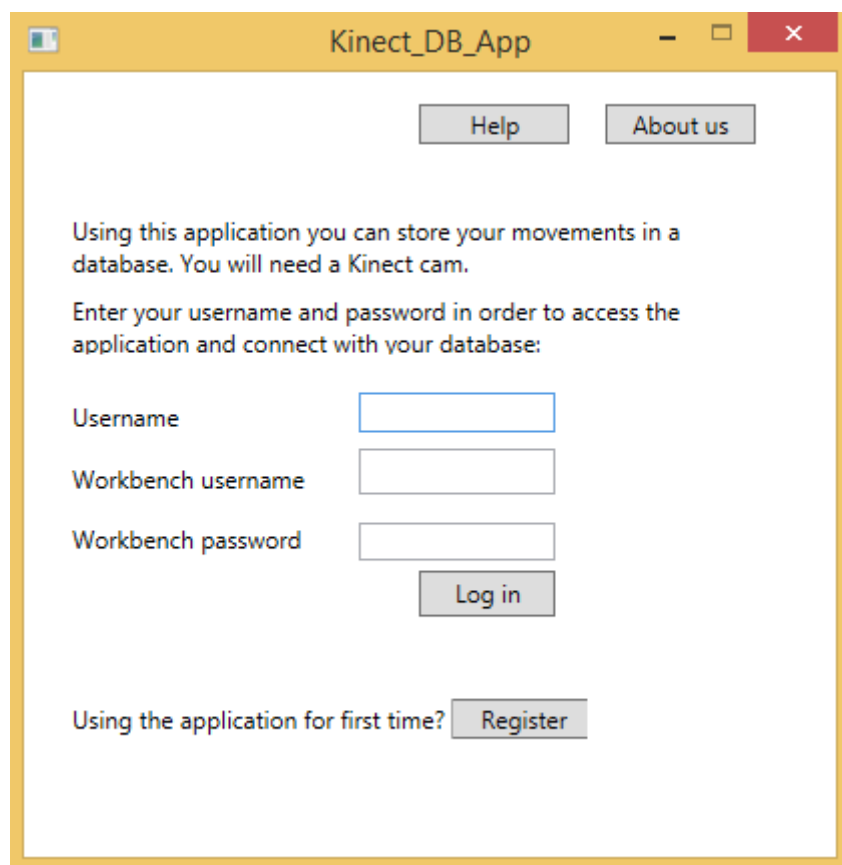
Ο ταξινομητής Bayes είναι ένας πιθανοτικός ταξινομητής βασιζόμενος στον κανόνα του Bayes [21].

4. ΥΛΟΠΟΙΗΣΗ ΚΑΙ ΤΕΧΝΟΛΟΓΙΕΣ

4.1 Περιγραφή υλοποίησης

Σε αυτή την υποενότητα θα περιγραφεί η υλοποίηση της εφαρμογής. Παρακάτω παραθέτονται τα αρχεία, η λειτουργικότητα του καθενός, καθώς και εικόνες των παραθύρων της εφαρμογής.

α) MainWindow.xaml

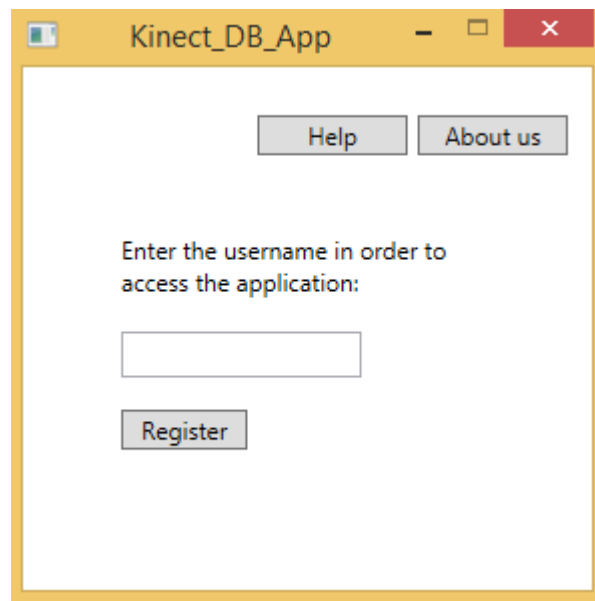


Εικόνα 4: Παράθυρο Log in.

Σε αυτό φαίνεται το πρώτο παράθυρο της εφαρμογής. Ο χρήστης καλείται να εισάγει το username του και τον κωδικό του στο workbench για να κάνει log in και να συνδεθεί η εφαρμογή με τη βάση.

Επίσης έχουμε τα κουμπιά Help, About us και Register, τα οποία περιγράφονται παρακάτω.

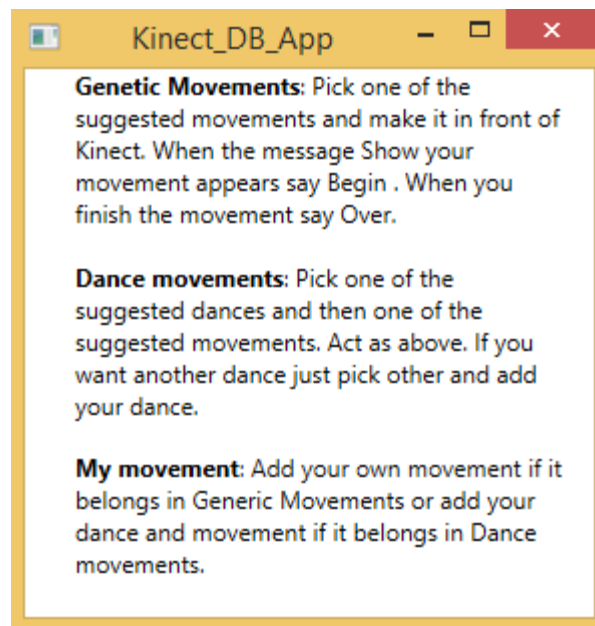
β) Register.xaml



Εικόνα 5: Παράθυρο Register.

Αν είναι η πρώτη επαφή του χρήστη με την εφαρμογή τότε πρέπει να κάνει Register. Ουσιαστικά αυτό γίνεται για να έχει ο κάθε χρήστης τις δικές του κινήσεις (κάτι που εξηγείται στην υποενότητα 4.3.α). Ο χρήστης λοιπόν εισάγει το όνομα που θέλει. Σε περίπτωση που το username που θέλει υπάρχει ήδη, τότε εμφανίζεται ανάλογο μήνυμα που προτρέπει το χρήστη να δοκιμάσει κάποιο άλλο.

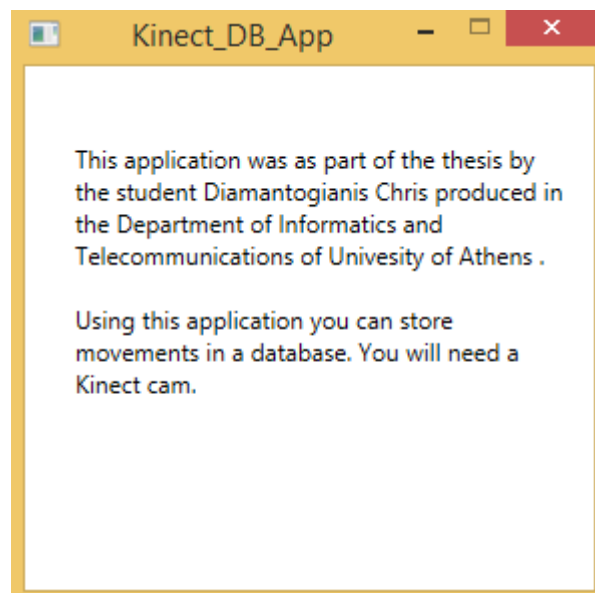
γ) Help.xaml



Εικόνα6: Παράθυρο για Βοήθεια.

Παρέχει πληροφορίες για τη λειτουργικότητα των κουμπιών.

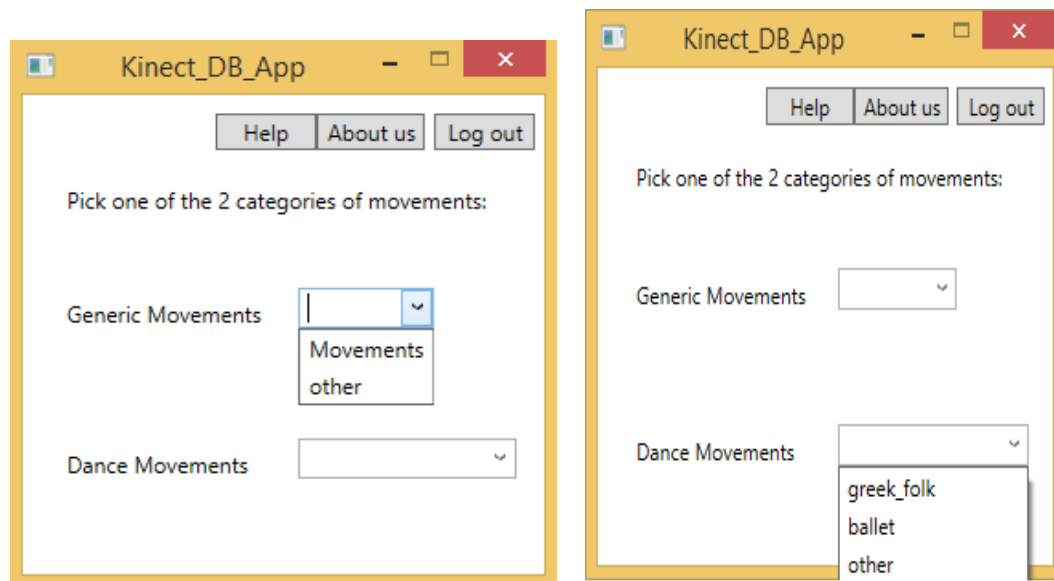
δ) About_us.xaml



Εικόνα7: Παράθυρο για πληροφορίες του δημιουργού.

Στο αρχείο αυτό υπάρχουν πληροφορίες για το δημιουργό της εφαρμογής.

ε) Gestures_Database.xml



Εικόνα 8: Παράθυρο όπου ο χρήστης επιλέγει αν θέλει να κάνει Generic ή Dance Movement.

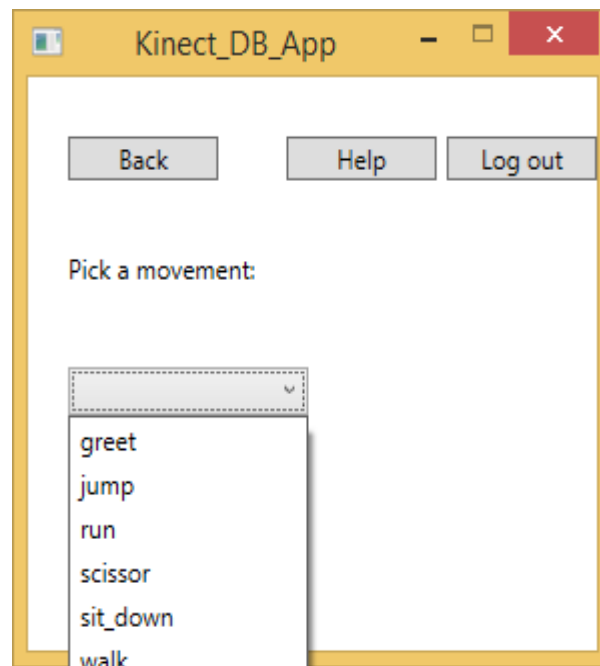
Σε αυτό το παράθυρο ο χρήστης διαλέγει αν θέλει να πραγματοποιήσει μια generic ή dance movement.

Στην πρώτη κατηγορία προτείνονται οποιαδήποτε ήδη κινήσεων (εκτός χορευτικών). Ο χρήστης έχει τη δυνατότητα να διαλέξει κάποια από τις προτεινόμενες κινήσεις ή ακόμα και να πραγματοποιήσει οποιαδήποτε δική του κίνηση επιθυμεί.

Στη δεύτερη κατηγορία ο χρήστης έχει τη δυνατότητα να επιλέξει κάποιον από τους προτεινόμενους χορούς και κάποια από τις προτεινόμενες κινήσεις για τον κάθε χορό. Επίσης μπορεί να εισάγει κάποια δική του κίνηση για συγκεκριμένο χορό. Ακόμα μπορεί να εισάγει και νέο χορό με δικές του κινήσεις.

Επίσης πέρα από τα κουμπιά Help και About us έχουμε και το Log out. Αυτό το κουμπί χρησιμοποιείται στην περίπτωση που κάποιος άλλος χρήστης θελήσει να πραγματοποιήσει κίνηση. Οπότε πατάει Log out και κάνει Log in με το δικό του username.

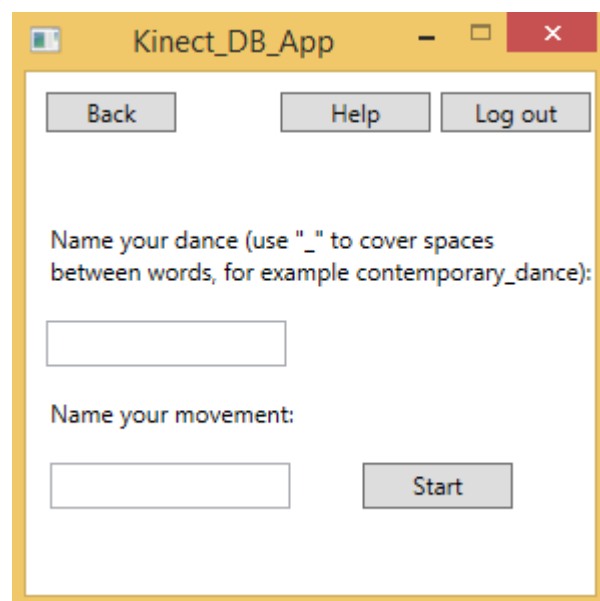
στ) Generic_Movements.xaml



Εικόνα 9: Παράθυρο όπου ο χρήστης διαλέγει ποια Generic Movement θα εκτελέσει.

Αν στο προηγούμενο παράθυρο ο χρήστης επιλέξει να κάνει κίνηση η οποία βρίσκεται ήδη στη βάση τότε εμφανίζεται αυτό το παράθυρο. Εδώ ο χρήστης επιλέγει ποια από τις προτεινόμενες κινήσεις θα πραγματοποιήσει.

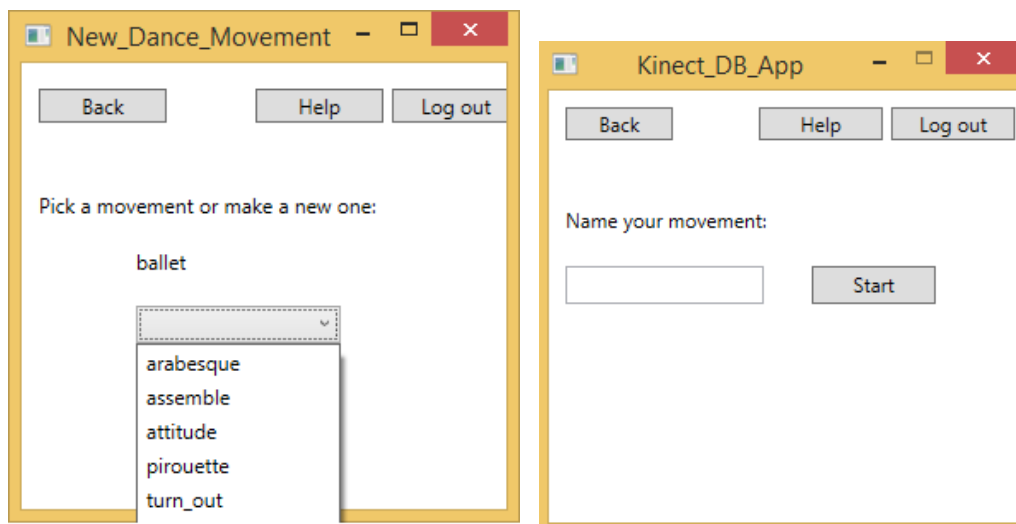
ζ) New_Dance_and_Gesture.xaml



Εικόνα 10: Παράθυρο όπου ο χρήστης εισάγει νέο χορό και την αντίστοιχη νέα κίνηση που θέλει να εκτελέσει.

Αν ο χρήστης επιλέξει να κάνει κάποιο καινούριο χορό τότε θα ανοίξει το παραπάνω παράθυρο και ο χρήστης τώρα καλείται να εισάγει το δικό του χορό καθώς και κάποια αντίστοιχη χορευτική κίνηση.

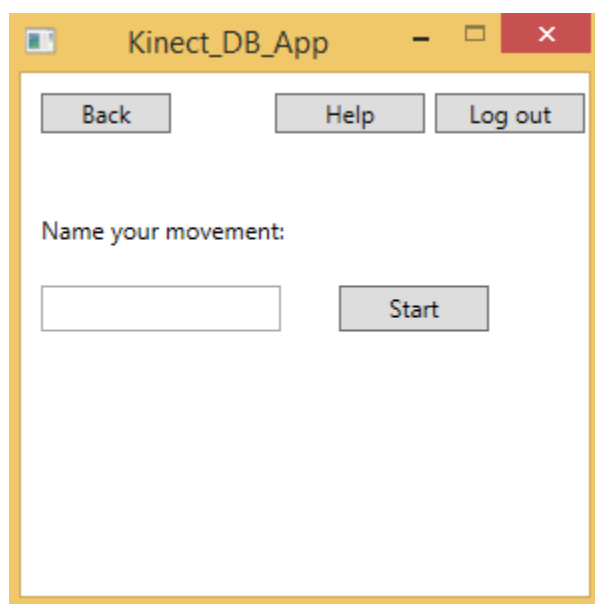
η) New_Dance_Movement.xaml



Εικόνα 11: Παράθυρο όπου ο χρήστης διαλέγει αν θα εκτελέσει μία προτεινόμενη κίνηση ή αν θα κάνει μία δική του.

Αν ο χρήστης επιλέξει να πραγματοποιήσει κίνηση για κάποιον χορό που υπάρχει ήδη στη βάση, τότε θα εμφανιστεί το παραπάνω παράθυρο το οποίο προτρέπει το χρήστη να εκτελέσει κάποια προτεινόμενη κίνηση για το συγκεκριμένο χορό ή να εισάγει τη δική του χορευτική κίνηση.

θ) My_Gesture.xaml



Εικόνα 12: Παράθυρο όπου ο χρήστης ονοματίζει τη νέα Generic κίνηση που θέλει να εκτελέσει.

Αν ο χρήστης στο παράθυρο `Gesture_Database.xaml` επιλέξει στην κατηγορία `Generic_Movements` το `other` τότε εμφανίζεται το παραπάνω παράθυρο που προτρέπει το χρήστη να εισάγει μία δική του κίνηση.

ι) `Global_Variables.cs`

Σε αυτό το αρχείο γίνονται τα πάντα όσων αφορά τη βάση και τη λειτουργία του Kinect.

Στις συναρτήσεις `create_connection` και `open_connection` δημιουργούμε και ανοίγουμε τη σύνδεση με το workbench αντίστοιχα.

Για το Kinect τώρα έχουμε την `public static KinectSensor initialize_kinect()`, στην οποία βρίσκουμε την κάμερα που έχουμε συνδέσει στον υπολογιστή μας, ενεργοποιούμε το `SkeletonStream` (η κάμερα επιστρέφει ακόμα τα `ImageStream` και `DepthStream`) και δημιουργούμε έναν `Speech recognizer`. Αφού γίνουν όλα αυτά, θέτουμε σε λειτουργία την κάμερα.

Πριν είπαμε ότι δημιουργούμε έναν `Speech recognizer`. Αυτό γίνεται στη συνάρτηση:

`public static SpeechRecognitionEngine create_speech_recognizer()`. Εδώ δημιουργούμε τη γραμματική που θα χρησιμοποιήσουμε. Στη συγκεκριμένη εφαρμογή έχω ορίσει να έχω 2 φωνητικές εντολές: `BEGIN`, `OVER` (θα εξηγηθεί η χρήση τους παρακάτω). Επίσης εδώ ορίζουμε και το γεγονός `SSpeechRecognized`, όπου εκεί ορίζεται η αντίδραση της εφαρμογής όταν αναγνωρίσει κάποια από τις δύο προαναφερθέντες λέξεις. Όταν ο χρήστης πει τη λέξη `BEGIN` τότε η εφαρμογή προτρέπει το χρήστη να ξεκινήσει την κίνηση του και το Kinect αρχίζει να στέλνει frames (της κίνησης που εκτελεί ο χρήστης). Όταν ο χρήστης τελειώσει την κίνησή του πρέπει να πει `OVER`. Τότε εκτελείται η συνάρτηση `insert_in_database()`.

Πριν μιλήσουμε για την `insert_in_database()` θα πούμε για κάποιες κλάσεις του προγράμματος, όπως επίσης και για τη συνάρτηση:

`public static void frame_insert_in_list(Skeleton skeleton, int frame_id)`. Έχω ορίσει μία κλάση `joint` η οποία ουσιαστικά έχει τρεις float αριθμούς, που αποτελούν τις συντεταγμένες του `joint`. Άλλη μια κλάση είναι η `frame`, η οποία περιέχει έναν αριθμό (αναγνωριστικό του κάθε `frame`) και έναν πίνακα με τα `joints` του (τα οποία είναι όσα αναγνωρίζει η κάμερα, δηλαδή 20). Επίσης έχω ορίσει και μία λίστα από `frames` τη `frame_list`.

Η `public static void frame_insert_in_list(Skeleton skeleton, int frame_id)` είναι η συνάρτηση στην οποία δημιουργεί τα `frames` της κίνησης (παίρνει από την `skelton` τις συντεταγμένες και για τα 20 `joints`) και τα εισάγει στη `frame_list`.

Η `insert_in_database()` κάνει τα εξής:

- 1) Εισαγωγή της κίνησης στη βάση.
- 2) Εισαγωγή των user_id και gesture_id για να μπορεί να ταιριάζει τον κάθε χρήστη με τις κινήσεις του.
- 3) Εισαγωγή του κάθε frame της frame_list.
- 4) Εισαγωγή των gesture_id και frame_id για να μπορεί να ταιριάζει την κάθε κίνηση με τα δικά της frames.
- 5) Εισαγωγή του κάθε joint του κάθε frame.
- 6) Εισαγωγή των frame_id και joint_id για να μπορεί να ταιριάζει το κάθε frame με τα αντίστοιχα joints.

Η `static void SensorSkeletonFrameReady(object sender, SkeletonFrameReadyEventArgs e)` ουσιαστικά καθορίζει το πώς θα αντιδρά το πρόγραμμα όταν το Kinect στέλνει frames. Έχουμε τη boolean μεταβλητή `start_frame_recognition`, η οποία θα έχει την τιμή `true` μόνο όταν ο χρήστης πει `BEGIN`. Όταν συμβεί αυτό τότε κάθε 10 frames (το Kinect έχει τη δυνατότητα να καταγράφει 30 frames/sec, άρα εδώ κρατάει 3 frames/sec), το πρόγραμμα παίρνει το frame που στέλνει το Kinect και το εισάγει σε μία προσωρινή λίστα μέχρι να έρθει η ώρα της εισαγωγής στη βάση που εξηγήθηκε παραπάνω.

Το σχήμα της βάσης θα εξηγηθεί στην υποενότητα 4.3 (ποιες οντότητες υπάρχουν και τι περιέχει η κάθε μία) .

Γενικά η λειτουργικότητα του κάθε παραθύρου (αρχείου .xaml) καθορίζεται από το αντίστοιχο αρχείο .cs. Για παράδειγμα αν έχουμε το αρχείο `x.xaml` η λειτουργικότητα των κουμπιών, μορφή παράθυρου, κτλ καθορίζεται από το αρχείο `x.xaml.cs` .

4.2 Τεχνολογίες που χρησιμοποιήθηκαν

Για την υλοποίηση της συγκεκριμένης εφαρμογής χρησιμοποιήθηκαν τα εξής:

α) Η κάμερα Kinect για την οποία μιλήσαμε στην ενότητα 2.

β) Το Microsoft Kinect for Windows SDK

Το SDK αναπτύχθηκε από τη Microsoft και κυκλοφόρησε για πρώτη φορά στην αγορά το 2011. Το SDK περιλαμβάνει όλα τα απαραίτητα εργαλεία που είναι απαραίτητα για την ανάπτυξη εφαρμογών συμβατές με το Kinect και δίνει τη δυνατότητα στον κάθε προγραμματιστή να αναπτύξει εφαρμογές σε C#, C++, Visual Basic χρησιμοποιώντας το Visual Studio. Γι αυτή την εφαρμογή χρησιμοποιήθηκε η έκδοση 1.8 για το SDK και η έκδοση 2013 του Visual Studio.

Το SDK στη συγκεκριμένη εφαρμογή χρησιμοποιήθηκε για την αναγνώριση του ανθρώπινου σκελετού (skeleton tracking) και τη μετέπειτα προσωρινή αποθήκευση στο πρόγραμμα, για τη μεταφορά τους στη βάση.

γ) Γλώσσα Προγραμματισμού C#

Η C# [13] είναι μία αντικειμενοστραφής γλώσσα προγραμματισμού που αναπτύχθηκε από τη Microsoft και χρησιμοποιήθηκε για να αναπτυχθεί το μεγαλύτερο μέρος του κώδικα. Ουσιαστικά σε C# έγιναν όλα τα κομμάτια του κώδικα εκτός του μέρους που ασχολείται με το πώς θα είναι τα παράθυρα της εφαρμογής.

δ) XAML

Η XAML (Extensive Application Markup Language) [14] είναι μία γλώσσα σήμανσης (Markup Language) που βασίζεται στην ευρέως διαδεδομένη γλώσσα σήμανσης XML. Η XAML αναπτύχθηκε από τη Microsoft. Χρησιμοποιείται κατά κύριο λόγο στην τεχνολογία .Net Framework και ειδικά σε εφαρμογές τύπου WPF(Windows Presentation Foundation) και Silverlight. Η παρούσα εφαρμογή είναι τύπου WPF. Η XAML στη συγκεκριμένη εφαρμογή χρησιμοποιήθηκε για τη δημιουργία των παραθύρων της εφαρμογής, τη δημιουργία πλαισίων κειμένου (textbox), τη δημιουργία κουμπιών καθώς επίσης και τα γεγονότα που θα συμβαίνουν κατά το πάτημά τους (events).

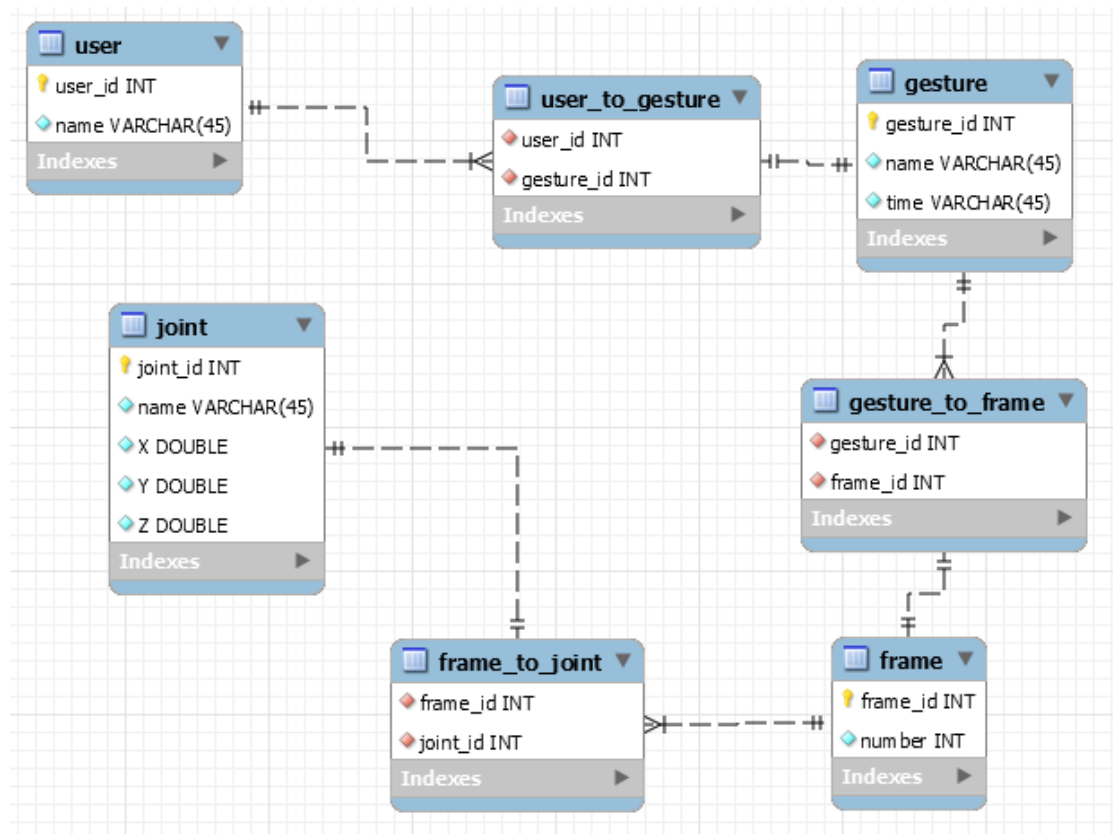
ε) MySQL Workbench

Το MySQL Workbench [15] είναι ένα εργαλείο για την ανάπτυξη βάσεων δεδομένων. Έχει χρησιμοποιηθεί από χιλιάδες προγραμματιστές και είναι πολύ βολικό από άποψη σχεδιασμού βάσεων δεδομένων ελέγχου καθώς και debugging. Υποστηρίζει τις εξής σχεσιακές βάσεις δεδομένων: Oracle Database, Microsoft SQL Server, SQL Anywhere, Firebird, NexusDB, InterBase, MySQL and MariaDB. Για τη συγκεκριμένη εφαρμογή χρησιμοποιήθηκε η έκδοση 6.3.

Το MySQL Workbench χρησιμοποιήθηκε για τη δημιουργία της βάσης δεδομένων που θα έχει τα frames της κάθε κίνησης (οποιοδήποτε χορού) του κάθε χρήστη. Στην επόμενη υποενότητα ακολουθούν το σχήμα της βάσης καθώς επίσης και η επεξήγησή του.

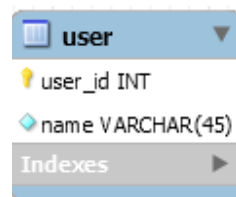
4.3 Σχήμα βάσης

Το σχήμα της βάσης φαίνεται στην εικόνα παρακάτω:



Εικόνα 13: Σχήμα βάσης.

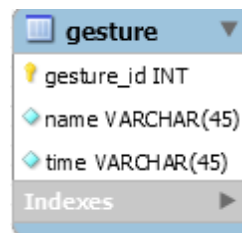
α) user



Εικόνα 14: Οντότητα user.

Κάθε χρήστης έχει τις δικές του κινήσεις. Δεν μπορούμε απλά να έχουμε τις κινήσεις σκέτες στη βάση γιατί από άνθρωπο σε άνθρωπο διαφέρουν. Για παράδειγμα κάποιος μπορεί να είναι πιο ψηλός, κάποιος πιο γυμνασμένος, διαφορετική απόσταση από την κάμερα (πράγμα που επηρεάζει τη συντεταγμένη Z των joints) με αποτέλεσμα οι αρθρώσεις (joints) του καθενός που αναγνωρίζει το Kinect να μην έχουν τις ίδιες συντεταγμένες (στον ψηλότερο χρήστη η Y συντεταγμένη του κεφαλιού θα έχει τιμή μεγαλύτερη από αυτή του κοντύτερου). Οπότε ο κάθε χρήστης έχει τις δικές του κινήσεις.

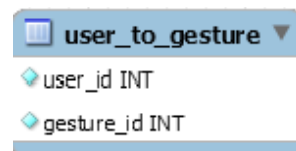
β) gesture



Εικόνα 15: Οντότητα gesture.

Gesture είναι η κίνηση που πραγματοποιεί ο εκάστοτε χρήστης και αποτελείται από το όνομά της και τη χρονική στιγμή που έγινε.

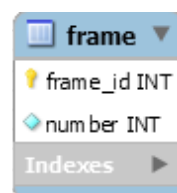
γ) user_to_gesture



Εικόνα 16: Το Table που ταιριάζει τους users με τις αντίστοιχες κινήσεις που εκτέλεσε ο καθένας.

Το user_to_gesture αντιστοιχίζει τους χρήστες με τις κινήσεις του. Όπως αναφέραμε και παραπάνω ο κάθε χρήστης έχει τις δικές του κινήσεις.

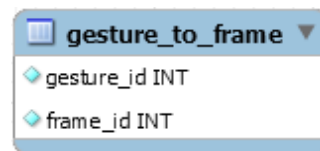
δ) frame



Εικόνα 17: Οντότητα frame.

Κάθε κίνηση αποτελείται από πλαίσια (frames). Το κάθε frame έχει τον αριθμό του. Κάθε πλαίσιο που επιστρέφει η κάμερα έχει το δικό του μοναδικό αριθμό.

ε) gesture_to_frame

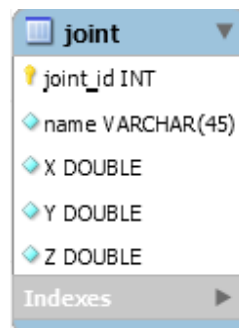


gesture_to_frame	
gesture_id	INT
frame_id	INT

Εικόνα 18: Το table που αντιστοιχίζει τις κινήσεις με τα αντίστοιχα frames.

Το gesture_to_frame αντιστοιχίζει τα frames στην κάθε κίνηση. Ουσιαστικά η εκάστοτε κίνηση έχει τα δικά της frames.

στ) joint



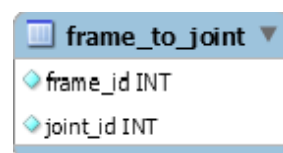
joint	
joint_id	INT
name	VARCHAR(45)
X	DOUBLE
Y	DOUBLE
Z	DOUBLE

Indexes

Εικόνα 19: Οντότητα joint.

Όπως αναφέραμε και στη δεύτερη ενότητα το Kinect μπορεί να κάνει tracking τον ανθρώπινο σκελετό αναγνωρίζοντας ουσιαστικά 20 joints. Το κάθε joint έχει το όνομά του καθώς επίσης και τις συντεταγμένες του. Για παράδειγμα το joint head θα έχει το όνομά του που θα είναι head και τις συντεταγμένες του (ανάλογα τον άνθρωπο).

ζ) frame_to_joint

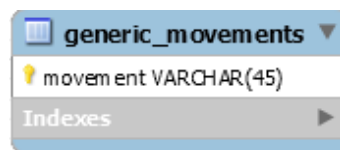


frame_to_joint	
frame_id	INT
joint_id	INT

Εικόνα 20: Το table που αντιστοιχίζει τα frames με τα joints τους.

Το frame_to_joint αντιστοιχίζει τα αντίστοιχα frames με τα joints. Κάθε frame ουσιαστικά θα έχει 20 joints, των οποίων οι συντεταγμένες περιγράφουν με ακρίβεια τη στάση που έχει ο ανθρώπινος σκελετός την κάθε χρονική στιγμή που τον βλέπει η κάμερα.

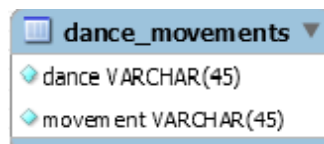
η) generic_movements



Εικόνα 21: Η οντότητα generic movements.

Τα generic_movements περιέχουν τα ονόματα των κινήσεων που ανήκουν σε αυτήν την κατηγορία (εξηγήθηκε στην υποενότητα 4.1). Κάθε όνομα κίνησης αναφέρετε μόνο μία φορά και χρησιμοποιείται για να δώσει παραπάνω επιλογές στο χρήστη κατά τη χρήση της εφαρμογής.

θ) dance_movements



Εικόνα 22: Η οντότητα dance movements.

Τα dance_movements περιέχουν τα ονόματα των χορών καθώς επίσης και τα ονόματα των κινήσεων που ανήκουν στον κάθε χορό. Για παράδειγμα dance: greek folk, ballet, movement: relever. Χρησιμοποιείται όπως και η παραπάνω οντότητα, δηλαδή για να δώσει παραπάνω επιλογές στον εκάστοτε χρήστη. Ακολουθούν κάποιες ενδεικτικές εικόνες για χορευτικές κινήσεις.

Greek folk

Ψαλιδάκι [19]



Εικόνα 23: Ψαλιδάκι

Ballet

Arabesque [17]



Εικόνα 24: Arabesque

Attitude [18]



Εικόνα 25: Attitude

5. ΔΥΣΚΟΛΙΕΣ ΚΑΤΑ ΤΗΝ ΥΛΟΠΟΙΗΣΗ

5.1 Προβληματισμοί

Κατά τη διάρκεια της ανάπτυξης της εφαρμογής υπήρξαν κάποια σημεία που απαιτήσαν ιδιαίτερη προσοχή και σκέψη μέχρι τη λήψη της τελικής απόφασης.

Ένα πρόβλημα ήταν ο αριθμός των frames που θα αποθηκεύονταν στη βάση. Προφανώς όσο πιο πολλά, τόσο καλύτερα περιγράφεται η κίνηση. Όμως αυτό κοστίζει σε χρόνο. Οπότε πρέπει να βρεθεί ο βέλτιστος αριθμός για τη συγκεκριμένη περίπτωση. Στην αρχή σκέφτηκα, την κάθε κίνηση να την αντιπροσωπεύουν 3 συνολικά frames(αρχή – μέση – τέλος). Και για μικρής χρονικής διάρκειας κινήσεις είναι όντως αντιπροσωπευτικός ο αριθμός. Όμως αυτό δεν αποδεικνύει, ότι με τον ίδιο αριθμό frames θα περιγράφονται εξίσου καλά κινήσεις μεγαλύτερης χρονικής διάρκειας. Δεδομένου αυτού λοιπόν αποφάσισα ότι θα ήταν καλύτερο να αποθηκεύονται 3 frames/sec. Το μειονέκτημα με αυτή την προσέγγιση είναι ότι αν για μεγάλης χρονικής διάρκειας κίνηση, η αποθήκευση των δεδομένων στη βάση είναι χρονοβόρα. Το πλεονέκτημα είναι ότι η κίνηση περιγράφεται αρκετά καλά, αφού δίνουμε περισσότερες πληροφορίες στη βάση.

Ένας άλλος προβληματισμός ήταν ο τρόπος με τον οποίο θα γινόταν ο χειρισμός του προγράμματος, όταν ο χρήστης πάταγε το κουμπί για να εκτελέσει κάποια κίνηση (προτεινόμενη ή μη, generic ή dance). Η πρώτη σκέψη ήταν, όταν ο χρήστης πάταγε το κουμπί για να κάνει την κίνηση, η εφαρμογή να δίνει ένα χρονικό περιθώριο (για παράδειγμα 3 sec) μέχρι ο χρήστης να προλάβει να σταθεί μπροστά από την κάμερα σε απόσταση 1 - 1,5m. Αυτή η λύση απορρίφθηκε, διότι αναγκάζοντας το χρήστη να πηγαίνει μπροστά από την κάμερα σε3 συγκεκριμένο χρονικό διάστημα μειώνεται η ευχρηστία του προγράμματος. Μία άλλη σκέψη πάνω σε αυτό το πρόβλημα ήταν να υπάρχει άλλος ένας άνθρωπος, ο οποίος θα βοηθούσε τον χρήστη που θα ήθελε να πραγματοποιήσει κάποια κίνηση. Οπότε ο χρήστης που θα ήθελε να εκτελέσει τις κινήσεις θα ήταν ήδη μπροστά από την κάμερα και ο άλλος θα διαχειριζόταν ανάλογα το πρόγραμμα. Απορρίφθηκε και αυτή η λύση. Εφαρμόζοντας την υποχρεώνεις το χρήστη να έχει οπωσδήποτε έναν “βοηθό”, κάτι το οποίο μειώνει αρκετά την ευχρηστία της εφαρμογής, αφού απαιτεί την ύπαρξη ενός ακόμα προσώπου.

Ο παραπάνω προβληματισμός υπήρξε γιατί με κάποιον τρόπο θα έπρεπε, αφού ο χρήστης τελειώσει την κίνησή του, να σταματήσει η κάμερα να δίνει frames στο πρόγραμμα. Κάπως λοιπόν θα έπρεπε να σταματήσουμε την αποστολή περιττών frames.

Μία άλλη λύση στον παραπάνω προβληματισμό ήταν να οριστεί μία θέση τερματισμού. Τη συγκεκριμένη θέση θα την έπαιρνε ο χρήστης όταν τελείωνε την κίνησή του. Όμως και αυτή η λύση απορρίφθηκε κυρίως για δύο λόγους. Ο πρώτος λόγος είναι ότι, όταν τελείωνε την εκτέλεση της κίνησης ο χρήστης θα χρειαζόταν κάποιο χρονικό διάστημα για να φτάσει στη θέση τερματισμού. Αυτό το χρονικό διάστημα (έστω και μικρό) αυτόματα σημαίνει καταγραφή παραπάνω frames τα οποία όμως δεν ανήκουν στην κίνηση που ήθελε να

πραγματοποιήσει αρχικά ο χρήστης. Σαν αποτέλεσμα στη βάση θα αποθηκεύονταν περίπου frames. Αυτό μπορεί να επηρεάζει μετέπειτα άλλα project που θα χρησιμοποιούσαν τη συγκεκριμένη βάση. Ο δεύτερος λόγος είναι ότι πολύ απλά η θέση που θα ορίζαμε σαν θέση τερματισμού μπορούσε να αποτελεί φυσιολογικά μέρος κάποιας κίνησης, από αυτές που θα επιθυμούσε να εκτελέσει ο χρήστης.

Τελικά αποφασίστηκε ο χειρισμός της εφαρμογής να γίνεται κατά αποκλειστικότητα από έναν και μόνο χρήστη. Όταν ο χρήστης πατάει το κουμπί για να εκτελέσει μία οποιαδήποτε κίνηση, τότε έχει όσο χρόνο θέλει να πάει μπροστά από την κάμερα. Μετά πρέπει να δώσει την εντολή BEGIN με φωνητικό μήνυμα, για να ξεκινήσει το Kinect να στέλνει frames στο πρόγραμμα. Όταν ολοκληρώσει την εκτέλεση της κίνησης πρέπει να δώσει την εντολή OVER με φωνητικό μήνυμα, έτσι ώστε να σταματήσει η αποστολή των frames από την κάμερα στο πρόγραμμα.

5.2 Μειονεκτήματα

α) Αν η κίνηση είναι μεγάλης χρονικής διάρκειας, τότε ο χρόνος για την αποθήκευση στη βάση δεδομένων είναι μεγάλος (ξεπερνάει τα 40 sec).

β) Υπάρχουν πολλοί χοροί που γίνονται μεταξύ ζευγαριών (για παράδειγμα τανγκό). Η συγκεκριμένη εφαρμογή όμως καταγράφει τις κινήσεις μόνο ενός χρήστη.

5.3 Πλεονεκτήματα

α) Υπάρχει η δυνατότητα, ο χρήστης να εκτελέσει είτε κινήσεις που θα του προτείνει η εφαρμογή είτε κινήσεις που θα εισάγει ο ίδιος.

β) Όλοι χρησιμοποιούν την ίδια. Οπότε όταν ένας χρήστης εκτελεί μία κίνηση δικιά του (που δεν υπάρχει δηλαδή ήδη στη βάση), τότε αυτή η κίνηση εισάγεται στη βάση. Αν αργότερα κάποιος άλλος χρήστης χρησιμοποιήσει την εφαρμογή, τότε στις προτεινόμενες κινήσεις θα έχει προστεθεί και η κίνηση του προηγούμενου.

γ) Η εντολή για να ξεκινήσει η κάμερα να στέλνει frames στο πρόγραμμα, γίνεται με φωνητικό μήνυμα, κάτι το οποίο διευκολύνει το χρήστη που δε χρειάζεται να αγχώνεται σε πόσο χρόνο πρέπει να βρεθεί μπροστά από την κάμερα ή να χρειάζεται οπωσδήποτε και κάποιον να το βοηθήσει κατά τη χρήση της εφαρμογής.

6. ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΗ ΕΡΓΑΣΙΑ

6.1 Ανασκόπηση

Στις παραπάνω ενότητες εξηγήθηκε η διαδικασία που ακολουθήθηκε για την ανάπτυξη της εφαρμογής. Αρχικά παρουσιάστηκαν το θέμα της εργασίας και ο σκοπός της υλοποίησής της. Στη συνέχεια έγινε μία πρώτη γνωριμία με την κάμερα Kinect, οι δυνατότητές της, πού χρησιμοποιείται, για ποιο λόγο χρησιμοποιείται. Μετά αναφέρθηκε σχετική εργασία πάνω σε αυτό και περιγραφή της σχετικής εφαρμογής. Ακολούθως έγινε περιγραφή της υλοποίησης της εφαρμογής που αναπτύχθηκε στα πλαίσια της πτυχιακής, αναφορά στο σχήμα της και μία περιγραφή των τεχνολογιών που χρησιμοποιήθηκαν για τη διεκπεραίωση του project. Μετά την περιγραφή αυτή ακολούθησαν οι δυσκολίες και οι προβληματισμοί που δημιουργήθηκαν κατά την ανάπτυξη της εφαρμογής καθώς επίσης και τα μειονεκτήματά της.

6.2 Συμπεράσματα

Κατά τη διάρκεια της υλοποίησης της εφαρμογής εξήχθησαν σημαντικά συμπεράσματα. Είναι αναγκαία η έρευνα για να επιλεχθεί ο κατάλληλος εξοπλισμός για την ανάπτυξη της εφαρμογής. Επίσης είναι πολύ σημαντικό να γίνουν κατανοητά τα τεχνικά χαρακτηριστικά του εξοπλισμού καθώς επίσης και ο τρόπος με τον οποίο λειτουργεί, ώστε να είναι πιο εύκολη η υλοποίηση της εφαρμογής και να αποφευχθούν τεχνικά τυχόν λάθη. Στην παρούσα εφαρμογή για παράδειγμα ήταν αναγκαία η κατανόηση της λειτουργίας και των χαρακτηριστικών του Kinect(ποια ρεύματα στέλνει, πόσα σημεία του ανθρώπινου σκελετού αναγνωρίζει, σε τι μορφή στέλνει τα frames, και γενικά πως μπορούμε να διαχειριστούμε τις πληροφορίες που στέλνει το Kinect στη C#).

6.3 Μελλοντική εργασία - ανοιχτά θέματα

Γενικά κάποια πράγματα ίσως μπορούσαν να γίνουν με άλλον τρόπο και προφανώς η εφαρμογή επιδέχεται βελτιώσεις.

Η παρούσα εφαρμογή έχει μόνο δύο είδη κινήσεων (generic, dance τα οποία αναφέρονται στο 1.2). Θα μπορούσε να επεκταθεί και για κινήσεις που ανήκουν σε άλλους τομείς (όπως για παράδειγμα στον τομέα του αθλητισμού).

Επίσης θα μπορούσε να αναγνωρίζει κινήσεις που θα κάνουν ταυτόχρονα 2 άτομα (έτσι θα χρησιμοποιούνταν στο μέγιστο οι δυνατότητες του Kinect αφού μπορεί να κάνει tracking μέχρι και δύο άτομα).

Μία μεγαλύτερη επέκταση αντίστοιχη με την παραπάνω θα ήταν η χρήση περισσότερων Kinect αισθητήρων για την καταγραφή των κινήσεων περισσότερων ανθρώπων. Σαν αποτέλεσμα θα μπορούσαν να καταγραφούν χοροί που απαιτούν περισσότερους ανθρώπους από δύο.

Στην παρούσα φάση όταν αποθηκεύουμε ένα frame στη βάση, ουσιαστικά αποθηκεύονται και τα 20 joints που αναγνωρίζει το Kinect. Όμως αυτό κάνει πολλές φορές το πρόγραμμα χρονοβόρο(εξαρτάται βέβαια και από την κίνηση καθώς και από τη χρονική διάρκειά της). Ίσως στο μέλλον να μπορεί να γίνει διάκριση μεταξύ των joints που χρειάζεται να αποθηκευτούν στη βάση ανάλογα με το όνομα της κίνησης που θα εισάγει ο χρήστης.

Για κινήσεις που έχουν χρονική διάρκεια πάνω από 3-4 sec, ο χρόνος αποθήκευσης είναι ικανοποιητικός (μέχρι 5 sec). Όμως για κινήσεις από 5 sec και πάνω ο χρόνος αποθήκευσης στη βάση είναι μεγάλος (μπορεί να ξεπερνάει και τα 40 sec). Στο μέλλον θα ήταν αναγκαίο (αν μελετούνται κινήσεις με μεγάλο μέσο όρο χρονικής διάρκειας), να βρεθεί τρόπος να μειωθεί ο χρόνος αποθήκευσης στη βάση (αυτό εκτός των άλλων επηρεάζεται και από τα χαρακτηριστικά του συστήματος και του εξοπλισμού που χρησιμοποιούνται για να αναπτυχθεί η εφαρμογή και να τρέξει ο κώδικας).

Επίσης χρειάζεται περισσότερη έρευνα πάνω στον αριθμό των frames που θα καταγράφονται για κάθε κίνηση. Είναι ικανοποιητικό να έχουμε τον ίδιο αριθμό frames για όλες τις κινήσεις; Μήπως είναι καλύτερα να έχουμε λιγότερα frames από τον αριθμό που αναφέρθηκε παραπάνω(βελτίωση χρόνου αποθήκευσης στη βάση); Ίσως θα μπορούσε ο χρήστης να εισάγει τη χρονική διάρκεια που απαιτεί η κίνηση που θα εκτελέσει και ανάλογα να οριζόταν και ο αριθμός των frames από το πρόγραμμα.

Ακόμα ίσως θα μπορούσαν να αντικατασταθούν τα κουμπιά με φωνητικές εντολές. Και επιπλέον όταν ο χρήστης θα ήθελε να εισάγει μια δική του κίνηση, να την ονοματίζει με τη φωνή του και να αποθηκεύεται στη βάση με το όνομα αυτό. Κάτι τέτοιο βέβαια απαιτεί τουλάχιστον πολύ μεγάλη ευαισθησία μικροφώνου και ικανότητα ελαχιστοποίησης του θορύβου από το περιβάλλον στο οποίο εκτελείται η εφαρμογή, καθώς επίσης και αναγνώριση λέξεων που ίσως δεν υπάρχουν ορισμένες στο σύστημα.

ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΑΣ

Ξενόγλωσσος όρος	Ελληνικός όρος
gesture	χειρονομία
scanner	σαρωτής
gesture recognition	αναγνώριση χειρονομίας
facial recognition	αναγνώριση προσώπου
speech recognition	αναγνώριση ομιλίας
tracking	ιχνηλατώντας
depth map	χάρτης βάθους
joints	κλειδώσεις
tracking modes	λειτουργίες ιχνηλασίας
gait recognition	αναγνώριση βαδίσματος
sign language	νοηματική γλώσσα
frames	πλαίσια
skeleton tracking	ιχνηλασία σκελετού
textbox	κουτί κειμένου
generic	γενικός

ΣΥΝΤΜΗΣΕΙΣ – ΑΡΚΤΙΚΟΛΕΞΑ – ΑΚΡΩΝΥΜΙΑ

RGB	Red Blue Green
CMOS	Complementary metal–oxide–semiconductor
XAML	Extensible Application Markup Language
XML	EXtensible Markup Language
SDK	Software Development Kit
WPF	Windows Presentation Foundation
DTW	Dynamic Time Warping

ΠΑΡΑΡΤΗΜΑ

Απαιτήσεις συστήματος

Ο χρήστης για να τρέξει την εφαρμογή θα πρέπει να διαθέτει τα εξής:

- α)** Λειτουργικό σύστημα: Windows 7, Windows 8 x32/x64
- β)** Επεξεργαστή Intel Core i3 CPU 2.40 GHz.
- γ)** Μνήμη RAM 4Gb.
- δ)** Εγκατεστημένο το Microsoft Kinect for Windows SDK.
- ε)** Εγκατεστημένο το MySQL Workbench.

ΑΝΑΦΟΡΕΣ

- [1] "Kinect", Internet: <https://en.wikipedia.org/wiki/Kinect> [Jul 5 2015].
- [2] "Gesture recognition",
Internet: https://en.wikipedia.org/wiki/Gesture_recognition [Jul 5 2015]
- [3] "Skeletal tracking",
Internet: <https://msdn.microsoft.com/en-us/library/hh973074.aspx>
[Jul 5 2015]
- [4] "JointType Enumeration", Internet:
<https://msdn.microsoft.com/en-us/library/microsoft.kinect.jointtype.aspx>
[Jul 5 2015].
- [5] "Tracking Modes (Seated and Default)",
Internet: <https://msdn.microsoft.com/en-us/library/hh973077.aspx>
[Jul 5 2015]
- [6] "Gait recognition", Internet: http://globalseci.com/?page_id=44
[Aug 22 2015]
- [7] "Kinect Sign Language Translator expands communication possibilities", Internet:
<http://research.microsoft.com/en-us/collaboration/stories/kinect-sign-language-translator.aspx> [Aug 22 2015]
- [8] G. J. Garcia, P. Gil, D. Llacer, F. Torres, "Guidance of Robot Arms using Depth Data from RGB-D Camera", Physics, Systems Engineering and Signal Theory Department, University of Alicante, San Vicente del Raspeig, Spain, 2010
- [9] Erdenetsogt Davaasambuu , Chia-Chi Chiang , John Y.Chiang , Yung-Fu Chen , Sukhbaatar Bilgee, "A Microsoft Kinect based virtual rehabilitation system", Department of Healthcare Administration, Central Taiwan University of Science and Technology, Taichung 40601, Taiwan, Department of Computer Science and Engineering, National Sun Yat-sen University, Kaohsiung 80424, Taiwan, School of Information and Telecommunication Technology, Mongolian University of Science and Technology, Ulaanbaatar, Mongolia, 2012
- [10] Alena KOPANIČÁKOVÁ, Mária VIRČÍKOVÁ, "Gesture Recognition using DTW and its Application Potential in Human-Centered Robotics", Dept. of Cybernetics and Artificial Intelligence, FEI TU of Košice, Slovak Republic, 2013
- [11] Lu Xia, Chia-Chih Chen, and J. K. Aggarwal, "View Invariant Human Action Recognition Using Histograms of 3D Joints", Computer & Vision Research Center / Department of ECE
The University of Texas at Austin, 2011
- [12] "Speech", Internet:
<https://msdn.microsoft.com/en-us/library/jj131034.aspx> [Aug 25 2015]
- [13] "C Sharp (programming language)", Internet:
[https://en.wikipedia.org/wiki/C_Sharp_\(programming_language\)](https://en.wikipedia.org/wiki/C_Sharp_(programming_language))
[Aug 22 2015]
- [14] "Extensible Application Markup Language", Internet:
https://en.wikipedia.org/wiki/Extensible_Application_Markup_Language
[Aug 22 2015]
- [15] "Database Workbench", Internet:
https://en.wikipedia.org/wiki/Database_Workbench [Jul 10 2015]

[16] G.A. ten Holt, M.J.T. Reinders E.A. Hendriks, “Multi-Dimensional Dynamic Time Warping for Gesture Recognition”, Information and Communication Theory Group Delft University of Technology, Mekelweg 4, 2628 CD, Delft, The Netherlands

[17] “ARABESQUE”, Internet:
<http://aballeteducation.com/2014/10/19/arabesque/> [Sep 1 2015]

[18] “The Dance Training Project”, Internet:
<http://danceproject.ca/your-hips-dont-lie/#.Vj4Y39LhCUk> [Sep 1 2015]

[19] “Παραδοσιακοί Χοροί”, Internet:
http://xoroi-ths-paradoshs.blogspot.gr/2012/01/blog-post_4479.html
[Sep 1 2015]

[20] Johannes Preis, Moritz Kessel, Martin Werner, Claudia Linnhoff-Popien,
“Gait Recognition with Kinect”,
Ludwig-Maximilians University Munich, Germany

[21] Bayes’ Theorem, Internet:
https://en.wikipedia.org/wiki/Bayes%27_theorem [Nov 4 2015]