Multidisciplinary Research Academic Journal (MDRAJ) ISSN : l-2467-4699/ e-2467-4834, Volume 3, Issue 1 (June 2018), pp.48-66, mdraj.openlu.org

KINECT DYNAMIC BEHAVIOR CLASSIFICATION MODEL FROM BODY GESTURES

Osée Muhindo Masivi[1]

ABSTRACT

This research proposes to determine the generic multimodal behavior model from body gestures sensed by Microsoft Kinect Technology. The model encompasses detecting, recognizing, discovering, learning, predicting and measuring human emotional behavior patterns. The skeletal data and voice signals from Kinect are used specific to recognize and group gestures and interjections depending on specified environmental context. Rule-based matching and Dynamic Time Warping alignment have been used respectively for static and dynamic body gestures. The systems established were the real time, independency from camera view point, fully automated, easy to train, and adaptable. Tests for gesture classifier revealed that the system performance was with precision at 94.67%, sensitivity 88.89%, specificity 99.24%, efficiency 92.91%, positive prediction accuracy 98.91%, reliability 94.40% and overall accuracy of 88.60%. The deterministic behavior model classified positively recognized the gesture patterns present in the database. The abundance of output data for this study make it recommendable for system in research, education, physical exercises training, and entertainment games.

Key words: Dynamic gesture behavior model, 3D gesture model classifier, Kinect Model.

INTRODUCTION

Since time immemorial, one of the most common humans' concerns in communication and other activities has been the problem of inferring the mental states of their interlocutors. But this challenge (behavior understanding) is hard due to the instantaneous correlation between involved signals evaluated 3,000 times faster than rational thought (Tang, et al., 2012). Yet, behavior analysis models remain practically important in solving many humans and animals' problems in medicine, security, epidemiology, conservation/protection monitoring, and many other applications.

These applications are so highly emotional that emotion behaviors rise in various contexts. For instance, happiness us shown after watching a great movie, completed a difficult task or preferred team success. In each context, the same emotion is generally displayed through different expressions. Face expressions have been intensively used to discriminate those expressions in near-distance face recognition. But, behavior classifiers

[1] Dr. Osee Muhindo Masivi, PhD. Is an Associate Professor at Institut Supérieur Pédagique de Muhangi à Butembo- République Démocratique du Congo, osmasivi@gmail.com

based on face recognition at a distance (FRAD) are affected by camera distance problems: resolution, focus, interlaces effect, and motion blur (Ao, et al., 2009; Phillips, et al., 2005).

For these reasons, multimodal behaviors patterns from body gestures constitute alternative clues for behavior understanding models. Their advantage is that they can be seen and heard from a relatively long distance by reasonably cheap sensors. This article focused on designing computational model automatizing the human emotion inference from gesture patterns sensed by Kinect technology.

METHODOLOGY

As shown in Figure 1, this research explored emotion behavior classification models in three steps:

- Behavior "modal action pattern" correlation: emotions gesture patterning and patterns structuring.
- Gesture temporal correlations: body gesture modeling from Kinect output signals and are linked as time signal patterns.
- Signal alignment: performer's incoming gesture alignment with known signals and patterns.

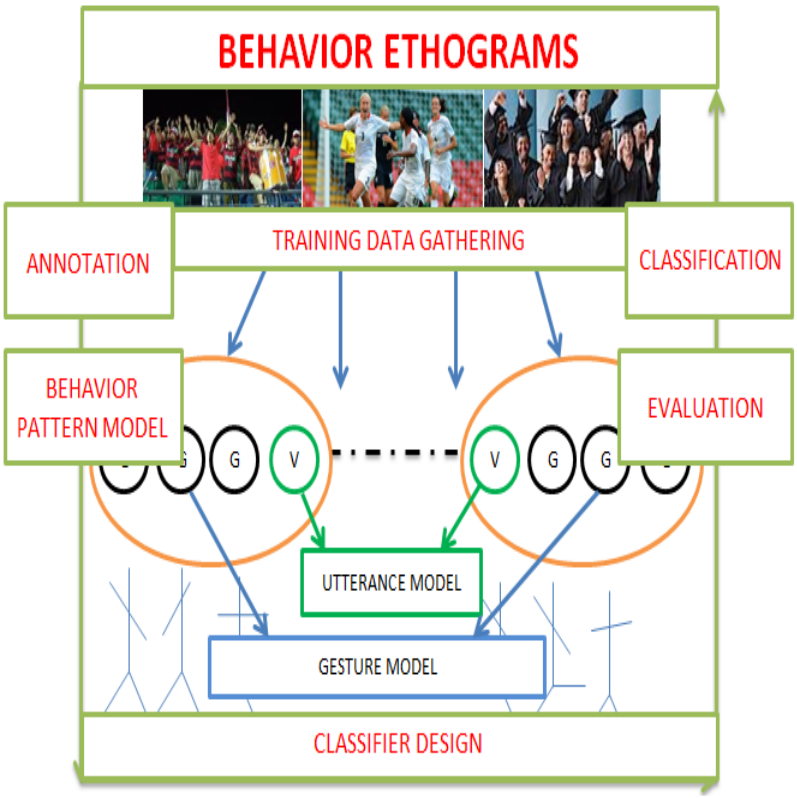


Figure1. Dynamic behavior classification framework

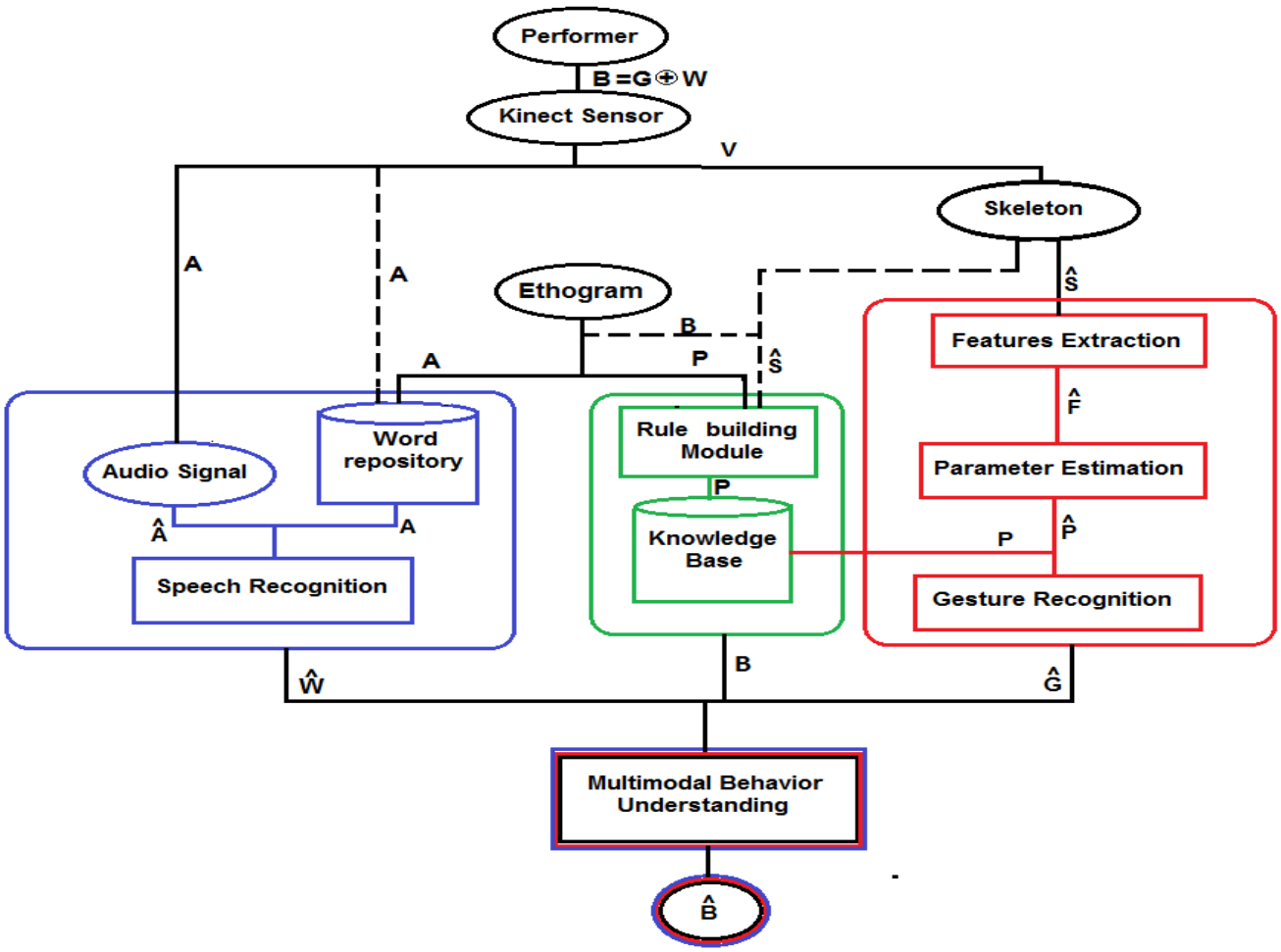


Figure 2. Overview of the proposed multimodal behavior models used in the research: in green the Knowledge base construction, in bleu the speech recognition, in red the gesture recognition and in bleu-red the multimodal behavior classification. The input/output data are B=Behavior displayed, G=gesture, W=Utterance, V=visual images, A=audio signal, $\hat{S}$ = user skeleton inferred from visual images V, P=gesture paramerers, $\hat{F}$= estimated features, $\hat{P}$=estimated parameters, $\hat{A}$=estimated audio signal, $\hat{W}$=estimated interjection, $\hat{G}$=estimated gesture and $\hat{B}$=estimated behavior.

Ethological approach (Klein, 2000) was used in the first stage for two objectives. The first was to identify and statistically analyze correlations between gesture-contexts patterns and the six universal emotions (happiness, surprise, sadness, disgust, fear and anger) and build a multimodal behavior model. The second is to find the biological structure of identified gestures and get their gesture kinematic models ready for computational modeling. Speech recognition was not considered in this article.

Computational modeling approach was used in the two last stages to convert the obtained ethological models into computational models composed of data structures and algorithms. The computational models with constitute the behavior classifier modules (Figure 2) were trained for classification. The classifier was evaluated in conformity to the annotated data.

a b c d e

Figure 3. a, b and c show "hand up" gesture with respectively disgust, neutral and happy emotion; d and e show"hand on hip" gesture with happy and angry emotion (George & Robert D., 1988).

EMOTION ETHOLOGICAL MODEL

An emotion is an episode of interrelated and synchronized changes in cognitive processing, subjective feeling, action tendencies, physiological changes, and motor expressions as response to an event of major significance to the organism (Kipp & Martin, 2009). Because the research focuses on behavior structure, motor behavior expressions have been used. In order to correlate emotions with body gesture patterns, we observed how characters express their emotions by body gestures from the movie "Coming to America". The film was chosen due its emotional richness. Observation results showed that a gesture pattern can convey two or more emotions in different environment contexts such as sport, achievement, relationship, etc. (Figure 2). This situation was partly explained by the fact that emotions don't have lives of their own, but mutually influence and overlap each other across time (Pe & Kuppens, 2012). Table 1 and Figure 3 show the percentage (for gesture-context pattern) of observed body gestuesto convey universal emotions where context is "Relationships".

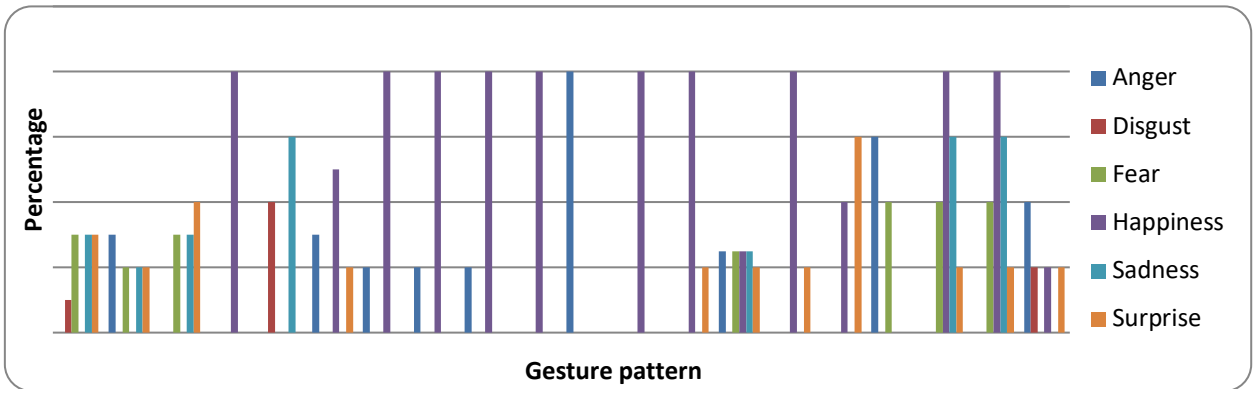


Figure 4. Gesture pattern-Emotion correlation in Relationship context

Table 1:

Percentage (confidence) for 20 Gesture Patter in universal emotions for computed from "Coming to America" movie

	Considered emotions Gesture Pattern	Happy	Surprise	Sad	Disgust	Fear	Anger
00	Nothing (Neutral)	0	0	0	0	0	0
11	One hand on head	0	10	30	0	30	30
12	Two hands on the head	30	0	20	0	20	20
13	Two hands together	0	0	30	0	30	40
14	Two hands departed and raised	0	0	0	80	0	0
15	Crossed arms	0	40	0	0	60	0
16	One arm extended in front	30	0	0	50	0	20
17	Two arms extended in front	20	0	0	80	0	0
18	Extend one fore arm in front	20	0	0	80	0	0
19	Wave one arm	20	0	0	80	0	0
29	Hands lifted up	0	0	0	80	0	0
21	Fore arms punch down	80	0	0	0	0	0
22	Fore arms punch up	0	0	0	80	0	0
23	Two hands meet	0	0	0	80	0	20
24	Motion with joined hands	25	0	25	25	25	20
25	Arms extended laterally	0	0	0	80	0	20
26	Mix extend hand and touch head	0	0	0	40	0	60
27	Hands departed at hip level	60	0	40	0	0	0
28	Two hands on opposed shoulders	0	0	40	0	49	20
29	Two hands on opposed shoulders	0	0	40	0	40	20
30	Two hands joined with hips	40	20	0	20	0	20

Emotion computational model

Let E, G, U and C be respectively the set of emotions, gestures and environment contexts. We denote:

- The user's overlapped emotional states by the set $\left(E_{s_i}^{p}\right)^{\#S}$ with S={happy, surprise, sad, disgust, fear, anger, …}, p ∈ [0,100] the confidence value, $\left(E_{s_i}^{0}\right)^{\#S}$ the neutral emotional state and $\left(E_{s_i}^{100}\right)^{\#S}$ the presence of all emotion states.
- The user's gesture multimodal emotion pattern by g-u ∈G×U.
- The multimodal emotion inference function f(g-u,c)= $\left(E_{s_i}^{c}\right)^{\#S}$ with c∈C, g∈G and u∈U.

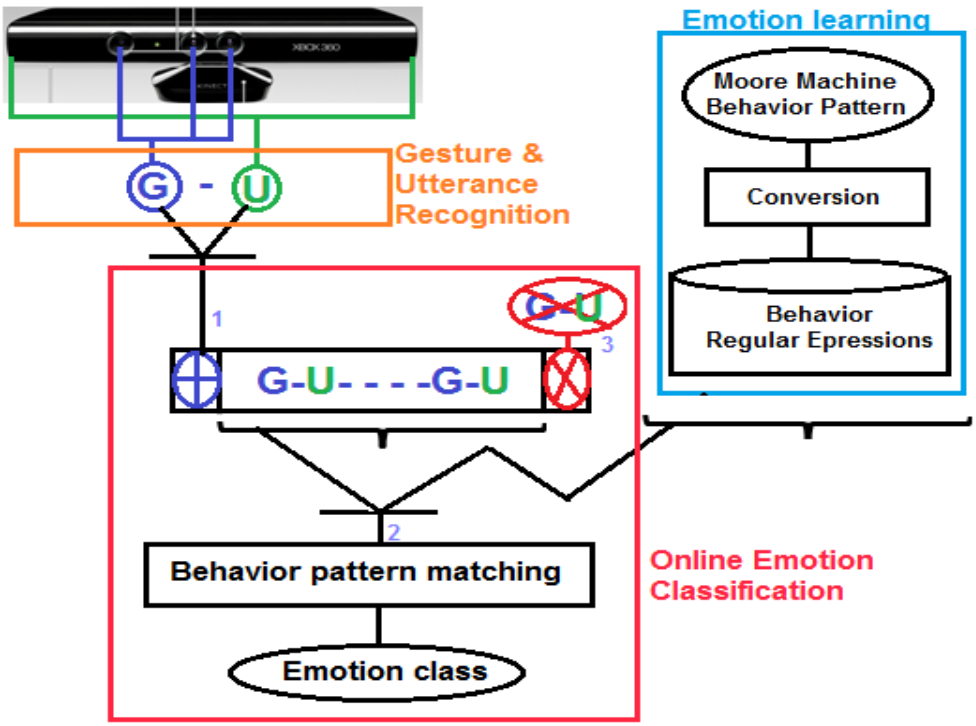


Figure 5. *Emotion behavior classifier framework*

Based on this mathematical emotion models, the classifier framework in Figure 5formalizes the gesture pattern by the Moore Machine in Figure 6 (Albacea E. A., 2006) which, after conversion (Algorithm 1) feeds the emotion classifier database. The emotion classifier (in red on Figure 5), which depends on gesture recognition iterates as follows (Algorithm 2):

- upon recognition in the frame, push gesture in the behavior link list,

- if the link list is full, match link list content (gesture pattern) with database regular expressions,
- upon idle time expiration or new incoming gesture, pop gesture from the list.

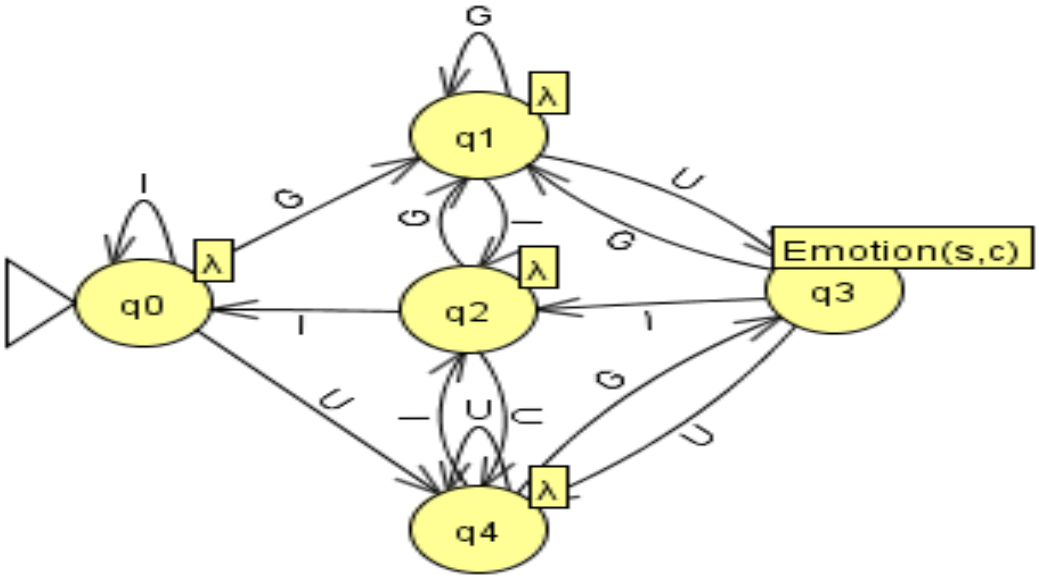


Figure 6. Gesture Moore Machine

Algorithm 1. DFAtoRE(MooreMachineStatTransList)

```
Foreach rule in gestureList Do
    If (rule.DeStateType = Final)
        final.Add(rule.DestState)
    If (rule.OrigineType = Start)
        start.Add(rule.OrigineState)
End
If (number_of_StartState n >1) Then
    Add S=Ai|…An to transitionFxList
    Foreach Ai in gestureList Do
        For aj to B  do
            If (equation = empty) Then equation=ajB
            Else          equation=equation + | + ajB
            Endif
            add A = equation to transitionFxList
        End
        Foreach final Ai in listFinalStates Do
            Add Ai=“ ” to transitionFxList
        End
    End
    Foreach equation in transitionFxlist Do
        If(origineEquat in equation) then
            origineEquat = ‘*’
```

```
    End
    For i= number_equationDownTO 1 Do
        s= origineEquat
        For j= i-1 TO 1 STEP -1 Do
                origineEquat (j)=equation(i)
        Next j
    Next i
        GestureList = equation
```

Algorithm 1.Brzozowski Algebraic method adaped from **(Neuman, 2005)**.

Algorithm 2.BehaviorMatching (BehDd, G/U, PatternLenght)

```
Patterm.Push (G/U)
If (PatternLength=patternQueue.Length)
For (i = 0ToBehDd.Count)
If (BehDd [i].Regexpression==Pattern
        Return BehDd [i] Descriptions and confidence
        ExitFor
      EndIf
    End
EndIf
```

Algorithm2. Multimodal Online Emotion ClassificationPseudo code

Gesture ethological model

A gesture is any limb(s) motion(s) that conveys information (Kramer, et al., 2012; Webb & Ashley, 2012). These limb motion(s) can be seen differently depending on sensor's view plan. Sensors restricted to two planes (2D) lose motion details from third view point. In 3D, human gestures are ethologically described based on three planes (McGinnis, 2005) of motion passing through the human body (sagittal, frontal and transverse or horizontal plane) in nine degrees of freedom (Figure 7). However, for simplicity in computer modeling, there is a tendency to refer to the motion in one particular plane whose motions are dominant. Hence, body joint motions can be measured in reference to one or more dominant planes (axis) and body degree of freedom.

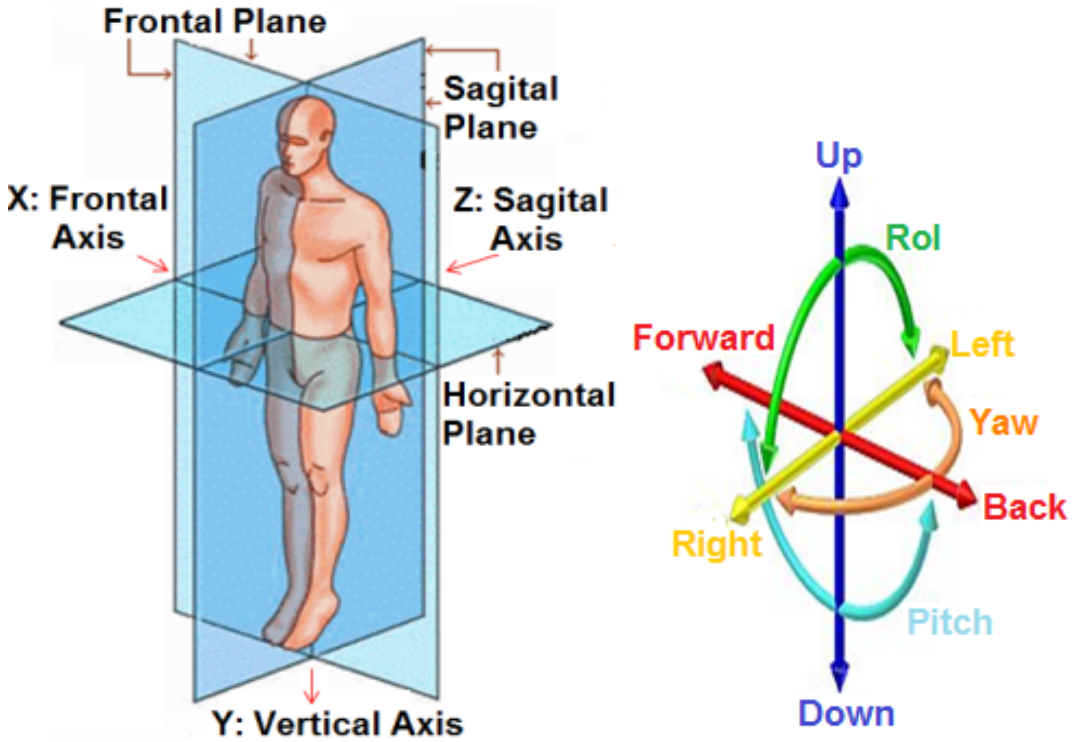


Figure 7. Motion planes, axis and degree of freedom

In this work, body gesture is defined by skeleton joint positions during (dynamic gesture) and after (static gesture) angular movements. Angular movements are based on the increase or decrease of the angle between 2 bones joined by a joint. McGinnis (2005), defined five principal 3D movements: flexion, extension, adduction, abduction and **Circumduction** (Figure 8).

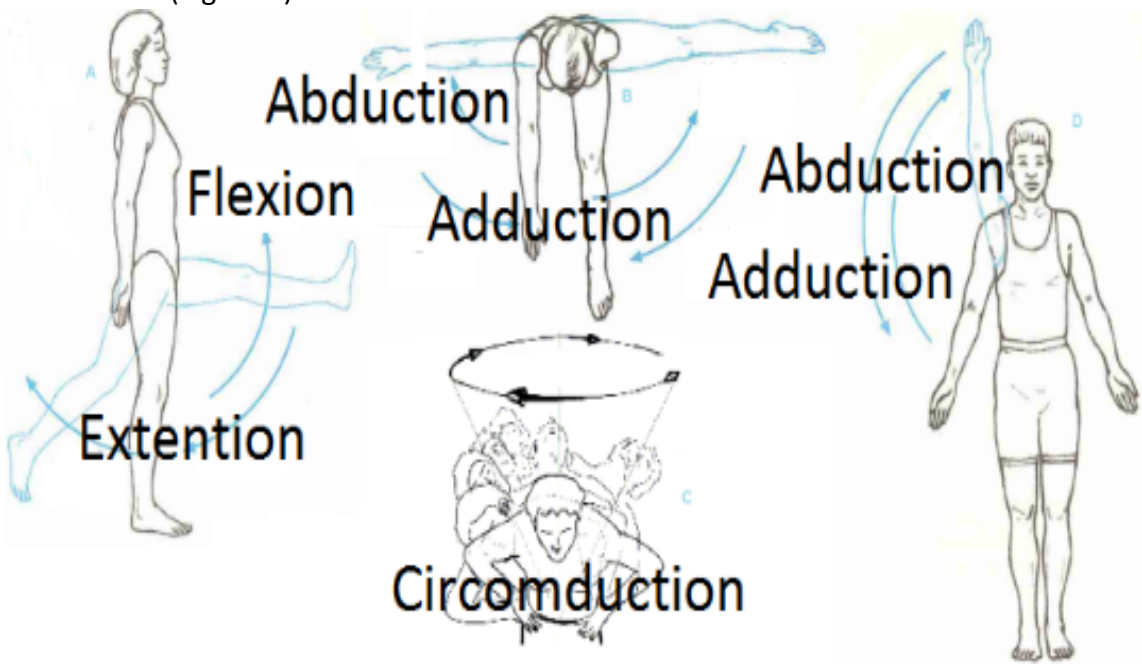


Figure 8. Some 3D dominant movements

Gesture computational model

Two approaches can be used to infer body gestures: appearance model which infer gestures directly from visual images and 3D-based approaches which transit by the motion reconstruction before the inference. This research applied 3D skeletal models for static and dynamic gestures. It reconstructs the motion using Kinect SDK skeleton data (Equation 2) returned as a set of 20 points (Microsoft Research, Kinect™ for Windows® Programming Guide: Getting Started with the Kinect for Windows SDK Beta from Microsoft Research, 2011).

$$M = T_{mg}G \quad (1) \qquad \hat{G} = T_{vg}^{-1}V \quad (4)$$
$$V = T_{vm}M \quad (2) \qquad \hat{H} = T_{vm}^{-1} \quad (5)$$
$$V = T_{vm}(T_{mg}G) = T_{vg}G \quad (3) \qquad \hat{G} = T_{mg}^{-1} \quad (6)$$

Equation 1. Gesture recognition equations

```
_NUI_SKELETON_DATA (eTrackingState, dwTrackingID, dwEnrollmentIndex,
    dwUserIndex, Position, SkeletonPositions[NUI_SKELETON_POSITION_INDEX],
    eSkeletonPositionTrackingState [NUI_SKELETON_POSITION_INDEX],
    dwQualityFlags).
```

Equation 2. Kinect SDK skeleton output data structure

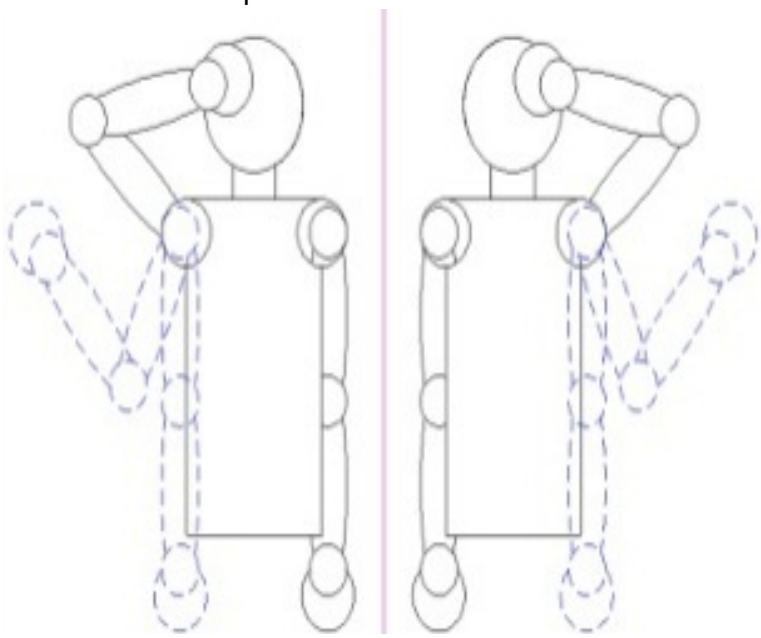


Figure 9. Hand on hand gesture representation

Static gestures (postures) model

A posture is defined as relative positions of body limbs at a given moment. In 3D, this positions (less, greater, equal) are found by comparing X, Y and Z coordinate values of target and referred limbs in dominant movement planes (inner part of Equation 3). As shown in

Figure 9, a posture can be represented by m different set of positions (outer part of Equation 3).

Let Ps (ps_1, ps_2, … ps_n) be the set of n limb positions. Let us call elementary rules (ElmtRule) each ps_i in Ps. Note that P can have m different set of limb positions Ps_1 to Ps_m. Each posture case is a clause (ps_1^ps_2^…^ ps_n). The posture is a logical statement for Ps_m clauses joined by Or (Equation 3). For instance, the hand on head posture (Figure 9) can be represented by the rule matrix in Equation4.

$$\text{Posture} = \underset{k=1}{\overset{m}{\text{OR}}} \left[\underset{l=1}{\overset{n}{\text{AND}}} \left[\text{ElmtRule}(k,l) \right] \right]$$

Equation 3.Posture equation : One posture can have m different descriptions linked by OR and description can be defined by a combination of n rules linked by AND where
ElmtRule=(PostureID,RuleID,SourceJoint,SourceCoordinate, ComparisonOperator,TargetJoint,TargetCoordinate,RuleScore)

Lhand.X=Head.X	OR	Rhand.X=Head.X
Lhand.Y=Head.Y		Rhand.Y=Head.Y
Lhand.Z=Head.Z		Rhand.Z=Head.Z

Equation 4. Posture sample Rule Matrix

Static gestures submission is automatically done by the algorithm 3 and their recognition (matching) is done by algorithm 4.

Algorithm 3. PostureSubmit(Cordinate, Joints_Source, Joints_Target)

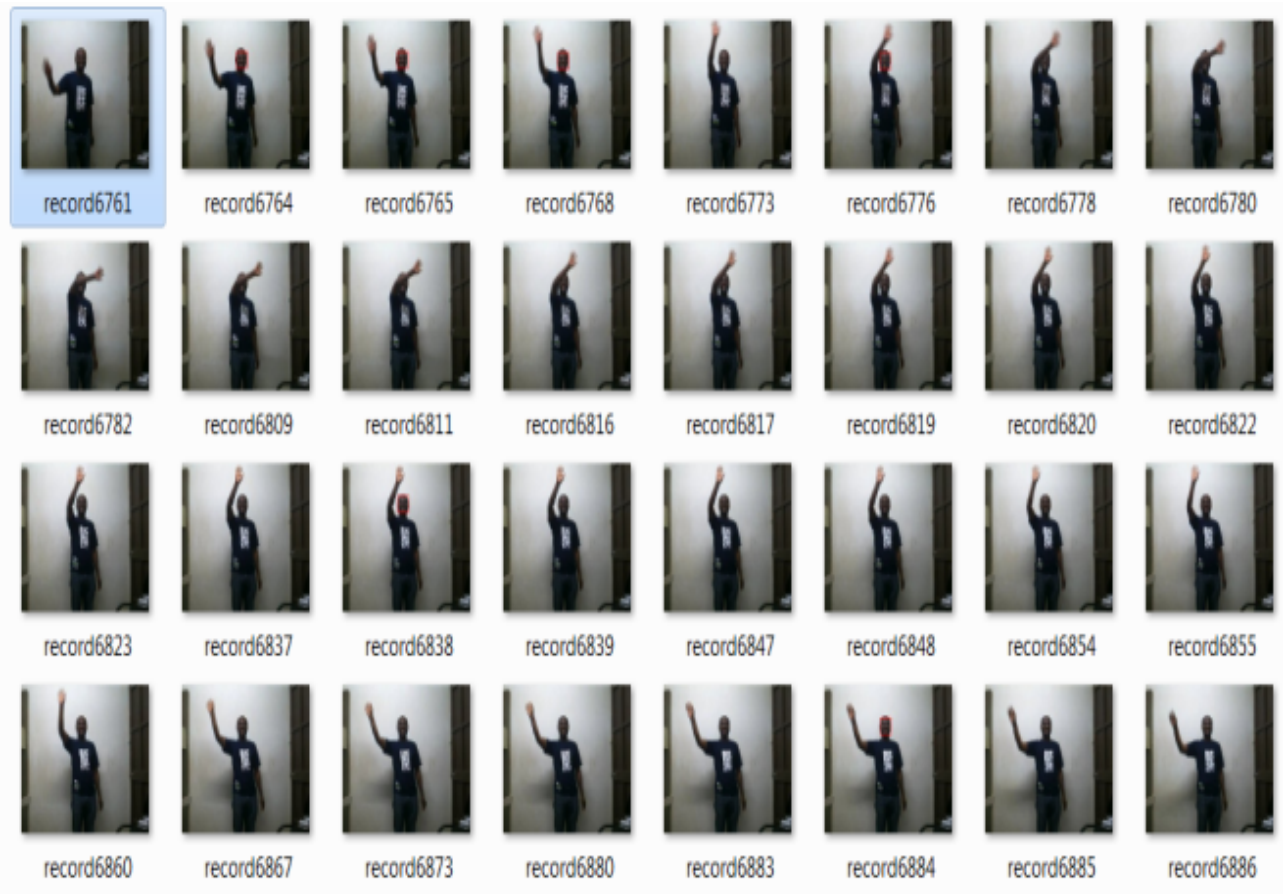
```
Foreach tracked JointS in Joints_Source Do
    ForeachJointT in Joints_Target Do
        Foreachcoord in Cordinate_Involved Do
            If (JointS.coord>JointT)Then
                Add new rule(Operator = Bigger)
Else if (JointS.coord<JointT)
        Add new rule (Operator = Smaller)
        Else
                Add new rule (Operator = Equal)
            EndIf
        EndFor
    EndFor
Endfor
```

Algo.3. Pseudo code for Posture Automatic rule submission algorithm

Algorithm 4 PostureRecognizer(Skeleton, RulesMatrix)

```
Load rules in RulesMatrix, Return if no rules
For each Posture in RulesMatrix Do
  For each rule in Posture Do
    Intensity = |sv – tv|
    If (rule.Operator = Bigger &&sv>tv + var)
            Score = Score + rule.Score;
    Else If (rule.Operator = Smaller &&sv<tv - var)
            Score = Score + rule.Score;
    Else If (rule.Operator = Equal && (|sv-tv|-var<=0)
            Score = Score + rule.Score;
    Endif
  End For
    If (Score>=Goal) Then
        Return (Posture, Goal)
        Store(FrameRGB)
End for
```

Algo.4. Pseudo code for Posture Recognition Algorithm



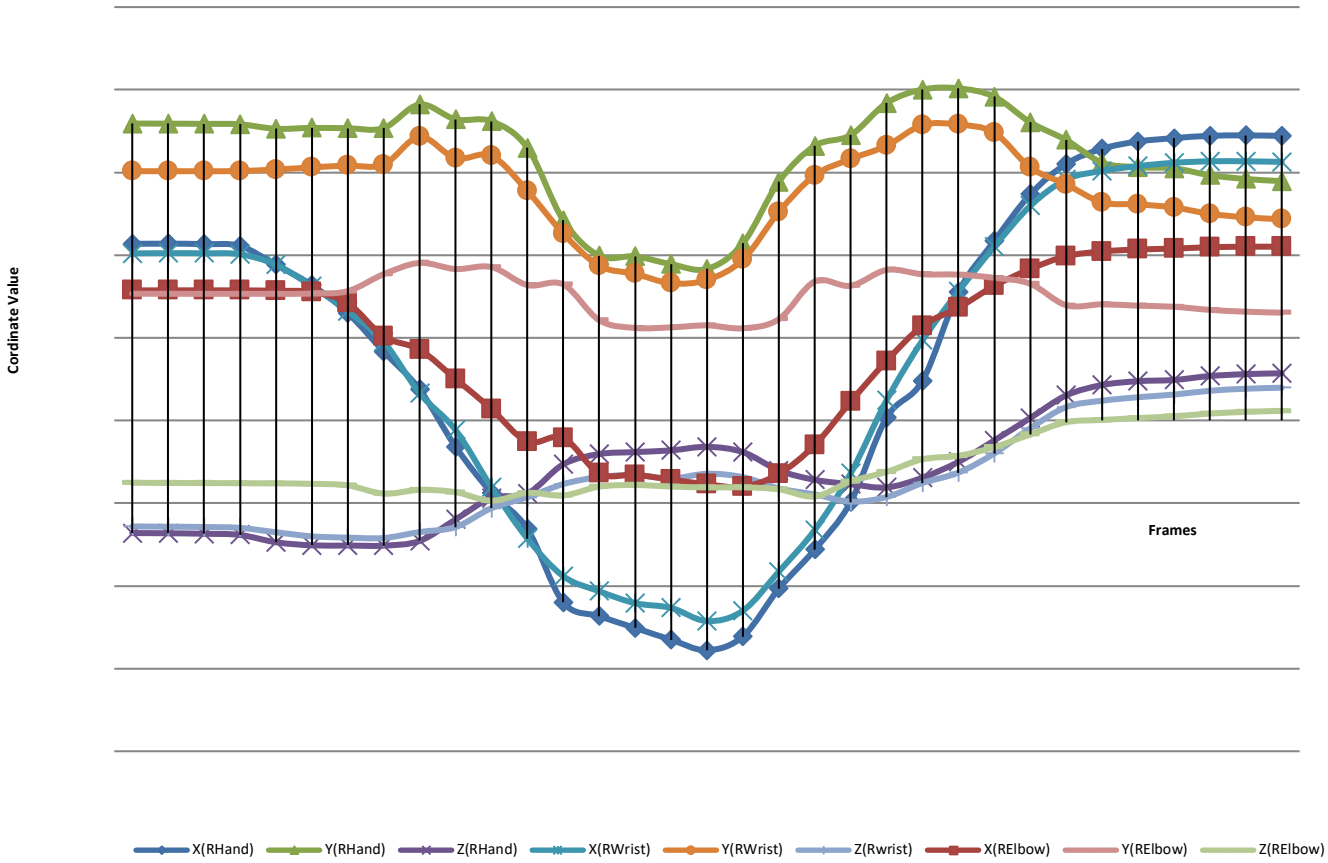


Figure 10. 32 frames from Wave Right hand gestures (left) and Trajectory pattern graphic (right)

Dynamic gestures

Dynamic gesture is modeled by the 3D trajectory patterns made by X, Y and Z coordinates of targeted body limbs in motion. For instance, the “wave right hand” gesture involves right hand, right wrist and right elbow limbs. Figure 10 shows the 33-frame captured from the sensor (left) and the 3D trajectory pattern for the three limbs. Any movement in which these three limbs display a trajectory pattern similar to the one in Figure 10 could be classified as “wave right hand” gesture. The alignment and recognition is done by algorithm 6.

Algorithm 5DTWDistance(InputSequence[InputLength], SampleGesture[SampleLength], maxSlope, bestMatch)

```
verticalStep[i,j]=0
horizontalStep[i,j]=0
For i=1 to InputLength Do
For j=1 to SampleLength Do
    If (DistanceMatrix(i,j-1)<DistanceMatrix(i-1,j-1) &&
      DistanceMatrix(i,j-1)<DistanceMatrix(i-1,j) &&
      horizontalStep[I,j-1]<maxSlope)
      DistanceMatrix(i,j)=EuclidianDistance(InputSequence(i-1),
              SampleGesture(j-1) + DistanceMatrix(i,j-1)
              verticalStep[i,j]= horizontalStep[i,j-1]
              horizontalStep[i,j]= verticalStep[i,j-1]+1
    Else If (DistanceMatrix(i-1,j)<DistanceMatrix(i-1,j-1)
    DistanceMatrix(i,j)=EuclidianDistance(InputSequence(i-1),
              SampleGesture(j-1)+ DistanceMatrix(i-1,j)
              verticalStep[i,j]= horizontalStep[i-1,j]+1
              horizontalStep[i,j]= verticalStep[i-1,j-1]
    Else DistanceMatrix(i,j)=EuclidianDistance(InputSequence(i-1),
              SampleGesture(j-1) + DistanceMatrix(i,j-1)
              verticalStep[i,j]= 0
              horizontalStep[i,j]= 0
    End If
  End For
End For
For i =1 to InputLength Do
        If (DistanceMatrix(I, SampleLength)<bestMatch) Then
              bestMatch= DistanceMatrix(I, SampleLength)
        End If
End For
Return bestMatch
```

Algo.5. Pseudo code for DTW distance computation

This algorithm implements a novel $O(n^2)$ version of Dynamic Time Warping applied to gesture recognition. It was from adapted from Muscillo, et al. (2001), Tormene, et al. (2009) and Vivaracho-Pascuala, et al. (2009) and runs as follows:

- Offline database 3D gesture signals construction,
- Input signal capturing frame by frame,
- Input signal and referred signals (from database) alignment, signal distance thresholding, and nearest offset gesture selection.

RESULTS EXPERIMENTAL EVALUATION

An application was implanted from the above models and algorithms for experimental evaluation. The system was tested by more than 200 volunteers. These tests showed that all regular expressions found the database (File 1) where successfully recognized (Figure 11). However, the behavior classifier depends on gesture recognition shown in confusion matrix (Table 3): average accuracy of 85%, average reliability of 94% and overall accuracy of 84%.

```
HAP@Happiness@1200@80
HAP@Happiness@1210@100
HAP@Surprise@1201@0
HAP@Nothing@0000@0
HAP@Happiness@0101@0
SUR@Surprise@1200@50
```

File 1. Emotion pattern Database with four fields separated by @: emotion code, emotion description, gesture-context pattern and confidence percentage.

Multimodal behavior classifiers resulted from this research have not rendered outputs of the type yes of not. They have outputted not only the classified gesture, interjection and behavior, but also have answered the six behavioral proximal and ultimate causation questions: what, who, when, where, how and why (Lehner, 1996).

These questions were solved in the system many results outputs files and screen forms. The files are automatically stored in the application directory ender appropriate folders created instantaneously on need. Among all other files, the files “Results”, “Performed Gesture”, “Emotion Classification” and “Behavior Results” can be cited.

Major information in these files are the posture description (what), the matrix of rules involved and the intensity of the posture (how), and the time of exhibition (when). The intensity of a posture is the distance (meters) or the angle (in degrees) formed by the involved joints. In addition, information in gives answer to the question what happened in the scene for a given period of time.

Who and where questions were answered by sample pictures showing details on the subject and the conditions in which the gesture was performed (surroundings, other limbs pose). Upon recognition, pictures relating to the gesture are automatically stored in the corresponding folder. Each picture’s name gives the gesture name followed by the date and time (when). The answer to the why question includes the proximate and the ultimate mechanisms of the gesture and interjection. The “Behavior Results” answered the why question by inferring the emotion state behind the gesture-interjection pattern. Furthermore, the environmental context added in the model gives more details to the answer of the why question.

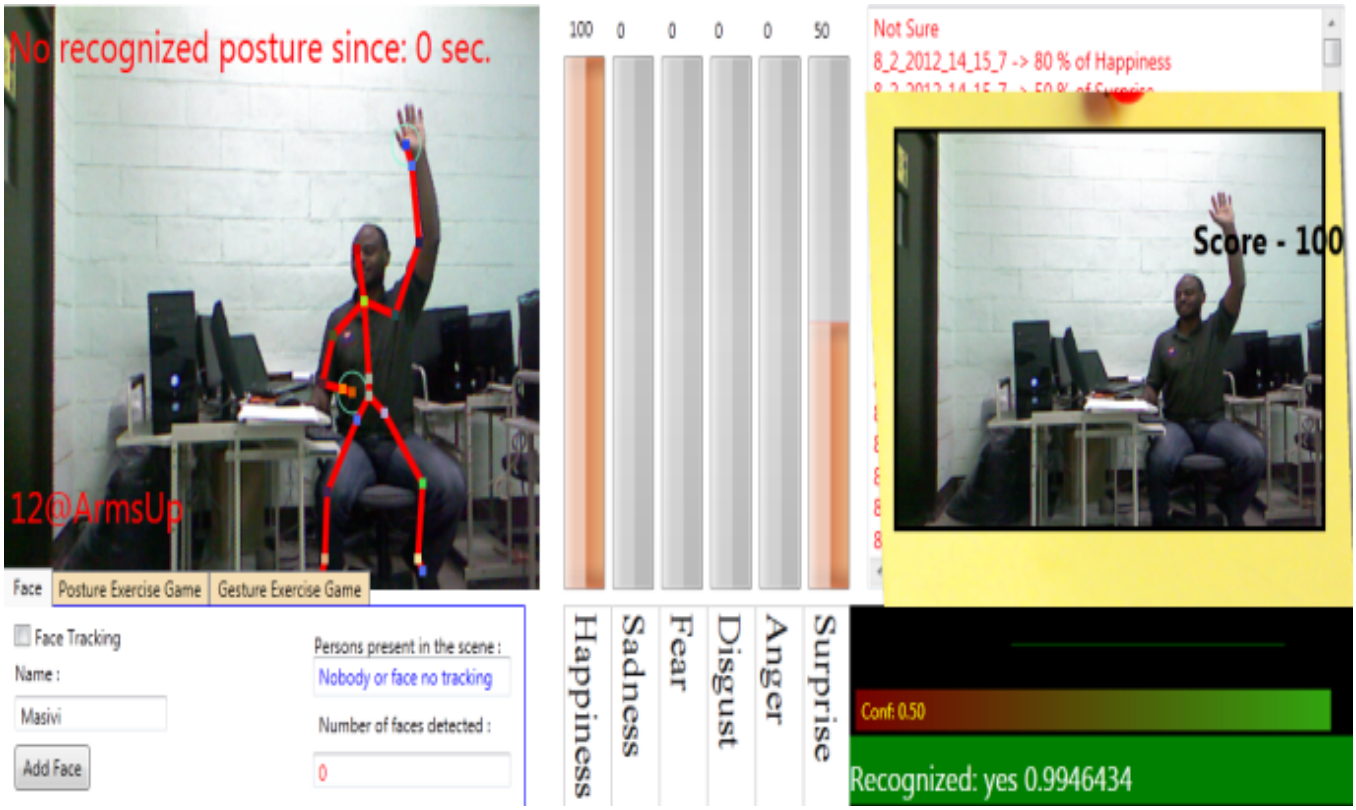


Figure 11. Behavior real time classification results from a volunteer preforming the hand rise gesture

Table 3.
Indicators of performance

Code	Precision	Sensitivity	FA	Specificity	Efficiency	Accuracy	FN rate
00-Idle	0.9986	1.0000	0.0859	0.9141	0.9571	0.9986	0.0000
11-One hand(Left or Right) on head	0.9000	0.8182	0.0078	0.9922	0.9052	0.9784	0.2222
12-Two hands on the head	0.9375	1.0000	0.0060	0.9940	0.9970	0.9945	0.0000
13-Two hands are put together	0.9412	0.9143	0.0039	0.9961	0.9552	0.9909	0.0938
14-Two hands departed and raised	0.9259	0.8621	0.0038	0.9962	0.9291	0.9891	0.1600
15-Crossed arms	0.9667	0.8529	0.0019	0.9981	0.9255	0.9891	0.1724
16-One arm extended in front the body	0.9545	0.8571	0.0040	0.9960	0.9266	0.9837	0.1667
17-Two arms extended in front the body	0.9375	0.7500	0.0039	0.9961	0.8731	0.9784	0.3333

18-Extend one fore arm in front	0.9189	0.8500	0.0059	0.9941	0.9221	0.9837	0.1765
19-Wave one arm	0.9388	0.8364	0.0060	0.9940	0.9152	0.9784	0.1957
20-Hands lifted up upward over the head	0.7778	0.8333	0.0193	0.9807	0.9070	0.9696	0.2000
21-One arms extended upward over the head	0.9167	0.9167	0.0038	0.9962	0.9564	0.9927	0.0909
22-One arms extended upward below the head	0.9259	0.9259	0.0038	0.9962	0.9610	0.9927	0.0800
23-Two hands meet	0.9211	0.9459	0.0059	0.9941	0.9700	0.9909	0.0571
24-Motion with joined hands	1.0000	0.8000	0.0000	1.0000	0.9000	0.9945	0.2500
25-Ams extended laterally	1.0000	0.9091	0.0000	1.0000	0.9545	0.9963	0.1000
26-Mix extend hand and touch head	0.9200	0.7931	0.0038	0.9962	0.8946	0.9855	0.2609
27-Hands departed at hip level	1.0000	0.9615	0.0000	1.0000	0.9808	0.9982	0.0400
28-One hand joined with opposed shoulder	1.0000	0.7500	0.0000	1.0000	0.8750	0.9927	0.3333
29-Two hands on opposed shoulders	1.0000	0.7143	0.0000	1.0000	0.8571	0.9891	0.4000
30-Hands departed at hip level	1.0000	0.8889	0.0000	1.0000	0.9444	0.9945	0.1250
AVG	0.9467	0.8657	0.0079	0.9921	0.9289	0.9886	0.1647

CONCLUSION AND DISCUSSION

The Behavior classifier proposed in this research was very simple and generic. General multimodal gesture behavior models based on Kinect output signals was designed, implemented and applied to emotion classification. Algorithms for gesture recognition by aligning targeted joints signals regardless positions and tracking condition of other joints using dynamic time warping were supported and aligned with other studies as indicated Muscillo, et al. 2001; Tormene, et al. 2009; Vivaracho-Pascuala, et al. 2009, and Webb & Ashley, 2012.

It was found that some vocal interjections lead to an acceptable unambiguous emotion classification. The skeletal data and voice signals from Kinect are used specific to recognize and group gestures and interjections depending on specified environmental context. Rule-based matching and Dynamic Time Warping alignment have been used respectively for static and dynamic body gestures. The system implanted is real time, independent from camera view point, fully automated, easy to train, and adaptable. Tests for gesture classifier revealed a system performance with precision 94.67%, sensitivity 88.89%, specificity 99.24%, efficiency 92.91%, positive prediction accuracy 98.91%, reliability of 94.40% and overall accuracy of 88.60%. The deterministic behavior model classified

positively all recognized gesture patterns present in the database. The abundance of output data makes the system recommendable in research, in education, in physical exercises training, and in entertainment games.

The findings of this study indicated the alternative models that faced expressions emotion classification suitable in many humans and animals' behavior applications. Proposed models were generic and hence constituted the contribution to the maturation of behavior classification from differentiated to general body gesture models of Olivier (2009). Furthermore, the models and the subsequent system show the contribution of the Kinect technology to migrate from X-box to PC applications and from entertainment games to behavior research, monitoring, surveillance, human-machine and other applications.

As illustrated, it was mentioned yes for happiness, wow for surprise, ouch for sadness, no for fear, bah for disgust and fuck for anger. Thus, vocal utterances were introduced in the model to form a gesture-utterance-context behavior multimodal model. Solely observed, body gesture patterns did not show an unambiguously discriminated emotion. Therefore, the ethological emotion model showed a multidimensional array associating to percentage of confidence to the threefold "gesture-utterance-context" for each emotion, the models were dependent to the Kinect tracking capability and running time. Kinect SDK estimation tracking of occluded limbs facilitated introduction of many mistakes made form restriction of the camera view point and distance from the player.

REFERENCES

Albacea, E. A. (2006). *Automata, Formal Languages and Computations* (3rd ed.). San Pedro, Banlat: JPVA Publishing House.

Ao, M., Yi, D., Lei, Z., & Li, S. (2009). Face Recognition at a Distance: System Issues. In S. Z. Massimo Tistarelli (Ed.), *Handbook of Remote Biometrics: for Surveillance and Security* (1st ed., pp. 155-168). Springer.

Kipp, M., & Martin, J.-C. (2009). Gesture and Emotion: Can basic gestural form features discriminate emotions?

Klein, Z. (2000). The ethological approach to the study of human behavior. *Neuroendocrinology Letters* (21), 477–481.

Kramer, J., Burrus, N., Florian, E., Herrera, C. D., & Parker, M. (2012). *Hacking the Kinect.* New York: Apress..

George, F. J., Robert D., W. (Producteurs), & Landis, J. (Réalisateur). (1988). *Coming to America* [Film]. Paramount Pictures.

Lehner, P. N. (1996). Handbook of ethological methods (Second ed.). Cambridge, United Kingdom: Cambridge University Press.

McGinnis, P. (2005). *Biomechanics of sport and exercise* (2nd ed.). Champaign, IL: Human Kinetics.

Microsoft Research. (2011, July 22). *Kinect™ for Windows® Programming Guide: Getting Started with the Kinect for Windows SDK Beta from Microsoft Research.* Retrieved November 24, 2011, from http://research.microsoft.com/kinectsdk

Microsoft Research. (2011a, July 22). *Kinect™ for Windows® Programming Guide: Getting Started with the Kinect for Windows SDK Beta from Microsoft Research.* Retrieved November 24, 2011, from http://research.microsoft.com/kinectsdk

Microsoft Research. (2011b, July 22). *SkeletialViewe Walkthrough: C++ and C#.* Retrieved November 24, 2011, from http://research.microsoft.com/kinectsdk

Muscillo, R., Schmid, M., Conforto, S., & D'Alessio, T. (2001). Early recognition of upper limb motor tasks through accelerometers: real-time implementation of a DTW-based algorithm. *Computers in Biology and Medicine, 41*, 164–172.

Neuman, C. (2005). *Converting Deterministic Finite Automata to Regular Expression*. Retrieved February 17, 2012, from Prof. K.R. Chowdhary, PhD: http://krchowdhary.com/toc/dfa-to-reg-exp.pdf

Olivier, M. S. (2009). *Information Technology Research - A Practical Guide for Computer Science and Informatics* (Third ed.). Van Schaik.

Pe, M. L., & Kuppens, P. (2012). The Dynamic Interplay Between Emotions in Daily Life: Augmentation, Blunting, and the Role of Appraisal Overlap. *Emotion*.

TANG, J., ZHANG, Y., SUN, J., RAO, J., YU, W., CHEN, Y., et al. (2012). Quantitative Study of Individual Emotional States in Social Networks. *IEEE TRANSACTIONS ON AFFECTIVE COMPUTING, 3* (2), 132-144.

Tormene, P., Giorgino, T., Quaglini, S., & Stefanelli, M. (2009). Matching incomplete time series with dynamic time warping: an algorithm and an application to post-stroke rehabilitation. *Artificial Intelligence in Medicine, 45*, 11-34.

Vivaracho-Pascuala, C., Faundez-Zanuyb, M., & Pascual, J. M. (2009). An efficient low-cost approach for on-line signature recognition based on length normalization and fractional distances. *Pattern Recognition, 42*, 183-193.

Webb, J., & Ashley, J. (2012). *Beginning Kinect Programming with the Microsoft Kinect SDK.* Apress