Computing and Informatics, Vol. 27, 2008, 837–851

MULTI-LEVEL REPRESENTATION OF GESTURE AS COMMAND FOR HUMAN COMPUTER INTERACTION

Radu-Daniel Vatavu

SC Instead Software SRL
2, Rulmentului Street
720263 Suceava, Romania
&
University Stefan cel Mare of Suceava
13, University Street
720229 Suceava, Romania
e-mail: radu@instead.ro

Ştefan-Gheorghe Pentiuc

University Stefan cel Mare of Suceava
13, University Street
720229 Suceava, Romania
e-mail: pentiuc@eed.usv.ro

Manuscript received 11 May 2007; revised 11 October 2007
Communicated by Milan Rusko

Abstract. The paper addresses the multiple forms of representation that human gesture takes at different levels for human computer interaction, ranging from gesture acquisition to mathematical model for analysis, pattern for recognition, record for database up to end-level application event triggers. A mathematical model for gesture as command is presented. We equally identify and provide particular models for four different types of gestures by considering both posture information and underlying motion trajectories. The problem of constructing gesture dictionaries is further addressed by taking into account similarity measures and dictionary discriminative features.

Keywords: Human gestures, gesture representation, dictionary of gestures

1 INTRODUCTION

Gesture has been given a lot of attention in the last decades as an effective means of human-computer interaction. [15, 16, 17, 21, 22, 32, 36] give overviews on the state-of-the-art in gesture interaction including gesture taxonomies for HCI, existing technologies, recognition and interpretation techniques. Human gestures and gesture recognition are terms often encountered in human-computer interaction and for many (or most) people gestures are perceived as the next upcoming interaction technology. The main arguments are those of naturalness (gestures are used every day in the real world to interact with real objects or to convey information) and efficiency (interaction becomes similar to what we have been training for all our lives). All this is true only if we consider an ideal gesture interaction technique that does not intrude, burden or add significant cognitive load [10, 30].

Use of gesture as means of interacting has thus been very rapidly proliferating and many systems emerged that allow users to touch, point, work and travel by means of head, hand, arm, gaze or whole body gestures. The most common areas of application are in sign language recognition [14], gesture to speech [11], virtual environments [10, 24], 3D modeling [23], human-robot dialogue [20], and the list goes on.

In this paper we follow human gesture in all its representations as it appears at different levels in the HCI domain (such as raw data streams for the low-level tasks of gesture acquisition, mathematical model for analysis, pattern for recognition, record for database/dictionary of gestures, event in a high-end programming language). We provide a mathematical formalization for gestures in order to develop further discussions on gesture dictionaries. With regard to gesture as command, we identify and provide models for four different types of gestures by taking into account both posture information (or static dimension of gesture) and underlying motion trajectories (dynamic component).

2 DEFINITIONS

A first look on a few terms and definitions that pertain to gesture and which are relevant to this paper is mandatory as it is a brief discussion on the possible meanings of the word gesture.

Prior to anything, gestures can be defined as a physical movement of hands, arms, face and body with the intent of conveying information and meaning. From a biological and sociological perspective, gestures are loosely defined and thus researchers are free to visualize and classify gestures as they see fit. For example, biologists define gestures broadly, stating "the notion of gesture is to embrace all kinds of instances where an individual engages in movements whose communicative intent is paramount, manifest, and openly acknowledged" [28].

First of all, distinction must be made between gesture and posture. There is the tendency [27] to capture the dynamic part in gesture while to consider posture as being static. Consulting [1, 26] we end up with the following definitions for posture and gesture:

Posture (noun): a position of the body or of body parts.

Posture (noun):

1. The position or bearing of the body whether characteristic or assumed for a special purpose (erect posture).
2. A conscious mental or outward behavioral attitude.

Gesture (noun): 1. A motion of the limbs or body made to express or help express thought or to emphasize speech.

2. The act of moving the limbs or body as an expression of thought or emphasis.
   - (verb intr.) To make gestures.
   - (verb tr.) To show, express or direct by gestures.

We are thus further considering posture as being the gesture represented by the position of body or of body parts as with reference to computer interaction. For example, making a fist and holding it in a certain position for a given amount of time is considered to be a posture.

Definition 1. By posture we understand a set of measurements $p = (p_1, p_2, \dots p_n)$ from a given values domain $p_i \in X$ that describe the pose of body or of body parts at one instant of time.

Definition 2. Let P be the set of all postures: $P = \{p = (p_1, p_2, \dots p_n), p_i \in X\}$.

The number of measurements being performed is denoted by n while p_i represents the value of the ith measurement or feature. The features may be quantitative (continuous, discrete) or qualitative (nominal, ordinal). X may be a real-valued domain such as $R, R^k, [0,1]^k$, etc. For example, we may refer to hand posture as to the relative position and orientation of fingers at one instant of time. Consequently, we may choose our measurements for describing a hand posture to be real-valued such as angles between fingers [37], relative distances between adjacent fingers, hand orientation [3], moments of different order [21]. We may as well use qualitative variables such as $\{true, false\}$ for given predicates [13].

A gesture is defined as a dynamic movement, such as waving good bye or describing the shape of a circle using hands. Combining postures and dynamic movements can lead up to different types of gestures as we will discuss them later in the article.

Equally important, we want to isolate for the purpose of our discussion only those interactions for which gestures are articulated and recognized. The definition of articulated gesture as in [19] is more appropriate in this case: "A gesture is a motion of the body that contains information. Waving goodbye is a gesture. Pressing a key on a keyboard is not a gesture because the motion of a finger on its way to hitting a key is neither observed nor significant. All that matters is which key was pressed."

3 GESTURE AS COMMAND

Interaction is performed with both motion trajectories (dynamics) and postures (static information). We may thus provide the following classification of gesture commands with regards to their structural pattern as given by the amount of posture or motion employed when performing the command.

3.1 Static Simple Gestures

Static simple gestures are gestures that convey the desired information only through the use of a single posture that is maintained for a certain amount of time (as an example, see the confirmation or acknowledge gesture as depicted in Figure 1):

$$g_{ss} = (\text{posture}, \text{time}) \in P \times [0, \infty) \tag{1}$$

where ss stands for static simple and P represents the set of all postures.

Let also G_{ss} be the set of all simple static gestures:

$$G_{ss} = \{g_{ss}/g_{ss} = (\text{posture}, \text{time}) \in P \times [0, \infty)\} . \tag{2}$$

For completeness we consider the empty gestures set $\Phi \subset G_{ss}$ as the set of simple static gestures for which the time component is 0:

$$\Phi = \{g_{ss} \in G_{ss}/g_{ss} = (\text{posture}, 0)\} . \tag{3}$$

Definition 3. Any element $\phi \in \Phi$ is an empty gesture.

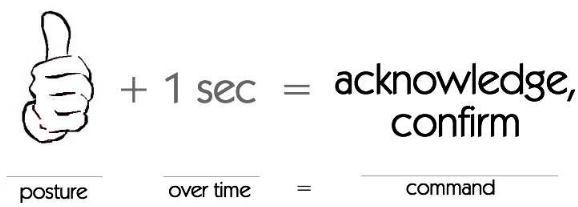

Fig. 1. Example of a simple static gesture: the "thumbs up" posture held for a period of say 1 second may be associated with user acknowledging in response of an application confirmation enquiry

3.2 Static Generalized Gestures

Static generalized gestures are gestures that are represented by a series of consecutive postures which are maintained for certain amounts of time. Again, only posture information is sufficient for grasping the meaning of the gesture command (for example see the "change scale" gesture as given in Figure 2).

$$g_{sg} = \left\{g_{ss}^1, g_{ss}^2, \ldots\right\} \quad (4)$$

where sg stands for static generalized. Equivalently, the generalized static gesture may be defined as a subset of the partition set of G_{ss}: $g_{sg} \subset P\left\{G_{ss}\right\}$.

A generalized gesture may be reduced to a simple gesture if the set order $|g_{sg}| = 1$. Also, a generalized gesture may include at limit an infinity number of postures.

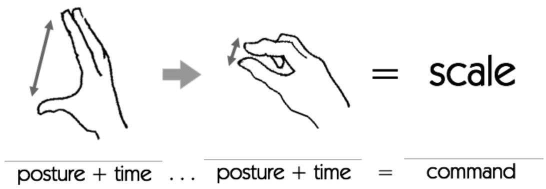

Fig. 2. Example of a generalized static gesture: the "change scale" gesture executed with one hand indicates a change in size or equivalently, a zoom operation that would be proportional to the distance between the index and thumb fingers. The gesture consists of several hand postures ranging say from a large distance to a small/null one between the user hands index and thumb fingers

3.3 Dynamic Simple Gestures

Dynamic simple gestures are gestures for which the posture information is not important as all the meaning lies within the underlying motion trajectory (for example the "undo" gesture as depicted in Figure 3).

A simple dynamic gesture may be defined as a function of time (either continuous or discrete) having as values the coordinates in R^d of the motion trajectory.

$$g_{ds} = g_{ds}(t) : R \to R^d \quad (5)$$

where ds stands for dynamic simple.

3.4 Dynamic Generalized Gestures

Dynamic generalized gestures are gestures for which both the motion trajectory and posture are equally important for grasping the meaning of the user's gesture command (for example the "drag & drop" command as presented in Figure 4). A generalized dynamic gesture may be defined as:

$$g_{dg} = g_{dg}(t) : R \to R^d \times P \quad (6)$$

where dg stands for dynamic generalized and P is the set of all postures.

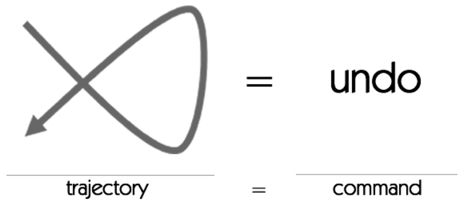

Fig. 3. Example of a simple dynamic gesture: the gesture represented by an X-cross may be associated with performing of an "undo" operation, closing the current window, a cut operation in an editing scenario, etc

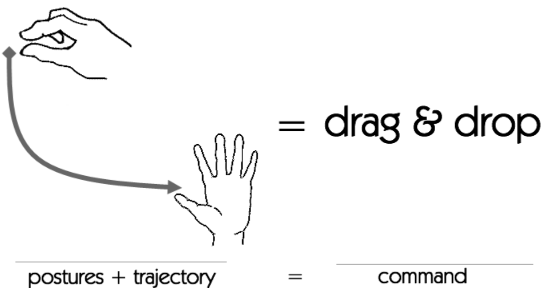

Fig. 4. Example of a generalized dynamic gesture : the "drag & drop" operation may be implemented using two hand postures ("grab" and "release") and a motion trajectory necessary for the start and end locations

Finally, we arrive at a general representation of gesture:

Definition 4. A gesture g is a function of time with values into the Cartesian product of the coordinates space R^d and the set of all postures P:

$$g = g(t) : R \to R^d \times P. \tag{7}$$

From this representation we can derive all the particular structural types:

- $d = 0$ and $P = p_0$ will give us the simple static gesture (p_0)
- for $d = 0$ we get the generalized static gesture
- $P = \Phi$ and $d > 0$ provides for the simple dynamic gesture
- $d > 0$ gives us the generalized dynamic gesture.

4 GESTURE DICTIONARIES

Definition 5. Let G be the set of all gestures $G = \left\{g/g(t) : R \to R^d \times P\right\}$.

Due to the fact that gesture is defined as a function over the whole range of time values, we further introduce the restriction of this function to a limited time range. This will allow us to select the "interesting" part of gesture, i.e. the moment when the gesture actually begins (for example the "grab" posture as in Figure 4) and the moment the gesture ends (the "release" posture), moments that match the start/stop sequence as perceived by the user.

Definition 6. Let $g \in G$ be a gesture and $a, b \in R, a \leq b$ two real values. Then $g|_{a,b}$ is the restriction of gesture g on the $[a, b]$ time interval.

The restriction of a gesture to a time interval allows for further definitions of amplitude of gesture $g(t)$, Ag and the interesting part of gesture, $intg$.

Definition 7. Let $g \in G$ be a gesture. Let $a = inf\,\{t \in R/g(t) \notin \Phi\}$ and $b = sup\,\{t \in R/g(t) \notin \Phi\}$. Then $Ag = b - a$ is the time amplitude of g and $intg : [0, Ag] \to R^d \times P$, $intg = g(t + a)$ is the interesting part of g.

Definition 8. Let $intG$ be the set of all interesting gestures, $intG = \{\tilde{g} \in G/\exists g \in G \text{ so that } intG = \tilde{g}\}$.

The amplitude and the interesting restriction let us define identical gestures as follows.

Definition 9. Two gestures $g, h \in G$ are identical and we denote $g = h \Leftrightarrow Ag = Ah$ and $intg = inth$.

We further introduce similarity and dissimilarity measures over the set of all interesting gestures as well as two propositions that allow for transforming a similarity function into a dissimilarity one and vice versa.

Definition 10. Let $s : intG \times intG \to R^+$. s is a similarity measure over $intG$ if the following conditions are met $\forall g, h \in intG$:

a) $s(g, h) = s(h, g)$
b) $s(g, g) = s(h, h) \geq s(g, h)$.

Definition 11. Let $d : intG \times intG \to R^+$. d is a dissimilarity measure over $intG$ if the following conditions are met $\forall g, h \in intG$:

a) $d(g, h) = d(h, g)$
b) $d(g, g) = 0$.

Proposition 1. Let $s : intG \times intG \to R^+$ be a similarity measure over $intG$ and $\max(s) = \max\{s(g, h)/g, h \in intG\}$. Then $d : intG \times intG \to R^+, d(g, h) = \max(s) - s(g, h)$ is a dissimilarity measure over $intG$.

Proposition 2. Let $d : intG \times intG \to R^+$ be a dissimilarity measure over $intG$ and $\max(d) = \max\{d(g, h)/g, h \in intG\}$. Then $s : intG \times intG \to R^+, s(g, h) = \max(d) - d(g, h)$ is a similarity measure over $intG$.

A set of interesting gestures and a dissimilarity measure permits definition of gesture dictionaries.

Definition 12. Let D be a set of interesting gestures, $D \subset P\{intG\}$ and $d : intG \times intG \to R^+$ a dissimilarity measure over $intG$. Then D is discriminative of degree 1 with respect to D if $d(g,h) > 0\ \forall g,h \in D$.

The degree 1 of discriminative power for a set D simply restricts that there are no two gestures in the set that are similar with respect to a dissimilarity measure. The definition below is more powerful and restricts the set D so that there are no two gestures in the set g,h where h would be similar to any part of g.

Definition 13. Let D be a set of interesting gestures, $D \subset P\{intG\}$ and $d : intG \times intG \to R^+$ a dissimilarity measure over $intG$. Then D is discriminative of degree 2 with respect to D if $d(g, h|_{a,b}) > 0\ \forall g,h \in D\ \forall a,b \in [0, Ah]$, $a \leq b$.

The next proposition states an inclusion level between the sets of discriminative degrees 1 and 2.

Proposition 3. Let $d : intG \times intG \to R^+$ a dissimilarity measure over $intG$, $D^1|_d$ be the set of all sets of interesting gestures that are discriminative of degree 1 with respect to d and $D^2|_d$ be the set of all sets of interesting gestures that are discriminative of degree 2 with respect to d. Then $D^2|_d \subset D^1|_d$.

We may now define a gesture dictionary.

Definition 14. Let D be a set of interesting gestures $D \subset P\{intG\}$ and $d : intG \times intG \to R^+$ a dissimilarity measure over $intG$. The pair (D,d) is a gesture dictionary if D is discriminative of degree 1 with respect to d.

Due to the fact that gestures are functions of time and considering the restrictions of HCI systems that need to process, recognize gestures and provide feedback in real time, we introduce the sequential gesture dictionary below.

Definition 15. Let D be a set of interesting gestures $D \subset P\{intG\}$ and $d : intG \times intG \to R^+$ a dissimilarity measure over $intG$. The pair (D,d) is a sequential gesture dictionary if D is discriminative of degree 2 with respect to d.

5 MULTI-LEVEL REPRESENTATION OF GESTURE

The complex nature of gesture recognition makes human gestures have different representations at different levels. We have identified four distinct processing layers (see Figure 5) corresponding to the actual gesture execution in the real world, acquisition process, modeling and application. We follow the gesture in all its representations starting from the raw format as provided by an acquisition device to the final actual interaction stage (where gesture triggers action) at the highest application level.

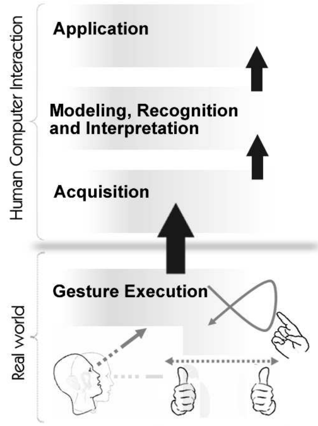

Fig. 5. Different levels of representation for gestures in human-computer interaction

5.1 Gesture in Action

Gestures express ideas, sentiments and intentions, sometimes replacing words and enhancing speech. Gestures convey information and are accompanied by content and semantics. Various psycholinguistic studies have been conducted in what concerns the understanding of gesture communication. All these studies [7, 9, 18, 25] provide an excellent starting material for the domain of human-computer interaction. For example, [7] identifies three types of gestures (or three different functional roles associated to gestures) hence we have a classification by gesture functions:

- Ergotic gesture that acts on the environment and derives from the notion of modeling the real world. It is the type of gesture that is considered for example when interacting with the virtual objects of a virtual environment [4, 6, 31, 35].
- Epistemic gesture that offers information with regard to temperature, pressure, shape, orientation, weights (the tactile sense). The environment discovery is achieved through tactile experience [5, 8].
- Semiotic gesture that produces an informational message for the environment with the role of conveying information. It is the type of gesture for yes/no, approve/deny actions for the human-computer dialogue [12]

5.2 Acquisition Level

At the acquisition level, a gesture is represented by a stream of data according to the technology used by the capture device. At this level we dispose of a raw representation of the gesture that may contain a large amount of noise. For example, a gesture may be represented by a single image that contains the user's hand with a specific posture in front of a working desk that may contain additional information:

parts of the desk, notebooks, the PC keyboard, the user's watch or the color of the shirt (Figure 6). We are dealing with the raw representation of gesture as provided by the acquisition device and we can easily note the huge amount of extra information that is not needed at all. Noise is heavily present in this basic gesture representation.

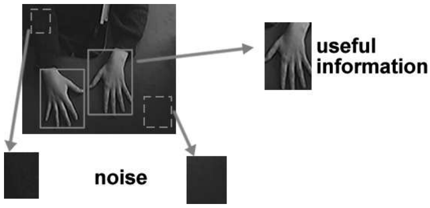

Fig. 6. Gesture is a raw stream of data at the acquisition level as outputted by the capture device (the picture represents a snapshot of lab developed hand gesture acquisition system)

If the interest is on dynamic gestures implying motion trajectories and the technology is video, the raw representation as outputted by the acquisition device would be a sequence of video frames, each containing a huge amount of extra not needed information that we can qualify as noise. Another raw representation of a gesture at this level could be a stream of data points given say as (x, y, z) pairs as output by a tracking device, for example [2, 33]. We would deal in this case with an incomplete representation of our gesture as posture had not been taken into account.

The gesture representation depends on the technology used for capturing. It may be a single image or a sequence of video frames for visual gesture acquisition; a stream of data points for a tracking device; a set of measurements (angles, flexions) for a data glove and many others. This is due to a large amount of non-traditional immersive devices that have been very rapidly proliferating. They include spatial input devices (or trackers), pointing devices and whole hand devices that allow for hand gestures input. The technology varies including: magnetic, mechanical, acoustic, inertial, vision/video camera based or hybrid [21, 35].

5.3 Modeling, Recognition and Interpretation

At this level we dispose of a mathematical model for gesture in close relation to the interaction purpose as stated in the gesture definition paragraph. We make distinction between posture (as static information) and gesture (as dynamics). Several stages may be encountered at this level (see Figure 7):

Gesture modeling: We defined a gesture as a function of time with values into the Cartesian product of the coordinates space R^d and the set of all postures P: $g = g(t) : R \to R^d \times P$.

Classification/Semantic Interpretation: Gesture is also a pattern after feature extraction has been performed on the mathematical model. Pattern recognition at this level aims to classify data (patterns) based on either a priori knowledge or statistical information extracted from the patterns. The patterns to be classified are usually groups of measurements or observations defining points in an appropriate multidimensional space. Semantic interpretation may further associate the gesture to a class of semantic types such as commands, gesticulation, etc. Production rules and formal logic or any other artificial intelligence techniques for association to knowledge sets may be used.

Gesture storage/dictionary of gestures: Gesture is also a record in a database that allows for storage of the gesture model (a bijection may be defined between the gesture model representation and the database storage specific format).

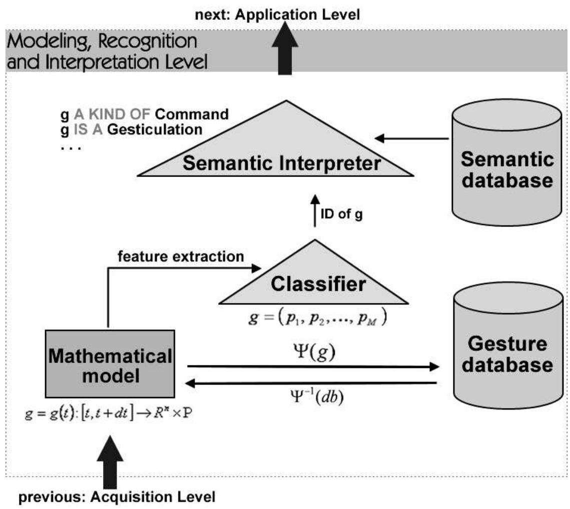

Fig. 7. Gesture is a mathematical model, a pattern or a database record at the Modeling, Recognition and Interpretation level

Real problems arise in what concerns choosing the best dictionary for human gestures for a given application. Although gestures are perceived as a natural means of interacting and conveying information (hence a gesture based interface would prove to be ideal), gestures may also be described as imprecise, not self revealing and also non ergonomic. A particular problem relates to finding the right gestures that would feel comfortable and natural from the user's experience point of view. Several attempts have been made on defining gesture dictionaries for application specific needs [27, 34, 38]. [30] conducted an ergonomic study for selecting the appropriate gesture commands for operations such as: selection, move, scale, copy, confirm, yes/no, undo using both single hand and two hands gestures in a video camera based top view of a working table.

Another problem arises from the fact that gesture commands have to be identified (or designed) with the particularity of assuring a natural and comfortable user experience, all this considering the existing GUIs and interaction paradigms, for example WIMP (Windows, Icon, Menu, Pointing). In this particular case, which would be the most suitable gesture for activating/closing a menu or for maximizing/minimizing an application window? A simple proposal is given in Figure 8 [35] but the question still remains: are the proposed gestures natural or are we looking for a compromise between natural and new to be learned gestures?

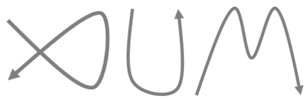

Fig. 8. Sample dynamic gestures for Undo/Redo/Open Menu commands

5.4 Application Level

Gesture is associated with the trigger of a certain action at the application level, according to the current selected working scenario. We can look at a gesture as an event that is fired whenever the gesture interface is enabled and allows for executing the associated action. This can be described using pseudo code language-like formalism as follows:

```
app.EventHandler += new EventHandler(OnGestureInputEvent);
...
result OnGestureInputEvent(EventParams e)
{
        switch(e.GestureType)
        {
                  ...
                  // take appropriate action
                  ...
        }
        ...
}
```

6 CONCLUSIONS

We discussed gestures as commands at several stages of processing as they appear in human-computer interaction: data acquisition, mathematics modeling and analysis, end-user application. We identified four types of gestures ranging from simple static to complex motion trajectories with associated postures. Making use of a mathematical model for gestures, we introduced a few definitions and propositions with regards to the structure and composition of gesture dictionaries. Future work will converge to further develop the already started gesture dictionary theoretical development with specifics on real-time systems implementations.

REFERENCES

[1] American Heritage Dictionary of the English Language, 4th Ed., Houghton Mifflin Company, 2000.

[2] Ascension Products, Flock of Birds. Available on: `http://www.ascension-tech.com/products/flockofbirds.php`, Last visited Oct. 2006.

[3] Bedregal, B. C.—Costa, A. C. R.—Dimuro, G. P.: Fuzzy Rule-Based Hand Gesture Recognition. In: Proc. of IFIP Conference on Artificial Intelligence, Springer, Berlin, 2006, pp. 1–10.

[4] Bolt, R. A.: Put-That-There: Voice and Gesture at the Graphics Interface. In: Proc. of the 7th Annual Conf. on Computer Graphics and Interactive Techniques, Seattle, United States, July 14–18, 1980, pp. 262–270.

[5] Buttolo, P.—Stewart, P. J.—Marsan, A. L.: A Haptic Hybrid Controller for Virtual Prototyping of Vehicle Mechanisms. In: Proc. of 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2002, pp. 249–254.

[6] Cabral, M. C.—Morimoto, C. H.—Zuffo, M. K.: On the Usability of Gesture Interfaces in Virtual Reality Environments. In: Proc. of the Latin American Conf. on Human-Computer Interaction, Cuernavaca, Mexico, 2005, pp. 100–108.

[7] Cadoz, C.: Le Geste Canal de Communication Homme/Machine. Technique et Science Informatique, Vol. 13, 1994, pp. 31–61.

[8] Campbell, C.—Zhai, S.—May, K.—Maglo, P.: What You Feel Must Be What You See: Adding Tactile Feedback to the Track Point. In: Proc. of IFIP Interact '99, IOS Press, 1999, pp. 383–390.

[9] Cassell, J.: A Framework for Gesture Generation and Interpretation. In: R. Cipolla and A. Pentland (Eds): Computer Vision in Human-Machine Interaction, Cambridge University Press, New York, 1996, pp. 191–215.

[10] Cerney, M. M.—Vance, J. M.: Gesture Recognition in Virtual Environments: A Review and Framework for Future Development. TR ISU-HCI-2005-01, Iowa State University, 2005.

[11] Craven, M. P.—Curtis, K. M.: GesRec3D: A Real-Time Coded Gesture-To-Speech System With Automatic Segmentation and Recognition Thresholding Using Dissimilarity Measures. Gesture-Based Communication in Human-Computer Interaction, No. 2915, 2004, LNCS, pp. 231–238.

[12] Danylak, R.: A Semiotic Reading of Gestures for Multimodal Interaction. In: Proc. of Gestural Interaction Workshop, Sydney, Australia, 2005.

[13] Eisenstein, J.—Ghandeharizadeh, S.—Golubchik, L.—Shahabi, C.—Yan, D.—Zimmermann, R.: Device Independence and Extensibility in Gesture Recognition. In: Proc. of IEEE Virtual Reality, 2003, pp. 207–214.

[14] Gineke ten Holt: Why Don't You See What I Mean? Prospects and limitations of current automatic sign recognition research. Sign Language Studies, No. 4, 2006, pp. 416–437.

[15] Hu, W.—Tan, T.—Wang, L.—Maybank, S.: A Survey on Visual Surveillance of Object Motion and Behaviors. IEEE Transactions On Systems, Man, and Cybernetics, Vol. 34, 2004, No. 3, pp. 334–352.

[16] Jaimes, A.—Sebe, N.: Multimodal Human Computer Interaction: A Survey. In: Proc. of IEEE Int. Workshop on Human Computer Interaction, Beijing, China, 2005, LNCS, 3766, pp. 1–15.

[17] Karam, M.—Schraefel, M. C.: A Taxonomy of Gestures in Human-Computer Interactions. TR ECSTR-IAM05-009, University of Southampton, 2005.

[18] Kendon, A.: Current Issues in the Study of Gesture. In: Nespoulous, Perron, and Lecours (Eds), The Biological Foundation of Gestures: Motor and Semiotic Aspects, Lawrence Erlbaum Assoc., Hillsdale, 1986, pp. 23–47.

[19] Kurtenbach, G.—Hulteen, E.: Gestures in Human-Computer Communications. In: B. Laurel (Ed.), The Art of Human Computer Interface Design, Addison-Wesley, 1990, pp. 309–317.

[20] Kim, K. K.—Kwak, K. C.—Ch, S. Y.: Gesture Analysis for Human-Robot Interaction. In: Proc. of the 8th Int. Conf. on Advanced Communication Technology, Vol. 3, 2006, pp. 1824–1827.

[21] LaViola, J.: A Survey of Hand Posture and Gesture Recognition Techniques and Technology. TR CS-99-11, Department of Computer Science, Brown University, Providence RI, 1999.

[22] Lenman, S.—Bretzner, L.—Thuresson, B.: Computer Vision Based Hand Gesture Interfaces for Human-Computer Interaction. TR TRITA-NA-D0209, CID Center for User Oriented IT Design, 2002.

[23] Looi, N. K.—Loo, C. K.—Pau, K. N.: Real Time Actions and Gestures As Intuitive User Interface for Creative 3D Modeling. In: Proc. of Int. Conf. on Active Media Technology, 2005, pp. 487–490.

[24] Mine, M.: Moving Objects in Space: Exploiting Proprioception in Virtual Environment Interaction. In: Proc. of SIGGRAPH '97, ACM Press, 1997, pp. 19–26.

[25] McNeill, D.: Hand and Mind: What Gestures Reveal About Thought. University of Chicago Press, 1992.

[26] Merriam-Webster's Medical Dictionary, Merriam-Webster Inc., 1995.

[27] Mulder, A.: Hand Gestures for HCI: Research on Human Movement Behavior Reviewed in the Context of Hand Centered Input. TR 96-1, Simon Fraser University, 1996.

[28] Nespoulous, J. L.—Perron, P.—Lecours, A. R.: The Biological Foundations of Gestures: Motor and Semiotic Aspects. Lawrence Erlbaum Associates, Hillsdale, MJ, 1986.

[29] Tsang, E. K. H.—Sun, H.: An Efficient Posture Recognition Method Using Fuzzy Logic. Virtual Reality, Vol. 3, 1998, No. 2, Springer London, pp. 112–119.

[30] Nielsen, M.—Moeslund, T.—Storring, M.—Granum, E.: A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI. In: Proc. of 5th Int. Workshop on Gesture and Sign Language based Human-Computer Interaction, Genova, Italy, 2003.

[31] O'Hagan, R.—Zelinsky, A.: Vision-Based Gesture Interfaces to Virtual Environments. In: Proc. of 1st Australasian User Interfaces Conf., Canberra, Australia, 2000, pp. 73–80.

[32] Pavlovic, V. I.—Sharma, R.—Huang, T. S.: Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review. IEEE Transactions on PAMI, Vol. 19, 1997, No. 7, pp. 677–695.

[33] Sensable Technologies, Products: PHANTOM devices. Available on: `http://www.sensable.com`, Last visited October 2006.

[34] Soares, J. M.: Contribution a la Communication Gestuelle dans les Environments Virtuels Colaboratifs. Ph. D. dissertation sustained 9th July 2004 at the Institut National des Telecomunications and Universite de Technologie de Troyes.

[35] Vatavu, R. D.—Pentiuc, S. G.—Chaillou, C.: On Natural Gestures for Interacting in Virtual Environments. Advances in Electrical and Computer Engineering, Vol. 5, 2005, No. 24, pp. 72–79.

[36] Watson, R.: A Survey of Gesture Recognition Techniques. TR TCD-CS-93-11, Department of Computer Science, Trinity College Dublin, 1993.

[37] Weissmann, J.—Salomon, R.: Gesture Recognition for Virtual Reality Applications Using Datagloves and Neural Networks. In: Proc. of Int. Joint Conf. on Neural Networks, Washington, DC, USA, 1999, pp. 2043–2046.

[38] Wexelblat: An Approach to Natural Gesture in Virtual Environments. ACM Transactions on Computer-Human Interaction TOCHI, Vol. 2, 1995, No. 3, pp. 179–200.

Radu-Daniel Vatavu received his diploma in computer science from the University of Suceava, Romania, in 2004. He is currently working on his Ph. D. on gesture-based interaction at the LIFL lab in Lille, France and at the University of Suceava, Romania. His research interests include human-computer interaction, computer vision and machine learning.

Ştefan-Gheorghe Pentiuc is a Professor at the Computer Science Department of University of Suceava, Romania since 1997 and the director of the Research Center in Computer Science of Suceava since 2004. His research interests include pattern recognition, distributed artificial intelligence, programming on the Internet, algorithms and Java technologies.