Gestures for your Workplace: Investigating Body Interaction for Everyday Desk Scenarios

Radu-Daniel Vatavu
Research Center in Computer Science, University "Stefan cel Mare" of Suceava
13, Universitatii Street
Suceava 720229, Romania
vatavu@eed.usv.ro

Ovidiu-Ciprian Ungurean
Research Center in Computer Science, University "Stefan cel Mare" of Suceava
13, Universitatii Street
Suceava 720229, Romania
ungurean.ovidiu@gmail.com

Stefan-Gheorghe Pentiuc
Research Center in Computer Science, University "Stefan cel Mare" of Suceava
13, Universitatii Street
Suceava 720229, Romania
pentiuc@eed.usv.ro

ABSTRACT

We discuss in this paper gesture-based interaction techniques that are appropriate for everyday office desk scenarios for which the constraints of the workspace and the seated position limit the available range of body motions. We are interested in gestures performed by hands and head and highlight our findings with regards to the acquisition and recognition of such gestures using vision techniques. Possible applications and implementations are discussed with focus on the intuitiveness and easy-of-use of the involved gesture vocabularies.

Categories and Subject Descriptors

H.5.2 **[User Interfaces]**: Interaction styles (e.g., commands, menus, forms, direct manipulation), Graphical user interfaces (GUI), Input devices and strategies (e.g., mouse, touchscreen), Ergonomics.

General Terms

Algorithms, Human Factors.

Keywords

Gesture-based interaction, hands detection, head acquisition, gesture recognition, desktop computing scenarios.

1. INTRODUCTION

The last decades have witnessed a considerable development of technologies that allow acquisition and recognition of body motions at various rates and resolutions with considerable effort from the computer vision community [1, 2]. Natural body movements and interacting with gestures are usually perceived as innovatory and ideal interfaces however questions must be raised and properly addressed on how the technology should be used in order to compete with the mouse & keyboard pair on the same comparison term of efficiency.

We are focusing in this paper on the most predominant scenario encountered today when interacting with computers: office desktop PC environments. Although systems of sensing computing machines have been put in place for dealing with full-body motion capture [1, 2] and mobile personal computing has already started to conquer the users' computing habits [3], the fact remains that the office desktop PCs with mice and keyboards are the most commonly encountered computing scenarios today. We investigate what are the best ways to enhance interaction at the desktop by means of gestures and how body interaction applies to desktop computing. The goal is not to replace the mouse and keyboard but rather to enhance the current input strategies and techniques in order to achieve fluent, intuitive and easy-of-use interfaces. We present our findings, describe application types and share the lessons we've learned.

2. CONSTRAINTS AT THE OFFICE

Working in front of a desk brings in several limitations and adds constraints on the availability of motions that can be efficiently captured and used. The seated position limits body motions to gestures performed by head and hands in the immediate space above the desk. Also, the popularity of mice and keyboards due to their extensive use makes gestures appear as complementary interaction. These constraints further lead to limitations in gesture acquisition technologies (e.g. putting on and removing sensor gloves just won't work however other objects may be used for interaction such as pens [10] or tangible cubes [11]).

Computer vision has developed in the recent years into an affordable technology that gives good results if used wisely. Web cams, ubiquitous now due to their low-cost, are providing real-time 25 fps at acceptable resolutions and image quality. Cameras may be placed on top of the computer screen so that they would up-front users [9], they may be oriented face down in order to capture the movements of the hands [4, 5, 6] or even

Published by the British Computer Society

hidden inside the desk for tabletop and interactive surfaces computing [12, 13] or simply under the table [17]. Video-based processing comes with many advantages yet it brings in drawbacks as well (among most challenging being the dependency on the environment such as lighting conditions and random motions). However, when wisely used, computer vision will give surprising and very interesting results.

2.1 Your Hands are your Input

Hands may be used to point and select, translate, rotate or generally manipulate objects on the screen while they move in the horizontal plane just above the keyboard or in the vertical plane in front of the monitor. Hands detection and tracking may be achieved using color-based processing such as skin filtering [14]; tracking of KLT features [15] or mixed techniques.

With regards to the interaction possibilities, the dominant hand may simply map the cursor and trigger click events [4] while both hands may work together to achieve complex manipulations. The main advantage of using the hands is the extra degrees of freedom (up to 5 fingers per hand) and the benefit of having two manipulation points (left and right hand) over the single interaction point provided by the mouse. This allows for one and two-hands manipulations to occur naturally as they do in the real world [5].

Two important notes to be made here are related to precision and occlusion which sometime interfere. Up to 5 fingers may be tracked for each hand but the question must be raised and investigated: is the precision really necessary? What are the applications or tasks that would indeed benefit of this high-level precision? Also, depending on the acquisition scenario (one or more cameras), some fingers may be occluded while in movement which makes the tracking process difficult. On the other hand (pun not intended) simple postures may be easily detected such as: pinch, index finger pointed, hand open or closed [4, 5, 7] which in turn give a variety of interaction possibilities depending on context, selected objects or handedness.

2.2 Use your Head

Head movements are a natural way to indicate changes in direction and viewing angle. Motion flow techniques are popular to detect the direction of motion [16]. Once detected, head movements may be mapped to the cursor [7] or included into other controlling metaphors (such as changing the camera view in a FPS game). The target is to achieve intuitive and natural interaction while keeping the efforts involved to a minimum (simple head movements with low amplitude should be considered).

3. INVESTIGATING APPLICATION OPPORTUNITIES

We briefly describe several applications that are immediately available once a simple hand and head tracking mechanism is put in place.

3.1 Take Control of Your Cursor

Simple head rotations (to the right, left, up and down) may be easily mapped to cursor movements while left and right clicks may be triggered by slightly tilting the head to the corresponding direction or by using a timer which automatically triggers an event according to the cursor position and context [7] (see Figure 1).

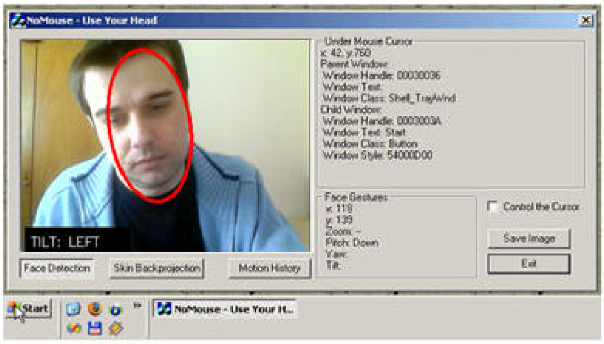

Figure 1. Controlling Windows tasks using head movements and nods.

3.2 Object Manipulation at your Fingertips

Manipulation of virtual objects may be achieved using simple postures that feel natural (such as point and pinch) combined with simple movements [5, 6, 18]. The main goal is not to definitely replace the mouse and keyboard but to add complementary interactions that feel right for certain operations (e.g. rotating an object may be done with two hands that control the axis of rotation while still using the keyboard to introduce text and numbers for various parameters).

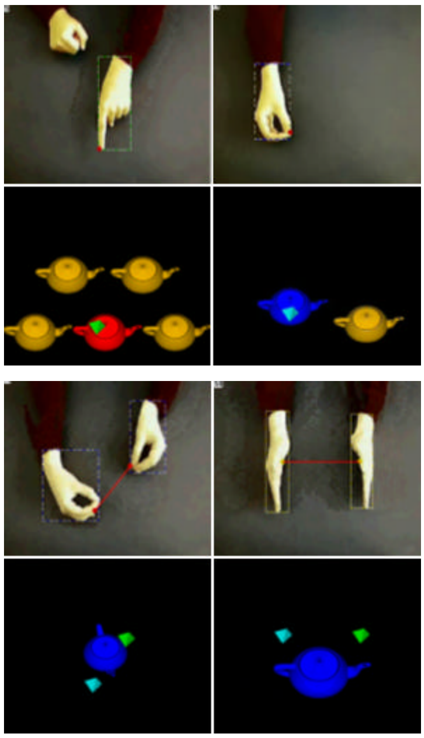

Figure 2. One and two-hands manipulation. Visual feedback is displayed below the video snapshots.

3.3 Do you Game?

Tracking head movements provides enhanced gaming experiences as the camera view changes in accordance with how the head moves [7, 9]. This brings realism into the perceived gaming experience (Figure 3 presents a user controlling the camera angle for a FPS game using left/right and up/down head movements).

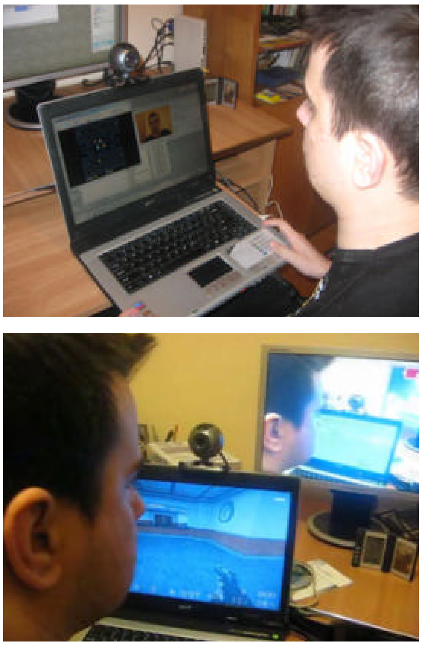

Figure 3. Mapping head movements to the camera view.

3.4 Out of Gestures? Make your Objects Interact for you

Sometimes gestures are not enough and interaction must be complemented with tangible real-world objects. Interaction at the desk may be enhanced beyond gesture commands by making use of objects near-by: real objects which are detected on the surface of the desk may be transformed into virtual graphical elements for which the logic is dictated by the current running application. GUIs become thus tangible with no cognitive effort or special training required, all with a haptic-like feedback that guides the correspondence between real and virtual [8].

Figure 4 illustrates various objects that may be used for interaction: from everyday objects as they may be found on a desk such as coffee cups or pens up to special designed pattern marked cardboards [19].

4. LESSONS WE'VE LEARNED

Whilst implementing the various interaction scenarios, we encountered several challenges and also got feedback from subjects that tested the techniques.

As a general guideline, we found that when dealing with computer vision based gestures acquisition, simplifying (not constraining) the scene as much as possible leads to better accuracies and low-cost algorithms. For example, when detecting hands above the surface of the desk, assuring a comfortable contrast between the desk and the skin color allows for robust and low-cost tracking with simple thresholding techniques. Many interaction opportunities are offered by the use of only two hand postures: point (index finger stretched) and pinch (index and thumb touching). The two postures are different enough to discriminate using simple geometric-based rules on blobs issued from the segmentation process. In turn, they provide a natural technique for selection (such as the point posture) and manipulation (the pinch posture "grabs" the object that will be subjected to the motions of the user's hands). The relative positioning and distance between the hands should be exploited. Visual feedback is mandatory in order for users to have their actions confirmed in real time.

Figure 4. Various objects may be used for interaction. Top: everyday desktop objects (coffee cup, pen, and sheet of paper). Bottom: special designed cardboards.

In what concerns head gestures, when mapping movements to various controls, users find accommodation rather easy and they appreciate the technique as easy to use after practice. However, they testify fatigue for precision movements after some usage hence head-based interaction should be carefully considered only for those operations that would really benefit of such an interaction type. There is great opportunity for games as the movements naturally embed into the action and do not require fine precision.

When using additional objects on the desk, the interaction techniques should be simplified as much as possible (placing an object on the surface of the desk should lead to it being spotted and tracked with real-time visual feedback). Additional objects bring in the advantage of direct manipulation with tangible feedback.

Probably the most important lesson to report may be put as: simple gestures go a long way. Simple and familiar postures together with natural head mappings to the virtual camera were found well suited for complementary desktop interaction.

5. ACKNOWLEDGEMENTS

This research has been supported by the national funding grant Ref. No. 131/2006, INTEROB, under the Research of Excellence national framework.

6. VIDEOS

Several videos demonstrating various interaction techniques such as hand manipulation, gesture recognition, head tracking and head motion recognition as well as object-based interaction are available at http://www.eed.usv.ro/~vatavu

7. REFERENCES

[1] Moeslund, T.B., Hilton, A., Krüger, V. A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding*, vol. 104, issue 2, (Nov. 2006), 90-126.

[2] Poppe, R. Vision-based human motion analysis: An overview. *Computer Vision and Image Understanding*, vol. 108, issues 1-2 (Oct. 2007), 4-18.

[3] Hong, S., Thong, J. Y., Moon, J., and Tam, K. Understanding the behavior of mobile data services consumers. *Information Systems Frontiers*, 10, 4 (Sep. 2008), 431-445.

[4] Wilson, A. Robust Vision-Based Detection of Pinching for One and Two-Handed Gesture Input. In *Proceedings of the 19th Symposium on User Interface Software and Technology (UIST'06)* (Montreux, Switzerland, October 15-18, 2006), ACM Press, New York, NY, 2006, 255-258.

[5] Vatavu R.D., Pentiuc, S.G., Chaillou, C., Grisoni, L., Degrande, S. Visual Recognition of Hand Postures for Interacting with Virtual Environments. In *Proceedings of the 8th International Conference on Development and Application Systems (DAS'06)* (Suceava, Romania), 2006, 477-482.

[6] Vatavu, R.D., Grisoni, L., Pentiuc, S.G. Gesture Recognition based on Elastic Deformation Energies. In *Proceedings of the 7th International Workshop on Gesture in Human-Computer Interaction and Simulation (GW'07)* (Lisbon, Portugal, May 23-25, 2007), 2007, 1-2.

[7] Pentiuc, S.G., Vatavu, R.D., Ungurean, C.O., Cerlinca, T.I. Techniques for Interacting by Gestures with Information Systems. In *Proceedings of the European Conference on the Use of Modern Information and Communication Technologies (ECUMICT'08)* (Gent, Belgium, March 14-18, 2008).

[8] Vatavu, R.D., Pentiuc, S.G., Cerlinca, T.I. Context into Play: Supporting Game Interaction through Real-Time Context Acquisition. In *Proceedings of the Workshop on Multimodal Interfaces in Semantic Interaction at ICMI 2007 (WMISI'07)* (Nagoya, Japan, November 15, 2007), ACM Press, New York, NY, 2008, 3-8.

[9] Ungurean, C.O., Pentiuc, S.G., Vatavu, R.D., Use Your Head: An Interface for Computer Games Using Head Gestures. *Accepted for the 8th International Gesture Workshop (GW'09), Gesture in Embodied Communication and Human-Computer Interaction 2009*, Bielefeld, Germany.

[10] LaViola, J.J. Sketching and gestures 101. In *ACM SIGGRAPH 2007 Courses (SIGGRAPH '07)* (San Diego, California, August 05 - 09, 2007). ACM, New York, NY, 2.

[11] Van Laerhoven, K., Villar, N., Schmidt, A., Kortuem, G., and Gellersen, H. Using an autonomous cube for basic navigation and input. In *Proceedings of the 5th international Conference on Multimodal interfaces (ICMI '03)* (Vancouver, British Columbia, Canada, November 05 - 07, 2003). ACM, New York, NY, 203-210.

[12] Wilson, A. D., Izadi, S., Hilliges, O., Garcia-Mendoza, A., and Kirk, D. Bringing physics to the surface. In *Proceedings of the 21st Annual ACM Symposium on User interface Software and Technology (UIST'08)* (Monterey, CA, USA, October 19 - 22, 2008). ACM, New York, NY, 67-76.

[13] Cao, X., A. Wilson, R. Balakrishnan, K. Hinckley, S. Hudson. ShapeTouch: Leveraging Contact Shape on Interactive Surfaces. In *Proceedings of the 3rd IEEE International Workshop on Horizontal Interactive Human-Computer Systems (Tabletop 2008)* (Amsterdam, the Netherlands, October 1-3), 2008.

[14] Jones, M.J., Rehg, J.M. Statistical color models with application to skin detection. Technical Report 98/11, Cambridge Research Laboratory, 1998.

[15] Kolsch, M., Turk, M. Fast 2d hand tracking with flocks of features and multi-cue integration. In *Proceedings of the IEEE Workshop on Real-Time Vision for Human-Computer Interaction (at CVPR'04)*, 2004.

[16] Bradski, G.R., Davis, J.W. Motion segmentation and pose recognition with motion history gradients. *Machine Vision and Applications*, 13 (2002), 174–184.

[17] Wigdor, D., Leigh, D., Forlines, C., Shipman, S., Barnwell, J., Balakrishnan, R., and Shen, C. 2006. Under the table interaction. In *Proceedings of the 19th Annual ACM Symposium on User interface Software and Technology* (UIST '06) (Montreux, Switzerland, October 15 - 18, 2006). ACM, New York, NY, 259-268.

[18] Vatavu, R.D., Grisoni, L., Pentiuc, S.G. Multiscale Detection of Gesture Patterns in Continuous Motion Trajectories. *Accepted for the 8th International Gesture Workshop (GW'09)*, 2009, Bielefeld, Germany.

[19] Kato, H., Billinghurst, M. Developing AR applications with ARToolKit. In *Proceedings of the 3rd IEEE/ACM international Symposium on Mixed and Augmented Reality (ISMAR'04)* (November 02 - 05, 2004). Symposium on Mixed and Augmented Reality. IEEE Computer Society, Washington, DC, 305-305.