Chapter 61
Human-Computer Interaction in Games Using Computer Vision Techniques

Vladimir Devyatkov
Bauman Moscow State Technical University, Russia

Alexander Alfimtsev
Bauman Moscow State Technical University, Russia

ABSTRACT

A primary goal of virtual environments is to support natural, efficient, powerful and flexible human-computer interaction. But the traditional two-dimensional, keyboard- and mouse-oriented graphical user interface is not well-suited for virtual environments. The most popular approaches for capture, tracking and recognition of different modalities simultaneously to create intellectual human-computer interface for games will be considered in this chapter. Taking into account the large gesture variability and their important role in creating intuitive interfaces, the considered approaches focus one's attention on gestures although the approaches may be used also for other modalities. The considered approaches are user independent and do not require large learning samples.

INTRODUCTION

A primary goal of virtual environments is to support natural, efficient, powerful, and flexible human-computer interaction. If the interaction technology is awkward, or constraining, the user's experience with the synthetic environment is severely degraded. If the interaction itself draws attention to the technology, rather than the task at hand, it becomes an obstacle to a successful virtual environment experience.

The traditional two-dimensional, keyboard- and mouse-oriented graphical user interface (GUI) is not well-suited for virtual environments. Instead, synthetic environments provide the opportunity to utilize several different sensing modalities and integrate them into the user experience. The cross product of communication modalities and sens-

DOI: 10.4018/978-1-4666-3994-2.ch061

ing devices begets a wide range of unimodal and multimodal interface techniques. The potential of these techniques to support natural and powerful interfaces is the future of game constructing and designing.

To more fully support natural communication, it has to not only track human movement, but to interpret that movement in order to recognize semantically meaningful gestures. While tracking user's head position or hand configuration may be quite useful for directly controlling objects or inputting parameters, because people naturally express communicative acts through higher-level constructs such as gesture or speech.

In this chapter, we shall consider the most popular approaches for capture, tracking and recognition of different modalities simultaneously to create intellectual human-computer interface for games. Taking into account the large gesture variability and their important role in creating intuitive interfaces, the considered approaches focus one's attention on gestures although the approaches may be used also for other modalities. The considered approaches are user independent and do not require large learning samples.

BACKGROUND

Gesture Modalities

The scientific interest in the verbal (speech) and nonverbal (gestures, mimicry, touches, etc.) behavior of people during the communication arose only in XX century. The theory of verbal and nonverbal communications for a long time was developed at an intuitive level. Serious scientific investigations of verbal and nonverbal communications began in the 1920s–1930s within the framework of the journalism theory. Psychologists established that a percentage of information transferred by nonverbal signals during the people interaction was from 60% up to 80% (Ekman & Friesen, 1969).

Moreover, most researchers adhere to an opinion that the verbal channel is used substantially for transferring the factographic information while the nonverbal channel is a means of transferring the interpersonal relations and only in rare cases it is used instead of verbal messages. This fact testifies to the important role of the nonverbal information transferred by gestures and mimicry for the people behavior analysis and developing human-machine interface in computer games. For the most part scientific matters deal with gestures performed by hands.

Generally, a gesture is the sign unit carried out by human body parts consciously and unconsciously for the purpose of communications. In order to decode the information incorporated in gestures it is needed to define their classification. Gestures are subdivided into natural and artificial ones. Natural gestures are inherent in a person by nature or are produced by the humanity during the evolution. Gesture classifiers describe images and senses of gestures to use them with a high degree of adequacy.

It should be noted, that all positions are connected with knowledge of environment properties in which the gesture is made or with knowledge of the context accompanying the gesture. And many cultures interpret the same gesture completely in a different way. Moreover, the frequency of gesticulation (a number of gestures made per unit of time) in West Europe is higher than in Russia, but gestures of West Europeans occupy less space, than those of Russians as West Europeans gesticulate with elbows pressed to the body. West European gestures do not intrude at all in the personal space of the interlocutor. As against the Russian tradition, in West Europe symmetric gestures prevail, a handshake is less long, than that in Russia, and gestures are made by the half-bent arm, instead of the outstretched arm as in Russia.

There is another class of gestures, named artificial gestures (Kyle & Woll, 1988). For example, two gesture alphabets for deaf people communication are presented in Figure 1.