Recognition of Chaining Gestures in the Flow in the Context of Virtual Theater

Ronan Billon, Alexis Nédélec, and Jacques Tisseau
Cerv: Centre Européen de Réalité Virtuelle
Laboratoire d'Informatique des SYstèmes Complexes
{billon,nedelec,tisseau}@enib.fr
http://www.cerv.fr/

Introduction. Our aim is to put on a short play featuring a real actor and a virtual actor, who will communicate through movements and choreography with mutual synchronization. Although the theatrical context is a good ground for multimodal communication between human and virtual actors, this paper mainly deals with the ability to perceive the gestures and chaining of gestures of a real actor. During the performance, our goal is to match real-time observation with recorded examples. Recognition events are sent to a virtual actor which replies with its own movements. We follow an approach similar to that proposed by Campbell [1] or Bobick [2]. In this paper, we will describe these two main aspects of gestures communication: representation and perception.

Gesture Representation and Perception. First, we must store an example of the gesture to be recognized. Our idea is that the data for movements from any motion-capture system can be reduced to a single artificial signature. Motion-capture systems generate a large amount of data and each type of sensor has its own characteristics, so we have decided to move away from those technologies and turn our system into a more generic recognition system. We decided to work on the variation during the gesture. For that purpose, we use the well-known PCA[1], that preserves variance, whether it is an angle or acceleration variation. Although we use PCA, our approach is quite different compared to [3] or [4]. Instead of creating a partitioned space with all the samples, each gesture generates its own signature.

With this signature we can perform the recognition. Our system is quite similar to multiagent systems in which each agent is a gesture (perception of the real-time flow, decision of the similarity recognition and action of sending an event). It is the gesture itself that signals when it recognizes itself in the stream. During the performance we have to compare the real-time observation with the stored examples. Two main problems are raised by recognition: firstly, segmentation, e.g., finding the beginning and the end; then the fidelity computation between the recorded gesture and the observed one. We use a forward spotting scheme that executes gesture segmentation and recognition simultaneously. The detection of

[1] Principal Component Analysis

the correct gesture is quick enough for real-time use. During experimentation, we found that the average detection gesture was done 0.7 seconds before the effective end of a 3 seconds gesture.

Gesture Chain Framework. Now that we are able to recognize gesture in real-time, we plan to understand a more abstract layer: the syntax. We can compare our work to language analysis [5] with different layers of abstraction. We can capture data from a device and transform these raw data to symbolic words. With words, we define a lexicon. The next step is to analyze the structure or syntax. There are three main advantages to rise up over simple lexicon analysis. First, for the recognition part: even with the multi-freedom degrees of our body, we cannot do all possible movements between two gestures. There is an inter-relation of the words [6]. This means that if we know the syntax we can eliminate false-positive recognition. Second, for the context understanding: the same gesture can be repeated at different time during the play. We can distinguish the current position in the scenario with the current context or syntax. Third, behavior understanding: rules can help to interpret the reaction of the user. For example: repetition of the same gesture, pair of gestures.

Conclusion. In this paper we presented a new method for perceiving gestures made by a real actor in real-time. These gestures are recorded during rehearsals and pre-processed to generate a signature. During the performance, the recognition system uses this signature and compares it with its observation. The basics of the gestural recognition system are achieved at this point. We have a versatile system that can record, load or unload gestures during execution time. Now we can increase our system with the integration of the syntax analysis.

References

1. Campbell, L.W.; Bobick, A.: Recognition of human body motion using phase space constraints. Computer Vision, 1995. Proceedings., Fifth International Conference on (1995) 624–630
2. Bobick, A., Wilson, A.: A state-based technique for the summarization and recognition of gesture. iccv **00** (1995) 382
3. Vasilescu, M.A.O.: Human motion signatures: Analysis, synthesis, recognition. In: ICPR '02: Proceedings of the 16 th International Conference on Pattern Recognition (ICPR'02) Volume 3, Washington, DC, USA, IEEE Computer Society (2002) 30456
4. Forbes, K., Fiume, E.: An efficient search algorithm for motion data using weighted pca. In: SCA '05: Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on Computer animation, New York, NY, USA, ACM (2005) 67–76
5. Kendon, A.: How gestures can become like words. In Poyatos, F., ed.: Cross-Cultural Perspectives in Nonverbal Communication. (1988) 131–141
6. Goldin-Meadow, S., Butcher, C., Mylander, C., Dodge, M.: Nouns and verbs in a self-styled gesture system: What's in a name? Cognitive Psychology **27**(3) (December 1994) 259–319