HAL
open science

Creating a library of gestures with variants

André-Marie Pez, Catherine Pelachaud, Zhenbo Li, Patrick Horain

To cite this version:

André-Marie Pez, Catherine Pelachaud, Zhenbo Li, Patrick Horain. Creating a library of gestures with variants. GW 2009 : 8th International Gesture Workshop, Feb 2009, Bielefeld Germany. pp.1 - 2. hal-01369596

HAL Id: hal-01369596

https://hal.science/hal-01369596v1

Submitted on 21 Sep 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Creating a Library of Gestures with Variants

André-Marie PEZ[1,2], Catherine PELACHAUD[1,2], Zhenbo LI[3], Patrick HORAIN[3],

[1] IUT de Montreuil, Université de Paris 8,
140 rue de la Nouvelle France, 93100 Montreuil, France

[2] Institut TELECOM, TELECOM ParisTech
46 rue Barrault, 75634 Paris Cedex 13, France

[3] Institut TELECOM; TELECOM & Management SudParis,
9 rue Charles Fourier, 91011 Evry, Cedex, France

andremarie.pez@telecom-paristech.fr, catherine.pelachaud@telecom-paristech.fr
zhenboli@gmail.com, Patrick.Horain@IT-SudParis.eu

Keywords: communicative behavior library, variety of gestures, tracking

1. Introduction

Avatars are virtual self representation. They evolve in a 3D world and interact with other virtual entities on our behalf. Avatars are animated by their human counterpart. One difficulty is the control of their behaviors. Selecting behaviors from a menu or by using icons is tedious. Moreover the avatar animation is not lively as it moves only on commands. Some attempts have been made to alleviate the user's role and endow avatars with some autonomy (Vilhjálmsson, 2003). These approaches are promising but do require computational time to parse what the user aims to say and compute the appropriate nonverbal behaviors that accompany this text. In our project we let the user fully control her avatar. The user is filmed through a simple web cam. The avatar reproduces the gestures of the user that have been tracked Horain et al, 2005).

Our tracking algorithm uses statistical gesture models as guides to constrain and disambiguate 3D tracking. To increase the precision of the tracking model, we need to have it trained on appropriate human gesture databases. Most of existing databases gather data on human actions such as running, jumping and the like. There are very few databases that groups communicative gestures. One difficulty is the variety that exists for communicative gestures. Even though there are some gestures with defined shape that can be linked to precise meaning, most of the communicative gestures are creative (Poggi, 2007); that is most of them are created on the spot. Building a library of such gestures is a tremendous enterprise. Thus, gesture shape can vary tremendously depending on their meaning and discourse context. To overcome this problem, we have decided to work on a given discourse context. Our context is interview recruiting. We gathered data of 4 dyads lasting each around 30mn.

In this paper we present how work toward building such a library of communicative gestures with the aim to use it to train the tracking algorithm. We describe also how we encompass user's communicative distinctiveness in our library.

2. Creating a library of gestures

Several dictionaries of emblematic gestures have been gathered. They are bound to a given culture. These dictionaries provide detailed information about the gesture shape and its associated meaning. Other attempts have looked to describe gesture shape in association with their physical meaning (Calbris, 1990). Raised hand with palm facing one's interlocutor carries the meaning to stop something or somebody to do something. It can be viewed as symbolizing a wall between both interlocutors.

In our work we are interested at all types of gestures as we aim to track any hand and arm movements done while communicating. We do not aim to recognize gestures or to interpret them. We are interested in detecting their shape and following their movement in view in reproducing them by the avatar. This reproduction does not require understanding the meaning of the gestures. With such an aim we decided to consider only one feature of the gestures: their shape.

2.1 Gesture variability

While communicating, people show large variability not only in their intentions but also in their way of expressing them. One can be characterized by a signature, a style one carries along in, basically, all circumstances (Gallaher, 1992).

We have developed a model of distinctive agent that encompasses variability in the modality preference used to communicate a given intention and on the behavior expressivity (Mancini & Pelachaud, 2008). In particular when modulating the last set of parameters, the agent can display gestures more or less extent, more or less fast and powerful, etc. These variations occur at the level of execution of behaviors and not on the type of behaviors to be displayed.

2.2 Gesture clustering

In our corpus we have gathered data from 8 interlocutors. The data was annotated using ANVIL (Kipp, 2001). Around 800 communicative gestures were found in the data. We gathered them into classes of gestures looking alike in their shape and movement.

Our aim is to create a library to train our gesture tracking algorithm (Horain et al, 2005). To ensure robustness of our tracking algorithm over a large population of users, we have enhanced the library of gestures of the training phase with gesture variability using the model of distinctive agent (Mancini & Pelachaud, 2008). The gesture of each class found in our corpus is reproduced by a virtual agent (see Fig. 1). The library contains the reproduced gesture as well as the same gesture with different expressivities. Thus each gesture in a class and its variations are present.

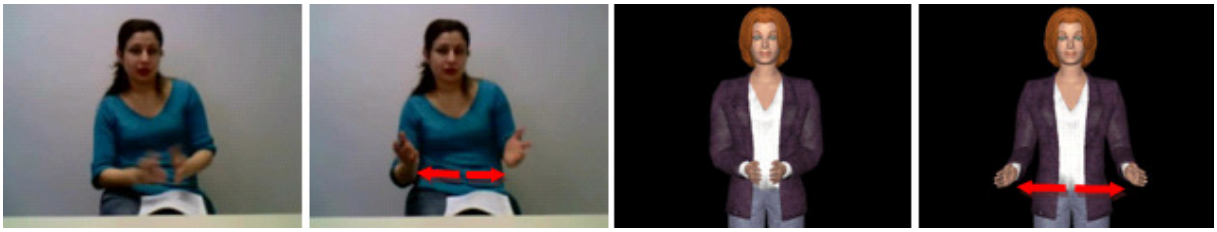

Fig.1: reproduction by the virtual agent of a gesture belonging to a gesture class.

3 First results

We have trained our tracking algorithm on the library of gestures done by our virtual agent. The first result shows an increase in the tracking of gestures; that is, our algorithm shows qualitative better tracking results.

We are currently increasing the class of gestures of the library by enhancing the animation capabilities of our virtual agent. In particular, our agent is now able to perform asymmetric and circular hand and arm motions.

In the very near future, we hope to show a refinement in the tracking algorithm when being trained with the enlarged library.

References

1. Calbris, G., The semiotics of French gestures, University Press, Bloomington: Indiana, 1990.
2. Gallaher, P. E., Individual differences in nonverbal behavior: Dimensions of style. Journal of Personality and Social Psychology, 63(1):133{145, 1992.
3. Horain, P., J. Marques Soares, P. K. Rai, A. Bideau, “Virtually enhancing the perception of user actions”, *Proceedings of the 15th International Conference on Artificial Reality and Telexistence (ICAT 2005)*, December 5th-8th, 2005, University of Canterbury, Christchurch, New Zealand, pp. 245-246.
4. Kipp M., “Anvil - A Generic Annotation Tool for Multimodal Dialogue”, Proceedings of the 7th European Conference on Speech Communication and Technology (Eurospeech), pp. 1367-1370, Aalborg, September 2001.
5. Mancini, M., C. Pelachaud, Distinctiveness in multimodal behaviors, Seventh International Joint Conference on Autonomous Agents and Multi-Agent Systems, AAMAS'08, Estoril Portugal, May 2008.
6. Poggi, I., Mind, Hands, Face and Body. A Goal and Belief View of Multimodal Communication, Weidler Verlag, Körper, Zeichen, Kultur, n° 19, 2007.
7. Vilhjálmsson, H, Avatar Augmented Online Conversation, Ph.D. thesis, Media Arts and Sciences, Massachusetts Institute of Technology, Media Laboratory, Cambridge, MA, 2003.