# Representation of human postures for vision-based gesture recognition in real-time*

Antoni Jaume-i-Capó, Javier Varona, and Francisco J. Perales

Unitat de Gràfics i Visió per Ordinador
Departament de Ciències Matemàtiques i Informàtica
Universitat de les Illes Balears
Edifici Anselm Turmeda
Ctra. de Valldemossa km 7,5
(07122) Palma de Mallorca, Spain
{antoni.jaume,xavi.varona,paco.perales}@uib.es
http://dmi.uib.es/~ugiv/

**Abstract.** In this paper, we present a Vision-Based Interface guided by the user gestures. The advantage of our system is that it is built over a motion capture system that recovers the body joints positions of the user's upper body in real-time. From the computed joints positions we make this data spatially invariant by normalizing limbs positions and sizes, only using the limbs orientations. From limbs orientations, the user posture is represented by an appropriate representation of all the limbs in a histogram. Cumulating the posture histograms we represent a gesture in a temporal invariant form. Finally, using this gesture representation, the performed gestures are classified for generating the desired computer events in real-time.

**Key words:** Vision-Based Gesture Recognition

## 1 Representation of human postures

Previously to recognition, the user's movements are obtained through a real-time vision-based motion capture system [1]. Using the computed 3D positions of the involved body joints, we address the main problems in the gesture recognition challenge: temporal, spatial and style variations between gestures. Temporal variations are due to different gesture speed between different users. Spatial variations are due to physical constraints of the human body such as different body sizes. Style variations are due to the personal way in which a user makes its movements.

To cope with spatial variations we represent each body limb by means of a unit vector. The idea is to represent the user's body posture as a feature vector

* This work has been supported by the Spanish M.E.C. under project TIN2004-07926 and the Ramon y Cajal fellowship of Dr. J. Varona.

composed by all the unit vectors of the user's limbs. Formally, the representation of the orientation of a limb, $l$, is

$$\mathbf{q}^l = (u_x^+, u_x^-, u_y^+, u_y^-, u_z^+, u_z^-), \tag{1}$$

where $u_x^+$ and $u_x^-$ are respectively the positive and negative magnitudes of the $x$-component of unit vector, $u_x$, note that $u_x = u_x^+ - u_x^-$ and $u_x^+, u_x^- \geq 0$. The same applies for components $u_y$ and $u_z$. Therefore, we build a histogram of limbs orientations which represent the complete user's limbs orientations. We propose two forms to build the histogram. The first one is by cumulative limbs orientations and the second one is by linking limbs poses. The main difference between the two representations depends on the considered gesture set. The cumulative representation is more robust to tracking errors, but the set of recognized gestures is much reduced. On the other hand, the linked representation allows the definition of more gestures, although it is more sensible to errors in the estimation of the limbs orientations.

Temporal variation is managed using a temporal gesture representation. A gesture is composed by several body postures, and then the gesture representation feature vector is composed by the cumulative postures involved in the gesture.

An important goal of this work is that the human-computer interaction should be performed using natural gestures. A gesture is natural depending on the user experience. To cope the style variations, before the recognition process starts the system asks the user to perform several of the allowable gestures in order to build a user's specific gestures models in real-time.

Finally, for the recognition phase, we choose the Bhattacharyya coefficient in order to complete the process, as a distance for comparison between the current gesture and a gesture model.

## 2 Conclusions

The most important contribution of this work is that we have defined two gesture representations, capable to cope with variations between gestures in different users and performances, making also possible the recognition in real-time. The complete system has been tested in a real-time application, a gesture-based videogame control, and the results obtained state that the presented approach for gesture recognition performs well (**84.95%** for the cumulated representation and **87.69%** for the linked representation).

## References

1. Boulic, R., Varona, J., Unzueta, L., Peinado, M., Suescun, A., Perales, F.: Evaluation of on-line analytic and numeric inverse kinematics approaches driven by partial vision input. Virtual Reality **10**(1) (2006) 48–61