Improving human movement recovery using qualitative analysis

Barbara Mazzarino[1], Manuel J. Peinado[2], Ronan Boulic[3], Marcelo Wanderley[4] & Antonio Camurri[1]
[1]InfoMus Lab-DIST- Università degli Studi di Genova, Italy
[2]University of Alcala, Spain
[3]Ecole Ploytechnique Fédérale de Lausanne, Switzerland
[4]McGill University, Canada
Barbara.Mazzarino@unige.it

This paper focuses on the use of qualitative analysis of human gesture for improving the believability of recovered human movements from a small number of sensors. In particular, the pilot study presented here compares expressive features extracted from a real human performance and the rendered animations from two reconstructions based on the measured sensor data.

Introduction

The work here described is part of the research carried out in the framework of the European Project EU-IST Network of Excellence ENACTIVE (Enactive Interfaces, strategic objective Multimodal interfaces). The present study aims at identifying which expressive factors are critical to explain the believability of reconstructed movements when displayed on virtual characters.

The underlying idea is to apply techniques previously developed by DIST for human gesture analysis also to virtual characters. In this way, it is possible to evaluate the quality of the motion of virtual characters by comparing with the corresponding real humans from which the motion has been extracted, and thus to identify the key differences between the same motions performed by a human and by the character.

The main objective of this work is to evaluate whether both the real and the virtual movements convey the same qualitative and expressive information, and to compare different rendering techniques with the aim of evaluating which method better maintains the high level content (Camurri et al. 2005) conveyed by a specific motion. In addition we want to minimize the number of sensors weared by the performer so that full-body motion capture can be integrated within the interface of a broader range of applications.

Methodology and Results

A clarinet player was recorded during a performance, with one camera in lateral position. Sensors were fixed on the musician's body in order to accurately track body motion for the subsequent 3D reconstruction, using an OptoTrack System, 100Hz sampling frequency. It is important to notice that the sensors on the musician's body were localized just on the right side of the body, on the opposite side of the video camera.

Two main factors were identified as conveyor of unbelievably in the reconstructed motion: (i) the occupation of the surrounding space with local unbelievable postures, and (ii) the low amount of fluidity of the motion. In order to measure such two factors, first we extracted the quantity of movement of the motion from all the movies. A snapshot of this work is visible in Figure 1.

The analysis showed that a too strict constraint was imposed on the center of mass position, i.e. it was only allowed to move freely on the vertical axis passing through a mid-ankle point. This led to the generation of a second set of reconstructed motions where the movement of the center of mass was guided by the recorded instrument movement along the forward-backward direction (but was prevented to project outside the supporting polygon). The feet constraints were also adjusted to better reflect the musician posture.

For this study the fluidity was considered in two different ways:
- as a results of the segmentation of the motion (many fragment of the motion means an overall reduction of the motion fluidity); in Table 1 (second column) it is possible to see an improvement of this segmentation in the subsequent set of movies.

- as a measure of the agreement of the motion of different parts of the body. In this study we considered the agreement between upper and lower parts of the body. This is a real time measure of the fluidity and it is more related to the observer point of view. If it is present a relevant disagreement between subparts motion, the ratio of the measured energy generates some peaks. The higher the value of these peaks and the longer their duration, the more disagreement is present: this means less fluidity. In Table 1 (last column) it is possible to see the major results of fluency improvement.

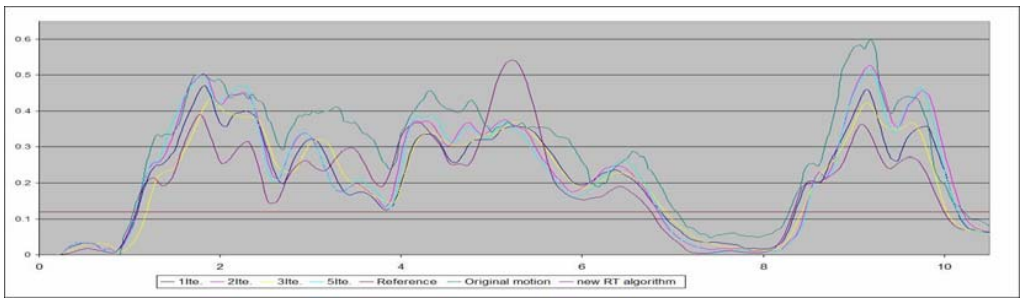

Figure 1. Analysis of the initial part of the clarinet player movements: the focus is on the first two gestures (the attack of the music performance). In the upper graph are represented the motion bells extracted from the real-time algorithms and the real movie are represented.

Table 1. This table gives an overview of the results obtained from the analysis of fluency

	Analysed movie	Number of motion phases	Average Value of Ratio	Max value
	REAL	10	4.75	56.18
First set of movies	Off Line Algorithm	17	31396.76	124,431.00
	Real Time algorithm	16	1571.12	59,626.70
Second set of movies	2 Iteraction Worst case	13	19.21	1,423.22
	3 Iteraction best case	13	18.06	776.17
	10 Iteraction , off line	13	20.45	1,444.72
Last set of movies with postures correction	Off Line Algorithm	12	5.52	195.02
	Real Time algorithm	16	10.78	169.64

The output of the first analysis provided precious hints to the conceptors of the Inverse Kinematics reconstruction technique for reducing the differences in quality. This led to propose an improved balance constraint. The resulting set of animations has been again analyzed. This new analysis demonstrated an increase in the quality of the virtual motion. A particularly significant criteria is the fluidity cue of motion, that in the first set motion was reduced due to a disagreement between motion of upper and lower body part. In the new set of motion this disagreement was significantly reduced and the related reconstructed motions appeared more fluent. Future tests will examine how to model, in the reconstructed human, also lateral movements of the center of mass to reflect the observed weight transfer in real musician performances.

Using techniques developed for human gesture analysis, we showed how it is possible to extract high-level motion features from reconstructed motion and to compare them with the same features extracted from the corresponding real motions. Moreover, these features allow a qualitative comparison between different rendering techniques. This resulted in a precious complementary tool to believability studies based on analyzing solely viewer feedback through questionnaires.

References

Camurri, A., De Poli, G., Leman, M., & Volpe, G. (2005). Toward Communicating Expressiveness and Affect in Multimodal Interactive Systems for Performing Art and Cultural Applications. *IEEE Multimedia Magazine, 12(1),* 43- 53, IEEE CS Press, 2005.

Peinado, M., Herbelin, B., Wanderley, M., Le Callennec, B., Boulic, R., Thalmann, D., & Méziat, D. (2004). Towards Configurable Motion Capture with Prioritized Inverse Kinematics. Proceedings of International Workshop in Virtual Rehabilitation (IWVR/SENSOR04), Lausanne Sept. 16-17th.