A Multi-Modal 3D Capturing Platform for Learning and Preservation of Traditional Sports and Games

François Destelle, Amin Ahmadi, Kieran Moran and Noel E. O'Connor
INSIGHT, Dublin City University, Ireland
francois.destelle@dcu.ie

Nikolaos Zioulis, Anargyros Chatzitofis, Dimitrios Zarpalas and Petros Daras
ITI-CERTH
Thessaloniki, Greece

Luis Unzueta, Jon Goenetxea, Mikel Rodriguez and Maria T. Linaza
Vicomtech-IK4
San Sebastián, Spain

Yvain Tisserand, Nadia Magnenat Thalmann
MIRALab, University of Geneva, Switzerland

ABSTRACT

We present a demonstration of a multi-modal 3D capturing platform coupled to a motion comparison system. This work is focused on the preservation of Traditional Sports and Games, namely the Gaelic sports from Ireland and Basque sports from France and Spain. Users can learn, compare and compete in the performance of sporting gestures and compare themselves to real athletes. Our online gesture database provides a way to preserve and display a wide range of sporting gestures. The capturing devices utilised are Kinect 2 sensors and wearable inertial sensors, where the number required varies based on the requested scenario. The fusion of these two capture modalities, coupled to our inverse kinematic algorithm, allow us to synthesize a fluid and reliable 3D model of the user gestures over time. Our novel comparison algorithms provide the user with a performance score and a set of comparison curves (i.e. joint angles and angular velocities), providing a precise and valuable feedback for coaches and players.

Categories and Subject Descriptors

I.4.5/7 [**Computing Methodologies**]: Image Processing and Computer Vision—*Reconstruction, Feature Measurement*

General Terms

Capture, Tracking, Measurement, Comparison.

Keywords

Motion capture, Sensor Fusion, Vision, Motion Retargeting.

1. INTRODUCTION

Traditional Sports and Games (TSG) are a strong part of a society's identity and a strong mechanism for the promotion of cultural diversity. European examples include Gaelic games (Hurling, Football, etc.) in Ireland, Basque Pelota (and variants) in Spain, among many others.

Our project [1] aims to preserve and promote these TSG, that may be under the threat of disappearing and losing their social and community value. This demonstration software platform implements three different scenarios: *Play&Learn*, *Coach&Train* and *Interact&Preserve*. The first one is dedicated to children at home and allows the user to learn TSG gestures and compare himself to local existing heroes. The second one is more technical and aims to provide coaches with a precise measurement tool in order to improve an athlete's performance. This scenario is close to the first one but it uses up to 4 Kinect sensors and up to 9 Wearable Inertial Measurement Units (WIMUs). The last one is intended for a museum or other exhibition place. It preserves the TSG sporting gestures into an accessible online database of captured gestures.

A key feature of our framework is the cost to achieve our goals. Given the budget-constrained conditions under which most TSG clubs operate, it was imperative that our aims be achievable in a cost-effective manner. To this end, we used open source software wherever possible, putting the creation of such digital exhibitions within reach of even modest budgets.

2. SYSTEM DESCRIPTION

The platform that includes all our scenarios is embedded in a single application, and the game engine Unity [2] was employed to allow us to create the 3D interactive environment.

Motion Capture Framework.

Our main platform handles four Kinect sensors and up

[1]RePlay project: www.fp7-replay.eu
[2]www.unity3d.com

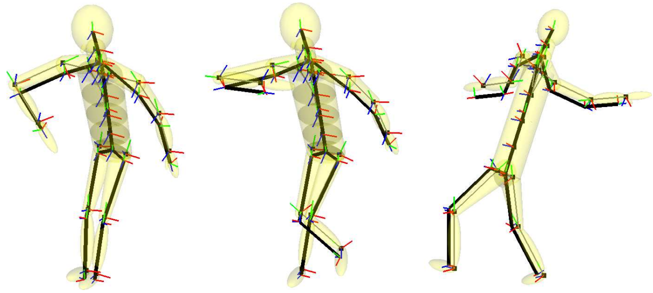

Figure 1: Transformation of a captured fused skeleton (black stickman) to the normalized avatar (yellow avatar) for motion retargeting.

to nine WIMUs. The multi-Kinect setup enables the creation of the athlete's full body 3D reconstruction, providing a photo-realistic visualization of the performed skill [3]. Each Kinect version 2 sensor requires a different machine, thus involving a network design. This client-server design uses a fast and reliable TCP/IP implementation, where each receiver machine is a client and where the main machine running the Unity 3D main software fulfills the server role. Moreover, as one machine alone can't easily handle more than 7 paired Bluetooth devices, especially at high framerate (100Hz or 256Hz), we use this multi-machine framework to also handle the Bluetooth reception of inertial data. Each client receives the WIMU data packet via Bluetooth at a fixed framerate, this packet is then used to compute the resulting local orientation changes of each sensor, using the efficient Madgwick rotational estimator [4].

Multi-modal Sensor Fusion.

The first step of our capturing pipeline is the fusion of the captured Kinect skeleton and the rotational information computed from the WIMUs. The whole process is described in [1], it uses one frame from the Kinect tracking algorithm to create a skeleton reference. This reference is then animated using solely the rotational information provided by the WIMUs. The fused Kinect-WIMUs skeleton provides a much more fluid, reliable and high-framerate model of the human gesture than the Kinect tracking algorithm.

Real-time Motion Retargeting.

The system could be used by a wide range of people, so a high variability of body part sizes is expected. Thus, we need to set a common integration framework for both the captured skeleton and the reference one from the elite ('gold standard') athlete to which they are compared. This representation constructs a normalized skeleton from any captured user, through real-time motion retargeting. The motion retargeting could be summarized as inverse kinematics (IK) solvers integrated with an objective function optimization for space-time constraints [2]. This kinematic representation of the virtual human has more segments and joints than those tracked by the sensors, see Fig. 1. It also insures a constant size for each bone and provides biomechanical constraints on the joint orientations to finally achieve a natural preservation of the performed skill properties.

[3] 3D Photorealistic reconstruction: www.youtube.com/watch?v=gu72_g2GOQk

[4] Sebastian O.H. Madgwick, 30/09/2010: www.x-io.co.uk/open-source-imu-and-ahrs-algorithms

Figure 2: Pictures taken from our live demo, a Gaelic (left) and a Basque player (right).

Motion Comparison Algorithms.

The last part of our pipeline is the evaluation of the user's gesture against the one performed by the elite athlete. It compares two skeletons obtained by our IK algorithms applied to both the reference skeleton captured using a full motion capture system (Vicon, UK) and the skeleton obtained by our sensor fusion. We extract a number of geometric features including hierarchical rotations, linear/angular velocities, joint angles and kinetic energy. A dynamic time warping technique is used for both global and local temporal alignment of the multi-spectral signal. We then use a novel weighted scheme to obtain a global scoring from the overall performance of the user. These important weights are defined for each gesture by real coaches and biomechanical experts. While the scoring is a global performance indicator, our graphical interface also provides a rendering of all the local information for each joint of the skeleton over time.

3. DEMONSTRATION

The demonstration will consist of the *Coach&Train* scenario, involving one Kinect sensor and nine WIMUs. Two laptops will be used in order to receive the information coming from the WIMUs via Bluetooth. The RePlay exhibition will be showcased on a main large screen in front of the user, while the Kinect camera will be placed below it. A short video detailing the main RePlay platform and showing graphically the steps involved will also be displayed.

4. ACKNOWLEDGMENTS

The work presented in this paper was supported by the European Commission under contract FP7-601170 RePlay.

5. REFERENCES

[1] F. Destelle, A. Ahmadi, N. E. O'Connor et al. Low-cost accurate skeleton tracking based on fusion of kinect and wearable inertial sensors. In *EUSIPCO '14*, pages 371–375, Sept 2014.

[2] L. Unzueta, J. Goenetxea, M. Rodriguez, and M. Linaza. Viewpoint-dependent 3d human body posing for sports legacy recovery from images and video. In *EUSIPCO '14*, pages 361–365, Sept 2014.