3DLIVE: A MULTI-MODAL SENSING PLATFORM ALLOWING TELE-IMMERSIVE SPORTS APPLICATIONS

Benjamin Poussard[1], Simon Richir[1], Jukka Vatjus-Anttila[2], Stylianos Asteriadis[3], Dimitrios Zarpalas[3], Petros Daras[3].

[1]Arts et Métiers ParisTech, LAMPA, 2 boulevard du Ronceray, Angers, France.
{benjamin.poussard, Simon.Richir}@ensam.eu
[2]Cyberlightning ltd. Hallituskatu 13-17 E 48, FI-90100 Oulu, Finland.
jukka.vatjus-anttila@cyberlightning.com
[3]Information Technologies Institute, Centre for Research and Technology Hellas,6th klm Charilaou-Thermi, GR-57001, Thessaloniki, Greece.
{stiast, zarpalas, daras}@iti.gr

ABSTRACT

3DLive* project is developing a user-driven mixed reality platform, intended for augmented sports. Using latest sensing techniques, 3DLive will allow remote users to share a three-dimensional sports experience, interacting with each other in a mixed reality space. This paper presents the multi-modal sensing technologies used in the platform. 3DLive aims at delivering a high sense of tele-immersion among remote users, regardless of whether they are indoors or outdoors, in the context of augmented sports. In this paper, functional and technical details of the first prototype of the jogging scenario are presented, while a clear separation between indoor and outdoor users is given, since different technologies need to be employed for each case.

Index Terms— Augmented Sports, Motion capture, Tele-Immersion, Activity assessment, 3DLive

1. INTRODUCTION

1.1. Multi-modal sensing for motion analysis in Augmented Sports applications

Since the beginning of the 2000s, Augmented Sports applications are witnessing an increasing success in different types of audiences. A lot of effort is being directed towards engaging stakeholders in latest technologies aiming at enhancing User Experience through novel tools of activity tracking, assessment and virtual environments. The purpose of Augmented Sports applications can vary depending on the activity and the targeted stakeholders, as it often involves people from the sports market (sportsmen and sports companies), while the game industry (gamers and game companies) is also entering the field through gamified applications targeting certain sports.

While there is a large gamut of applications, only few sensing technologies have made their way to the market, mainly due to usability and cost-related reasons. Below is a table listing representative Augmented Sports R&D projects, each utilizing different sensing technologies:

Project	**Sport**	**Sensing Technologies**
Virku [1]	VTT	Record rotation of an indoor bicycle wheel.
Kick Ass Kung Fu [2]	Martial Arts	Computer vision and background subtraction. (Cameras).
Sports Over a Distance [3]	Soccer	2 cameras allowing the tracking of ball shot against a wall.
Bouncing Stars [4]	Ball game	IR camera and IR markers on a ball. Accelerometers inside the ball.
Tilt'n'Roll [5]	Skate-boarding	Inertial Measurement Units on a skateboard + GPS tracking helping the detection of tricks.
Virtual Archery [6]	Archery	IR Markers on the bow and the arm + IR Cameras
ExoInterfaces [7]	Any	Sensing of muscle efforts thanks to belt and motor on the arm.
Flying Sports Assistant [8]	Any outdoors	ARDrone following sportsmen thanks to camera colour tracking.
Enhanced interactive Gaming [9]	Any indoors	Tracking of the body with Kinect depth sensors.
Realtime sonification [10]	Ski	Force sensors under feet and audio feedbacks.

Table 1. Augmented Sports Projects in the last 15 years.

The sensing solutions could be categorized into two groups, depending on the application environment: indoors and outdoors.

Regarding projects related to indoor sports, the physical area that needs to be analysed is often limited, thus is well controlled. Consequently, most of the applications use cameras and employ customized strategies for motion detection and tracking. Infrared cameras are among the most com-

* This work was supported by the EU funded project 3DLive, GA 318483.

monly used solutions in Virtual Reality systems and their usability has been showcased in the Bouncing Stars and Virtual Archery projects [4-6]. The underlying principles allow robust tracking due to markers attached on athletes. Several LEDs are used, emitting infrared light, which is subsequently reflected by the markers that can be tracked, with the help of proper cameras. While this solution offers good accuracy, it can be cumbersome for the users. Colour tracking can also be used for determining 3D positions, a solution which is usually adequate in many applications where clearly defined objects need to be tracked (e.g. a ball). One or several cameras can be used (e.g. Sports over a Distance [3]). Another family of techniques consists in background subtraction [2] with human silhouettes forming part of the foreground, which is subsequently analyzed. Such an approach has been followed in Kick Ass Kung Fu project, where users fight against virtual 2D characters. In addition to cameras, indoor Augmented Sports often use depth sensors in order to detect the positions of moving elements in a scene [9].

On the other side, outdoor Augmented Sports have to meet different constraints, which usually relate to spatial configurations. Wearable sensors are usually required to capture outdoor users. The use of cameras is still a possibility if implemented with the use of dynamic systems, although it restricts usability. One of the most representative examples is the "Flying Sports Assistant" project [8], where an embedded camera can track user's trajectory using colour information. On the contrary, an increasing number of applications is making use of Inertial Measurement Units, through which acceleration and orientation information of moving elements can be retrieved, among other types of information. Typical example is the "Tilt'n'roll" project [5], which uses inertial sensors to capture tricks performed by a skateboarder. Force or Pressure sensors are also used in board sports, in order to get the overall posture of the user. For instance, the "Real-time Sonification" project generates sound feedbacks to skiers who try to align their centres of gravity. In addition to the above, common sensors like GPS to track the positions of users in outdoor settings have started to be employed in the last years, mainly thanks to their widespread availability on common devices (e.g. smartphones).

1.2. The 3DLive project

Within the context of Augmented Sports, the 3DLive project aims at co-designing, developing and experimenting on a Tele-immersive platform allowing several users to interact with each other, regardless of whether they are indoors or outdoors (Fig. 1). The users are virtually immersed in a shared mixed reality environment, where they can communicate, see each other and perform a specific sport, namely jogging, skiing and golfing. One of the main objectives of the project is to investigate the proper Quality of Service (QoS) and Quality of Experience (QoE) metrics when users are fully immersed in mixed environments, and to identify new potential requirements of Future Internet (FI) research. To achieve the goals of the project, a first platform has been implemented, incorporating already several technical aspects intended to maximise the sense of smooth and unconstrained co-existence in shared virtual spaces. Moreover, 3DLive is making extensive use of input from potential end users, who have brought a series of requirements in terms of User Experience maximization, accessibility, usability and the raise of issues regarding safety and "edutainment".

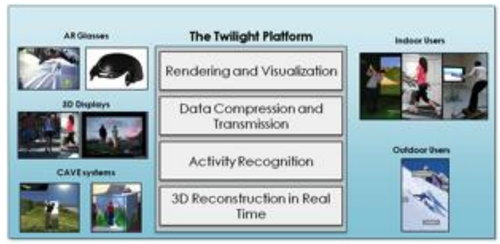

Fig. 1. The 3DLive concepts allowing Tele-Immersive Augmented Sports.

2. TELE-IMMERSION IN SPORTS ACTIVITIES

2.1 The 3DLive Platform

The first prototype of the platform allows several users to take part in the 3DLive experience for the three scenarios. As mentioned above, it already incorporates a series of modules intending to, either directly or indirectly, contribute to the maximization of sense of co-existence with remote peers. Namely, these modules are: 1) Environment reconstruction for measuring weather-related data at the outdoor site; these measurements are transferred to the virtual environment and shape the visual effect accordingly. 2) A human motion reconstruction module takes care of the three dimensional reconstruction of real users' motion and leads to the replication of human motions on dedicated avatars. 3) An activity recognition module uses human motion in order to recognize and assess the sport activity of users. 4) A voice communication module handles the social voice interactions between the users. 5) A compression and transmission module collects data flows, compresses them if needed and transmits them through the 3DLive servers, in order to achieve efficient use of network resources, guaranteeing, at the same time, high levels in terms of User Experience. 6) A Rendering and Visualization module which retrieves data flows and produces the final, mixed environment, common for all users, regardless of the platform they use (Android mobile phones and desktops have been tested, so far).

Since the first prototype release, the 3DLive platform has achieved high levels of sense of tele-immersion, mainly thanks to its multimodal approach. The above are illustrated by thoroughly describing the representative example of the

Jogging case, with ski and golf following similar technical and functional principles. Details on human motion estimation and activity recognition are given, in order to highlight the impact of the chosen strategies, on achieving high levels of sense of co-existence and immersion.

2.1 The Jogging Scenario

Similar to the ski and golf cases, indoor and outdoor joggers co-exist in the virtual space of the shared experience of jogging. The indoor user's activity is tracked and employed in the gameplay, through the use of an avatar, which replicates the user's movements and runs at the same speed (Fig. 2). On the other side, an outdoor user is tracked thanks to a GPS embedded in a Smartphone and is, at the same time, reconstructed in the mixed environment, which is a virtual representation of the city of Oulu (Fig. 3).

Fig. 2. An indoor 3DLive jogger running on a treadmill in Greece, reconstructed in the shared virtual environment.

.

Fig. 3. An outdoor 3DLive jogger running in Finland, reconstructed in the shared virtual environment.

In the first version of the prototype, the indoor user can visualize the experience on a wide screen while running on a treadmill. On the other side, the outdoor user has only one Smartphone and no visualisation system allowing him to see his remote peer (future prototypes will make use of AR glasses for this scope).

3. MULTI-MODAL MOTION CAPTURING

3.1 Indoor Motion Capturing

For raw joint positions and orientation extraction, the OpenNi 2 [13] framework has been chosen. OpenNi 2 delivers a set of joints for the whole body, in the form of skeleton, at a rate of 30 frames per second.

3.1.1 Skeleton merging

One of the core components of the 3DLive architecture involves accurate extraction of human motion parameters. With one of the toughest bottlenecks being related to non-wearable sensors, though, due to their significant sensitivity to occlusions or self-occlusions (Fig. 5). Trackers fail to make proper detections and thus fall into local minima, from which most of the times is difficult to recover. For the above reasons, in 3D-Live a multiple Kinect sensors approach has been adopted: each user is being tracked by two or more Kinect sensors, from different points of view. A population of n candidate positions $p_{j,n,k}$ around each joint j, for every Kinect sensor k is initialized, and each of these units is translated accordingly in the 3D space. In subsequent frames, each point $p_{i,n,k}$ is assigned a weight $w_{j,n,k}$, which depends on two factors (equation 1): the distance of each candidate position from the corresponding observation $s_{j,k}$ (the actual measurement of Kinect sensor k for joint j) and a series of confidence values, modelled as non-linear functions of heuristics, namely the joints' expressivity, their distances on the z-axis from the corresponding values of sensor's depth map and the possibility that $s_{j,k}$ belongs to a limb with expected posture. These factors are non-linearly fused by following a Mamdani Fuzzy-Inference-Scheme [14] delivering an overall confidence equal to $\lambda_{j,k.}$. Confidence and distance values define weight $w_{j,n,k}$:

$$w_{j,n,k} = \sum_k \{\lambda_{j,k}\}^{-1} e^{-|s_{j,k}-p_{j,n,k}|}$$

After each iteration t, a roulette wheel selection scheme [15] is followed, encouraging more promising positions. The algorithm converges to a few candidates after a number of iterations, which give the final joint position.

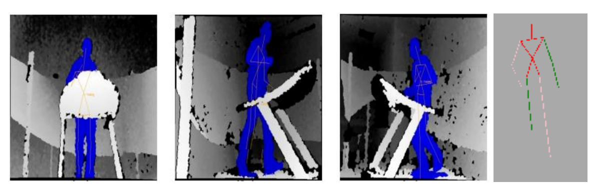

Fig. 4. Human Skeleton from 3 different angles with occlusion, and the resulting skeleton, using joints from proper sensors.

3.1.2 Recognizing and Evaluating Activity in 3D-Live

3DLive platform provides the user, not only with a sense of (tele-)immersion, but can also act as a tool for evaluating one's own ability to perform an exercise and, at the same time, receive suggestions for improving in future interactions. To this aim, a real-time Activity Recognition and Evaluation module has been developed, offering the following functionalities:

Estimate indoor user's speed: Prior to evaluating a jogger's activity, the type of activity has to be defined. In 3D-Live, we make an automatic distinction among four basic categories: 1. Running, 2.Jogging, 3.Walking 4.Standing still. To this aim, a training set of known pairs *(z_f, u)* has been recorded, with z_f being feet's average depth position, and *u* corresponding to actual treadmill speeds. A first-order regression model, mapping first order positive derivatives *dz/dt* to *u* is found using least-squares and this is subsequently used, for unknown data, in order to calculate user speed in a generic way and avoid complex solutions for extracting it from the treadmill itself. Thresholds are defined in order to infer whether user is standing still, walking, jogging or running and, for the two last cases, proper heuristics and fuzzy machines are triggered for activity evaluation.

Evaluate activity performance: Multi-skeleton motion extraction delivers joints positions, which can subsequently be used in order to make proper suggestions to the user regarding posture (arms, bending of knees, leaning, etc.). These rules may differ regarding whether the user is jogging or running and, transforming these variables into fuzzy sets may give an overall estimate of user's performance, at any time during interaction. Overall suggestions and scores may be given at the end of the interaction. Figure 6 is representative of the feedback user receives while interacting with the platform.

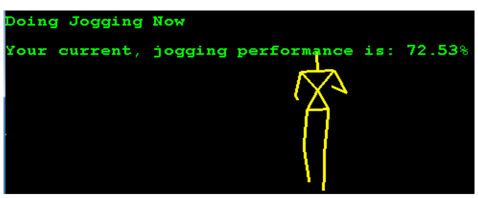

Fig. 5. Real-time activity recognition and evaluation.

3.2 Outdoor motion capturing

In this first prototype two components are required to replicate the outdoor jogger into the virtual venue. Standard issue Android 4.0+ smartphone with GPS-sensor and a webservice to handle the communication between the android application and 3DLive service.

3.2.1 Position tracking

Outdoor jogger's GPS-location, extracted from the smartphone, is transmitted via a web-service to the 3DLive server hosting the virtual Oulu venue. This virtual venue has a reference point and orientation mapped according to its real world venue counterpart. Therefore when jogger's position moves along the latitude axis, this is translated to value change on X-axis in the virtual venue. Same translation happens for longitude to Y-axis value change. With this mapping, it is easy to translate latitude and longitude values to XY-plane used in the virtual venue.

GPS data inaccuracies are eliminated by calculating the closest possible location for avatar on top of the road, which effectively disallows the avatar to be located inside buildings or other structures (Figure 6). This is crucial because city venue contains lots of tall buildings and running on the sidewalk close to these buildings causes GPS sensor to have multipath issues.

Fig. 6: Path of the GPS location updates with closest-to-road vectors.

Running velocity of the avatar in the virtual Oulu venue is calculated using two of the latest corrected GPS position updates and time delta between them. This allows to display the actual speed in the user interface for other users and also to modulate the avatar speed accordingly so it matches expected and realistic running animation. 3DLive platform handles the running and walking animation interpolations according to the position and speed of the avatar.

3.2.2 Webservice

An intermediate webservice component links the Android application user to the 3DLive service. When jogger launches the application, a universally unique identifier (UUID) is created. This UUID is used to separate the positional data of the different joggers, when sent to the webservice through HTTP-protocol. 3DLive service also uses HTTP-protocol to query the positional data of the currently active joggers in the venue from the webservice as shown in Figure 7.

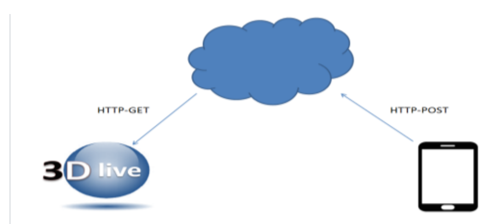

Fig. 7: Android application connection to 3DLive system through webservice.

5. CHALLENGES AND FUTURE WORK

Future prototypes of the 3DLive platform will include full body realistic 3D dynamic mesh reconstruction for the indoor jogger, while outdoor body tracking technologies will be integrated, as well. Specifically, regarding the indoor jogger, the multiple Kinect sensors that surround the jogger on the treadmill for capturing his/her motion, will be used for the 3D reconstruction of his/her appearance in real time. Advanced algorithms that combine the multiple depth maps and associated textured images will be used to offer the actual 3D reconstruction of the indoor user in real time. The realistic 3D replication of the user will then be visualized in a 3D display by the rest of the indoor users. To allow for real-time interaction among the users and their immersive experience, solutions for real-time compression and transmission of the textured 3D meshes should also be implemented and integrated in the platform. Such actions are envisaged in 3DLive to be integrated in the 2nd prototype of the project.

Realistic animations of avatars corresponding to the outdoor runners necessitate information related to joints orientations. To perform a real time motion capture of the outdoor jogger, inertial sensor-based motion capture devices will be used. This kind of solutions already exists in various indoor Virtual Reality applications or movies. However, 3DLive's intention is to utilize smart phones for retrieving this information in outdoor spaces. Such information is vital for evaluation of sport-related activity, which will incorporate inertial sensors for outdoor users, moving beyond motion sensing through Kinect sensors. The outdoor user will equipped with AR glasses and proper feedback can be visualized on the fly.

6. CONCLUSION

The co-existence, evaluation of activity and replication of movements among distant users are the basic ingredients for increased levels of User Experience in a tele-immersive application. Reports from experiments denote a high user acceptance, which is attributed to these already implemented and tested modules: Jogger running indoors, on a treadmill, is tracked and replicated to all the connected users in the service, while outdoor user is also replicated to the virtual venue according to data retrieved from the GPS sensor. Virtual venue is also dynamically updated in real time with the real weather data from the location. Venue can also be observed through 3DLive observer application if a person only wants to participate socially in the experience.

REFERENCES

[1] P. Vaïkkynen et. Al, "Using exercise cycle as a haptic input device in a virtual environment," *EGVE'01 Proceedings of the 7th Eurographics conference on Virtual Environments & 5th Immersive Projection Technology*, pp. 229-235, 2001.

[2] P. Hämäläinen et. Al, "Martial Arts in Artificial Reality," *CHI '05 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 781-790, 2005.

[3] F. 'Floyd' Mueller and S. Agamanolis, "Sports over a distance," *Computers in Entertainment*, vol. 3, no. 3, pp. 4-4, July 2005.

[4] O. Izuta et. Al, "Bouncing Star project: design and development of augmented sports application using a ball including electronic and wireless modules," in *Proc. The South African Institute for Computer Scientists and Information Technologists Conference*, Pretoria, South Africa, 2012.

[5] E. Reynell and H. Thinyane, "Hardware and software for skateboard trick visualisation on a mobile phone," in *Proc. The South African Institute for Computer Scientists and Information Technologists Conference*, Pretoria, South Africa, 2012.

[6] S. Göbel, C. Geiger, C. Heinze, and D. Marinos. "Creating a virtual archery experience," in *Proc. of The International Conference on Advanced Visual Interfaces (AVI '10)*, Rome, Italy, 2010.

[7] D. Tsetserukou, K. Sato, and S. Tachi, "ExoInterfaces: novel exoskeleton haptic interfaces for virtual reality, augmented sport and rehabilitation," in *Proc. The 1st Augmented Human International Conference (AH '10), Megève, France, 2010.*

[8] K. Higuchi, T. Shimada and J. Rekimoto, "Flying Sports Assistant: External Visual Imagery Representation for Sports Training," in *Proc. The 2nd Augmented Human International Conference*, Tokyo, Japan, 2011.

[9] A. Bleiweiss et. Al, "Enhanced interactive gaming by blending full-body tracking and gesture animation," *SIGGRAPH ASIA*, 2010

[10] S. Hasegawa, S. Ishijima, F. Kato, H. Mitake, and M. Sato, "Realtime sonification of the center of gravity for skiing," In *Proc. The 3rd Augmented Human International Conference (AH '12)*, Megève, France. 2012.

[11] N. Miller et al, "Motion Capture from Inertial Sensing For Untethered Humanoid Teleoperation," *International Journal of Humanoid Robotics*, 2004.

[12] S. A. Skogstad and K. Nmoen. Comparing Inertial and Optical MoCap technologies for Synthesis Control," Creative Commoncs, 2011.

[13] http://www.openni.org/openni-sdk/

[14] E. H. Mamdani, and S. Assilian. "An experiment in linguistic synthesis with a fuzzy logic controller." *International Journal of Man-Machine studies* 7, no. 1 (1975): 1-13.

[15] A. J. Chipperfield and P. J,. Fleming. "The MATLAB genetic algorithm toolbox." In *Applied Control Techniques Using MATLAB, IEE Colloquium on*, pp. 10-1. IET, 1995.