Serious Games with Human-Object Interactions using RGB-D Camera

Hubert P. H. Shum*
Northumbria University

Abstract

Commercial RGB-D cameras typically require a clear posture without occlusion. This hugely limits the usability of the device for serious applications that require manipulation of external objects. In this paper, we propose an integrated framework to track motion and object during human-object interactions. We implement a data-driven posture reconstruction algorithm to correct wrongly tracked body parts during occlusions, as well as a computer vision based object tracking algorithm using the depth image. We demonstrate preliminary results in which the system tracks a user playing with a basketball.

1 Introduction

With the hardware improvement of RGB-D cameras, motion-based applications have become more and more popular. The major shortfall of these motion tracking systems is the incapability of capturing human movement with real-object interaction, due to the occlusions between the body parts and the objects. This limitation explains why RGB-D cameras are still not popular for serious games such as sport training, in which external equipment is usually involved. An ideal system should allow the user to interact with real-objects, such that the user not only performs the target motion, but also produces the necessary forces to handle the objects. The system should accurately identify the human movement and the objects' trajectories in order to evaluation the motion and produce feedback. For example, a basketball training system should allow the user to play with a real basketball and evaluate the user's performance.

In this paper, we present a prototype of a serious game system that tracks the human movement and the interacting objects at the same time. On one hand, we implement the posture reconstruction system proposed by [Shum et al. 2013] in order to estimate the real human posture while performing complex movement that involves object manipulation and occlusion. On the other hand, we implement a vision-based tracking algorithm using the depth image to track the moving objects, and apply [Shum and Ho 2012] to synthesize the movement trajectories in a physical world. Our system is a step towards reliable and automatous motion analysis for serious applications such as sport training and rehabilitation analysis.

2 Tracking Human-Object Interactions

Here, we explain the algorithm to track the human movement and the trajectory of the moving objects, and present some preliminary experimental results.

We implement the posture reconstruction framework proposed by [Shum et al. 2013] to perform human motion tracking. The purpose of the algorithm is to estimate the true posture of a human user from the noisy and partially incorrect tracked postures. The idea is to (1) estimate the reliability of each tracked body part using kinematics features of the human body and movement characteristics, and (2) to apply a data-driven lazy learning algorithm to reconstruct a true posture, relying more on the body parts with high reliability values. This allows the system to automatically estimate body postures with

---

occlusions and incorrect tracking. The readers are referred to the [Shum et al. 2013] for further details.

We implement a computer vision based algorithm to perform object tracking. Given a depth image obtained by the depth camera, we first apply background subtraction to identify the pixels with movement. We filter the pixels that are due to the human movement. This is achieved by comparing the 3D positions of the pixels with each bone of the 3D body skeleton, and removing the pixels if the distance is smaller than a pre-defined value. We then threshold the resultant image to obtain a binary representation, and apply morphological erosion and dilation to filter random noise. Finally, we obtain the K regions that contain the most movement pixels, where K is manually set. The centroids of these regions are considered as the tracked positions of the objects. As a post-process of the object trajectories, we implement a positional based PD controller [Shum and Ho 2012] to synthesize the object movement.

Some experimental results are shown in Figure 1. The synthesized character and object are placed side by side with the source images for easy comparison. Notice the robustness of the system during human-object occlusions.
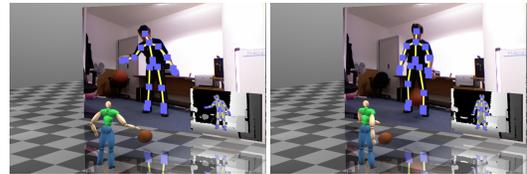

Figure 1: *The user interacting with a basketball.*

3 Conclusions and Future Work

In this paper, we propose a prototype system for serious games to identify human movement and object trajectories during human-computer interaction. We demonstrate tracking for challenging scenes that involve complex movement and occlusions.

One potential problem is the reliability of the object tracking algorithm. While we can reconstruct occluded human posture using kinematics features of the human body, object movement is general and difficult to be reconstructed. A future direction is to implement a spatio-temporal tracking algorithm [Shum and Komura 2004] that considers the continuity of the whole trajectory when deciding the position of the object from possible choices.

References

SHUM, H., AND HO, E. S. 2012. Real-time physical modelling of character movements with microsoft kinect. In *Proceedings of the 18th ACM symposium on Virtual reality software and technology*, ACM, New York, NY, USA, VRST '12, 17–24.

SHUM, H. P. H., AND KOMURA, T. 2004. A spatiotemporal approach to extract the 3d trajectory of the baseball from a single view video sequence. In *Multimedia and Expo, 2004. ICME '04. 2004 IEEE International Conference on*, vol. 3, 1583–1586 Vol.3.

SHUM, H., HO, E. S. L., JIANG, Y., AND TAKAGI, S. 2013. Real-time posture reconstruction for microsoft kinect. *IEEE Transactions on Cybernetics*.

---

*e-mail:hubert.shum@northumbria.ac.uk