

Evaluating User's Energy Consumption using Kinect Based Skeleton Tracking

Zhenbao Liu

Sicong Tang Hongliang Qin
Northwestern Polytechnical University
Xi'an, China

Shuhui Bu*

*bushuhui@nwpu.edu.cn

ABSTRACT

By utilizing the dataset provided by 3DLife/Huawei Challenge of ACM Multimedia, we propose a refreshing application that automatically evaluates player's energy consumption in gaming scenarios by a model with tracked skeleton, which may help users to know their exercise effects and even diet or reduce their weights. We develop a program to compute the energy consumption in real time by analyzing data captured from Microsoft Kinect, and also give a cue in the dynamic interaction. We model 3D human skeleton by joining different body parts with 15 nodes, and decompose player action into rigid body motions of these parts. Amount of energy consumed in the action is calculated as the sum of powers required to overcome gravity of each part. Experimental results show that instantaneous and total energy consumption of different dancers can be stably calculated. The hardware system is based on low-price Kinect, and easily accepted by users. The proposed application also provides a quantitative approach which help users to control their dining and exercise intensity.

Categories and Subject Descriptors

I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism-Virtual Reality; I.4.8 [Image Processing and Computer Vision]: Scene Analysis-Motion, Tracking

General Terms

Algorithms, Experimentation, Performance.

Keywords

Kinect, Skeleton tracking, Energy consumption

1. INTRODUCTION

Home-oriented virtual reality technologies become increasingly important in real-time realistic interaction games. They bring players rich experience by placing players in virtual environments. As a new kind of devices based on infrared structured light, depth sensors with low price such as Microsoft Kinect have attracted much attention among not only game users but also researchers and developers. It is possible to use depth sensors like Microsoft Kinect to capture human motion in an easy and robust way. This will help

many interactive games to animate an avatar, and promote user experience in games.

Some researchers have begun to use depth sensor data to construct 3D virtual applications. Tong et al. [1] present a novel scanning system for accurately capturing 3D full human body model by using multiple Kinects, which could be used to provide personalized avatars for everyday users. Bleiweiss et al. [2] blend player's actual movements tracked using a depth sensor with pre-defined animation sequences. They aim at visually enhancing the player's motion to display exaggerated and super-natural motions. Suma et al. [3] provide us a middleware to facilitate integration of full-body control with virtual reality applications and video games using OpenNI-compliant depth sensors. Alexiadis et al. [4] propose an interesting application and evaluate dance performances of students against a gold-standard performance, and provide visual feedback to the student dancer in a 3D virtual environment.

Different from above applications, we propose a novel system that automatically evaluates player's energy consumption in gaming scenarios, which may help users to know their exercise effects and even diet or reduce their weights. The dancing action is seen as the process of burning calories or fat. We stably obtain quantitative energy consumption of each dancer, and convert it to the value of burned fat. Players could interactively know the exercise effects during dance.

2. KINECT SKELETON TRACKING

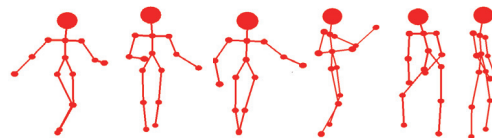


Figure 1: Skeletons extracted from depth images.

We use the provided dancer dataset captured using a Microsoft Kinect by 3DLife/Huawei Challenge of ACM Multimedia [5]. We take 10 dancers' videos from the set of dance videos. Each video is composed of a sequence of depth images, which are encoded with ONI format. By means of the high-level skeleton tracking module provided by OpenNI Develop Kit, we remove the background, and extract the skeletons from the captured images. It help us to accurately obtain the changed skeleton in each frame, as shown in the

head	forearm	upper arm	thigh	shank	torso
23.1%	1.8%	3.5%	9.4%	4.2%	58%

Table 1: Every part’s percentage of the body weight.

Figure 1. We define a typical human model with 175cm height and 65kg weight by a normal relation $W = H - 110$ between height H and weight W . According to human factors engineering, every body part has a percentage of the body weight as listed in the Table 1.

3. ENERGY CONSUMPTION MODEL

In this section, we describe the energy consumption model adopted in the application. The body is composed of 10 mass parts $\{m_i\}$ between 15 joint nodes. Energy consumption of each motion is approximately equivalent to the required power to overcome gravitational potential energy of moving parts. The power is obtained by tracking the position change of each body part, and the powers of all the parts are summed up. Certainly we view each body part as a rigid body with uniform mass distribution.

In the model, the vertical coordinate axis is defined as y axis, horizontal coordinate axis is defined as x axis, and z axis points outside the display. In order to acquire the kinetic energy of each moving part, its motion is firstly decomposed into two directional movements in vertical plane and horizontal plane. In the vertical plane, the height change of each part between the frame $t + 1$ and the frame t is defined as $\Delta h(t)$. The power $E_v(i)$ of moving part m_i in the vertical plane is given by $E_v(i) = m_i g \Delta h(t)$. Similarly, we deduce the required power E_h during movement in the horizontal plane, and represent it via the kinetic energy of each mass part in the horizontal plane $E_h(i) = E_x(i) + E_z(i)$. Each item of kinetic energy is computed with the aid of movement speed and mass of each part. Here we differentiate positions in two continuous frames for real time speed.

Finally, the total power $E(i)$ of the i th part between two frames is obtained by summing up the powers $E_v(i)$ and $E_h(i)$ in the two planes. We consider intervals between two frames are uniform, and hence we see movement of rigid part as uniform linear motion. The total energy consumption E of the body at time t is the energy summation of all the parts.

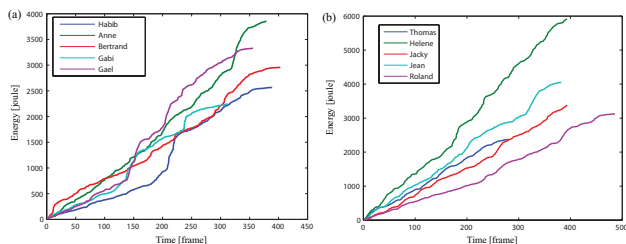


Figure 2: The cumulative energy consumption curves of all ten dancers.

4. EXPERIMENTAL RESULTS

We implemented the application based on C++, OpenNI, and OpenGL. OpenNI is first used to track the skeleton from

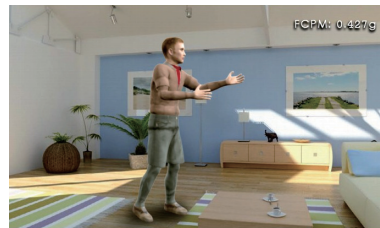


Figure 3: Visualization result in one frame.

the provided dataset. The skeleton movements are visualized and confirmed in OpenGL. We programmed the computation process of energy consumption using C++, and visualize the interaction in a gaming development software, Unity. We totally experimented ten dancers’ depth images, and extracted their skeletons. The instantaneous and total energy consumption are calculated in real time. We obtained energy consumption in each frame, and illustrate the relationship between energy consumption and time. The cumulative energy consumption curves of the ten dancers are drawn in the Figure 2(a)(b). The horizontal axis is frame (30 frames per second), and vertical axis is joule. Curves with larger slope show that these dancers consume more energy at this frame because their dances need more motions of body parts. In order to interactively visualize the result, we also map the dancer’s motions onto a photo-realistic avatar, and place the avatar into Unity development environment. And then the system provides a vivid description of energy consumption, that is, we suggest a conversion of dancers’ energy consumption to fat consumption per minute (FCPM). It is known that 1g fat contains the energy of 37.67 kilojoule. Player’s motion interactively gives the dynamic value of fat consumption per minute, for example, 0.427g, as illustrated in the Figure 3.

5. ACKNOWLEDGMENTS

This work was partly funded by NSFC (61003137), SRF-ROCS, NWPU Fundamental Fund(JC201220)(JC201202), NSBRP in Shanxi (2012JQ8037), and Fund from State Key Lab of CAD & CG, ZJU.

6. REFERENCES

- [1] J. Tong, J. Zhou, L. Liu, Z. Pan, and H. Yan, Scanning 3D full human bodies using kinects, IEEE Transactions on Visualization and Computer Graphics, 18(4), 2012.
- [2] A. Bleiweiss, D. Eshar, G. Kutliroff, A. Lerner, Y. Oshrat, and Y. Yanai, Enhanced interactive gaming by blending full-body tracking and gesture animation, Proceedings of ACM SIGGRAPH Asia Sketches, 2010.
- [3] E. A. Suma, B. Lange, A. Rizzo, D. M. Krum, M. Bolas, FAAST: The flexible action and articulated skeleton toolkit, Proceedings of IEEE Virtual Reality Conference, 2011.
- [4] D. Alexiadis, P. Daras, P. Kelly, N. E. O’Connor, T. Boubekeur, and M. B. Moussa, Evaluating a dancer’s performance using kinect-based skeleton tracking, Proceedings of ACM Multimedia, 2011.
- [5] 3DLife/Huawei ACM MM Grand Challenge. <http://perso.telecom-paristech.fr/~essid/3dlife-gc-12/>