Real-time Avatar Motion Synthesis by Replacing Low Confidence Joint Poses

Rina Tanaka * Hiroshi Mori † Fubito Toyama Kenji Shoji

Utsunomiya University

1 Introduction

In recent years, the availability of low-cost depth sensors has facilitated motion capture applications for personal use. Human motion can be captured to control an animated avatar, with no requirement to wear the dedicated sensors. However, users must contend with avatar pose errors resulting from user pose recognition failures, which can be caused by usage environment problems or measurement errors. For these problems, the accurately capture system using a single depth camera [Wei et al. 2012] have been proposed. In this paper, we propose a simple method for synthesizing seemingly natural avatar motion based on the user's body movements, including user pose recognition failures (Figure 1). First, we calculate the degree of confidence for each joint's pose parameters that are captured using a depth sensor. Next, the low confidence joint poses are replaced with a similar pose that is calculated based on high confidence joint poses. In addition, the joints that are not detected are complemented with a calculated similar pose.

2 Estimation of false recognition joints pose

Body motion is recorded utilizing a Kinect™sensor, which records each joint's rotation during real-time user movements. We calculate the degree of confidence for joints $\mathbf{p}^i, i \in \{head, neck, \cdots\}$ using the structure of the human model to detect user pose recognition failures. We evaluate each joint pose from (i) the continuity of pose transition, (ii) the rotation limit, (iii) the collision between joints, and (iv) the hierarchical structure of each joint.

3 Replacing the low confidence joints poses

We search the database of sample motions for a similar pose $\hat{\mathbf{q}}$ that can replace a low confidence joint pose. The evaluation value employed for the similar pose search is calculated using the weighted squared distance of the rotation coordinates of the joints from the two poses, according to the following formula:

$$\hat{\mathbf{q}} = \arg\min_{\mathbf{q}} \left[\alpha \sum_c w_c^{dist} ||\mathbf{p}^c - \mathbf{q}^c||^2 + \beta \sum_j w_j^{diff} ||\mathbf{r}_t^j - \mathbf{r}_{t-1}^j||^2 \right] \quad (1)$$

$$\mathbf{r}^j = \begin{cases} \mathbf{p}^j & if \;\; j \in M^{conf} \\ \mathbf{q}^j & if \;\; j \in \overline{M^{conf}} \end{cases} \quad (2)$$

$\{\mathbf{q}^c | c \in M^{conf}\}$ is the sample motion joint corresponding to the estimated high confidence joint$\{\mathbf{p}^c | c \in M^{conf}\}$. $\{\mathbf{r}^j | j \in M^{output}\}$ is the output joint pose after replacing the low confidence joint with the similar pose joint.

The low confidence joint pose is replaced with the similar pose $\hat{\mathbf{q}}$. In addition, the joints that are not detected are complemented with a similar pose $\hat{\mathbf{q}}$. Finally, the output pose of the avatar model is synthesized after integration processing (Figure 2).

*e-mail: tanaka@image.is.utsunomiya-u.ac.jp

†e-mail: hmori@is.utsunomiya-u.ac.jp

SIGGRAPH 2014, August 10 – 14, 2014, Vancouver, British Columbia, Canada.

ACM 978-1-4503-2958-3/14/08

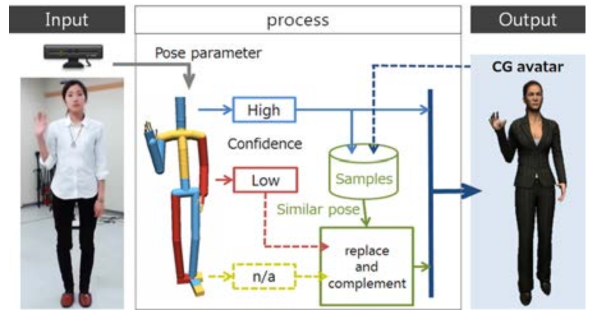

Figure 1: *Overview of the proposed method*

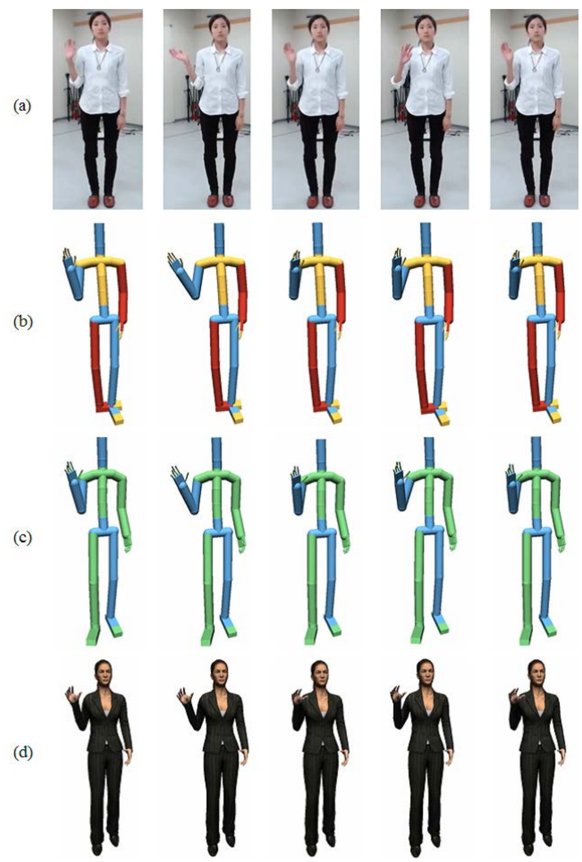

Figure 2: *Results of the synthesized avatar animation. (a) User performance. (b) Estimation pose errors. Blue indicates the high confidence joints; Red indicates the low confidence joints; Yellow indicates the joints that are not detected. (c) Replaced with similar pose. (d) The applied avatar model.*

References

WEI, X., ZHANG, P., AND CHAI, J. 2012. Accurate realtime full-body motion capture using a single depth camera. *ACM Trans. Graph. 31*, 6 (Nov.), 188:1–188:12.