Animations of Real Swimming via Motion Reconstruction

Seiya Takahashi[1] and Shigeru Kuriyama[1]

[1]Toyohashi University of Technology, Toyohashi, Aichi, Japan

***Abstract*— CG animations synthesized with captured data are widely utilized for visually analyzing the skills of various kinds of exercises and sports. Animations of swimming, however, are very difficult to create because its motions are hard to be captured with existing devices on an underwater condition.**

For overcome this difficulty, this paper proposes a reconstruction method of swimming motions using inertial sensors for estimating the acceleration and angular velocity of bodily parts. Only six sensors are attached to wrists, ankles, head and waist for capturing the motions while minimizing water resistance. We also construct motion dataset captured with optical markers by imitating crawling strokes on the ground, and the captured motions are referenced for calibrating initial poses and estimating joints' movements where the inertial sensors cannot detect. We demonstrate the capability of this hybrid motion capture system by visually comparing estimated crawling movements in the water.

***Index Terms*—Computer Graphics, Three-Dimensional Graphics and Realism, Animation, Motion Reconstruction, Motion Capture, Inertial Sensor, Swimming.**

I. Introduction

Visual analysis of swimming motions with observed data can supply an effective training tool. Underwater photography is widely used to capture the swimming movements in the water owing to its handiness. The motion in 3D space, however, cannot be obtained with one video camera alone. Photography with multiple cameras can capture 3D movements with some computer vision technology; capturing swimming motions are, however, very difficult because of the optical reflection and refraction caused by bubbles in the water.

Optical motion capture systems with many infrared cameras are widely used in accurately capturing 3D movements. However, setting many cameras in the water is usually very difficult and most systems supply no water-resistant capability.

Meanwhile, inertial sensors are often utilized in analyzing swimming motions by capturing accelerations and angular velocities of bodily parts, and they can measure small differences of motion that are difficult to distinguish from the movie. They are, however, seldom used in generating motions due to the following problems:

- It is very difficult to detect initial limbs' poses.
- Even if the initial pose is detected, the drift of acceleration and angular velocities accumulates errors in estimating limbs' poses.
- The number of attached sensors is usually insufficient to estimate the motions of all joints.

The target of this research is to capture and visualize swimming motion using inertial sensors. We propose a method for estimating swimming motions with only 6 inertial sensors and the supplementary motion data captured with a conventional motion capture system. The complementary motion data is accurately captured on the ground by imitating crawling movements, and it is converted to angular velocities of moving limbs for finding closest match. This dataset is prepared for estimating motions undetectable with the inertial sensors.

Our system introduces handy inertial sensors that are popularly utilized in motion analysis and no other devices are introduced for measuring 3D motions in the water. Unlike conventional optical motion capture systems, our handy capture system requires no limitation on measurement area. These properties are very suited to develop the visual analysis system of swimming for the training of beginners at reasonable cost.

II. Related Work

Inertial sensors have been utilized in 3D motion capture systems, and some products [1] demonstrate good performance in accuracy. However, they often require integrating with other kinds of sensors [2] for compensating the effect of drift. In addition, the attachment of sensors needs a special tailor-made suit, which imposes stresses in moving a body.

A performance capture system generates motions from only attached acceleration sensors for the interactions in computer games [3] and the visualization of athlete performances [4]. They often generate motions on the fly by retrieving most similar movements from pre-captured motion data set. We introduce a similar approach to this performance capture for swimming motion.

In the field of analysis of swimming motions, inertial sensors are often utilized for detecting swimmers' movements. Bächlin et al. [5] used accelerometers to estimate the roll and pitch components of swimming motions. Nakashima et al. [6] used acceleration and angular rate sensors to estimate wrists' movements. This method requires the initial poses to be appropriately given because no static states are supplied from the sensors.

S. Takahashi is with Toyohashi University of Technology, Aichi, Toyohashi, Japan (e-mail: takahashi@val.cs.tut.ac.jp).

S. Kuriyama is with Toyohashi University of Technology, Aichi, Toyohashi, Japan (e-mail: kuriyama@cs.tut.ac.jp).

ISBN: 978-602-8944-11-3

III. METHODS

A. Representation of Motion data

Optical motion capturing system measures the 3D positions of markers attached around joints, and they are usually converted to rotational joint angles by considering the skeletal configuration of a human body. We generate motion data represented by rotational matrices of 14 bones whose separated components are represented in Fig.1. We place a root position at the origin and align X- and Y- axes of right-handed coordinate system with the swimming and gravitational directions, respectively, where Z-axis is automatically determined to be orthogonal to both X- and Y- axes.

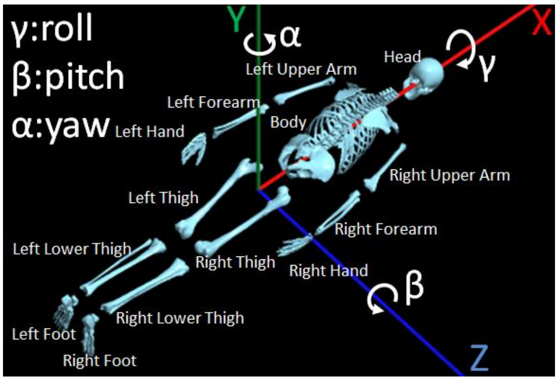


Fig. 1. Skeletal model of bodily parts and coordinates system.

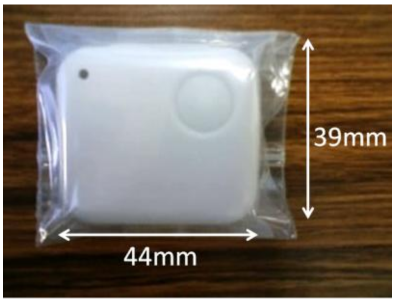


Fig. 2. WAA-006.

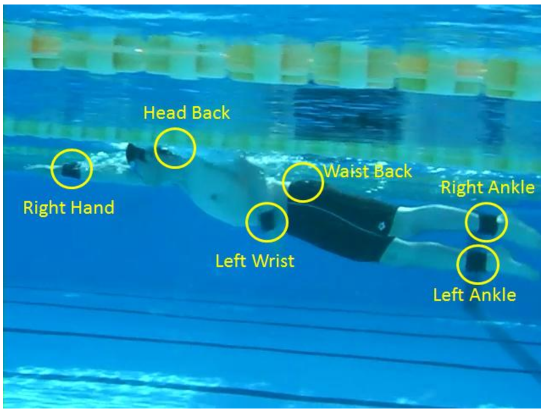


Fig. 3. Attached locations of 6 sensors.

(a) Lying on stomach on a tabletop

(b) Hanging with a rope

(c) Standing on a floor

Fig.4. Three different conditions for ground MoCap.

B. Inertial sensors in the water

We use inertial sensors in Fig.2, that consist of tri-axial accelerometer of ±2G/4G and tri-axial angular velocity meter of ±500dps(degree per sec) for X,Y axes and ±300dps for Z axis, where all components are measured at the frequency of 333Hz. The sensors are waterproofed by doubly covering them with cheap vinyl. We reconstruct swimming motions based on the data obtained with 6 sensors attached to a body as shown in Fig. 3. From now on, we call the motion data captured with the inertial sensors in the water by *underwater MoCap*.

C. Mimicking motions on the ground

The swimming motions cannot be fully reconstructed from the inertial sensors alone, and the missing parts of the motions are predicted from the motion data captured on the ground where we use the optical motion capture system that has 12 infrared

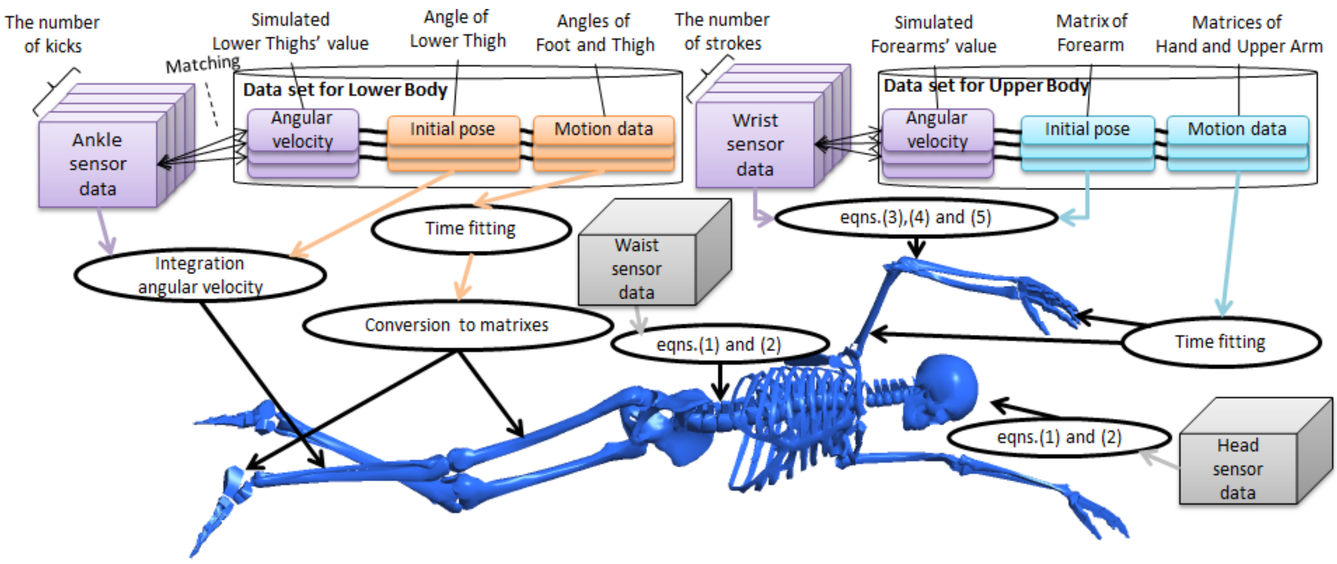


Fig. 5. System Overview.

high-resolution cameras and 27 optical reflective markers. From now on, we call this motion data captured by mimicking swimming motions by *ground MoCap*.

We experimentally captured the ground MoCap on three different physical conditions, as shown in Fig.4. The equipments for supporting a body in (a) *lying on a stomach* and (b) *hanging with a rope* have the advantage in that they can measure full-body movements. However, the roll rotations of a trunk are hard to be capture because of the contact force with the object supporting a body. This improper physical condition also propagates to the movements of the arms linked to the trunk. Moreover, the condition (b) generates unnatural yaw rotations caused by swinging arms as physical reactions.

The condition (c) cannot reconstruct the movements of a lower body; it, however, can reliably reconstruct arm swings because the subjects can roll their waist easily and stably. On the other hand, the rotational axes of kicking motions are fixed because we can approximate all joints of legs as hinge joints without losing the accuracy in reconstructing motions.

For taking advantageous property of these conditions, we separately capture the movements of the lower and upper parts of a body with the conditions (a) and (c), respectively.

D. Motion synthesis for each bodily parts

The diagram of our motion reconstruction is shown in Fig. 5. We first explain about motions for a trunk and head that are reconstructed from the underwater MoCap. The output of accelerometers includes the acceleration of gravity G, and its component can be easily detected with the inertial sensors attached to waist back and head back (see Fig.3), because their motions are usually very slow and the output is mostly occupied by the acceleration of gravity. For this reason, the rotational components of the trunk and waist are computed by the change of acceleration because the gravitational components are changed according to their orientations. The values of yaw rotations is, however, hard to be obtained because its axis is parallel to the gravitational direction, and we therefore neglect its variations by resetting them with zeros.

Given the output of accelerometer, $\mathbf{a}=[a_x \quad a_y \quad a_z]$ the components of roll and pitch are computed as follows:

$$\gamma = Sin^{-1}\frac{g_z}{\sqrt{a_y^2+a_z^2}} - Tan^{-1}\frac{a_z}{a_y}, \qquad (1)$$

$$\beta = Sin^{-1}\frac{a_x}{\sqrt{g_x^2+g_y^2}} - Tan^{-1}\frac{g_x}{g_y}, \qquad (2)$$

where $\mathbf{g}=[g_x \quad g_y \quad g_z]$ denotes the output when swimmers have a neutral pose, and they usually take $g_y \fallingdotseq G >> g_x, g_z$ if the sensors are properly attached.

We next explain about the motion reconstruction of arms. Their motions are synthesized with the ground MoCap segmented per stroke consisting of three types of data: 1) the simulated values of the angular velocity sensors attached to the wrists, 2) the pose matrices of upper arms and hands, and 3) the initial poses of the forearms.

The pose of the forearm can be computed from the angular velocity of the sensor attached to the wrist. Let $\mathbf{A}_i$ the coordinate system of the forearm at the i-th frame, and let $\mathbf{w}_i$ the rotational vector obtained from the sensor, the rotational matrix $\mathbf{R}_i$ is then computed with eqns. (3) and (4), to compute the pose at the $(i+1)$-th frame as

$$\mathbf{w}_i' = \frac{1}{\mathrm{Freq}}\mathbf{w}_i\mathbf{A}_i, \qquad (3)$$

$$\mathbf{R}_i = \begin{bmatrix} 1 & w_z' & -w_y' \\ -w_z' & 1 & w_x' \\ w_y' & -w_x' & 1 \end{bmatrix}, \qquad (4)$$

$$\mathbf{A}_{i+1} = \mathbf{A}_i\mathbf{R}_i, \qquad (5)$$

where **Freq** denotes the sampling frequency of the sensor. The initial pose $\mathbf{A}_1$ of the forearm is given by adequately selecting a corresponding pose from the ground MoCap by matching the motion segments as described below. Our method computes motions separately for each interval of stroke, and the underwater MoCap is segmented by referencing the wave pattern of the angular velocity. The motion of crawl is segmented by dividing it at the timing where the norm of the angular velocity becomes minimum after sinking of the arms, at which one arm is stretched out for a short while (see Fig. 6). We also segment the ground MoCap by computing the angular velocities of forearms for estimating the initial poses. The motions of upper arms and hands are automatically synthesized by adopting the rotational matrices taken from the ground MoCap on condition (c) every frame. Notice that the coordinates defined in the condition (c) is transformed to match with those defined in Fig.1.

The motions of legs are adopted from the rotational angels of the ground MoCap along fixed axis that is determined from the line connecting two markers of both sides of waist. Only the motions of calves are computed from the underwater MoCap with the initial pose taken form the ground MoCap. We simulate the angular velocity of the sensors at ankles with ground MoCap in order to segment both types of MoCap at the timing where the joint velocity along the fixed axis become 0, at which the movements of kicking up and down are exchanged. Fig.7 shows the angular velocity of the calf and the initial angle. We set the initial angles at which the angular velocity changes from negative to positive because the crossing points correspond to the start of kicking up. The pose matrices of calves are then obtained by integrating the angular velocity with these initial states. We similarly compute the pose matrices of thighs and feet by taking rotational angles along the fixed axis.

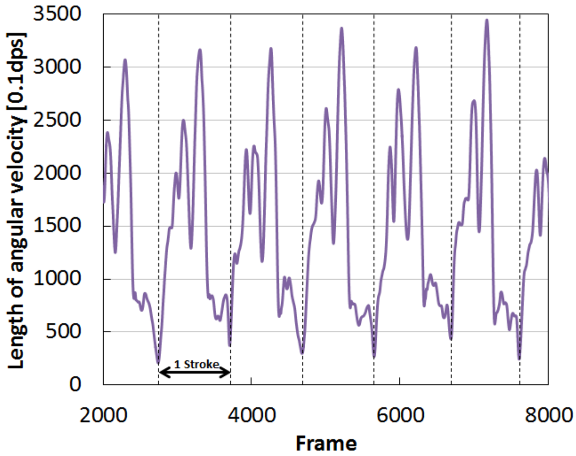


Fig. 6. Angular velocity of a right forearm.

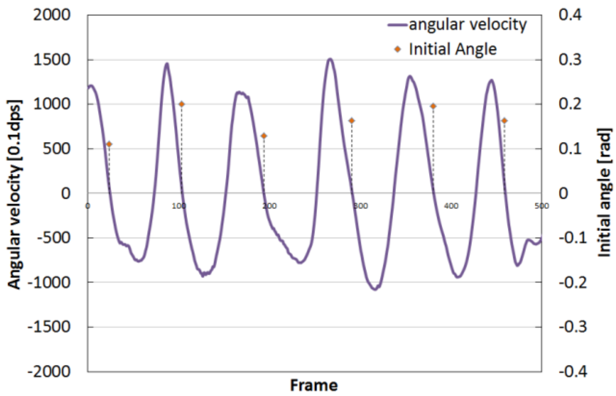


Fig. 7. Angular velocity and initial angles of lower thigh.

IV. Results

The examples of the estimated motions are demonstrated in Fig.8 and Fig.9 from two viewpoints, compared with the images taken from a digital video of similar viewpoints, respectively. Fig.8 reveals the unnatural head motions when breathing. The resulting movies also show discontinuous motions because they are independently reconstructed per strokes or kicking. Smart interpolation of these gaps should be developed.

The comparative experiments against the action capture methodology is demonstrated in Fig.10. The figure (a) denotes the posed synthesized by using our hybrid method. The figure (b) shows the pose synthesized by directly adapting the ground MoCap so as to be time-aligned with the underwater MoCap using segmentation. This approach is similar to existing action capture systems that extract accurate and high dimensional motion fragments from the low dimensional signals. The figure (c) is the snapshot selected to have the same timing as (a) and (b). This shows that the orientation of the left forearm in (a) is more similar to the actual pose in (c) than the pose in (b). This derives from the difference of timing or the behavior of bending arms, between motions on the ground and underwater MoCap.

V. Discussion and Conclusion

We have proposed the motion reconstruction using inertial sensors. The missing movements of the underwater MoCap are adaptively replaced with the ground MoCap. Since the ground MoCap does not perfectly match the underwater one, the resulting movements often include implausible complement, for example, the orientation of the left wrist in Fig.10 (a) is slightly different from the actual pose in (c). These defects could be improved by increasing the size and variety of the ground MoCap for increasing motion fragments that can complement various styles and levels of skills.

Our reconstruction is based on the simple motion matching between the ground and underwater motions; more sophisticated technique, however, should be explored for increasing the estimation accuracy. This experiment only demonstrates the reconstruction of crawling motions, and the motion data is measured from only one person. From a practical viewpoint, the ground MoCap should be adaptively recycled for different swimmers, and the validity for such inter-person reconstruction should be investigated, for different types of swimming motions, which are included in our future works.

References

[1] Moven: moven.com. 2010

[2] D. Vlasic, R. Adelsberger, G. Vanucci, J. Barnwell, M. Gross, W. Matusik, J. Popovic, “Practical motion capture in everyday surroundings” In ACM Transactions on Graphics(2007), vol.26, p.35.

[3] R. Slyper, J. Hodgins, “Action capture with accelerometers.” In 2008 ACM SIGGRAPH / Eurographics Symposium on Computer Animation (2008).

[4] P. Kelly, C. Conaire, J. Hodgins, N. O’Connor, “Human Motion Reconstruction using Wearable Accelerometers.” In 2010 ACM SIGGRAPH / Eurographics Symposium on Computer Animation (2010).

[5] M. Bächlim, K. Förster, G. Tröster, “SwimMaster: A Wearable Assistant for Swimmer” In 11th international conference on Ubiquitous computing (2009)

[6] M. Nakashima, Y. Ohgi, E. Akiyama, N. Kazami, “Development of a swimming motion display system for athlete swimmers’ training using a wristwatch-style acceleration and gyroscopic sensor device” In 8th Conference of the International Sports Engineering Association(2010)

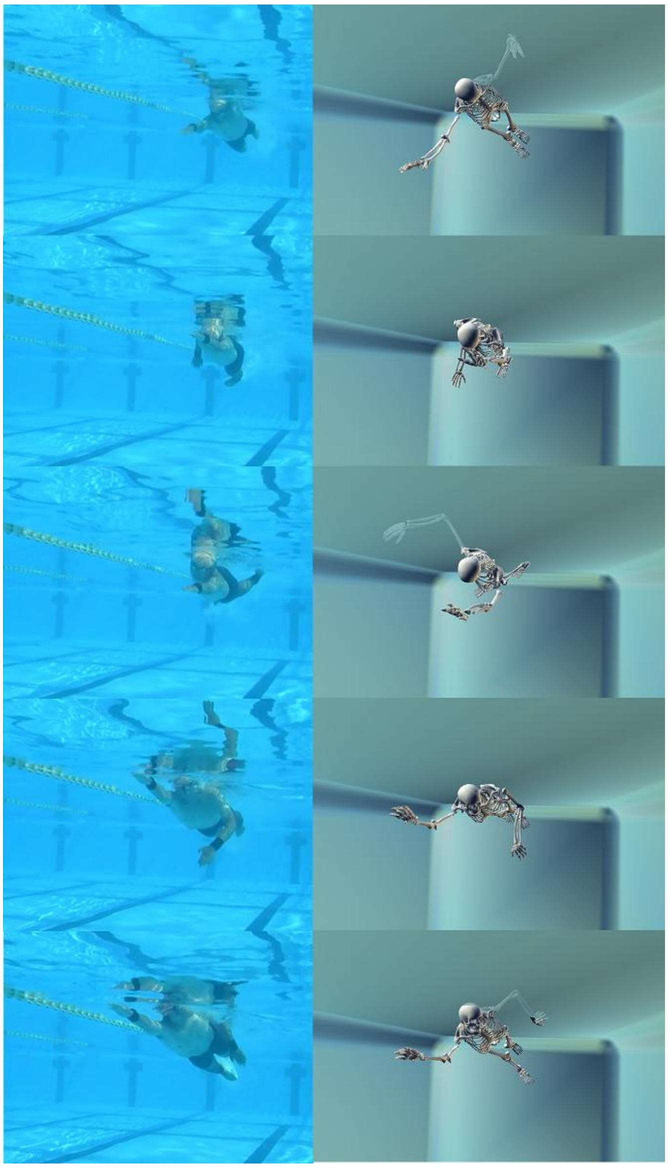


Fig. 8. Comparison between underwater video and reconstructed CG animations.

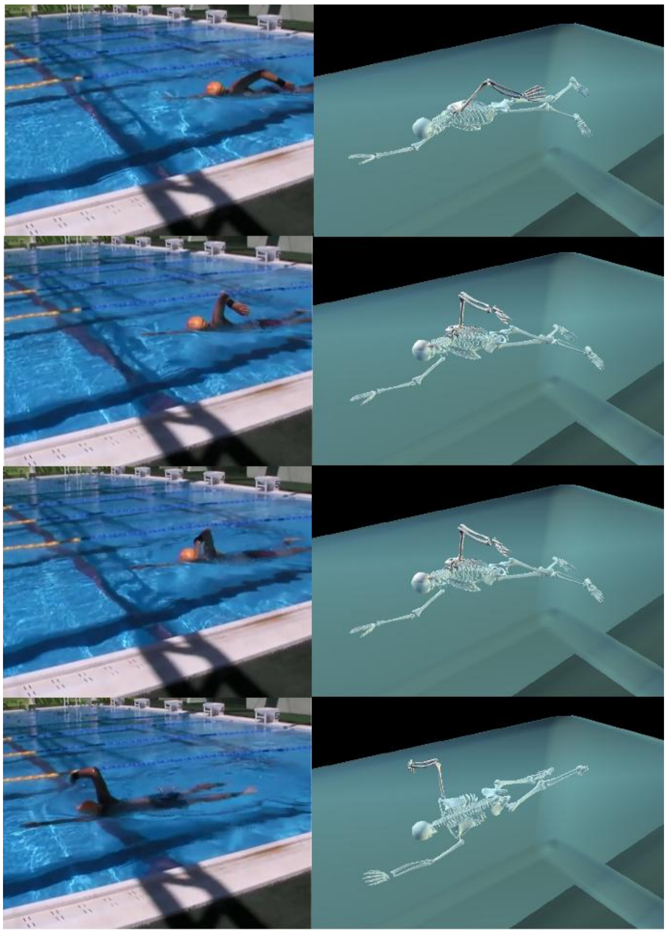
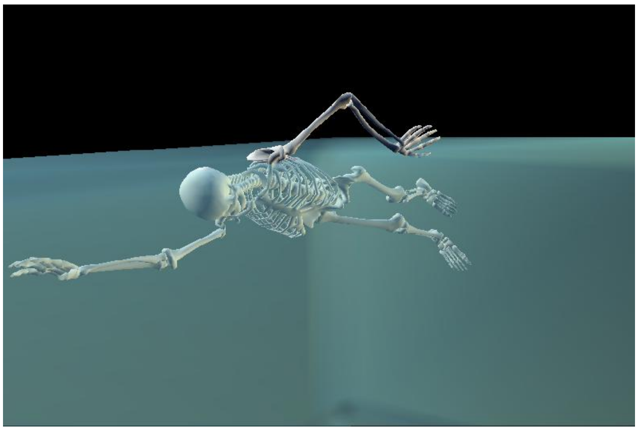
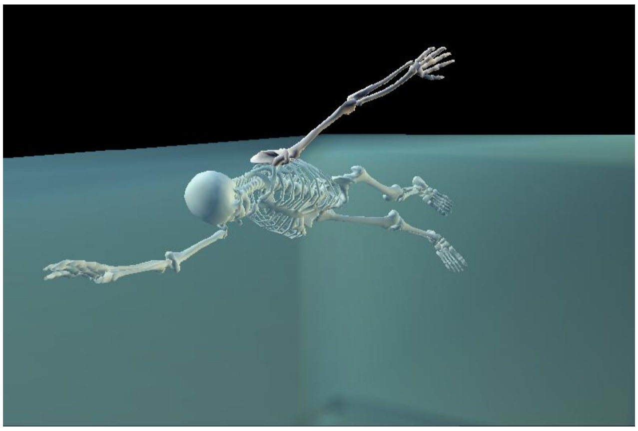


Fig. 9. Comparative results from another viewpoint.



(a) Motion reconstruction using our method



(b) Motion synthesis using the approach of action capture

(c) Snapshot from video images

Fig. 10. Comparison of the forearm motions for different types of approaches.