Journal of Information Processing Vol.23 No.5 704–707 (Sep. 2015)

[DOI: 10.2197/ipsjjip.23.704]

Technical Note

Sensor Number Reduction in Skeleton Estimation from Magnetic Motion Capture Data

Takeshi Miura[1,a)] Takaaki Kaiga[2] Katsubumi Tajima[1] Takeshi Shibata[3] Hideo Tamamoto[4]

Received: February 20, 2015, Accepted: May 9, 2015

Abstract: In this paper, we try to reduce the number of sensors used in skeleton estimation from magnetic motion capture data. A disadvantage of limiting the subject's motion caused by wiring for sensors can thereby be reduced. We remove the sensors attached to the lower legs and forearms. The parameters related to non-sensor body segments are estimated based on the biomechanical structure of the human body. Experimental results showed that the parameters were properly estimated by the present method, even though the number of sensors was reduced.

Keywords: motion capture, magnetic motion capture system, skeleton, sensor number reduction

1. Introduction

Nowadays, motion capture (Mocap) systems are used to accurately record human-body motions. Magnetic Mocap systems are known as those having several advantages; e.g., sensors give the information on both the position and the orientation, and are not occluded by non-metallic objects [1]. However, there is also a disadvantage that the wiring for sensors can limit a subject's motion [1].

Meanwhile, raw Mocap data are often converted into those giving the motion of an articulated biomechanical model, i.e., a skeleton [2], [3]. In a skeleton, rigid-body segments are connected via rotating joints. The above conversion allows us to represent the state of body motion systematically with a skeletal structure. As for magnetic systems, several researchers have proposed methods to automatically estimate the parameters of a skeleton from acquired Mocap data. O'Brien et al. [4] proposed an approach to estimate joint positions by linear least squares fitting (LLSF). Knight and Semwal [5] improved the computation speed of joint-position estimation by adopting the Generalized Delogne-Kåsa Estimator (GDKE). Both methods require the setup of two sensors for estimating the position of a joint sandwiched between them. This means that every segment in a skeleton must have at least one sensor corresponding to itself.

In this paper, we try to reduce the number of sensors. The disadvantage of limiting the subject's motion can thereby be reduced. We remove the sensors attached to the lower legs and forearms. The positions of joints connected to a non-sensor segment are estimated based on the biomechanical structure of the human body. As for the positions of remaining joints each sandwiched between two sensors, we use any of the methods previously proposed, i.e., LLSF or GDKE. Meanwhile, we also present a method to estimate the orientation of each segment; to our knowledge, there are few reports explicitly presenting a biomechanically valid approach for orientation estimation from magnetic Mocap data.

[1] Graduate School of Engineering and Resource Science, Akita University, Akita 010–8502, Japan
[2] Digital Art Factory, Warabi-za Co., Ltd., Semboku, Akita 014–1192, Japan
[3] Venture Business Laboratory, Akita University, Akita 010–8502, Japan
[4] Tohoku University of Community Service and Science, Yamagata 998–8580, Japan
[a)] miura@ipc.akita-u.ac.jp

2. Methods

2.1 Joint-position Estimation

Figure 1 shows the skeleton used in this paper. As mentioned in Section 1, the sensors attached to the specific segments constituting the limbs (i.e., dotted boxes in Fig. 1) are removed. The parameters L_{T}, L_{TH}, L_{LL}, L_{UA} and L_{FA} will be used later in Eq. (4).

First, the positions of the following joints (i.e., joints in the torso, each sandwiched between two sensors) are estimated by LLSF or GDKE: W (waist), N (neck), LH (left hip), RH (right hip), LS (left shoulder) and RS (right shoulder). Next, the position of R (root) is obtained by calculating the centroid of W, LH and RH. Then, the position of C (chest) is obtained as follows:

$$\boldsymbol{q}_{\mathrm{C,O}}(i) = \frac{\boldsymbol{q}'_{\mathrm{C,O}}(i) - \boldsymbol{q}_{\mathrm{W,O}}(i)}{|\boldsymbol{q}'_{\mathrm{C,O}}(i) - \boldsymbol{q}_{\mathrm{W,O}}(i)|}\alpha + \boldsymbol{q}_{\mathrm{W,O}}(i) \tag{1}$$

$$\alpha = \frac{1}{M}\sum_{i=1}^{M} |\boldsymbol{q}'_{\mathrm{C,O}(i)} - \boldsymbol{q}_{\mathrm{W,O}}(i)|$$

where $\boldsymbol{q}_{\mathrm{C,O}}(i)$ is the position of C (at the ith frame in the global coordinate system having the origin O, the same hereinafter), $\boldsymbol{q}'_{\mathrm{C,O}}(i)$ is the middle point between the two points (one is on Line W-N and the other on LS-RS, giving the shortest distance between W-N and LS-RS), $\boldsymbol{q}_{\mathrm{W,O}}(i)$ is the position of W and M is the total number of frames, respectively. Finally, the positions of the joints constituting the limbs are estimated as follows.

Figure 2 shows the arrangement of sensors and joints with respect to the limb. The jth and $(j+1)$th joints are connected to a non-sensor segment. Under the assumption that each segment is

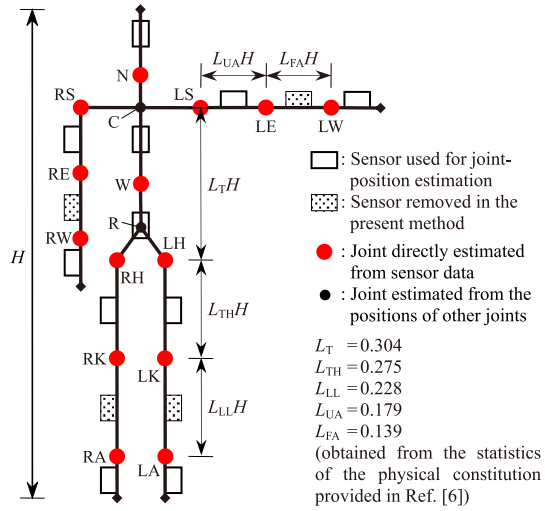


Fig. 1 Arrangement of sensors and joints in the skeleton.

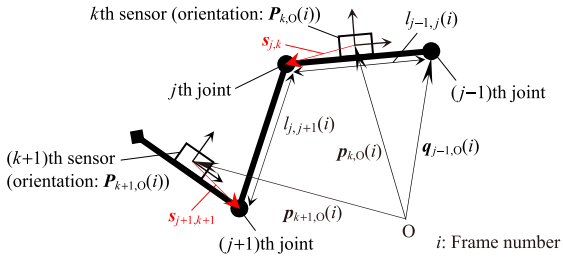


Fig. 2 Arrangement of sensors and joints with respect to the limb.

a rigid body, the distance between these joints becomes constant. As a result, the variance of the distance throughout an entire time series, $v_{j,j+1}$, should become zero:

$$v_{j,j+1} = \frac{1}{M}\sum_{i=1}^{M}\{l_{j,j+1}(i) - \bar{l}_{j,j+1}\}^2 \tag{2}$$

$$\bar{l}_{j,j+1} = \frac{1}{M}\sum_{i=1}^{M} l_{j,j+1}(i),$$

$$l_{j,j+1} = |\{\boldsymbol{p}_{k+1,O}(i) + \boldsymbol{P}_{k+1,O}(i)\boldsymbol{s}_{j+1,k+1}\} - \{\boldsymbol{p}_{k,O}(i) + \boldsymbol{P}_{k,O}(i)\boldsymbol{s}_{j,k}\}|$$

where $\boldsymbol{p}_{k,O}(i)$, $\boldsymbol{p}_{k+1,O}(i)$, $\boldsymbol{P}_{k,O}(i)$ and $\boldsymbol{P}_{k+1,O}(i)$ are the positions and orientations (rotation matrices) of the kth and $(k+1)$th sensors at the ith frame in the global coordinate system, and $\boldsymbol{s}_{j,k}$ and $\boldsymbol{s}_{j+1,k+1}$ are the positions of the jth and $(j+1)$th joints in the local coordinate systems of the kth and $(k+1)$th sensors, respectively. We can estimate appropriate joint positions by adjusting $\boldsymbol{s}_{j,k}$ and $\boldsymbol{s}_{j+1,k+1}$ so as to minimize $v_{j,j+1}$.

To prevent the deterioration of estimation accuracy caused by the lack of the motion amount of the limb, we also introduce the evaluation indices $e_{Sj,j+1}$ and $e_{Lj,j+1}$ in addition to $v_{j,j+1}$:

$$e_{Sj,j+1} = |\boldsymbol{s}_{j,k}|^2 + |\boldsymbol{s}_{j+1,k+1}|^2 \tag{3}$$

$$e_{Lj,j+1} = \frac{1}{M}\sum_{i=1}^{M}[\{l_{j-1,j}(i) - L_{j-1,j}\}^2 + \{l_{j,j+1}(i) - L_{j,j+1}\}^2] \tag{4}$$

$$l_{j-1,j}(i) = |\boldsymbol{q}_{j-1,O}(i) - \{\boldsymbol{p}_{k,O}(i) + \boldsymbol{P}_{k,O}(i)\boldsymbol{s}_{j,k}\}|,$$

$$L_{j-1,j} = L_\beta L'_T/L_T \quad (\beta\text{: TH for the leg, UA for the arm}),$$

$$L_{j,j+1} = L_\gamma L'_T/L_T \quad (\gamma\text{: LL for the leg, FA for the arm}),$$

$$L'_T = \frac{\sum_{i=1}^{M}\{|\boldsymbol{q}_{W,O}(i) - \frac{\boldsymbol{q}_{LH,O}(i)+\boldsymbol{q}_{RH,O}(i)}{2}| + |\boldsymbol{q}_{C,O}(i) - \boldsymbol{q}_{W,O}(i)|\}}{M}$$

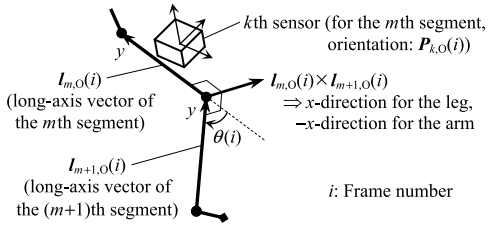


Fig. 3 Joint rotation of the limb.

where $\boldsymbol{q}_{j-1,O}(i)$ is the position of LH (for the left leg), RH (for the right leg), LS (for the left arm) or RS (for the right arm). Minimizing $e_{Sj,j+1}$ prevents that the distance between a sensor and a joint becomes too large, whereas minimizing $e_{Lj,j+1}$ suppresses an extreme deviation from the standard physical constitution. We eventually adopt the index E as a cost function to be minimized:

$$E = v_{j,j+1} + w_S e_{Sj,j+1} + w_L e_{Lj,j+1} \tag{5}$$

where w_S and w_L are the weights for $e_{Sj,j+1}$ and $e_{Lj,j+1}$, respectively. We give w_S and w_L small values to take a tolerance margin ($w_S = w_L = 0.01$). We use the simplex method [7] to minimize E.

2.2 Segment-orientation Estimation

We define the axes of the local coordinate system of each segment as follows; x: leftward, y: upward and z: forward (in an upright posture). As for each of the segments constituting the limbs, its long axis gives the y-direction. We identify the x- and z-directions using the information on joint rotation. **Figure 3** shows a joint rotation of the limb. $\boldsymbol{l}_{m,O}(i) \times \boldsymbol{l}_{m+1,O}(i)$ in Fig. 3 gives both the mth and $(m+1)$th segments the x-direction at the ith frame (for the leg, $-x$-direction for the arm). However, the direction of the above outer product becomes indefinite when $\theta(i)\approx 0$ ($\theta(i)$: rotation angle). Therefore, we use the vector $\boldsymbol{x}(i)$ (for the leg, $-\boldsymbol{x}(i)$ for the arm) as that giving the x-direction at the ith frame:

$$\boldsymbol{x}(i) = \boldsymbol{P}_{k,O}(i)\bar{\boldsymbol{x}}_k\{\pi - \theta(i)\} + \frac{\boldsymbol{l}_{m,O}(i) \times \boldsymbol{l}_{m+1,O}(i)}{|\boldsymbol{l}_{m,O}(i) \times \boldsymbol{l}_{m+1,O}(i)|}\theta(i) \tag{6}$$

$$\bar{\boldsymbol{x}}_k = \frac{\boldsymbol{x}_k}{|\boldsymbol{x}_k|}, \quad \boldsymbol{x}_k = \sum_{i=1}^{M}\boldsymbol{P}_{k,O}(i)^{-1}\frac{\boldsymbol{l}_{m,O}(i) \times \boldsymbol{l}_{m+1,O}(i)}{|\boldsymbol{l}_{m,O}(i) \times \boldsymbol{l}_{m+1,O}(i)|}\theta(i)$$

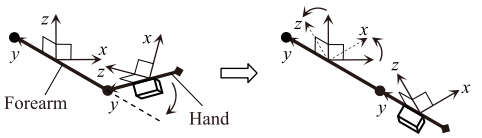
where $\boldsymbol{P}_{k,O}(i)$ is the orientation of the kth sensor attached to the mth segment. When $\theta(i)\approx 0$, the direction of $\bar{\boldsymbol{x}}_k$ (i.e., the average direction of the outer product weighted by $\theta(i)$) is dominant, whereas the instant direction of the outer product itself becomes dominant as $\theta(i)$ approaches π. The z-direction is finally obtained from the outer product of $\boldsymbol{x}(i)$ (or $-\boldsymbol{x}(i)$) and $\boldsymbol{l}_{m,O}(i)$ or $\boldsymbol{l}_{m+1,O}(i)$.

Flexion of the knee is reproduced by the above method. As for the arm, the elbow is regarded as a joint giving a twist in addition to flexion. We obtain the twist angle as follows. First, we obtain a time series of the orientation of the hand by giving it a particular orientation at a particular frame (i_0th frame):

$$\boldsymbol{R}_{H,O}(i) = \boldsymbol{P}_{H,O}(i)\{\boldsymbol{P}_{H,O}(i_0)^{-1}\boldsymbol{R}_{H,O}(i_0)\} \tag{7}$$

where $\boldsymbol{R}_{H,O}(i)$ and $\boldsymbol{P}_{H,O}(i)$ are the orientations of the hand and the sensor attached to the hand at the ith frame, respectively; the orientations of the foot and head are determined in the same way [*1].

[*1] In the present method, giving the hands, feet and head particular orientations at a particular frame is the only procedure users have to carry out in advance.



1. Rotate the hand. (Match the y-directions.)
2. Twist the forearm. (Match the x- and z-directions.)

Fig. 4 Determination of the twist angle of the forearm.

Then, the procedure of **Fig. 4** is executed. The forearm is twisted so as to match its x- and z-directions to those of the hand, after the y-direction of the hand is matched to that of the forearm.

Meanwhile, we obtain the orientation of the pelvis as follows:

$$\boldsymbol{R}_{\mathrm{P,O}}(i) = [\boldsymbol{x}_{\mathrm{P}}(i)/|\boldsymbol{x}_{\mathrm{P}}(i)| \quad \boldsymbol{y}_{\mathrm{P}}(i)/|\boldsymbol{y}_{\mathrm{P}}(i)| \quad z_{\mathrm{P}}(i)/|z_{\mathrm{P}}(i)|] \tag{8}$$
$$\boldsymbol{x}_{\mathrm{P}}(i) = \boldsymbol{y}_{\mathrm{P}}(i) \times z_{\mathrm{P}}(i),$$
$$\boldsymbol{y}_{\mathrm{P}}(i) = \boldsymbol{q}_{\mathrm{W,O}}(i) - \{\boldsymbol{q}_{\mathrm{LH,O}}(i) + \boldsymbol{q}_{\mathrm{RH,O}}(i)\}/2,$$
$$z_{\mathrm{P}}(i) = \{\boldsymbol{q}_{\mathrm{RH,O}}(i) - \boldsymbol{q}_{\mathrm{W,O}}(i)\} \times \{\boldsymbol{q}_{\mathrm{LH,O}}(i) - \boldsymbol{q}_{\mathrm{W,O}}(i)\}$$

where $\boldsymbol{q}_{\mathrm{LH,O}}(i)$ and $\boldsymbol{q}_{\mathrm{RH,O}}(i)$ are the positions of LH and RH. To determine the orientations of the segments constituting the chest, we first obtain the orientation of the whole chest:

$$\boldsymbol{R}_{\mathrm{C,O}}(i) = [\boldsymbol{x}_{\mathrm{C}}(i)/|\boldsymbol{x}_{\mathrm{C}}(i)| \quad \boldsymbol{y}_{\mathrm{C}}(i)/|\boldsymbol{y}_{\mathrm{C}}(i)| \quad z_{\mathrm{C}}(i)/|z_{\mathrm{C}}(i)|] \tag{9}$$
$$\boldsymbol{x}_{\mathrm{C}}(i) = \boldsymbol{y}_{\mathrm{C}}(i) \times z_{\mathrm{C}}(i),$$
$$\boldsymbol{y}_{\mathrm{C}}(i) = \{\boldsymbol{q}_{\mathrm{LS,O}}(i) + \boldsymbol{q}_{\mathrm{RS,O}}(i)\}/2 - \boldsymbol{q}_{\mathrm{W,O}}(i),$$
$$z_{\mathrm{C}}(i) = \{\boldsymbol{q}_{\mathrm{LS,O}}(i) - \boldsymbol{q}_{\mathrm{W,O}}(i)\} \times \{\boldsymbol{q}_{\mathrm{RS,O}}(i) - \boldsymbol{q}_{\mathrm{W,O}}(i)\}$$

where $\boldsymbol{q}_{\mathrm{LH,O}}(i)$ and $\boldsymbol{q}_{\mathrm{RH,O}}(i)$ are the positions of LS and RS. The orientation of the segment W-C is given as follows:

$$\boldsymbol{R}_{\mathrm{WC,O}}(i) = \boldsymbol{R}_{\mathrm{C,O}}(i)\boldsymbol{R}(\theta_{\mathrm{WC}}(i), \boldsymbol{n}_{\mathrm{WC}}(i)) \tag{10}$$
$$\theta_{\mathrm{WC}}(i) = \cos^{-1}([0\ 1\ 0]^{\mathrm{T}} \cdot \boldsymbol{n}'_{\mathrm{WC}}(i)),$$
$$\boldsymbol{n}_{\mathrm{WC}}(i) = [0\ 1\ 0]^{\mathrm{T}} \times \boldsymbol{n}'_{\mathrm{WC}}(i),$$
$$\boldsymbol{n}'_{\mathrm{WC}}(i) = \boldsymbol{R}_{\mathrm{C,O}}(i)^{-1}\{\boldsymbol{q}_{\mathrm{C,O}}(i) - \boldsymbol{q}_{\mathrm{W,O}}(i)\}/|\boldsymbol{q}_{\mathrm{C,O}}(i) - \boldsymbol{q}_{\mathrm{W,O}}(i)|$$

where $\boldsymbol{R}(\theta_{\mathrm{WC}}(i), \boldsymbol{n}_{\mathrm{WC}}(i))$ is the rotation with the angle $\theta_{\mathrm{WC}}(i)$ and axis $\boldsymbol{n}_{\mathrm{WC}}(i)$. The orientation of the segment C-N is obtained in the same way. The orientations of the segments C-LS and RS-C are obtained by replacing $[0\ 1\ 0]^{\mathrm{T}}$ with $[1\ 0\ 0]^{\mathrm{T}}$.

3. Results

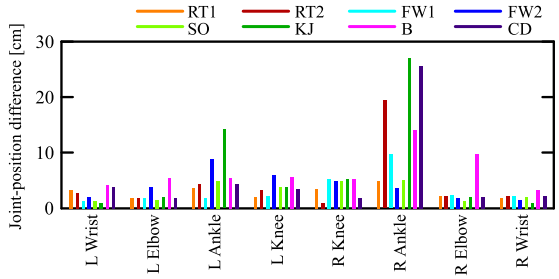
We report experimental results in this section. The Mocap data used are shown in **Table 1**. We selected them from various motion categories. All the data in Table 1 are acquired using 15 sensors configured as shown in Fig. 1 (including the dotted boxes). In the cases that the present method is applied, we use only the data of the 11 sensors represented as the white boxes in Fig. 1. The obtained results are compared with those given by LLSF or GDKE both using the data of all the 15 sensors.

Figure 5 shows the difference of joint positions between the present method and LLSF or GDKE. The values shown in the figure are those of the distance between the same joints averaged over an entire time series. Since the positions of the joints in the torso are given by LLSF or GDKE also in the present method, we compare only the joints of the limbs. According to Ref. [4], an error in joint-position estimation from magnetic Mocap data can reach several centimeters. In Fig. 5, the difference in joint positions with respect to the arms is within the above range in almost all the cases. On the other hand, the difference in ankle position was remarkably deviated from the above range in many cases.

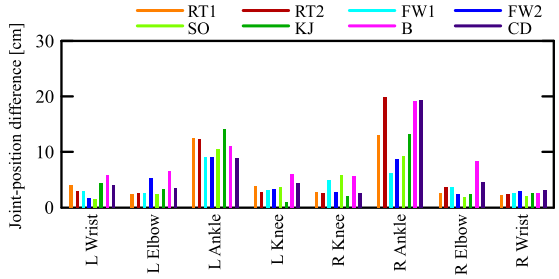
Table 1 Motion capture data (frame rate: 30 fps).

Label	Category	Frames	Subject	System
RT1	*Rajio Taisō*, No.1 (Japanese radio calisthenics)	6,283	A	A
RT2	*Rajio Taisō*, No.2 (part) (Japanese radio calisthenics)	3,944	B	A
FW1	Farm working (cultivating a field, rice planting)	2,389	C	B
FW2	Farm working (rice reaping, sawing)	2,443	C	B
SO	*Sansa Odori* (Japanese folk dance)	3,124	C	B
KJ	*Jinku Odori* from *Kemanai Bon Odori* (Japanese folk dance)	2,199	D	B
B	Ballet (échappé, passé, entrechat quatre, piqué en dehors)	1,508	B	A
CD	Contemporary dance (release technique)	2,754	B	A

Subject A: male, B, C and D: female
System A: MotionStar Wireless™ (Ascension Technology Corporation)
B: MotionStar Wireless™ with LIBERTY™ (Polhemus) ×2



(a) Difference between LLSF and LLSF (torso) + Present (limbs) .

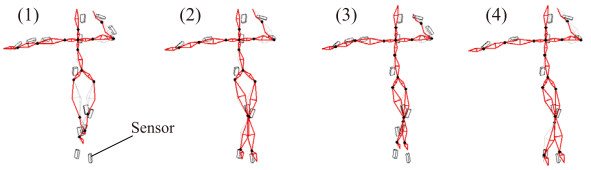


(b) Difference between GDKE and GDKE (torso) + Present (limbs) .

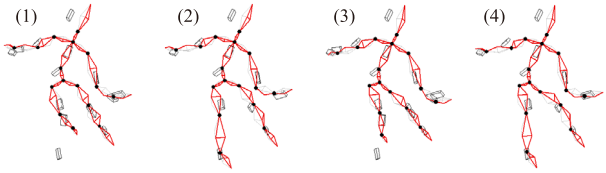
Fig. 5 Average joint-position difference between the present method and the other methods.

Figure 6 shows two examples of the obtained skeletons. These gave especially a large difference in ankle position. In the cases that only LLSF or GDKE were applied, the position of the ankle (or both the ankles) was too far away from the sensor attached to the foot. In the cases that the present method was used, in contrast, the positions of both the ankles were properly estimated, even though the number of sensors was reduced. The same tendency is seen in all the cases in which a significant difference in the ankle position is shown. As for the orientation of each segment, an appropriate estimation according to the rotation of each joint was carried out as long as joint positions were properly estimated. Errors caused by improper joint-position estimation are seen in the orientations of the thighs in (1) of (a) in Fig. 6.

It should also be pointed out that the present method has a disadvantage in calculation time. As shown in **Table 2**, the calculation time of the present method is several times longer than that of



(a) *Jinku Odori* from *Kemanai Bon Odori* (Japanese folk dance, subject: D).



(b) Contemporary dance (release technique, subject: B).

(1): LLSF, (2): LLSF (torso) + Present (limbs),
(3): GDKE, (4): GDKE (torso) + Present (limbs).
Planes surrounded by red lines: parallel to the frontal plane [8] in an upright posture

Fig. 6 Examples of the obtained skeletons.

Table 2 Calculation time.

Label	Calculation time [s]			
	LLSF	LLSF+Present	GDKE	GDKE+Present
RT1	6	27	2	25
RT2	4	13	1	15
FW1	2	8	1	8
FW2	2	9	1	8
SO	3	10	1	9
KJ	2	7	1	7
B	2	6	1	5
CD	3	9	1	9

CPU: Intel Core i3-350M

the other methods. This is caused by iterative calculations to minimize the cost function used for joint-position estimation. Further research efforts are needed to address this issue.

4. Conclusions

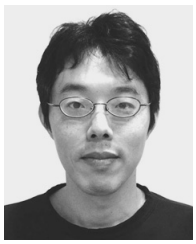
The main contribution of this paper is to reduce the number of sensors in skeleton estimation from magnetic Mocap data. The experimental results showed the effectiveness of the present method. However, the issue of long calculation time still remains unresolved. This will be the subject of future work.

Acknowledgments We would like to thank Associate Professor Naho Matsumoto, Akita University, for her advice about the description of dance names. This study was supported by Grants-in-Aid for Scientific Research (No.26370942) of Japan Society for the Promotion of Science.

References

[1] Kitagawa, M. and Windsor, B.: *MoCap for Artists*, Focal Press (2008).
[2] Kirk, A.G., O'Brien, J.O. and Forsyth, D.A.: Skeletal Parameter Estimation from Optical Motion Capture Data, *CVPR 2005*, pp.782–788 (2005).
[3] Xiao, Z., Nait-Charif, H. and Zhang, J.J.: Real Time Automatic Skeleton and Motion Estimation for Character Animation, *Comp. Anim. Virtual Worlds*, Vol.20, No.5-6, pp.523–531 (2009).
[4] O'Brien, J.O., Bodenheimer, R.E. Brostow, G.J. and Hodgins, J.K.: Automatic Joint Parameter Estimation from Magnetic Motion Capture Data, *Proc. Graphics Interface 2000*, pp.53–60 (2000).
[5] Knight, J.K. and Semwal, S.: Fast Skeleton Estimation from Motion Capture Data using Generalized Delogne-Kåsa Method, *WSCG '2007 Full Papers Proceedings*, pp.225–232 (2007).
[6] Kouchi, M., Mochimaru, M., Iwasawa, H. and Mitani, S.: *Anthropometric Database for Japanese Population 1997-98*, Japanese Industrial Standards Center (AIST, MITI) (2000).
[7] Nelder, J.A. and Mead, R.: A Simplex Method for Function Minimization, *The Computer Journal*, Vol.7, No.4, pp.308–313 (1965).
[8] Bartlett, R.: *Introduction to Sports Biomechanics*, 2nd ed., Routledge (2008).

Takeshi Miura received his D.Eng. degree in electrical engineering from Hokkaido University in 1998. He is currently an associate professor in the Department of Electrical and Electronic Engineering, Graduate School of Engineering and Resource Science, Akita University.



Takaaki Kaiga received his M.E. degree in mechanical engineering from Ibaraki University in 1995. Since 1996, he has been with the computer division of Digital Art Factory, Warabi-za Co., Ltd.

Katsubumi Tajima received his D.Eng. degree in electrical engineering from Tohoku University in 1998. He is a professor in the Cooperative Major in Life Cycle Design Engineering, Graduate School of Engineering and Resource Science, Akita University.

Takeshi Shibata received his D.Eng. degree in electrical engineering from Akita University in 2012. He is currently a post-doctoral researcher in the Venture Business Laboratory, Akita University. His research interests include virtual reality technique and archiving and handing-down technique for traditional folk dances.

Hideo Tamamoto received his D.Eng. degree in electrical engineering from the University of Tokyo in 1976. He is currently a professor/dean of Tohoku University of Community Service and Science. His research interests include design-for-testability of logic circuits, archiving and passing-down technique of traditional folk dances, and e-learning system.