# Human Spine Posture Estimation from Video Images Based on Connected Vertebra Spheres Model

Daisuke FURUKAWA, Kensaku MORI and Yasuhito SUENAGA
Dept. of Computational Science and Engineering,
Graduate School of Engineering,
Nagoya University, Japan
{dfuru,mori,suenaga}@suenaga.cse.nagoya-u.ac.jp

## Abstract

*This paper reports a method for estimating human spine posture using the front and side views of a human body taken by a video camera. We present here a new 3D model to estimate the spine posture using connected vertebra spheres. In this model, each vertebra forming the spine is approximated as a sphere. The spine is approximated as a series of spheres connecting each other. Each sphere has the control points to represent the outer shape of the body. The spine posture is estimated by moving the sphere and the control points. The estimation is performed by calculating matching ratio between a projected image of the model and an input image. X-ray CT slices are used for the construction of the model. We applied the proposed method to real human images. The experimental results showed that our 3D model worked reasonably for estimation of human spine posture based on real human images.*

## 1 Introduction

In the computer animation field, posture estimation of the human body is a useful technique for generating animations automatically. Suppose that we generate an animation of a human motion. Current systems generally generate motion sequences by arranging a basic data of motions such as walking or running. The basic data should be prepared on the system in advance. Users generate complex motions by modifying these data manually. However it is usually difficult to make appropriate data for generating naturally looking motion. Also it spends a lot of time or cost. One approach to solve this problem is to use a motion capture technology [5]. For using typical motion capture systems, the user should wear special suits including markers for detecting motion by an optical equipment. If we could directly estimate the human motion from video images, these inconvenient situation would be solved. It could be possible to generate a real animation based on a actual human motion easily. In such a system, the posture estimation based on conventional video images plays an important role.

Many human models are proposed for estimating the human posture. In these models, each part of the human body is approximated as a cylinder or an ellipsoid [10, 7], a polygon mesh [2, 11] and more complex object such as metaball [8]. Postures are generated by controlling the positions of cylinders or ellipsoids and the joint angles. Some models are available for tracking the human motion from video sequences [1, 3, 4, 8, 9]. Boulic et al. [9] proposed the anatomically correct human body model. This model has 31 anatomic joints and totally 62 degrees of freedom.

But these models are not available for estimating spine posture with a non-rigid or a flexible structure, since each segment is controlled under simple kinematic constraint. In particular, it is impossible to consider the structure of the spine such as the shape of a vertebra. To estimate the spine posture, the model needs to approximate the human spine precisely. In addition, the model should include the outer shape of the human body since the human spine is invisible.

This paper describes a new model and a method for estimating a spine posture from human images taken by video cameras. In this model, the spine is approximated as connecting spheres. Each sphere has control points that defines the outer shape of a body. By constraining the movement of spheres and associated control points, it is possible to estimate a spine posture by performing matching between the outer shape of the generated model and the input image. In Section 2, we briefly describe the concept of the spine posture model and its construction procedure. This section also provides basic rules to constrain the movement of the vertebra sphere and associated control points and methods for generating actual postures such as bending and twist postures. Section 3 shows processing procedures for estimating postures from video images by using the proposed model.

0-7803-7237-9/01/$10.00 ©2001 IEEE

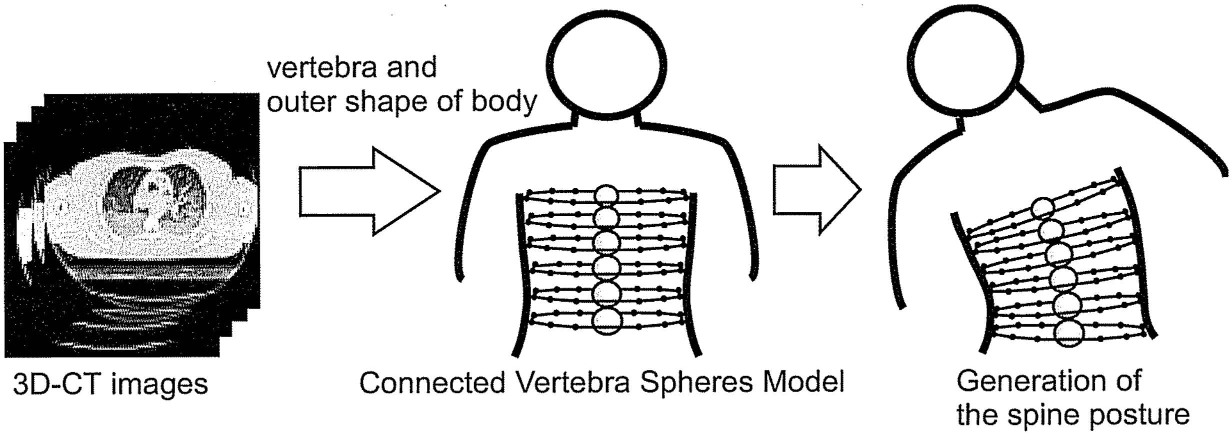



**Figure 1. Concept of the spine posture estimation model.**

In Section 4, we show experimental results of application to real human images and discussion.

## 2 Connected Vertebra Spheres Model

### 2.1 Overview

Our Spine posture estimation model consists of connected spheres and control points as shown in Fig.1. The connected spheres are formed by a series of vertebra spheres that approximate the vertebrae. Postures are generated by rotating these spheres and moving control points under the constraints defined in the following section.

Each vertebra sphere is associated with control points for representing the outer shape of a human body. These control points are placed on a control plane associated with each vertebra sphere as shown in Fig.2. As a vertebra sphere rotates, the associated control plane rotates, and the control points on it rotate accordingly. This operation is repeated for all the vertebra spheres. The 3D shape of a human body is obtained by forming polygons based on the control points on adjacent control planes.

We use three dimensional (3D) X-ray CT images of a whole human body to construct an initial model. It is difficult to take CT images for only the purpose of the posture estimation because of an influence of radiation exposure and the cost. Therefore, we construct a generic model by using a set of CT images provided by the Visible Human Project [6]. A personal model is generated by deforming the generic model in the estimation process. The generic model is constructed by the following processes. Firstly, bone regions are extracted from input CT images by thresholding. We obtain the front and side views by projecting the extracted regions. The center and the radius of each vertebra sphere are determined from these views manually. The control points of each sphere are located on the contour of the subject person every 30 degrees manually.

As shown in Fig.2, we denote the $i$-th vertebra sphere as $S_i (i = 1, 2, ..., n)$, where $S_1$ and $S_n$ correspond to the waist and the neck position, respectively.

### 2.2 Rotating vertebra spheres and control points

We give the limitation against the movement of the vertebra spheres and control points. Only two primitive operations are allowed to our model under the constraint.

#### (a) Rotation of the control points around a certain axis

This operation is to rotate control points associated with a vertebra sphere around an axis. Now let us consider the rotation of the control points of $S_{i+1}$ through the angle $\theta$ as shown in Fig.2(a). The rotation axis is defined as the line that passes through the center point of $S_{i+1}$. This axis is perpendicular to the plane made by control points of $S_{i+1}$. The control points of $S_j (j = i+1, ..., n)$ are rotated around this axis.

#### (b) Rotation of both the vertebra sphere and the associated control points

This operation is to rotate both the vertebra sphere and control points. The movement of each vertebra is limited so that it can roll only on the adjacent spheres. A vertebra has two degrees of freedom under this constraint. Consequently, this operation needs two parameters, the direction and the angle of rotation of a vertebra sphere.

Now let us rotate $S_{i+1}$ through the angle $\phi$ against $S_i$. $S_{i+1}$ is rotated through the angle $\phi$ around the center of $S_i$ as shown in Fig.2(b). The control points of $S_{i+1}$ rotate around a center of $S_{i+1}$ by the angle $\phi'$ that is calculated by the following equation

$$\phi' = \frac{r_i + r_{i+1}}{r_{i+1}} \times \phi, \qquad (1)$$

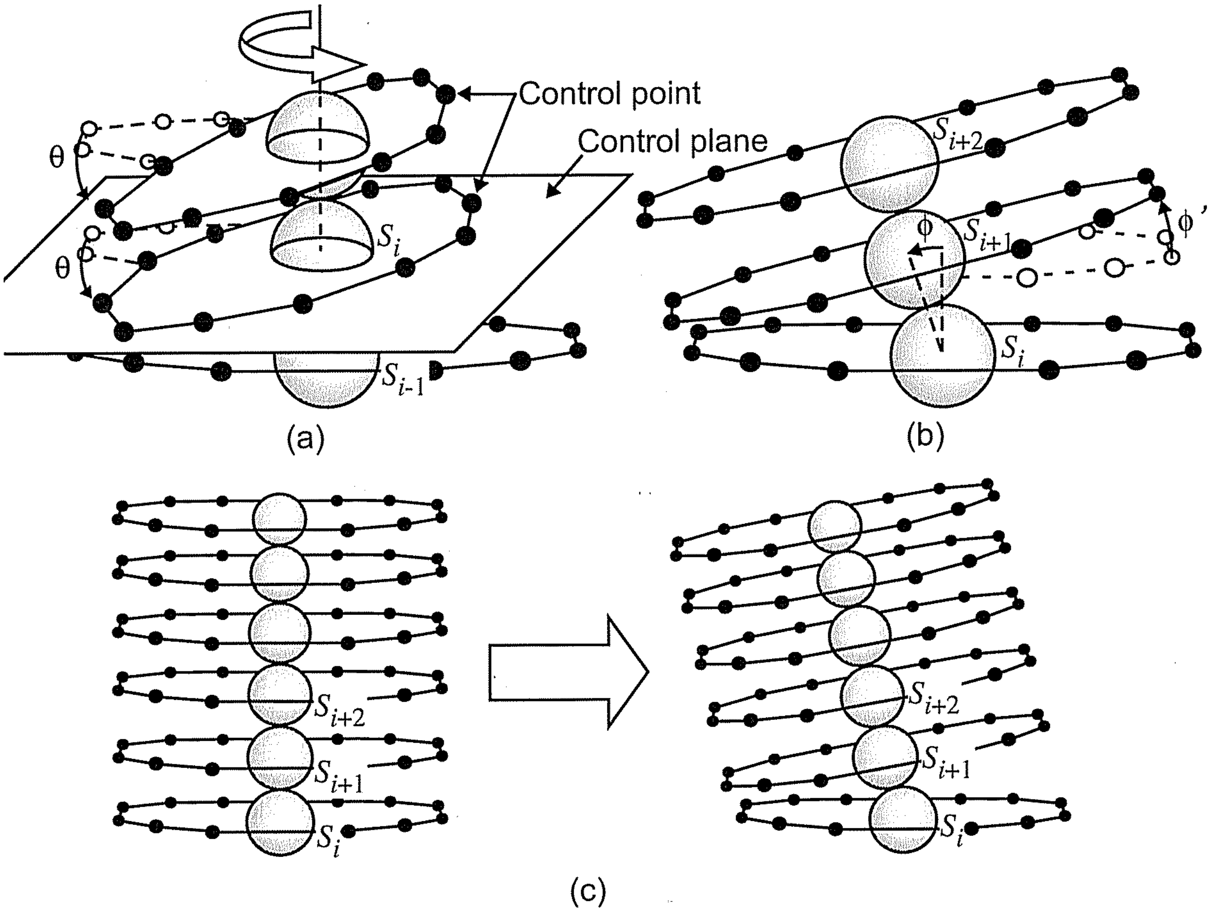



**Figure 2. Rotational operation of vertebra spheres and control points.**

where $r_i$ and $r_{i+1}$ are the radius of $S_i$ and $S_{i+1}$, respectively. The whole set of spheres $S_j (j = i + 2, ..., n)$ is treated as if it were a solid object in this rotation operation. As the solid object rotates, the associated control points also rotate accordingly as shown in Fig.2(c).

## 3 Posture estimation

Fig.3 shows the processing flow of the spine posture estimation.

### 3.1 Generation of the personal model

The model for estimating the spine posture is generated based on a subject person's CT images. We need to deform the generic model so as to fit it to a subject person's outer shape.

The input is given to the system as two images of the front and side views of human body taken while the subject person keeps standing posture. These images are taken by two video cameras as shown in Fig.4(a). The width $D_1$, the

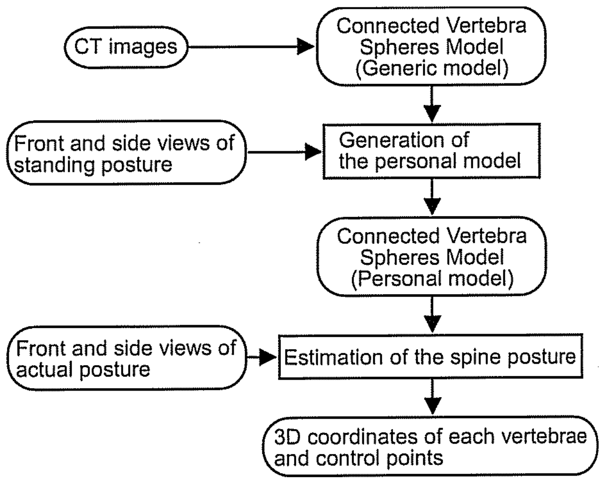



**Figure 3. Flow of the spine posture estimation**

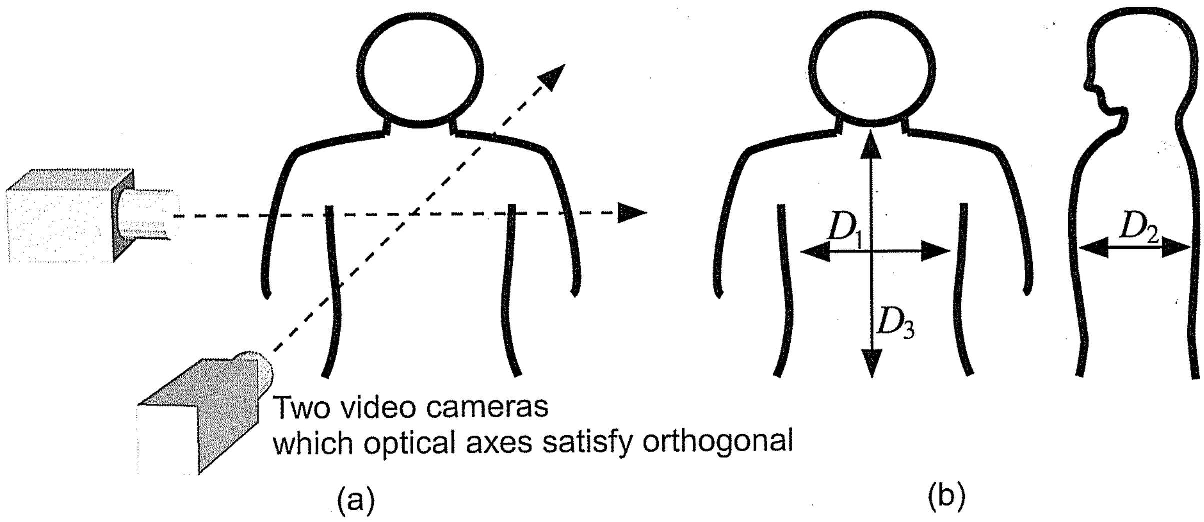



**Figure 4. Generation of the personal model.**

thickness $D_2$, and the height $D_3$ are measured on the acquired images (Fig.4(b)). The generic model is transformed to the personal model by adopting the following deformation operation,

$$\begin{pmatrix} x_i \\ y_i \\ z_i \\ r_i \end{pmatrix} \leftarrow \begin{pmatrix} D_1/d_1 & & & 0 \\ & D_2/d_2 & & \\ & & D_3/d_3 & \\ 0 & & & D_3/d_3 \end{pmatrix} \begin{pmatrix} x_i \\ y_i \\ z_i \\ r_i \end{pmatrix}, \quad (2)$$

$$\begin{pmatrix} c_x^{ij} \\ c_y^{ij} \\ c_z^{ij} \end{pmatrix} \leftarrow \begin{pmatrix} D_1/d_1 & & 0 \\ & D_2/d_2 & \\ 0 & & D_3/d_3 \end{pmatrix} \begin{pmatrix} c_x^{ij} \\ c_y^{ij} \\ c_z^{ij} \end{pmatrix}, \quad (3)$$

where $(x_i, y_i, z_i)$ is the position of $S_i$, $r_i$ the radius of $S_i$, $(c_x^{ij}, c_y^{ij}, c_z^{ij})$ the position of the $j$-th control point of $S_i$. $d_1, d_2$ and $d_3$ are the width, the thickness and the height of the model, respectively.

### 3.2 Generation of postures

The process of the posture generation consists of two stages: (a)generation of postures by applying two primitive operations to the model, (b)generation of the outer shape of the model by making the rectangle polygons with adjacent four control points. The actual postures are divided into two types: basic type and composite type.

The basic type postures are composed of flexion or extension, lateral bending and twist. The basic type postures are generated by iterating the primitive operation. The posture with twist is generated by using the operation described in 2.2(a). The operation described in 2.2(b) supports the generation of two postures with flexion or extension and lateral bending by exchanging the rotational direction of the vertebra sphere. Suppose that we generate the posture of flexion. First, $S_{i+1}$ is rotated against $S_i$ by using the operation described in 2.2(b). Next, $S_{i+2}$ is rotated against $S_{i+1}$ in the same way. Thus, the posture of flexion is generated by applying this operation to the vertebra sphere from $S_{i+1}$ to $S_n$ sequentially.

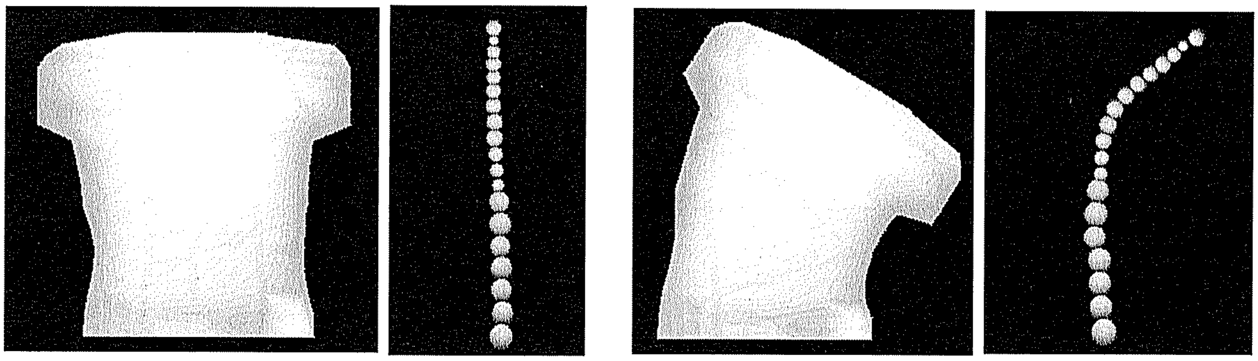
The composite type postures are formed by combining basic type postures. First, the posture with flexion or extension is generated by the operation described in 2.2(b). Next, we generate the composite posture with flexion or extension and lateral bending by applying the operation described in 2.2(b) to the posture obtained by the previous step. Finally, the combination posture is generated by applying the operation described in 2.2(a).

### 3.3 Posture estimation

Our method uses two images as shown in Fig.4(a) since the portions of the body invisible in one view are obtained in another view. Two cameras taking the input images are placed at the same position and direction when they were used in the preprocessing. The method for estimating the spine posture is described as follows.

We use the number of pixels in the exclusive regions between the two regions, the external regions of the model projected onto the image plane and the silhouette region to measure the matching ratio of the posture.

First, silhouette images corresponding to two input images are generated. At this time, the position and the orientation of the model are arranged where the model's waist corresponds to the waist on the silhouette images. Next, the posture candidates are generated by using methods described in 3.2. The rotational angle parameters $\theta_{fb}, \theta_{side}$ and $\phi$ needed for using the primitive operations are set once before the posture estimation procedure. From the given pa-



(a) (b)

**Figure 5. Connected Vertebra Spheres Model. (a)Model's outer shape and spine posture with standing. (b)Model's outer shape and spine posture with left lateral bending.**

rameters, totally 729 sets of the parameters are made by the following equations,

$$\Theta_{fb} = \theta_{fb} \times n, \quad (4)$$

$$\Theta_{side} = \theta_{side} \times n \quad (-4 \leq n \leq 4), \quad (5)$$

$$\Phi = \phi \times n. \quad (6)$$

From each set of the parameters, posture is generated by applying the method of posture generation. Finally, the outer shapes of all the models are generated, and are projected onto the silhouette images. The matching ratio between the silhouette and the outer shape is evaluated, and the posture with the highest score is estimated as the current posture. The outputs are the 3D coordinates of each vertebrae and the associated control points.

## 4 Experimental results and discussion

### 4.1 Generation of spine posture estimation model

Fig.5 (a) shows the spine posture estimation model generated by the method described in 2. Fig.5 (b) shows a left lateral bending posture generated by applying the method described in 3.2.

### 4.2 Evaluation of spine posture estimation

We applied our proposed method to the real human images to evaluate efficiency of spine posture estimation with our model. The input are front and side view images taken by the two digital video cameras (SONY DCR-VX1000). These are located in the position so that their optical axes are orthogonal. The image size is 640 x 480 pixels. Totally 69 couples of images were examined for evaluating the proposed method. As our model currently does not include arms, a subject person has to fold his arms on head to avoid its influence on the performance of the estimation. Fig.6 shows an example of the tested images.

**Table 1. An estimation result.This table shows the number of times that we have estimated the posture correctly.**

| | Number of times |
|---|---|
| Flexion | 10 (43.5%) |
| Left lateral bending | 9 (39.1%) |
| Twist | 1 (4.3%) |

Since we can not currently find out the really correct spine posture, we compare the obtained body silhouette with the subject's silhouette in order to evaluate the estimation result.

Table 1 shows the experimental result and Fig.6 shows an example of estimation result. The gray area at the center of image corresponds to the external silhouette of the model, and the connected spheres through the gray area correspond to the estimated spine.

Our method was run on an SGI Octane (R12000 300MHz), and the computation time is about 23 seconds.

### 4.3 Discussion

It turned out that the flexion posture can be estimated correctly as shown in Fig.6. Since the spine is modelled as the connected spheres, our model is able to represent a bending spine with a flexible structure. Also we can estimate the invisible spine posture by using the video camera images that are available easily. This is the main advantage of our method.

Fig.7 shows a failure example for estimating the bending posture. We can see that the bottom portion of the model is exposed. Since the model does not currently include a under parts of the waist, an orientation of the bottom portion of

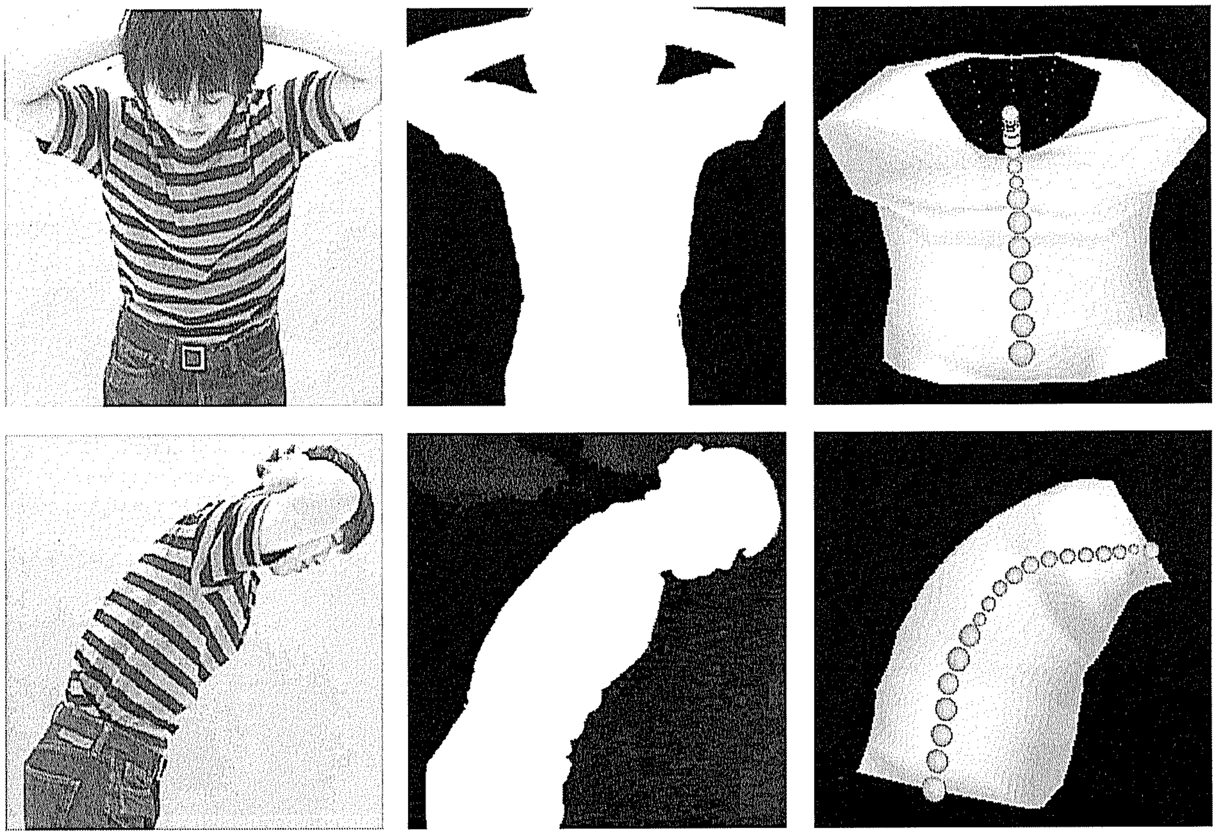


**Figure 6. Experimental result to estimating forward bending. Left side column shows a couple of the input images. Center column shows silhouette images corresponding the left side column, Right side column shows the estimation results. This image displays by overlapping a spine posture on the image of the model's outer shape.**

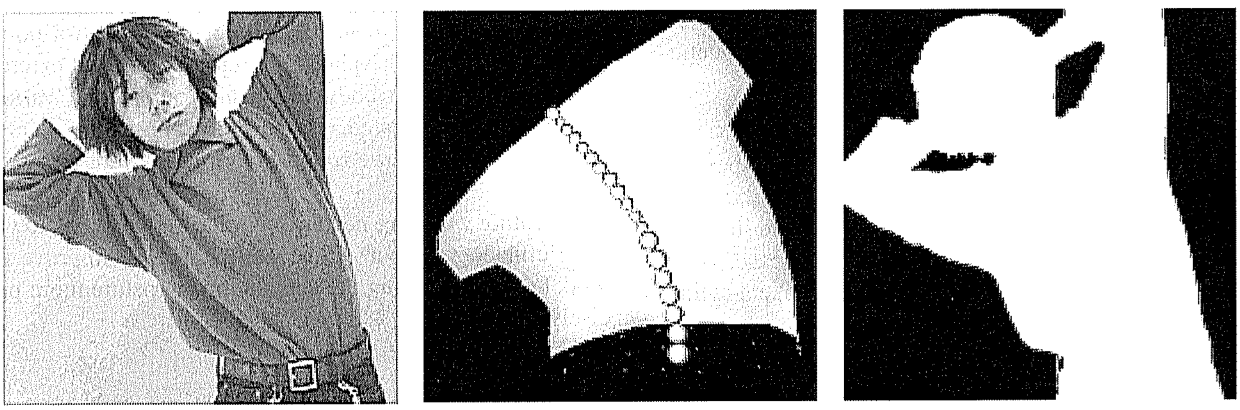


**Figure 7. Examples of failure estimation of bending posture.**

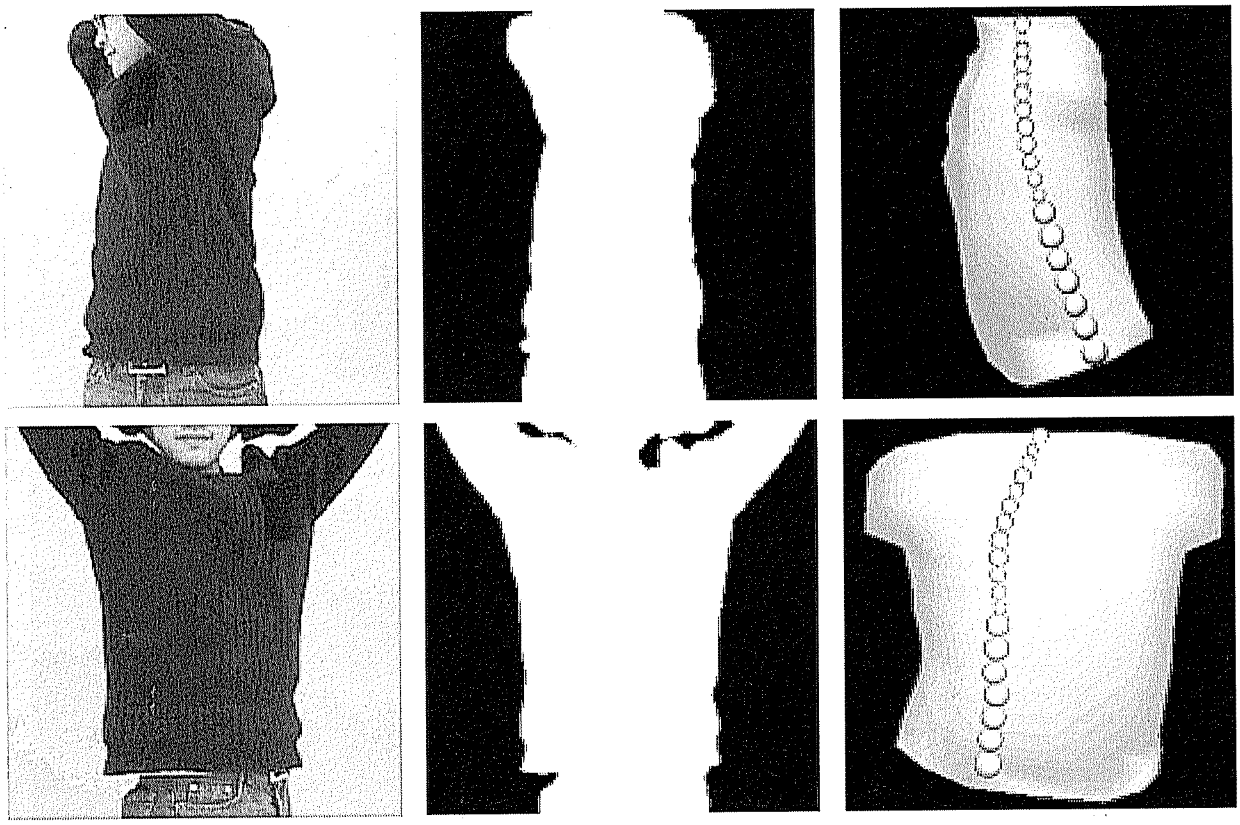


**Figure 8. Examples of failure estimation of twist posture.**

the model affects the measurement of matching. To avoid this influence, we used a couple of images so that the occluded area can be complemented mutually. However, the estimation result was not improved. This indicates that the measurement of matching is affected by the bottom portion of the model, rather than the difference between silhouette and the external of the model.

It is difficult to estimate the twist posture. Since people usually wears the clothes and his clothes occlude the contour of his own body, its silhouette vary only little in twist posture. Fig.8 shows a failure example in this case. The body of the subject in two silhouette images is like a cylinder, and the estimation is failed for that influence.

From the discussion described above, it is reasonable to use the neck and the waist position for the estimation in addition to the silhouette information, In this approach, we first extract the neck and the waist position from input images by a method applicable to estimating a posture of whole human body. Based on the information thus obtained, the connected spheres are deformed under the following constraints, (1) both ends of the connected spheres are located on the positions of the neck and the waist, (2) the external silhouette of the model fits to the input silhouette.

## 5 Conclusion

In this paper, we have proposed a model and a method for estimating a spine posture from two images taken by two video cameras. This model consists of the connected spheres that approximate the spine and the control points associated with each sphere for representing the external shape of the human body. By using our model, various postures were generated. In the processing of 69 pairs of real human images, we found fifteen pairs were estimated correctly.

As future work, we need to examine more images to validate efficiency of our estimation method, to improve the model for representing the actual human posture more precisely.

## References


[1] Adrian Hilton, Daniel Beresford, Thomas Gentils, Raymond Smith and Wei Sun. Virtual people: Capturing human models to populate virtual worlds. *Proc. of Computer Animation 1999*, 1999.

[2] Ioannias A.Kakadiaris and Dimitris Metaxas. Model-based estimation of 3D human motion with occlusion based on active multi-viewpoint selection. *Proc. of the IEEE Computer Vision and Pattern Recognition Conf.*, 1996.

[3] Christoph Bregler and Jitendra Malik. Tracking people with twists and exponential maps. *Conference on Computer Vision and Pattern Recognition*, 1998.

[4] Quentin Delamarre and Olivier Faugeras. 3D articulated models and multi-view tracking with physical forces. *International Conference on Computer Vision*, 1999.

[5] L.Hedra, P.Fua, R.Plänkers, R.Boulic and D.Thalmann. Skeleton-based motion capture for robust reconstruction of human motion. *Proc. of Computer Animation 2000*, 2000.

[6] NATIONAL LIBRARY OF MEDICINE. *The Visible Human Project*. http://www.nlm.nih.gov/research/visible/visible_human.html.

[7] James M.Rehg and Takeo Kanade. Model-based tracking of self-occluding articulated objects. *International Conference on Computer Vision*, pages 612–617, 1995.

[8] Ralf Plänkers and Pascal Fua. Tracking and modeling people in video sequences. *International Journal of Computer Vision and Image Understanding*, 81(3), 2001.

[9] Ronan BOULIC, Pascal FUA, Lorna HERDA, Marius SILAGHI, Jean-Sébastien MONZANI, Luciana NEDEL and Daniel THALMANN. An anatomic human body for motion capture. *EMMSEC'98*, 1998.

[10] Tadahiko Kimoto, Akihiko Kajigaya and Yasuhiko Yasuda. A method of analyzing a human walker from monocular moving pictures based on stick models. *Transactions of the institute of electronics, Information and communication engineers(D-II)*, J74-D-II(3):376–387, 1991.

[11] Yoshinari Kameda, Michihiko Minoh and Katsuo Ikeda. A pose estimation method for an articulated object from its silhouette images. *Transactions of the institute of electronics, Information and communication engineers(D-II)*, J74-D-II(1):26–35, 1996.