MVA2002 IAPR Workshop on Machine Vision Applications, Dec. 11-13, 2002, Nara-ken New Public Hall, Nara, Japan

6—1

Human Spine Posture Estimation from 2D Frontal and Lateral Views Using 3D Physically Precise Spine Model

Daisuke FURUKAWA* Kensaku MORI† Yasuhito SUENAGA‡
*‡Graduate School of Engineering, Nagoya University
†Research Center for Advanced Waste and Emission Management, Nagoya University

Abstract

This paper describes a physically precise spine model and its application to estimate three dimensional spine posture from the frontal and lateral views of a human body taken by two conventional video cameras. In our method, rough position of each part is estimated by using a simple model applicable to images. Based on the obtained features from the fitting process of the simple model, the forces acting on the precise model are computed. The precise posture is estimated by deforming the model under these forces. Since the forces acting on intervertebral disks are easily calculated, the proposed method is useful for giving good suggestions to a user to achieve healthy and safe body posture. From the experimental results of application to actual human images, we confirmed that the proposed method estimated 3D posture of the human spine reasonably.

1 Introduction

In the computer animation field, posture estimation is a useful technique for generating human motion automatically. Several human models for posture estimation have been proposed in the past [1, 2, 4, 5, 6]. However the resulting postures are not accurate in physical sense, since physical properties of each part of a human body and interactions among the others are not considered. Moreover, even if such a physically precise model is constructed, it would be difficult to estimate posture by applying the model to human images directly.

This paper describes a physically precise spine model and its application to estimate three dimensional spine posture from the frontal and lateral views of human body taken by two conventional video cameras. In order to solve the problems mentioned above, we first obtain the neck and waist positions by employing a simple silhouette-based matching method from the successive two sets of human images. Then the forces acting on the models are calculated from the obtained displacements of the neck and the waist positions and the time interval. Finally, the deformation of the spine model is determined based on the dynamic simulation.

*Address: Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8603 Japan. E-mail: dfuru@suenaga.cse.nagoya-u.ac.jp
†E-mail: mori@cse.nagoya-u.ac.jp
‡E-mail: suenaga@cse.nagoya-u.ac.jp

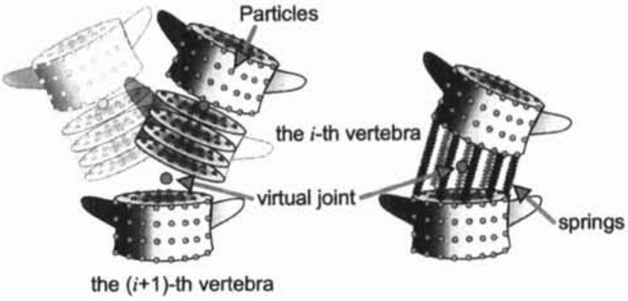

Figure 1: Accurate spine model. Each vertebra is modeled as a rigid body composed of a set of particles. The upper and lower vertebrae are connected by a set of a virtual joint and eight springs

2 Methods

2.1 Accurate spine model

Human spine is mainly composed of a set of vertebrae and intervertebral disks. We model each vertebra as a rigid body composed of a set of particles as shown in Fig.1. Physical parameters, such as the relative mass property, a center of mass, and components of an inertia tensor characterizing the physical properties of a vertebra can be calculated from the positions and the number of the particles. We use three dimensional X-ray CT images of a whole human body to construct an initial model. The position of a particle is determined by the position of the voxel in the corresponding bone region in a set of CT slices.

An intervertebral disk is modeled as a set of a joint with 3 DOF (degrees of freedom) and springs connecting the upper and the lower vertebrae. The movement of two vertebrae in the deformation process are constrained by the joint and the springs. The initial position of the virtual joint located is determined manually. The spring constant is determined experimentally.

2.2 Spine posture estimation

Figure 2 shows the processing flow of the spine posture estimation method using the accurate human spine model. The inputs of the method are the frontal and lateral views of a human body taken by two video cameras located at the places where their optical axes satisfy orthogonal.

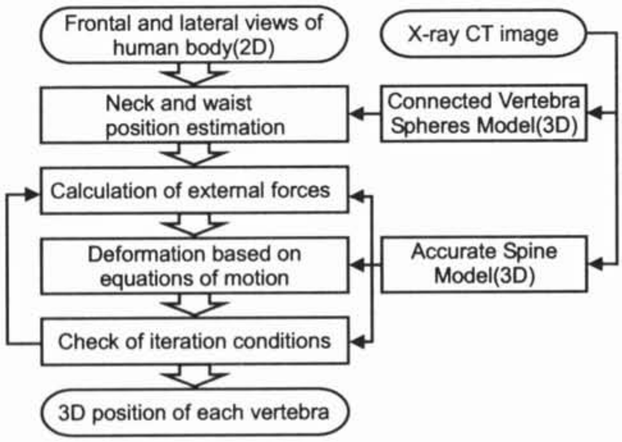

Figure 2: Processing flow of the spine posture estimation using the accurate spine model.

2.2.1 Rough estimation

We estimate the rough positions of the neck and waist on $\mathrm{I_f}(t)$ and $\mathrm{I_l}(t)$. Here, we denote $\mathrm{I_f}(t)$ and $\mathrm{I_l}(t)$ as the frontal and lateral views taken at time t. Silhouette images corresponding to $\mathrm{I_f}(t)$ and $\mathrm{I_l}(t)$ are first generated by using basic image processing operations including thresholding. Then posture candidates are generated using Connected Vertebra Spheres Model [3]. The outer shape of the model with respect to each candidate is generated and is projected onto the silhouette images. The number of pixels in the exclusive regions between the projected regions and the silhouette regions are calculated to measure a matching rate of the actual and the generated postures. The candidate with the highest matching rate is estimated as the current posture. We simply regard the top and bottom vertebrae of the estimated spine as the positions of the neck and waist, $\mathbf{p}_{\mathrm{neck}}(t)$ and $\mathbf{p}_{\mathrm{waist}}(t)$.

2.2.2 Precise estimation

(a) Deformation of the accurate spine model

The accurate spine model is deformed so as to locate the top and bottom vertebrae of the model on $\mathbf{p}_{\mathrm{neck}}(t)$ and $\mathbf{p}_{\mathrm{waist}}(t)$, respectively. The external force $\mathbf{F}_{\mathrm{ext}}$ acting on each vertebra is calculated. The force $\mathbf{F}_{\mathrm{ext}}$ is expressed as the sum of the gravitational force $\mathbf{f}_{\mathrm{g}}$ and the intervertebral disk force $\mathbf{f}_{\mathrm{disk}}$. Two virtual forces $\mathbf{f}_{\mathrm{neck}}$ and $\mathbf{f}_{\mathrm{waist}}$ are working on the top and the bottom vertebrae. After calculating the external forces, a set of equations of motion are solved under the constraint that each vertebra is connected to its upper and lower vertebrae by two virtual joints. Lagrange multipliers method is used for solving these equations. The velocity and the angular velocity with respect to each vertebra are computed and then the position of each vertebra is calculated. The deformation process is carried out iteratively until the both following conditions are satisfied,

$$|\mathbf{p}_{\mathrm{neck}}(t) - \mathbf{x}_{\mathrm{neck}}(t')| < t_{\mathrm{neck}}, \quad (1)$$

$$|\mathbf{p}_{\mathrm{waist}}(t) - \mathbf{x}_{\mathrm{waist}}(t')| < t_{\mathrm{waist}}, \quad (2)$$

where t_{neck} and t_{wiast} are predetermined thresholds. The outputs are 3D position and the orientation of each vertebra.

(b) Forces acting on vertebrae

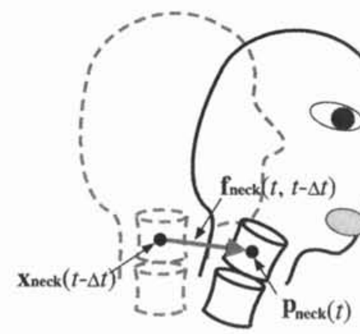

Figure 3: The force $\mathbf{f}_{\mathrm{neck}}(t,\ t-\Delta t)$ acting on the top vertebra of the accurate spine model. $\mathbf{x}_{\mathrm{neck}}(t-\Delta t)$ is the neck position obtained by the precise estimation process at time $t-\Delta t$. $\mathbf{p}_{\mathrm{neck}}(t)$ is the neck position obtained by the rough estimation process at time t.

Two types of forces acting on one vertebra are considered here during the deformation process. The first type is the gravitational force $\mathbf{f}_{\mathrm{g}}$ and the other is the intervertebral disk force $\mathbf{f}_{\mathrm{disk}}$. The force $\mathbf{f}_{\mathrm{disk}}$ occurs when the disk is compressed by the upper and lower vertebrae. In our model, each vertebra is connected to the upper and lower vertebrae by eight springs, respectively. The force $\mathbf{f}^{i}_{\mathrm{disk}}(t)$ acting on the i-th vertebra at time t is expressed by the following equation,

$$\mathbf{f}^{i}_{\mathrm{disk}}(t) = \sum_{j=1}^{8} k(\mathbf{d}^{j}_{i,i-1}(t) - \mathbf{d}^{j}_{i,i-1}(0)) + \sum_{j=1}^{8} k(\mathbf{d}^{j}_{i,i+1}(t) - \mathbf{d}^{j}_{i,i+1}(0)), \quad (3)$$

where $\mathbf{d}^{j}_{i,i+1}(t)$ and $\mathbf{d}^{j}_{i,i+1}(0)$ are the lengths of the j-th spring connecting the i-th vertebra and the $(i+1)$-th vertebra, respectively. The variable k is a spring constant.

Besides, in order to deform the accurate spine model so as to locate the top and the bottom vertebrae on the neck and the waist position, respectively, we introduce two virtual forces $\mathbf{f}_{\mathrm{neck}}$ and $\mathbf{f}_{\mathrm{waist}}$ acting on two vertebrae as shown in Fig. 3. Let $\mathbf{x}_{\mathrm{neck}}(t-\Delta t)$ and $\mathbf{x}_{\mathrm{waist}}(t-\Delta t)$ be the positions of two vertebrae at time $t-\Delta t$, and $\mathbf{p}_{\mathrm{neck}}(t)$ and $\mathbf{p}_{\mathrm{waist}}(t)$ the positions of the neck and the waist obtained by the rough estimation at time t. Then these forces are expressed as

$$\mathbf{f}_{\mathrm{neck}} = \mathrm{M}_{\mathrm{neck}} * (\mathbf{p}_{\mathrm{neck}}(t) - \mathbf{x}_{\mathrm{neck}}(t-\Delta t))/\Delta t^2, \quad (4)$$

$$\mathbf{f}_{\mathrm{waist}} = \mathrm{M}_{\mathrm{waist}} * (\mathbf{p}_{\mathrm{waist}}(t) - \mathbf{x}_{\mathrm{waist}}(t-\Delta t))/\Delta t^2, \quad (5)$$

where $\mathrm{M}_{\mathrm{neck}}$ and $\mathrm{M}_{\mathrm{waist}}$ are the masses of two vertebrae. The external force is derived from the forces $\mathbf{f}_{\mathrm{g}}$, $\mathbf{f}_{\mathrm{disk}}$, $\mathbf{f}_{\mathrm{neck}}$, and $\mathbf{f}_{\mathrm{waist}}$. Suppose that the model at time t_0 is deformed based on $\mathbf{p}_{\mathrm{neck}}(t)$ and $\mathbf{p}_{\mathrm{waist}}(t)$. In the n-th iteration of the deformation process, the force $\mathbf{F}^{i}_{\mathrm{ext}}$ acting on the i-th vertebra is expressed by the following equations,

$$\mathbf{F}^{i}_{\mathrm{ext}} = \begin{cases} \mathbf{f}_{\mathrm{g}} + \mathbf{f}^{i}_{\mathrm{disk}}(t') + \mathbf{f}_{\mathrm{neck}}(t, t-t'), & \text{(if top vertebra)} \\ \mathbf{f}_{\mathrm{g}} + \mathbf{f}^{i}_{\mathrm{disk}}(t') + \mathbf{f}_{\mathrm{waist}}(t, t-t'), & \text{(if bottom vertebra)} \\ \mathbf{f}_{\mathrm{g}} + \mathbf{f}^{i}_{\mathrm{disk}}(t'), & \text{(otherwise)} \end{cases} \quad (6)$$

where $t' = t_0 + n\Delta t$ and Δt is a time step in a dynamic simulation.

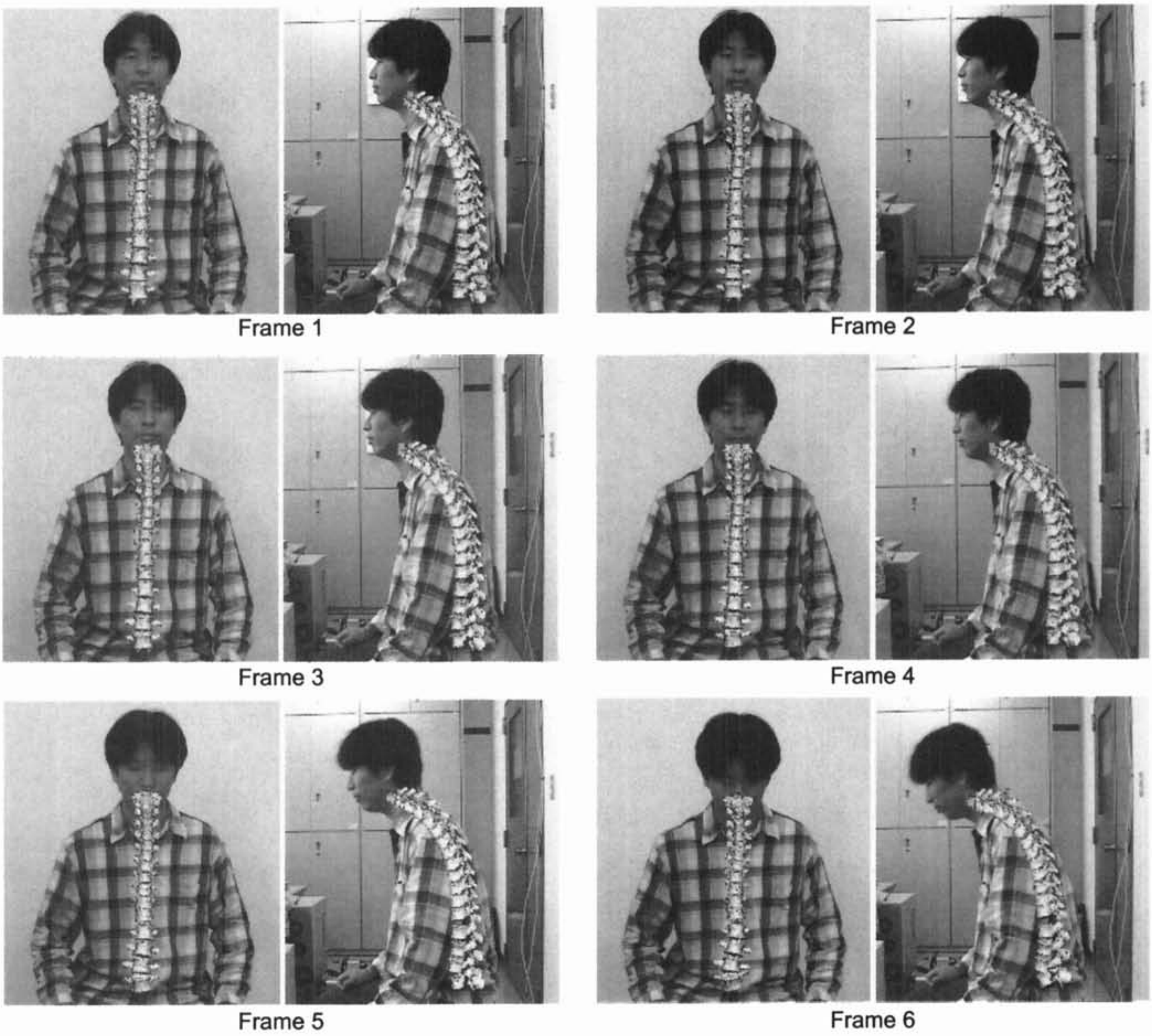

Figure 4: Estimation results of the flexion posture using accurate spine model.

3 Experimental results and discussion

We applied the proposed method to real human images for evaluating efficiency of spine posture estimation. The inputs are frontal and lateral views taken by conventional video cameras. The size of images is 320×240 pixels. The subject was sitting on a chair, and he was straightening his back in the first frame of input sequences. Then he bent his body slowly. Figures 4 and 5 show the estimated results of the spine postures. The estimated spine postures are overlaid on figures. We have tested our method on a PC (Pentium 4 1.5GHz). Total computation time for one posture was 159.3 seconds, and it took 7.10 seconds per one iteration of the procedure of the accurate spine model fitting.

The human spine posture is estimated reasonably as shown in Figs.4 and 5. It could be possible to estimate flexion and lateral bending posture. In the case of the twisting posture estimation, the body direction, which can be obtained by the positions of both shoulders, would be needed as additional constraint. Our approach is effective for fitting the model with articulated structure to an actual human image. It may be applicable to estimate an arm and leg posture or motion. In our method, only the interaction between the vertebra and the intervertebral disk are currently modeled. However, a variety of forces such as pressure from surrounding muscles and tissues are acting on an actual vertebra. In order to estimate the spine posture more precisely, we need to improve the method for computing the external force.

4 Conclusion

In this paper, we have proposed an accurate human spine model and a method for estimating the spine posture from frontal and lateral views of human body taken by two video cameras. Each vertebra is approximated as a rigid body composed of a set of particles and an intervertebral disk is modeled as a set of a joint and springs connecting the upper and the lower vertebra in our model. Rough positions of the neck and the waist are estimated using simple model from the images. Two virtual forces acting the top and the bottom vertebrae are computed based on the obtained posi-

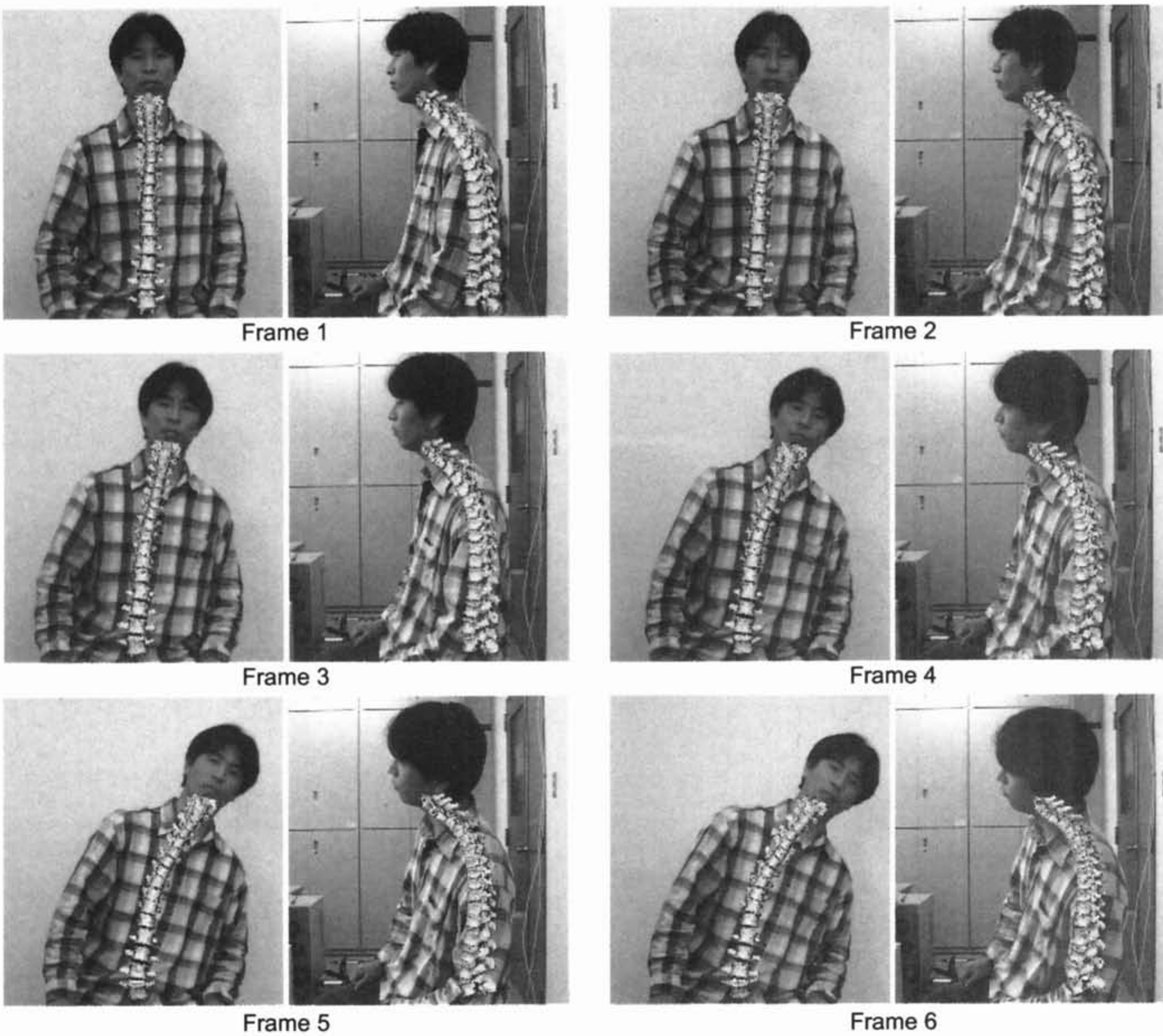

Figure 5: Estimation results of the lateral bending posture using accurate spine model.

tions. The precise posture of the spine is estimated by deforming the model under these forces. The model is deformed by solving the equations of motion. From the experimental results of application to actual human images, we showed that our proposed method estimated 3D postures of the human spine reasonably.

Future works include: (1) application to a large set of human images to validate efficiency of the method, (2) quantitative evaluation of the estimated spine posture using Open MRI images, (3) development of a method that generates a personal model from the generic spine model constructed from a set of X-ray CT images, and (4) development of a method for calculating the loads acting on the spine.

References

[1] Christoph Bregler and Jitendra Malik, "Tracking People with Twists and Exponential Maps," International Conference on Computer Vision and Pattern Recognition, pp.8-15, 1998

[2] Quentin Delamarre and Olivier Faugeras, "3D Articulated Models and Multiview Tracking with Physical Forces," International Journal of Computer Vision and Image Understanding, vol. 81, no. 3, pp.328-357, 2001

[3] Daisuke Furukawa, Kensaku Mori, and Yasuhito Suenaga, "Human Spine Posture Estimation from Video Images Based on Connected Vertebra Spheres Model," Proceedings of Computer Animation 2001, pp.174-181, 2001

[4] Ioannis Kakadiaris and Dimitris Metaxas, "Model-Based Estimation of 3D Human Motion," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp.1453-1459, 2000

[5] Ralf Plänkers and Pascal Fua, "Tracking and Modeling People in Video Sequences," International Journal of Computer Vision and Image Understanding, vol 81, no. 3, pp.285-302, 2001

[6] James M. Rehg and Takeo Kanade, "Model-Based Tracking of Self-Occluding Articulated Objects," Proceedings of Fifth International Conference on Computer Vision, pp.612-617, 1995