Finding Articulated Body
in Time-series Volume Data

Tomoyuki MUKASA, Shohei NOBUHARA, Atsuto MAKI,
Takashi MATSUYAMA

Graduate School of Infomatics, Kyoto University, Japan

Abstract. This paper presents a new scheme for acquiring 3D kinematic structure and motion from time-series volume data, in particular, focusing on human body. Our basic strategy is to first represent the shape structure of the target in each frame by using aMRG, augmented Multiresolution Reeb Graph [6], and then deform each of the shape structures so that all of them can be identified as a common kinematic structure throughout the input frames. Although the shape structures can be very different from frame to frame, we propose to derive a unique kinematic structure by way of clustering some nodes of graph, based on the fact that they are partly coherent. The only assumption we make is that human body can be approximated by an articulated body with certain number of end-points and branches. We demonstrate the efficacy of the proposed scheme through some experiments.

1   Introduction

The description of moving articulated objects, e.g. human figure, is important in many applications including technical analysis of sports and dance, or production of video contents. Motion-capture system is well-known and available for such purposes. However, the scope of description is limited to the case that the precise structure is given as articulated rigid body. Meanwhile, for the production of video contents, it is desirable to be able to describe non-rigid objects such as human skin or clothes, in motion. Aiming for the realization of comprehensive scheme for describing rigid and non-rigid objects in motion, we thus employ time-series volume data for the input of motion description acquisition scheme. In this paper, we propose a scheme for acquiring kinematic structure of articulated rigid body as an opening outset of above purpose.

   Conventional approaches for acquisition of kinematic structure from volume data are either top-down or bottom-up. Top-down approach uses a specific model, cylinder-model for example, to match with volume data. It is impossible to acquire a description of non-rigid motion by this approach. Moreover, in this approach, we need employ a somewhat elaborate model for each observation target. On the other hand, there are some bottom-up methods which acquire kinematic structure from observed motion without specific model [4, 3]. In these methods, unit of motion description is voxel or vertex on surface. Therefore, it is impossible to detect correspondence of their motion units before and after an

osculation of arthromeres(primitive segments of body). To avoid this problem, some partial sets should be prepared, to which some units related by closeness belong. However, in each method [4, 3], the partial sets are constructed based only on the motion units, and thus are not always that kind of sets when the osculation occurs.

In the proposed method, we prepare the global description of shape which represents the partial sets described above, and acquire a kinematic structure accordingly. That is, we first represent the shape structure of the target in each frame, and then deform each of the shape structures so that all of them can be identified as a common kinematic structure throughout the input frames.

2 Acquisition of Kinematic Structure

2.1 Overview

We acquire kinematic structure through an off-line process as shown below.

(1) Acquire time-series visual hull [1] from multi-viewpoints videos.
(2) Obtain surface mesh by applying marching cubes method [7] to each visual hull.
(3) Construct Reeb Graphs [5, 6] based on geodesic distance on the surface.
(4) As initial models, pick up some graphs that appear to be relatively close to the targeted kinematic structure.
(5) Deform each initial model so that it fits to reeb graphs in a certain number of neighboring frames in what we call "fitting interval". We define the size of the fitting intervals so that the sum of them can cover all the input frames.
(6) For each interval, cluster some nodes of deformed reeb graph based on their motion, and acquire kinematic structure.
(7) Integrate kinematic structures that are acquired for different fitting intervals.

(1),(2),(3) are processes for acquiring global shape structure. (4),(5) are processes for making correlations between shape structure and kinematic structures which enable us to acquire the kinematic structure in process (6). Last of all, by process (7), we can acquire kinematic structure that reflects the diversity of motion in the input visual hull.

2.2 Reeb Graph

We employ the reeb graph for the global shape description. To construct a reeb graph, we first segmentalize surface S of an object on the basis of a continuous function, $\mu(\mathbf{v})$ ($\mathbf{v}$ is an arbitrary vertex on the surface S), defined on the surface, and represent each segmented surface by a node, and finally link the nodes based on the connectivity between the segmented surfaces. We use geodesic function as $\mu(\mathbf{v})$. Now, the surface of the object is represented by a mesh model. The

geodesic distance $g(\mathbf{v}, \mathbf{p})$ is defined by shortest path between $\mathbf{v}$ and another vertex $\mathbf{p}$ on the mesh. Then, function $\mu(\mathbf{v})$ can be represented as

$$\mu(\mathbf{v}) = \sum_{\mathbf{p} \in S} g(\mathbf{v}, \mathbf{p}). \tag{1}$$

This indicates the sum of distance from vertex $\mathbf{v}$ to all vertices on S.

The reeb graph based on the geodesic distance is stable when there are no osculation between any arthromeres. However, even when observing an identical object, the graph structure may easily vary due to an osculation, and the coherence between the graph structure and the kinematic structure will break. Considering this problem, we represent the shape structure by a unique graph structure in a certain definite time range. To acquire such unique graph structure, we regard partial invariability of reeb graph.

We denote the vertices on a partial region surface $S_{partial}$ by $\mathbf{p}_i, (i = 0, 1, ..., N_{\mathbf{p}}-1)$, the vertices on a partial region $S_{complement}$, which is complement region of $S_{partial}$, by $\mathbf{p}_j^c, (j = 0, 1, ..., N_{\mathbf{p}^c}-1)$. We also denote the vertices connected to an arbitrary vertex $\mathbf{p}_i^c$ and also belonging to $S_{partial}$ by $\mathbf{p}_k^m, (k = 0, 1, ...N_{\mathbf{p}^m} - 1)$, and represent the partial region to which $\mathbf{p}_k^m$ belonging as $S_{marginal}(\subset S_{partial})$ (see Figure 1)　The function $\mu(\mathbf{p})$ defined at arbitrary $\mathbf{p}_i$ can be represented as
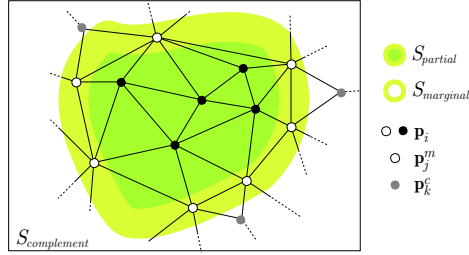
Fig. 1. Partial region on the surface and its mesh structure

$$\mu(\mathbf{p}_i) = \mu_{partial}(\mathbf{p}_i) + \mu_{complement}(\mathbf{p}_i), \tag{2}$$

where $\mu_{partial}(\mathbf{p}_i)$ is caused by $\mathbf{p}_i$, $\mu_{complement}(\mathbf{p}_i)$ is caused by $\mathbf{p}_j^c$, and

$$\mu_{partial}(\mathbf{p}_i) = \sum_{i'=0}^{N_{\mathbf{p}}-1} g(\mathbf{p}_i, \mathbf{p}_{i'}), \tag{3}$$

$$\mu_{complement}(\mathbf{p}_i) = \sum_{j=0}^{N_{\mathbf{p}^m}-1} \mu_{complement}(\mathbf{p}_j^m) = \sum_{j=0}^{N_{\mathbf{p}^m}-1} \sum_{k=0}^{N_{\mathbf{p}^c}-1} g(\mathbf{p}_j^m, \mathbf{p}_k^c). \tag{4}$$

The condition for that arbitrary vertices $\mathbf{p}_{i1}$ and $\mathbf{p}_{i2}$ on $S_{partial}$ correspond to a same node in the reeb graph is represented as

$$|\mu(\mathbf{p}_{i1}) - \mu(\mathbf{p}_{i2})| < \mu_{span}, \tag{5}$$

where μ_{span} is a constant. Regarding Equation (4), this equation can be transformed to

$$|\mu_{partial}(\mathbf{p}_{i1}) - \mu_{partial}(\mathbf{p}_{i2})| < \mu_{span}. \tag{6}$$

Equation (6) means that the structure of the partial reeb graph corresponding to $S_{partial}$ depends only on the vertices in $S_{partial}$. Therefore, the partial reeb graph is stable while all $\mathbf{p}_i$ are identical.

We can find almost invariable structures at the neighborhood of the end-points of the reeb graphs in a certain number of neighboring frames, because there are no osculation between corresponding arthromeres, and all $\mathbf{p}_i$ are thus almost identical. For these reeb graphs, it is reasonable to pick up a reeb graph as an initial model and deform it so that it fits to the other graphs. After this deformation, the shape structure is represented by a unique graph structure.

We thus can estimate the motion of the nodes in the graph, and then acquire the kinematic structure based on the motion.

2.3   Selection of Initial Model

The reeb graph which we select as an initial model should have just an appropriate number of branches and no loop. These properties are equivalent to the following conditions for the graph structure.

(i) "The graph has five end-points".
(ii) "The graph has two branch-points, one has four branches, and the other has three branches", or "The graph has three branch-points each of which has three branchs".

The later is due to the fact that the reeb graph is not necessarily symmetric.

2.4   Deformation of Initial Model

We deform initial models so as to deal with the changes in shape structure over time. In order to deform an initial model and fit it to other reeb graphs, we shift the nodes of initial model to the position where the nodes of other reeb graph are, while maintaining connective relations of the nodes in initial model.

We take the following constraints into account for the movement of nodes.

(a) Maintain the distance between nodes as much as possible
(b) Move a node in initial model to its nearest neighbor node in the target reeb graph

We enforce these constraints to work on each node. (a) can be realized by elastic force between nodes (we call the force as "internal force"), and (b) by the external force to move these nodes. Given these forces working at each nodes, we can compute how initial model deforms its shape by solving Newtonian equation with backward Euler integration method.

External Force. The nodes of the initial model should be moved to the neighborhood of the nearest nodes of the target reeb graph without any crossing in pathways of nodes' movement. We define the external force which urges such nodes' movement based on the point-set deformation algorithm [2].

We denote the reeb graph at time a and b by N and M respectively. First, we find nearest neighbor nodes in M for each node in N, and make a pair (m_a, m_b). We also make another pair (n_a, n_b) for each node in M in a similar way. Note that (m_a, m_b) and (n_a, n_b) are not always identical generally. Now, we denote the distances from an arbitrary point $p(x, y)$ to node n_a and m_a as $d_a(n)$ and $d_b(m)$, respectively. We also denote the drift force of node n_a to n_b, node m_a to m_b by $f_a(n)$ and $f_b(m)$, respectively. Then, we can define the external force $f_{external}(x, y)$ working on $p(x, y)$ as

$$f_{external}(x, y) \equiv \frac{\sum_n \alpha_n f_a(n)}{\sum_n \alpha_n} + \frac{\sum_m \beta_m f_b(m)}{\sum_m \beta_m}, \qquad (7)$$

$$\alpha_n = exp(\frac{-d_a(n)}{\sigma}), \qquad \beta_m = exp(\frac{-d_b(m)}{\sigma}). \qquad (8)$$

Internal Force. We introduce the internal force, $F_{internal}(v)$, as

$$F_{internal}(v) \equiv \sum_1^n k_j(\|\boldsymbol{q}_{v_j} - \boldsymbol{q}_v\|) \frac{\boldsymbol{q}_{v_j} - \boldsymbol{q}}{\|\boldsymbol{q}_{v_j} - \boldsymbol{q}\|} - k'_j \boldsymbol{q}_v, \qquad (9)$$

where k_j denotes a constant, v_j a node connected to v within 2 hops, $\boldsymbol{q}_{v_j}$ its 3D position, and n the number of v_j, respectively. The internal force works so as to maintain the smoothness of reeb graph's shape.

Updating Initial Model Based on Correlation of Reeb Graph. The temporal sampling rate of the input volume date can often be too low to follow the rapid motion of the target object. That is, it is not always true that temporally neighboring graphs are strongly-correlated.

On every deformation, we select a frame in which the reeb graph has the highest correlation to the target reeb graph among the frames in which the reeb graph has already deformed, and use its deformed reeb graph as a new initial model. We utilize a correlation computing method of reeb graphs that is proposed in [6].

2.5 Acquisition of Kinematic Structure Based on the Motions of Nodes

We first acquire kinematic structures in the corresponding fitting intervals, starting with initial models. We call the kinematic structures "piecewise kinematic structure". We then choose the most detailed partial kinematic structures for each branch corresponding to arm, leg, or head. Finally, we deform the partial

kinematic structures so as to fit them all over the time-series reeb graphs, and integrate them. This integrated kinematic structure is the result of our method, which reflects various kinematic motion pattern in the input volume data.
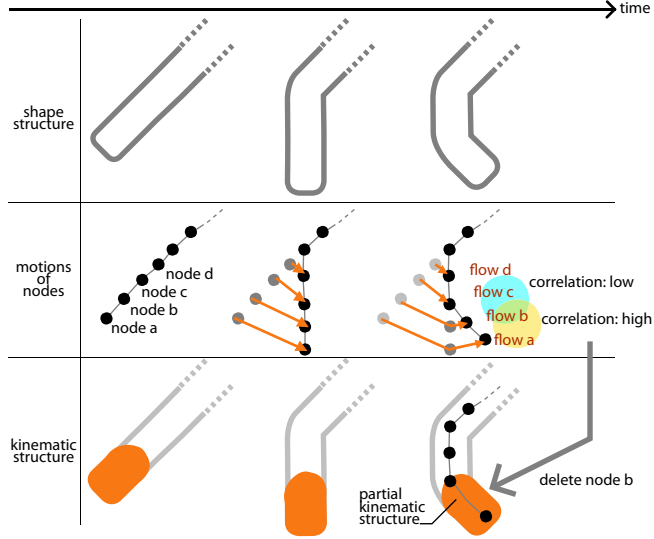
Fig. 2. Clustering of nodes based on their motion

Acquisition of Piecewise Kinematic Structure. We acquire kinematic structure by clustering nodes of reeb graph based on cross-correlation of motion between neighboring nodes(see Figure 2). We denote the number of nodes in the initial model by n, the velocity of node p_i, $(i = 1, 2, ..., n)$ at time t by $\boldsymbol{v}_t^i$, and the cross-correlation function of nodes p_i, p_j at time interval $[t_b, t_e]$ as $r(i, j)$ can then be represented as

$$r(i, j) = \frac{\sum_{k=t_b}^{t_e}(\boldsymbol{v}_k^i - \boldsymbol{v}_{mean}^i)(\boldsymbol{v}_k^j - \boldsymbol{v}_{mean}^j)}{\sqrt{\sum_{k=t_b}^{t_e}(\boldsymbol{v}_k^i - \boldsymbol{v}_{mean}^i)^2}\sqrt{\sum_{k=t_b}^{t_e}(\boldsymbol{v}_k^j - \boldsymbol{v}_{mean}^j)^2}}, \qquad (10)$$

where

$$\boldsymbol{v}_{mean}^i = \frac{\sum_{t=t_b}^{t_e}\boldsymbol{v}_t^i}{t_e - t_b}, \qquad \boldsymbol{v}_{mean}^j = \frac{\sum_{t=t_b}^{t_e}\boldsymbol{v}_t^j}{t_e - t_b}. \qquad (11)$$

Regarding a set of neighboring two nodes, the higher the cross-correlation is, the more probable that both of the nodes belong to the same rigid body.

Integration of Piecewise Kinematic Structure. We choose the most detailed partial kinematic structure for each branch, corresponding to arm, leg, or head. Then, we deform the partial kinematic structures so as to fit them all over the time-series reeb graphs, and integrate them(see Figure 3).
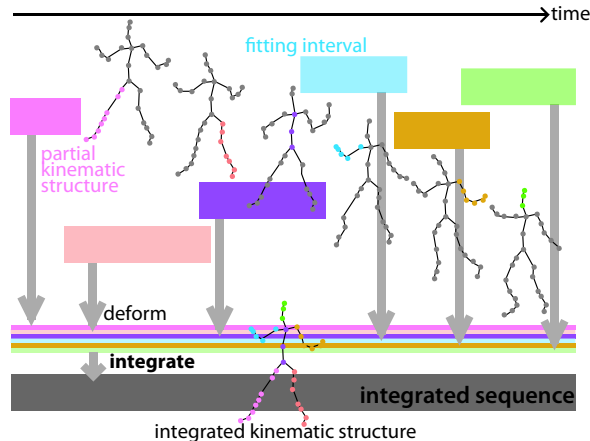
Fig. 3. Integration of piecewise kinematic structures

To coordinate partial kinematic structures in different fitting intervals, We use proximity of partial kinematic structures at overlapping fitting intervals.

3 Experiments

3.1 Real Data

We used nine cameras circumnavigating the target object (a dancing lady), and acquired time-series visual hull from multi-viewpoints videos. Applying proposed method, we acquired the kinematic structure. This acquisition process is exemplified at frame 147 as in Figure 4 and Figure 5. as an example.

In Figure 5, we can observe that human like kinematic structure is acquired from real data by the proposed method.

3.2 CG Data

We generate CG scene of a walking human by applying a motion capture data [1] to a CG model. We then shot the scene by 15 virtual cameras circumnavigating the CG object, and acquired its time-series voxel data using a technique based on

---

[1] The source of data anonymous. Detailed information will be provided in the final draft of this paper.
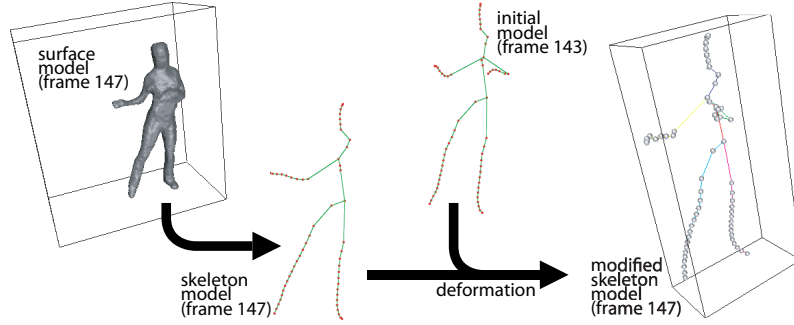
Fig. 4. Modification of reeb graph at frame 147
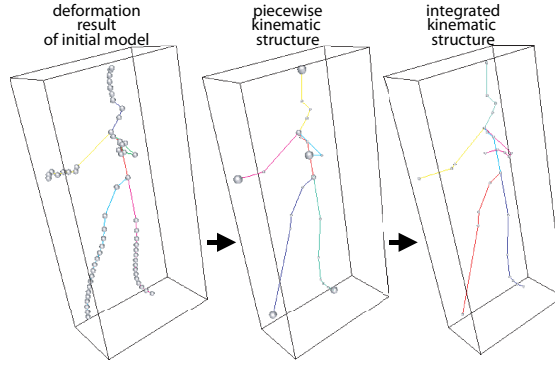
Fig. 5. Acquisition process of kinematic structure at frame 147

visual hull. Finally, we acquired kinematic structure from the time-series visual hull by the proposed method.

We evaluated the acquired kinematic structure by comparing it with the original motion capture data. In this evaluation, as a criterion, we computed the mean value of distance between joint positions in the motion capture data to their nearest neighbor nodes in the acquired kinematic structure. We denote the motion capture data by M, acquired kinematic structure by N, joint points in structures M and N by $m_i, (i = 0, 1, ..., i_{max})$ and $n_j, (j = 0, 1, ..., j_{max})$, respectively. We also denote the distance between the joint points in M and N by $dist(m_i, n_j)$. Then, we can define above described criterion by

$$E_{dist} \equiv \frac{\sum_i dist(m_i, n_{i'})}{i_{max} + 1} \tag{12}$$

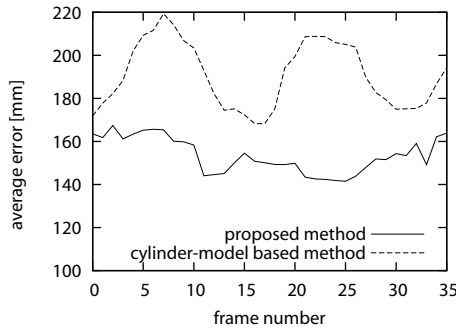while

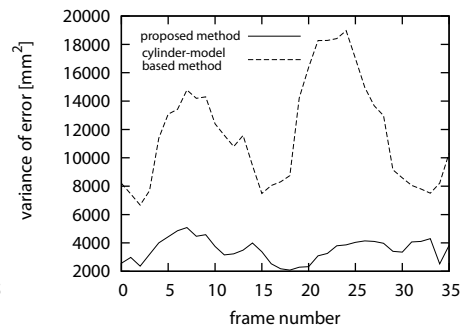$$dist(m_i, n_{i'}) = \min\{dist(m_i, n_j)\}. \tag{13}$$
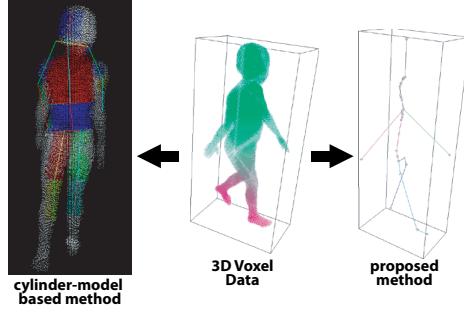
Fig. 6. Acquisition results of kinematic structure.

Fig. 7. Evaluation value E_{dist}. **Fig. 8.** Evaluation value V_{dist}.

Moreover, we also use another criterion V_{dist}, which is the variance of the distances between the joint points in M and N.

$$V_{dist} = \frac{\sum_i \{dist(m_i, n_{i'}) - E_{dist}\}^2}{i_{max} + 1}. \tag{14}$$

For the purpose of comparative evaluation, we also implemented a simple model-based human posture estimation method. This method employs cylindrical articulated human-figure model. First, the hip position of the cylinder-model is matched with the centroid of voxel data. Then, each joint angle of the model is adjusted so as to maximize the overlapping volume of the voxel data and the group of cylinders corresponding to arthromeres from the hip to each end-points.

In Figure 6, we show examples of kinematic structures acquired by the cylinder-model based method and proposed method, respectively.

Figure 7 and 8 show the evaluated value, E_{dist} and V_{dist}, respectively throughout the input frames. Solid line represents the values of the proposed method, and dot by the cylinder-model based method. In both of Figure 7 and 8, the values by the cylinder-model based method is affected by changes of object's posture

whereas those by the proposed method undergo in small range of shifting, showing that the kinematic structure can be acquired by the proposed method more stably.

However, the mean value of estimation error, E_{dist}, hover around 150mm whereas object's height defined by motion capture data is 1800mm. There is a room for improvement in terms of estimation accuracy.

4   Conclusion

In this paper, we proposed a scheme for acquiring kinematic structure from time-series voxel data using the reeb graph for the global shape description. For the case of human figure in motion, we have presented our early results that show the stable characteristic of the proposed scheme compared to a model-based method.

However, it is not guaranteed that the kinematic structure which reflects proper structure of the observed object can be always acquired. Accuracy of the acquired kinematic structure may not be sufficient for practical use. One of the solution to this problem will be to introduce a countermeasure in the presented bottom-up approach, for example, by restricting the edges connecting the nodes to lie inside the volume of the corresponding arthromeres.

Finally, since the proposed method works on time-series voxel data, we plan to expand the scope of the scheme to deal with non-rigid body in motion.

Acknowledgements

References

1. A. Laurentini, "How Far 3D Shapes Can Be Understood from 2D Silhouettes" *IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.17, No.2* pp. 188–195, 1995.
2. D.J.Burr, "A Dynamic Model for Image Registration", *Computer Graphics and Image Processing 15*, pp. 102–112, 1981.
3. S. Nobuhara, T. Matsuyama, "Heterogenieous Deformation Model for 3D Shape and Motion Recovery from Multi-Viewpoint Images", *2nd International Symposium on 3D Data Processing, Visualization, and Transmission*, pp. 566–573, 2004.
4. I. Masaaki, K. Yoshinari, M. Michihiko, "Estimation of the Location of Joint Points of Human Body from Successive Volume Data" *ICPR2000*, pp. 699–702, 2002.
5. M. Hilaga, Y. Shinagawa, T. Kohmura, T.L. Kunii, "Topology Matching for Fully Automatic Similarity Estimation of 3D Shapes", *Proc. of SIGGRAPH 2001*, pp. 203–212, 2001.
6. Tony Tung at el., "The augmented multiresolution Reeb graph approach for content-based retrieval of 3D shapes", *International Journal of Shape Modeling (IJSM), Vol.11, No.1*, pp. 91–120, June 2005.
7. Y. Kenmochi, K. Kotani, A. Imiya, "Marching Cubes Method with Connectivity", *Proc. of International Conference on Image Processing*, pp. 361–365, 1999.