Method for unsupervised non-rigid registration of motion capture markers on the surface of 3D mesh

Elena Martinova, Maria Lyashko
Intel Russia Research Centre,
Nizhny Novgorod, Russia
Elena.Martinova@intel.com

Abstract

Motion capture (MC) data are the most reliable source of information about human body motion. Such data are widely used in different kinds of applications, for example, movies creation, game development, physical simulation of human body etc. The task of markers registration on the surface of 3D mesh model, describing human body, is a necessary step of all mentioned applications. In publicly available MC databases, like Carnegie Mellon University (CMU) Graphics Lab Motion Capture Database, each marker has unique name and some semantic information about marker position on actor body; MC data contain time series of 3D marker positions and skeleton bones transformations. We present method for unsupervised search of correspondence between motion capture markers and 3D mesh vertices. Template mesh is a model of human in initial (T-shirt) pose, for registration any frame with arbitrary pose may be chosen. Correlated correspondence of markers to mesh vertices is reached by optimizing of joint probabilistic distribution over correspondence variables, represented as Markov Random Field (MRF). We define potentials, preserving geodesic distances between pairs of markers and correspondent vertices. Markers are added into consideration in portions by several steps. For each marker subset we create a Graphic Model as a minimum spanning tree, and execute Pearl inference procedure from Intel Open-Source Probabilistic Network Library (PNL). The obtaining results may be employed for pose deformation model learning in human animation and also in all mentioned applications.

Keywords: *Motion capture, markers registration, mesh registration, body animation, Markov Random Field, correlated correspondence.*

1. INTRODUCTION

Within last 10 years human animation (both face and body animation) have made impressive progress from fully handmade models, deformed by artists into a set of key poses, to automatic model creation and deformation by use of laser scans and MC data. Though nowadays synthetic personages are widely employed in movies, games, and entertainment, the role of artists in creation of really impressive and lifelike animation still remains important.

Further automation and improvement of body animation quality aren't possible without research of human motion. There are two main directions of body movement investigation and modeling: geometry-based approaches and physical simulation. The first large group has made several important steps and now remains the main direction of human animation development. Such approaches rely upon using laser scanner or motion capture data.

Physical and biophysical simulation of body movements still suffers from serious problems like inadequate deepness of simulation and difficulty of real phenomenon detailed description [1]. It is necessary to point that deep physical modeling is still a hard task for current computer architecture generation. At the same time processing of collision response and virtual actors interaction demands dynamic simulation of motion. Currently for getting appropriate result physical simulation uses motion capture data for mapping to dynamic-driven model [2]. Some applications use dynamic simulation for a short time interval to process interaction, and then returns to pseudo-physical movement control under guidance of MC data.

These data also used for manipulation like retargeting, edition and combining; result information also applied for virtual personages animation.

In our method we employ data base of Carnegie Mellon University, known as CMU Graphics Lab Motion Capture Database [3]. It contains more than 1600 trials in 6 categories of motion, from hundreds to thousands of frames, frame rate is 120 fps. For each frame there are a set of 3D markers coordinates and a set of skeleton bones rotation angles.

For use of such data in animation it is necessary to resolve a problem of finding correlated positions of markers on the surface of mesh body model. Usual mesh registration methods couldn't be directly applied to this task taking into account the next reasons:

- there is no information about surface to which markers belong, therefore it's impossible to calculate and employ surface normal, spin or exact distance along the surface;

- marker data are sparse.

From the other hand, MC data have qualities, which may be employed:

- there exists some semantic information about markers, and it may be used as prior data;

- for each frame data about skeleton bones are provided;

- since the main task is to obtain the best correlated correspondence between marker positions on the mesh surface, we are able to chose any frame with relatively simple pose or use several frames for control of results.

Presented unsupervised method for markers registration on the surface of 3D mesh establishes the best correspondent vertices for a set of markers. The method gives possibility to obtain the data for learning of body mesh model deformation on a set of variable poses. Since in different sequences markers are placed on varied positions, the set of correspondent vertices from different trials, taken together, presents motion of some mesh area. It is important to point, that in common case we can't use information about skeleton angles directly for calculation of marker position, since it contains both component from rigid body part transformation and non-rigid one, determined by muscle bulging and irregular stretching.

We shortly present related work in the section 2, give detailed description of our probabilistic model and method in section 3, and discuss results in section 4.

2. RELATED WORK

The group of earliest 3D human animation methods were based on interpolation between key poses. Quality of obtained animation depends mainly on model appearance in key deformations. If such poses were manually created by artists, final animation might look perfect: the more key poses – the better. Improved with some tools for artist work automation, this technique still is actively used.

The next step was made with creation of methods referenced as skinning, vertex blending, bones blending or enveloping. The idea was introduced by Nadia Magnenat Thalmann [4] for modeling of various joints in the hand, and later a number of similar methods appeared [5-8]. The common idea was: each surface vertex moves according to underlying skeleton movement. Vertex position is calculated as a weighted blending of limb transformations. For resolving well-known problem like elbow collapse or twisting some special processing was proposed. For instance, in work of Weber [5] for prevention of undesirable surface deformation additional bones were added to skeleton.

Some approaches tried to combine advantages of both mentioned methods [6, 7]. Pose-space deformation [6] overcomes artifacts for elbow or shoulders by use ready shapes, but needs artist work for desirable poses creation. New body deformations are calculated as a result of scattered interpolation. The obvious frustration is that some human poses aren't reachable as a combination of predefined ones: extrapolation may give bad results. In fact it is only one step to automation, but nothing new was added to the understanding of human body deformation process.

In the work [9], presented SCAPE method, the problem of animation by laser scans is treated as a shape completion task. The SCAPE method uses a set of laser 3D scans for learning of pose deformation and shape deformation models. Then, having a single scan of a person and time series of 50-60 markers, the method predict for each frame a full shape of the person, in a pose consistent with the observed marker positions. The contribution of this work looks important, since the new approach to description of body surface deformation was proposed. Authors introduced separate transformation terms for 3 main sources of deformation: (1) rigid part, defined by skeleton bones rotation, (2) nonrigid part due to muscle bulging and tissue deformation, (3) 'body shape' deformation, determined by personal body type and defined by weight, height, muscularity, etc. One of the steps in the SCAPE method is estimation of correspondence between template mesh and each instance mesh, obtaining from scans.

The next work, closely related to ours, describes 'correlated correspondence' method [10]. The method is applied for 3D scans registration of deformed objects and registers two meshes by optimizing a joint probabilistic model over point-to-point correspondences between them. Authors use Markov Network with potentials, penalizing difference in distances and transformations for correspondent point pairs, a set of markers is employed. For each marker in the network all others (not only neighboring) are considered as linked by edges; some additional constraints for geodesic distance decrease the number of neighbors in network. In fact correlated correspondence is an evolution of extended version Iterative Closest Point algorithm (ICP) [11], applied for nonrigid object models and used for simultaneously registering scans and recovering the surface configuration.

We present implementation of the method, which is related to the ideas of correlated correspondence, but is developed for registration of scattered motion capture markers on 3D mesh surface. Our contribution is a new scheme of Graphical Model creation as a minimum spanning tree. We also propose simplified form of pairwise Gaussian potentials in MRF, which makes the method fast and efficient. The potentials preserve only distance information, while the data about affine transformation of body surface are defined directly by skeleton angles from motion capture data files. We developed original algorithm of rough marker positions estimation and geodesic distance calculation with use of these positions. We employ Pearl inference procedure from Probabilistic Network Library (PNL) for search of final distribution.

3. METHOD OF MARKERS REGISTRATION ON 3D MESH

In the description of presented method we use, when possible, the terms and notations, introduces in related works [9, 10].

3.1 Problem statement

The input is:

- template mesh $X = (V^X, E^X)$, where $V^X = (x_1, \ldots, x_N)$ denotes vertices and E^X is a set of links between them;
- a set of markers $Z = (m_1, \ldots, m_k)$.

Correspondence of marker m_n to mesh is described as a correspondence variable c_n, where $c_n = i$, if m_n corresponds to mesh vertex x_i.

The task of registration is estimation the set of optimal correspondences $C = (c_1, \ldots, c_k)$ of markers Z to vertices of mesh X.

3.2 Probabilistic model

Z provides information about relative 3D markers position. Though some distances possibly were changed as a result of pose deformation, some of them still remains unchanged (for example, if two markers are on the same limb), and others got only small changes. We describe probability distribution of correspondence variables as $\psi(c_k)$, where

$$\psi(c_k = i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-d_i^2 / 2\sigma^2}, \tag{1}$$

d_i is Euclidian distance between x_i and point O_k, correspondent to the marker position in undistorted model.

We model joint distribution of correspondence variable C as a pairwise Markov Random Field (MRF). For each pair of adjacent markers m_k and m_l we define a probabilistic potential $\psi(c_k, c_l)$ that preserve distances between adjacent markers and penalize body limbs deformation:

$$\psi(c_k = i, c_l = j) = \frac{1}{\delta\sqrt{2\pi}} e^{\frac{-(D^Z_{i\to j} - D^X_{i\to j})^2}{2\delta^2}}, \quad (2)$$

where $D^Z_{i\to j}$ is scaled to body dimension geodesic distance between adjacent markers, and $D^X_{i\to j}$ is geodesic distance between vertices i and j in X, correspondent to the markers. We search the optimal value for C as a giving maximum likelihood for joint distribution in form

$$p(C) = \frac{1}{\Omega}\prod_k \psi(c_k)\prod_{k,l}\psi(c_k, c_l),$$

Ω is a normalizing constant.

The most probable values for correspondence set C is determined using Pearl inference, also known as belief propagation algorithm [12]; we exploited its implementation from Intel Open-Source Probabilistic Network Library (PNL) [13].

3.3 Search of correspondent vertices

Calculations are executed in several steps, start with 7-11 markers and put into consideration more on the each next step. The correspondences, calculated on the previous steps, are used as evidence on the next ones. In current realization the order of markers consideration is fixed and chosen with the intention to have first of all positions of markers, using for calculation of geodesic distances (see below).

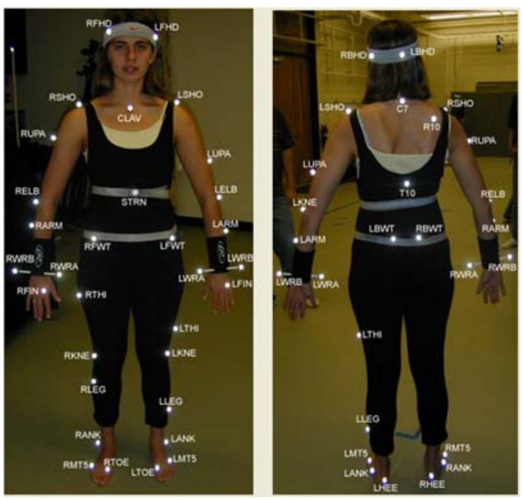

Figure 1: CMU motion capture markers.

We have some prior information about position of markers on the mesh (we denote it as O_k). To initialize the process, we define manually O_k for all markers on the first step: we point mesh vertices, which are approximately in suitable positions. These rough positions later are corrected as a result of inference. It is important, that we need to define centers O_k manually only once for one mesh model; these points may be used for any sequences later. After getting results for the step, if was calculated that $c_k = i$, than vertex x_i replaces estimated vertex O_k.

For the next steps O_k is estimated using position of already placed markers. We define simplified structure of the body as a tree with root in pelvis and 7 other body parts (back of the model is one of these parts), each part has the marker, connecting this body part to the parent one. For instance, for right arm such marker is RSHO, which connects it to torso; for left leg LFWT marker connects it to pelvis. On the first step at least one marker for each body part is defined by O_k. If now the next marker m_n is added to consideration, to estimate position of O_n the algorithm works in the next way:

1) define body part number p_n, corresponding to m_n;
2) find marker m_s, which belongs to the same body part p_n and already have value O_k or x_k;
3) define marker m_p, connecting p_n to parent body part;
4) $\mathbf{z}_s$ and $\mathbf{z}_n$ are vectors from m_p to m_s and from m_p to m_n, respectively, calculate angle $\alpha = \arccos\frac{\mathbf{z}_s \cdot \mathbf{z}_n}{\|\mathbf{z}_s\| \cdot \|\mathbf{z}_n\|}$ and axis of rotation $\mathbf{z}_r = \frac{\mathbf{z}_s \times \mathbf{z}_n}{\|\mathbf{z}_s\| \cdot \|\mathbf{z}_n\| \cdot \sin\alpha}$, $\mathbf{M}_\alpha$ is rotation matrix around the axis $\mathbf{z}_r$ by α, $\mathbf{R}$ is matrix of uniform scaling by ration $\frac{\|\mathbf{z}_n\|}{\|\mathbf{z}_s\|}$;
5) $\mathbf{x}_s$ and $\mathbf{x}_n$ are vectors from x_p to x_s and from x_p to O_n, $\mathbf{M}_p$ is matrix of rotation for body part p_n from CMU data;
6) calculate $\hat{O}_n$ as end point of vector $\mathbf{x}_n = \mathbf{R}\mathbf{M}_p^{-1}\mathbf{M}_\alpha\mathbf{M}_p\mathbf{x}_s$;
7) find O_n as vertex in X, nearest to $\hat{O}_n$.

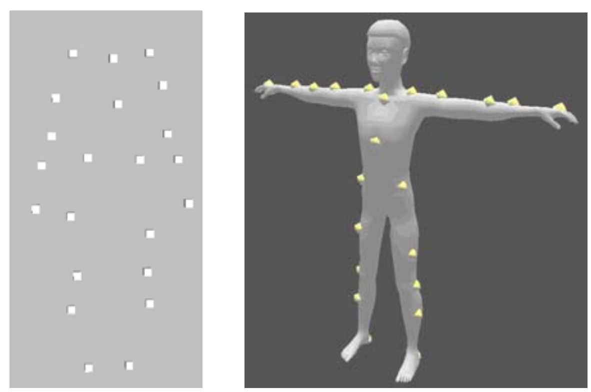

Figure 2: a – marker positions (front) in CMU data, b – result for correspondent vertices calculation.

This procedure gives approximate estimation, and then probabilistic part of the method calculates more precise marker position. Figure 3 presents an example of such situation.

The method employs minimum spanning tree (MST) as a Graphical Model for MRF, since in this case Pearl inference gives exact result. The weight of MST edge is geodesic distance between correspondent markers. For MST creation marker CLAV is used as a start point, and then for each step all valid for current step markers are added one by one in order of increase the minimal geodesic distance of a new marker to all others, already presented in MST (Boruvka's algorithm, also known as 'Sollin's' algorithm).

Since we haven't information about surface to which markers belong, it is impossible to calculate geodesic distance between markers along the surface. To resolve the problem we use body tree and markers, "connecting" each body part to parent. The algorithm of geodesic distance calculation finds the way in tree to connect markers and calculated geodesic distance along this way. Geodesic distance in X also is calculating in the same way, but instead of "connecting" markers correspondent vertices from mesh are used.

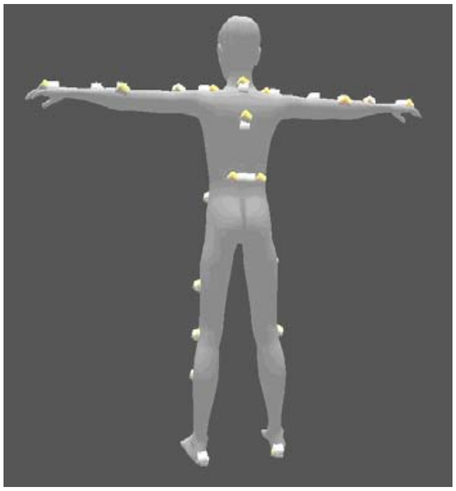

Figure 3: white boxes – estimated positions of O_k, color rotated boxes – calculated marker positions. It is seen, that inference gives plausible result, even having shifted estimation of O_k.

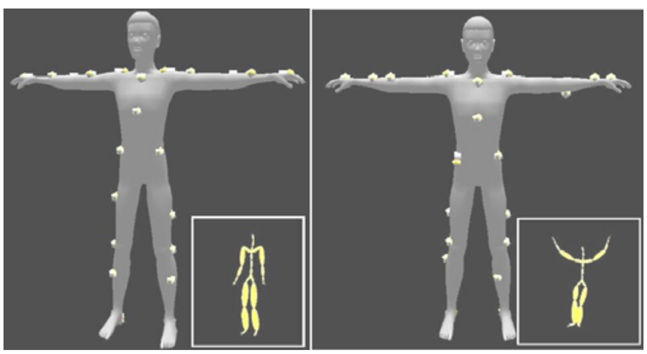

Figure 4: pose influence on result, in the right corner correspondent actor pose is shown.

4. APPLICATIONS AND DISCUSSION

The results of method application are presented on figures 2, 3. Source markers coordinates are obtained from converted c3d files. The order of markers in sequences is controlled via Deep Exploration (Right Hemisphere, version 3.5.10). We used mesh model of about 24K vertices in 'T-shirt' pose in our experiments. For getting stable output different parameters were investigated: influence of object poses in employed frame, value δ in functions (1), (2), and different schemes of markers involving into consideration. We find out that the best results were obtained for the frames with small flexion of limbs, see fig. 4 for an example. The best value of δ is about 0.02 of mesh model "height", and this value also may influences to proper final distribution of markers. It is necessary to point, that in some cases part of markers looks placed well, but another part looks shifted. It is desirable to distinguish well and incorrectly placed markers, in order to use only reliable data. This work is on the way now.

5. CONCLUSION

Proposed method of motion capture markers registration may be employed in combination with and as a part in any geometry-based approach to human animation, if the approach exploits motion capture data. Some limitations of current method implementation like predefined order of markers consideration aren't critical and will be a point of future improvement. We also intend to work on of proper evaluation of the obtained result quality and separation of marker position by their reliability.

6. REFERENCES

[1] Zordan, V. B., Majkowska, A., Chiu, B., Fast, M., *Dynamic Response for Motion Capture Animation, ACM Transactions on Graphics (ACM SIGGRAPH 2005), 24, 3, 697-701.*

[2] Zordan, V., Van Der Horst, N., *Mapping Optical Motion Capture Data to Skeletal Motion Using a. Physical Model, In proceedings of ACM SIGGRAPH/EUROGRAPHICS Symposium on Computer Animation, pp.245-250,2003.*

[3] CMU Graphics Lab Motion Capture Database. http://mocap.cs.cmu.edu/.

[4] N. Magnenat-Thalmann, R. Laperri`ere, and D. Thalmann, *Joint-dependent local deformations for Hand Animation and Object Grasping. Proc. Graphics Interface,1988, pp. 26-33*.

[5] Weber, J. *Run-Time Skin Deformation. Intel Architecture Labs, 2000.*

[6] John Lewis, Matt Cordner and Nickson Fong. *Pose Space Deformation: A Unified Approach to Shape Interpolation and Skeleton-Driven Deformation. SIGGRAPH 2000, pp. 165-172.*

[7] A.Mohr, M.Gleicher, *Building Efficient, Accurate Character Skins from Examples, SIGGRAPH'03 Conference Proceedings, pp 562-568, 2003*.

[8] Peter Sand, Leonard McMillan, and Jovan Popovic. *Continuous Capture of Skin Deformation. ACM Transactions on Graphics (TOG) 22, 3(2003), 578-586.*

[9] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, J. Rodgers, J. Davis *SCAPE: Shape Completion and Animation of People. ACM SIGGRAPH 2005, 24, 3, 408-416.*

[10] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, H. Pang, J. Davis, 2005. *The correlated correspondence algorithm for unsupervised registration of nonrigid surfaces. In Advances in Neural Information Processing Systems 17, 33–40.*

[11] D. Hahnel, S. Thrun, and W. Burgard. *An extension of the ICP algorithm for modeling nonrigid objects with mobile robots. In Proc. IJCAI, Acapulco, Mexico, 2003.*

[12] J.S.Yedidia, W.T.Freeman, Y.Weiss. *Generalized Belief Propagation. Advances in Neural Information Processing Systems, volume 13, pages 689-695, MA, 2001, MIT Press.*

[13] Intel Open-Source Probabilistic Network Library (PNL) http://www.intel.com/technology/computing/pnl/index.htm

About the authors

Elena Martinova and Maria Lyashko are research scientists in Intel Russia Research Centre. Contact phone (8312) 16244,

contact emails: Elena.Martinova@intel.com,
MariaLyashko@intel.com .