

Direct Shape Carving: Smooth 3D Points and Normals for Surface Reconstruction

Kazuki MATSUDA[†], Nonmember and Norimichi UKITA^{†a)}, Senior Member

SUMMARY This paper proposes a method for reconstructing a smooth and accurate 3D surface. Recent machine vision techniques can reconstruct accurate 3D points and normals of an object. The reconstructed point cloud is used for generating its 3D surface by surface reconstruction. The more accurate the point cloud, the more correct the surface becomes. For improving the surface, how to integrate the advantages of existing techniques for point reconstruction is proposed. Specifically, robust and dense reconstruction with Shape-from-Silhouettes (SfS) and accurate stereo reconstruction are integrated. Unlike gradual shape shrinking by space carving, our method obtains 3D points by SfS and stereo independently and accepts the correct points reconstructed. Experimental results show the improvement by our method.

key words: human shape reconstruction, shape from silhouettes, stereopsis, 3D surface reconstruction

1. Introduction

Human motion capture by multiple cameras is useful for various applications such as producing 3D video contents, computer-assisted physical rehabilitation, and so on. For human motion capture, several techniques are required (e.g. camera calibration, 3D reconstruction, texture mapping, pose estimation). Among all, we focus on 3D reconstruction.

3D reconstruction from multiviews is still a hot issue in machine vision. It is briefly classified into two classes: Shape-From-Silhouettes (SfS) and Stereopsis.

In SfS, multiview silhouettes of a target object are projected to a 3D space and then their intersection is regarded as the volume of the object, which is called a visual hull. While SfS is fast, robust, and able to get dense and smooth points, the visual hull might include false-positives in the concave regions of the object shape as shown in Fig. 1 and 2. The concave regions are often observed in a complex motion sequence of a moving person. These false-positives are unavoidable due to the nature of SfS.

In stereopsis, on the other hand, image windows that match between multiviews are found in order to compute the distance to the 3D points that projected onto those windows. In principle, every point where multiview matching is established can be reconstructed. Difficulty in matching is caused in textureless or uniquely-textured regions. This difficulty results in sparse and incorrect points.

Manuscript received October 6, 2011.

Manuscript revised January 4, 2012.

[†]The authors are with the Graduate School of Information Science, Nara Institute of Science and Technology, Ikoma-shi, 630-0192 Japan.

a) E-mail: ukita@is.naist.jp

DOI: 10.1587/transinf.E95.D.1811

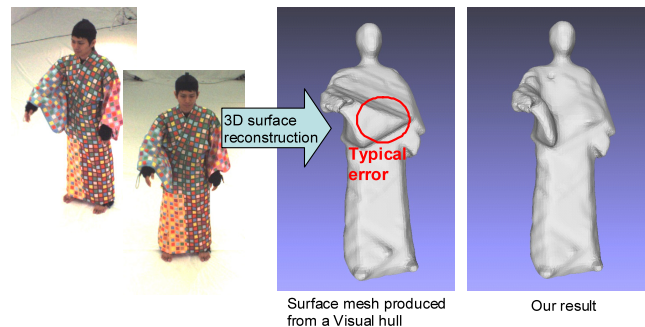


Fig. 1 Our result compared with the result of Shape from Silhouette. SfS produces huge error in concave regions.

This paper proposes how to integrate the advantages of the above two schemes; silhouette constraint of SfS and photo-consistency of stereopsis. In particular, our method focuses on how to carve huge ghost volumes, which are Fig. 1, efficiently. In the proposed approach, a visual hull from SfS is refined by using a point cloud reconstructed by multiview stereo. In other words, points obtained by SfS and multiview stereo are partly selected so that the visual hull is replaced by the stereo point cloud in sufficiently textured surfaces where photo-consistency works well.

While existing algorithms refine the visual hull in an iterative manner using 3D positions and photo-consistency, our new contribution is 1) direct shape carving (i.e. replacing the visual hull with the stereo point cloud) for avoiding local optima in iteration and carving huge ghost volumes efficiently and 2) using surface normals for robust shape carving.

2. Basic Methods for 3D Surface Reconstruction

This section introduces existing algorithms for 3D surface reconstruction and their known problems, which are resolved in our proposed method.

2.1 Shape-from-Silhouettes

Figure 2 illustrates a visual hull reconstructed by SfS [1]–[3]. Even if the correct silhouettes of a target object are extracted from multiview images, the visual hull includes false-positives as well as the real shape of the object. The false-positives are called *ghost volumes*. While the amount of them is reduced as the number of the cameras grows, it is essentially impossible to remove them in the concave re-

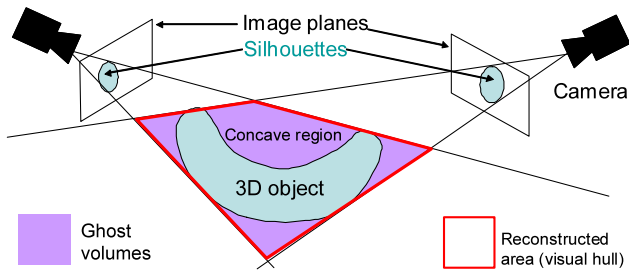


Fig. 2 Visual hull reconstructed by shape from silhouettes. The visual hull includes false-positive volumes, which are called ghost volumes, as well as real shape.

gions of the object.

Despite the ghost volumes, SfS is widely used for human shape reconstruction. This is because silhouette extraction is easier than stereo point correspondence in a studio, and dense and smooth surface points are obtained.

The smooth points allow us to estimate their surface normals as follows. The surface points, where their one or more neighboring points are outside the visual hull, are extracted. Then, at each surface point, an outward vector that is perpendicular to the local tangent plane defined by the surface points is regarded as the normal.

2.2 Accurate Multiview Stereo

Although early works in multiview stereo matches all points independently, recent approaches find the points on the surface that minimizes a global photo-consistency with smoothness constraints (e.g. optimized by level sets [19], and EM [20]). Novel techniques can reconstruct normals as well as 3D points; for example, patch-based multiview stereo [21].

While multiview stereo can reconstruct accurate 3D positions, it has a big disadvantage for human shape reconstruction; most cloth has less texture, which makes point correspondence difficult. This difficulty causes incorrect reconstruction and lack of the surface.

2.3 Existing Approaches for improving SfS and Multiview Stereo

The most popular approach for refining a visual hull is space carving [4]–[6]. The visual hull, which is an initial shape, is carved gradually until photo-consistency is satisfied between multiviews. Other constraints such as smoothness can be also optimized (e.g. using continuous local optimization by gradient descent [7], [8], discrete global optimization by graph cuts [9]–[12], and continuous global optimization [13]). Testing reliability of photo-consistency [14] can improve accuracy of carving. Despite the additional constraints and reliability test, all of the above approaches are based on space carving [4], which might fall into local optima.

Photo-consistency becomes more powerful by employing artificial textures on a target surface. For example, col-

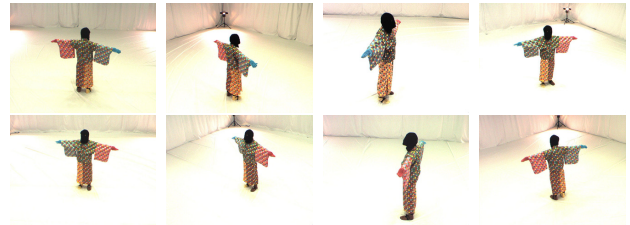


Fig. 3 Images captured from multi viewpoints used for reconstructing the surfaces shown in Fig. 4.

ored patches on the cloth surface [22], [23] can be used for human shape reconstruction.

As well as silhouette constraint and photo-consistency, other constraints such as smoothness can be also useful; for example, mesh deformation using multiple constraints [16].

For reconstructing a structurally-known shape such as a human body, the shape template model (e.g. articulated mesh [17], [18]) is useful for improving robustness to noise.

The approaches mentioned in this section have the following limitations:

- Gradual carving from a visual hull tends to fall into local optima.
- Specially-colored cloth is unavailable in natural scenarios.
- The articulated model with rigid motion of limbs is inapplicable to non-rigid loose-fitting clothing such as a skirt, floaty dress, or Japanese kimono.

2.4 Poisson Surface Reconstruction

3D surface reconstruction from 3D points is a popular problem in Computer Graphics for producing movies and so on. In contrast to simple meshing such as Marching Cubes [24], an implicit function framework is good at filling of surface holes and smooth meshing of existing points. In this paper, Poisson surface reconstruction [25] is employed as the implicit function framework. Compared with other implicit surface fitting algorithms that heuristically segment a surface into regions for local fitting and integrate them using blending functions, Poisson reconstruction [25] is a global solution that can obtain smooth surfaces with robust approximation of noisy data. In Poisson reconstruction [25], the Poisson formulation is applied to the oriented points. Therefore, as well as the 3D position, the orientation, which is equal to the surface normal, is required.

Figure 4 shows the results of Poisson surface reconstruction from a visual hull (left) and a point cloud reconstructed by multiview stereo [21] (right) using eight-viewpoint images shown in Fig. 3. For emphasizing the limitations of each method, specially-colored clothing was used for easy reconstruction. While the visual hull produces the feasible surface with no deleted body-regions, several body-regions (e.g. feet) are missing in the one reconstructed from the stereo point cloud. This is because no valid point correspondence was obtained in these regions. From the stereo

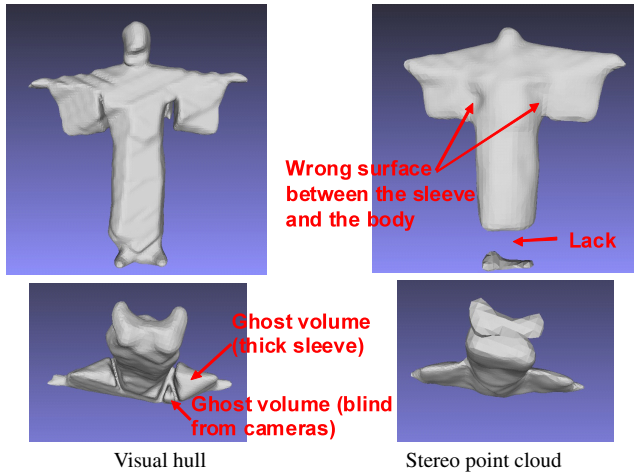


Fig. 4 Poisson surface reconstruction from a visual hull and a point cloud using multiview images shown Fig. 3.

point cloud, on the other hand, thin sleeves are reconstructed correctly although ghost volumes make the sleeves thicker than the real shape.

Our goal is to acquire 3D oriented points that allow Poisson surface reconstruction to generate the accurate surface of the whole human body.

3. Direct Shape Carving Using Shape from Silhouettes and Multiview Stereo

Existing approaches in space carving gradually remove surface points in a visual hull until photo-consistency is satisfied. Photo-consistency could be satisfied before the reconstructed surface reaches the real surface.

Instead of gradual update, our method directly accepts surface points reconstructed by multiview stereo [21] and then carve the visual hull points outside the accepted points. To this end, technical issues are as follows:

- Accept only reliable oriented points reconstructed by multiview stereo, which are not inconsistent with the visual hull.
- Find and carve visual hull points that interrupt the view lines from cameras to the accepted stereo points.

Again note that the essential difference between space carving and our method is how to acquire an optimized shape. In space carving, gradual carving from the initial shape (i.e. visual hull) might halt at local optima. In our method, on the other hand, surface points obtained by multiview stereo, where brute-force matching among all feature points is achieved for acquiring a global optimum, are employed. While it is practically impossible that the brute-force optimization is used for searching for the global optimum in a huge search space (i.e. all points/voxels between the real surface and the one of the visual hull)[†], the brute-force matching among the limited number of the feature points used in multiview stereo can be achieved.

Contrary to our method (i.e. carving the surface of a

visual hull), surface segments of the visual hull can be attached to a stereo point cloud for filling holes in the point cloud [15]. While it is difficult to find the holes that should be filled from the sparse point cloud and to determine the shape of the surface segment so that it fits with the hole^{††}, our method carves the surface of the visual hull, which is attached to the stereo point cloud, by a ray intersection test from multiviews (will be described in Sect. 3.2).

3.1 Pruning Multiview Stereo Points

First of all, SfS and multiview stereo are applied to images independently. We rely on points reconstructed by multiview stereo, except the ones whose normals are inconsistent with those of the visual hull. This is because:

- accuracy of the normals gets much lower where the points reconstructed by multiview stereo are sparse (while dense points are obtained in every surface by SfS, stereo reconstructs sparse points in textureless regions) and
- inconsistent normals, which might be observed where stereo points are reconstructed sparsely, make it difficult to apply the Poisson formula for surface reconstruction.

Based on the criteria above, a stereo point is pruned if 1) it is close to the surface of the visual hull and 2) its normal is significantly different from the normal of the nearest visual hull surface. Specifically, if 1) the distance to the nearest visual hull surface is shorter than $w_l l$, where l and w_l denote the side length of a voxel and a constant, respectively, and 2) the angle between the normals of the stereo point of interest and the nearest visual surface is larger than a threshold, θ_a , the stereo point of interest is pruned. In our experiments, $w_l = 8$ and $\theta_a = 30$ degrees^{†††}.

3.2 Carving Visual Hull

Like ray tracing, each of the remaining stereo points is projected onto the image planes in which the point is extracted. If the ray hits one or more points reconstructed by SfS, these points are carved as ghost volumes. Actually, the bounding box around each point is prepared for this intersection test [26]; if the ray crosses the box, its respective visual hull

[†]Although the search space can be reduced by constraints such as smoothness and the Epipolar constraints, it is difficult to optimize simultaneously whether or not each of massive points/voxels should be carved.

^{††}If a hole is far from its nearest surface of a visual hull because the hole is caused in a huge concave region of a target object (i.e. in the case of a huge ghost volume), it is difficult 1) to determine whether or not the hole should be filled by the surface of the visual hull and 2) to extract the good shape of the surface segment that fits with the hole.

^{†††}These values should be adjusted depending on reliability of multiview stereo. For example, if more cameras are used for multiview stereo, w_l should be smaller for accepting more points reconstructed by stereo.

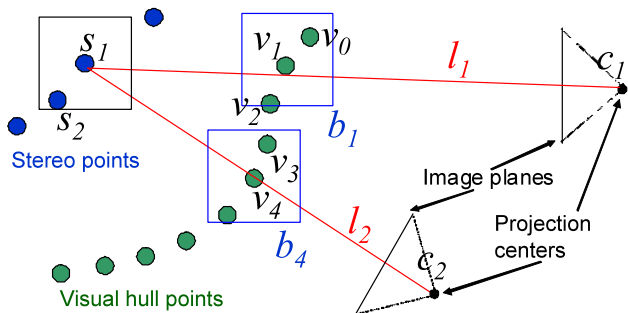


Fig. 5 Shape carving by testing ray intersection with a bounding box.

point is carved.

For this intersection test, the size of the bounding box is critical. If the size is smaller/larger, visual hull points that must be carved/retained are retained/carved incorrectly. The size is determined in accordance with the distance between the stereo point of interest and its nearest stereo point. Figure 5 illustrates this intersection test. Given a stereo point, s_1 and its nearest point, s_2 , the size of the bounding box is equal to the cube whose center is s_1 and that passes s_2 . This bounding box is located in every visual hull point (e.g. box b_1 is located in point v_1 in Fig. 5). Assume that s_1 is reconstructed by cameras 1 and 2, whose projection centers are c_1 and c_2 , respectively. Since the ray from s_1 to c_1 (denoted by l_1) passes through b_1 , v_1 is carved. v_0 and v_2 are also carved because l_1 passes through their bounding boxes. Similarly, v_3 and v_4 are also carved because their bounding boxes are on the way of l_2 that is from s_1 to c_2 .

4. Smooth Surface Reconstruction

Finally, with the remaining stereo points and visual hull points, the 3D surface is acquired by Poisson surface reconstruction [25]. This surface reconstruction can produce smooth surface patches while filling large gaps between the stereo points and the surface points of a visual hull (Fig. 6).

As mentioned above, our method consists of pruning stereo points and carving visual hull points. Carving is essentially required as in the way of space carving. On the other hand, one might use all stereo points with no pruning because careful optimization in multiview stereo obtains correct 3D points.

To see the effectiveness of pruning, surface reconstruction from the points with/without the pruning was conducted. From multiview images in Fig. 7, the surface shown in Fig. 8 was reconstructed. This is a typical example that proves the demand of the pruning. The left-hand image is the surface reconstructed from all stereo points, the remaining visual hull points, and their normals, which are shown in the right-hand image. Although the points of the left hand were reconstructed, it was over-carved. This was caused due to sparse scattered normals in that region; see the right-hand image. In the right-hand image of Fig. 8, it can be seen that the normals are scattered in the head. These scattered normals make the Poisson formulation difficult to solve.

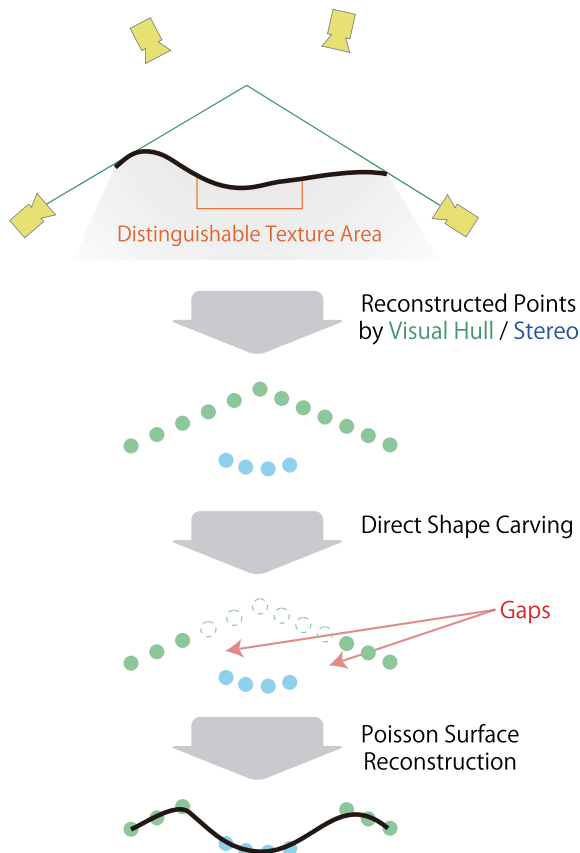


Fig. 6 Large gaps between stereo points and surface points of a visual hull. The gaps are smoothly concatenated with 3D surface patches by Poisson surface reconstruction [25].



Fig. 7 Images captured from multi viewpoints used for reconstructing the surfaces shown in Fig. 8 and Fig. 9.

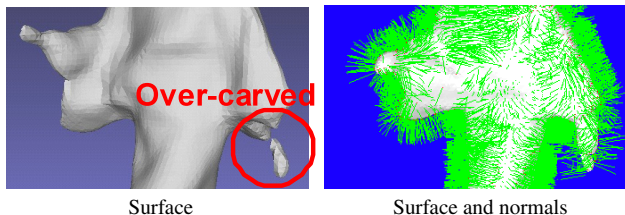


Fig. 8 Poisson surface reconstruction from all stereo points and the remaining visual hull points. Green lines indicate normals.

In contrast to Fig. 8, the result of our method (Fig. 9) looks better because of smooth normals even in the hand region. This is why pruning scattered normals is required

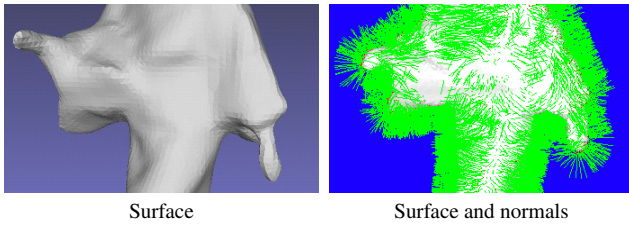


Fig. 9 Poisson surface reconstruction using our method. The difference from Fig. 8 is that our method prunes inconsistent stereo points before applying Poisson surface reconstruction to the reconstructed points.



Fig. 11 Dance sequence: observed images.

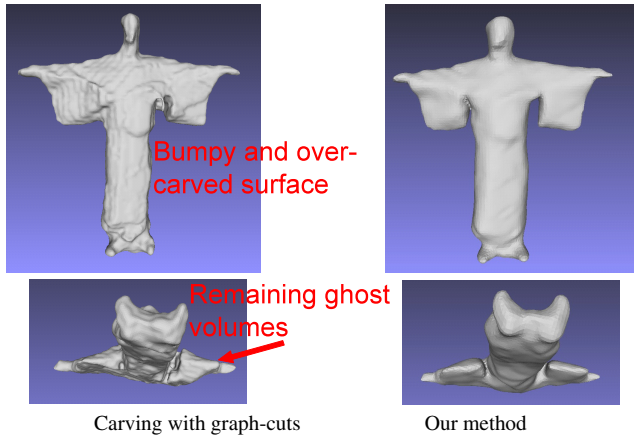


Fig. 10 Poisson surface reconstruction from the results of space carving and our method using multiview images shown Fig. 3.

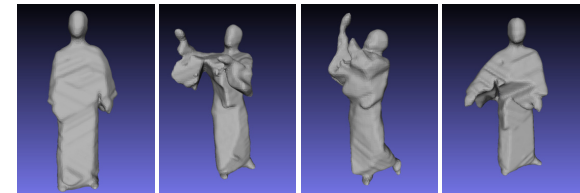


Fig. 12 Dance sequence: shape from silhouettes.

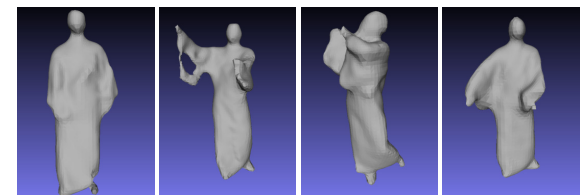


Fig. 13 Dance sequence: multiview stereo [21].

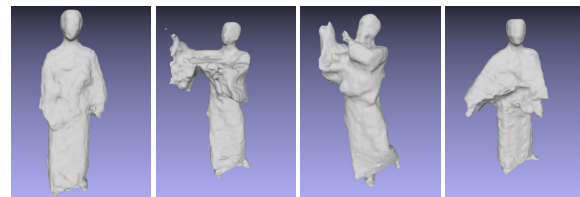


Fig. 14 Dance sequence: space carving [12].

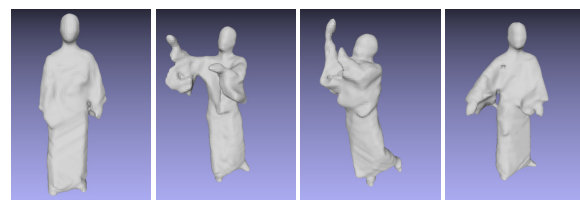


Fig. 15 Dance sequence: our method.

before surface reconstruction.

To prove the performance of our method in carving ghost volumes, the surfaces reconstructed by space carving with graph-cuts [12] and our method were compared. The results are shown in Fig. 10: space carving with graph-cuts (left) and our method (right). The ghost volumes remained in the result of space carving, while our method could successfully carve them. It can be also seen that the surface of space carving was very bumpy. This might be caused due to local optima in carving.

5. Experiments with Image Sequences

Our method was applied to multiview image sequences for validating the effectiveness of our method.

Figures 11 – 30 show observed images and the results of surface reconstruction from them. For comparison, as well as the results of our method, those obtained by SfS, multiview stereo [21], space carving [12] are also shown. For validating the results of various shapes and motions, four kinds of image sequences were used: dance motion with kimono clothes (Figs. 11 – 15), exercise motion (Figs. 16 – 20), pitching motion (Figs. 21 – 25), and batting motion (Figs. 26 – 30).

All the results were obtained from eight cameras. The cameras were located around a subject. If a camera was located so that it looked down on the subject from the above, it could reduce ghost volumes efficiently by SfS, especially

those surrounded by the arms. But no camera was located right above the subject for examining the performance of carving them under stringent conditions.

All silhouettes used by SfS and parameters in Poisson surface reconstruction were same in all the methods. The parameters used in carving with graph-cuts [12] were determined so that they were suitable for reconstructing human motion in [12].

In each sequence, the following observations can be seen:

- In the sequence of the dance motion, a space between



Fig. 16 Exercise sequence: observed images.



Fig. 21 Throwing sequence: observed images.

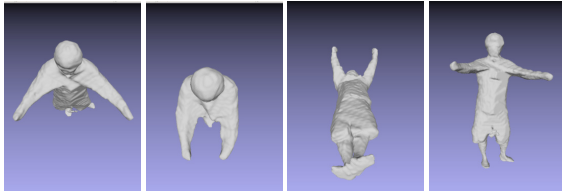


Fig. 17 Exercise sequence: shape from silhouettes.

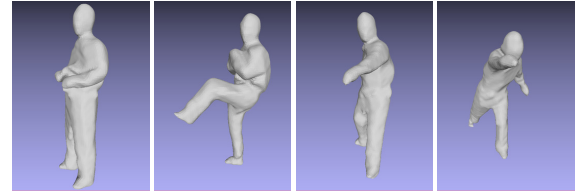


Fig. 22 Throwing sequence: shape from silhouettes.

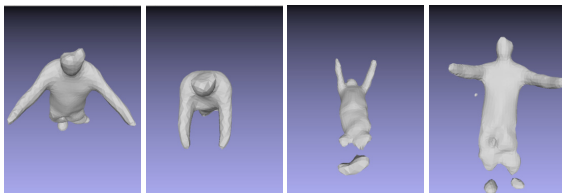


Fig. 18 Exercise sequence: multiview stereo [21].

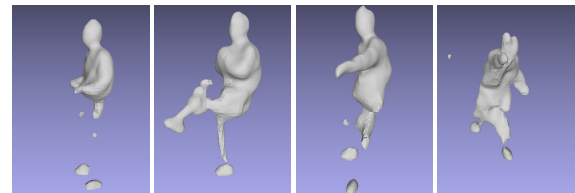


Fig. 23 Throwing sequence: multiview stereo [21].

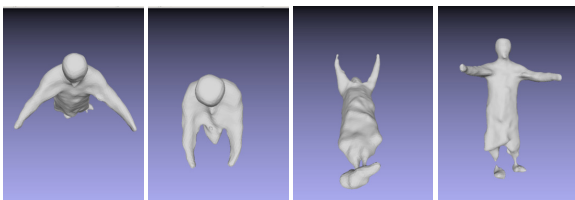


Fig. 19 Exercise sequence: space carving [12].

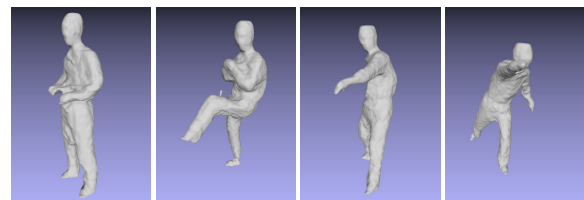


Fig. 24 Throwing sequence: space carving [12].

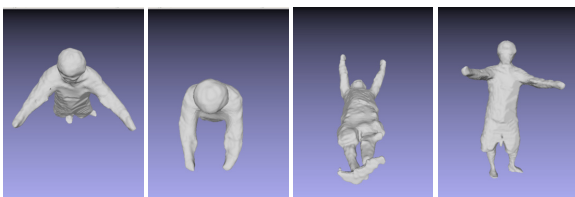


Fig. 20 Exercise sequence: our method.

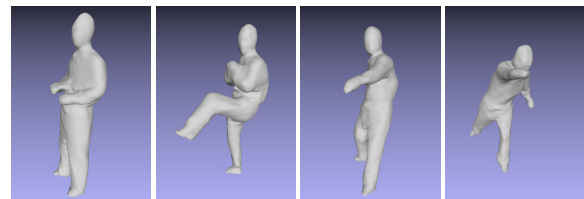


Fig. 25 Throwing sequence: our method.

the large sleeves of the kimono was occluded from the viewpoints of the cameras. That space was included in the visual hull as ghost volumes.

- While tight-fitting clothing was captured in the exercise motion, the raising arms produced large ghost volumes between the arms. Furthermore, textureless pants were difficult to reconstruct by multiview stereo.
- The human shape in the pitching motion is relatively easy to reconstruct by SfS. But multiview stereo could get only a small number of surface points because the clothing had less textures. In the sequence of the bat-

ting motion also, the clothing had less textures. In addition, a space surrounded by the arms was occluded from the cameras.

- The batting sequence is a difficult one in terms of less textures and a small loop with the arms. A bat was not reconstructed because its silhouette was not extracted due to its color-similarity to a background.

It can be seen that our method could reconstruct smooth surfaces with less ghost volumes in contrast to other methods as follows:

- Shape-from-silhouettes and space carving produced

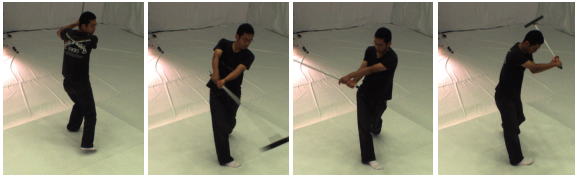


Fig. 26 Batting sequence: observed images.

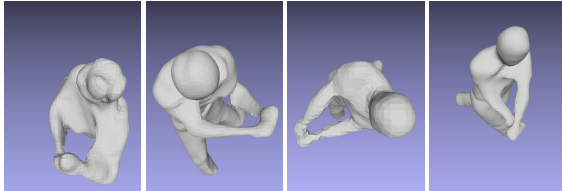


Fig. 27 Batting sequence: shape from silhouettes.

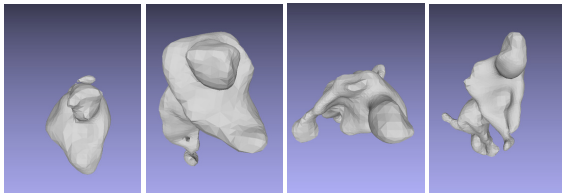


Fig. 28 Batting sequence: multiview stereo [21].

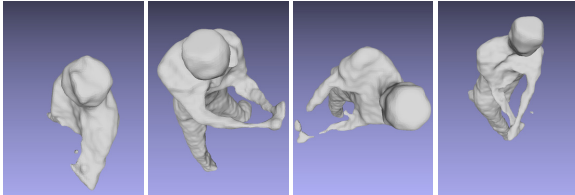


Fig. 29 Batting sequence: space carving [12].

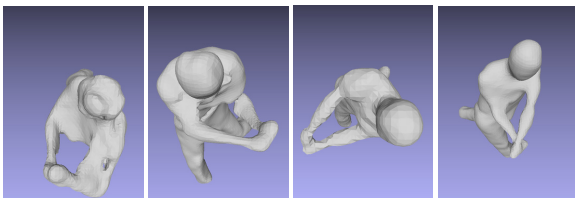


Fig. 30 Batting sequence: our method.

large ghost volumes due to the large sleeves (i.e. ghost volumes between the torso and the right sleeve in the rightmost image in Fig. 12). Multiview stereo produced a hole in the sleeve. This is because no points were reconstructed by multiview stereo in that region and it is impossible to make a surface mesh in the region with no points.

- In the exercise sequence also, shape-from-silhouettes and space carving produced large ghost volumes between the arms. Note that the ghost volumes were produced, though 1) the arms, which produced the ghost

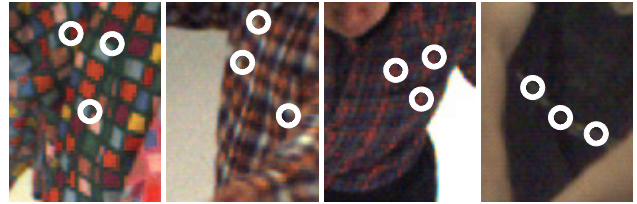


Fig. 31 Pixels used for manual matching and 3D point reconstruction for quantitative evaluation. In the dance, exercise, and pitching sequences, the corners of checked patterns were used. In the batting sequence, thin gray lines on the shirts were used.

volumes, were much smaller than the sleeves in the dance sequence and 2) theoretically speaking, by space carving, the ghost volumes could be carved because sufficient textures were on the upper body, but actually couldn't. This fact says that space carving is easily trapped by local minima.

- Incomplete surfaces were obtained by multiview stereo, in particular in textureless regions. Due to the incomplete surfaces, the results of our method are almost same with those of shape-from-silhouettes and space carving.
- Space carving stopped short of the real surface (e.g. surface between the torso and the right sleeve in the rightmost image in Fig. 14) and over-carved the volume of a slim region (e.g. arms in Fig. 29).

The computational cost of the proposed method is compared with that of space carving [12]. With eight 1024×768 pixel images and 1 cm^3 voxel size, space carving [12] and the proposed method needed around 30 secs and 20 secs on average by using Xeon 2.4 GHz. The breakdown of the computational time of our method is 10 secs in multiview stereo [21], less than 1 sec in SfS, and 10 secs in point pruning and carving.

For quantitative evaluation of reconstruction accuracy, pixelwise matching was manually achieved between multiview images and used for 3D point reconstruction of the ground truth. In one frame of each sequence, three pairs of matching pixels were selected from regions where huge ghost volumes were reconstructed. Circles in Fig. 31 shows the selected pixels. From the surface reconstructed by each method, 3D points corresponding to the manually selected pixels were extracted by ray intersection with the reconstructed surface. Then, the distance between the ground truth and the reconstructed surface was evaluated. RMS errors over all frames in SfS, multiview stereo, space carving, and the proposed method were 7.6, 5.3, 6.1, and 2.6 cm, respectively. While multiview stereo could get good results in textured regions such as around the selected pixels, the RMS error was worse than our method because the reconstructed surfaces were corrupted due to sparse reconstruction in textureless regions. Space carving could get good results in shallow concave regions, but in deep concave regions in the dance and batting sequences (i.e. regions around the selected pixels), the RMS error became larger.

6. Concluding Remarks

This paper proposed 3D oriented point reconstruction from multiviews. The method employs two kinds of point sets reconstructed by shape from silhouettes and optimized multiview stereo. For sorting out these two kinds of the point sets, a two-phased point removal is achieved:

1. Pruning based on irregularity of the normals of the points.
2. Carving along the rays from the optimized stereo points to the cameras that observe the points of interest.

Future work includes the following aspects:

- Smoothness optimization for avoiding isolated points in the result of shape carving.
- Adaptive parameter selection for shape carving depending on the spatial resolution of voxels and the density of points reconstructed by multiview stereo.
- Reoptimization around boundaries between the points reconstructed by shape from silhouettes and multiview stereo for a more smooth surface.

The source codes of PMVS [21], Poisson surface reconstruction [25], and space carving with graph-cuts [12] were given by Y. Furukawa, M. Kazhdan, and S. Nobuhara, respectively.

References

- [1] G. Cheung, T. Kanade, J. Bouguet, and M. Holler, "A real time system for robust 3D voxel reconstruction of human motions," CVPR, 2000.
- [2] X. Wu, O. Takizawa, and T. Matsuyama, "Parallel pipeline volume intersection for real-time 3D shape reconstruction on a PC cluster," 4th IEEE International Conference on Computer Vision Systems (ICVS), 2006.
- [3] G. Cheung, S. Baker, and T. Kanade, "Shape-from-silhouette of articulated objects and its use for human body kinematics estimation and motion capture," CVPR, 2003.
- [4] K.N. Kutulakos and S.M. Seitz, "A theory of shape by space carving," IJCV, vol.38, no.3, pp.199–218, 2000.
- [5] M. Goesele, B. Curless, and S.M. Seitz, "Multi-view stereo revisited," CVPR, 2006.
- [6] S.M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A comparison and evaluation of multi-view stereo reconstruction algorithms," CVPR, 2006.
- [7] Y. Furukawa and J. Ponce, Carved visual hulls for imagebased modeling, ECCV, vol.1, 2006.
- [8] C.H. Esteban and F. Schmitt, "Silhouette and stereo fusion for 3D object modeling," CVIU, vol.96, no.3, 2004.
- [9] S. Paris, F. Sillion, and L. Quan, A surface reconstruction method using global graph cut optimization, ACCV, Jan. 2004.
- [10] S. Tran and L. Davis, 3D surface reconstruction using graph cuts with surface constraints, ECCV, 2006.
- [11] G. Vogiatzis, P.H. Torr, and R. Cipolla, "Multi-view stereo via volumetric graph-cuts," CVPR, 2005.
- [12] T. Tung, S. Nobuhara, and T. Matsuyama, "Simultaneous super-resolution and 3D video using graph-cuts," CVPR, 2008.
- [13] K. Kolev, M. Klodt, T. Brox, and D. Cremers, "Continuous global optimization in multiview 3D reconstruction," IJCV, vol.84, no.1, pp.80–96, 2009.
- [14] J. Arnabat, S. Casanovas, and G.G. Medioni, "3D modeling from turntable sequences using dense stereo carving and multi-view consistency," ICPR, 2004.
- [15] I. Feldmann, N. Atzpadin, O. Schreer, J.-C. Pujol-Acolado, J.-L. Landabaso, and O.D. Escoda, "Multi-view depth estimation based on visual-hull enhanced Hybrid Recursive Matching for 3D video conference systems," ICIP, 2009.
- [16] S. Nobuhara and T. Matsuyama, "Deformable mesh model for complex multi-object 3D motion estimation from multi-viewpoint video," 3DPVT, 2006.
- [17] E. de Aguiar, C. Stoll, C. Theobalt, N. Ahmed, H.-P. Seidel, and S. Thrun, "Performance capture from sparse multi-view video," ACM SIGGRAPH, 2008.
- [18] D. Vlastic, I. Baran, W. Matusik, and J. Popovic, "Articulated mesh animation from multi-view silhouettes," ACM SIGGRAPH, 2008.
- [19] J.-P. Pons, R. Keriven, and O.D. Faugeras, "Modelling dynamic scenes by registering multi-view image sequences," CVPR, 2005.
- [20] C. Strecha, R. Fransens, and L.V. Gool, "Combined depth and outlier estimation in multi-view stereo," CVPR, 2006.
- [21] Y. Furukawa and J. Ponce, "Accurate, dense, and robust multiview stereopsis," PAMI, vol.32, no.8, pp.1362–1376, 2010.
- [22] V. Scholz, et al., "Garment motion capture using color-coded patterns," Computer Graphics Forum, vol.24, no.3, pp.439–448, 2005.
- [23] R. White, K. Crane, and D. Forsyth, "Capturing and Animating Occluded Cloth," ACM TOG (SIGGRAPH), vol.26, no.3, 2007.
- [24] W.E. Lorensen and H.E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," Computer Graphics, vol.21, no.4, 1987.
- [25] M. Kazhdan, M. Bolitho, and H. Hoppe, "Poisson surface reconstruction," SGP, 2005.
- [26] T. Moller and B. Trumbore, "Fast, Minimum Storage Ray-Triangle Intersection," J. Graphic Tools, vol.2, no.1, pp.21–28, 1997.



Kazuki Matsuda is a master course student of Nara Institute of Science and Technology. His research theme is human shape reconstruction.



Norimichi Ukita received the B.E. and M.E. degrees in information engineering from Okayama University, Japan, in 1996 and 1998, respectively, and the Ph.D degree in Informatics from Kyoto University, Japan, in 2001. After working as an assistant professor at Nara Institute of Science and Technology (NAIST), he became an associate professor in 2007. He was a research scientist of PRESTO, Japan Science and Technology Agency (JST) from 2002 to 2006, and a visiting research scientist at the Robotics Institute, Carnegie Mellon University from 2007 to 2009. He is now working also as a visiting researcher at the Intelligent Robotics and Communication Laboratories, ATR. His main research interests are object detection/tracking and pose/shape estimation of human body and clothing. He has received the best paper award from the IEICE in 1999.