

# Texture Overlay onto Deformable Surface for Virtual Clothing

Jun Ehara, Hideo Saito

Department of Information and Computer Science, Keio University, Yokohama, Japan  
*{eha, saito}@ozawa.ics.keio.ac.jp*

## Abstract

In this article, we describe a method for overlaying arbitrary texture image onto surface of T-shirt worn by a user. In this method, the texture image is previously divided into a number of patches. On the T-shirt, markers are printed at the positions corresponding to the vertices of the patches. The markers on the surface of the T-shirt are tracked in the motion image taken by a camera. The texture image is warped according to the tracked positions of the markers, which is overlaid onto the captured image. This article presents experimental results with the pilot system of virtual clothing implemented based on the proposed method.

**Key words:** Virtual Clothing, Deformable Surface, Texture, Marker Tracking

## 1. Introduction

Recently, Augmented Reality(AR) is used for many applications in various fields. AR, which is the technique of overlaying a virtual object onto real scene, is used mainly for rendering a virtual image on an image of real scene. TV advertisement[1], car navigation system[2], municipal engineering[3], entertainment[4] are the examples of AR applications.

In such AR applications, virtual object is frequently overlaid on particular positions in real scene. For example, virtual objects are imposed on a planar surface by recognizing a specific pattern attached to a surface of a book[5]. Another approach renders virtual texture on curved surface of a book using HMD[6]. This approach recognizes 2D matrix pattern and small markers on a book in an image captured with a camera on HMD, and warps a saved texture image. By projecting the warped image to an object by a projector, an user can feel as if the pattern of surface of the object is a projected texture [7][8]. As getting correspondences between a camera image and a projector image, no explicit calibration of a camera or a projector is needed in these works.

When we want to overlay a virtual object on a real curved surface seamlessly, precise shape of the surface needs to be measured. We can categorize methods for obtaining surface shape from motion images.

The first approach is to use scene features. This approach

does not require certain maintenance of visual markers. Natural feature points on an object are tracked after priori construction of 3D model structure of the object by locating visual markers[9]. Features on non-rigid deformable surface of an object are tracked using precise low-level optical flow without prior knowledge of the 3D shape of the object[10]. Another method detects deformable surfaces using matches between an image of the object with no deformation and a detected image[11]. According to distinctive features called as SIFT, which are invariant to image scale and rotation, they perform reliable matching between different views of an object or scene[12].

The second approach is to use visual markers that are applied on a surface of an object. Such a marker-based tracking increases robustness and reduces computational time. A grid pattern on a flexible moving surface such as cloth can properly tracked even with the presence of self-occlusions, self-shadowing, and wrinkles of the cloth[13]. A small number of self-identifiable markers are placed on a non-rigid paper or cloth, and a texture is augmented onto a mesh created according to the positions of the markers[14].

The idea of textured augmentation on a surface of a non-rigid object can be applied to a virtual clothing system. This system enables a user to look at him/her wearing clothes even if he/she does not have real clothes. Only data of a texture of the clothes is needed. We can mention that a surface of clothes is more difficult to overlay the texture on it than a surface of a rigid object, because the surface of the clothes is complicated with cloth stretching, wrinkles, and self-occlusion.

There are some related works on virtual clothing. First, an approach to overlay texture on a surface of a user by using computer graphics techniques is proposed[15,16,17]. In this approach, 3D shape data must be preliminarily obtained with 3D measurement instruments or installed as 3D model of a person, and data of clothes is transformed in 3D space using computer graphics techniques. However, a large amount of 3D data is hard to deal with and prevents real-time computation.

As other works that realize virtual clothing, an approach using motion estimation of a person from an image sequence is also studied. Motion estimation is the way to

transform certain model to a corresponding image of the person and recognize the human motion by the model shape. As works using motion estimation, 3D human body model and time-space gradient of motion image are used[18], or dynamic simulation of cloth with spring model is used[19]. However, this approach does not recover deformation of the surface of the object because it uses global shape fitting of the human model.

In order to realize the real-time virtual clothing with rendering texture which recovers deformation of the surface, some related works use specific patterns on the surface of clothes such as markers on T-shirt. A number of markers with various colors are attached on the center of T-shirt and the positions of the markers are acquired with their colors[20]. This approach moves in real time because the texture is transformed to align to an area consisted with markers according not to the 3D shape but to the 2D shape. Since the approach needs no 3D model in advance, it reduces the burden to construct the model of each object or person. However, sleeves are significantly deformed and waist part of T-shirt is considerably self-occluded compared to the center of T-shirt. Therefore, rendering texture on overall area of T-shirt is considered to be difficult.

In this article, we propose a method to overlay arbitrary texture image onto surface of T-shirt worn by a user. In order to realize realistic virtual clothing, the texture is rendered on overall area of T-shirt. In this method, the texture image is preliminary divided into number of planar patches. Markers are attached at the positions which correspond to the vertices of the planar patches. The markers are tracked in motion image and the texture image is warped according to the positions of the markers and rendered on the captured image. As long as a marker is not occluded by wrinkles or self-occlusion, texture is warped to proper position. Our approach deals with the shadow on T-shirt in 2D image rather than 3D deformation of T-shirt as stretching and wrinkles of clothes.

The rest of this article is arranged as following. Section 2 briefly describes outline of our approach. Section 3 shows how markers on T-shirt and texture image are corresponded in the first image, and section 4 describes how to track the markers and to render the texture on the T-shirt. We show experimental results, and discuss the effectiveness and the future work of our approach in section 5, and reach the conclusion in section 6.

## 2. Outline of Proposed Approach

Fig.1 shows the overview of our system of virtual clothing. A user wears the T-shirt with markers attached on its surface, when the user moves in front of the blue curtain. The image on which the texture of the virtual clothes is rendered according to the positions of the markers on the motion image taken by a camera is shown on a display. The size of an output image is  $320 \times 240$  pixels.

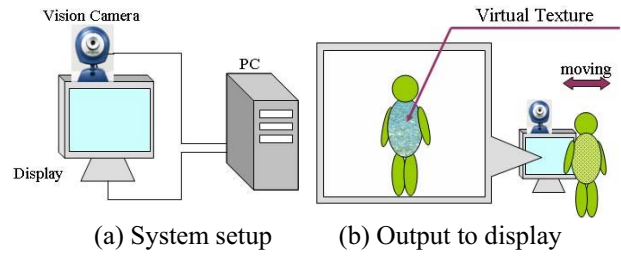


Figure 1: Outline of proposed system

Fig.2 shows the T-shirt and the example of clothes texture used in our system. Markers consist of one cross pattern and 28 small dots. The cross pattern is symmetrically attached in the center of T-shirt, and the dots are attached all over T-shirt.

If there are too few markers on T-shirt, rendered texture does not recover deformation of the surface of T-shirt, such as clothes stretching or wrinkles. On the contrary, if there are too many markers on T-shirt, markers are likely to be recognized as wrong markers near the real markers. Since shoulder parts of clothes deform considerably with human motion, the density of the small markers is higher at the shoulders than at other parts of T-shirt. The size of the texture image is  $1600 \times 1200$  pixels.

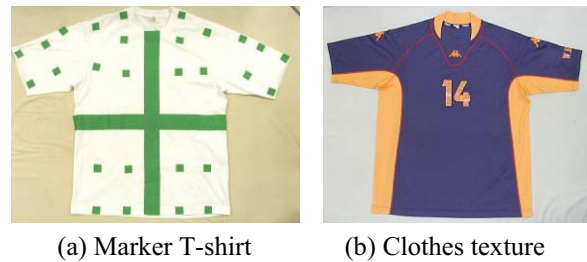


Figure 2: T-shirt and texture

Fig.3 shows the outline of our method. Our method can be divided into two steps; the first frame, and the frames after the first frame. In the first step, all markers of T-shirt in an input image are initially corresponded with all markers of T-shirt in a texture image. In the second step, the markers are tracked according to the positions of markers in the previous frame. This process of marker tracking enables accurate marker corresponding even if the markers' positions are largely deformed in motion image, and is one feature of our method.

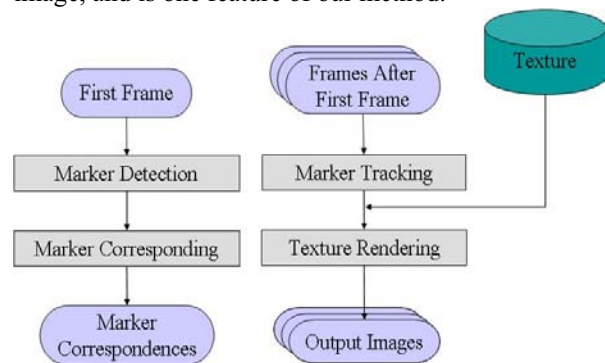


Figure 3: Outline of proposed method

### 3. Corresponding Initial Marker

In the first frame, all markers on T-shirt need to be seen, because all markers in an input image and in a texture image are corresponded. This method in the first frame is divided into two processes. The first process is detection of markers and the second process is corresponding markers in an input image to markers in a texture image.

#### 3.1. Marker Detection

Since a shape of clothes is complicated, tones of markers of the same type are different in an input image. All colors are expressed as color vectors in RGB space, as shown in Fig.4(a). R, G, and B represent three primary colors. We determine whether the color of each pixel in the input image is similar to the color of the marker by comparing color vectors. The result image of marker detection is shown in Fig.4(c).

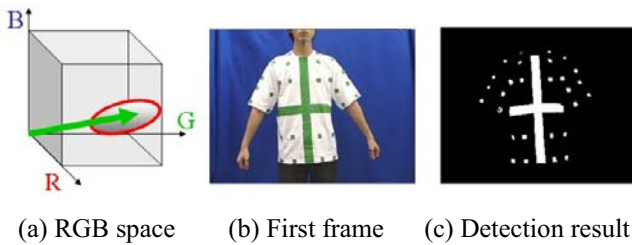


Figure 4: Marker detection

#### 3.2. Marker Corresponding

Detected markers need to be corresponded to markers in a texture image for marker tracking in frames after the first frame.

Positions of markers in the input image are different from those of the texture image, because the texture image is the image that clothes are spread out on a plane. Therefore, we use the standard appearance model as shown in Fig.5. The standard appearance model is the image that clothes are worn by a user and includes positions of the markers. The markers in the standard appearance model are preliminary corresponded to those in a texture image by hand.

Next, we automatically correspond the markers in the input model to those in the standard appearance model by using the cross pattern. Markers in the model are divided into four areas by cross pattern, upper-left, upper-right, lower-left and lower-right. Each marker is included in any of the area. According to the shape of cross pattern detected in the input image, homographies are computed for each area in the model. Positions of all markers in the model are warped into the input image with each homography. The detected marker is found at the neighborhood of each warped marker. Then all markers in the input image are corresponded to markers in the standard appearance model.

By these two correspondences, all markers in the input image and in the texture image are corresponded.

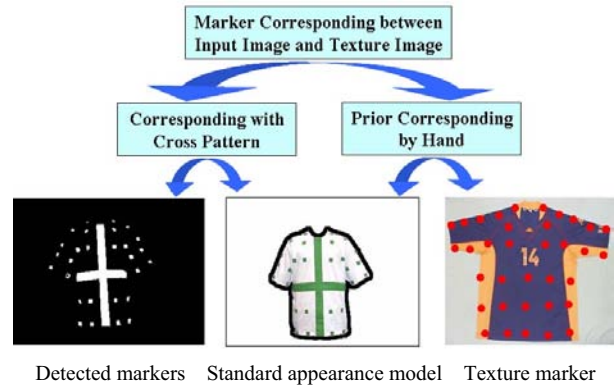


Figure 5: Marker Corresponding

### 4. Marker Tracking and Rendering Texture

In frames after the first frame, all markers do not need to be seen in an input image. Detected markers are tracked using positions of markers in previous frame and texture is rendered according to positions of markers.

#### 4.1. Marker Tracking

Marker tracking means corresponding between a marker in previous frame and that in recent frame. At the neighborhood of position of detected marker in recent frame, tracked markers previous frame are searched as shown in Fig.6. The nearest tracked marker from the position of detected marker in recent frame is tracked. If there is no tracked marker in neighborhood of position of detected marker, the detected marker is not tracked at this stage.

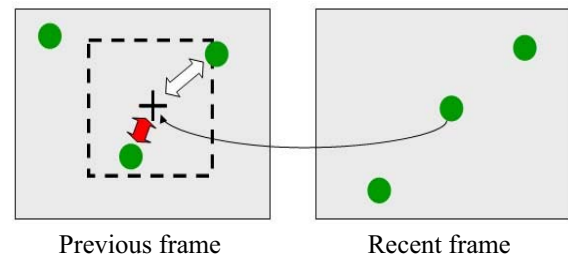


Figure 6: Marker tracking

#### 4.2. Overlapping of Marker Tracking

After the stage of marker tracking, several detected markers can be corresponded to the same tracked marker in previous frame. As markers in previous frame correspond one-to-one to those in recent frame, this marker overlapping must be avoided. When marker overlapping happens, the nearest detected marker in recent frame from position of the overlapped tracked marker is corresponded as seen in Fig.7. The detected markers that are not yet corresponded try to correspond the other tracked markers. If the detected marker is not corresponded, it is not tracked at this stage.

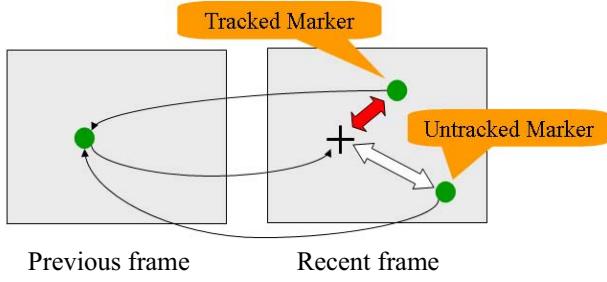


Figure 7: Marker overlapping

#### 4.3. Marker Re-corresponding

Some detected markers may not be corresponded after the processes of section 4.1 and section 4.2. These non-corresponded markers are considered that they have disappeared in any previous frame and appear in recent frame again because of cloth stretching, wrinkles, or self-occlusion. They are re-corresponded to markers in a texture image that are not yet corresponded, because all detected markers in recent frame should be corresponded to any of markers in a texture image.

As shown in Fig.8, the nearest tracked marker from not corresponded marker is defined as the neighbor marker. Then position of non-corresponded marker in standard appearance model is computed according to the cross pattern and the position of the neighbor marker in standard appearance model. At the neighborhood of the computed position of non-corresponded marker, non-corresponded marker in standard appearance model is found and re-corresponded.

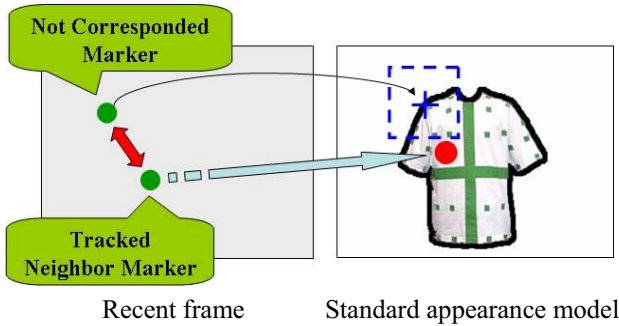
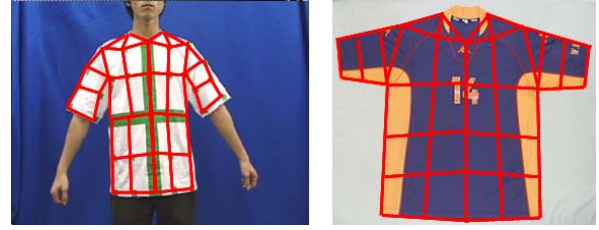


Figure 8: Marker re-corresponding

#### 4.4. Texture Rendering

In our method, texture pattern is rendered onto T-shirt by warping each planar patch of texture pattern which corresponds that of T-shirt in an input image.

A surface of T-shirt is approximated by set of small planar patches that are made by markers, shown in Fig.9(a). Positions of markers in a texture image are preliminary set by hand. Also the texture pattern consists of planar patches that are constructed with markers. Each planar patch is a quadrangle or a triangle.



(a) Input image

(b) Texture image

Figure 9: Planar patches

Each pixel in a planar patch of a texture image is geometrically transformed to a pixel in a corresponding planar patch of an input image. Then the planar patch of the texture image is warped to the input image.

If a planar patch is a quadrangle, we use homography. Homography is a  $3 \times 3$  matrix and has 8 degrees of freedom, so we can determine it with four or more corresponding points between the two images. Fig.10 shows texture rendering about a quadrangle patch. If all four vertices of the patch are tracked in the input image, homography is determined and the texture image is warped. If a planar patch is a triangle, we use affine transformation instead. Since affine matrix is a  $3 \times 3$  matrix with 6 degrees of freedom, it is determined with three or more corresponding points. If all three vertices of the patch are tracked, affine matrix is computed and the texture image is warped.

After rendering the texture image to the planar patches, the circumference of T-shirt is still not rendered. Then we interpolate the circumference with the color of rendered neighbor texture.

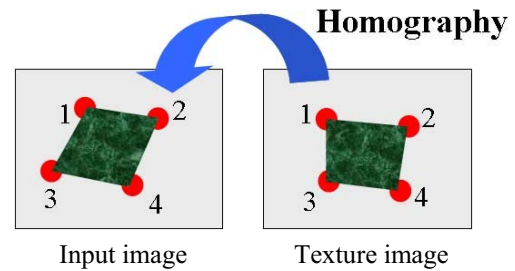


Figure 10: Texture rendering of a quadrangle patch

### 5. Experimental Results

In this article, we performed four types of experiments described below.

- Marker tracking about basic motion
- Case of motion with large deformation
- Case of motion with large rotation
- Texture shadowing

These experiments are performed on a PC with 2.8GHz CPU and 1024MB RAM. Texture image in these experiments is shown in Fig.11.





Figure 11: Texture image for experiments

### 5.1. Marker Tracking about Basic Motion

We applied our approach to the images with rotation and cloth stretching as basic motion. Rotation means that a user turns around in the spot. Several markers are sometimes occluded by rotation. Cloth stretching means that a user stretches clothes and changes the 3D shape of them. Input images and output images with rotation are shown in Fig.12. Input images and output images with cloth stretching are shown in Fig.13.

As is shown in Fig.12, the texture image is properly warped to the planar patches. If a marker is not detected in input image, the planar patch constructed with the marker is not rendered. However, the marker is re-corresponded and texture is rendered on the planar patch when a user turns around and the marker appears in input image again.

Fig.13 shows that our approach can be applied to the change of 3D shape of clothes by 2D marker tracking. Virtual clothing system with 2D marker tracking and 2D texture transforming typically takes shorter computational time than the one with 3D texture transforming and computer graphics. Computational time is an important feature for interactive virtual clothing system. Minimum frame rate of our system is 15.6 fps and average frame rate is 16.2 fps. We think that  $320 \times 240$  is appropriate image size for our system according to the trade-off between computational time and image size.

### 5.2. Case of Motion with Large Deformation

When a user stretched his/her arms out, the markers significantly moves with large deformation of the clothes. Input images and output images in such a case are shown in Fig.14.

If we attached markers on clothes, the markers' positions are deformed with human motion. Since it is considered that corresponding markers with clothes deformation is difficult, some related works of virtual clothing using markers render a texture only on the center of clothes[20]. To solve this problem, we track markers after corresponding all of them. Fig.14 shows that markers are corresponded and a texture is naturally rendered even if the markers' positions are largely changed from those in the first frame. Thus, we can say that our approach is robust to the large deformation of

the markers' positions and that marker tracking is effective for virtual clothing system using markers.

### 5.3. Case of Motion with Large Rotation

In case of large rotation of a user, a significant number of markers are occluded, so the tracking can be failed. For avoiding bad effects caused by such occlusion, the markers are re-corresponded as described in section 4.3.

In order to verify effectiveness of marker re-corresponding, we applied our approach to the images that several markers are occluded at one time. Several markers are self-occluded when a user turns sideways. Input images and output images with a user turning around more widely than input images of the experiment in section 5.1 are shown in Fig.15.

As is shown in Fig.15, texture is not properly rendered onto both sleeves. Planar patches on the left sleeve are not rendered because of the self-occlusion. As is described in section 4.4, a planar patch of the texture image is not warped on input image when not all markers of the planar patch are tracked. Planar patches on the right sleeve are not rendered properly, because markers on right sleeve are wrongly re-corresponded. If a number of markers have been occluded at one time, the marker which appears again in recent input image can be corresponded to wrong marker in texture image. Thus our approach can sometimes fail in the marker tracking when the user rotates him/herself widely.

To improve marker re-corresponding, we can place restraints on positions of markers and refresh marker correspondence when texture is rendered on the same pixel.

### 5.4. Texture Shadowing

After warping a texture image onto the whole area of T-shirt, the texture is not blended in an input image well. To exhibit the output image as if a user wears other clothes in real scene, the 3D deformation of clothes need to be considered. However, we aim at the realization of the interactive virtual clothing. Therefore, we increase the realism of output image by adding shadows to the texture image rather than using 3D shape of the object. The average brightness of pixels in green or white area is computed, and the brightness of each pixel has the difference from the average brightness as the shadow. We simply overlay the shadow of the input image on each pixel of texture image. The result of adding shadows is showed in Fig.16.

You can see that the shadows of T-shirt are reflected to the texture image rendered on the output image. However, the texture image in the area of markers on T-shirt is not properly rendered and not seamless. It is considered that the difference of reflection properties of white and green fabric is one of the causes of this problem. Additionally, it is difficult to figure out which area a pixel on T-shirt in the input image is included in.

## 6. Conclusion

We have proposed the method of interactive virtual clothing with markers on clothes. We have also shown that marker tracking is effective against the deformation of markers' positions with human motion by the experimental results. Markers in an input image are corresponded to markers on clothes in a texture image according to the shape of the cross pattern on clothes in the first frame. Since the surface of clothes moves flexibly, we divide it into the multiple planar patches and render a texture image on overall area of clothes by warping the image onto each planar patch. The experimental results show that marker re-corresponding possibly fails when a number of markers are occluded at one time.

Future work is to increase the realism of the rendered texture with shadows in real scene. Moreover, we think of using a T-shirt with no marker in order not to prepare the special T-shirt for virtual clothing system. By using T-shirt with markers on it, we preliminary make our system to learn the relation of silhouette and positions of markers in learning stage. After that, the system renders a texture image on a T-shirt with no marker which a user wears in recognition stage.

## References

- [1] <http://www.orad.tv/>
- [2] Zhencheng Hu, Keiichi Uchimura, Hanqing Lu, " Fusion of Realities for Vehicle Navigation," International Symposium on Computer Vision, Object Trackin and Recognition, 2004.
- [3] M. Billinghurst, D. Belcher, A. Gupta, K. Kiyokawa, " Communication Behaviors in Co-located Collaborative AR Interfaces," International Journal of Human Computer Interaction (IJHCI), 2002.
- [4] Y.Okuno, H.Kakuta, T.Takayama and K.Asai, " Jellyfish Party, Blowing Soap Bubbles in Mixed Reality Space," ISMAR 2003, pp. 358-359, 2003.
- [5] Mark Billinghurst, Hirokazu Kato, Ivan Poupyrev, " The MagicBook, a transitional AR interface," Computers & Graphics, vol.25, pp. 745-753, 2001.
- [6] Mototsugu Emoro, Hideo Saito, " Texture Overlay onto Deformable Surface Using Geometric Transformation," ICAT 2003, 2003.
- [7] Shinichiro Hirooka, Hideo Saito, " Displaying Digital Documents on Real Paper Surface with Arbitrary Shape," ISMAR 2003, pp. 278-279, 2003.
- [8] Tardif, J.-P., Roy S., Trudeau M., " Multi-projectors for arbitrary surfaces without explicit calibration nor reconstruction," The 4th International Conference on 3-D Digital Imaging and Modeling (3DIM'2003), pp. 217-224, 2003.
- [9] Y. Genc, S.Riedel, F. Souvannavong, C. Akinlar, N. Navab, " Marker-less Tracking for Augmented Reality, A Learning-Based Approach," ISMAR 2002, pp. 295-304, 2002.
- [10] L. Torresani, D. Yang, E. Alexander, and C. Bregler, " Tracking and modeling non-rigid objects with rank constraints," Proc. IEEE Conference on computer Vision and Pattern Recognition, 2001.
- [11] J. Pilet, V. Lepetit and P. Fua, " Real-Time Non-Rigid Surface Detection", Conference on Computer Vision and Pattern Recognition, 2005.
- [12] David G. Lowe., " Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91-110, 2004.
- [13] I. Guskov, L. Zhukov, " Direct Pattern Tracking on Flexible Geometry," The 10th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision 2002 (WSCG 2002), Journal of WSCG, vol. 10, p. 203, 2002.
- [14] Bradley, D., Roth, G., " Augmenting Non-Rigid Objects with Realistic Lighting," NRC/ERB-1116, 2004.
- [15] Michael Keckeisen, Stanislav L. Stoev, Matthias Feurer, Wolfgang Straser, " Interactive Cloth Simulation in Virtual Environments," IEEE Virtual Reality 2003, pp.71-78, 2003.
- [16] Vassilev T., " Dressing Virtual People", SCI'2000 conference, 2000.
- [17] K. Bhat, C. Twigg, J. Hodgins, P. Khosla, Z. Popovic and S. Seitz, " Estimating Cloth Simulation Parameters from Video," Proc. ACM SIGGRAPH 2003, pp. 37-51, 2003.
- [18] Jun'ichi Hoshino, Hirofumi Saito, Masanobu Yamamoto, " A Match Moving Technique for Merging CG Cloth and Human Video," Journal of Visualization and Computer Animation, vol. 12, no.1, pp. 23-29, 2001.
- [19] Yoshikazu Tagawa, Tomokazu Kakimi, Ryugo Kijima, Takeo Ojika, " Method for Dynamic Cloth Simulation and its Application in the Virtual Fashion System," Proc. Int. Conf. on VSMM '98, pp. 110-115, 1998.
- [20] Akinori Taguchi, Terumasa Aoki, Hiroshi Yasuda, " MIRACLE , A study on Virtual Mirror system and its Application of T-shirt Clothing", 2003 International Conference on Software, Telecommunications and Computer Networks, pp. 666-670, 2003.

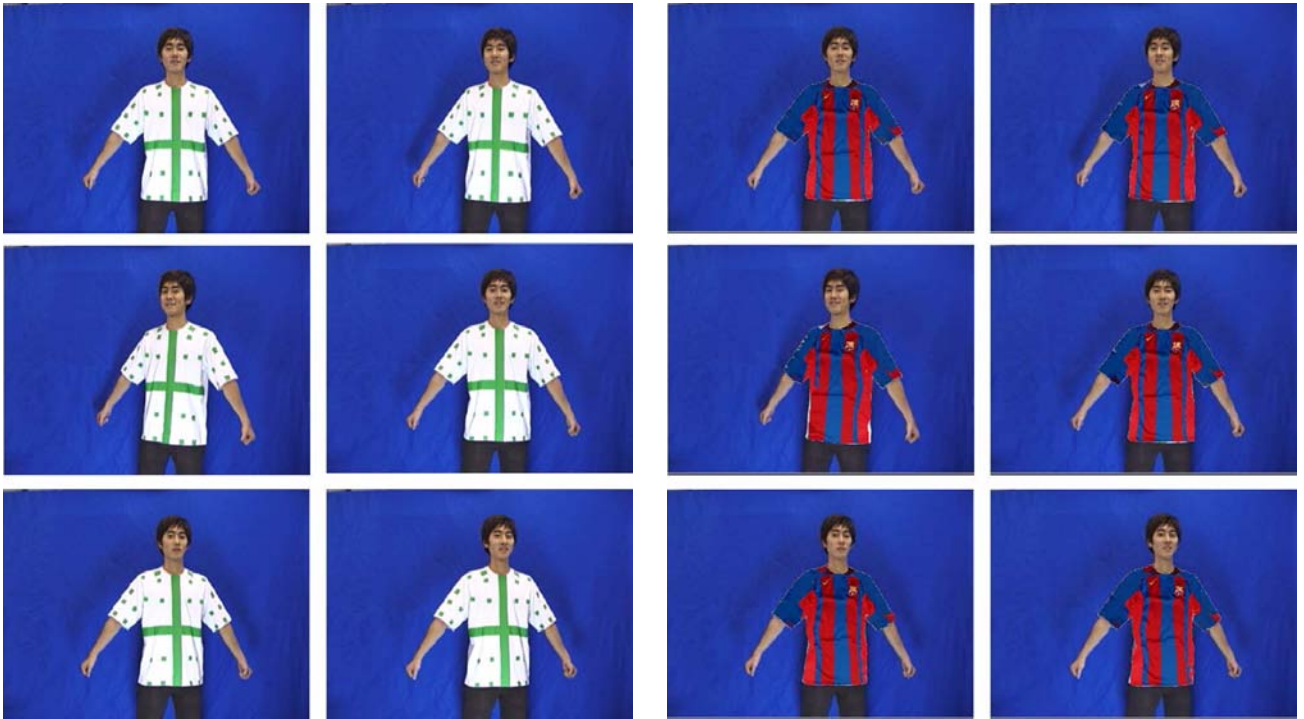


Figure 12: Basic motion with rotation of a user. Left: input images. Right: output images.

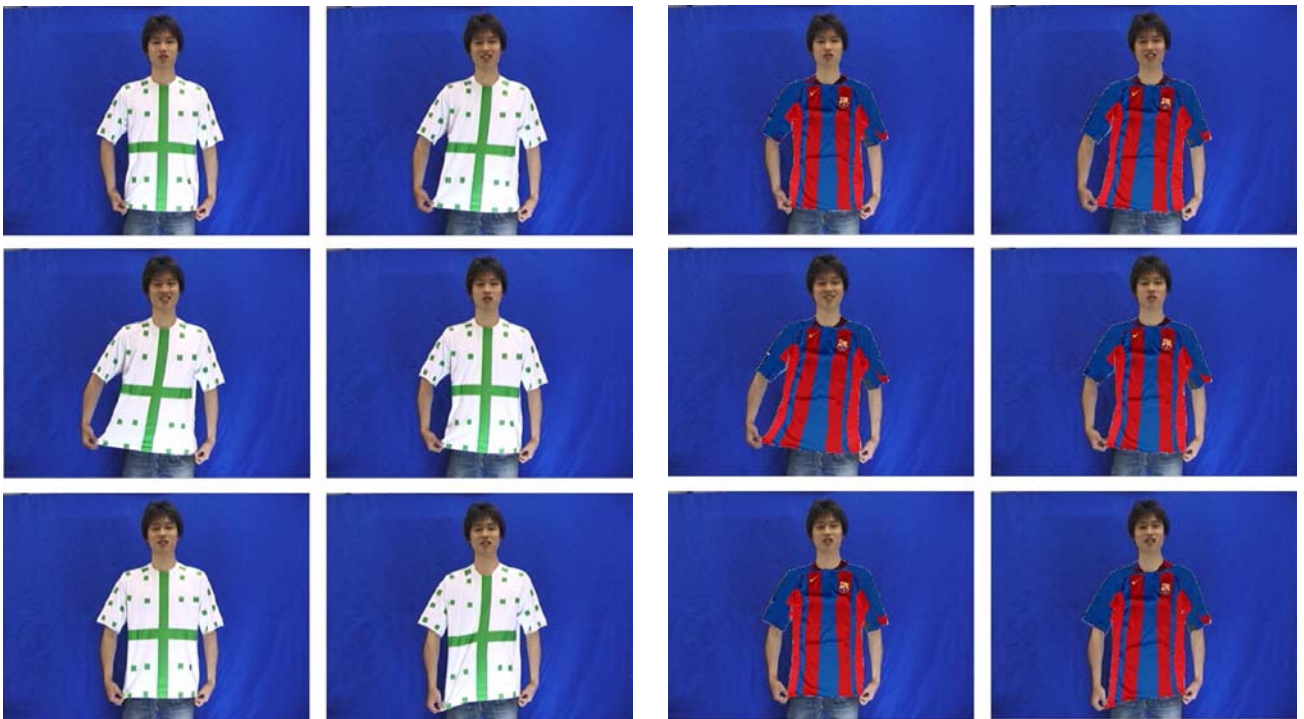


Figure 13: Basic motion with cloth stretching. Left: input images. Right: output images.



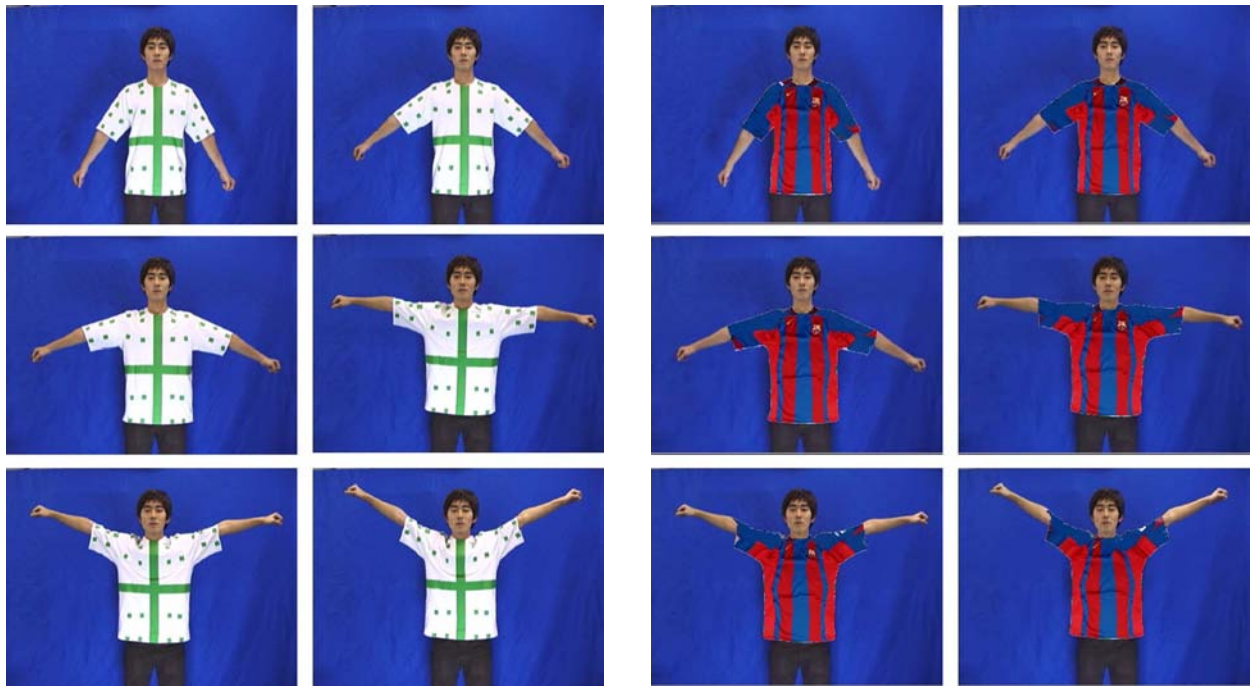


Figure 14: Motion with large deformation of clothes. Left: input images. Right: output images.

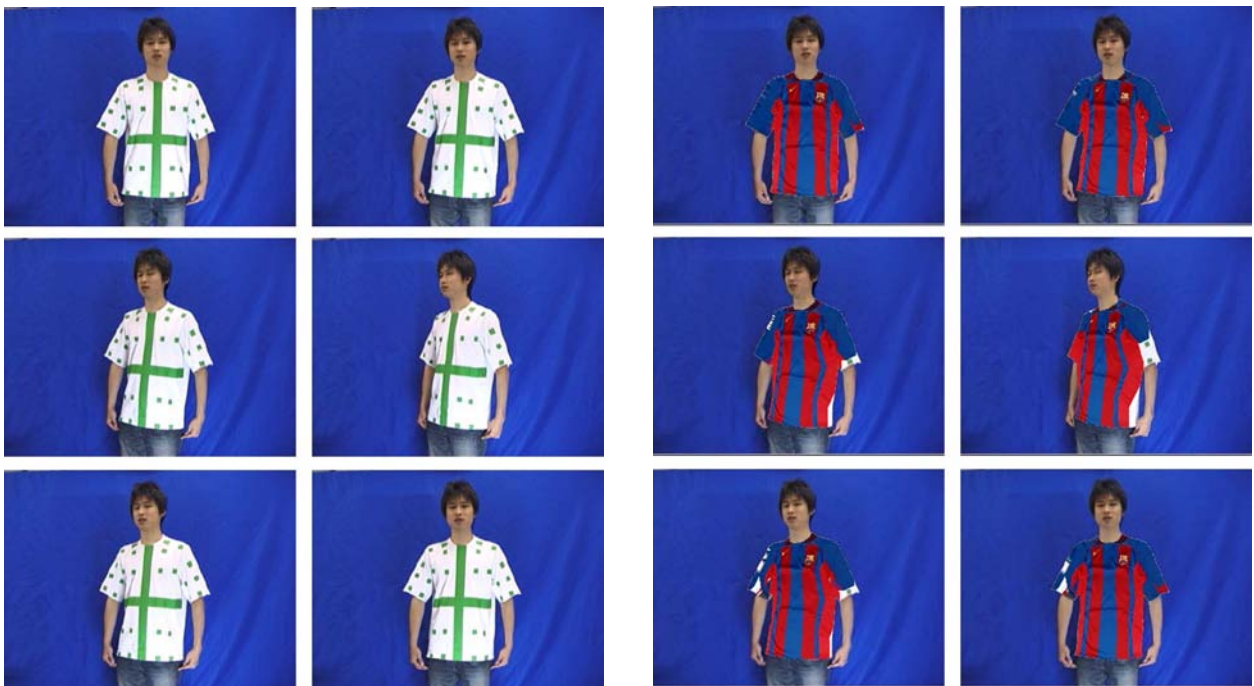


Figure 15: Motion with wide rotation of a user. Left: input images. Right: output images.



(a) Input image

(b) Output image without shadowing

(c) Output image with shadowing

Figure 16: Texture shadowing