Texture Overlay for Virtual Clothing Based on PCA of Silhouettes

Jun Ehara*   Hideo Saito†

Graduate School of Science and Technology, Keio University, Yokohama, Japan

ABSTRACT

In this paper, we propose a method for overlaying an arbitrary texture image onto a surface of a plain T-shirt worn by a user. For overlaying arbitrary textures onto the surface of the T-shirt, we need to know the deformation of the surface. For estimating the deformation of the surface from the input images, we use a two-phase process: learning and searching. In the learning phase, the system learns the relationship between the deformation of the surface and the silhouette of the T-shirt region in the image. A database of a number of training images in which a person wearing a T-shirt with markers moves through a variety of positions is used for this learning. Using the database, the system can learn the relationship between the shape of the silhouette and the surface deformation that is provided by the 2D positions of the markers on the surface of the T-shirt. In the searching phase, the silhouette of the user's T-shirt is extracted from the input image, and then, a search for a similar silhouette in the database is conducted in the subspace of the silhouette, which is computed using a PCA of the database. By using the proposed method for estimating the deformation of the surface of the T-shirt, we perform experiments for overlaying virtual clothing.

CR Categories: I.4.9 [Image Processing and Computer Vision]: Applications

Keywords: augmented reality, virtual clothing, texture mapping, deformable surface, silhouette, PCA, learning

1 INTRODUCTION

In Augmented Reality (AR) techniques, virtual and real objects are integrated in a real environment. AR systems need to run interactively and in real time, and an alignment of virtual and real objects is one of the most significant issues[1]. In recent years, AR has been used for a broad range of applications such as car navigation systems[10], municipal engineering[4], and entertainment[11].

We intend to apply an AR technique to a virtual clothing system, in which a user can watch himself/herself wearing clothes even if the user is not wearing those clothes. In the virtual clothing system, an AR technique is used to overlay arbitrary texture patterns onto the surface of clothes. Related systems based on overlaying virtual patterns onto real surfaces have been proposed previously, such as overlaying virtual patterns onto a planar surface[3] and curved surface[7]. A clothing surface is very complicated because it is non-rigid and its shape changes considerably when the cloth stretches, wrinkles, and self-occludes. Therefore, overlaying a texture onto the surface of clothes is more difficult than overlaying that onto the surface of a rigid object.

There are several kinds of studies on virtual clothing. An example of an approach is to overlay texture onto the surface of a user by using computer graphics techniques[2, 5]. In this approach, 3D

shape data must be previously obtained with 3D measurement instruments or installed as a 3D model of a person, and clothing data is transformed in a 3D space using computer graphics techniques. However, a large amount of 3D data is difficult to handle, which prevents real-time computation.

An approach using motion estimation of a person from an image sequence was also proposed. Motion estimation is a method to transform a certain model into a corresponding image of the person and recognize the human motion with the transformed model shape. An example of research using motion estimation for 3D human body modeling using the spatial and temporal gradient method is described in [9]. In [13], a dynamic simulation of a cloth is described. However, the approach using motion estimation does not recover the deformation of the surfaces of objects because it uses global shape fitting of human models.

Some related studies use patterns on the surface of clothes such as colorful markers[14] or natural features[12] on a T-shirt to determine the deformation of the T-shirt. These approaches do not need a 3D model in advance, so they reduce the computational burden of constructing a model of each object or person and run in real time. However, sleeves are significantly deformed and the waist area of T-shirts is considerably self-occluded compared to the chest area of a T-shirt. Rendering a texture on the overall area of a T-shirt is considered to be difficult. Therefore, we have proposed a method to overlay a texture image onto the overall area of a T-shirt in real time by tracking markers on the T-shirt[6]. In this method, the deformation of the clothes is provided by the 2D positions of the markers in a captured image. Then, an arbitrary texture image is warped according to the 2D positions. However, this method requires the user to wear the T-shirt with the specific markers.

In this paper, we propose a method to overlay an arbitrary texture image onto the surface of a plain T-shirt worn by a user. In overlaying the texture image onto the surface of the T-shirt, the deformation of the surface needs to be known. However, knowing the deformation of the surface of the plain T-shirt in an input image is difficult. In the proposed method, we use the silhouette of the T-shirt, which is easy to extract from the input image. For estimating the deformation of the surface of the T-shirt from the extracted silhouette, we prepare a database of the relationship between the silhouette and the deformation of the T-shirt by using a number of training images. In the subspace of the shape of the silhouette computed using principal component analysis (PCA) of the database, a search for a shape similar to that of the silhouette of the input image is conducted. The deformation of the input image can be estimated by using the 2D positions of the markers in the training images that were found in the search.

2 PROPOSED METHOD

2.1 Method outline

The setup of our virtual clothing system is shown in Fig.1. A web camera is connected to a PC and set on a display. A user wears a plain T-shirt and stands in front of a blue curtain. An image of the user is taken from the camera. Then, a prepared texture image of the clothes that the user wants to try on is transformed and overlaid onto the captured image. When the overlaid image is shown on the

---

*e-mail: eha@ozawa.ics.keio.ac.jp
†e-mail: saito@ozawa.ics.keio.ac.jp

display, the user can feel as if he/she is trying on other clothes and looking at himself/herself in a mirror.
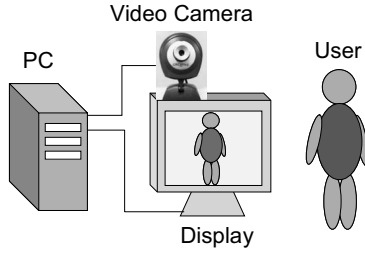

Figure 1: System setup

An outline of the proposed method is shown in Fig.2. Our method is divided into two phases: the learning phase and the searching phase. The former is an off-line process, and the latter is an on-line process. In the learning phase, we train our system on the relationship between the positions of the markers and the silhouettes of the T-shirt with various shapes. In the searching phase, our system is used. The texture image of the clothes is appropriately transformed and overlaid onto the input images according to the deformation of the T-shirt. The silhouettes similar to those of the training images are selected in the subspace computed with PCA of the silhouettes. Then, the selected training images are used to estimate the deformation of the T-shirt in the input image. Here, we assume that the deformation of the T-shirt depends on the silhouette of the T-shirt, which can be clearly extracted from the input image.
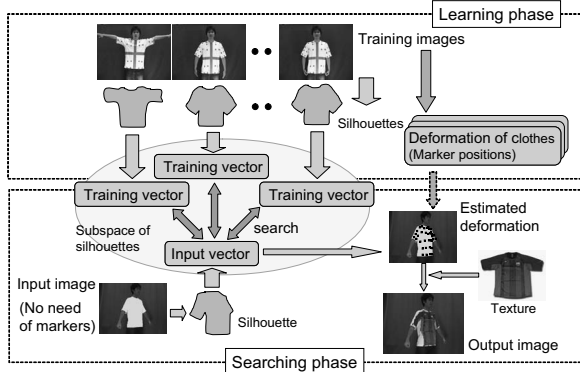

Figure 2: Method outline

2.2 Learning phase

In the learning phase, a person wears the special marked T-shirt shown in Fig.3. Then, the person makes various gestures in front of a web camera. A number of images are captured by the web camera. We call these training images. We trained our system with the silhouettes of the T-shirt shown in the training images used in this phase.

First, we extract the silhouette features from the silhouette of the T-shirt in the training image, as described below. After extracting the silhouette of the T-shirt from the training image, we transform the outline of the silhouette into 360 sample points in the polar coordinate system, as shown in Fig.4(a). In our implementation, the sampling interval of the angle θ is one degree. For example, we perform this process on the outline shown in Fig.4(b), and then, the 360 sample points shown in Fig.4(c) are obtained. We calculate

Figure 3: Marker T-shirt

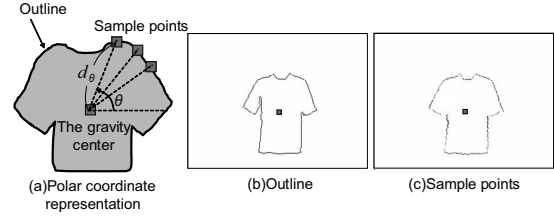the 360 distances from the center of gravity of the outline to every sample point.

Figure 4: Feature extraction

Next, we scale the size of the outline so that the average value of the 360 distances of the scaled outlines of each training image is equivalent to each other. This scaling process enables us to use training images with any outline size. The 360 distances of the scaled outline $d_\theta\,(\theta = 1, 2, \cdots, 360)$ are the silhouette features of the training image in this paper. We use PCA to represent the silhouette feature d_θ of the training image with a small number of principal component scores. The pth principal component score s_p is represented as follows:

$$s_p = \sum_{\theta=1}^{360} v_{p\theta} \left(d_\theta - \bar{d}_\theta \right), \qquad (1)$$

where $V_p = \left(v_{p1}, v_{p2}, \cdots, v_{p360} \right)$ is the pth eigenvector of PCA, and $\bar{d}_\theta$ is the average of d_θ of the training images. The dimension of the silhouette features is reduced when the sum of contribution ratios is greater than 90%.

Then, the silhouette features are converted into feature vectors, which consist of a small number of principal component scores in the subspace. In this paper, we call the feature vectors of the training images the training vectors. Moreover, as well as extracting the silhouette from the training image, the deformation is constructed in the learning phase. The deformation of the input image is provided by the 2D positions of the markers in the training images. These positions are manually retrieved. If some markers are self-occluded, their positions are fixed on the neighbor markers which are showing in the training images. Then the relationship between the training vector and the deformation of the T-shirt is stored in the database. To reduce the burden of retrieval of the positions of the markers, we attach recognizable patterns like ARTag[8] on the T-shirt.

2.3 Searching phase

In the searching phase, a user wears a plain T-shirt. The color of the plain T-shirt can be arbitrary as long as it is different from the other colors shown in the input image captured by the camera. The texture image of the clothes a user wants to try on is transformed to fit the silhouette in the input image and overlaid onto it. As the

overlaid image is shown on the display, a user can look at himself/herself wearing other clothes.

In the same way as that of the learning phase, the silhouette is extracted from the input image and transformed into the 360 distances, which are the silhouette features of the input image. These features are transformed into the feature vector in the subspace computed with PCA. In this paper, we call this feature vector of the input image the input vector. Then, a search for the training vectors, which are within a certain distance of the input vector in the subspace, is conducted, and a predetermined number of training vectors are selected beginning with the closest one to the input vector. The silhouettes that correspond to the selected training vectors are supposed to be similar in shape to the silhouette in the input image.

To deform the texture image, we need to know the deformation of the T-shirt of the input image. This deformation is represented by the 2D positions of the markers. The 2D positions of the markers in the input image are estimated by integrating the positions of markers in training images that correspond to the selected training vectors. The integration weighting is inversely proportional to the distance from the input vector to each selected training vector in the subspace. The texture image is divided into a number of planar patches according to the positions of the markers. Each planar patch in the texture image is warped to the corresponding planar patch constructed with the estimated markers in the input image. However, the texture image is overlaid onto only the silhouette in the input image.

3 EXPERIMENT

We demonstrate the effectiveness of the proposed method via experiments. In these experiments, we used training and input images with a 320×240-pixel size. The training images consist of two types of images: 48 images of a person who leans his body sideways moving his arms up and down, and 18 images of a person who turns around standing in one spot. The former are shown in Fig.5 and the latter are shown in Fig.6. The 2D positions of 37 markers are manually retrieved from those training images. The silhouettes of the T-shirts in both training and input images are converted to ten-dimensional feature vectors that consist of ten principal component scores. There are up to three training vectors that are selected in the subspace in the searching phase. These experiments were performed on a PC with a 2.8-GHz CPU and 1024 MB of RAM. Windows XP was installed on the PC. Though a user can wear a plain T-shirt without any marker on the surface in the searching phase, we use a T-shirt with markers like that used in the learning phase in these experiments to verify that the deformation of the T-shirt in the input image can be estimated properly.
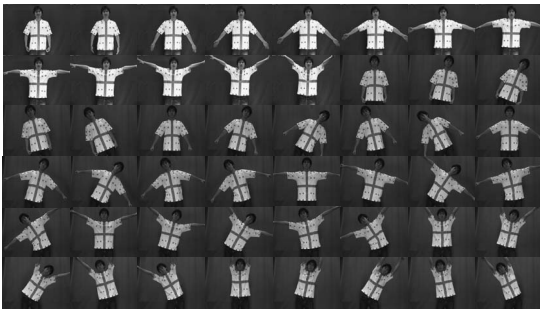

Figure 5: Training images with body leaning sideways and arms moving up and down

Figure 6: Training images with rotation of body

3.1 Experiment for silhouettes of different sizes

In this paper, the size of the silhouette of the input image is widely different from those of the silhouettes of the training images to verify that the texture image is properly transformed to fit the various silhouette sizes. The silhouette in the input image shown in Fig.7(a) is much smaller than those in training images. The 360 sample points extracted from the outline are shown in Fig.7(b). In Fig.7(c), black points are estimated markers that represent the deformation. They approximately coincide with the real markers. This means that the deformation of the T-shirt in the input image has been properly estimated. Though a user of our system usually wears a plain T-shirt, the user in this experiment wears a T-shirt with markers to verify whether the deformation has been correctly estimated. The output image with the selected texture is shown in Fig.7(d).

The training images that are selected in the subspace and their integration weighting for estimating the deformation of the input image of Fig.7(a) are shown in Fig.8. The integration weighting is proportional to the similarity of the shape of the silhouette. Our system runs at 15.6 frames per second on average.
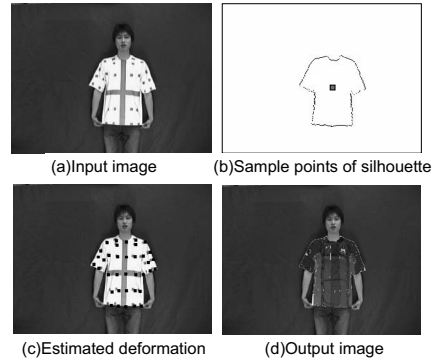

(a)Input image   (b)Sample points of silhouette

(c)Estimated deformation   (d)Output image

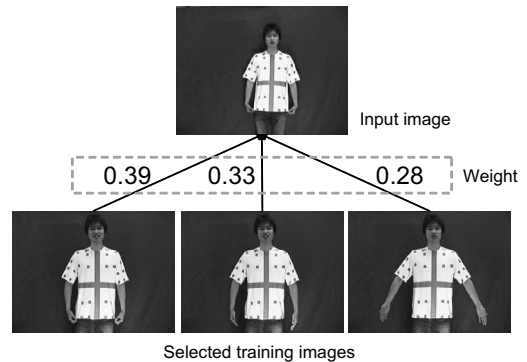Figure 7: Result for silhouettes of different sizes (1)

Figure 8: Result for silhouettes of different sizes (2)

3.2 Experiment for interpolating pose

A user turns his body sideways in the input image in this experiment. The result for the input image of a user who turns his body is shown in Fig.9. The integration weightings of training images 1 and 2 in estimating the deformation are heavier than that of training image 3. Moreover, the angle of rotation of the body in the input image is between those in training images 1 and 2. This means that the deformation of the input image can be estimated by interpolating the deformations of these training images, even if the database of the training images does not have exactly the same silhouette as that of the input image. No marker is attached on the body side of the T-shirt, so the texture image cannot be overlaid on the area of the body side in the output image. However, the texture image can be overlaid onto the proper position because the deformation of the input image is properly estimated by interpolating that of the training images.
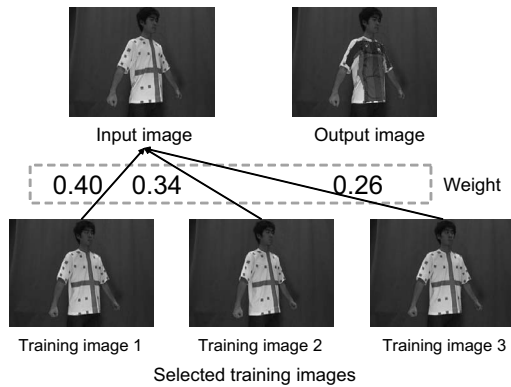

Figure 9: Result for interpolation of pose

3.3 Experiment for extrapolating pose

The result for the input image of a user who leans his body to the side and stretches his hands upward is shown in Fig.10. In Fig.10(a), he is leaning his body at a greater angle than that of any pose in the training images. Estimated markers are shown in Fig.10(c) as black points that represent the deformation of the T-shirt in the input image. In the red circle in Fig.10(c), estimated markers on the right sleeve are plotted on different positions than the markers on the T-shirt. In addition, the direction of vertical stripes in the overlaid texture is different from the proper direction. The deformation of the input image is wrongly estimated. This is because the input vector cannot be properly composed by extrapolation with the training vectors. The input vector cannot be composed of any combination of training vectors, so the estimated markers are not correct. To solve this problem, we need (1) to have a larger number of training images with different silhouettes, (2) to change the method of extracting features from images, and (3) to have a new method of integrating the deformation of the T-shirt in training images during the searching phase.

4 CONCLUSION

We proposed a method for overlaying arbitrary texture images onto a surface of a plain T-shirt in real time for a virtual clothing system. The deformation of the surface of the T-shirt was estimated from the silhouette of the T-shirt taken from the input image based on a pre-collected database of a number of shapes of the T-shirt. We performed experiments for virtual clothing based on the proposed
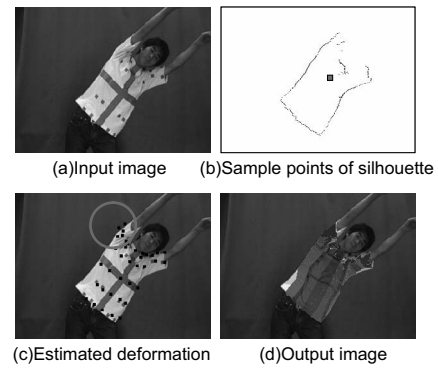

Figure 10: Result for extrapolation of pose

method for estimating the deformation of the surface of the T-shirt. The experimental results demonstrate that the proposed method can estimate the deformation of the T-shirt and an arbitrary texture can be successfully overlaid onto the surface of the deformed T-shirt.

REFERENCES

[1] Ronald Azuma, Yohan Baillot, Reinhold Behringer, Steven Feiner, Simon Julier, and Blair MacIntyre. Recent advances in augmented reality. *IEEE Computer Graphics and Applications*, 21(6):34–47, 2001.

[2] Kiran Bhat, Christopher D. Twigg, Jessica K Hodgins, Pradeep Khosla, Zoran Popovic, and Steven Seitz. Estimating cloth simulation parameters from video. In *Proc. of the SIGGRAPH*, pages 37–51, 2003.

[3] Mark Billinghurs, Hirokazu Kato, and Ivan Poupyrev. The Magic-Book: a transitional AR interface. *Computers & Graphics*, 25(5):745–753, 2001.

[4] Mark Billinghurst, Daniel Belcher, Arnab Gupta, and Kiyoshi Kiyokawa. Communication behaviors in colocated collaborative AR interfaces. *International Journal on Human Computer Interaction*, 16(3):395–423, 2003.

[5] Luca Chittaro and Demis Corvaglia. 3d virtual clothing: from garment design to web3d visualization and simulation. In *Proc. of the eighth international conference on 3D Web technology*, pages 73–84, 2003.

[6] Jun Ehara and Hideo Saito. Texture overlay onto deformable surface for virtual clothing. In *Proc. of the ICAT*, pages 171–178, 2005.

[7] Mototsugu Emori and Hideo Saito. Texture overlay onto deformable surface using geometric transformation. In *Proc. of the ICAT*, pages 58–65, 2003.

[8] Mark Fiala. ARTag, a fiducial marker system using digital techniques. In *Proc. of the CVPR*, volume 2, pages 590–596, 2005.

[9] Jun'ichi Hoshino, Masanobu Yamamoto, and Hirofumi Saito. A match moving technique for merging CG cloth and human movie sequences. *j-J-VIS-COMP-ANIMATION*, 12(1):23–29, 2001.

[10] Zhencheng Hu, Keiichi Uchimura, and Hanqing Lu. Fusion of realities for vehicle navigation. In *International Symposium on Computer Vision, Object Trackin and Recognition*, 2004.

[11] Yasuhiro Okuno, Hiroyuki Kakuta, Tomohiko Takayama, and Kazuhiro Asai. Jellyfish Party: Blowing soap bubbles in mixed reality space. In *Proc. of the ISMAR*, pages 358–359, 2003.

[12] Julien Pilet, Vincent Lepetit, and Pascal Fua. Real-time non-rigid surface detection. In *Proc. of the CVPR*, pages 822–828, 2005.

[13] Yoshikazu Tagawa, Tomokazu Kakimi, Ryugo Kijima, and Takeo Ojika. Method for dynamic cloth simulation and its application in the virtual fashion system. In *Proc. of the VSMM*, pages 110–115, 1998.

[14] Akinori Taguchi, Terumasa Aoki, and Hiroshi Yasuda. MIRACLE : A study on virtual mirror system and its application of T-shirt clothing. In *International Conference on Software, Telecommunications & Computer Network*, pages 666–670, 2003.