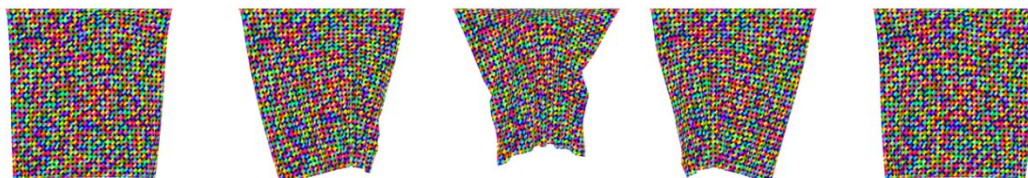


A Dense Surface Motion Capture System for Accurate Acquisition of Cloth Deformation

Alexandros NEOPHYTOU, Jean-Yves GUILLEMAUT, Adrian HILTON
CVSSP, University of Surrey Guildford, Surrey, UK

DOI: 10.15221/15.120 <http://dx.doi.org/10.15221/15.120>



Abstract

In this paper, a system for dense surface capture of cloth deformation is presented. In the proposed pipeline cloth motion and dynamics are captured and reconstructed from a multi-view camera set-up. To allow precise tracking of a high number of surface points a tailored pattern is printed on the cloth. Several existing approaches make use of a printed pattern where point correspondence is determined using colour-coded vertex neighborhoods [4, 5]. In this paper we show that point correspondence can be improved and refined using a Laplacian mesh fitting process in the image domain. Results show an average increase of $\approx 20\%$ in the number of correctly labelled markers.

Keywords: cloth deformation, surface capture.

1. Introduction

Physical based simulations of cloth can generate beautiful animations taking into considerations various complex effects such as collisions, friction, air resistance and various cloth fabric properties. These methods are often used in the film industry due to the high quality output but are not so frequently used in games and on-line fashion applications due the high computational cost and inability to produce outputs in real time. An alternative to physical based modeling is motion capture where the motion is measured directly. This paper presents a system for accurate dense cloth motion capture.

2. Related Work

In order to capture precise, high frequency details of cloth dynamics, accurate acquisition systems are required. Over recent years, a number of approaches have been proposed with a common strategy; the use a printed pattern on the surface of the cloth.

In a typical pipeline, a pattern is designed and printed on the surface of the cloth such that local regions of the cloth geometry are uniquely identifiable and can be matched across multiple views. Using triangulation methods [7] the 3D location of each region can then be determined from corresponding image sets. Finally, the resulting point clouds are triangulated, hole filled and temporally smoothed to produce triangle meshes with static connectivity for each frame.

The primary challenge of cloth capture is increasing marker density while maintaining correspondence between markers. This becomes a difficult task, especially in the case where the surface is heavily occluded. While the reconstruction of a high detailed mesh requires a significant number of correspondences, marker mismatches can cause improbable cloth deformations.

Cloth markers are generally painted on the surface of the cloth with special care taken to avoid altering the material properties of the cloth. Cloth markers as used in existing work can be classified into two main categories: complex surface gradients which are typically detected using SIFT descriptors [3] and markers which are based on regions of constant colour [1-5].

Pritchard and Heidrich [3] used cloth with existing textures as markers. In their approach stereo correspondence is used in combination with SIFT descriptors [8] to establish correspondence. They make use of a novel seed-and-grow algorithm that extends the original SIFT algorithm to handle deformable geometry.

Guskov et al. [1] use a custom black and white checkerboard pattern that allows for easy marker localization. However, locally, markers are indistinguishable from each other and correspondence requires global reasoning. To solve this problem Guskov et al. [2] extend this work to introduce distinctive markers of different colour enabling correct handling of occlusions.

Scholz et al. [4] advance the technique proposed by Guskov et al. [2] by creating a non-repeating grid of colour markers using five colours. Each marker is associated with a unique neighborhood of three by three. Since each marker can be uniquely identified, correspondence errors are greatly reduced.

To maximize marker to image pixel ratio White et al. [4] use the full colour space instead of a small number of colours in the pattern. They generate a random pattern of coloured triangles which are matched to a reference mesh in an iterative procedure. They show how incorrect correspondences can be pruned using a strain minimization algorithm.

As mentioned above, several approaches make use of a printed pattern where point correspondence is determined using colour-coded vertex neighbourhoods [4, 5]. However, colour-coded local regions are not always easy to detect and identify especially in the case of occlusions due to cloth folding. Additionally, due to restrictions from the camera and printer, colour regions will not always be classified accurately, leading to incorrect marker matches. In this paper, we show that point correspondence can be improved and refined using a Laplacian mesh fitting process in the image domain.

2. Colour coded pattern

The proposed pattern is designed such that it satisfies the following criteria: 1) marker detection should be invariant to various illumination changes in the scene; 2) each marker should be uniquely identifiable by its immediate local neighborhood without requiring global reasoning; 3) marker detection should be able to handle motion blur due to the fast motion of moving cloth; and finally 4) the number of colours used to construct the pattern should not exceed 8. Higher number of colours means higher number of unique markers but also makes identifying and classifying distinct colours a difficult task due to restrictions imposed by the printer and camera.

The pattern consists of a mesh with colour coded triangle neighbourhoods centered at each vertex (figure 1). Each vertex can be identified from its unique colour combination of six neighboring triangles in a clockwise order. In order to detect individual triangle colours reliably, neighboring triangles cannot have the same colour. In addition, to account for rotations of the pattern, we do not allow vertex neighborhood combinations that are rotated versions of other neighborhood combinations. For example: Consider a neighborhood combination of 012345 (0=Red, 1=Yellow etc.). This combination is the same as 501234, 450123, 345012, 234501, and 123450.

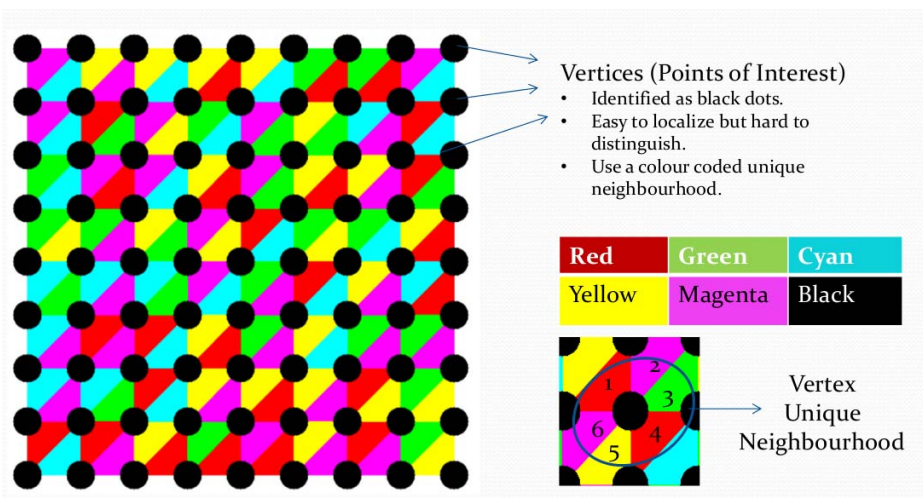


Fig. 1. Proposed colour coded pattern with 6 colours: red, green cyan, yellow, magenta and black.

3. Cloth motion capture overview

An overview of the proposed pipeline is illustrated in figure 2. Firstly, each frame is processed for marker detection and identification. Colour regions in each frame are identified using RGB colour models learnt from multiple examples in various illuminations. Once the colour coded vertex neighbourhoods are extracted initial correspondences can be calculated by matching to a reference triangle mesh. Marker correspondences can be refined by fitting the template mesh to the unlabelled points using a Laplacian mesh fitting technique the image domain (figure 5). Once the 2D mesh is refined to better fit the observe data, triangulation is performed to obtain the final mesh for each frame.

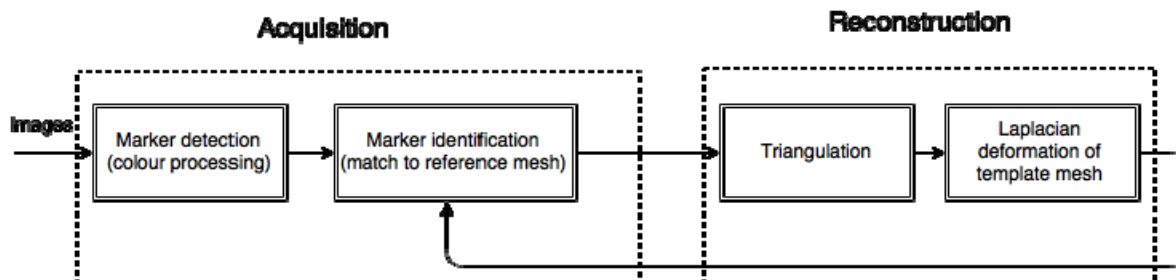


Fig. 2. Overview of the proposed pipeline. In the acquisition step markers are detected, identified and matched across multiple views. In the reconstruction step, triangulation is performed to obtain the 3D location of each point of the reference mesh. Then a Laplacian deformation process is used to reconstruct the whole mesh. This process is repeated until no more correspondences can be found.

4. Marker detection

In order to find vertex correspondences, firstly, the colour-coded neighbourhoods need to be identified (figure 3). This is achieved using Gaussian RGB colour models to classify each individual pixel of the input image into one of the eight colour classes. Each pixel is assigned a colour based on the smallest Mahalanobis distance to each colour distribution. Finally, regions or blobs of the same colour class are extracted using a blob detection algorithm. Blobs are refined, using an area threshold to remove very large regions and regions that are too small. The next step involves matching the unlabelled mesh vertices to the reference mesh by considering the colour coded triangle neighbourhoods of each vertex.

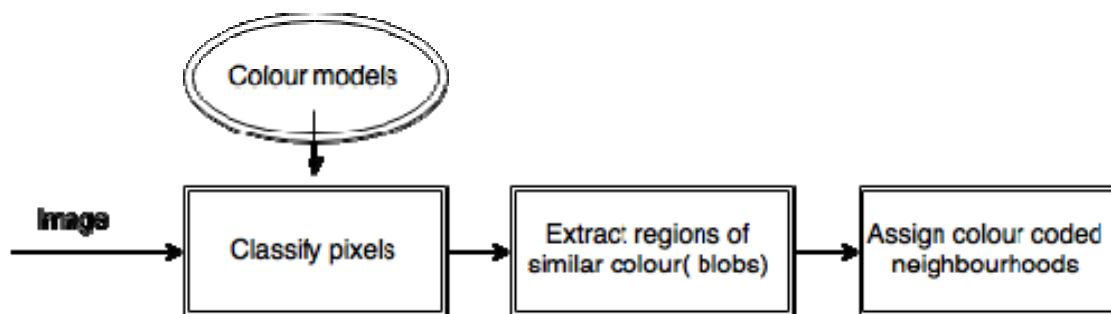


Fig. 3. Marker detection procedure. Colour regions in each frame are identified using RGB colour models learnt from multiple examples in various illuminations. Once the colour coded vertex neighbourhoods are extracted initial correspondences can be calculated by matching to a reference triangle mesh.

5. Marker labelling

For each vertex, the triangle neighborhood is extracted by taking the six closest coloured regions (figure 4). To account for foreshortening, regions are selected in a circular manner by considering the angle between the center of each vertex blob and the center of adjacent regions. Finally, mesh vertices are labelled by matching the triangle neighbourhoods to the reference mesh. Incorrect correspondences can be pruned by ensuring that labelled vertices have the same vertex neighbors as in the reference mesh.

Marker identification fails in the following two cases: 1) marker is not visible; and 2) marker colour neighborhood is not extracted accurately. This may occur due to colour miss-classification or when parts of the triangle neighborhood are occluded.

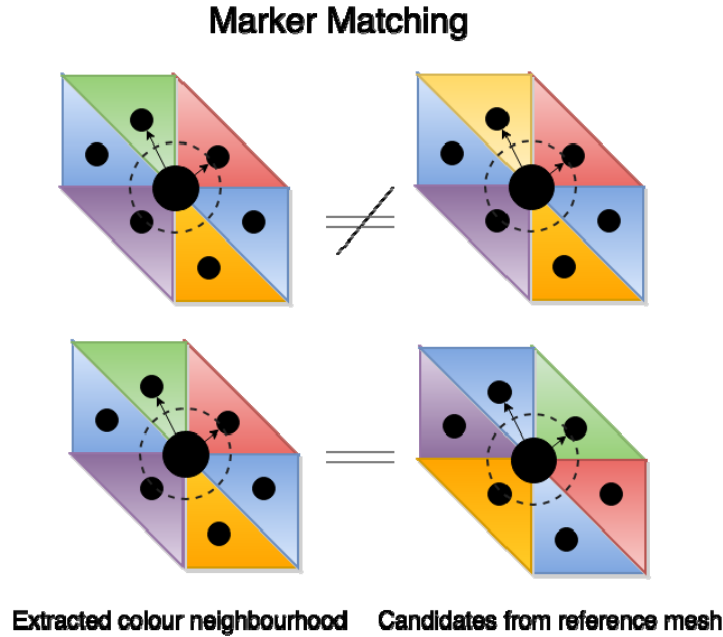


Fig. 4. Marker labelling. Mesh vertices are labelled by matching the triangle neighborhoods to the reference mesh.

6. Laplacian Mesh Fitting

Once initial marker correspondences are calculated, triangulation is performed to obtain the 3D location of each labelled point of the reference mesh. A Laplacian deformation framework in 3D [6] is used to approximate the position of all mesh vertices for each frame using the position of the labelled vertices as boundary constraints:

$$E(V') = \sum_{i=1}^{N_v} \|L v'_i - \delta_i\|^2 + \sum_{l=1}^{N_c} \|v'_l - c_l\|^2 \quad (1)$$

where L is the discrete Laplace-Beltrami operator, v'_i are the deformed mesh vertices δ_i are the differential coordinates and c_l are labelled vertices. This approach involves finding the optimal set of vertices that best preserve the geometrical properties of the local areas of the mesh.

By back-projecting the 3D mesh to each view we can observe how far the mesh lies from the actual data. Typically, the resulting 2D mesh lies far away from the observed data in areas with high number of unlabelled vertices (figure 5). The mesh can be refined in an iterative procedure by fitting to the unlabelled mesh vertices in the image domain (2D). In each step, vertex correspondences are calculated between the reference mesh and the unlabelled vertices using closest point. The 2D mesh is then deformed using the unlabelled mesh vertices as boundary constraints in equation 1. As shown in figure 5 this technique can significantly increase the number of labelled mesh vertices.

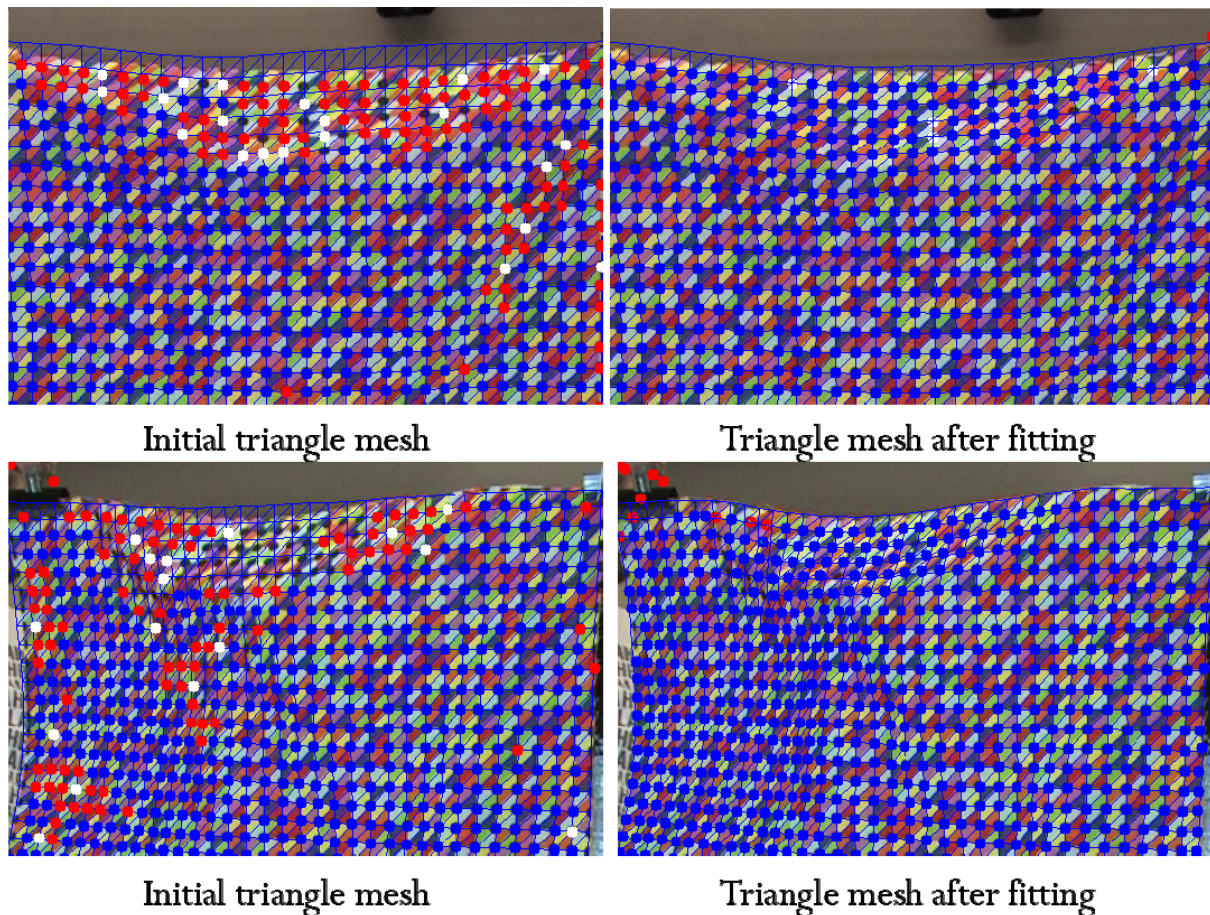


Fig. 5. Laplacian mesh fitting. In an iterative procedure the 2D mesh is fitted to the unlabelled mesh vertices. Blue, white and red dots represent labelled, incorrectly labelled, and unlabelled mesh vertices respectively.

7. Results

This proposed pipeline is evaluated using two cloth sequences with varying motions. The pattern is printed on a 70x70cm cotton drill cloth with 1296 markers. Three front view cameras are arranged in a semi-circle at about 2 meters distance from the target (figure 6). A frame rate of 25 frames per second is used with an image resolution of 1920x1080 pixels.

We compare the average number of observed markers obtained with and without the mesh fitting step and by varying the number of cameras used for reconstruction. As shown in table 1 the Laplacian mesh fitting process can significantly increase the number of observed markers which results in a more detailed and accurate mesh reconstruction. The highest discrepancies occur in areas where the clothing is heavily folded and therefore vertex neighborhoods are occluded from all viewpoints. Figure 7 illustrates some example cloth surface reconstructions obtained after using the Laplacian mesh fitting process.

8. Limitations

One limitation of the system is the number of unique markers that can be generated which roughly 3000. An obvious solution is to increase the number of colours used to design the pattern. However, this may introduce additional problems when classifying colour regions due to limitations from the printer and camera. A more suitable solution is to apply global reasoning. In other words duplicate markers can exist as long as their immediate vertex neighbors are different.

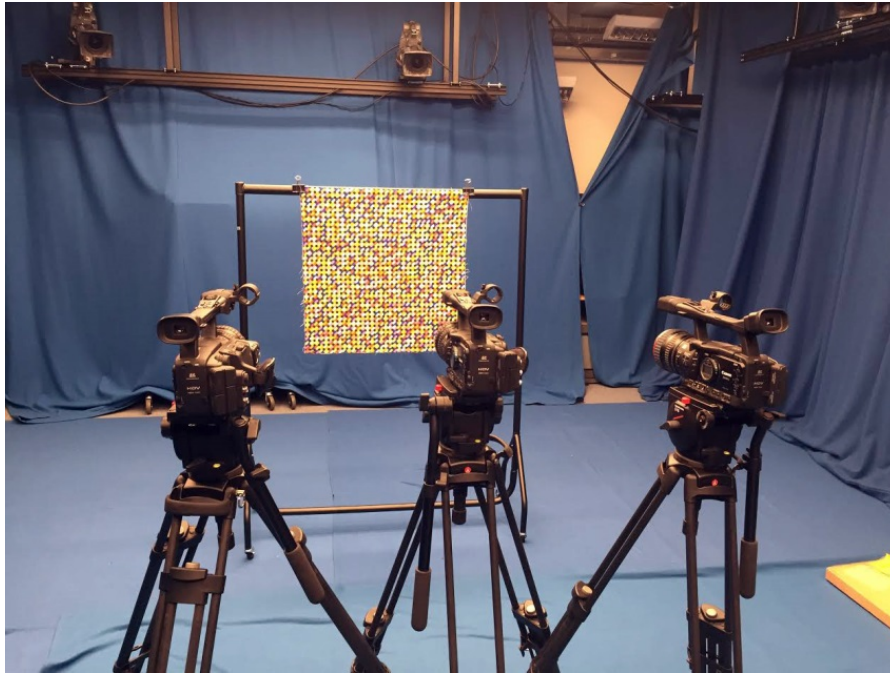


Fig. 6. Camera set-up using three frontal cameras.

Table 1a. Results without Laplacian mesh fitting.

	Sequence 1	Sequence 1	Sequence 2	Sequence 2
#cameras	2	3	2	3
total markers	1296	1296	1296	1296
recovered	955	1001	898	973
percentage	73.6%	77.2%	69.2%	75%

Table 1b. Results with Laplacian mesh fitting. An average accuracy of $\approx 20\%$ is achieved.

	Sequence 1	Sequence 1	Sequence 2	Sequence 2
#cameras	2	3	2	3
total markers	1296	1296	1296	1296
recovered	1201	1226	1188	1220
percentage	92.6%	94.5	91.6%	94.1%

9. Conclusion and future work

This paper presented a dense motion capture system for acquisition of cloth deformation. Using unique colour-coded local regions printed on the surface of the cloth, detected markers can be matched to a reference triangle mesh. Results show that point correspondences can then be refined using a Laplacian deformation framework to fit to unlabeled markers.

The proposed capture system has been evaluated on square pieces of cloth of a single fabric. Future work will focus on capturing and reconstructing a range of different garments with varying material properties. Capturing full garments, will introduce new a set of challenges, including seams, occlusions and self-shadowing.

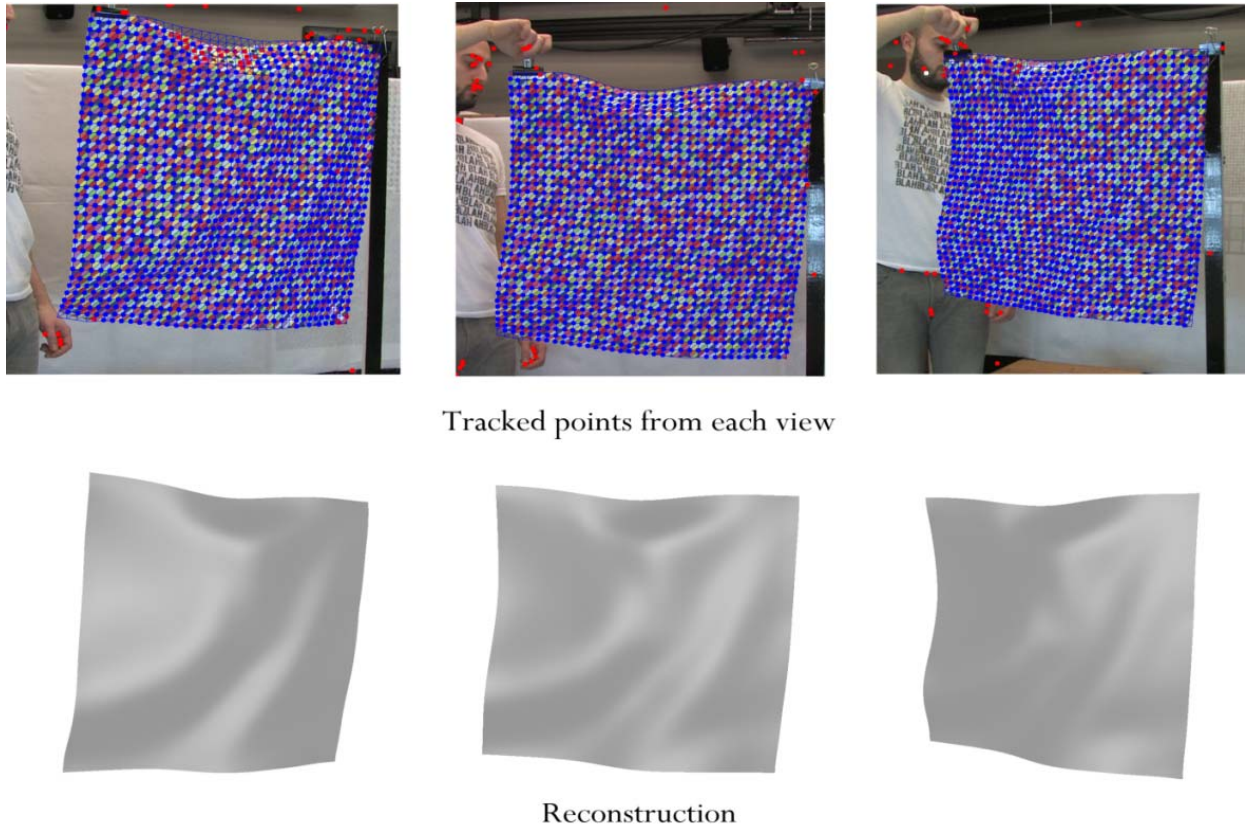


Fig. 7. Reconstruction results.

10. Acknowledgments

Current research project was supported by an EPSRC grant entitled: Fashion garment design, e-tailing and manufacturing with zero prototyping (Project reference: EP/M506801/1).

References

- [1] Guskov, I. Efficient tracking of regular patterns on non-rigid geometry. Object recognition supported by user interaction for service robots 2 (2002).
- [2] Guskov, I., Klibanov, S., and Benjamin, B. Trackable Surfaces. In Eurographics/SIGGRAPH Symposium on Computer Animation (2003).
- [3] Pritchard, D., and Heidrich, W. Cloth Motion Capture. In Computer Graphics Forum (2003), vol. 22, pp. 263271.
- [4] Scholz, V., Stich, T., Keckeisen, M., Wacker, M., and Magnor, M. Garment motion capture using color-coded patterns. Computer Graphics Forum 24, 3 (2005), 439-447.
- [5] White, R., Crane, K., and Forsyth, D. A. Capturing and animating occluded cloth, 2007.
- [6] Sorkine, O. Laplacian Mesh Processing. Processing, 84(Section 4):53-70, 2005.
- [7] Hartley, Richard I., and Peter Sturm. "Triangulation." Computer vision and image understanding 68.2 (1997): 146-157.
- [8] Lowe, David G. "Object recognition from local scale-invariant features." Computer vision, 1999. The proceedings of the seventh IEEE international conference on. Vol. 2. IEEE, 1999.