A DCV & USIR Method implemented for Reconstructed of Un-calibrated captured MVS Images

Gorikapudi Usha Rani [1] G. Madhusudhana Rao [2] P.Jayababu[3]
[1]M.Tech scholar, [2]ECE HOD, [3]Asst. Prof.,
Nannapananeni Venkatrao College of Engineering & Technology, Tenali.

Abstract **- This paper presents a Detail-Storing and Content-aware difference (DCV) multi-view stereo method and a novel algorithm for un-calibrated stereo image pair rectification under the constraint of geometric distortion, called USIR is presented in this work. Although it is straightforward to define a rectifying transformation given the epipolar geometry. Preserve boundaries details of the reconstructed surface and builds a connection between guided image filtering and image registration. To obtain rectified images with reduced geometric distortions while maintaining a neutralized rectification error. We analyse the homographs by considering the effects of various kinds of geometric distortions. Next, we propose several geometric measures and incorporate them into a new cost function for description optimization. A content-aware Lp-minimization algorithm by adaptively estimating the p-value and regularization parameters based on the current input . Finally, we propose a constrained adaptive optimization scheme to allow a balanced performance between the rectification error and the geometric error. Extensive experimental results are provided to demonstrate the data Storing and rectifying of image by DCV&USIR method.**

Index Terms **- Projective rectification, homographs, epipolar geometry, fundamental matrix, geometric distortion, constrained optimization, Multi-view stereo,L_0 Minimization.**

Introduction:

The performance of existing MVS methods is limited due to factors such as violation of the Lambertian reflectance model, inaccurate camera calibration, lack of textures on the object, and false matches. Therefore, noises are inevitable for the reconstructed 3D surface, resulting in degraded accuracy and visually unpleasant artefacts. A number of methods, e.g., weighted minimal surface models have been proposed to suppress noises. However, this line of methods usually impose isotropic smoothness prior on 3D models, and tend to over-smooth Out-line(Boundries) details and sharp features. To overcome these limitations, various methods have been developed to suppress noise while Storing sharp features.

(DCV) method for MVS.

An inter-image similarity measure is proposed to preserve Out-line (Boundaries) details of the reconstructed surface. The proposed similarity measure also builds a connection between guided image filtering [34] and image registration, making our measure have promising edge-Storing performance.A content-aware Lp-minimization algorithm is proposed for mesh denoising. By adaptively estimating a suitable p value and regularization parameters, our algorithm works very well in mesh smoothing while Storing sharp features.

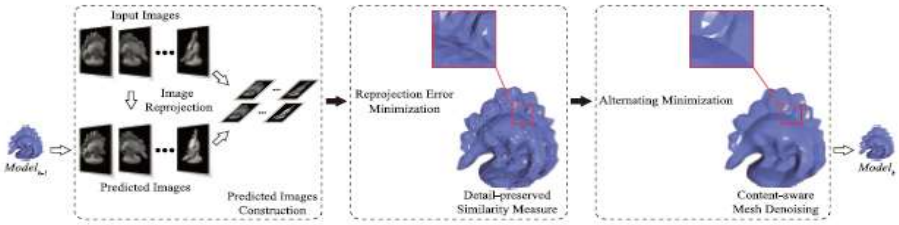


Fig. 1: Overview of the proposed DCV method.

In this work, we propose a novel rectification algorithm for uncalibrated stereo images, which demands no prior knowledge of camera parameters. Although quite a few methods are proposed to reduce unwanted warping distortions in rectified images with different homography parameterization schemes, there is no clear winner among them. Additionally, only two geometric measures (namely, orthogonality and the aspectratio) are used as geometric distortion criteria while they are not sufficient in characterizing the subjective quality of all rectified images. Here, we analyze the effect of various geometric distortions on the quality of rectified images comprehensively, and take them into account in algorithmic design. The proposed USIR-algorithm minimizes the rectification error while keeping errors of various geometric distortion types below a certain level.

The main contributions of this work are summarized below.

An uncalibrated stereo rectification algorithm is proposed to minimize the rectification error with constrained geometric distortions. A variety of geometric distortions such as the aspect-ratio, rotation, skewness and scale-variance are introduced and incorporated in the new cost function, then it is minimized by our novel optimization scheme. A parameterization scheme for rectifying transformation is developed. The parameters include the focal length difference between two cameras, the vertical

displacement between optical centers, etc. This new scheme helps reduce the rectification error by adding more degrees of freedom to the previous Euclidean model .

We provide a synthetic database that contains six geometric distortion types frequently observed in uncalibrated stereo image pairs. It allows a systematic way to analyze the effect of geometric distortions and parameterize the rectifying transformations accordingly. We also provide a real world stereo database with various indoor and outdoor scenes of full-HD resolution. The performance of several algorithms can be easily evaluated with these two databases

DCV model:

DCV model consists of two terms, i.e., data fidelity Eim and surface regularization Ereg. The energy functional of our model can be formulated as:

E(S) = Eim(S) + λEreg(S)

where S denotes the reconstructed surface of the object, and λ is the trade-off parameter. Note that Eim usually is differentiable while Ereg is non-smooth.

The model can be solved by extending the proximal gradient algorithm , which iteratively performs the following two steps.

Step 1. Gradient Descent.

Given the current estimate S k , the gradient descent algorithm is adopted to minimize the data fidelity term

Eim: S k+0.5 = S k+0.5 − η∂Eim(S)/∂S,

where η is the stepsize.

Step 2. Surface Denoising.

Given S k+0.5 , the reconstructed surface S is further refined by solving the following mesh denoising problem:

S k+1 = arg

min S 1 2 kS − S k+0.5 k 2 + ληEreg(S).

Given the nonsmooth convex function Ereg and the smooth convex function Eim with Lipschitz constant L, when the stepsize η ≤ 1/L and the surface denoising problem has the global solution, the algorithm can converge to the global optimum .

For our case, even Ereg is nonconvex, our algorithm empirically converges to a satisfactory solution. In this work, we propose a detail-Storing similarity measure for Step 1 and propose a content-aware mesh denoising algorithm for Step 2.

UNCALIBRATED RECTIFICATION for MVS

The pinhole camera model consists of optical centre C, image plane R, object point W, and image point M that is the intersection of R and the line containing C and W. The focal length is the distance between C and R, and the optical axis is the line that is orthogonal to R and contains C, where its intersection with R is the principal point. Let w and m be the coordinates of W and M, respectively. They are related by a perspective projection matrix P of dimension 3 × 4 as

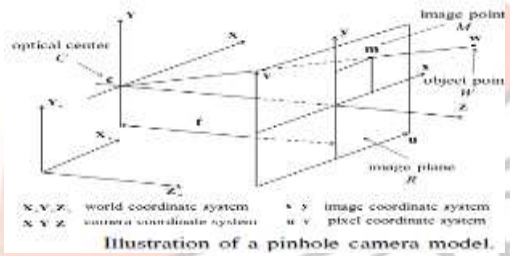


Illustration of a pinhole camera model.

$$\mathbf{m} = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \simeq \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \mathbf{P}\mathbf{w},$$

where 'indicates the equal up to scale. Matrix P can be decomposed into

$$\mathbf{P} = \mathbf{K}\,[\mathbf{R} \mid \mathbf{t}]\,,$$

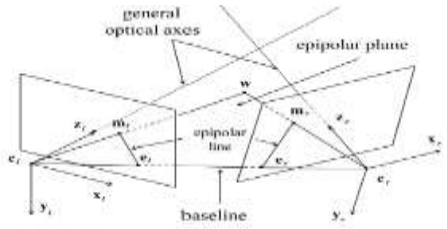
where K and [R | t] are called the camera intrinsic matrix and the camera extrinsic matrix, respectively. Matrix K is in form of

$$\mathbf{K} = \begin{bmatrix} \alpha_u & \gamma & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix},$$

where αu = suf and αv = svf are focal lengths in the u and v axes, respectively (f is the physical focal length of the camera in the unit of millimetres while su and sv are the scale factors), and (u0, v0) are the coordinates of the principal point, γ is the skew factor when the u and the v axes of the model are not orthogonal. For simplicity, it is often assumed that the horizontal and vertical focal lengths are the same and there is no skew between u and v axes. Thus, we have

$$\mathbf{K} = \begin{bmatrix} \alpha & 0 & \frac{w}{2} \\ 0 & \alpha & \frac{h}{2} \\ 0 & 0 & 1 \end{bmatrix},$$

where w and h are the width and the height of the image, respectively. The camera extrinsic matrix of dimension 3 × 4 is concerned with camera's position and orientation. It consists of two parts: a rotation matrix R of dimension 3 × 3 and a displacement vector t of dimension 3 × 1. The plane that contains optical center C and is parallel to the image plane is the focal plane. According to [14], the Cartesian coordinates ˜c of C is given by



: The epipolar geometry of a pair of stereo images.

$$\tilde{\mathbf{c}} = \mathbf{R}^{-1}\mathbf{t}.$$

Then, any optical ray that passes through M and C can be represented by the set of points w:

$$\tilde{\mathbf{w}} = \tilde{\mathbf{c}} + \alpha \mathbf{R}^{-1}\mathbf{K}^{-1}\mathbf{m},$$

$$\mathbf{m}_l^T \mathbf{F} \mathbf{m}_r = \mathbf{m}_r^T \mathbf{F}^T \mathbf{m}_l = 0,$$

Where α is a constant. Next, we consider two stereo pinhole cameras as shown in Fig. Let ml and mr be point correspondences that are the projections of the same 3D object point w on images Il and Ir, respectively. el and er are called epipoles that are intersection points of the baseline with the left and right image planes. The plane containing the baseline and object point w is called the epipolar plane, and the intersection lines between the epipolar plane and each of the two image planes are epipolar lines. The intrinsic projective geometry between the corresponding points in the left and right images can be described by the epipolar constraint as where F is the fundamental matrix, which is a 3 × 3 matrix with rank 2, and 0 = [0 0 0]T is a zero column vector. The epipole, which is the null space of F, satisfies the following condition:

$$\mathbf{F}\mathbf{e}_l = \mathbf{F}^T\mathbf{e}_r = \mathbf{0}.$$

Fundamental matrix F maps a point, ml , in one image to the corresponding epipolar line, Fml = lr, in the other image, upon which the corresponding point mr should lie. Generally, all epipolar lines are not in parallel with each other and passing through the epipole (namely, l T l el = lr T er = 0). Image rectification as shown in Fig. 3 is the process of converting the epipolar geometry of a given stereo image pair into a canonical form that satisfies two conditions: 1) all epipolar lines are parallel to the baseline, 2) there is no vertical disparity between the corresponding epipolar lines in both images. This can be done by applying homographies to each of image planes or, equivalently, mapping the epipoles to a point at infinity as e∞ = [1 0 0]T . Especially, the fundamental matrix of a pair of rectified images can be expressed in form of

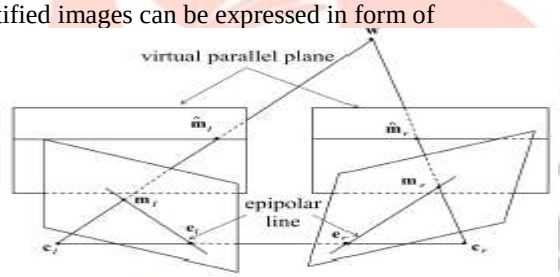


Illustration of image rectification.

$$\mathbf{F}_\infty = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Let Hl and Hr be two rectifying homographies of the left and right images, respectively, and (ˆml , ˆmr) be the corresponding points in the rectified images. Then, we have

$$\hat{\mathbf{m}}_\mathbf{l} = \mathbf{H}_l\mathbf{m}_l, \qquad \hat{\mathbf{m}}_\mathbf{r} = \mathbf{H}_r\mathbf{m}_r.$$

$$\hat{\mathbf{m}}_\mathbf{l}^T \mathbf{F}_\infty \hat{\mathbf{m}}_\mathbf{r} = \mathbf{0}.$$

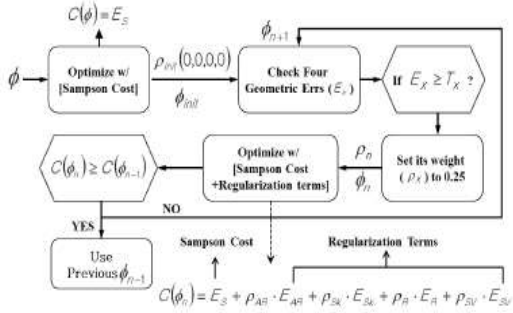
By incorporating above Eq, we obtain

$$\mathbf{m}_l^T \mathbf{H}_l^T \mathbf{F}_\infty \mathbf{H}_r \mathbf{m}_r = \mathbf{0}.$$

As a result, the fundamental matrix of the original stereo image pair can be specified as F = HT l F∞Hr. The fundamental matrix is used to calculate the rectification error in the process of parameter optimization. Thus, the way to parameterize Hl and Hr is critical to the generation of good rectified images.

PROPOSED USIR:

Iterative Optimization



The block diagram of the proposed iterative optimization procedure.

Based on this design, the cost is a dynamically changing function that always contains the rectification error. A geometric distortion term will be included in the cost function (or "being turned-on") only when it is larger than a threshold. Otherwise, it is turned off. To solve this problem, we propose an iterative optimization procedure as shown in Fig. It consists of the following steps.

We begin with a cost function that contains the Sampson error (Es) only. The optimization procedure offers the initial set of rectification parameters, which is denoted by φinit.

We update four weight ρ = (ρAR, ρSk, ρR, ρSV) from φinit, which is initially set to ρinit = (0, 0, 0, 0).

Under the current set of rectification parameters, φ, and the current set of weights, ρ, we solve the optimization problem with respect to Eq. Then, both rectificatioin parameters and weights are updated accordingly. Step 3 is iterated until the the cost of the current round is greater than or equal to that of the previous round. Mathematically, if C(φn) ≥ C(φn−1), we choose φn−1 as the converged rectification parameters.

When we compare the current cost C(φn) with the cost in the previous round C(φn−1) in Step 3, the number of geometric terms in Eq. may change. If this occurs, we should compensate it for fair comparison.

That is, the cost should be normalized by the sum of weight via

Cnormalized(φn) = C(φn)/(1 + P X ρX).

The choice of a proper threshold value for each geometric error is important in reaching balanced performance. In this work, we set threshold values of geometric errors to the following:

$0.8 \le EAR \le 1.2$, $ESk \le 5^\circ$,

$0.8 \le ESV \le 1.2$,

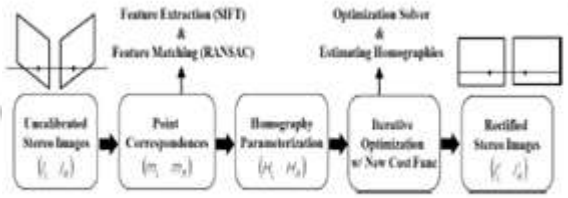
$|ER| \le 30^\circ$

Furthermore, we normalize four geometric errors by considering their value range.

The normalizing factors are:

NAR = 1.5, NSk = 6.5, NR = 18.5, NSV = 2.5,

which can be absorbed in their respective weight; i.e. the new weight becomes ρX = 0.25/NX when the term is on. Last, the minimization of the cost function is carried out using the nonlinear least square method, Trust Region, starting with all unknown variables set to zero.



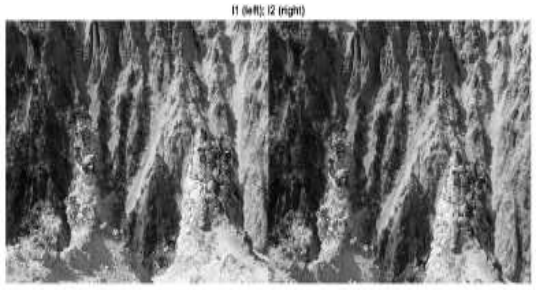
The block-diagram of the proposed USIR system.

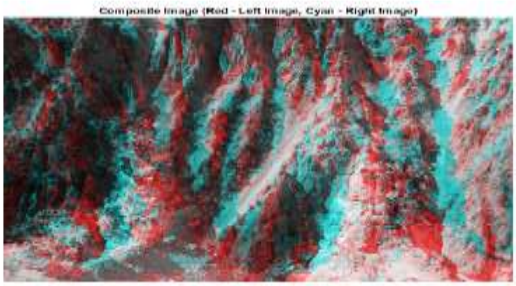
The block-diagram of the proposed USR-CGD system is shown in Fig. This system is fully automatic since there is no need to estimate the fundamental matrix. To establish the point correspondence between the left and right images, we extract the SIFT feature and find the initial putative matching points. We also apply RANSAC to remove outliers. It is noteworthy that the number of the correspondences strongly affects the rectification performance because the homography is estimated based on their errors, and the optimal number varies with the image resolution. A special case of the USR-CGD algorithm is to turn off all geometric distortion terms, which is called the USR algorithm.

Simulation Results

Step 1: Read Stereo Image Pair

Read in two color images of the same scene, which were taken from different positions. Then, convert them to grayscale. Colors are not required for the matching process. Display both images side by side. Then, display a color composite demonstrating the pixel-wise differences between the images.



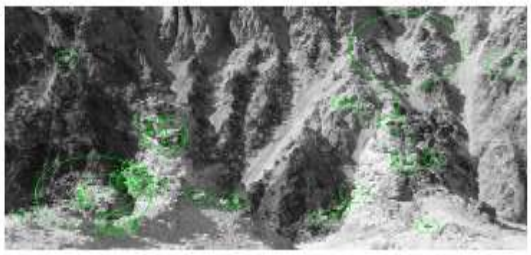


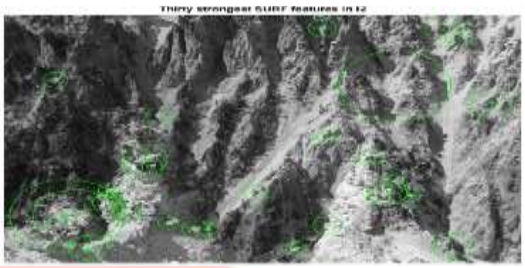
There is an obvious offset between the images in orientation and position. The goal of rectification is to transform the images, aligning them such that corresponding points will appear on the same rows in both images.

Step 2: Collect Interest Points from Each Image

The rectification process requires a set of point correspondences between the two images. To generate these correspondences, you will collect points of interest from both images, and then choose Potential matches between them. Use detectSURFFeatures to find blob-like features in both images.

Visualize the location and scale of the thirty strongest SURF features in I1 and I2. Notice that not all of the detected features can be matched because they were either not detected in both images or because some of them were not present in one of the images due to camera motion.





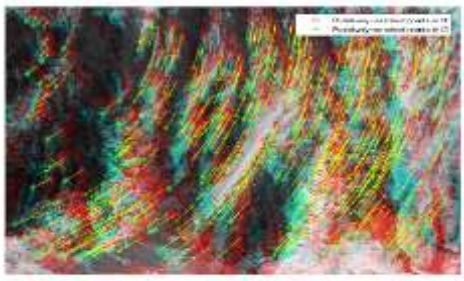
Step 3: Find Putative Point Correspondences

Use the extractFeatures and matchFeatures functions to find putative point correspondences. For each blob, compute the SURF feature vectors (descriptors).

Use the sum of absolute differences (SAD) metric to determine indices of matching features.

Retrieve locations of matched points for each image

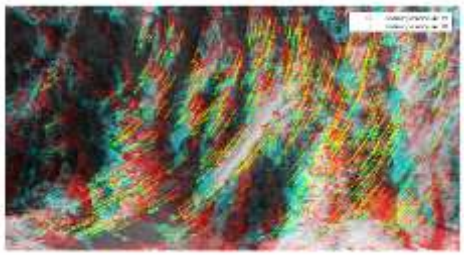
Show matching points on top of the composite image, which combines stereo images. Notice that most of the matches are correct, but there are still some outliers.



Step 4: Remove Outliers Using Epipolar Constraint

The correctly matched points must satisfy epipolar constraints. This means that a point must lie on the epipolar line determined by its corresponding point. You will use the

estimateFundamentalMatrix function to compute the fundamental matrix and find the inliers that meet the epipolar constraint.

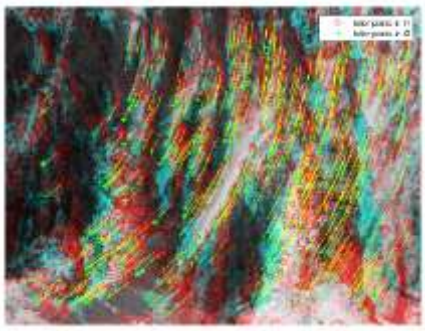
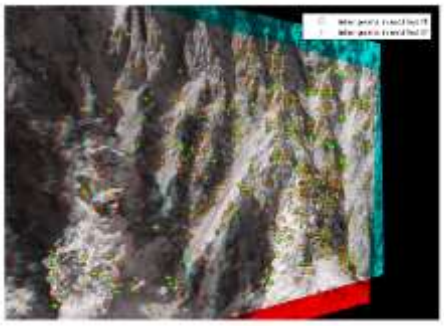


Step 5: Rectify Images

Use the estimateUncalibratedRectification function to compute the rectification transformations. These can be used to transform the images, such that the corresponding points will appear on

the same rows.

Rectify the images using projective transformations, tform1 and tform2. Show a color composite of the rectified images demonstrating point correspondences.

Crop the overlapping area of the rectified images. You can use red-cyan stereo glasses to see the 3D effect.

Step 6: Generalize the Rectification Process

The parameters used in the above steps have been set to fit the two particular stereo images. To process other images, you can use the cvexRectifyStereoImages function, which contains

additional logic to automatically adjust the rectification parameters. The image below shows the result of processing a pair of images using this function.

CONCLUSION AND FUTURE WORK

In this paper, we proposed a detail-Storing and contentaware variational (DCV) method and a new rectification algorithm, called USIR, was proposed for uncalibrated stereo images in this work. image filtering with image registration, a novel similarity measure was proposed to protect the boundries details in reconstruction. It adopts a generalized homographs model to reduce the rectification error like reprojection error and incorporates several practical geometric distortions in the cost function as regularization, neutralization transform terms, which prevent severe point of view distortions in rectified images. It was shown by experimental results that the proposed USIR-algorithm outperforms existing algorithms in both objective and subjective quality measures. In the future, we would like to study the stereo matching problem for depth estimation based on the current work on uncalibrated stero image rectification.

References:

[1] S. Seitz, B. Curless, J. Diebel, D. Scharstein and R. Szeliski, "A comparison and evaluation of multi-view stereo reconstruction algorithms," in Proc. CVPR, pp. 519-526, 2006. [2] C. Strecha, W. von Hansen, L. Van Gool, P. Fua and U. Thoennessen, "On benchmarking camera calibration and multiview stereo for high resolution imagery," in Proc. CVPR, pp. 1-8, 2008.

[3] P. Tanskanen, K. Kolev, L. Meier, F. Camposeco, O. Saurer and M. Pollefeys, "Live Metric 3D Reconstruction on Mobile Phones," in Proc. ICCV, pp. 1-8, 2013.

[4] I. Kostrikov, E. Horbert and B. Leibe, "Probabilistic Labeling Cost for High-Accuracy Multi-View Reconstruction," in Proc. CVPR, 2014.

[5] A. Delaunoy and M. Pollefeys, "Photometric Bundle Adjustment for Dense Multi-View 3D Modeling," in Proc. CVPR, 2014.

[6] M. Meyer, M. Desbrun, P. Schrder and A. H. Barr, "Discrete differentialgeometry operators for triangulated 2-manifolds," Visualization and Mathematics III,Part I, pp 35-57,

[7] Y. Duan, L. Yang, H. Qin and D. Samaras, "Shape Reconstruction from 3D and 2D Data Using PDE-Based Deformable Surfaces," in Proc. ECCV, pp 238-251,

[8] G. Vogiatzis, P. Torr, S.M. Seitz and R. Cipolla, "Reconstructing relief surfaces," in Proc. BMVC, 2004

[9] J. Isidoro and S. Sclaroff, "Stochastic Refinement of the Visual Hull to Satisfy Photometric and Silhouette Consistency Constraints," in Proc. ICCV, pp. 1335-1342, 2003

[10] D. Bradley, T. Popa, A. Sheffer, W. Heidrich and T. Boubekeur, "Markerless garment capture," ACM Trans. Graphics, vol. 27, no. 3, pp. 538-551, 2008.

[11] P. Yan, S.M. Khan and M. Shah, "3d model based object class detection in an arbitrary view," in Proc. ICCV, pp. 1-6, 2006.

[12] A. Kushal and J. Ponce, "Modeling 3d objects from stereo views and recognizing them in photographs," in Proc.