ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

Particle Filtering Framework and Occlusion Handling Using Region Based Tracking Technique for 2-D and 3-D Image Sequences

J.Cynthia Esther Rani[1], R.Sathya[2], R.D. Sathiya[3]

M.Tech, School of Computing, SASTRA University, Tanjore, Tamilnadu, India[1, 2]

Assistant Professor, School of Computing, SASTRA University, Tanjore, Tamilnadu, India[3]

ABSTRACT: A method is introduced to address the estimation of pose for 2D-3D images. This approach involves the tracking of 2D images in 3D spaces which results in the estimation of pose with the continuous process of particle filtering and handling the scheme in the presence of occlusion. In order to maintain the track and to detect the exact region, a unique method Region Based Tracking method is used in this paper. It also involves the degree of dependencies between predictions and measurements model which determines the position and estimation of the object sustaining particle filters and handling occlusions. Particle filtering can propagate via autoregressive model for tracking 2D images but also estimating a 3D pose. Finally, this method is considered to be more effective in both robustness and speed compared to similar video tracking and pose estimation methods. Thus the conversion of pose estimation from 2D Silhouette curve to 3D Occluding Curve is finally illuminated in video.

KEYWORDS- Pose estimation, Particle filtering, Occlusion handling

I. INTRODUCTION

Numerous algorithms have been proposed to locate and track objects of interest in recent years. 2D-3D pose estimation aims to determine the pose of a 3D object relative to a calibrated camera from a unique or a collection of 2D images. By knowing the mapping between the world coordinates and image coordinates from the camera calibration matrix, and after establishing correspondences between 2D features in the image and their 3D counterparts on the model, it is then possible to solve the pose transformation.

This problem consists of determining object's pose in 3D relative to a calibrated camera, from a 2D image sequence the relative pose can be described in terms of trajectories on a transformation matrix [1]. The MonteCarlo based sampling method is to estimate a 3D transformation matrix for 2D object tracking and pose estimation. This allows the tracker to capture various aspects of an object with respect to its dynamic motion. The methods based on learning a collection of 2D shape priors have difficulties in completely describing them.

The Bayesian framework, all relevant information about $\{x_0,x_1,...x_k\}$ given observations up to and including time k can be obtained from the posterior distribution $p(x_0,x_1,...x_k / y_0,y_1,...y_k)$. In many applications we are interested in estimating recursively in time this distribution, and particularly one of its marginals, the so-called filtering distribution $p(X_k / y_0,y_1,...y_k)$. Given the filtering distribution one can then routinely proceed to filtered point estimates such as the posterior mode or mean of the state. This problem is known as the Bayesian filtering problem or the optimal filtering problem. Practical applications include target tracking [2], blind deconvolution of digital communications channels [3], estimation of stochastic volatility [4] and digital enhancement of speech and audio signals. The recent literatures describes with 3D pose estimation is very large and a complete survey is beyond the scope of this paper. However, most methods can be

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

distinguished by the type of local image features used to establish correspondences, such as points [5], lines or segments [6, 7], multi-part curve segments [8], or complete contours [9, 10]. Finally, we show how to extend these methods to compute the prediction and measurement model in the fixed-interval.

II. RELATED WORK

In this section, we review some basic notions from the theory of level set evolutions. Analyzation of pose recovers body poses and motion parameters excepting position coordinates from static images or video sequences. There exist two problems under such cases of the estimation of pose in regardance to its views: a) camera pose estimation and b) object pose estimation. Many algorithms are applied for the solution of both the problems. Ohayon formulated the recovery of head pose as a camera pose estimation problem. Recently used techniques undergo two categories:1)feature based methods and 2)appearance-based methods.

Appearance-based methods rely over the premise of training set of images for the estimation of object pose. Usually, appearance based methods depend on training set of images only. Thus, a small change in the image leads to the repetition of training of the image. Image normalization is used for the appearance models in order to make the image more robust to illumination changes.

Feature-based methods rely over the premise of corresponding feature points from different images. The sufficed methods are of two types: 1) Linear methods2) Nonlinear methods. Those are the solutions with redundant data and they are more robust to noise. As linear methods require less computation, it requires lack of accuracy and robustness.

A region-based approach is subtended to continuously drive the estimation of pose . The pose estimation algorithm fore comes two main categories. Firstly, finding the similarities between the poses with the intention of prediction and estimation problem. Secondly, image features occurring as local features does not rely upon reliable and robustness concepts. Furthermore, simplifying assumptions usually need to be made on the class of shapes that a 2D-3D pose estimation technique can be handled. Many techniques which are limited to simple shapes can be described using geometric initiatives such as corners, lines, circles or cylinders. Recent work focused on free-form objects, which features a manageable parametric description as in [11]. However, even this type of algebraic approaches can become unmanageable for objects of complex shape. Our approach can deal with rigid object of complex shape, represented by a 3D set [12] or a 3D points.

Variation approaches that address the problem of structure from motion and reconstruction from multiple cameras, also from surveillance cameras. We present a method to reconstruct the 3D shape of an object from multiple 2D views obtained from calibrated cameras as well as surveillance cameras from a certain place or the estimated distance from the camera to the variated pose. The present contribution aims at: Given the 3D model of an object, perform the segmentation of 2D images and recover the 3D pose of the object relative to a unique camera. Thus, the framework is adapted and employed in the specific context of segmenting 2D images from a unique camera, using a 3D model.

III. PRELIMINARIES

3.1. Particle filtering framework

Particle filtering methods are used for characterization in video sequences based Baye's rules:

$p(x_k/z_{1:k})\,\alpha\, p(z_k/x_k)p(x_k/z_{1:k-1})$

where integration is confined to the following equation and the resultant is,

$p(x_k/z_{1:k-1})=p(x_k/x_{k-1})p(x_{k-1}/z_{1:k-1})dx_{k-1}$

Here, x and z are used to denote the state and observed data respectively.
For this method a significant state vector is used

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

$$x = (O_x, O_y, a, b, r_x, r_y, r_z) \qquad (1)$$

where (O_x, O_y) represent the position coordinates in the center of ellipse, (a,b) denotes the size of the ellipse (i.e).,short or long and (r_x, r_y, r_z) represent the rotational axis around the axis respectively. The Euler's theorem confining where the rotational matrices to its corresponding angles. The probability function are is denoted using Monte Carlo sampling, where the particle filters are considered to be of great usage.

The approximation is given as,

$$p(x_k|z_{1:k}) \approx \sum_{i=1}^{N} \omega_k^i \delta(x_k - x_k^i) \qquad (2)$$

The unnormalised weights are determined as,

$$\omega_k^i \propto \omega_{k-1}^i \frac{p(z_k|x_k^i)p(x_k^i|x_{k-1}^i)}{q(x_k^i|x_{k-1}^i, z_k)} \qquad (3)$$

The discrete weight approximation is given as,

$$\hat{x}_k = \sum_{i=0}^{n} w_k^i x_k^i \qquad (4)$$

By this method a less expensive computation is effected.

IV. PROPOSED METHOD

4.1. Initialization

For the purpose of initialisation of system, a known pose and the estimation of pose is represented to the target of the first image. This can be made manually or automatically. In our simulations we mark points directly around the object.Thus, the nature of objects based on their characteristics are predominantly defined.

4.2 Perspective Projection from 2-D to 3-D Sequences

From the matched points between two successive images, the rotation and the translation vector is determined which is considered to be a problematic feature. Thus camera pinpoint model is specified here. The distance to the camera from the object is relied here.

The positions before and after motion is defined as:

$$\begin{pmatrix} x_{t+1}^j \\ y_{t+1}^j \\ z_{t+1}^j \end{pmatrix} = R_{3x3} \begin{pmatrix} x_t^j \\ y_t^j \\ z_t^j \end{pmatrix} + T_{3x1} \qquad (5)$$

where,R_{3x3} is a 3x1 rotation matrix,T_{3x1} is a 3x1 translation vector. Thus, rotation and translation matrix is to be calculated.

As the perspective projection model is considered, the points of projection in the image with respect to (u_t^j, v_t^j) and (u_{t+1}^j, v_{t+1}^j) are defined as,

$$\frac{x_k^j}{z_k^j} = \frac{u_k^j}{f}, \frac{y_k^j}{z_k^j} = \frac{v_k^j}{f}, k = t, t+1 \qquad (6)$$

where f is the focal length.

Integrating both the equations, we get,

$$\begin{pmatrix} u_{t+1}^j \\ v_{t+1}^j \\ f \end{pmatrix} = \frac{z_t^j}{z_{t+1}^j} R_{3x3} \begin{pmatrix} u_t^j \\ v_t^j \\ f \end{pmatrix} + T'_{3x1} \qquad (7)$$

where $T'_{3x1} = T_{3x1}$.Also, the ratio $\frac{z_t^j}{z_{t+1}^j}$ is unknown.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

Here, some depth information is lost on projection of images from 3-D to 2-D images. Thus, we consider, $\frac{z_t^j}{z_{t+1}^j} = 1$. This is termed to be an unusual sequence for which the distance between the camera and the object is too large. Also, the distance between the frames are too small. Thereby, the accuracy seems to be degraded to a greater extend over here.

In other case, if the distance over the frames is considered to be in unity, the resultant equation is simplified from the equation (3) as,

$$\begin{pmatrix} u_{t+1}^j \\ v_{t+1}^j \\ f \end{pmatrix} = R_{3\times3} \begin{pmatrix} u_t^j \\ v_t^j \\ f \end{pmatrix} + T'_{3\times1} \qquad (8)$$

On rearranging the first two rows of the above equation, we obtain,

$$\begin{pmatrix} u_{t+1}^j \\ v_{t+1}^j \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} \begin{pmatrix} u_t^j \\ v_t^j \end{pmatrix} + \begin{pmatrix} l_x \\ l_y \end{pmatrix} \qquad (9)$$

where, l_x and l_y are new translation parameters.

4.3 Occlusion Handling

Occlusions always hinders the tracking of a sequence and estimating the pose of an object. Occlusion detection is an effective and considerable task before performing occlusion handling. In general, occlusion is detected using the relative change of the object's size compared to the average object's size as well as the distance between the object and an obstacle. However, this method may give ambiguous results if an object's size changes due to camera zooming or in surveillance camera. Here, we propose a histogram based occlusion detection technique, which is performed by checking the variation of the histogram of the object during tracking. For color-based imagery, a histogram is calculated by the mean of color intensities with the presence of basic colors or color model such as RGB. Here, the RGB color space is normalized to remove the effect of variations in intensity. The evaluation of the histogram change is achieved by computing the Bhattacharyya coefficient between two appearance models of the silhouette curve also of the template model.

V. EXPERIMENTAL RESULTS

Extensive synthetic and real sequences of different rigid objects were used to demonstrate the robustness of the proposed method to noise, cluttered environments. 3Dmodels used in this paper consist ofan elephant as shown in Figs. 1 and 2. In this section, we provide qualitative and quantitative results of various tracking scenarios including a comparison to the algorithms. In particular, in the quantitative experiments, two quantitative results regarding the robustness to noise and occlusion of the proposed method are provided. We also should note that because code of other joint 2D–3D pose estimation algorithms were not readily available, our experiments are focused on highlighting the advantages and limitations of exploiting dynamics in visual tracking. However, before doing so, we briefly mention some numerical details associated with the experiments performed. Implementation details: In these experiments, the parameters used were held fixed across all sequences.

5.1. Tracking in noisy and cluttered environments

In this subsection, we show quantitative and qualitative results regarding the robustness of the proposed method to noise on synthetic data in fig.1. In generating the synthetic data, we first construct a basic elephant sequence, and then add several noise levels of Gaussian noise whose variance ranges from $\sigma^2 n$=1% to $\sigma^2 n = 100\%$. The translation and rotation parameters linearly increase and decrease throughout the sequences of 200 frames to produce a large variation for the

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

aspect of the object. The size of the sequence images ranges over 242 - 322. To quantitatively evaluate the tracking results, absolute errors are computed for both translation and rotation over each level of noise sequences.

$$\% \text{ - absolute error} = \frac{||Vmeasured - Vtruth||}{||Vtruth||} X\ 100 \qquad (15)$$

where Vmeasured and Vtruth are measured and ground-truth of translation and rotation vectors, respectively. This is due to the probability of replica of local minima.

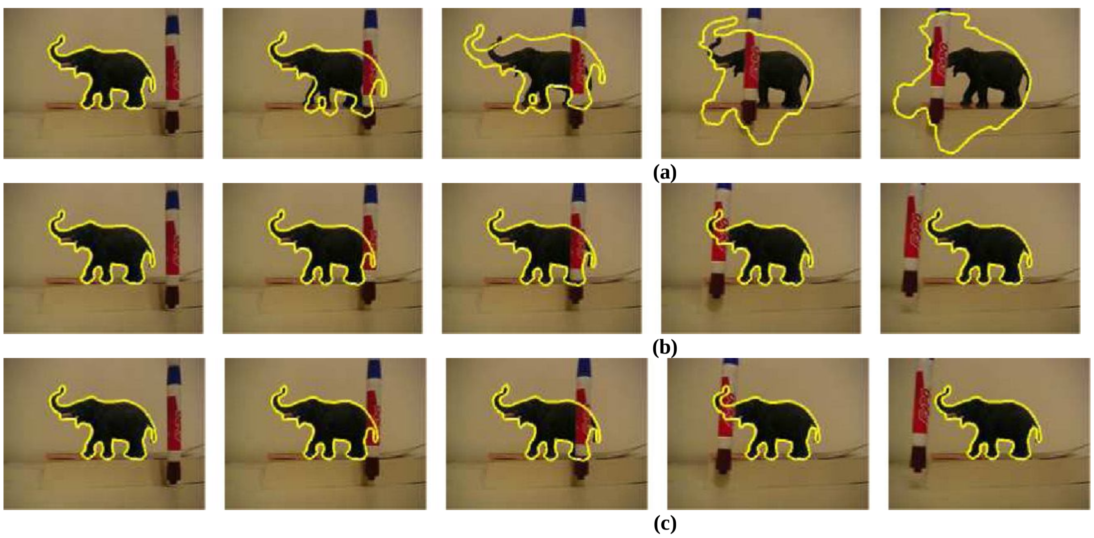
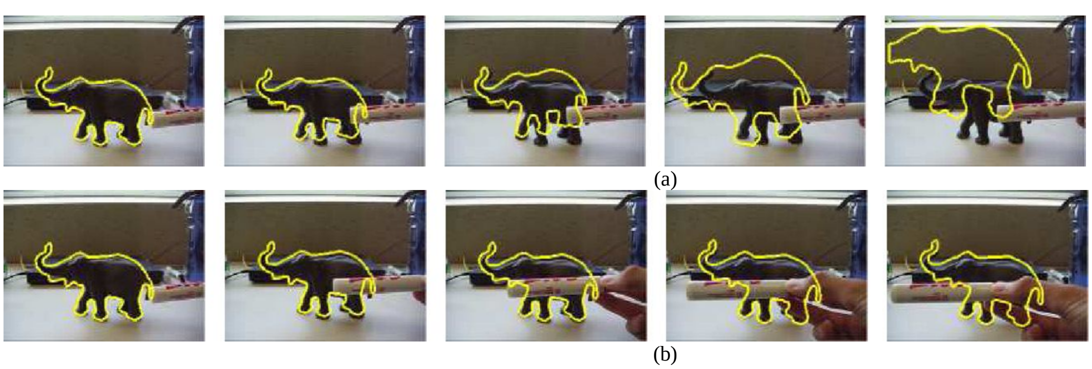


Fig.1. Elephant sequence with occlusion in a cluttered environment. Tracking results: (**a**) using particle filtering method, (**b**) using occlusion method and (**c**) using the proposed method



ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

(c)

Fig. 2. Elephant sequence II with occlusion in a cluttered environment. Tracking results: (**a**) using particle filtering method, (**b**) using occlusion method and (**c**) using the proposed method

5.2. Tracking in the presence of occlusion

In contrast to the previous sequences, the scenarios in this subsection not only include dynamic changes of an object's pose, but other objects too, which occlude the object of interest, and provide an added difficulty to the overall tracking problem. Moreover, this subsection demonstrates how the proposed method outperformed the approaches of [14,15] in dealing with occlusion handling. For the robustness of the proposed method in handling occlusions we generate a set of synthetic sequences, in which an obstacle bar exhibiting different levels of gray-scale intensity is added.

Figs. 1 and 2 show a red marker and a white marker that pass by the gray elephant from right to left in a cluttered background, respectively. Here, the algorithms in [14] works well with the occlusions that are similar in nature to that of the interested object. However, due to the fact that the occlusion contains a statistically different intensity from that of the object of interest, the methods in [14] is not able to maintain track as shown in Figs. 1 and 2. The movement of the marker acts as if it pushes or blows the silhouette curve off of the elephant. This is simply due to the statistical difference between the object and the occlusion. From a robust statistics point of view, the flow regarding the integration about the occluding curve excludes the possible points on the white marker because they are viewed as outliers. In turn, one cannot properly estimate the 3D pose or maintain track. However, if one exploits the underlying dynamics as done in the proposed algorithm, one achieves a more robust result. Compared to the sequences, Figs. 1 and 2 show sequences with moving objects in addition to static occlusions.

These results show that the approach of [15] is more vulnerable to occlusions when it tracks a moving object than a stationary object. This is because the work of [15] disregards control of predictions and measurements of the system, which was taken into account in the present occlusion handling scheme. In addition, in the proposed method, in contrast to the work of [15], the separately distributed samples of pose parameters and the embedded variational technique aid the tracker in finding the optimum in a filtering distribution.

Video Tracking

Considering the given video, the tracking of video depicts the monitoring of objects motion from the video sequences. Thus, the video tracking locates the object both in static and dynamic sequences. First the frames from the video are analysed. The existing frame is compared to the incoming frame from the video by the process of tracking. Finally, it locates the moving targets within the video frame, also considered here. The predominant feature of this approach is that, the coordinates regarding to the position of the image, with its rotational angles are considered. The camera parameters along with the head geometry from trained image set is obtained. Also an approach is presented for the tracking of head alone and estimating its pose with the particle filtering framework.

They also use training set of images. Thus, Gaussian and Gabor filters are used in this method, for the estimation of the pose of head. Particle filtering methods are used for characterization in video sequences. From the matched points between two successive images, the rotation and the translation vector is determined which is considered to be a problematic feature. Thus camera pinpoint model is specified here. The distance to the camera from the object is relied here. The Euler's theorem confining where the rotational matrices to its corresponding angles. This rule characterises objects in video sequences.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) **Vol.2, Special Issue 1, March 2014**

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

VI. CONCLUSION

A method to estimate tracking and parameters used for the determination of pose are estimated within the particle filtering framework. This approach can directly estimate 3-D rotation without the consideration and construction of model. Thus, the 2D visual tracking and 3-D pose estimation is obtained using particle filters. Occlusion handling scheme strengthens the entire dynamicity of the tracking of object. Degree of trust is enforced in order to ensure predictions and measurements. Finally, tracking performance is ensured even under several challenging environments with the presence of noise, clutter etc. Moreover it incorporates the particle filtering approach, for both the process of video tracking and the estimation of pose. The resulting algorithm provides the process of video tracking and also for pose estimation which specifies the location coordinates and the parameters involved in the pose. For the accumulation of errors and the recovery of the system, the incorporation of learning-based template is to be effected. Finally, tracking performance is ensured even under several challenging environments with the presence of noise, clutter etc.

REFERENCES

[1] Y. Ma, S. Soatto, J. Kosecka, and S. Sastry, An Invitation to 3D Vision, Spinger-Verlag, 2003.
[2] Gordon N.J., Salmond D.J., and Smith A.F.M. 1993. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. IEE-Proceedings-F 140: 107–113.
[3] Clapp T.C. and Godsill S.J. 1999. Fixed-lag smoothing using sequential importance sampling. In: Bernardo J.M., Berger J.O., Dawid A.P., and Smith A.F.M. (Eds.), Bayesian Statistics, Vol. 6, Oxford University Press, pp. 743–752.
[4] Pitt M.K. and Shephard N. 1999. Filtering via simulation: Auxiliary particle filters. Journal of the American Statistical Association 94: 590–599.
[5] Quan, L., Lan, Z.D.: Linear n-point camera pose determination. IEEE Transactions on Pattern Analysis and Machine Intelligence 21 (1999) 774–780
[6] Dhome, M., Richetin, M., Lapreste, J.T.: Determination of the attitude of 3d objects from a single perspective view. IEEE Trans. Pattern Anal. Mach. Intell. 11 (1989) 1265–1278
[7] Marchand, E., Bouthemy, P., Chaumette, F.: A 2d-3d model-based approach to real-time visual tracking. Image and Vision Computing 19 (2001) 941–955
[8] Zerroug, M., Nevatia, R.: Pose estimation of multi-part curved objects. In: ISCV '95: Proceedings of the International Symposium on Computer Vision. (1995) 431
[9] Rosenhahn, B., Perwass, C., Sommer, G.: Pose estimation of free-form contours. IJCV 62 (2005) 267–289.
[10] Drummond, T., Cipolla, R.: Real-time tracking of multiple articulated structures in multiple views. In: Proc. 6th European Conf.on Computer Vision, ECCV. (2000) 20–36
[11] Rosenhahn, B., Perwass, C., Sommer, G.: Pose estimation of free-form contours. IJCV 62 (2005) 267–289.
[12] Osher, S., Fedkiw, R.: Level Set Methods and Dynamic Implicit Surfaces. Springer Verlag (2003)
[13] Unal, G., Yezzi, A., Soatto, S., Slabaugh, G.: A variational approach to problems in calibration of multiple cameras. Trans. Pattern Analysis and Machine Intelligence 29 (2007) 1322–1338.
[14] S. Dambreville, R. Sandhu, A. Yezzi, A. Tannenbaum, A geometric approach to joint 2D region-based segmentation and 3D pose estimation using a 3D shape prior, SIAM Journal on Imaging Sciences 3 (1) (2010) 110–132.
[15] J. Lee, R. Sandhu, A. Tannenbaum, Monte Carlo sampling for visual pose tracking, in: IEEE International Conference on Image Processing, 2011, pp. 509–512.