N-View Human Silhouette Segmentation in Cluttered, Partially Changing Environments *,**

Tobias Feldmann[1], Björn Scheuermann[2],
Bodo Rosenhahn[2], and Annika Wörner[1]

[1] Karlsruhe Institute of Technology (KIT), Germany
{feldmann,woerner}@kit.edu
[2] Leibniz Universität Hannover, Germany
{scheuermann,rosenhahn}@tnt.uni-hannover.de

Abstract. The segmentation of foreground silhouettes of humans in camera images is a fundamental step in many computer vision and pattern recognition tasks. We present an approach which, based on color distributions, estimates the foreground by automatically integrating data driven 3d scene knowledge from multiple static views. These estimates are integrated into a level set approach to provide the final segmentation results. The advantage of the presented approach is that ambiguities based on color distributions of the fore- and background can be resolved in many cases utilizing the integration of implicitly extracted 3d scene knowledge and 2d boundary constraints. The presented approach is thereby able to automatically handle cluttered scenes as well as scenes with partially changing backgrounds and changing light conditions.

1 Introduction

The problem of segmenting foreground silhouettes of humans in camera images is a fundamental problem in computer vision. High quality silhouettes are an essential prerequisite for dense camera based 3d reconstruction or image based human pose estimation. Camera based dense 3d reconstruction of humans can, hence, be partitioned into three main blocks: Modality of image acquisition, foreground estimation and -separation and finally, dense 3d reconstruction.

Regarding the modality, image based human motion capture has been done monocular [1], with stereo [2], multi view [3,4,5] or multi view stereo setups [6].

The estimation of foreground from the image data can be realized by image differencing over time [7], by using color model coherence by integrating appropriate a priori knowledge of fore- and background [8,9] or by integrating 3d knowledge from stereo [2] or n-view reconstructions [4,5]. If colors are used for foreground segmentation, the approaches can be separated in simple per channel differencing approaches, codebooks models [10] and mixture models [11].

* This work was partially supported by a grant from the *Ministry of Science, Research and the Arts of Baden-Württemberg*.
** This work is partially funded by the German Research Foundation(RO 2497/6-1).

M. Goesele et al. (Eds.): DAGM 2010, LNCS 6376, pp. 363–372, 2010.

Based on the color models, the segmentation is usually performed based on probabilities [5] or energy minimization, e.g. level set approaches [12,13] or graph cuts [14]. The 2d information of multiple cameras can be fused by the voxel carving approach [3], by probabilistic data based fusion [15], probabilistic fusion with integration of a priori knowledge about the appearance of human silhouettes [1] or energy based formulations [4,6].

Most of the mentioned approaches for detailed 3d reconstruction focus on the segmentation and 3d reconstruction in artificial or laboratory like scenarios with homogenous but mostly disjunct colors of fore- and background. The approaches, which take into account clutter and occlusions, often need a manual initialization step [6,9,16]. We, thus, present an approach, which conjoins probabilistic 3d fusion and energy based level set approaches, which enables auto initialization and adaptivity to scenes with cluttered, moderately changing backgrounds. Based on our recordings with a calibrated multi camera setup in realistic, cluttered and partially changing environments, we can show that our approach is able to produce high quality foreground segmentation results of human silhouettes. Utilizing these silhouettes for dense 3d reconstructions gains convincing results in these difficult scenarios.

2 Segmentation by Probabilistic 3d Fusion

The segmentation via probabilistic 3d fusion proposed in [5] is based on two ideas: First, a probabilistic 2d segmentation of fore- and background in all camera images of a static, calibrated multi camera setup is performed based on color distribution models. To make this segmentation more robust and adaptive, the second part integrates 3d scene information reconstructed from all cameras. The 3d information is used as a feedback mechanism to the segmentation task. Hereby the color distributions are adapted automatically to achieve better segmentation results. The basic assumption is that observed objects are surrounded by multiple cameras to obtain complete 3d reconstructions of the foreground.

The steps of the approach are depicted in Fig. 1. First, coarse fore- and background models are generated. They are used with the current camera images to create a probabilistic 3d voxel reconstruction of the scene. Probabilistic in this context means that each reconstructed voxel has a specific occupation probability derived from the probabilities of the corresponding pixels in all views to be foreground. The 3d reconstruction is projected into the camera images, thresholded and in this way provides a masked area of foreground in the images. Image areas which are not covered by this mask are used to update the background model. By utilizing this updated model a segmentation is performed to precisely determine the foreground silhouettes. The silhouettes are used to update the foreground model accurately in a succeeding step. The fore- and background models are then used to create a probabilistic 3d reconstruction of the foreground by using the next camera frame and the loop restarts.

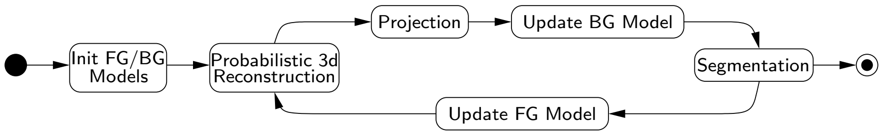

Fig. 1. Segmentation loop utilizing probabilistic 3d fusion as data driven feedback mechanism to enhance the segmentation by automatically adapted color distributions.

2.1 Fore- and Background Model

To model fore- and background, the random variable $\mathcal{F} \in \{0,1\}$ decides whether a pixel at a given time t is fore- or background ($\mathcal{F} = 1$ respectively $\mathcal{F} = 0$). Based on a given color vector $\boldsymbol{c}$ the color distribution $p(\boldsymbol{c}|\mathcal{F} = 1)$ models the foreground and is used to infer the conditional probability $P(\mathcal{F} = 1|\boldsymbol{c})$. The foreground model is generated based on the foreground segment for each frame separately and consists of two parts A and B:

$$p(\boldsymbol{c}|\mathcal{F} = 1) = \underbrace{(1 - P_{\mathrm{NF}}) \sum_{k=1}^{K_{\mathrm{fg}}} \omega^k \eta(\boldsymbol{c}, \boldsymbol{\mu}^k, \boldsymbol{\Sigma}^k)}_{A} + P_{\mathrm{NF}} \underbrace{\mathcal{U}(\boldsymbol{c})}_{B} \ . \tag{1}$$

The first part A models known foreground in terms of a Gaussian Mixture Model (GMM) with the density function $\eta(\boldsymbol{c}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ where $\boldsymbol{\mu}^k$ and $\boldsymbol{\Sigma}^k$ are mean and variance of the k^{th} of K_{fg} components of the mixture and ω^k is the component's weight. B models a uniform color distribution which is necessary to integrate suddenly arising new foreground. Both parts are coupled by the probability $P_{\mathrm{NF}} = \frac{1}{2}$ of new foreground. The model is generated continuously by utilizing k-means clustering of the colors of the foreground silhouette during consecutive frames. The background model consists of two parts as well:

$$p(\boldsymbol{c}_t|\mathcal{F}_t = 0) = \underbrace{(1 - P_{\mathrm{S}}) \sum_{k=1}^{K_{\mathrm{bg}}} \omega_t^k \eta(\boldsymbol{c}_t, \boldsymbol{\mu}_t^k, \boldsymbol{\Sigma}_t^k)}_{C} + \underbrace{P_{\mathrm{S}} \sum_{k=1}^{K_{\mathrm{bg}}} \omega_t^k p(\boldsymbol{c}_t|\mathcal{S}_t^k = 1)}_{D} \ . \tag{2}$$

Part C models the color distribution of the background similar to the model in eq. 1 with K_{bg} components. In contrast to eq. 1, the model is updated over the whole observation time t. The second part D models the occurrence of shadows and highlights. Both parts are again coupled with an additionally probability of shadows $P_{\mathrm{S}} = \frac{1}{2}$. The shadow and highlight model D is modeled in analogy to the background color model C, i.e. the weightings of C are reused. To determine shaded areas or areas of highlights, the colors are examined in the YUV color space. A luminance ratio λ is calculated in the Y channel: $\lambda = \frac{Y_t}{Y_B} = \frac{c_t^1}{\mu_t^{k,1}}$. Two thresholds are introduced to detect shadows, if $\tau_S < 1$, and highlights, if $\tau_H > 1$. The resulting shadow model is:

$$p(\boldsymbol{c}_t|\mathcal{S}_t^k = 1) = \begin{cases} \frac{1}{(\tau_H - \tau_S)\mu_t^{k,1}} \prod\limits_{d=2,3} \eta(c_t^d, {\mu'}_t^{k,d}, \Sigma_t^{k,d}) & \text{if } \tau_S \leq \lambda_t^k \leq \tau_H \\ 0 & \text{else} \end{cases} . \quad (3)$$

The scale factor $\frac{1}{(\tau_H - \tau_S)\mu_t^{k,1}}$ is needed to achieve the density's integration to result in 1. The background model in eq. 2 is updated continuously by integration of all previous frames over time by utilizing an online Expectation Maximization (EM) approach as presented in [5].

2.2 Probabilistic 3d Fusion

To update fore- and background models, a method is needed to reliably identify foreground in the camera images. In case of multi camera setups it is feasible to exploit the strong prior of geometric coherence of the scene observed from multiple views by using the approach of a bayesian probabilistic 3d reconstruction [15]. The volume seen by the cameras is discretized into voxels $\mathcal{V} \in \{0, 1\}$. For each voxel the probability of being foreground is derived from the foreground probabilities of the corresponding pixels in all cameras according to the model definition in [5]. Four a priori probabilities are introduced into the reconstruction model. First, the probability of voxel occupation: $P(\mathcal{V}) = \frac{1}{2}$. Additionally, three error probabilities P_{DF}, P_{FA} and P_{O}. P_{DF} means a *detection failure*, i.e. a voxel should be occupied but is not due to e.g. camera noise. P_{FA} means a *false alarm*, i.e. a voxel should not be occupied but erroneously is, e.g. due to shadows. Finally, P_{O} means an *obstruction*, i.e. a voxel should not be occupied but is on the same line of sight as another voxel which is occupied and, hence, classified incorrectly. The conditional probability of foreground of an unoccupied voxel is, thus, $\mathcal{V}$: $P(\mathcal{F}_n = 1|\mathcal{V} = 0) = P_{\mathrm{O}}(1 - P_{\mathrm{DF}}) + (1 - P_{\mathrm{O}})P_{\mathrm{FA}}$. The conditional probability of background of an unoccupied voxel is $\mathcal{V}$: $P(\mathcal{F}_n = 0|\mathcal{V} = 0) = 1 - [P_{\mathrm{O}}(1 - P_{\mathrm{DF}}) + (1 - P_{\mathrm{O}})P_{\mathrm{FA}}]$. Values of 5% for P_{DF}, P_{FA} and P_{O} provide reasonable results. We use the joint probability distribution defined in [5], and marginalize over the unknown variables $\mathcal{F}_n$ by observing the colors $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_N$ at the corresponding pixels in the images of the cameras $1, \ldots, N$ by eq. 4:

$$P(\mathcal{V} = 1|\boldsymbol{c}_1, \ldots, \boldsymbol{c}_N) = \frac{\prod\limits_{n=1}^{N} \sum\limits_{f \in \{0,1\}} P(\mathcal{F}_n = f|\mathcal{V} = 1) p(\boldsymbol{c}_n|\mathcal{F}_n = f)}{\sum\limits_{v \in \{0,1\}} \prod\limits_{n=1}^{N} \sum\limits_{f \in \{0,1\}} P(\mathcal{F}_n = f|\mathcal{V} = v) p(\boldsymbol{c}_n|\mathcal{F}_n = f)} . \quad (4)$$

The resulting probabilistic 3d reconstruction is backprojected into the camera images and then used to identify fore- and background segments (cf. sec. 2).

2.3 Probabilistic Foreground Detection

By using the probability densities $p(\boldsymbol{c}|\mathcal{F} = 1)$ and $p(\boldsymbol{c}|\mathcal{F} = 0)$ (sec. 2.1) the conditional probability $P(\mathcal{F} = 1|\boldsymbol{c})$ that a pixel belongs to the foreground based on an observed color value $\boldsymbol{c}$ can be calculated using Bayes' rule which under assumption of no a priori knowledge about the unconditional probabilities $P(\mathcal{F} = f)$ and a resulting uniform distribution cancels out to:

$$P(\mathcal{F} = 1|\boldsymbol{c}) = \frac{p(\boldsymbol{c}|\mathcal{F} = 1)}{\sum\limits_{f \in \{0,1\}} p(\boldsymbol{c}|\mathcal{F} = f)} \,. \tag{5}$$

3 Variational Segmentation

The problem of segmentation has been formalized by Mumford and Shah as the minimization of a functional [17]. The level set method was introduced by Osher and Sethian [18] to implicitly propagate hypersurfaces by evolving an appropriate embedding function to find minimizers to such a functional. The variational approach used in our segmentation framework is based on the works of [12,13]. In this section we will shortly review this variational framework and the way the different information is fused. The basis of our segmentation framework is a variation of the very well known energy functional for image segmentation:

$$\begin{aligned} E(\varphi) = &- \int_{\Omega} H(\varphi) \sum_{j=1}^{k} \log p_{1,j}(\boldsymbol{c})\, d\Omega - \int_{\Omega} (1 - H(\varphi)) \sum_{j=1}^{k} \log p_{2,j}(\boldsymbol{c})\, d\Omega \\ &+ \nu_1 \int_{\Omega} |\nabla H(\varphi)|\, d\Omega \,, \end{aligned} \tag{6}$$

where $\boldsymbol{c} \in \mathbb{R}^k$ is the image feature vector, $H(\varphi)$ is a regularized Heaviside function and $p_{i,j}$ are specific, independent object ($i = 1$) and background ($i = 2$) distributions for the different image feature channels j. These distributions can be inferred from the respective regions (divided by $\varphi(x)$) by fitting parametric distributions or by performing the nonparametric Parzen density estimates [19] to histograms of the feature channels.

Instead of multiplying the different probabilities arising from the feature channels, which leads to the above formulation of the segmentation energy, we generalized this approach and use Dempster-Shafer theory of evidence [20] to fuse information arising from different feature channels. The key idea, which makes it different from other Bayesian frameworks, is the use of Dempster's rule of combination to fuse different information [20]. This allows to favor feature channels that support a specific region instead of favor channels with low support for a region. We will make use of this property to fuse the image data of traditional segmentation frameworks and the information arising from segmentation by probabilistic foreground detection.

The energy functional, which uses evidence theory can be expressed as follows:

$$\begin{aligned} E(\varphi) = &- \int_{\Omega} H(\varphi) \log m(\Omega_1)\, d\Omega - \int_{\Omega} (1 - H(\varphi)) \log m(\Omega_2)\, d\Omega \\ &+ \nu_1 \int_{\Omega} |\nabla H(\varphi)|\, d\Omega \,, \end{aligned} \tag{7}$$

where $m = m_1 \otimes m_2 \otimes \ldots \otimes m_k$ is the mass function, fusing k feature channels with Dempster's rule of combination. The single mass functions m_j are defined by the object and background distributions $p_{i,j}$:

$$\begin{aligned} m_j(\emptyset) &= 0\,, & m_j(\Omega) &= 1 - (p_{1,j}(\boldsymbol{c}) + p_{2,j}(\boldsymbol{c}))\,, \\ m_j(\Omega_1) &= p_{1,j}(\boldsymbol{c}), & m_j(\Omega_2) &= p_{2,j}\boldsymbol{c}\,, \end{aligned} \tag{8}$$

for $j \in \{1, \ldots, k\}$. The mass $m_j(\Omega) = m_j(\{\Omega_1, \Omega_2\})$ introduces a way to represent inaccuracy and uncertainty of the feature channels, while the mass $m_j(\Omega_i)$ can be interpreted as the belief strictly placed on foreground- (Ω_1) and background (Ω_2) regions. Dempster's rule of combination is defined by:

$$m(\rho_1) = m_1(\rho_1) \otimes m_2(\rho_1) = \frac{\displaystyle\sum_{\rho_2 \cap \rho_3 = \rho_1} m_1(\rho_2) m_2(\rho_3)}{1 - \displaystyle\sum_{\rho_2 \cap \rho_3 = \emptyset} m_1(\rho_2) m_2(\rho_3)}\,, \tag{9}$$

where $\rho_1, \rho_2, \rho_3 \in \wp\{\Omega_1, \Omega_2\} = \{\emptyset, \Omega_1, \Omega_2, \{\Omega_1, \Omega_2\}\}$.

The minimization of the energy (7) with respect to φ can be performed using variational methods and a gradient descent [16]. Thus, the segmentation process works according to the EM principle with an initial partitioning.

4 Integrating Probabilistic 3d Fusion into Variational Segmentation

Given the probabilities $P(\mathcal{F} = 1|\boldsymbol{c})$ for each feature vector $\boldsymbol{c}$ arising from the probabilistic foreground detection (5) we build the following mass function:

$$\begin{aligned} m_{\text{fg}}(\emptyset) &= 0\,, & m_{\text{fg}}(\Omega) &= 1 - \nu_2\,, \\ m_{\text{fg}}(\Omega_1) &= \nu_2 \cdot P(\mathcal{F} = 1|\boldsymbol{c}), & m_{\text{fg}}(\Omega_2) &= \nu_2 \cdot (1 - P(\mathcal{F} = 1|\boldsymbol{c}))\,, \end{aligned} \tag{10}$$

with a weighting parameter $\nu_2 \in [0, 1]$. This parameter can be interpreted as the belief we put on the probabilistic foreground detection. With a parameter $\nu_2 < 1$ we integrate inaccuracy. As a consequence, the evolving boundary is directly driven by the intensity information of the image and the result of the probabilistic 3d fusion.

The mass function m_{fg} is now integrated into the variational approach for image segmentation (7) using Dempster's rule of combination:

$$m_{\text{new}} = m \otimes m_{\text{fg}} = m_1 \otimes m_2 \otimes \ldots \otimes m_k \otimes m_{\text{fg}}\,. \tag{11}$$

The energy functional for segmentation fusing image features and probabilistic foreground detection can be written as:

$$\begin{aligned} E(\varphi) = & \underbrace{- \int_\Omega H(\varphi) \log m_{\text{new}}(\Omega_1)\, d\Omega - \int_\Omega (1 - H(\varphi)) \log m_{\text{new}}(\Omega_2)\, d\Omega}_{\text{fusion of image features and probabilistic foreground detection}} \\ & + \nu_1 \int_\Omega |\nabla H(\varphi)|\, d\Omega\,. \end{aligned} \tag{12}$$

Compared to the Bayesian approach the proposed framework is able to correct wrong classifications coming from the probabilistic foreground detection and vice versa, because channels with a strong support are favored.

5 Evaluation

We present a qualitative and a quantitative analysis of our algorithm based on the images of the *Dancer Sequence* in [5] and our own recordings of gymnasts with seven Prosilica GE680C cameras in a circular setup. The quantitative analysis is performed based on hand labeled data.

In a qualitative analysis we compare the results of the approach of [5] with the results of a variational segmentation, with GrabCut [14] and the results of the proposed combined approach. The probabilistic segmentation of [5] is initialized with a priori recorded background images. These images varied in lighting and details which was automatically compensated by the presented approach. In case of the variational segmentation and GrabCut, the result of the probabilistic 3d fusion is used as the initial boundary. In the combined approach the information from the probabilistic 3d fusion is used as the initial boundary and integrated into the variational segmentation framework as proposed in eq. (12).

In Fig. 2 we present exemplary results of all four approaches performed on a difficult scene with very similar color distributions of fore- and background. It is clearly observable that neither the variational approach nor the segmentation by probabilistic fusion are able to fully cope with that ambiguity. The variational approach integrates large parts of the wooden background into the foreground silhouette while the approach of [5] leads to very low probabilities of foreground in the ambiguous areas. Solely, the proposed approach leads to satisfying results in such difficult scenarios. As an alternative to variational segmentation, the results of the probabilistic segmentation could also be used as initialization for GrabCut. But we found that only the combination of initialization by probabilistic segmentation and fusing this information utilizing the Dempster-Shafer approach can close erroneous holes and, thus, recover from false classifications.

Due to the convincing results of Fig. 2 we performed a quantitative analysis of the three approaches and measured the error compared to hand labeled data. Exemplary results of the cameras 6 and 7 are presented in Fig. 3. Camera 6 has been chosen because this view contains background motion and we want to demonstrate that the adaptivity of [5] is not compromised by the presented approach. The results of Camera 7 are selected to link the qualitative results in Fig. 2 with quantitative results to clarify the benefits of the presented approach. In all cases the proposed approach provides better results over the full sequence.

Finally, we performed a qualitative analysis of the proposed approach on the dancer from [5]. We were able to show, that again, our approach gains better segmentation results (cf. Fig. 4) than the probabilistic segmentation. We could also additionally demonstrate that the proposed approach is applicable in these kinds of difficult scenarios with occluding noise and, thus, unites the benefits of robust segmentation and robust dense 3d reconstruction results.

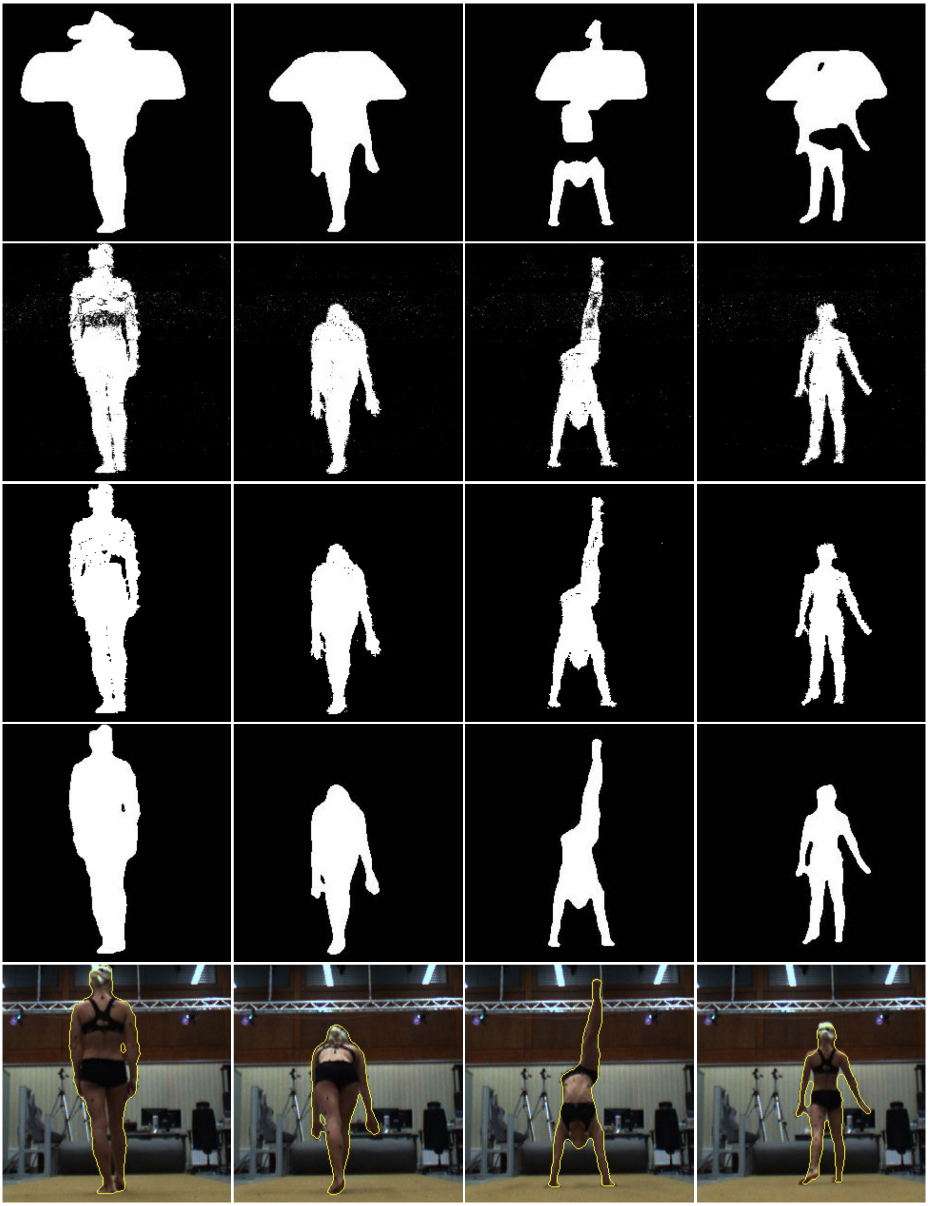

Fig. 2. Frames 100, 300, 500 and 700 of camera 7. First row: Variational segmentation only. Second row: Segmentation by probabilistic fusion only without post processing. Third row: Combined approach with GrabCut segmentation. Fourth row: Proposed combined approach with variational segmentation. Fifth row: Input image and detected contour of combined approach. All single approaches have difficulties in areas with nearly identical color distributions of fore-/background. Only the proposed combined approach is able to cope with these kinds of ambiguities.

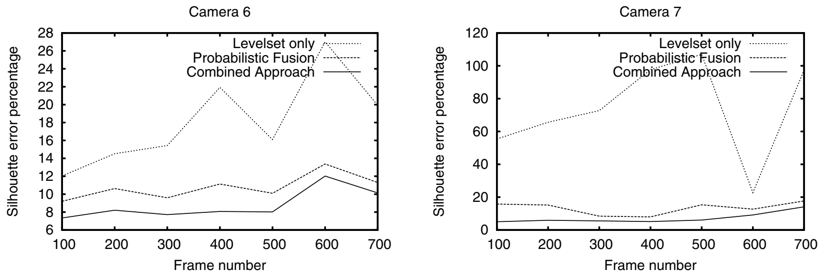

Fig. 3. Silhouette error percentage in steps of 100 frames in cameras 6 and 7 of the gymnast sequence. The proposed approach generates the best results with the fewest errors compared to the variational approach and the approach of [5].

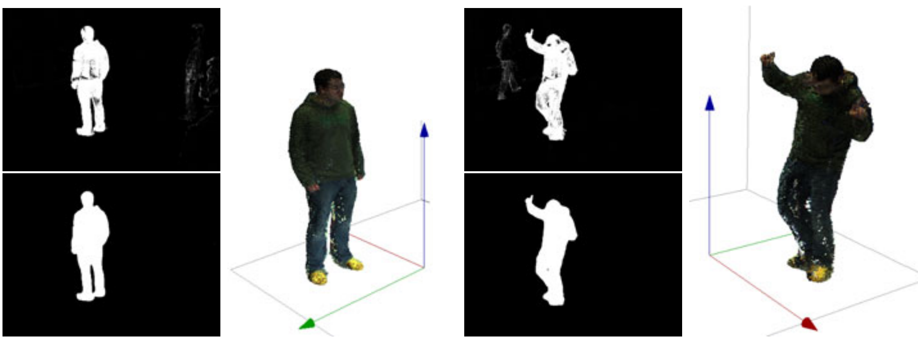

Fig. 4. Column 1, top: Probabilistic segmentation of the second frame (after first model update) of camera 6 of the dancer sequence; Bottom: Segmentation of proposed approach. Column 2: Resulting 3d reconstruction of proposed approach. Column 3, top: Probabilistic segmentation of frame 645; Bottom: Segmentation of the proposed approach. Column 4: Resulting 3d reconstruction.

6 Conclusion

In the presented work we developed a new approach for color based foreground segmentation with multi camera setups which implicitly integrates geometric priors of the used camera setup and energy based constraints to allow an adaptive, purely data driven high quality segmentation of foreground in cluttered, changing and, thus, realistic scenarios. The new approach combines the segmentation by probabilistic 3d fusion and the variational approach of level set segmentation based on Dempster-Shafer theory of evidence. We revealed that both algorithms on their own have massive difficulties in scenarios with very similar color distributions of fore- and background. However, we were able to show, that with our specific approach the integration of both methods allows tremendous improvements of the segmentation results in these kinds of scenarios. To attest the impact of our method, we performed qualitative as well as quantitative evaluations on natural image sequences. We showed that the combination of probabilistic 3d fusion and the level set segmentation based on Dempster-Shafer theory of evidence produces much better foreground extractions, which is an important prerequisite for many tasks in computer vision.

References

1. Grauman, K., Shakhnarovich, G., Darrell, T.: A bayesian approach to image-based visual hull reconstruction. In: IEEE CVPR, vol. 1, pp. 187–194 (2003)
2. Gordon, G., Darrell, T., Harville, M., Woodfill, J.: Background estimation and removal based on range and color. In: IEEE CVPR, vol. 2, pp. 24–59 (1999)
3. Cheung, G.K., Kanade, T., Bouguet, J.Y., Holler, M.: A real time system for robust 3d voxel reconstruction of human motions. In: IEEE CVPR, vol. 2, pp. 714–720 (2000)
4. Kolev, K., Brox, T., Cremers, D.: Robust variational segmentation of 3d objects from multiple views. In: Franke, K., Müller, K.-R., Nickolay, B., Schäfer, R. (eds.) DAGM 2006. LNCS, vol. 4174, pp. 688–697. Springer, Heidelberg (2006)
5. Feldmann, T., Dießelberg, L., Wörner, A.: Adaptive foreground/background segmentation using multiview silhouette fusion. In: Denzler, J., Notni, G., Süße, H. (eds.) DAGM 2009. LNCS, vol. 5748, pp. 522–531. Springer, Heidelberg (2009)
6. Vogiatzis, G., Torr, P.H.S., Cipolla, R.: Multi-view stereo via volumetric graph-cuts. In: IEEE CVPR, vol. 2, pp. 391–398 (2005)
7. Lim, S.N., Mittal, A., Davis, L.S., Paragios, N.: Fast illumination-invariant background subtraction using two views: Error analysis, sensor placement and applications. In: IEEE CVPR, pp. 1071–1078 (2005)
8. Lee, W., Woo, W., Boyer, E.: Identifying foreground from multiple images. In: Yagi, Y., Kang, S.B., Kweon, I.S., Zha, H. (eds.) ACCV 2007, Part II. LNCS, vol. 4844, pp. 580–589. Springer, Heidelberg (2007)
9. Schmaltz, C., Rosenhahn, B., Brox, T., Weickert, J.: Localised mixture models in region-based tracking. In: Denzler, J., Notni, G., Süße, H. (eds.) DAGM 2009. LNCS, vol. 5748, pp. 21–30. Springer, Heidelberg (2009)
10. Kim, K., Chalidabhongse, T., Harwood, D., Davis, L.: Background modeling and subtraction by codebook construction. In: ICIP, vol. 5, pp. 3061–3064 (October 2004)
11. Stauffer, C., Grimson, W.: Adaptive background mixture models for real-time tracking. In: IEEE CVPR, vol. 2, pp. 2246–2252 (1999)
12. Chan, T., Vese, L.: Active contours without edges. IEEE TIP 10(2), 266–277 (2001)
13. Zhu, S.C., Yuille, A.: Region competition: unifying snakes, region growing, and bayes/mdl for multiband image segmentation. IEEE TPAMI 18(9), 884–900 (1996)
14. Rother, C., Kolmogorov, V., Blake, A.: "grabcut": interactive foreground extraction using iterated graph cuts. ACM Trans. Graph. 23(3), 309–314 (2004)
15. Franco, J.S., Boyer, E.: Fusion of multi-view silhouette cues using a space occupancy grid. Technical Report 5551, INRIA (April 2005)
16. Scheuermann, B., Rosenhahn, B.: Analysis of numerical methods for level set based image segmentation. In: Bebis, G., Boyle, R., Parvin, B., Koracin, D., Kuno, Y., Wang, J., Pajarola, R., Lindstrom, P., Hinkenjann, A., Encarnação, M.L., Silva, C.T., Coming, D. (eds.) ISVC 2009, Part II. LNCS, vol. 5876, pp. 196–207. Springer, Heidelberg (2009)
17. Mumford, D., Shah, J.: Boundary detection by minimizing functionals. In: IEEE CVPR, San Francisco, CA, pp. 22–26 (1985)
18. Osher, S., Sethian, J.: Fronts propagating with curvature dependent speed: Algorithm based on hamilton-jacobi formulation. J. Comput. Phys. 79, 12–49 (1988)
19. Kim, J., Fisher III, J., Yezzi Jr., A., Cetin, M., Willsky, A.: Nonparametric methods for image segmentation using information theory and curve evolution. In: ICIP, pp. 797–800 (2002)
20. Dempster, A.P.: A generalization of bayesian inference. Journal of the Royal Statistical Society, Series B (Methodological) 30(2), 205–247 (1968)