On-line, Incremental Learning of a Robust Active Shape Model

Michael Fussenegger[1], Peter M. Roth[2], Horst Bischof[2], Axel Pinz[1]

[1] Institute of Electrical Measurement and Measurement Signal Processing
Graz University of Technology
Kopernikusgasse 24/IV, 8010 Graz, Austria
{fussenegger,axel.pinz}@tugraz.at
[2] Institute for Computer Graphics and Vision
Graz University of Technology
Inffeldgasse 16/II, 8010 Graz, Austria
{pmroth,bischof}@icg.tu-graz.ac.at

Abstract. Active Shape Models are commonly used to recognize and locate different aspects of known rigid objects. However, they require an off-line learning stage, such that the extension of an existing model requires a complete new retraining phase. Furthermore, learning is based on principal component analysis and requires perfect training data that is not corrupted by partial occlusions or imperfect segmentation. The contribution of this paper is twofold: First, we present a novel robust Active Shape Model that can handle corrupted shape data. Second, this model can be created on-line through the use of a robust incremental PCA algorithm. Thus, an already partially learned Active Shape Model can be used for segmentation of a new image in a level set framework and the result of this segmentation process can be used for an on-line update of the robust model. Our experimental results demonstrate the robustness and the flexibility of this new model, which is at the same time computationally much more efficient than previous ASMs using batch or iterated batch PCA.

1 Introduction

Prior knowledge of the object contour/shape is used to improve the result in many computer vision approaches dealing with segmentation, object detection or tracking. A common approach, to model different aspects of rigid objects in a shape prior formalism, is the use of Active Shape Models (ASMs) proposed by Cootes et al. [4, 5]. The standard ASM framework consists of two stages: (1) the modeling/learning and (2) the segmentation/detection stage.

In this paper we use a level set segmentation framework. Level set representation [14] is an established technique for image segmentation. Over the years several different level set models of ASMs have been presented (e.g., [6, 16]). In particular, we use the level set representation of Rousson and Paragios [16]. To avoid unnecessary computation and numerical errors, we work with the level set shape representation and avoid a conversion to the classical landmark representation used in [5].

In the learning stage, a set of registered training shapes is used to model different aspects of a model. Generally ASM approaches use a batch (off-line) version of Principal Component Analysis (PCA) [11] for learning, that has two main disadvantages: (1) the approach is not robust in the recognition nor in the training [12, 18] (but, we might receive data that is corrupted by partial occlusions or imperfect segmentation) and (2) all training data has to be provided a priori, thus, hand segmentation is required and the shape model cannot be extended as new segmentation results become available.

Various approaches have been proposed to introduce robustness in the recognition stage (e.g., [1, 12, 15]). For these methods it is assumed that the samples in the learning stage are undisturbed. Robust learning is a more difficult problem, since there is no previous knowledge that can be used to estimate outliers. Several methods have been proposed to robustly extract the principal axes in the presence of outliers [7, 21]. Other approaches use robust M-estimator [7] or are based on the EM formulation of PCA [17, 18, 20]. Using a robust approach in our framework has two advantages: (1) the robust reconstruction from the ASM allows a much better segmentation of occluded objects and (2) robust learning on the improved segmentation results provides a better shape representation.

We use an incremental PCA approach in our ASM. Applying an incremental method, we can efficiently build and update an ASM that is used for the segmentation process, i.e., we can use the partially learned ASM to perform segmentation and use the segmentation result to retrain the ASM. Different incremental PCA approaches have been proposed that are based on incremental SVD-updating (e.g., [2, 9]). Recently even robust and incremental [13, 19] approaches have been proposed. In particular we apply a simplified version of the approach of Skočaj and Leonardis [19] to learn the ASM that will be explained in Section 2.2.

Applying this incremental and robust PCA method, we need a priori only a small hand segmented data set to initialize our ASM. This first model provides shape priors for the segmentation process. Furthermore, the Active Shape Model can be successively updated with new data from the segmentation process.

The outline of the paper is as follows: Section 2 explains our system and gives a short description of its components. In Section 2.1, we describe the shape registration. In Section 2.2, we introduce in detail the Robust Incremental PCA. Experiments are presented in Section 3 and finally, conclusions are drawn in Section 4.

2 Incremental Robust Active Shape Model

Fig. 1 depicts our proposed method, which is split into two components: (i) the segmentation module and (ii) our novel ASM module. For the segmentation, we use the approach proposed by Fussenegger et al. [8] which is based on the level set formulations by Rousson and Paragios [16] and Brox and Weickert [3].

The output of the segmentation module is the distance function $\Phi(\mathbf{x})$, $\Phi : \Omega \rightarrow \mathbb{R}$, with $\Phi(\mathbf{x}) > 0$, if $\mathbf{x}$ lies in the shape and $\Phi(\mathbf{x}) < 0$, if $\mathbf{x}$ lies out of the shape. Ω is the image domain and $\mathbf{x}$ denotes a pixel in Ω. In order to avoid unnecessary computation and numerical errors, we use directly this distance function as the shape representation instead of the landmark representation used in [4, 5]. However, our ASM can be easily adapted to alternative shape representations for use with other segmentation approaches.

In a first step, the ASM module is initialized with a training set of non corrupted, aligned shapes characterizing different aspects of an object to learn a first Active Shape Model. This learned ASM is then used in the segmentation process. After each level set iteration step the current result Φ_i is passed from the segmentation module to the ASM module. A registration process (Section 2.1) applies a similarity transformation $\mathcal{A}$ on Φ_i to map it with Φ_M in the best way. Φ_M is the mean shape calculated over the already learned shapes. Φ_i is then projected to the eigenspace and robustly reconstructed (Section 2.2). The reconstruction $\tilde{\Phi}_i$ is passed to the segmentation module and used as a shape prior in the next level set iteration step. This is repeated until the segmentation process ends. The final result is used to update and improve our ASM.

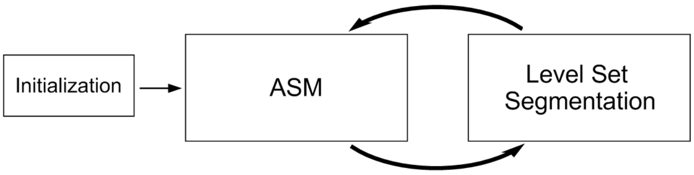


Fig. 1. Our System consisting of two interacting components: The level set segmentation [8] and our novel Active Shape Model.

2.1 Shape registration

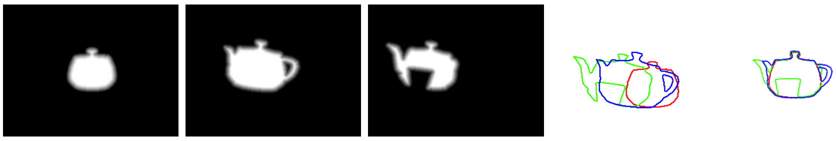
For the shape registration, we assume a global deformation $\mathcal{A}$ between Φ_M (the mean shape) and Φ (the new shape) that involves the parameters $[\mathcal{A} = (s; \theta; \mathbf{T})]$ with a scale factor s, a rotation angle θ and a translation vector **T** [16]. The objective function

$$E(\Phi_M, \Phi(\mathcal{A})) = \int_\Omega (s\Phi_M - \Phi(\mathcal{A}))^2 d\mathbf{x} \tag{1}$$

can be used to recover the optimal registration parameters. The rigid transformation $\mathcal{A}$ is dynamically updated to map Φ_M and Φ in the best way. Thus, the calculus of variations for the parameters of $\mathcal{A}$ yields the system

$$\begin{aligned}
\frac{\partial s}{\partial t} &= 2\int_\Omega (s\Phi_M - \Phi(\mathcal{A}))(\Phi_M - \nabla\Phi(\mathcal{A})\frac{\partial}{\partial s}\mathcal{A})d\mathbf{x}\\
\frac{\partial \theta}{\partial t} &= 2\int_\Omega (s\Phi_M - \Phi(\mathcal{A}))(-\nabla\Phi(\mathcal{A})\frac{\partial}{\partial \theta}\mathcal{A})d\mathbf{x}\\
\frac{\partial \mathbf{T}}{\partial t} &= 2\int_\Omega (s\Phi_M - \Phi(\mathcal{A}))(-\nabla\Phi(\mathcal{A})\frac{\partial}{\partial \mathbf{T}}\mathcal{A})d\mathbf{x}.
\end{aligned} \tag{2}$$

Fig. 2(a-c) shows three example shapes. Fig. 2(d) and 2(e) show all three shape contours before and after the registration process.



(a) (b) (c) (d) (e)

Fig. 2. Three example shapes before and after registration.

2.2 Robust Incremental PCA

For batch PCA all training images are processed simultaneously. A fixed set of input images $\mathbf{X} = [\mathbf{x}_1, \ldots \mathbf{x}_n] \in \mathbb{R}^{m \times n}$ is given, where $\mathbf{x}_i \in \mathbb{R}^m$ is an individual image represented as a vector. It is assumed that $\mathbf{X}$ is mean normalized. Let $\mathbf{Q} \in \mathbb{R}^{m \times m}$ be the covariance matrix of $\mathbf{X}$, then the subspace $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_n] \in \mathbb{R}^{m \times n}$ can be computed by solving the eigenproblem for $\mathbf{Q}$ or more efficiently by solving SVD of $\mathbf{X}$.

For incremental learning, the training images are given sequentially. Assuming that an eigenspace was already built from n images, at step $n+1$ the current eigenspace can be updated in the following way [19]: First, the new image $\mathbf{x}$ is projected in the current eigenspace $\mathbf{U}^{(n)}$ and the image is reconstructed: $\tilde{\mathbf{x}}$. The residual vector $\mathbf{r} = \mathbf{x} - \tilde{\mathbf{x}}$ is orthogonal to the current basis $\mathbf{U}^{(n)}$. Thus, a new basis $\mathbf{U}'$ is obtained by enlarging $\mathbf{U}^{(n)}$ with $\mathbf{r}$. $\mathbf{U}'$ represents the current images as well as the new sample. Next, batch PCA is performed on the corresponding low-dimensional space $\mathbf{A}'$ and the eigenvectors $\mathbf{U}''$, the eigenvalues $\boldsymbol{\lambda}''$ and the mean $\boldsymbol{\mu}''$ are obtained. To update the subspace the coefficients are projected in the new basis $\mathbf{A}^{(n+1)} = \mathbf{U}''^T \left(\mathbf{A}' - \boldsymbol{\mu}''\mathbf{1}\right)$ and the subspace is rotated: $\mathbf{U}^{(n+1)} = \mathbf{U}'\mathbf{U}''$. Finally, the mean $\boldsymbol{\mu}^{(n+1)} = \boldsymbol{\mu}^{(n)} + \mathbf{U}'\boldsymbol{\mu}''$ and the eigenvalues $\boldsymbol{\lambda}^{(n+1)} = \boldsymbol{\lambda}''$ are updated. In each step the dimension of the subspace is increased by one. To preserve the dimension of the subspace the least significant principal vector may be discarded [10]. To obtain an initial model, the batch method may be applied to a smaller set of training images. Alternatively, to have a fully incremental algorithm, the eigenspace may be initialized using the first training image $\mathbf{x}$: $\boldsymbol{\mu}^{(1)} = \mathbf{x}$, $\mathbf{U}^{(1)} = \mathbf{0}$ and $\mathbf{A}^{(1)} = \mathbf{0}$.

This method can be extended in a robust manner, i.e., corrupted input images may be used for incrementally updating the current model. To achieve this, outliers in the current image are detected and replaced by more confident values: First, an image is projected to the current eigenspace using the robust approach [12] and the image is reconstructed. Second, outliers are detected by pixel-wise thresholding (based on the expected reconstruction error) the original image and its robust reconstruction. Finally, the outlying pixel values are replaced by the robustly reconstructed values.

3 Experiments

For the experiments, we have created several different data sets: *teapot*, *African man*, *elephant* and *octopus* (Fig. 3). The first one (Fig. 3(a)) was created artificially by using 3D-MAX. The others are representing real world objects where the images were acquired in two different ways: a smaller number of images was obtained using a turntable and a homogeneous background, such that we can use the level set segmentation without any shape-prior (Fig. 3(b)-(d)). These views are used to build the initial consistent model. Depending on the complexity of the object, i.e., the complexity of the object's shape, 10 up to 80 views are needed. The consistent model is necessary for robust reconstruction of outlying values in the input shapes resulting from over-segmentation and under-segmentation in further steps. Additionally, more complex images (hand held presentations of the objects with cluttered background) are acquired and used to demonstrate the incremental update and robustness of the method (Fig. 3(e)-(f)).

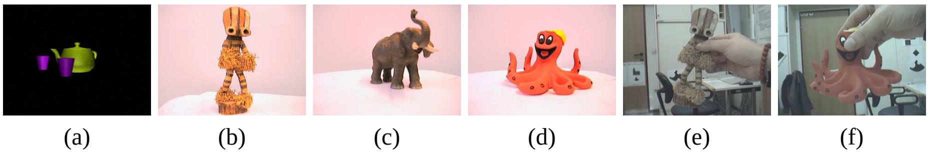


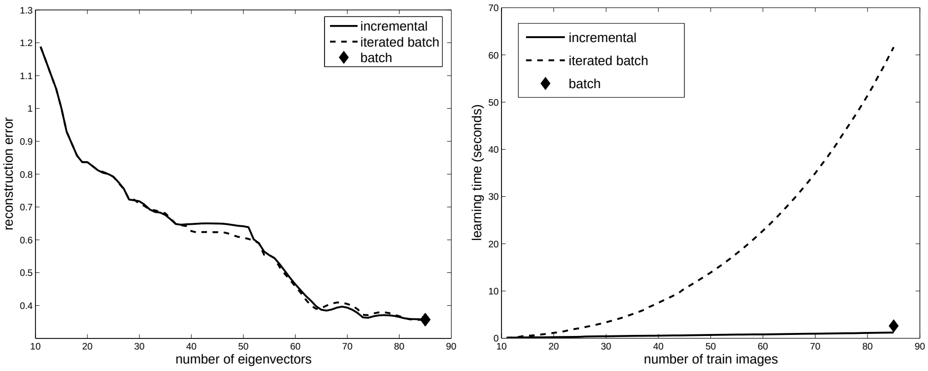
Fig. 3. Examples of our data sets: *teapot*, *African man*, *elephant* and *octopus*.

To show the benefit of the incremental method we trained classifiers using different PCA approaches on the *teapot* data set. In total 85 images were processed in the training stage where only 10 eigenvectors were used for reconstruction. The obtained classifiers were evaluated on an independent test set of 10 images.

First, Fig. 4(a) shows that the incremental on-line method yields similar results as the "iterated batch" method (batch PCA is applied when a new image arises) that is applied in most applications. The reconstruction errors of the incremental PCA, the "iterated batch" PCA and the batch PCA are compared for an increasing number of training shapes and eigenvectors. The reconstruction error is similar for both, the incremental and the iterated batch method. Both error curves are continuously decreasing when the number of training shapes is increased and they are approaching the results of the batch method (trained from the full set of training shapes). Thus, there is no real loss of accuracy (for our application) in using the incremental approach. But as can be seen in Fig. 4(b) there are huge differences in the computational costs for the different methods; the learning times were obtained by evaluating the training in MATLAB on a 3GHz machine. The results for the whole training set containing 85 images are summarized in Table 1. Since the matrix operations are performed on smaller matrices only (less memory has to be allocated) for this data set the incremental method is even computationally cheaper than the batch method. But, as the main point, compared to the iterated batch "incremental" approach the computational costs of the incremental method are only approximately $1/40$!

method	incremental PCA	batch PCA	iterated batch PCA
time	4.72s	6.55s	205.38s

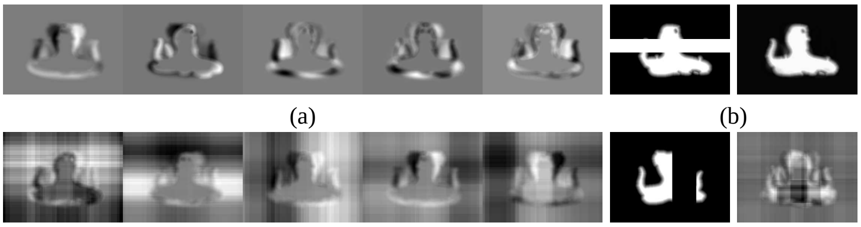
Table 1. Performance evaluation.



(a) (b)

Fig. 4. Incremental PCA approach evaluated on the *teapot* data set: (a) incremental PCA performs similar as incremental batch PCA, (b) incremental PCA is computationally much cheaper than incremental batch PCA.

As the applied incremental PCA method can easily be extended in a robust manner we want to show the advantages of robust incremental learning. For this purpose an initial classifier for the octopus is trained using only 15 clean shapes to get a consistent starting model. Later on, the training is continued from corrupted data. To simulate over-segmented and under-segmented shapes the corrupted data is created by randomly adding black and white bars occluding 25% of the image. By adding these shapes the non-robust model gets more and more corrupted (see (Fig. 5(c) for the first 5 eigenimages) while a stable model is estimated by using the robust approach (see Fig. 5(a) for the first 5 eigenimages). Examples of reconstructions are shown in Fig. 5(b) (robust eigenspace) and Fig. 5(d) (non-robust eigenspace).



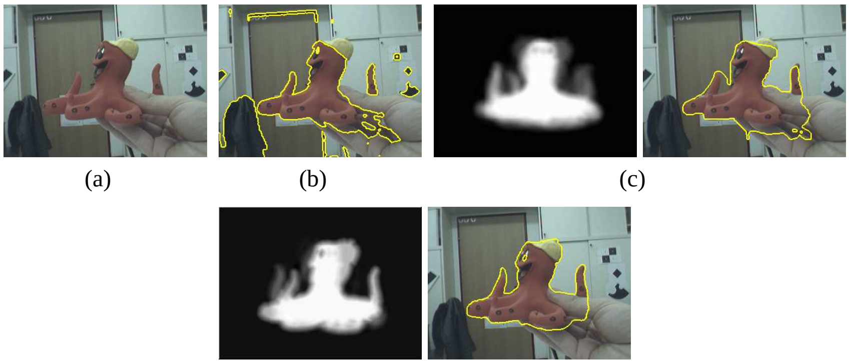
(a) (b)

(c) (d)

Fig. 5. Robust incremental vs. plain incremental approach: (a) eigenspace obtained by robust incremental learning from noisy data, (b) reconstruction from robust eigenspace, (c) eigenspace obtained by incremental learning from noisy data, (d) reconstruction from non-robust eigenspace.

To show the increasingly better segmentation results when incrementally updating the current model with newly obtained shapes, the more complex *octopus* data set was evaluated. In total 85 training shapes were processed where 35 eigenvectors were used for reconstruction. Fig. 6 shows different level set segmentation results using our

ASM in different training stages. For Fig. 6(b), the segmentation is done without a trained ASM. In this case the segmentation fails completely. In Fig. 6(c), we show the final shape prior provided from the initialized ASM (40 “off-line” training shapes) and the corresponding segmentation. The segmentation has been improved significantly but still some errors are present. Afterwards, our ASM is incrementally updated with new “on-line” training shapes. Fig. 6(d) shows the results after 40 additional incrementally obtained shapes. The segmentation is perfect and the segmentation result depicted in Fig. 6 can then be used to add a new aspect to our ASM.



(a) (b) (c)

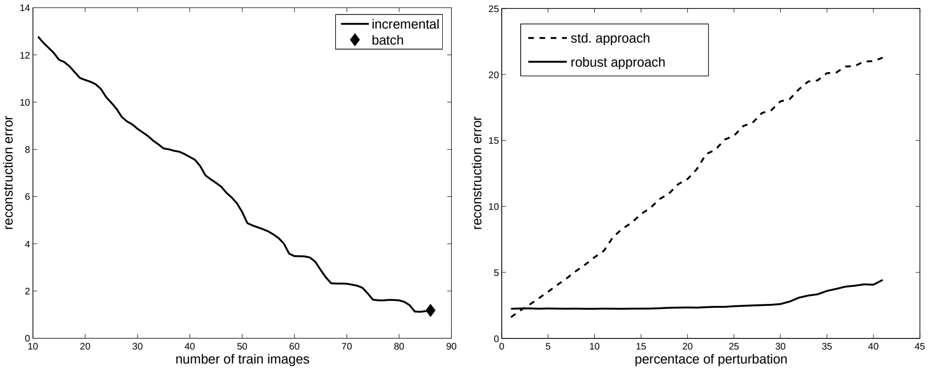
(d)

Fig. 6. Original image (a) and level set segmentation without an ASM (b). Estimated shape prior, with an ASM learned from 40 training shapes and corresponding level set segmentation (c). Estimated shape prior, with an ASM learned from 80 training shapes and corresponding level set segmentation (d).

For a more general evaluation different models (varying the number of processed training shapes) were evaluated on a set of 10 independent test images. For this purpose the shapes in the test set were reconstructed using the previously trained models and the reconstruction errors were analyzed. According to the number of processed shapes (10 to 85) up to 35 eigenvectors were used for reconstruction. The results are shown in Fig. 7(a). It can be seen that the reconstruction error is decreasing if the number of learned shapes (and thus the number of eigenimages used for reconstruction) is increased; a better model is obtained! Examples of continuously improving ASMs are shown in Fig. 8.

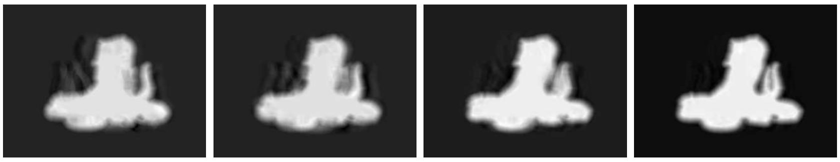
Furthermore, the robust extension of the approach was evaluated on the *African man* data set. The reconstruction error[3] was analyzed when the portion of noise is increased. The results are shown in Fig. 7(b). While the reconstruction error is continuously growing for the standard approach the performance of the robust method is not decreased even in cases of up to 25% occlusion. Thus, these robustly obtained reconstructions

[3] The reconstruction error was computed from the undisturbed original images. Alternatively, the distance between the learned eigenspace and the projected shape may be used as measurement (which will yield similar curves).



(a) (b)

Fig. 7. Better segmentation results by robust incremental ASM: (a) increasing the size of the model decreases the reconstruction error, (b) smaller reconstruction error when applying the robust method even for corrupted data.



(a) (b) (c) (d)

Fig. 8. Improving ASM by updating with new shapes: (a) ASM learned from 10 frames, (b) ASM learned from 20 frames, (c) ASM learned from 40 frames, (d) ASM learned from 80 frames,

can be used to update the current ASM as it can be seen in Fig. 9. The object was presented in a realistic scenario with background clutter by hand from varying views and under different illumination conditions. The registered segmentation without using a shape prior in Fig. 9(b) contains holes and over-segmentations. Thus, the standard reconstruction depicted in Fig. 9(c) is corrupted. But as shown in Fig. 9(e) the robust approach provides a perfect reconstruction that can be used to update the ASM. In addition, Fig. 9(d) shows the corrupted segmentation result obtained by using the standard reconstruction as shape prior while the segmentation result obtained by using the robust reconstruction shown in Fig. 9(f) is perfect.

Finally, we show some examples of segmentations using previously trained Active Shapes Models. In Fig. 10(a), we use the level set segmentation based on [8] without any shape information. The other three Figures 10(b)-(d) are segmented with three different ASMs. On all three objects, we achieve excellent segmentation results, even for Fig. 10(d), where the lower part of the object is highly occluded, our robust ASM is able to segment the object correctly.

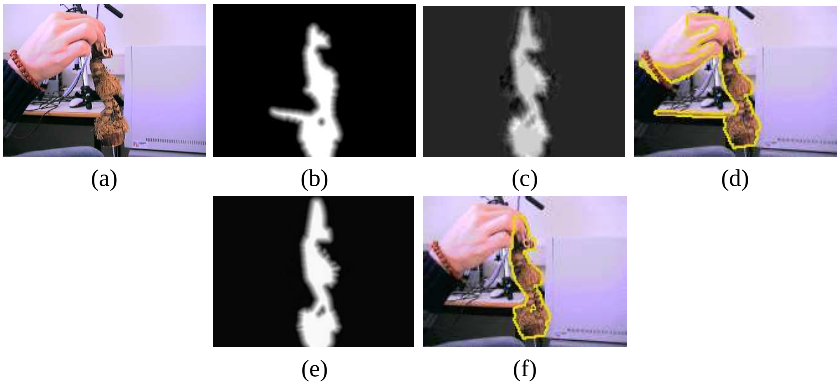


Fig. 9. Robust PCA on *African man* data set: (a) original data, (b) segmentation and registration, (c) reconstruction, (d) segmentation result with (c) as shape prior, (e) robust reconstruction, (f) segmentation result using (e) as shape prior.

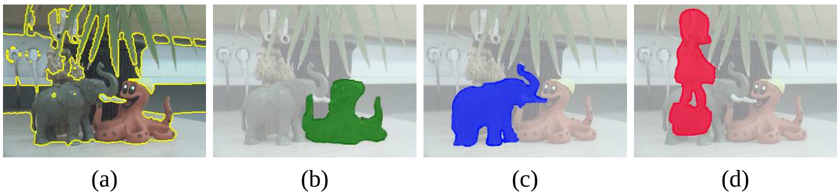


Fig. 10. Level set segmentation results based on [16]: (a) segmentation without a shape prior; (b)-(d) segmentation using different ASMs.

4 Conclusion

We have introduced a novel robust Active Shape Model, that can be updated on-line. Using a robust, incremental PCA allows a successive update of our ASMs even with non perfect data (i.e., corrupted data by partial occlusions or imperfect segmentation). For the segmentation and shape representation, we use the work of Rousson et al. [16] but our ASM can easily be adapted to other segmentation approaches and shape representations. We performed experiments on various data sets with different objects. The advantages of the robust, incremental PCA over the standard batch PCA were shown, and we also showed excellent results using different ASMs for segmentation. Even highly occluded objects in a cluttered background can be segmented correctly. Compared to the standard approach the computational costs can be dramatically reduced. Moreover the user interaction is reduced to taking a smaller number of images on a turn table under perfect conditions, i.e., manual segmentation can be completely avoided!

Acknowledgment

This work has been supported by the Austrian Joint Research Project Cognitive Vision under projects S9103-N04 and S9104-N04.

References

1. M. J. Black and A. D. Jepson. Eigentracking: Robust matching and tracking of articulated objects using a view-based representation. In *Proc. European Conf. on Computer Vision*, pages 329–342, 1996.
2. M. Brand. Incremental singular value decomposition of uncertain data with missing values. In *Proc. European Conf. on Computer Vision*, volume I, pages 707–720, 2002.
3. T. Brox and J. Weickert. Level set based image segmentation with multiple regions. In *Proc. DAGM Symposium*, pages 415–423, 2004.
4. T. F. Cootes, D. H. Cooper, C. J. Taylor, and J. Graham. A trainable method of parametric shape description. pages 289–294, 1992.
5. T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Gaham. Active shape models - their training and application. 61(1):38–59, 1995.
6. D. Cremers, N. Sochen, and C. Schnoerr. Towards recognition-based variational segmentation using shape priors and dynamic labeling. In *Proc. of Scale-Space*, pages 388–400, 2003.
7. F. de la Torre and M. J. Black. Robust principal component analysis for computer vision. In *Proc. IEEE Intern. Conf. on Computer Vision*, volume I, pages 362–369, 2001.
8. M. Fussenegger, R. Deriche, and A. Pinz. A multiphase level set based segmentation framework with pose invariant shape priors. In *Proc. of the Asian Conference on Computer Vision*, pages 395 – 404, 2006.
9. P. Hall, D. Marshall, and R. Martin. Incremental eigenanalysis for classification. In *Proc. British Machine Vision Conf.*, volume I, pages 286–295, 1998.
10. P. Hall, D. Marshall, and R. Martin. Merging and splitting eigenspace models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(9):1042–1049, 2000.
11. H. Hotelling. Analysis of a complex of statistical variables with principal components. *Journal of Educational Psychology*, 24:417–441, 1933.
12. A. Leonardis and H. Bischof. Robust recognition using eigenimages. *Computer Vision and Image Understanding*, 78(1):99–118, 2000.
13. Y. Li. On incremental and robust subspace learning. *Pattern Recognition*, 37(7):1509–1518, 2004.
14. S. J. Osher and J. A. Sethian. Fronts propagation with curvature depend speed: Algorithms based on Hamilton-Jacobi formulations. In *Journal of Comp. Phys.*, volume 79, pages 12–49, 1988.
15. R. Rao. Dynamic appearance-based recognition. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pages 540–546, 1997.
16. M. Rousson and N. Paragios. Shape priors for level set representations. In *Proc. of European Conf. of Computer Vision*, volume 2351 of LNCS, pages 78–92, 2002.
17. S. Roweis. EM algorithms for PCA and SPCA. In *Proc. Conf. on Neural Information Processing Systems*, pages 626–632, 1997.
18. D. Skočaj, H. Bischof, and A. Leonardis. A robust PCA algorithm for building representations from panoramic images. In *Proc. European Conf. on Computer Vision*, volume IV, pages 761–775, 2002.
19. D. Skočaj and A. Leonardis. Weighted and robust incremental method for subspace learning. In *Proc. IEEE Intern. Conf. on Computer Vision*, volume II, pages 1494–1501, 2003.
20. M. E. Tipping and C. M. Bishop. Probabilistic principal component analysis. *Journal of the Royal Statistical Society B*, 61:611–622, 1999.
21. L. Xu and A. L. Yuille. Robust principal component analysis by self-organizing rules based on statistical physics approach. *IEEE Trans. on Neural Networks*, 6(1):131–143, 1995.