

# A Robust Matching Algorithm for Active Appearance Models

Reinhard Beichel<sup>1</sup>, Horst Bischof<sup>1</sup>, Georg Langs<sup>1,2</sup> and Milan Sonka<sup>3</sup>

<sup>1</sup>Institute for Computer Graphics and Vision, Graz University of Technology,  
Inffeldgasse 16/2, A-8010 Graz, Austria

<sup>2</sup>Pattern Recognition and Image Processing Group, Vienna University of Technology,  
Favoritenstr. 9, A-1040 Vienna, Austria

<sup>3</sup>Department of Electrical and Computer Engineering, The University of Iowa,  
Iowa City, IA 52242 USA  
Email: beichel@icg.tu-graz.ac.at

**Abstract.** This paper proposes a method for robustly matching active appearance models (AAMs) on images with gross disturbances (outliers). The method consists of two steps. First, an initial residual is calculated by comparing model and image appearance, and modes of the residual are analyzed. Second, all possible mode combinations are tested by evaluating an objective function. The objective function allows the selection of an outlier-free mode combination. Experiments demonstrate the ability of the robust matching method to successfully cope with outliers – compared to standard AAM matching, no degeneration of the model during matching occurs.

## 1 Introduction

Active Appearance Models (AAMs), developed by Cootes et al. [1, 2], have proven their usefulness for a wide variety of different computer vision problems like the segmentation and interpretation of faces or the tracking of objects (see [2] for an overview). Especially in medicine, AAMs are frequently used for image analysis tasks like the segmentation of cardiac MRI data [3, 4], diaphragm segmentation in CT data [5], the segmentation of ventricles of the human brain [1], or the segmentation of metacarpals in radiographs [6], to give a few examples. One drawback is that AAM matching frequently fails in cases where object appearance is significantly changed due to gross disturbances (e.g., imaging artifacts, implants, etc.) or if parts of the target object are occluded. In this paper we propose a novel robust AAM matching algorithm that is capable of matching AAMs correctly even when gross outliers perturb object appearance.

Robustness is an important issue in computer vision. However, in the case of AAMs, it has not been addressed extensively. Attempts to make AAMs more robust have been reported by Edwards et al. [7], Stegmann et al. [6], and Gross et al. [8]. Edwards et al. [7] proposed to learn the usual residual magnitudes during the training phase of the AAM. Pixels from the input image with a deviating residual are ignored during model matching. Stegmann et al. [6] and Gross et al.

[8] use the same principle – the quadratic error measure is replaced by a robust error measure. Basically, all three approaches discard outliers based on the magnitude of the residual. However, it has been completely overlooked that a large residual during AAM matching is not an information that should be discarded a priori. For example, the error might be due to an initial model displacement. In this case, the residual provides valuable information for positioning the AAM. If discarded, a slower convergence or a complete failure to segment image data might result.

## 2 Methods

The robust AAM matching method presented below builds on the AAM framework described by Cootes et al. [2]. During standard AAM matching, updates of model parameter vector  $\mathbf{p}$  are obtained by evaluating

$$\delta\mathbf{p} = - \left( [\mathbf{J}(\mathbf{r})^T \mathbf{J}(\mathbf{r})]^{-1} \mathbf{J}(\mathbf{r})^T \right) \mathbf{r}(\mathbf{p}) = -\mathbf{R}\mathbf{r}(\mathbf{p}) \quad (1)$$

based on the observed actual residual  $\mathbf{r}$ , calculated by comparing the gray-values of the model and the image data underneath [2]. In Eq. (1),  $\mathbf{J}(\mathbf{r}) = \partial\mathbf{r}/\partial\mathbf{p}$  denotes the Jacobian of  $\mathbf{r}$ . To increase the robustness of AAMs to outliers, “misleading” coefficient updates (incorrect  $\delta\mathbf{p}$ ) must be avoided. If the outliers are known, Eq. (1) can be adjusted accordingly. Let  $\mathbf{v}_{sel} = (v_1, \dots, v_M)^T$  be the selection vector with respect to the residual  $\mathbf{r} = (r_1, \dots, r_M)^T$  with the property that  $v_i = 1$  if  $r_i$  is an inlier and  $v_i = 0$  if  $r_i$  belongs to outliers. Then the rows of the Jacobian  $\mathbf{J}(\mathbf{r})$  are rearranged into the vector  $\mathbf{J}(\mathbf{r})_{sel}$ .  $\mathbf{J}(\mathbf{r})_{sel}$  denotes the components of  $\mathbf{J}(\mathbf{r})$  for which  $\mathbf{v}_{sel}$  is equal to one. According to Eq. (1), a new matrix  $\mathbf{R}_{sel}$  and a new parameter update vector  $\delta\mathbf{p}_{sel}$  are calculated. By using  $\delta\mathbf{p}_{sel}$  instead of  $\delta\mathbf{p}$ , only inliers are used for the update of model parameters  $\mathbf{p}$  during matching. A degeneration of the model, due to outliers, can be avoided. Note, that  $\mathbf{R}_{sel}$  has to be recalculated in each iteration of the matching procedure, since the residual  $\mathbf{r}(\mathbf{p})$  changes during matching. The crucial step is the classification of the residual into inliers and outliers. Since classification of outliers only according to the magnitude of the residual  $\mathbf{r}(\mathbf{p})$  is not sufficient, a novel robust matching algorithm is proposed:

- a) **Initialize the AAM:** Initialize the AAM with the parameter vector  $\mathbf{p}_0$  based on an initial estimate.
- b) **Calculate an initial residual:** Calculate the initial residual  $\mathbf{r}(\mathbf{p}_0)$  by roughly aligning the texture vector of the model  $\mathbf{g}_m$  to the image texture vector  $\mathbf{g}_i$  (model frame):

$$\mathbf{r}(\mathbf{p}_0) = \mathbf{g}_i - a(\mathbf{g}_m), \quad (2)$$

where  $a(\mathbf{g})$  denotes the alignment function.

- c) **Mean shift based mode detection:** To partition the residual  $\mathbf{r}(\mathbf{p}_0)$  into modes, the mean shift algorithm is utilized [9]. Mode boundaries are detected

by finding the valleys between residual modes. This is done by following the mean shift iteration

$$\mathbf{m}_{h,G}(\mathbf{x}) = \mathbf{x} - \left[ \sum_{i=1}^n g \left( \left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right) \mathbf{x}_i \right] / \left[ \sum_{i=1}^n g \left( \left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right) \right] \quad (3)$$

where  $h$  is the bandwidth parameter of the radial symmetric kernel  $G$  with the profile function  $g$ . The found modes are stored in the set  $M_X = \{\mathcal{M}_1, \dots, \mathcal{M}_o\}$  where each mode  $\mathcal{M}_i$  is limited by the corresponding valley points.

- d) Evaluate mode combinations:** To find a mode combination which is only based on inlier modes, an objective function is used. The objective function is evaluated for the set  $S$  of all possible mode combinations with at least one mode selected:  $S = \wp(M_X) \setminus \{\emptyset\} = \{\mathcal{S}_1, \dots, \mathcal{S}_l\}$ , where  $\wp$  denotes the power set and  $l = |S| = 2^o - 1$ ;  $o$  is the number of modes. An evaluation step of a mode combination  $\mathcal{S}_i$  consists of:

1. A conventional AAM matching with selective parameter updates is applied. A selection vector  $\mathbf{v}_{sel}$  is generated and utilized for calculating  $\mathbf{r}_{sel}$ ,  $\mathbf{R}_{sel}$  and  $\delta\mathbf{p}_{sel}$ , respectively. The generation of  $\mathbf{v}_{sel}$  is based on an estimate  $\tilde{\mathbf{r}}$  for the residual  $\mathbf{r}$ .  $\tilde{\mathbf{r}}$  is calculated using Eq. (2) and the rough alignment function  $a(\mathbf{g})$ . The components of  $\mathbf{v}_{sel}$  are set to one, if the corresponding value in  $\tilde{\mathbf{r}}$  is covered by the actual mode selection under evaluation.

In addition, a z-score function  $z(\mathbf{x}) = [\mathbf{x} - \mu(\mathbf{x})\mathbf{i}]/\sigma(\mathbf{x})$  is used for the alignment of the image and the model texture vectors

$$\mathbf{r}(\mathbf{p}) = z(\mathbf{g}_i(\mathbf{p})) - z(\mathbf{g}_m(\mathbf{p})) \quad , \quad (4)$$

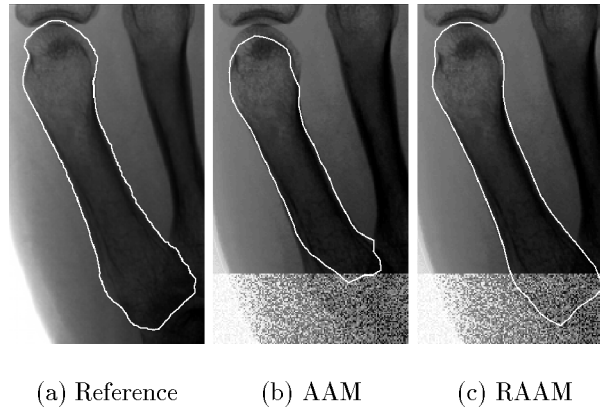
where the mean of the components of vector  $\mathbf{x}$  is denoted as  $\mu(\mathbf{x})$ , the standard deviation as  $\sigma(\mathbf{x})$  and the unit vector as  $\mathbf{i}$ . Based on  $\mathbf{r}(\mathbf{p})$  and  $\mathbf{v}_{sel}$ ,  $\mathbf{r}_{sel}(\mathbf{p})$  is generated by taking only components of  $\mathbf{r}(\mathbf{p})$  for which  $\mathbf{v}_{sel}$  is equal to one and used in combination with  $\mathbf{R}_{sel}$  for a parameter update:  $\delta\mathbf{p}_{sel} = -\mathbf{R}_{sel}\mathbf{r}_{sel}(\mathbf{p})$ .

2. Evaluation of the objective function  $Q(\mathcal{S}_i) = Y_{max}/(1 + |X_{max}|)$  where  $Y_{max}(\mathbf{r}_{\mathcal{S}_i})$  denotes the maximum histogram value of the residual  $\mathbf{r}_{\mathcal{S}_i}$  (Eq. 4), and the residual corresponding to  $Y_{max}(\mathbf{r}_{\mathcal{S}_i})$  is denoted as  $X_{max}(\mathbf{r}_{\mathcal{S}_i})$ . The objective function measures the quality of the AAM matching result based on the selected mode combination  $\mathcal{S}_i$ . It is maximal if the residual vector  $\mathbf{r}$  is equal to  $\mathbf{0}$ , the zero-vector.
- e) Select the final mode combination:** Finally, the best mode combination  $\mathcal{S}_j$  is selected:  $Q(\mathcal{S}_j) = \max_{i=1, \dots, l} \{Q(\mathcal{S}_i)\}$ . The AAM matching obtained by this selection is taken as final result.

### 3 Results

A comparison of AAM (with z-score alignment) and robust AAM (RAAM) is used for evaluation. A leave-one-out approach based on 40 metacarpal bone X-ray images of the small finger is applied. Model based segmentation of metacarpal

**Fig. 1.** Example of segmentation results. (a) Undisturbed image with reference tracing. (b) AAM and (c) RAAM matching result on disturbed image (partially additive noise, normally distributed within the range of  $[0, 200]$ ). A histogram equalization was applied for a better visualization.



bone images is of particular interest for the automated assessment of rheumatoid arthritis [10]. Reference tracings were made manually by a physician, and landmarks were placed automatically by using the method proposed by Thodberg [11]. The gray-values of the images were scaled to  $[0, 255]$ . Image data is disturbed artificially (see Fig. 1). RAAM parameters were as follows:  $g(x) = \exp(-x/2)$  and  $h = 8$ . For model building, only eigen vectors corresponding to 99% of the largest eigen values have been used. Matching performance is measured by the relative overlap error  $E_{rol\%} = 100\%( |A_{ref} \oplus A_{model}| ) / |A_{ref}|$  where  $A_{ref}$  and  $A_{model}$  are object masks for the reference segmentation and the matched model, respectively. The operator  $\oplus$  denotes the XOR operation between masks and the number of object pixels is denoted as  $|A|$ . Both, AAM and RAAM matching are started with identical initial model parameters.

The mean and standard deviation of the relative overlap error was  $12.87 \pm 2.43\%$  for the AAM and  $5.25 \pm 3.71\%$  for the RAAM. The occluded part was excluded from error calculation. The matched models are shown in Fig. 1. The AAM (Fig. 2(b)) is severely influenced by the added noise. In comparison, the RAAM (Fig. 2(c)) does not show such a behavior and delivers a good match.

## 4 Discussion and Conclusion

A novel robust AAM matching method was presented and compared to the standard AAM matching algorithm. It was shown, using a leave-one-out approach on 40 cases, that the robust method successfully copes with outliers and shows no degeneration of the border accuracy. Compared to other methods, no assumptions regarding “normal” residuals are made. This translates into a higher

flexibility regarding types of disturbances that can be handled without adjusting the algorithm.

## Acknowledgment

This work was supported in part by the Austrian Science Fund (FWF) under Grants P14897-N04, P17066-N04, and P17083-N04; and by NIH-NHLBI R01-HL071809. We are grateful to Philipp Peloschek of the Department of Diagnostic Radiology, Medical University of Vienna for contributing segmented X-ray images of the small finger.

## References

1. Cootes TF, Taylor CJ. Statistical models of appearance for medical image analysis and computer vision. In: Sonka Milan, Fitzpatrick JMichael, editors. SPIE: Medical Imaging: Image Processing. vol. 4322; 2001. p. 236–248.
2. Cootes TF, Taylor CJ. Statistical models of appearance for computer vision. Division of Imaging Science and Biomedical Engineering, University of Manchester; 2004. Available at <http://www.isbe.man.ac.uk/~bim/refs.html>.
3. Stegmann MB. Analysis of 4D Cardiac Magnetic Resonance Images. Journal of The Danish Optical Society, DOPS-NYT 2001;(4):38–39.
4. Mitchell SC, Bosch JG, Lelieveldt BPF, van der Geest RJ, Reiber JHC, Sonka M. 3-D Active Appearance Models: Segmentation of Cardiac MR and Ultrasound Images. IEEE Trans Med Imaging 2002;21:1167–1178.
5. Beichel R, Gotschuli G, Sorantin E, Leberl F, Sonka M. Diaphragm dome surface segmentation in CT data sets: A 3D active appearance model approach. In: Sonka Milan, Fitzpatrick JMichael, editors. SPIE: Medical Imaging: Image Processing. vol. 4684; 2002. p. 475–484.
6. Stegmann MB, Fisker R, Ersbøll BK. Extending and applying active appearance models for automated, high precision segmentation in different image modalities. In: Austvoll Ivar, editor. Proc. 12th Scandinavian Conference on Image Analysis - SCIA 2001, Bergen, Norway. Stavanger, Norway: NOBIM; 2001. p. 90–97.
7. Edwards GJ, Cootes TF, Taylor CJ. Advances in Active Appearance Models. In: Proc. International Conference on Computer Vision; 1999. p. 137–142.
8. Gross R, Matthews I, Baker S. Constructing and Fitting Active Appearance Models With Occlusion. In: Proceedings of the IEEE Workshop on Face Processing in Video; 2004.
9. Comaniciu D, Meer P. Mean shift: a robust approach toward feature space analysis. IEEE Trans Pattern Anal Machine Intell 2002;24(5):603–619.
10. Langs G, Peloschek P, Bischof H. ASM Driven Snakes in Rheumatoid Arthritis Assessment. In: Proceedings of 13th Scandinavian Conference on Image Analysis, SCIA 2003, Goeteborg, Sweden, Lecture Notes in Computer Science, LNCS2749; 2003. p. 454–461.
11. Thodberg HH. Shape Modelling and Analysis Minimum Description Length Shape and Appearance Models. In: Information Processing in Medical Imaging - IPMI2003. Springer; 2003. p. 51–62.