

# Wedgelet Enhanced Appearance Models

Sune Darkner, Rasmus Larsen, Mikkel B. Stegmann and Bjarne K. Ersbøll

Informatics and Mathematical Modelling, Technical University of Denmark

Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, Denmark

Email: {rl,mbs,be}@imm.dtu.dk, sune@darkner.dk, <http://www.imm.dtu.dk/image/>

**Abstract**—Statistical region-based segmentation methods such as the Active Appearance Model (AAM) are used for establishing dense correspondences in images based on learning the variation in shape and pixel intensities in a training set. For low resolution 2D images correspondences can be recovered reliably in real-time. However, as resolution increases this becomes infeasible due to excessive storage and computational requirements. In this paper we propose to reduce the textural components by modelling the coefficients of a wedgelet based regression tree instead of the original pixel intensities. The wedgelet regression trees employed are based on triangular domains and estimated using cross validation. The wedgelet regression trees are functional descriptions of the intensity information and serve to 1) reduce noise and 2) produce a compact textural description. The wedgelet enhanced appearance model is applied to a case study of human faces. Compression ratios of the texture information of 1:40 is obtained without sacrificing segmentation accuracy notably, even at compression ratios of 1:150 fair segmentation is achieved.

## I. INTRODUCTION

The Active Appearance Model (AAM) framework [1] has since its introduction been applied successfully to segmentation of many types of deformable objects in images (e.g. faces, cardiac ventricles, brain structures [1]–[4]). It is based on the estimation of linear models of shape and texture variation by the use of principal components analysis of landmarks coordinates and pixel intensities and subsequent inference of model parameters from unseen images by a tangent plane approximation of the image manifold.

Modelling every pixel intensity is manageable for low-resolution 2D images. But moving to high-resolution 2D images, 3D and even 3D time-series, this approach is rendered at best very slow and at worst infeasible due to excessive storage and computational requirements.

In order to overcome this problem various alternatives to modelling the raw pixel intensities have been considered. Cootes et al. [5] used a sub-sampling scheme to reduce the texture model by a ratio of 1:4. The scheme selected a subset of the pixel intensities based on the ability of each pixel to predict corrections of the model parameters. When exploring different multi-band appearance representations Stegmann and Larsen [6] studied the segmentation accuracy of facial AAMs at different scales in the range  $10^3 - 10^5$  pixels obtained by pixel averaging.

In this paper we will take the path of using linear functional descriptions of the underlying intensity patterns and carrying out truncated principal component decomposition of the parameter set of the functionals in order to extract a texture model. These parameter sets will typically be of much lower

dimensionality than the number of pixels or voxels in the images. In particular, we will use linear functionals based on a wedgelet basis representation of the texture.

Donoho [7] suggested the wedgelet representation for the texture as a means of edge detection and image compression. An image is represented by a collection of dyadically organised indicator functions with a variety of locations, scales and orientations. The classification and regression tree (CART) algorithm [8] uses sequential binary splitting of the spatial domain parallel to the coordinate axes, with splits allowed at every data point. In contrast to this the wedgelet regression tree obey special constraints. Only dyadic partitioning (i.e. recursive midpoint splitting) is allowed, with the added feature that at each terminal node a set of affine splits are also applicable. The wedgelet tree is a quadtree [9] with terminal nodes being either a dyadic (degenerate wedgelet) or an affinely split dyadic (non-degenerate wedgelet). The constrained splitting leads to fast algorithms. Within each resulting image terminal node (wedge or square) the pixel values are regressed to their mean value.

Another popular set of basis expansions for image texture are wavelet bases. Wavelets as well as wedgelets are able to represent piecewise continuous functions. This is an important property when modeling real world images. Wolstenholme and Taylor [10] incorporated a truncated Haar wavelet basis into the AAM framework and evaluated it on a brain MRI data set at a compression ratio of 1:20. Later, Stegmann, Forchhammer, and Cootes [11] further evaluated the use of the Haar wavelet as well as the Cohen-Daubechies-Feauveau [12] wavelet family in the AAM framework. Compression rates of 1:40 without compromising segmentation accuracy were obtained on a set of face images.

We will generalize the wedgelet transform to triangulated domains (cf. triangulated quadtrees [13]). This has the major advantage of rendering the wedgelet representation independent of piece-wise affine warps of the triangulated domain. Such piece-wise affine warps are customarily chosen in AAM for their speed [4] and the triangulated wedgelet representation thus embraces this choice. The wedgelet transform results in a truncated change of basis for the texture and is represented by a regression tree. The regression tree is estimated using the minimization of the cross validation prediction error across the training set.

The segmentation accuracy in a wedgelet based AAM is evaluated for a case of human face segmentation using cross validation.

## II. ACTIVE APPEARANCE MODELS

AAMs establish a compact parameterization of object variability as learned from a training set by estimating a set of latent variables. The modeled object properties are usually shape and pixel intensities. The latter is hence forward denoted texture. By exploiting prior knowledge of the nature of the optimization space, these models of shape and texture can be rapidly fitted to unseen images, thus providing image interpretation through synthesis.

Training examples are defined by marking up each example image with points of correspondence (i.e. landmarks) over the set either by hand, or by semi- to completely automated methods. From these landmarks a shape model [14] is built. Further, given a suitable warp function a dense (i.e., per-pixel) correspondence is established between the convex hull of the landmarks in each training example. Thus allowing for modeling of texture variability.

Joint variability in shape and texture is modeled by a set of truncated principal components, estimated by an eigen-analysis of the dispersions of shape and texture across the training set. The shape examples are aligned to a common mean using a Generalized Procrustes Analysis (GPA) [15], [16] where all effects of translation, rotation and scaling are removed. The obtained Procrustes shape coordinates are subsequently projected into the tangent plane to the shape manifold, at the pole given by the mean shape. The texture examples are warped into correspondence using a piece-wise affine warp and subsequently sampled from this shape-free reference. Typically, this geometrical reference shape is the Procrustes mean shape.

Let  $\mathbf{s}_i = \text{vec}\{x_{ijk}\}$ ,  $i = 1, \dots, I$ ,  $j = 1, \dots, J$ ,  $k = 1, \dots, K$  be  $J$  landmarks coordinates in  $K$  dimensions sampled from a training set of  $I$  images, and let  $\mathbf{t}_i = \text{vec}\{y_{ilm}\}$ ,  $l = 1, \dots, L$ ,  $m = 1, \dots, M$  be pixel intensities sampled at  $L$  sites in  $M$  color components for the same  $I$  training images. Furthermore, let  $\bar{\mathbf{s}}$  and  $\bar{\mathbf{t}}$  denote the mean shape and texture. Synthesized examples are parameterized by  $\boldsymbol{\theta}$  and generated by

$$\begin{aligned} E\{\mathbf{s}\} &= \bar{\mathbf{s}} + \boldsymbol{\Phi}_s \boldsymbol{\theta} \\ E\{\mathbf{t}\} &= \bar{\mathbf{t}} + \boldsymbol{\Phi}_t \boldsymbol{\theta} \end{aligned} \quad (1)$$

where  $\boldsymbol{\Phi}_s$  and  $\boldsymbol{\Phi}_t$  contain the first  $p$  eigenvectors of the estimated joint dispersion matrix of the shape and texture vectors,  $\mathbf{s}_i$  and  $\mathbf{t}_i$ . Eq. (1) constitutes the appearance model.

In order to infer the parameters  $\boldsymbol{\theta}$  as well as the four scalar parameters,  $\boldsymbol{\psi}$  – scale, orientation, and translation – of a 2D Euclidean similarity group, of a previously unseen image a Gaussian error model between model and pixel intensities is assumed. Furthermore, a linear relation between changes in parameters and difference between model and image pixel intensities  $\Delta \mathbf{t}$  is assumed, i.e.

$$\Delta \mathbf{t} = \mathbf{X} \begin{bmatrix} \Delta \boldsymbol{\psi} \\ \Delta \boldsymbol{\theta} \end{bmatrix}. \quad (2)$$

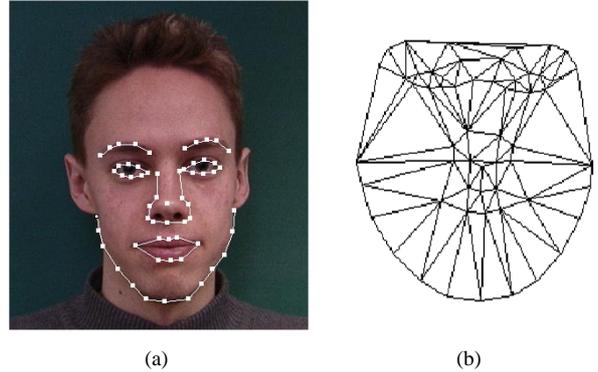


Fig. 1. (a) Human face annotated with 58 landmarks. (b) Mean shape.

$\mathbf{X}$  may be estimated by weighted averaging over perturbations of model parameters and training examples. For an in depth description of AAM and the software implementation used the reader is referred to [1], [4], respectively.

The relation in Eq. (2) is inverted using the least squares solution

$$\begin{bmatrix} \widehat{\Delta \boldsymbol{\theta}} \\ \widehat{\Delta \boldsymbol{\psi}} \end{bmatrix} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \Delta \mathbf{t} = \mathbf{Q} \Delta \mathbf{t}. \quad (3)$$

The computational problem lies in the repeated application of this relation in the innermost loop of the fitting algorithm.  $\mathbf{Q}$  is a non-sparse matrix of dimensions  $(p+4) \times (LM)$ .  $LM$  increase exponentially with spatial dimension. To reduce the computational burden we propose to use a truncated basis for the representation of the pixel intensities. This introduces the added overhead of transforming between image pixel intensities and this new representation. However, if this transform is based on a sparse matrix, as is the case with wavelet and wedgelet transforms the computational burden can be considerably reduced.

## III. WEDGELET DECOMPOSITION

The wedgelet approach is a way of representing images locally, orientation adaptively and at the appropriate scale. It involves a very simple basis used at different scales. The formulation in [7] is for a dyadic domain but the nature of the AAM imposes a shift of basis from dyadic to triangulated domains.

For each node in the wedgelet tree we may consider the following three wedgelets types:

- 1) a degenerate wedgelet, this is a terminal node without an affine split, cf. Fig. 2(b) and 2(f);
- 2) a non-degenerate wedgelet, this is a terminal node with an affine split, cf. Fig. 2(c) and 2(g);
- 3) an interior node corresponding to a step through scale space, cf. Fig. 2(d) and 2(h).

The wedgelet decomposition is seen to be embedded into a quadtree structure. In this structure the templates are the nodes and the step through scale space the branches. Furthermore, the terminal nodes are all either degenerate or non-degenerate

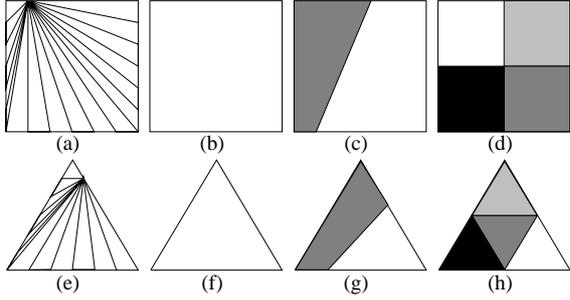


Fig. 2. Templates on the dyadic domain (top row) and the triangulated domain (bottom row). Plots (a) and (e) show the applicable affine split beginning at a particular perimeter point.

wedgelets. The resulting structure is a regression tree [8]. The corresponding regression model predicts the pixel intensity - greyscale, RGB or other - at the  $l$ th image coordinate  $\mathbf{x}_l$  in the  $i$ th image,  $y_{il}$ , with a constant  $\mu_{ir}$  in each region  $r$  for each image

$$f_i(\mathbf{x}_l) = \sum_r \mu_{ir} I\{\mathbf{x}_l \in r\}. \quad (4)$$

$I\{\mathbf{x}_l \in r\}$  is an indicator function returning 1 if  $\mathbf{x}_l$  belongs to  $r$ . Note that for notational simplicity we have dropped the index  $m$  for color component on the intensity values,  $y_{il}$ . For  $M > 1$ ,  $y_{il}$  is a vector of intensities at the  $M$  color components modelled.

For the sum of squared error loss criterion  $\sum_l \|(y_{il} - f_i(\mathbf{x}_l))\|^2$  it is easy to see that that optimal  $\mu_{ir}$  is just the average of  $y_{il}$  in region  $r$ ,  $\hat{\mu}_{ir} = \text{ave}(y_{il} | \mathbf{x}_l \in r)$ .

The optimal partitioning is found by a bottom-up approach. For each triangle at each level we seek the model that minimizes the  $C$ -fold cross-validation estimate of the prediction error across all affine splits/no split.

Let  $\kappa : \{1, \dots, I\} \mapsto \{1, \dots, C\}$  be an indexing function that indicates the partition to which training object (image)  $i = 1, \dots, I$  is allocated by randomization, and denote by  $\hat{s}^{-\kappa(i)}$  the split estimated with the  $\kappa(i)$ 'th part removed.  $\hat{a} = a(\hat{s}^{-\kappa(i)})$  and  $\hat{b} = b(\hat{s}^{-\kappa(i)})$  are the regions resulting from splitting a dyadic or a triangle by this affine split  $\hat{s}^{-\kappa(i)}$ , and  $\hat{c}$  is the entire dyadic or triangle. Furthermore, let the regression parameters from the  $i$ th image resulting from applying the split  $\hat{s}^{-\kappa(i)}$  be

$$\hat{\mu}_{ir} = \text{ave}(y_{il} | \mathbf{x}_l \in r), \quad r \in \{\hat{a}, \hat{b}, \hat{c}\}. \quad (5)$$

Then the cross validation errors become

$$\begin{aligned} \text{CVE}_{\text{split}} = & \sum_{i=1}^I \left[ \sum_{\mathbf{x}_l \in \hat{a}} \|y_{il} - \hat{\mu}_{i\hat{a}}\|^2 \right. \\ & \left. + \sum_{\mathbf{x}_l \in \hat{b}} \|y_{il} - \hat{\mu}_{i\hat{b}}\|^2 \right] \end{aligned}$$

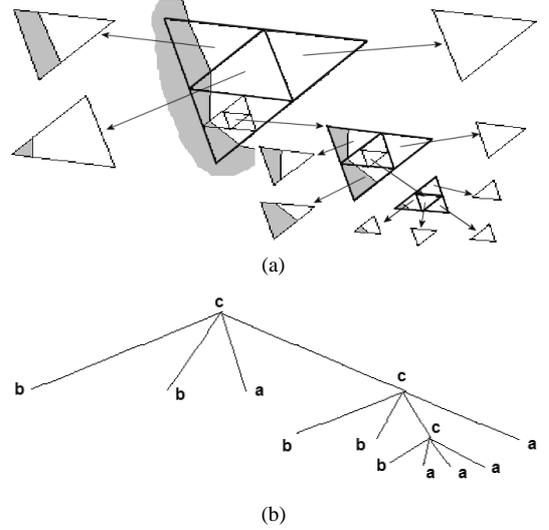


Fig. 3. A representation of (a) a binary image and (b) the resulting tree structure; the tree nodes are of types  $a$  degenerate,  $b$  non-degenerate,  $c$  interior.

$$\text{CVE}_{\text{no split}} = \sum_{i=1}^I \sum_{\mathbf{x}_l \in \hat{c}} \|y_{il} - \hat{\mu}_{ic}\|^2 \quad (6)$$

The optimal split/no split cross validation errors for a triangle and its three siblings are then compared to those for their parent in order to determine if a non-degenerate or a degenerate wedgelet should be declared or if the four siblings should be merged into a triangle or a dyadic at the next higher (parent) level.

In order to be able to control the compression ratio obtained, we add a complexity penalty to the error term over which we carry out cross validation. This complexity penalty is proportional to the image variance,  $\sigma^2$ , the area of the root triangle,  $A_{\text{root}}$  and inverse proportional to the area of the triangle under consideration,  $A_{\text{triangle}}$ , i.e.

$$\text{CP}(\lambda) = \lambda^2 \cdot \sigma^2 \frac{A_{\text{root}}}{A_{\text{triangle}}}. \quad (7)$$

In comparison Donoho [7] proposed a similar complexity penalized residual sum of squares criterion for the case of wedgelet compression over a single image, i.e.

$$\text{CPRSS}(\mathcal{P}, \lambda) = \sum_l \|y_l - f(\mathbf{x}_l; \mathcal{P})\|^2 + \lambda^2 \text{card}(\mathcal{P}) \quad (8)$$

where  $f$  is the regression model in Eq. (4),  $\mathcal{P}$  is the partition, and  $\text{card}(\mathcal{P})$  is the cardinality of  $\mathcal{P}$ . Our complexity penalty in Eq. (7) uses a relative area weighting to compensate for local size differences across the training set.

The optimization over all affine splits is conducted as an exhaustive search over a discretization (cf. Figs. 2(a) and 2(e)) corresponding to the pixel size. The indexing of pixels within a triangle and computation of areas are conveniently done by the use of barycentric coordinates [17] (cf. the appendix). Fig. 3 shows how a result on a binary image might look.

#### IV. USING WEDGELETS WITH AAM

When working on an AAM, an initial triangulation is available from the annotation and Delaunay triangulation on the mean shape (cf. Fig. 1(b)). This initial collection of triangles will be the root of the tree. From here each branch will be equivalent to the type of tree shown in Fig. 3.

After having grown a common wedgelet tree as described in the previous section for the training set, we can proceed to train the wedgelet enhanced appearance model. As before the shape is described by the landmark coordinates, i.e.  $s_i = \text{vec}\{x_{ijk}\}$ . However, for the texture we substitute the wedgelet coefficients for the original intensity samples, i.e. we use  $t_i = \text{vec}\{\hat{\mu}_{irm}\}$ ,  $r = 1, \dots, R$ ,  $m = 1, \dots, M$ .  $R$  is the number of regions used by the wedgelet tree.

The active part of the wedgelet enhanced appearance model is trained using Eqs. (2) and (3). The resulting  $Q$  matrix has dimensions  $(p + 4) \times (RM)$ . In each iteration of model to image fitting this is also the number of multiplications and additions to be carried out. However, we must also take into consideration the added overhead of calculating the wedgelet representation of the image patch that is covered by the model in each iteration. Letting every pixel fall into just one region of the wedgelet tree the computational load in computing the regional means is essentially  $LM$  additions (and  $R$  multiplications). Let the compression ratio of the wedgelet representation be  $\lambda = L/R$ , then the reduction in computational load is better than a factor

$$\frac{LM + (p + 4)RM}{(p + 4)LM} = \frac{1}{p + 4} + \frac{1}{\lambda} \quad (9)$$

As we shall see compression ratios of 1:100 are achievable and therefore the expression is usually dominated by the first term yielding a reduction factor equal to the number of parameters in the model.

We have here ignored the computational load related to warping the image patch to the model. This is conveniently and fast accomplished using modern graphics hardware [4].

#### V. RESULTS AND DISCUSSION

Data for the experiments are an image database of 37 annotated faces. Each image is a  $640 \times 480$  RGB image of a face of an adult human. The data set consists of images of 7 female and 30 male faces. Each face has been manually annotated with 58 corresponding landmarks (see [4] for a detailed analysis).

Fig. 4 shows the result of compressing a single image using the approach described above. The original Delaunay triangulation of the face data set (cf. Fig. 1(b)) is subdivided using the penalized complexity criterion in Eqs. (6) and (7) above. In Fig. 4 wedgelet compression at ratios 1:3 and 1:40 for a single face are shown.

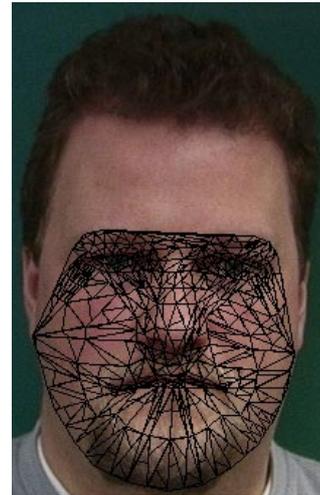
In Fig. 5 the combined principal components of the texture descriptors  $\mu_{irm}$  and the tangent space aligned landmark coordinates are shown. Comparing Figs. 5(a) and 5(c) with a compression ratio of 1:3 and Figs. 5(b) and 5(d) with a ratio



(a)



(b)



(c)

Fig. 4. Images compressed using triangulated wedgelets. (a) 1:3 ratio and (b) 1:40 using the triangulation from Fig. 1(b). (c) Result (b) superimposed with the subdivided original mesh.

of 1:40, we see that the the first principal components contain the same variations independent of the compression ratios. This leads us to conjecture that the wedgelet representation indeed dismisses irrelevant noise components and retains original signal information.

In Fig. 6 examples of the segmented facial features using a wedgelet enhanced AAM with compression ratios 1:3 and 1:40 are shown. Again the results are indistinguishable.

In order to compare the segmentation quality of the wedgelet enhanced AAM and its ability to perform with increasing compression ratio we have conducted a cross-validation study across the construction of both the wedgelet decomposition of the training set and the construction of the AAM. The average landmark distance from model to ground truth has been used to measure the performance.

The models were initialized using a displacement of 10% of the width and the height of the mean AAM model from the optimal position in the  $x$  and  $y$  direction.

For face segmentation we show compression rates of 1:150 with a decrease in segmentation accuracy of 8%. For brain segmentation the wavelet compressed AAMs reports compression rates of 1:20 with a decrease in segmentation accuracy of 7% [10]; and a compression rate of 1:40 with a decrease in segmentation accuracy of 8% [11]. These experiments were carried out on grayscale images.

As expected a slight decrease in performance is seen as the compression ratio increases (cf. Fig. 7(a)). Furthermore, using wedgelets yields good results all the way up to 1:150, (cf. Fig. 7(b)). Since the method proposed can easily be extended to 3-D, this makes it possible to apply the AAM onto high resolution medical images such MR or CT images both in 2 and 3 dimensions. The major differences between wedgelets and wavelets are in the synthesized image and in the computational time. The truncated wavelet representation yields nice and smooth synthetic images very pleasing to the eye. However, it should be noted that the purpose of the wedgelet enhanced appearance model is segmentation with minimum storage and computational cost and not image reconstruction. Therefore the model should not be evaluated on the "blockyness" of Fig. 4 and Fig. 5. Fig. 4 serves to demonstrate where in the images important information regarding segmentation is present. Fig. 5 demonstrates that the low order principal components of the uncompressed and the compressed data set are similar. On the other hand the truncated wavelet transform is computationally more demanding than the truncated wedgelet transform for the same number of basis functions. Wedgelets produce visually more coarse results than wavelets. However, they reduce the texture descriptor size and hereby reduce the computational cost and storage cost significantly due to the reduction of  $\mathbf{Q}$  in Eq. (3). Since the number of latent variables is almost unchanged, the size reduction of  $\mathbf{Q}$  is approximately the same as the texture compression ratio. This increases the speed and the storage cost of the AAM.

We have used the wedgelet based functional representation of the image intensity patterns in order to obtain a more

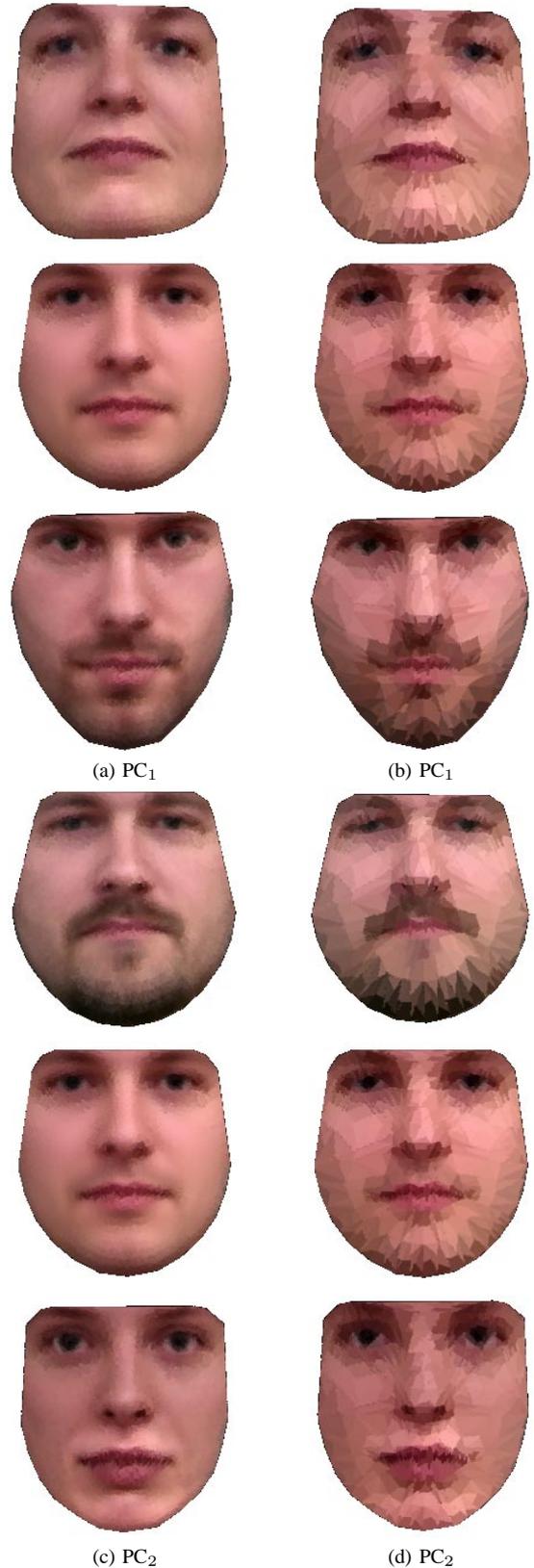


Fig. 5. 1st and 2nd principal components shown at +3 standard deviation, mean and -3 standard deviation for (a-b) wedgelet compression ratio 1:3, (c-d) wedgelet compression ratio 1:40.

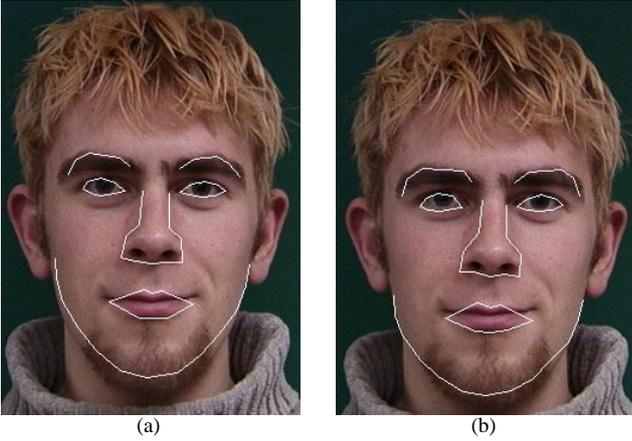


Fig. 6. (a) Wedgelet AAM using a wedgelet compression rate of 1:3, (b) 1:40.

parsimonious description of the texture. However, we can also think of such a functional representation as a way if regularising the principal components solution.

## VI. CONCLUSION AND FUTURE WORK

Different criteria have been reported in the AAM literature for choosing model complexity. In the original formulation a threshold at 95–98% was used on the cumulated variance described by the principal components. Alternatives include 1) "elbow" identification in scree-plots of covariance matrix eigenvalues [18]; 2) statistical comparison of covariance matrix scree-plot with scrambled data covariance scree-plots [19]; 3) probabilistic PCA modelling using AIC, BIC or cross validation [20]. For the compression AAM models the following methods have been applied. For the subsampled AAM [5] the subsampling is given by the union of the  $u\%$  elements (pixels) with largest absolute value in each column of  $Q$  (since large regression coefficients does not necessarily mean high significance for multiple regression this may not be a good idea). For the wavelet compression AAMs [10], [11] the criterion used is based on retaining as much variance across the training set as possible for a given compression ratio. The original wedgelet representation [7] uses a complexity penalized (CP) RSS criterion. For the multi object (image) situation we introduce the complexity penalized cross-validation error in Eqs. (4) and (5). The complexity penalty in Eq. (5) favours relative larger wedges/triangles (compensating for scaling differences/warps between images). The multiplication of the CP with the image variance compensates for (gain) differences between training images.

We have defined a 2D wedgelet transform on triangulated domains. We have demonstrated how cross-validation can be used to arrive at a truncated wedgelet representation of the texture in an active appearance model setting. The triangulated wedgelet transform embraces the triangulated domains used in AAMs in contrast to previous attempts based on wavelet transforms. The wavelet transforms require pixelated handling of boundaries because they are inherently based on

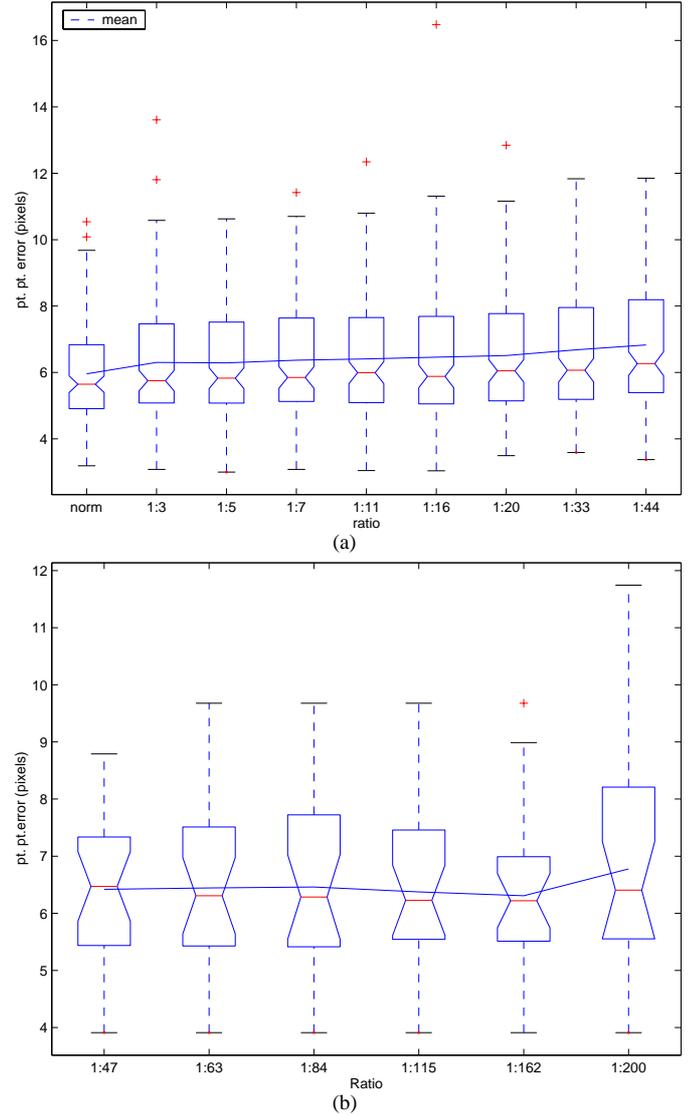


Fig. 7. Segmentation accuracy determined by cross validation across wedgelet compression and AAM segmentation. The abscissae are the ratio between the number of pixels and the number of wedges in the model. This may be interpreted as a compression ratio. The ordinate is the mean point to point prediction error evaluated at the original 58 landmarks for all faces in the cross validation study.

rectangular domains. The wedgelet scheme can readily be extended to higher dimensions and we aim to pursue this in near future. Also, additional constraints such as enforcing connecting edges between neighboring wedgelets as well as allowing perturbations of semi-landmarks introduced on the edges between their parent landmarks are considered.

In applying the wedgelet transform to ensembles of face images we arrive at compression rates up to 1:150 with only subtle degradation of the segmentation accuracy. Even higher compression rates will apply for 3D and 3D+time.

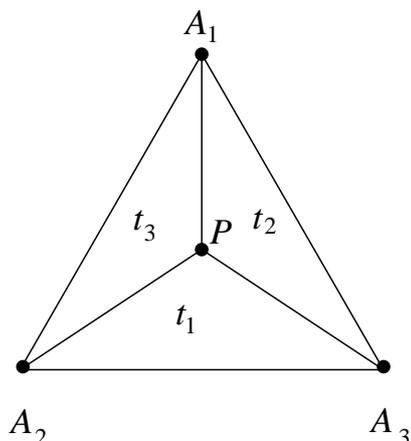


Fig. 8. Barycentric coordinates.

#### APPENDIX BARYCENTRIC COORDINATES

Homogeneous barycentric coordinates known from computer graphics are introduced [17]. A barycentric coordinate system is a local coordinate system for triangles in two dimensions, but can easily be extended to higher dimensions. Barycentric coordinates are triples of numbers  $(t_1, t_2, t_3)$  corresponding to masses placed at the vertices of a reference triangle  $\triangle A_1 A_2 A_3$ . These masses determine a point  $P$ , which is the geometric centroid of the three masses. The vertices of the triangle are given by  $A_1 = (1, 0, 0)$ ,  $A_2 = (0, 1, 0)$  and  $A_3 = (0, 0, 1)$ . The areas of  $\triangle A_2 P A_3$ ,  $\triangle A_1 P A_3$  and  $\triangle A_1 P A_2$  are proportional to  $t_1, t_2$  and  $t_3$ . For homogeneous barycentric coordinates the following is true  $t_1 + t_2 + t_3 = 1$  and every point  $P$  with coordinate  $c = (t_1, t_2, t_3)$  where  $0 \leq t_1, t_2, t_3 \leq 1$  lie within  $\triangle A_1 A_2 A_3$ . Moving from image coordinates to barycentric coordinates constitutes a shift of basis.

#### REFERENCES

- [1] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681–685, 2001.
- [2] S. Mitchell, B. Lelieveldt, R. Geest, J. Schaap, J. Reiber, and M. Sonka, "Segmentation of cardiac MR images: An active appearance model approach," in *Medical Imaging 2000: Image Processing, San Diego CA, SPIE*, vol. 1. SPIE, 2000.
- [3] S. C. Mitchell, B. P. F. Lelieveldt, R. J. van der Geest, H. G. Bosch,

- J. H. C. Reiber, and M. Sonka, "Multistage hybrid active appearance model matching: Segmentation of left and right ventricles in cardiac MR images," *IEEE Transactions on Medical Imaging*, vol. 20, no. 5, pp. 415–423, May 2001.
- [4] M. B. Stegmann, B. K. Ersbøll, and R. Larsen, "FAME – a flexible appearance modelling environment," *IEEE Transactions on Medical Imaging*, vol. 22, no. 10, pp. 1319–1331, may 2003.
- [5] T. F. Cootes, G. Edwards, and C. J. Taylor, "A comparative evaluation of active appearance model algorithms," in *BMVC 98. Proc. of the Ninth British Machine Vision Conf.*, vol. 2. Univ. Southampton, 1998, pp. 680–689.
- [6] M. B. Stegmann and R. Larsen, "Multi-band modelling of appearance," *Image and Vision Computing*, vol. 21, no. 1, pp. 61–67, Jan. 2003.
- [7] D. Donoho, "Wedgelets: Nearly minimax estimation of edges," *Annals of Statistics*, vol. 27, pp. 859–897, 1999.
- [8] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and regression trees*. Monterey, California: Wadsworth & Brooks/Cole advanced books & software, 1984, 358 pp.
- [9] R. A. Finkel and J. L. Bentley, "Quad trees: A data structure for retrieval on composite keys," *Acta Informatica*, vol. 4, pp. 1–9, 1974.
- [10] C. B. H. Wolstenholme and C. J. Taylor, "Wavelet compression of active appearance models," in *Medical Image Computing and Computer-Assisted Intervention, MICCAI, 1999*, pp. 544–554.
- [11] M. B. Stegmann, S. Forchhammer, and T. F. Cootes, "Wavelet enhanced appearance modelling," in *International Symposium on Medical Imaging 2004, San Diego CA, SPIE*. SPIE, 2004. [Online]. Available: <http://www.imm.dtu.dk/pubdb/p.php?2807>
- [12] A. Cohen, I. Daubechies, and J.-C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Comm. Pure and Applied Mathematics*, vol. 45, pp. 485–560, 1992.
- [13] M. W. Bern, D. Eppstein, and J. R. Gilbert, "Provably good mesh generation," in *Proc. 31st Symp. Foundations of Computer Science*, vol. I. IEEE, Oct. 1990, pp. 231–241.
- [14] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models – their training and application," *Computer Vision, Graphics and Image Processing*, vol. 61, no. 1, pp. 38–59, Jan. 1995.
- [15] J. C. Gower, "Generalized Procrustes analysis," *Psychometrika*, vol. 40, pp. 33–50, 1975.
- [16] J. M. F. ten Berge, "Orthogonal Procrustes rotation for two or more matrices," *Psychometrika*, vol. 42, pp. 267–276, 1977.
- [17] A. F. Möbius, *Der barycentrische Calcul, ein neues Hilfsmittel zur analytischen Behandlung der Geometrie / dargestellt und angewendet*. Leipzig, Germany, 1827.
- [18] T. Hastie, R. Tibshirani, and J. Friedman, *Elements of Statistical Learning: data mining, inference and prediction*. Springer-Verlag New York, 2001.
- [19] R. R. Paulsen, R. Larsen, S. Laugesen, C. Nielsen, and B. K. Ersbøll, "Building and testing a statistical shape model of the human ear canal," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2002, 5th Int. Conference, Tokyo, Japan.*, ser. Lecture Notes in Computer Science, vol. 2488. Berlin Heidelberg: Springer Verlag, 2002, pp. II–373–II–380.
- [20] R. Larsen and K. B. Hilger, "Probabilistic generative modelling," in *13th Scandinavian Conference on Image Analysis (SCIA), Gothenburg, Sweden*, ser. Lecture Notes in Computer Science, J. Bigün and T. Gustavsson, Eds., vol. 2749. Springer, jul 2003, pp. 861–868. [Online]. Available: <http://springerlink.metapress.com/openurl.asp?genre=article&issn=0302-9743&volume=2749&page=861>