Elastic Appearance Models

Mads Fogtmann Hansen
http://www.imm.dtu.dk/~mfh
Jens Fagertun
http://www.imm.dtu.dk/~jenf
Rasmus Larsen
http://www.imm.dtu.dk/~rl

DTU Informatic
Technical University of Denmark
Kgs. Lyngby, Denmark

AAM creates a shape model from set of training shapes by modeling them in a linear subspace. Linear generative shape models will in general not be able to capture complex shape variations as biological shapes rarely can be embedded in a linear subspace. This may result in folding shapes and poor registration/segmentation. Instead, we will assign an individual displacement to each point in the mean shape and construct a statistical local deformation prior which measures how well a given local deformation matches the local deformations in the training set using an energy density function. We use the Riemannian elasticity framework [3] to model this prior. Riemannian elasticity priors can capture complex deformations as they are locally rotation and translation invariant.

To protect against folding we want an energy density function, which approaches infinity when the approaches towards the black-hole deformation. This implies that the at least one of the eigenvalues of the strain tensor should go to infinity when the determinant of the Jacobian of the deformation gradient approaches zero. The Hencky strain tensor, $\boldsymbol{E}_0$, fulfils this requirement. Pennec *et al*. [3] proposed the *statistical Riemannian elasticity energy*, $\boldsymbol{E}_0 = \frac{1}{4}\|\text{vect}(\boldsymbol{E}_0 - \bar{\boldsymbol{E}}_0)\|^2_{\boldsymbol{\Sigma}^{-1}}$, where Σ and $\bar{\boldsymbol{E}}_0$ are calculated from a set of n previously observed deformations.

Similar to AAM, the EAM matches the appearance model to the template image using the sum of squared residuals with the addition of the statistical elasticity energy, i.e.

$$\underset{\boldsymbol{p},\Delta\boldsymbol{v}}{\text{argmin}}\|\boldsymbol{g}(\boldsymbol{p}) - \boldsymbol{T}\circ\boldsymbol{\phi}(\Delta\boldsymbol{v})\|^2 + \alpha_1\|\boldsymbol{p}\|^2_{C^{-1}} + \alpha_2 r(\Delta\boldsymbol{v}). \tag{1}$$

where r is the elasticity energy which is computed by integration of the energy density function.

The advantages of Eqn. 1 compared to the regular AAM formulation are that it cannot produce folding warps, and that it is not necessary to optimize rotation and translation separately due to local invariance. Eqn. 1 is not scale invariant which implies that Eqn. 1 cannot match the shape model to an image if the scale between them is too large or small. Global scaling effects can be filtered out of the shape displacements prior to calling the regularizer, i.e.

$$\underset{\boldsymbol{p},\Delta\boldsymbol{v}}{\text{argmin}}\|\boldsymbol{g}(\boldsymbol{p}) - \boldsymbol{T}\circ\boldsymbol{\phi}(\Delta\boldsymbol{v})\|^2 + \alpha_1\|\boldsymbol{p}\|^2_{C^{-1}} + \alpha_2 r\left(\frac{\Omega_T(\bar{\boldsymbol{v}})}{\Omega_T(\bar{\boldsymbol{v}}+\Delta\boldsymbol{v})}(\bar{\boldsymbol{v}}+\Delta\boldsymbol{v}) - \bar{\boldsymbol{v}}\right) \tag{2}$$

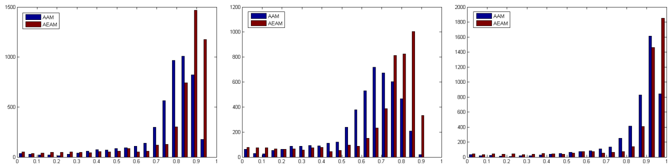
where $\Omega_T(\boldsymbol{v})$ computes the area of a shape defined by the triangulation T and the vertices $\boldsymbol{v}$.

1) Face labeling: This experiment used the subset of neutral facial expression images in the AR database [1] where annotation markup data were available from FGnet – in total 119 images. The performance of the methods was evaluated using leave-one-out cross-validation with the annotations as ground truth and the mean Euclidian point-to-point distance as the error measure. Table 1 presents the results for the shape (annotation) displacement error. From the results it can be seen that EAM labels the face on average more than 10 per cent better than AAM.

Method	Displacement error in pixel			
	Mean	25th percentile	75th percentile	97.5th percentile
AAM	4.8	3.1	4.7	15.2
EAM	4.3	3.1	4.5	8.8

Table 1: Displacement error, initialized with Viola-Jones face/eye detection.

2) Face labeling with poor initialization: In this experiment the AAM and EAM methods were initialized several time by regular perturbation of the Viola-Jones alignment (experiment 1) in intervals $\sqrt{2}[-20pixel, 20pixel]$, $[-15°, 15°]$ and $[90\%, 110\%]$ for diagonal translation, rotation and scaling, respectively. When making a poor initialization of the models it can be seen that EAM really outperforms AAM. By labeling the face on average 39 per cent better than AAM. Table 2 presents the displacement error results.



(a) Dice (b) Jaccard (c) Overlap

Figure 1: Histograms of the Dice, Jaccard and overlap coefficients for Corpus Callosum.

Method	Displacement error in pixel			
	Mean	25th percentile	75th percentile	97.5th percentile
AAM	13.1	3.9	16.8	52.8
EAM	8.0	3.3	7.8	34.0

Table 2: Displacement error, initialized with poor face detection.

3) Corpus Callosum segmentation: This experiment used 62 two dimensional MR images (with different subjects) of the mid-sagittal cross-section of the corpus callosum brain structure recorded at the Danish Reserach Centre for Magnetic Resonance, Hvidovre Hospital. Furthermore, each corpus callosum have manually been annotated with 72 landmarks by a clinician. The AAM and EAM methods were initialized several time by perturbation of the alignment between the mean shape and ground truth in the intervals $\sqrt{2}[-3pixel, 3pixel]$, $[-10°, 10°]$ and $[95\%, 105\%]$ for diagonal translation, rotation and scaling, respectively. The performance of the methods was evaluated with 6-fold cross-validation using the Dice, Jaccard and overlap coefficients to measure the quality of the segmentations. The results of the experiment are listed in Table 3 and visualized in Figure 1. As the Dice and Jaccard coefficients after matching are 0.89 and 0.81 for AAM and 0.90 and 0.82 for EAM using only the unperturbed ground truth alignments for initialization the large differences in Dice and Jaccard coefficients primarily indicate that the EAM is more robust and secondarily that is it more accurate. This conclusion is supported by the histograms of the coefficients shown in Figure 1.

Method	Mean		
	Dice	Jaccard	Overlap
AAM	0.75	0.63	0.85
EAM	0.80	0.70	0.86

Table 3: Means for the Dice, Jaccard and overlap coefficients for Corpus Callosum segmentation.

[1] A. M. Martinez and R. Benavente. The ar face database. Technical Report 24, FGNet, 1998.

[2] L. Pantoni, A. M. Basile, G. Pracucci, K. Asplund, J. Bogousslavsky, H. Chabriat, T. Erkinjuntti, F. Fazekas, J. M. Ferro, M. G. Hennerici, J. O´brien, P. Scheltens, M. C. Visser, L. O. Wahlund, G. Waldemar, A. Wallin, and D. Inzitari. Impact of age-related cerebral white matter changes on the transition to disability–the LADIS study: rationale, design and methodology. *Neuroepidemiology*, 24(1-2):51–62, 2004. ISSN 0251-5350.

[3] X. Pennec, R. Stefanescu, V. Arsigny, P. Fillard, and N. Ayache. Riemannian Elasticity: A Statistical Regularization Framework for Non-linear Registration. *Proceedings of the 8th International Conference on Medical Image Computing and Computer-Assisted Intervention (LNCS)*, 3750:943–950, 2005.