Dense Active Appearance Models Using a Bounded Diameter Minimum Spanning Tree

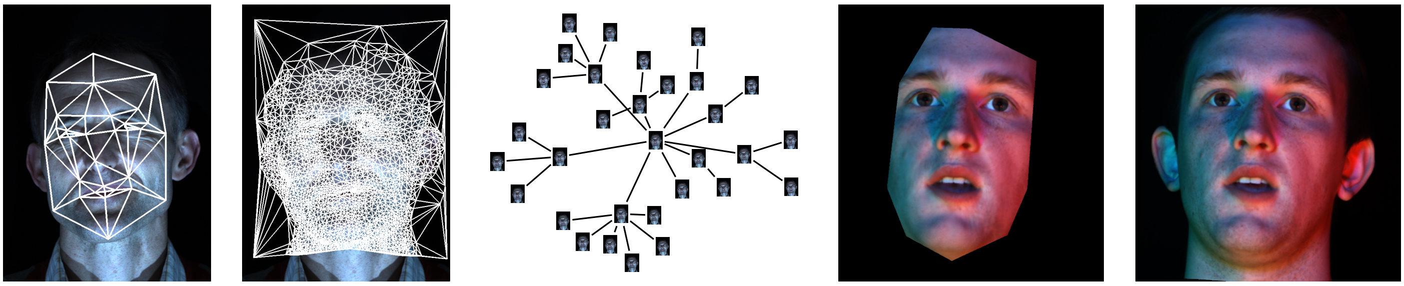
Robert Anderson[1]
ra312@cam.ac.uk

Bjorn Stenger[2]
bjorn.stenger@crl.toshiba.co.uk

Roberto Cipolla[1]
cipolla@eng.cam.ac.uk

[1] Department of Engineering
Cambridge University
Cambridge, UK

[2] Cambridge Research Laboratory
Toshiba Research Europe
Cambridge, UK



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

Figure 1: A sparse AAM (a), containing 37 vertices, is automatically refined to give a much denser AAM (b), containing 1000 vertices, by bringing all training images into joint alignment using a Bounded Diameter Minimum Spanning Tree (c). Synthesis of novel sequences using a sparse AAM (d) produces a blurred output that does not capture the same details as a dense AAM (e).

Active Appearance Models (AAMs) can provide an efficient method for synthesising novel video sequences. Currently AAMs are built using only a small number of vertices (<100) due to the time required to hand label training images, fig.1(a). This small number of vertices means that fine detail is not captured by the model as non-rigid deformation of the target occurs at a smaller scale than the model's mesh. In order to produce high quality synthesis the model must be densified by adding a much larger number of vertices, fig.1(b).

Previous attempts at automatic construction of AAMs [1, 7] have been encouraging but have not been demonstrated to capture the fine detail necessary for high resolution models. The key problem in this task is one of joint image alignment, as all training images must be brought into a dense correspondence. State-of-the-art results on this problem are currently achieved by methods such as that of Cootes *et al*. [2], however registering to a mean image results in some features not being correctly aligned. We propose a registration method which is based on a bounded diameter minimum spanning tree (BDMST), fig.1(c). In a similar way to other tree-based registration methods [3, 5] our approach aims to make each pairwise registration as accurate as possible by only aligning between similar images.

Our densification process consists of the three key components outlined below;

1. a joint alignment approach using a bounded diameter minimum spanning tree,
2. a method of optical flow refinement suited to regions which contain fine texture, and
3. a method for densifying an AAM given dense correspondences between training images.

Registration through a BDMST: To register all training images to a common frame we build a BDMST with one node for each image and where edge weights are determined by the residual error between two images after pairwise alignment. Pairwise registration is calculated between each pair of connected images in the tree in the form of an optical flow field. Each image is then registered to the base image (the image at the root of the tree) by concatenating flow fields along the path from the image to the tree root. This ensures that each pairwise registration is made between similar images, increasing the quality of the registration. Experimentally we find that a suitable diameter for the tree is 4, which trades off the maximum length of path from any image to the base image against the amount of difference between images between which pairwise registrations are computed.

Optical flow computation: To calculate pairwise registrations between images in the tree we use optical flow. To find an approximate initial flow we use the implementation of Liu [6]. This fails to align some fine detail in the images and we propose a refinement step to correct for this. We assume that the initial registration is approximately correct and so limit the range of displacements in our refinement to ± 15 pixels. We minimise an energy function consisting of a photoconsistency term and a regularisation term. To find a globally optimal solution we use a Markov Random Field formulation and in order to keep the number of labels for each pixel in the MRF to a reasonable number (31, instead of 961) we optimise horizontal and vertical displacement separately.

Mesh densification: Given a set of registered images we wish to densify the original AAM by adding additional vertices. We wish to add vertices in such a way that we can model the observed deformations between each image and the base image using as few vertices as possible. To do so we follow a similar approach to that used in the construction of digital terrain models. Instead of trying to minimise the difference between a scalar field (height) and its approximation given by interpolating over a mesh we wish to minimise the difference between a vector field (the flow from each training image to the base image) and its approximation interpolating on the mesh. To achieve this we use a greedy point insertion algorithm based on that of DeFloriani [4].

Results are shown which demonstrate that dense AAMs built using the proposed approach have an improved texture model compactness over the original AAMs. We also demonstrate qualitatively the improvement in synthesis resulting from using dense AAMs over sparse ones (fig.1(e) versus fig.1(d)). More information can be found in the paper and accompanying video.

[1] S. Baker, I. Matthews, and J. Schneider. Automatic construction of active appearance models as an image coding problem. *PAMI*, 26 (10):1380–1384, 2004.

[2] T. Cootes, C. Twining, V. Petrović, K. Babalola, and C. Taylor. Computing accurate correspondences across groups of images. *PAMI*, 32 (11):1994–2005, 2010.

[3] D. Cristinacce and T. Cootes. Facial motion analysis using clustered shortest path tree registration. *MLVMA Workshop (ECCV)*, 2008.

[4] L. De Floriani. A pyramidal data structure for triangle-based surface description. *IEEE Comput. Graph. Appl.*, 9(2):67–78, 1989.

[5] J. Hamm, D. Hye Ye, R. Verma, and C. Davatzikos. Gram: A framework for geodesic registration on anatomical manifolds. *Medical Image Analysis*, 14(5):633–642, 2010.

[6] C. Liu. Beyond pixels: Exploring new representations and applications for motion analysis. *Doctoral Thesis, MIT*, 2009.

[7] K. Ramnath, S. Baker, I. Matthews, and D. Ramanan. Increasing the density of active appearance models. *CVPR*, 2008.