Analysis of Colour Space Transforms for Person Independent AAMs

Tadas Baltrusaitis
University of Cambridge

Peter Robinson
University of Cambridge

Figure 1: Convergence of AAM fitting on unseen subjects tested on both IMM and Multi-PIE databases under different training conditions

1 Introduction

Statistical models of non-rigid deformable objects such as Active Appearance Models (AAM) are a popular means of both registration, tracking and synthesis of faces. Due to rapid fitting and good accuracy they are used extensively for facial expression tracking and analysis. A problem facing AAM based face tracking, is their inability to generalise well to unseen faces especially from unseen databases. One way to overcome this problem is through the combination of colour spaces, as some of them capture lighting variations, others pose or expression. In addition, this allows us to capture more variation in ethnicity, gender and age. Use of multiple databases gives us a better opportunity to create person, expression, and pose independent models. A problem arises because of the heterogeneity of the available databases due to use of different lenses, exposure times, external lightning or shadows etc. We describe an approach that leads to improved convergence of AAM fitting at close range when training a model on two different databases. In addition, this approach offers a substantial improvement when fitting images from unseen databases.

AAM fitting usually uses greyscale, RGB or HSV colour spaces. Like [Jonita et al. 2009], our approach is based on l1l2l3 colour space. It was shown that it has better fitting accuracy than RGB or greyscale, especially when fitting to images from unseen databases. Although, its effect on models trained on several databases and its effect on radius of convergence were not investigated in their paper.

2 Experiments

For fitting we use a 3-level multi-resolution AAM fitting algorithm outlined by Cootes et al. [Cootes et al. 2001] with photometric colour normalisation (global in the case of RGB, and per channel in the case of l1l2l3). In all of the experiments the model was initialised displaced from known optimum at no rotation, mean shape and appearance, and at correct scale (to deal with different zoom levels across two databases). We used 65 displacements for each test image, these were: \{-30, -20, -10, 0, 10, 20, 30\} pixels in both x and y directions, and for a better resolution close to optimum \{-7, -3, 3, 7\}. A fit was assumed to have converged if on termination of AAM fit the mean Point-to-Curve distance from the manually annotated image to the fit is less than 5 pixels.

We used two standard and publicly available databases for our experiments, namely IMM [Stegmann and Larsen 2002] and Multi-PIE [Gross et al. 2008]. All of the images used were of frontal pose under uniform frontal lighting conditions with two facial expressions per subject. They were divided into training and testing subsets in such way that training and testing images would never have the same subject in them. We chose 40 images from IMM for training and 34 for testing, and 40 from Multi-Pie for training and 40 for testing. Each image was labelled with 58 landmarks.

We carried out three experiments in total under three different colour spaces: training the model on one database and seeing the convergence rate on the same database but on unseen images (Figure 1b), the convergence rate on a different database (Figure 1a), and training the model on both databases and seeing the convergence rate on both databases but on unseen images (Figure 1c).

l1l2l3 model performs better than RGB and greyscale at short range on unseen images and when combining databases, no effect was noticed for model trained on a seen database at close range, although RGB model performed better at greater ranges. It is possible that the convergence results of RGB colour space are inflated at larger displacements due to the two databases used having RGB intense backgrounds (Multi-Pie blue, and IMM green) which are not captured by greyscale or l1l2l3, thus they perform worse under greater displacements. This highlights the deficiencies of training and testing of AAM fitting on seen databases, as the model learnt is quite database specific and might not generalise well.

Our results show the relative advantages of l1l2l3 colour space to RGB or greyscale for person independent AAM fitting.

References


