Enhanced Active Shape Models with Global Texture Constraints for Image Analysis

Shiguang Shan, Wen Gao, Wei Wang, Debin Zhao, Baocai Yin

Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, 100080
College of computer science, Beijing polytechnic university, Beijing, China, 100022
{sgshan, wgao, wwang, dbzhao}@jdl.ac.cn, ybc@bjpu.edu.cn

Abstract. Active Shape Model (ASM) has been widely recognized as one of the best methods for image understanding. In this paper, we propose to enhance ASMs by introducing global texture constraints expressed by its reconstruction residual in the texture subspace. In the proposed method, each landmark is firstly matched by its local profile in its current neighborhood, and the overall configure of all the landmarks is re-shaped by the statistical shape constraint as in the ASMs. Then, the global texture is warped out from the original image according to the current shape model, and its reconstruction residual from the pre-trained texture subspace is further exploited to evaluate the fitting degree of the current shape model to the novel image. Also, the texture is exploited to predict and update the shape model parameters before we turn to the next iterative local matching for each landmark. Our experiments on the facial feature analysis have shown the effectiveness of the proposed method.

Keywords: Active Shape Models (ASMs), Active Appearance Models (AAMs), Enhanced Active Shape Models (EASMs), Global Texture Constraints (GTC)

1. Introduction

In most pattern recognition and computer vision tasks, the localization and alignment of target object from an image is a task of great importance. To deal with the problem, many methods have been presented in recent years including active contour models (snake)[1], deformable template [2], elastic bunch graph matching [3], Gabor wavelet networks [4], Active Shape Models (ASMs) [5] and Active Appearance Models (AAMs)[6] etc. Among them, ASMs and AAMs are both based on statistical models, which are demonstrated to be efficient and effective for image interpretation. This paper presents an Enhanced ASM (EASM) by combining global texture constraints into ASM, which is essentially a strategy by combining ASM's local profile matching with AAM's global texture models. To integrate the local profile matching, the overall shape constraints, and the global texture constraints, the reconstruction residual of the global texture in its subspace is exploited to evaluate the fitting degree of the current model for the novel image. In addition, the global texture is also exploited to predict and update the shape model parameters. Therefore, with such an interleaving iteration of local profile matching, overall shape constraints, and global texture constraints, our method takes advantages of both ASMs and AAMs, meanwhile avoids their deficiencies.

2. EASMs with Global Texture Constraints

2.1 Fitting Degree Measurement in EASM

As is well known, the goal of the analysis-by-synthesis is to optimize the parameters of the model in order that the synthesized model matches the novel image to the utmost. Therefore, it is one of the most important tasks to define the objective function measuring the fitting degree between the novel image and the model, which should be minimized during the optimization.

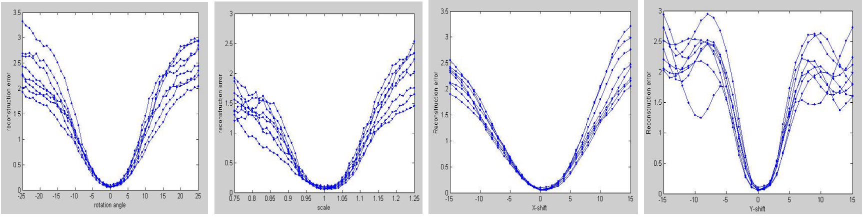
As illustrated in Fig.1, mis-matched shape implies distorted texture. Therefore, the mis-fitting degree is equivalent to the distortion degree in some sense. In this paper, the reconstruction residual of the current texture in the texture subspace is exploited as the fitting degree to be minimized.

Fig. 1. Mis-matched shape model and corresponding distorted texture

As in AAM, statistical texture model is pre-learned from a number of manually annotated images by applying PCA. The texture, g, warped from an image, can then be reconstructed as $g' = \bar{g} + P_g b_g$, where texture parameters, $b_g = P_g^T(g - \bar{g})$. So, the reconstruction residual can then be computed by:

$$\varepsilon = \left| g - g' \right| \tag{1}$$

If the model shape is well matched to the target object image, the texture is expected to be a normal patch without background and distortion. Therefore, according to the theory of the Eigen-analysis, the texture should be reconstructed "perfectly", that is, the reconstruction residual should be relatively small. Otherwise, the residual would be very large. To demonstrate this point, we systematically perturb the ground truth shape in scale (from 0.75 to 1.25), rotation (from –25° to +25°), and translation (from –15 pixels to 15 pixels along the X and Y axis respectively). Figure 1 illustrates the corresponding relationship between the reconstruction residual and the parameters displacements in our experiment. Each curve represents a series of parameter variations for one image. From these figures, approximate monotony relationship between the reconstruction residual and the shape parameters displacement can be observed in a certain range. Especially, when the displacement ranges within a certain limit, the monotony attribute is quite ideal. So, it is reasonable to use the reconstruction residual as the fitting degree measurement, i.e. the objective function, to be minimized.



(a) Rotation angle (b) Scale (c) X-shift (d) Y-shift

Fig. 2. Relationship between the shape parameters variations and the reconstruction residual

2.2 Updating Shape Model According to the Global Texture

Similar to AAMs, we also make use of the approximate linear relationship assumption between the texture displacement, δg , i.e. the shifting from the pre-trained mean texture, and the displacement of the shape parameter δb (However, we do not go all the way as AAMs, where appearance parameter is used.) [6]. That is, δb can be predicted by linear regression: $\delta b = R_b \delta g$, $\delta t = R_t \delta g$, where δb and δt are the shape parameter displacement and the affine transformation parameters respectively. The linear regression mapping R_b and R_t is computed in the same way as AAMs.

2.3 EASM with Global Texture Constraints for Image Feature Analysis

To sum up, the enhanced ASM with global texture constraints are outlined as follows:

1. *Initialize the mean shape in the novel image;*
2. *Move each landmark locally to its "best position" by profile matching as in ASMs;*
3. *Adjust the 2D affine parameters, and update the shape parameters under the statistical shape models constraints as in ASMs;*
4. *Warp the novel image to the mean shape according to the new shape to obtain the global texture* g*, and compute the fitting degree measurement, i.e. the reconstruction residual* ε *in the pre-learned texture subspace;*
5. *If* ε *is small enough, convergence is declared;*
   else {if ε *is smaller than its previous value, go to step 2;*
   else, go to the next step;}
6. *Predict and update the shape parameters according to* g*;*
7. *Perform the same operation as Step 4;*
8. *If* ε *is smaller than its previous value, go to step 6 again; otherwise go to step 2;*
9. *The above iteration is continued until no further improvement is observed.*

In the above iteration, each landmark is firstly matched by its local profile in its neighborhood, and the overall configure of all the landmarks is re-shaped by the statistical shape constraint as in the ASMs. Then, the global texture is warped out from the original image according to the current shape model, and its reconstruction residual from the pre-trained texture subspace is computed to measure the fitting degree of the current shape model to the novel image. Also, the texture is exploited to predict and update the shape model parameters before we turn to the next local matching for each landmark. Thus, by such an iteration interleaving the local profile matching, shape constraints and texture constraints, our method integrates the global texture constraints.

3. Experiments

By taking face image analysis for example, we conducted experiments to demonstrate the effectiveness of the proposed EASMs. A face database containing 500 face images manually labeled is exploited, in which all the faces possess different expressions including laugh, surprise and angry and all the faces are near frontal with slight illumination variation. To analyze the face image, we defined 103 landmarks on the face, whose coordinates form the face shape in our experiments.

In our experiments, the performance is evaluated by calculating the mean point distance between the resulting shapes and the labeled ones (ground-truth). The distance can be denoted as:

$$d = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{n}\sum_{j=1}^{n} dist(P_{ij}, P_{ij}^{'})\right) \tag{2}$$

where N is the total number of the test images, n is the number of the landmark points in the shape (for our face case, n=103), P_{ij} and $P_{ij}^{'}$ are respectively the labeled coordinate and the resulting coordinate for the j-th landmark of the i-th test image. The function $dist(p_{ij}, p_{ij}')$ is the Euclidean distance between the two points.

In the experiment, 350 images are randomly selected from the 500 images as the training set to build both the statistical texture model and the statistical shape model. The remaining 150 images are used for testing. Similar experiments are conducted 10 times for statistics. Table 1 shows the average experimental results of both ASMs and the proposed EASMs, from which some improvement of our method over ASMs can be observed. Figure 3 shows some example results of both ASM and EASM. As illustrated in this figure, the matching result of the mouth for ASM is not ideal due to the variation of expression, but the EASM works better.

Table 1. Performance comparison of different methods

Method	*d* (pixel)	Variance of *d* (pixel)
ASMs	3.05	3.23
EASMs	2.59	2.71

Results of ASM Results of our EASM

Fig. 3. Comparison between our method and ASM

4. Conclusion

In this paper, we propose the Enhanced ASMs by introducing global texture constraints expressed by its reconstruction residual in the texture subspace. The global texture is also exploited to predict and update the shape models. The method is essentially a combination of the ASM and AAM. Experiments have shown that our proposed method performs better than standard ASM.

Acknowledgement

This research is partly sponsored by National Hi-Tech Program of China (No.2001AA114190 and No. 2002AA118010), and IS' VISION Corp.

Reference

[1] M.Kass, A.Witkin,and D.Terzopoulos. Snakes: Active contour models. Int. Journal of Computer Vision, pp.321-331, 1988.

[2] A.L.Yuille, Deformable templates for face detection, J. Cogn. neurosci. 3, 59-70, 1991.

[3] L. Wiskott, J.M. Fellous, N. Kruger, C.v.d.Malsburg, Face Recogniton by Elastic Bunch Graph Matching, IEEE Trans. On PAMI, Vol.19, No. 7, pp775-779, 1997

[4] V. Kr"uger and G. Sommer. Gabor wavelet networks for object representation. Technical Report 2002, Institute of Computer Science, University of Kiel, 2000.

[5] T.F.Cootes, C.J.Taylor, D.Cooper, and J. Graham. Active shape models--their training and application. Computer vision and image understanding, 61(1): pp38-59, 1995.

[6] T.F.Cootes, C.J.Taylor. "Statistical models of appearance for computer vision", http://www.isbe.man.ac.uk/~bin/refs.html, 2001