Direct Appearance Models

XinWen Hou[1]*, Stan Z. Li[2], HongJiang Zhang[2], QianSheng Cheng[1].
[1] Institute of Mathematical Sciences, Peking University, Beijing, China
[2] Microsoft Research China, Beijing Sigma Center, Beijing 100080, China
Contact: szli@microsoft.com, http://research.microsoft.com/∼szli

Abstract

Active appearance model (AAM), which makes ingenious use of both shape and texture constraints, is a powerful tool for face modeling, alignment and facial feature extraction under shape deformations and texture variations. However, as we will show through our analysis and experiments, there exist admissible appearances that are not modeled by AAM and hence cannot be reached by AAM search; also the mapping from the texture subspace to the shape subspace is many-to-one and therefore a shape should be determined entirely by the texture in it.

In this paper, we propose a new appearance model, called direct appearance model (DAM), without combining from shape and texture as in AAM. The DAM model uses texture information directly in the prediction of the shape and in the estimation of position and appearance (hence the name DAM). In addition, DAM predicts the new face position and appearance based on principal components of texture difference vectors, instead of the raw vectors themselves as in AAM. These lead to the following advantages over AAM: (1) DAM subspaces includes admissible appearances previously unseen in AAM, (2) the convergence and accuracy are improved, and (3) the memory requirement is cut down to a large extent. The advantages are substantiated by comparative experimental results.

1 Introduction

The appearance based approach [16, 18, 2, 14] avoids difficulties in 3D modeling by using images of example appearances. It has become the dominant approach in face analysis and many other applications. The appearance of a face in an image is initially represented as a patch of image intensities (namely, texture) enclosed by the facial outline (namely, shape). The high dimensional initial representation is reduced to a low dimensional one by subspace analysis.

Subspace analysis techniques, such as principal component analysis (PCA) [10], help to reveal low dimensional structures of patterns observed in high dimensional spaces. A specific pattern of interest can reside in a low dimensional sub-manifold in the original input data space of an unnecessarily high dimensionality. Consider the case of $N \times M$ image pixels, each taking a value in $\{0, 1, \ldots, 255\}$; there is a huge number of possible configurations: $256^{N \times M}$. This space is capable of describing a wide variety of patterns or visual object classes. However, for a specific pattern, such as the human face, the number of admissible configurations is a only tiny fraction of that. In other words, the intrinsic dimension is much lower than $N \times M$.

Both the shape and texture provide important information useful for characterizing the face appearance [1]. Alignment of a given face to a canonical face enables extraction of refined shape and texture parameters in the coordinate system of the canonical face model. It is crucial for high accuracy face recognition and synthesis [13, 8, 9, 3, 11].

The active appearance model (AAM) [5] is a powerful model for face alignment, recognition [9] and synthesis [3]. It makes ingenious use of subspace analysis techniques, PCA in particular, to model both shape variation and texture variation, and the correlations between them. The idea is to warp the image patch enclosed by each training shape into a "shape free patch" enclosed by the mean shape. Statistical models of the warped shape and texture are then learned, and combined to form an appearance model by removing correlations between shape and texture. See also other papers of Cootes *et al.* at `www.isbe.man.ac.uk/~bim/refs.html` and thesis of Stegmann [17] at `www.imm.dtu.dk/~aam/`

Another merit of AAM is a smart search strategy: AAM assumes linear relationships between appearance variation and texture variation and between texture variation and position variation. It learns the two line regression models from training data. The two models facilitate the minimizations in high dimensional space. The AAM has been extended to multi-view faces using piecewise linear modeling [6, 7] or a single nonlinear model [15]

*The work presented in the paper was carried out at Microsoft Research, China.

In this paper, we propose a new appearance model, called direct appearance model (DAM), for aligning and estimating face appearances. The new appearance model is motivated by our findings of a flaw of AAM modeling and difficulties in training AAM in our analysis and experiments. Our analysis on mutual dependencies of shape, texture and appearance parameters in the AAM subspace models shows that there exist admissible appearances that are not modeled and hence cannot be reached by AAM search. The DAM model overcomes this problem by its proper subspace modeling based on our other findings: the mapping from the texture subspace to the shape subspace is many-to-one and therefore a shape can be determined entirely by the texture in it. From these relationships, the DAM model considers an appearance, which is composed of both shape and texture, to be determinable by using just the corresponding texture. DAM uses the texture information *directly* to predict the shape and to update the estimates of position and appearance (hence the name DAM); in contrast to AAM's crucial idea of modeling the AAM appearance subspace from shape and texture combined. This way, DAM includes admissible appearances previously unseen by AAM, and improves the convergence and accuracy.

Another problem with AAM is that its training of the two prediction models is based on texture difference vectors and is therefore very memory consuming because the training data for the two models are generated in a rapidly multiplicative way. The memory explosion makes AAM training very difficult even with a moderate number of images. To avoid this problem, DAM predicts the new face position and appearance based on principal components of texture difference vectors, instead of the raw vectors themselves as in AAM. This cuts down the memory requirement to a large extent, and further improves the convergence and accuracy. The claimed advantages of DAM are substantiated by comparative experimental results.

The rest of the paper is organized as follows: In Section 2, we analyze the AAM model and point out its shortcomings after a brief introduction of AAM. Then we propose the DAM model and search algorithm. Experimental results are presented in Section 3.

2 Direct Appearance Model

Assume that a training set be given as $\mathbf{A} = \{(S_0, T_0)\}$ where a shape $S_0 = ((x_1, y_1), \ldots, (x_K, y_K)) \in \mathbb{R}^{2K}$ is a sequence of K points in the 2D image plane, and a texture T_0 is the patch of image pixels enclosed by S_0. Let $\bar{S}$ be the mean shape of all the training shapes. $\bar{S}$ is calculated after the shapes are aligned to the tangent space of the mean shape $\bar{S}$, which can be implemented as an iterative procedure [5]. After the shape warping, the texture T_0 is warped correspondingly to $T \in \mathbb{R}^L$, where L is the number of pixels in the mean shape $\bar{S}$, by pixel value interpolation *e.g.* using a triangulation or thin plate spline method.

2.1 Introduction to AAM

In AAM, the shape is modeled by k ($< 2K$) principal modes learned from the training shapes using PCA. By this, a shape, which is originally in $\mathbb{R}^{2K}$, is represented as a point or vector s in the low dimensional shape subspace in $\mathbb{R}^k$

$$S = \bar{S} + \mathbf{U}s \tag{1}$$

where $\mathbf{U}$ is the matrix consisting of k principal orthogonal modes of variation in $\{S_0\}$. Because the training shapes have been aligned to the tangent space of $\bar{S}$, the eigenvectors in $\mathbf{U}$ is orthogonal to the mean shape $\bar{S}$, i.e. $\mathbf{U}^T\bar{S} = 0$, and the projection from S to s is

$$s = \mathbf{U}^T(S - \bar{S}) = \mathbf{U}^T S \tag{2}$$

The above defines AAM's shape subspace $\mathbb{S}_s$.

After deforming each training shape S_0 to the mean shape, the corresponding texture T_0 is warped to T. All the warped textures are aligned to the tangent space of the mean texture $\bar{T}$ by using an iterative approach as described in [5]. The PCA texture model is obtained as

$$T = \bar{T} + \mathbf{V}t \tag{3}$$

where $\mathbf{V}$ is the matrix consisting of ℓ principal orthogonal modes of variation in $\{T\}$, t is the vector of texture parameters. The projection from T to t is

$$t = \mathbf{V}^T(T - \bar{T}) = \mathbf{V}^T T \tag{4}$$

By this, the L pixel values in the mean shape is represented as a point in the ℓ dimensional texture subspace $\mathbb{S}_t$.

Since there may be correlations between the shape and texture variations, a further appearance model is built from $\{s\}$ and $\{t\}$. The appearance of each example is a concatenated vector

$$A = \begin{pmatrix} \mathbf{\Lambda}s \\ t \end{pmatrix} \tag{5}$$

where $\mathbf{\Lambda}$ is a diagonal matrix of weights for the shape parameters allowing for the difference in units between the shape and texture variation. One may simply set $\mathbf{\Lambda} = r\mathbf{I}$ where r^2 is the ratio of the total intensity variation to the total shape variation. Again, by applying PCA on the set $\{A\}$, one gets

$$A = \mathbf{W}a \tag{6}$$

where $\mathbf{W}$ is the matrix consisting of principal orthogonal modes of variation in $\{A\}$. By projecting from A to a, AAM models its appearance subspace $\mathbb{S}_a$ by

$$a = \mathbf{W}^T A \tag{7}$$

Consider the difference between the texture T_{im} in the image patch and the texture T_a reconstructed from the current appearance parameters

$$\delta T = T_{im} - T_a \tag{8}$$

In AAM, the search for a face in an image is guided by minimizing the norm $\|\delta T\|$. The AAM assumes that the appearance displacement δa and the position (including coordinates (x, y), scale s and rotation parameter θ) displacement δp are linearly correlated to δT. It predicts the displacements as

$$\delta a = \mathbf{A}_a \delta T \tag{9}$$

$$\delta p = \mathbf{A}_p \delta T \tag{10}$$

where the prediction matrices $\mathbf{A}_a, \mathbf{A}_p$ are to be learned from the training data by using linear regression. In order to estimate $\mathbf{A}_a$, we need to systematically displace a to get δa and the induced δT for each training image.

2.2 Motivations for DAM

The following analysis of relationships between the shape, texture and appearance subspaces in AAM shows defects of the AAM model. Thereby we suggest a property that an ideal appearance model should have, which motivates us to propose the DAM.

First, let us look into relationship between shape and texture from an intuitive viewpoint. A texture (*i.e.* the patch of intensities) is enclosed by a shape (before aligning to the mean shape); the same shape can enclose different textures (*i.e.* configurations of pixel values). However, the reverse is not true: different shapes can not enclose the same texture. So the mapping from the texture space to the shape space is many-to-one. The shape parameters should be determined completely by texture parameters but not vice versa.

Then, let us look further into the correlations or constraints between the linear subspaces $\mathbb{S}_s, \mathbb{S}_t$ and $\mathbb{S}_a$ in terms of their dimensionalities or ranks. Let denote the rank of space $\mathbb{S}$ by dim($\mathbb{S}$). We have the following analysis:

1. When dim($\mathbb{S}_a$)=dim($\mathbb{S}_t$)+dim($\mathbb{S}_s$), the shape and texture parameters are independent of each other, and there exist no mutual constraints between the s and t parameters.

2. When dim($\mathbb{S}_t$)<dim($\mathbb{S}_a$)<dim($\mathbb{S}_t$)+dim($\mathbb{S}_s$), not all the shape parameters are independent of the texture parameters. That is, one shape can correspond to more than one texture configuration in it, which conforms our intuition.

3. One can also derive the relationship dim($\mathbb{S}_t$)<dim($\mathbb{S}_a$) from Eq.(5) and (6) the formula

$$\mathbf{W}a = \begin{pmatrix} \mathbf{\Lambda}s \\ t \end{pmatrix} \tag{11}$$

when that s contains some components which are independent of t.

4. However, in AAM, it is often the case where dim($\mathbb{S}_a$)<dim($\mathbb{S}_t$) if the dimensionalities of $\mathbb{S}_a$ and $\mathbb{S}_t$ are chosen to retain, say 98%, of the total variations, which is reported by Cootes [5] and also observed by us. The consequence is that some admissible texture configurations cannot been seen in the appearance subspace because dim($\mathbb{S}_a$)<dim($\mathbb{S}_t$), and therefore cannot be reached by the AAM search. We consider this a flaw of AAM's modeling of its appearance subspace.

From the above analysis, we conclude that the ideal model should be such that dim($\mathbb{S}_a$)=dim($\mathbb{S}_t$) and hence that s completely linearly determinable by t. In other words, the shape should be linearly dependent on the texture so that dim($\mathbb{S}_t \cup \mathbb{S}_s$)=dim($\mathbb{S}_t$). The DAM model is proposed mainly for this purpose.

Another motivation of DAM is the space consumption: the regression of $\mathbf{A}_a$ in AAM is very memory consuming. AAM prediction needs to model linear relationship between appearance and texture difference according to (9). However, both a and δT are high dimensional vectors, and therefore the storage size of training data generated for learning (9) increases very rapidly as the dimensions increase. It is very difficult to train AAM for $\mathbf{A}_a$ even with a moderate number of images. Learning in a low dimensional space will relieve the burden.

2.3 DAM Modeling and Training

DAM consists of a shape model, a texture model and a prediction model. Abandoning AAM's crucial idea of combining shape and texture parameters into an appearance model, it predicts the shape parameters directly from the texture parameters. The shape and texture models are built based on PCA in the same way as in AAM. The prediction model includes two parts: prediction of position and prediction of texture.

Recall the conclusions we made earlier: (1) an ideal model should have dim($\mathbb{S}_a$)=dim($\mathbb{S}_t$) and (2) shape should be computable uniquely from texture but not vice versa. We propose the following prediction model by assuming a linear relationship between shape and texture

$$s = \mathbf{R}t + \varepsilon \tag{12}$$

where $\varepsilon = s - \mathbf{R}t$ is noise and $\mathbf{R}$ is a $k \times l$ projection matrix. Denoting the expectation by $E(\cdot)$, if all the elements in the variance matrix $E(\varepsilon\varepsilon^T)$ are small enough, the linear assumption made in Eq.(12) is approximately correct. This is true as will be verified later by experiments. Define the objective cost function

$$C(\mathbf{R}) = E(\varepsilon^T\varepsilon) = \mathrm{trace}[E(\varepsilon\varepsilon^T)] \tag{13}$$

$\mathbf{R}$ is learned from training example pairs $\{(s,t)\}$ by minimizing the above cost function. The the optimal solution is (see Appendix)

$$\mathbf{R}^* = E(st^T)[E(tt^T)]^{-1} \tag{14}$$

The minimized cost is the trace of the following

$$E(\varepsilon\varepsilon^T) = E(ss^T) - \mathbf{R}^* E(tt^T)\mathbf{R}^{*T} \tag{15}$$

Instead of using δT directly as in the AAM search (*cf.* Eq.(10), we use principal components of it, $\delta T'$, to predict the position displacement

$$\delta p = \mathbf{R}_p \delta T' \tag{16}$$

where $\mathbf{R}_p$ is the prediction matrix learned by using linear regression. To do this, we collect texture differences induced by small position displacements in each training image, and perform PCA on this data to get the projection matrix $\mathbf{H}^T$. A texture difference is projected onto this subspace as

$$\delta T' = \mathbf{H}^T \delta T \tag{17}$$

$\delta T'$ is about 1/4 of δT in dimensionality and this makes the prediction more stable. The DAM regression in Eq.(16) requires much less memory than the AAM regression in Eq.(9). This is because p is of much lower dimension than a and $\delta T'$ much lower than δT. This will be illustrated by numbers later.

Note that there is a variant of basic AAM [4], which uses texture difference to predict shape difference. The prediction of shape is done by $\delta s = \mathbf{B}\delta T$. However, this variant is not as good as the basic AAM [4].

2.4 DAM Search

The DAM prediction models leads to the following search procedure: The DAM search starts with the mean shape and mean texture, equivalent to the mean appearance with $a_0 = 0$, at a given initial position p_0. The texture difference δT is computed from the current shape patch at the current position, and and its principal components are used to predict and update p and s using the DAM linear models described above. If $\|\delta T\|$ calculated using the new appearance at the position is smaller than the old one, the new appearance and position are accepted; otherwise the position and appearance are updated by amounts $\kappa\delta_a$ and $\kappa\delta p$ with varying κ values. The search algorithm is summarized below:

1. Initialize position parameters p_0, and set shape parameters $s_0 = 0$;
2. Get texture T_{im} from the current position, project it into the texture subspace $\mathbb{S}_t$ as *t*, reconstruct the texture T_a, and compute texture difference $\delta T_0 = T_{im} - T_a$ and the energy $E_0 = \|\delta T_0\|^2$;
3. Compute $\delta T' = \mathbf{H}^T \delta T$, and get the position displacement $\delta p = \mathbf{R}_p \delta T'$;
4. Set step size $\kappa = 1$
5. Update $p = p_0 - \kappa\delta p$, $s = \mathbf{R}t$;
6. Compute the difference texture δT using the new shape at the new position, and its energy $E_0 = \|\delta T_0\|^2$;
7. If $|E - E_0| < \epsilon$, the algorithm is converged; exit;
8. If $E < E_0$, then let $p_0 = p, s_0 = s, \delta T_0 = \delta T, E_0 = E$, goto 3;
9. Change κ to the next smaller number in $\{1.5, 0.5, 0.25, 0.125, \ldots, \}$, goto 5;

The above DAM search can be performed with a multi-resolution pyramid structure to improve the result.

3 Experimental Results

3.1 Computation of Subspaces

A total of 80 images of size 128x128 are collected. Each image contains a different face in an area of about 64x64 pixels. The images set is randomly partitioned into a training set of 40 images and a test set of the other 40. Each image is mirrored and this doubles the total number of images in each set.

$K = 72$ face landmark points are labeled manually (see an example in Fig.1. The shape subspace is $k = 39$ dimensional, which retains 98% of the total shape variation. The mean shape contains a texture of $L = 3186$ pixels. The texture subspace is $\ell = 72$ dimensional, as the result of retaining 98% of total texture variation. These are common to both AAM and DAM.

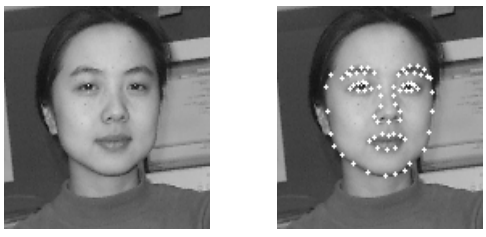

Figure 1: A face image and the landmark points.

For AAM, an appearance subspace is constructed to combine both shape and texture information: A concatenated shape and texture vector is 39+72 dimensional, where the weight parameter is calculated as $r = 7.5$ for $\mathbf{\Lambda} = r\mathbf{I}$

in Eq.(5). It is reduced to a 65 dimensional appearance subspace which retains 98% of total variation of the concatenated features.

For DAM, the linearity assumption made for the model $s = \mathbf{R}t + \varepsilon$ of Eq.(12) is well verified because all the elements in $E(\varepsilon\varepsilon^T)$ calculated over the training set are smaller than 10^{-5}.

The original texture difference δT, which is used in AAM for predicating position displacement, is 3186 dimensional; it is reduced to 724 dimensional $\delta T'$, which is used in DAM for the prediction, to retain 98% of variation over the 1920 training examples.

DAM requires much less memory during the learning of the prediction matrices $\mathbf{R}_p$ in Eq.(16) than AAM for learning $\mathbf{A}_a$ in Eq.(9). For DAM, there are 80 training images, 4 parameters for the position: $(x, y, \theta, scale)$, and 6 disturbances for each parameter to generate training data for the training $\mathbf{R}_p$. So, the size of training data for DAM is $80 \times 4 \times 6 = 1920$. For AAM, there are 80 training images, 65 appearance parameters, and 4 disturbances for each parameter to generate training data for training $\mathbf{A}_a$. The size of training data for $\mathbf{A}_a$ is $80 \times 65 \times 4 = 20800$. Therefore, the size of training data for AAM's prediction matrices is $20800 + 1920 = 22720$, which is 11.83 times that for DAM. On a PC, for example, the memory capacity for AAM training with 80 images would allow DAM training with 946 images.

3.2 Alignment and Appearance Estimation

Table 1 compares DAM and AAM in terms of the quality of position and texture parameter estimates, and the convergence rates. The effect of using $\delta T'$ instead of δT is demonstrated through DAM', which is DAM minus the PCA subspace modeling of δT. The initial position is a shift from the true position by $dx = 6, dy = 6$. The $||\delta p||$ is calculated for each image as the averaged distance between corresponding points in the two shapes, and therefore it is also a measure of difference in shape. The convergence is judged by the satisfaction of two conditions: $||\delta T||^2 < 0.5$ and $||\delta p|| < 3$.

	$E(\|\|\delta T\|\|^2)$	std($\|\|\delta T\|\|^2$)	$E(\|\|\delta p\|\|)$	std($\|\|\delta p\|\|$)	cvg rate
DAM	0.156572	0.065024	0.986815	0.283375	100%
DAM'	0.155651	0.058994	0.963054	0.292493	100%
AAM	0.712095	0.642727	2.095902	1.221458	70%
DAM	1.114020	4.748753	2.942606	2.023033	85%
DAM'	1.180690	5.062784	3.034340	2.398411	80%
AAM	2.508195	5.841266	4.253023	5.118888	62%

Table 1: Comparisons of DAM, DAM' and AAM in terms of errors in estimated texture (appearance) parameters δT and position δp and convergence rates for the training images (first block of three rows) and test images (second block).

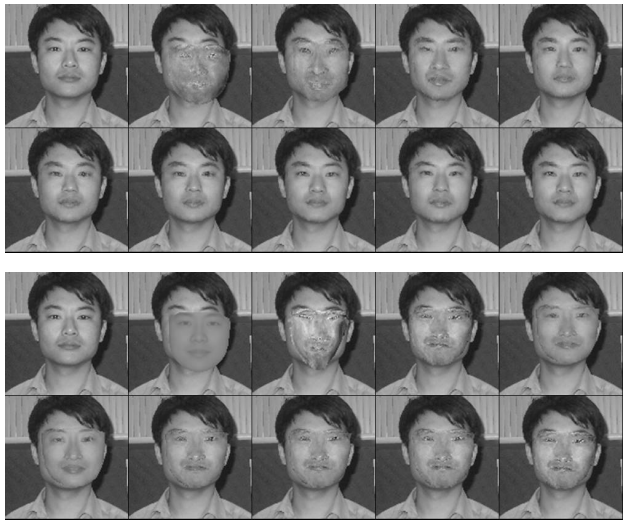

Figure 2: Scenarios of DAM (top) and AAM (bottom) alignment.

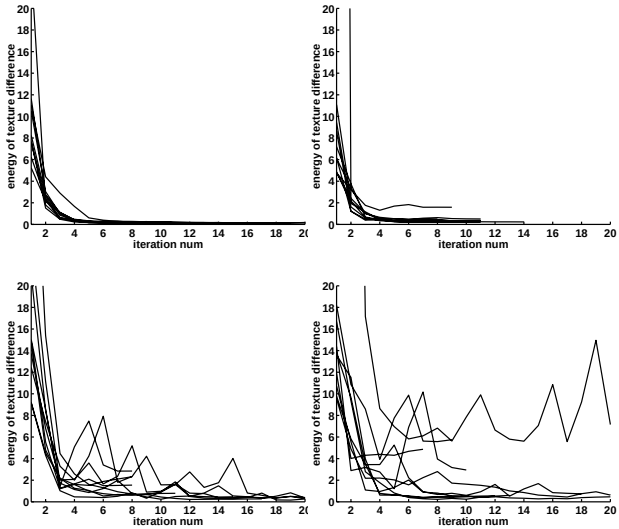

Figure 3: The evolution of total δT for the DAM (top) and AAM (bottom) as a function of iteration number for the training (left) and test (right) images.

Fig.2 illustrates average scenarios of DAM and AAM alignment. Fig.3 illustrates the dynamics of total error δT for 10 images randomly selected from the training set and 10 from the test set. We see that DAM has faster convergence and smaller error than AAM.

4 Conclusion

In this paper, we have proposed a novel appearance method, direct appearance model (DAM), to overcome defects of AAM. Subspace modeling in DAM overcomes a deficiency of AAM in appearance subspace modeling so that all admissible appearances can be seen in the modeled subspaces, and thus reachable in DAM search. This improves convergence and solution accuracy. Also, the use of subspace modeling in searching for DAM solution greatly improves the efficiency in memory and time for both training and performing and makes it possible to learn prediction matrix from a large number of training images. DAM has been extended for multi-view faces [12].

Appendix

Consider variation $\delta C(\mathbf{R})$ caused by $\delta\mathbf{R}$

$$\begin{aligned}
&\delta C(\mathbf{R}) \qquad (18)\\
&= \mathrm{trace}\{E([s-(\mathbf{R}+\delta\mathbf{R})t][s-(\mathbf{R}+\delta\mathbf{R})t]^T)\}\\
&\quad -\mathrm{trace}[E\{[s-\mathbf{R}t][s-\mathbf{R}t]^T\}]\\
&= \mathrm{trace}\{E[\mathbf{R}tt^T\delta\mathbf{R}^T+\delta\mathbf{R}tt^T\mathbf{R}\\
&\quad -st^T\delta\mathbf{R}^T-\delta\mathbf{R}ts^T]\}\\
&= \mathrm{trace}\{\mathbf{R}E(tt^T)\delta\mathbf{R}^T+\Delta\mathbf{R}E(tt^T)\mathbf{R}\\
&\quad -E(st^T)\Delta\mathbf{R}^T-\delta\mathbf{R}E(ts^T)\}
\end{aligned}$$

Letting $\delta C(\mathbf{R})=0$, we get

$$\begin{aligned}
&\mathrm{trace}\{\delta\mathbf{R}E(tt^T)\delta\mathbf{R}^T+\delta\mathbf{R}E(tt^T)\mathbf{R}\} \qquad (19)\\
&\quad = \mathrm{trace}\{E(st^T)\Delta\mathbf{R}^T+\Delta\mathbf{R}E(ts^T)\}
\end{aligned}$$

for any $\|\delta\mathbf{R}\|\to 0$. Substituting $\delta\mathbf{R}$ by $\epsilon 1_{i,j}$ for any (i,j) where $\epsilon\to 0$ and $1_{i,j}$ is the matrix in which entry (i,j) is 1 and 0 elsewhere, we arrive at $\mathbf{R}E(tt^T)=E(st^T)$, and hence obtain the optimal solution

$$\mathbf{R}=E(st^T)[E(tt^T)]^{-1} \qquad (20)$$

References

[1] D. Beymer. "Vectorizing face images by interleaving shape and texture computations". A. I. Memo 1537, MIT, 1995.

[2] D. Beymer, A. Shashua, and T. Poggio. "Example based image analysis and synthesis". A. I. Memo 1431, MIT, 1993.

[3] V. Blanz and T.Vetter. "A morphable model for the synthesis of 3d faces". In *SIGGRAPH'99 Conference Proceedings*, pages 187–194, 1999.

[4] T. F. Cootes, , and C. J. Taylor. Statistical models of appearance for computer vision. Technical report, `www.isbe.man.ac.uk/~bim/refs.html`, 2001.

[5] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. In *ECCV98*, volume 2, pages 484–498, 1998.

[6] T. F. Cootes, K. N. Walker, and C. J. Taylor. View-based active appearance models. In *Proc. Int. Conf. on Face and Gesture Recognition*, pages 227–232, 2000.

[7] T. F. Cootes, G. V. Wheeler, K. N. Walker, and C. J. Taylor. Coupled-view active appearance models. In *Proc. British Machine Vision Conference*, volume 1, pages 52–61, 2000.

[8] F. de la Torre, S. Gong, and S. McKenna. "View alignment with dynamically updated affine tracking". In *Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, Nara, Japan", month=14-16, April, year=1998.

[9] G. Edwards, T. Cootes, and C. Taylor. "Face recognition using active appearance models". In *Proceedings of the European Conference on Computer Vision*, volume 2, pages 581–695, 1998.

[10] K. Fukunaga. *Introduction to statistical pattern recognition*. Academic Press, Boston, 2 edition, 1990.

[11] S. Gong, S. McKenna, and A. Psarrou. *Dynamic Vision: From Images to Face Recognition*. World Scientific Publishing and Imperial College Press, April 2000.

[12] S. Z. Li, S. C. Yan, H. J. Zhang, and Q. S. Cheng. "Multi-view face alignment using direct appearance models". In preparation, 2001.

[13] B. Moghaddam and A. Pentland. "Probabilistic visual learning for object representation". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7:696–710, July 1997.

[14] H. Murase and S. K. Nayar. "Visual learning and recognition of 3-D objects from appearance". *International Journal of Computer Vision*, 14:5–24, 1995.

[15] S. Romdhani. "Multi-view nonlinear active shape model using kernel pca". In *Proceedings of the British Machine Vision Conference*, 1999.

[16] L. Sirovich and M. Kirby. "Low-dimensional procedure for the characterization of human faces". *Journal of the Optical Society of America A*, 4(3):519–524, March 1987.

[17] M. B. Stegmann. Active appearance models: Theory, extensions & cases. Master's thesis, Department of Mathematical Modelling, Technical University of Denmark, Lyngby, Denmark, 2000.

[18] M. A. Turk and A. P. Pentland. "Face recognition using eigenfaces.". In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 586–591, Hawaii, June 1991.