Appendix A
Some Statistical Analysis of the Angles between the Eigenvectors of Two Expression Subspaces

In this Appendix, we show the difference between a happy subspace and a neutral subspace by comparing their eigenvectors. In order to compare two eigenvectors, we calculated the angle between them.

Two subspaces might be the same, but their eigenvectors might not be in the same order (in terms of the order of their eigenvalues). For this reason, we considered the minimum angle between each eigenvector of one subspace and all the eigenvectors of the other subspace. Fig. 1 shows the minimum angles between the happy and neutral subspaces for their first ten eigenvectors. For this experiment, the neutral subspace was constructed using 300 neutral images from 300 subjects in the FRGC database. The happy subspace was constructed using 100 happy images from 100 subjects in the same database (the number of happy images is much less than the number of the neutral images in this database.). We repeated constructing the subspaces with different sets of images for many times. Fig. 1 shows the average of the minimum angles.

In order to make sure that the angles are not due to the different sets of images used to create the happy and neural subspaces, we also compared the eigenvectors of the neutral subspace with another *neutral* subspace created from a different set of neutral images. Fig. 1 shows the minimum angles between the two neutral subspaces, which were averaged over many trials. As seen in this figure, the minimum angle between a happy and a neutral subspace is considerably greater than the minimum angle between two neutral subspaces, which shows the difference between neutral and happy subspaces.

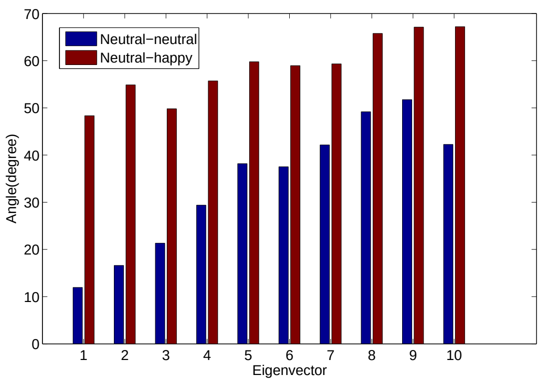

Fig. 1. The minimum angle between the eigenvectors of a neutral subspace and a happy subspace and the minimum angle between the eigenvectors of the same neutral subspace and another neutral subspace.

Appendix B
Proof of the Theorem

We prove the theorem for $n_i = 2$ and $n_i = 3$. The proof can be easily extended for $n_i > 3$. When $n_i = 2$ the scatter matrix of class i is

$$\mathbf{S}_w^i = (\mathbf{x}_{i1}-\mathbf{m}_i)(\mathbf{x}_{i1}-\mathbf{m}_i)^T+(\mathbf{x}_{i2}-\mathbf{m}_i)(\mathbf{x}_{i2}-\mathbf{m}_i)^T \quad (1)$$

By substituting $\mathbf{m}_i$ with $\frac{\mathbf{x}_{i1}+\mathbf{x}_{i2}}{2}$, (1) can be written as

$$\mathbf{S}_w^i = \left(\frac{\mathbf{x}_{i1}-\mathbf{x}_{i2}}{2}\right)\left(\frac{\mathbf{x}_{i1}-\mathbf{x}_{i2}}{2}\right)^T+\left(\frac{\mathbf{x}_{i2}-\mathbf{x}_{i1}}{2}\right)\left(\frac{\mathbf{x}_{i2}-\mathbf{x}_{i1}}{2}\right)^T \quad (2)$$

Hence,

$$\mathbf{S}_w^i = \frac{1}{2}(\mathbf{x}_{i1}-\mathbf{x}_{i2})(\mathbf{x}_{i1}-\mathbf{x}_{i2})^T = \frac{1}{2}\mathbf{d}_{i1}\mathbf{d}_{i1}^T \quad (3)$$

which proves the theorem for $n_i = 2$.
When $n_i = 3$

$$\begin{aligned}\mathbf{S}_w^i = &\left(\mathbf{x}_{i1}-\frac{\mathbf{x}_{i1}+\mathbf{x}_{i2}+\mathbf{x}_{i3}}{3}\right)\left(\mathbf{x}_{i1}-\frac{\mathbf{x}_{i1}+\mathbf{x}_{i2}+\mathbf{x}_{i3}}{3}\right)^T\\ &+\left(\mathbf{x}_{i2}-\frac{\mathbf{x}_{i1}+\mathbf{x}_{i2}+\mathbf{x}_{i3}}{3}\right)\left(\mathbf{x}_{i2}-\frac{\mathbf{x}_{i1}+\mathbf{x}_{i2}+\mathbf{x}_{i3}}{3}\right)^T\\ &+\left(\mathbf{x}_{i3}-\frac{\mathbf{x}_{i1}+\mathbf{x}_{i2}+\mathbf{x}_{i3}}{3}\right)\left(\mathbf{x}_{i3}-\frac{\mathbf{x}_{i1}+\mathbf{x}_{i2}+\mathbf{x}_{i3}}{3}\right)^T\end{aligned} \quad (4)$$

which can be written as

$$\begin{aligned}\mathbf{S}_w^i = &\frac{1}{9}(\mathbf{x}_{i1}-\mathbf{x}_{i2}+\mathbf{x}_{i1}-\mathbf{x}_{i3})(\mathbf{x}_{i1}-\mathbf{x}_{i2}+\mathbf{x}_{i1}-\mathbf{x}_{i3})^T\\ &+\frac{1}{9}(\mathbf{x}_{i2}-\mathbf{x}_{i1}+\mathbf{x}_{i2}-\mathbf{x}_{i3})(\mathbf{x}_{i2}-\mathbf{x}_{i1}+\mathbf{x}_{i2}-\mathbf{x}_{i3})^T\\ &+\frac{1}{9}(\mathbf{x}_{i3}-\mathbf{x}_{i1}+\mathbf{x}_{i3}-\mathbf{x}_{i2})(\mathbf{x}_{i3}-\mathbf{x}_{i1}+\mathbf{x}_{i3}-\mathbf{x}_{i2})^T\end{aligned} \quad (5)$$

We can then write (5) as

$$\begin{aligned}\mathbf{S}_w^i = &\frac{1}{9}[(\mathbf{x}_{i1}-\mathbf{x}_{i2})(\mathbf{x}_{i1}-\mathbf{x}_{i2})^T+\underbrace{(\mathbf{x}_{i1}-\mathbf{x}_{i2})(\mathbf{x}_{i1}-\mathbf{x}_{i3})^T}_{.}\\ &+\underbrace{(\mathbf{x}_{i1}-\mathbf{x}_{i3})(\mathbf{x}_{i1}-\mathbf{x}_{i2})^T}_{...}+(\mathbf{x}_{i1}-\mathbf{x}_{i3})(\mathbf{x}_{i1}-\mathbf{x}_{i3})^T]\\ &\frac{1}{9}[(\mathbf{x}_{i2}-\mathbf{x}_{i1})(\mathbf{x}_{i2}-\mathbf{x}_{i1})^T+\underbrace{(\mathbf{x}_{i2}-\mathbf{x}_{i1})(\mathbf{x}_{i2}-\mathbf{x}_{i3})^T}_{.}\\ &+\underbrace{(\mathbf{x}_{i2}-\mathbf{x}_{i3})(\mathbf{x}_{i2}-\mathbf{x}_{i1})^T}_{..}+(\mathbf{x}_{i2}-\mathbf{x}_{i3})(\mathbf{x}_{i2}-\mathbf{x}_{i3})^T]\\ &\frac{1}{9}[(\mathbf{x}_{i3}-\mathbf{x}_{i1})(\mathbf{x}_{i3}-\mathbf{x}_{i1})^T+\underbrace{(\mathbf{x}_{i3}-\mathbf{x}_{i1})(\mathbf{x}_{i3}-\mathbf{x}_{i2})^T}_{...}\\ &+\underbrace{(\mathbf{x}_{i3}-\mathbf{x}_{i2})(\mathbf{x}_{i3}-\mathbf{x}_{i1})^T}_{..}+(\mathbf{x}_{i3}-\mathbf{x}_{i2})(\mathbf{x}_{i3}-\mathbf{x}_{i2})^T]\end{aligned} \quad (6)$$

By summing up the terms identified by one dot, we get

$$\begin{aligned}&(\mathbf{x}_{i1}-\mathbf{x}_{i2})(\mathbf{x}_{i1}-\mathbf{x}_{i3})^T+(\mathbf{x}_{i2}-\mathbf{x}_{i1})(\mathbf{x}_{i2}-\mathbf{x}_{i3})^T\\ &=(\mathbf{x}_{i1}-\mathbf{x}_{i2})(\mathbf{x}_{i1}-\mathbf{x}_{i3})^T+(\mathbf{x}_{i1}-\mathbf{x}_{i2})(\mathbf{x}_{i3}-\mathbf{x}_{i2})^T\\ &=(\mathbf{x}_{i1}-\mathbf{x}_{i2})(\mathbf{x}_{i1}-\mathbf{x}_{i3}+\mathbf{x}_{i3}-\mathbf{x}_{i2})^T\\ &=(\mathbf{x}_{i1}-\mathbf{x}_{i2})(\mathbf{x}_{i1}-\mathbf{x}_{i2})^T\end{aligned} \tag{7}$$

Similarly, the other two pairs of terms can be simplified and the scatter matrix of class i can be written as

$$\begin{aligned}\mathbf{S}_w^i=&\frac{1}{3}(\mathbf{x}_{i1}-\mathbf{x}_{i2})(\mathbf{x}_{i1}-\mathbf{x}_{i2})^T+\frac{1}{3}(\mathbf{x}_{i2}-\mathbf{x}_{i3})(\mathbf{x}_{i2}-\mathbf{x}_{i3})^T\\ &+\frac{1}{3}(\mathbf{x}_{i3}-\mathbf{x}_{i1})(\mathbf{x}_{i3}-\mathbf{x}_{i1})^T=\frac{1}{3}\sum_{j=1}^{3}\mathbf{d}_{ij}\mathbf{d}_{ij}^T\end{aligned} \tag{8}$$

Appendix C
Orthogonality of the Synthesis Error and Synthesized Pairwise Distance

Since the neutral subspace is an affine subspace, the difference between any two points on this subspace can be written as a linear combination of the neutral eigenvectors, i.e.,

$$\exists a_1,...,a_k\in\mathbb{R}|\mathbf{n}_1-\mathbf{n}_2=a_1\mathbf{u}_1+...+a_k\mathbf{u}_k \tag{9}$$

where $\mathbf{n}_1$ and $\mathbf{n}_2$ are two points on the neutral subspace. Therefore, since $\mathcal{N}(\mathbf{h})$ and $\mathbf{n}$ are two points on the neutral subspace, $\mathbf{e}_s$ ($\mathbf{e}_s=\mathbf{n}-\mathcal{N}(\mathbf{h})$) can be written as a linear combination of the neutral eigenvectors.

On the other hand, using the definition of the projection, it is straightforward to show that the synthesized pairwise distance $\tilde{\mathbf{d}}$ is orthogonal to the neutral eigenvectors,

$$\tilde{\mathbf{d}}\perp\{\mathbf{u}_1,...,\mathbf{u}_k\} \tag{10}$$

As a result, $\tilde{\mathbf{d}}$ is orthogonal to any linear combination of neutral eigenvectors, and therefore, it is orthogonal to the synthesis error,

$$\tilde{\mathbf{d}}\perp\mathbf{e}_s \tag{11}$$