FACE HALLUCINATION REVISITED: A JOINT FRAMEWORK

Yonggang Jin and Christos Bouganis

Department of Electrical and Electronic Engineering
Imperial College London
London SW2 7AZ, UK

ABSTRACT

The paper presents a joint framework for face hallucination incorporating face deblurring and registration. The joint framework not only directly hallucinates low resolution faces, but also deblurs and aligns low resolution faces iteratively to improve the performance of face hallucination. Without the need for accurate face registration and prior knowledge of blurring kernels, it is robust to errors in face registration and blurring kernel. Experimental results demonstrate the robust performance of the proposed method.

Index Terms— High-resolution imaging, face hallucination, image restoration, deconvolution, image registration.

1. INTRODUCTION

Face hallucination is a very active research area in signal processing, image processing and computer vision with the objective to enhance facial image resolution. There has been much previous work carried out on face hallucination [1, 2, 3, 4, 5, 6, 7, 8], since the pioneering work of Baker and Kanade [9]. A face subspace-based approach was presented in [1], which incorporates Principal Component Analysis (PCA) prior based on two formulations, FS-MAP and IS-MAP. In [2], PCA-based global constraints, of which the soft-constraint is essentially the same as the FS-MAP in [1], are combined with a patch-based local constraint from [10] for face hallucination. In [3], Nonnegative Matrix Factorization (NMF) was adopted instead of PCA as the global constraint, which is then combined with sparse coding for face hallucination. Tensor-based face hallucination methods were introduced in [7, 8] using multi-modal faces. The asymmetry of registration of low resolution (LR) faces was pointed out in [5], Resolution-Aware Fitting (RAF) was proposed for joint registration and super-resolution (SR) and it avoids interpolating LR images, which is common for face alignment by the Lucas-Kanade (LK) algorithm [11] and the Active Appearance Model(AAM)[12].

The paper presents a joint framework for face hallucination incorporating face deblurring and registration. The joint framework not only directly hallucinates LR faces without interpolating LR images as suggested in [5], but also deblurs and aligns LR faces iteratively to improve the performance of face hallucination. It will be shown later that the FS-MAP[1]/Soft-Constraint[2] reduces to one iteration step in the proposed method. Without the need for accurate face registration and prior knowledge of blurring kernels, it is robust to errors in face registration and blurring kernel. Note that in previous work on face hallucination [1, 2, 5, 6, 7], LR faces are either aligned manually or automatically beforehand and there may be initial alignment errors, which actually may degrade performance of face hallucination. In addition, fixed and small blurring is often assumed in previous work, be it Gaussian or square, which may not be valid in practice.

The remainder of the paper is organized as follows. Face hallucination is formulated in section 2. Section 3 describes a joint framework for face hallucination incorporating face deblurring and registration. Experimental results are presented in section 4 and the paper is concluded in section 5.

2. FACE HALLUCINATION FORMULATION

2.1. Observation model

The process of image acquisition can be modeled by an observation model,which relates an original high resolution (HR) image $\mathbf{x}$ to an observed LR image $\mathbf{z}$, both vectors in lexicographical order,

$$\mathbf{z} = \mathbf{S}(s)\mathbf{B}(\mathbf{k})\mathbf{W}(\boldsymbol{\theta})\mathbf{x} + \mathbf{n} \quad (1)$$

where $\mathbf{W}(\boldsymbol{\theta})$ warps an HR image via transformation $\boldsymbol{\theta}$, $\mathbf{B}(\mathbf{k})$ represents blurring effect caused by blurring kernel $\mathbf{k}$, $\mathbf{S}(s)$ is the sub-sampling operator with fixed factor s, and $\mathbf{n}$ is additive sensor noise which is often assumed to be zero mean Gaussian with covariance $\mathbf{R}$. In the process of image acquisition, loss of image spatial resolution is mainly caused by blurring (optical blurring, motion blurring and sensor Point Spread Function (PSF)), sub-sampling and additive sensor noise. (1) can be rewritten as

$$\mathbf{z} = \mathbf{H}(\mathbf{k}, \boldsymbol{\theta})\mathbf{x} + \mathbf{n}$$

where $\mathbf{H}(\mathbf{k}, \boldsymbol{\theta}) = \mathbf{S}(s)\mathbf{B}(\mathbf{k})\mathbf{W}(\boldsymbol{\theta})$ combines warping, blurring and sub-sampling operations. The observation model provides data constraint [2] for an HR image to be estimated.

It can be noted that except sub-sampling factor s, both blurring kernel $\mathbf{k}$ and transformation $\boldsymbol{\theta}$ are unknown in practice. The following approaches can be found in previous work. In [1], an LR face is manually aligned by selecting eyes and mouth centre and blurring kernel is modelled as a Gaussian with a fixed variance. LR faces can be aligned automatically in [2, 7] by interpolating LR images, which however is undesirable as suggested in [5] and blurring kernel is assumed to be square. In [5], joint face registration and super-resolution is achieved without LR image interpolation however blurring kernel is also fixed to be square.

2.2. Prior

The observation model can be regarded as likelihood $p(\mathbf{z}|\mathbf{x})$, and to estimate an HR image with an LR image is equivalent to *Maximum Likelihood* (ML) estimation, $\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} p(\mathbf{z}|\mathbf{x})$. In practice, the ML solution is often under-determined and it is essential to apply prior $p(\mathbf{x})$ to constrain the solution and use *Maximum a Posteriori* (MAP) estimation instead. For PCA prior [1, 2], $\mathbf{x}$ is generated using eigenfaces $\mathbf{D}$ and mean face $\boldsymbol{\mu}$, which can be formulated in a generative model with latent coefficients $\boldsymbol{\alpha}$

$$\begin{aligned} p(\mathbf{x}|\boldsymbol{\alpha}) &= \mathcal{N}(\mathbf{x}; \boldsymbol{\mu} + \mathbf{D}\boldsymbol{\alpha}, \boldsymbol{\Sigma}) \\ p(\boldsymbol{\alpha}) &= \mathcal{N}(\boldsymbol{\alpha}; \mathbf{0}, \mathbf{P}) \end{aligned} \tag{2}$$

where prior of latent coefficients $\boldsymbol{\alpha}$ is a zero-mean Gaussian with diagonal covariance $\mathbf{P}$ based on eigenvalues of corresponding eigenfaces.

2.3. MAP estimation

Combining likelihood (1) and prior (2), the posterior is

$$p(\mathbf{x}, \mathbf{k}, \boldsymbol{\theta}, \boldsymbol{\alpha}|\mathbf{z}) \propto p(\mathbf{z}|\mathbf{x}, \mathbf{k}, \boldsymbol{\theta}) p(\mathbf{x}|\boldsymbol{\alpha}) p(\boldsymbol{\alpha}) p(\mathbf{k}) p(\boldsymbol{\theta}) \tag{3}$$

and if both blurring kernel $\mathbf{k}^*$ and transformation $\boldsymbol{\theta}^*$ are known beforehand as in [1, 2], then the posterior can be simplified to

$$p(\mathbf{x}, \boldsymbol{\alpha}|\mathbf{z}, \mathbf{k}^*, \boldsymbol{\theta}^*) \propto p(\mathbf{z}|\mathbf{x}, \mathbf{k}^*, \boldsymbol{\theta}^*) p(\mathbf{x}|\boldsymbol{\alpha}) p(\boldsymbol{\alpha}) \tag{4}$$

and MAP estimation is given by

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}, \boldsymbol{\alpha}} \|\mathbf{z} - \mathbf{H}(\mathbf{k}^*, \boldsymbol{\theta}^*)\mathbf{x}\|^2_{\mathbf{R}} + \|\mathbf{x} - \boldsymbol{\mu} - \mathbf{D}\boldsymbol{\alpha}\|^2_{\boldsymbol{\Sigma}} + \|\boldsymbol{\alpha}\|^2_{\mathbf{P}} \tag{5}$$

A general solution can be obtained for (5) in the form of Kalman filter,

$$\hat{\mathbf{x}} = \boldsymbol{\mu} + \mathbf{K}\left[\mathbf{z} - \mathbf{H}\boldsymbol{\mu}\right] \text{ and } \hat{\boldsymbol{\alpha}} = \mathbf{P}\mathbf{D}^T\tilde{\mathbf{P}}^{-1}\mathbf{K}\left[\mathbf{z} - \mathbf{H}\boldsymbol{\mu}\right] \tag{6}$$

where $\mathbf{K} = \tilde{\mathbf{P}}\mathbf{H}^T\left[\mathbf{R} + \mathbf{H}\tilde{\mathbf{P}}\mathbf{H}^T\right]^{-1}$ is the Kalman gain and $\tilde{\mathbf{P}} = \boldsymbol{\Sigma} + \mathbf{D}\mathbf{P}\mathbf{D}^T$. Both the FS-MAP/IS-MAP in [1] and the soft-constraint/hard-constraint in [2] can be viewed as special cases, as are shown below.

FS-MAP/Soft-Constraint/Hard-Constraint If $\boldsymbol{\Sigma} \to \mathbf{0}$, then $p(\mathbf{x}|\boldsymbol{\alpha})$ becomes a Dirac delta function and (4) can be further simplified to $p(\boldsymbol{\alpha}|\mathbf{z}, \mathbf{k}^*, \boldsymbol{\theta}^*) \propto p(\mathbf{z}|\mathbf{k}^*, \boldsymbol{\theta}^*, \boldsymbol{\alpha})p(\boldsymbol{\alpha})$, so MAP estimation becomes exactly the same as FS-MAP [1]/Soft-Constraint[2],

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\mathbf{z} - \mathbf{H}(\mathbf{k}^*, \boldsymbol{\theta}^*)(\boldsymbol{\mu} + \mathbf{D}\boldsymbol{\alpha})\|^2_{\mathbf{R}} + \|\boldsymbol{\alpha}\|^2_{\mathbf{P}} \tag{7}$$

and the solution is given by (6) by setting $\boldsymbol{\Sigma} = \mathbf{0}$. Futhermore, if both $\boldsymbol{\Sigma} \to \mathbf{0}$ and $\mathbf{R} \to \mathbf{0}$, then it further reduces to hard-constraint[2] and the solution is given by (6) by setting both $\boldsymbol{\Sigma} = \mathbf{0}$ and $\mathbf{R} = \mathbf{0}$, which can be shown the same as the result obtained in [2] by QR decomposition.

IS-MAP If $\mathbf{P} \to \infty$, then $p(\boldsymbol{\alpha})$ becomes non-informative and (4) reduces to $p(\mathbf{x}, \boldsymbol{\alpha}|\mathbf{z}, \mathbf{k}^*, \boldsymbol{\theta}^*) \propto p(\mathbf{z}|\mathbf{x}, \mathbf{k}^*, \boldsymbol{\theta}^*)p(\mathbf{x}|\boldsymbol{\alpha})$, so MAP estimation is

$$\begin{aligned} \hat{\mathbf{x}} &= \arg\min_{\mathbf{x}, \boldsymbol{\alpha}} \|\mathbf{z} - \mathbf{H}(\mathbf{k}^*, \boldsymbol{\theta}^*)\mathbf{x}\|^2_{\mathbf{R}} + \|\mathbf{x} - \boldsymbol{\mu} - \mathbf{D}\boldsymbol{\alpha}\|^2_{\boldsymbol{\Sigma}} \\ &= \arg\min_{\mathbf{x}} \left\{ \min_{\boldsymbol{\alpha}} \|\mathbf{z} - \mathbf{H}(\mathbf{k}^*, \boldsymbol{\theta}^*)\mathbf{x}\|^2_{\mathbf{R}} + \|\mathbf{x} - \boldsymbol{\mu} - \mathbf{D}\boldsymbol{\alpha}\|^2_{\boldsymbol{\Sigma}} \right\} \\ &= \arg\min_{\mathbf{x}} \|\mathbf{z} - \mathbf{H}(\mathbf{k}^*, \boldsymbol{\theta}^*)\mathbf{x}\|^2_{\mathbf{R}} + \left\|\left[\mathbf{I} - \mathbf{D}\mathbf{D}^T\right]\left[\mathbf{x} - \boldsymbol{\mu}\right]\right\|^2_{\boldsymbol{\Sigma}} \end{aligned} \tag{8}$$

where $\boldsymbol{\Sigma} = \varepsilon^2\mathbf{I}$ and it is exactly the IS-MAP in [1].

3. A JOINT FRAMEWORK FOR FACE HALLUCINATION

In the previous section, both blurring kernel and transformation are assumed to be known. However in practice, they are unknown and so both $p(\mathbf{k})$ and $p(\boldsymbol{\theta})$ are non-informative, posterior (3) becomes

$$p(\mathbf{x}, \mathbf{k}, \boldsymbol{\theta}, \boldsymbol{\alpha}|\mathbf{z}) \propto p(\mathbf{z}|\mathbf{x}, \mathbf{k}, \boldsymbol{\theta}) p(\mathbf{x}|\boldsymbol{\alpha}) p(\boldsymbol{\alpha}) \tag{9}$$

which can be optimized to estimate all the unknowns directly by $\arg\min_{\mathbf{x}, \mathbf{k}, \boldsymbol{\theta}, \boldsymbol{\alpha}} \|\mathbf{z} - \mathbf{H}(\mathbf{k}, \boldsymbol{\theta})\mathbf{x}\|^2_{\mathbf{R}} + \|\mathbf{x} - \boldsymbol{\mu} - \mathbf{D}\boldsymbol{\alpha}\|^2_{\boldsymbol{\Sigma}} + \|\boldsymbol{\alpha}\|^2_{\mathbf{P}}$. However as it will be shown later, instead of directly optimizing the posterior, a better approach is to optimize marginalized posterior,

$$p(\mathbf{k}, \boldsymbol{\theta}|\mathbf{z}) = \int\int p(\mathbf{x}, \mathbf{k}, \boldsymbol{\theta}, \boldsymbol{\alpha}|\mathbf{z}) d\boldsymbol{\alpha} d\mathbf{x} \tag{10}$$

where high dimensional HR image $\mathbf{x}$ and PCA coefficients $\boldsymbol{\alpha}$ are regarded as latent variables. Similar approach by marginalizing latent variables can be found in PCA-based tracking [13] where latent PCA coeffients are marginalized, blind deconvolution [14] where a latent image to be estimated is marginalized, and in multi-image super-resolution [15] where a latent HR image to be estimated is marginalized and [16] where transformation is marginalized instead.

Marginalized posterior $p(\mathbf{k}, \boldsymbol{\theta}|\mathbf{z})$ can be approximated by its lower bound,

$$\begin{aligned}
&\log p(\mathbf{k}, \boldsymbol{\theta}|\mathbf{z}) = \log \int \int p(\mathbf{x}, \mathbf{k}, \boldsymbol{\theta}, \boldsymbol{\alpha}|\mathbf{z}) d\mathbf{x} d\boldsymbol{\alpha} \\
&= \log \int q(\mathbf{x}, \boldsymbol{\alpha}) \tfrac{p(\mathbf{x}, \mathbf{k}, \boldsymbol{\theta}, \boldsymbol{\alpha}|\mathbf{z})}{q(\mathbf{x}, \boldsymbol{\alpha})} d\boldsymbol{\alpha} \\
&\geq \int \int q(\mathbf{x}, \boldsymbol{\alpha}) \log p(\mathbf{x}, \mathbf{k}, \boldsymbol{\theta}, \boldsymbol{\alpha}|\mathbf{z}) d\mathbf{x} d\boldsymbol{\alpha} \\
&- \int \int q(\mathbf{x}, \boldsymbol{\alpha}) \log q(\mathbf{x}, \boldsymbol{\alpha}) d\mathbf{x} d\boldsymbol{\alpha} \\
&= \mathcal{L}(\log p(\mathbf{k}, \boldsymbol{\theta}|\mathbf{z}))
\end{aligned} \tag{11}$$

where $\mathcal{L}(\log p(\mathbf{k}, \boldsymbol{\theta}|\mathbf{z}))$ is the lower bound of $\log p(\mathbf{k}, \boldsymbol{\theta}|\mathbf{z})$. Let $q(\mathbf{x}, \boldsymbol{\alpha}) = \delta(\mathbf{x} - \boldsymbol{\mu} - \mathbf{D}\boldsymbol{\alpha}) q(\boldsymbol{\alpha})$, then

$$\begin{aligned}
&\mathcal{L}(\log p(\mathbf{k}, \boldsymbol{\theta}|\mathbf{z})) = \int q(\boldsymbol{\alpha}) \log \left[p(\mathbf{z}|\mathbf{k}, \boldsymbol{\theta}, \boldsymbol{\alpha}) p(\boldsymbol{\alpha}) \right] d\boldsymbol{\alpha} \\
&- \int q(\boldsymbol{\alpha}) \log q(\boldsymbol{\alpha}) d\boldsymbol{\alpha} + c
\end{aligned} \tag{12}$$

which can be optimized by updating $q(\boldsymbol{\alpha})$, $\mathbf{k}$ and $\boldsymbol{\theta}$ iteratively, which is in fact doing super-resolution, deblurring and registration iteratively.

Super-Resolution by updating $q(\boldsymbol{\alpha})$

Given $\mathbf{k}$ and $\boldsymbol{\theta}$, it can be shown that

$$q(\boldsymbol{\alpha}) = p(\boldsymbol{\alpha}|\mathbf{z}, \mathbf{k}, \boldsymbol{\theta}) \propto p(\mathbf{z}|\mathbf{k}, \boldsymbol{\theta}, \boldsymbol{\alpha}) p(\boldsymbol{\alpha}) = \mathcal{N}(\boldsymbol{\alpha}; \hat{\mathbf{m}}, \hat{\mathbf{P}}) \tag{13}$$

which is the posterior given blurring kernel and transformation, i.e. essentially FS-MAP[1]/Soft-Constraint[2], with mean $\hat{\mathbf{m}}$ given in (6) by setting $\boldsymbol{\Sigma} = \mathbf{0}$ and covariance $\hat{\mathbf{P}} = \mathbf{D}^T \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{D} + \mathbf{P}^{-1}$, and $\hat{\mathbf{x}} = \boldsymbol{\mu} + \mathbf{D}\hat{\mathbf{m}}$. Note that in FS-MAP[1]/Soft-Constraint[2], only mean $\hat{\mathbf{m}}$ is used as the MAP estimation. It will be shown later that in addition to the mean, covariance $\hat{\mathbf{P}}$, which represents uncertainties of the MAP estimation, is important and will be used to estimate the blurring kernel and transformation.

Deblurring by updating $\mathbf{k}$

Given $q(\boldsymbol{\alpha})$ and $\boldsymbol{\theta}$, it can be shown that the lower bound can be rewritten in quadratic form

$$\mathcal{L}(\mathbf{k}) = \mathbf{k}^T \mathbf{A}(\hat{\mathbf{P}}) \mathbf{k} + 2\mathbf{b}^T(\hat{\mathbf{m}}) \mathbf{k} + c \tag{14}$$

where $\mathbf{A}(\hat{\mathbf{P}})$ and $\mathbf{b}(\hat{\mathbf{m}})$ involve the previously estimated mean $\hat{\mathbf{m}}$ and covariance $\hat{\mathbf{P}}$. It can be optimized using constraints of $\mathbf{k}$, L1 norm constraint $|\mathbf{k}|_1 = 1$ and non-negativity constraint $\mathbf{k} \geq \mathbf{0}$ via quadratic programming with constraints. For Gaussian blurring kernel, additional symmetric constraint can be applied so that the dimensionality of the blurring kernel can be further reduced.

Registration by updating $\boldsymbol{\theta}$

Given $q(\boldsymbol{\alpha})$ and $\mathbf{k}$, it can be shown that the lower bound reduces to

$$\mathcal{L}(\boldsymbol{\theta}) = \|\mathbf{z} - \mathbf{H}(\boldsymbol{\theta})(\boldsymbol{\mu} + \mathbf{D}\hat{\mathbf{m}})\|^2 + Tr\left[\mathbf{D}^T \mathbf{H}^T(\boldsymbol{\theta}) \mathbf{H}(\boldsymbol{\theta}) \mathbf{D} \hat{\mathbf{P}}\right] + c \tag{15}$$

Note covariance $\hat{\mathbf{P}}$ can be eigen-decomposed by $\hat{\mathbf{P}} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^T$, where $\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_L]$ and $\boldsymbol{\Lambda} = \mathrm{diag}([\lambda_1, \cdots, \lambda_L])$, $\{\mathbf{v}_l\}_{l=1}^L$ and $\{\lambda_l\}_{l=1}^L$ are eigenvectors and eigenvalues of $\hat{\mathbf{P}}$ respectively. Now the second term becomes $\sum_{l=1}^{L} \lambda_l \|\mathbf{H}(\boldsymbol{\theta}) \mathbf{D} \mathbf{v}_l\|^2$, so

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

Fig. 1. Iterative update. (a) LR $\mathbf{z}$. (b) 1st iteration $\hat{\mathbf{x}}'$. (c) 10th iteration $\hat{\mathbf{x}}'$.(d) Final iteration $\hat{\mathbf{x}}'$. (e) Final iteration $\hat{\mathbf{x}}'$ without using covariance.

$$\mathcal{L}(\boldsymbol{\theta}) = \|\mathbf{z} - \mathbf{H}(\boldsymbol{\theta})(\boldsymbol{\mu} + \mathbf{D}\hat{\mathbf{m}})\|^2 + \sum_{l=1}^{L} \lambda_l \|\mathbf{H}(\boldsymbol{\theta}) \mathbf{D} \mathbf{v}_l\|^2 + c \tag{16}$$

It can be observed that optimal transformation needs to not only reduce the first term which is due to the observation model, but also reduce the second term which involves covariance $\hat{\mathbf{P}}$, which represents uncertainties of $\hat{\mathbf{m}}$. (16) can be optimized in a similar way as the LK algorithm but in the same spirit of [5] to avoid interpolating LR images.

Iterative updating is shown in Fig. 1, where homography is used as the transformation. Note that given an LR image $\mathbf{z}$, the actual super-solved LR image is the warped HR image $\hat{\mathbf{x}}' = \mathbf{W}\hat{\mathbf{x}}$, which will be used in experiments in section 4, rather than the latent HR image $\hat{\mathbf{x}}$. It can be seen from Fig. 1 the importance of covariance $\hat{\mathbf{P}}$ in the estimation.

4. RESULTS

Experiments were carried out to test the performance of proposed joint framework for face hallucination. HR faces from FERET database [17] are used as the training set while 100 random HR faces from BioID database [18] as the testing set, where all faces are normalized by affine transformation using eyes and mouth centres. Given HR face images $\mathbf{x}$ shown in Fig. 2(a), they are perturbed by random homography to HR faces $\mathbf{x}'$ in Fig. 2(b), which are then blurred with Gaussian kernel, sub-sampled by $s = 4$ to 15×15 and added with Gaussian noise to obtain LR faces $\mathbf{z}$ in Fig. 2(c).

Results of FS-MAP/Soft-Constraint using the ground-truth (GT) blurring kernel and identity transformation is shown in Fig. 2(d). Results of FS-MAP/Soft-Constraint using a different blurring kernel from the GT blurring kernel and the GT transformation is shown in Fig. 2(e). Results of FS-MAP/Soft-Constraint using a different blurring kernel from the GT kernel and identity transformation is shown in Fig. 2(f). It can be noted that FS-MAP/Soft-Constraint is sensitive to errors in both the blurring kernel and the transformation.

Results of the proposed method in Fig. 2(g) are visually close to results in Fig. 2(h), which uses both the GT blurring kernel and the GT transformation. In addition, in some cases, as shown in the last two rows of Fig. 2, the proposed method performs even better, in terms of PSNR and SSIM, in comparison with FS-MAP/Soft-Constraint using both the GT kernel and the GT transformation. It is due to inaccura-

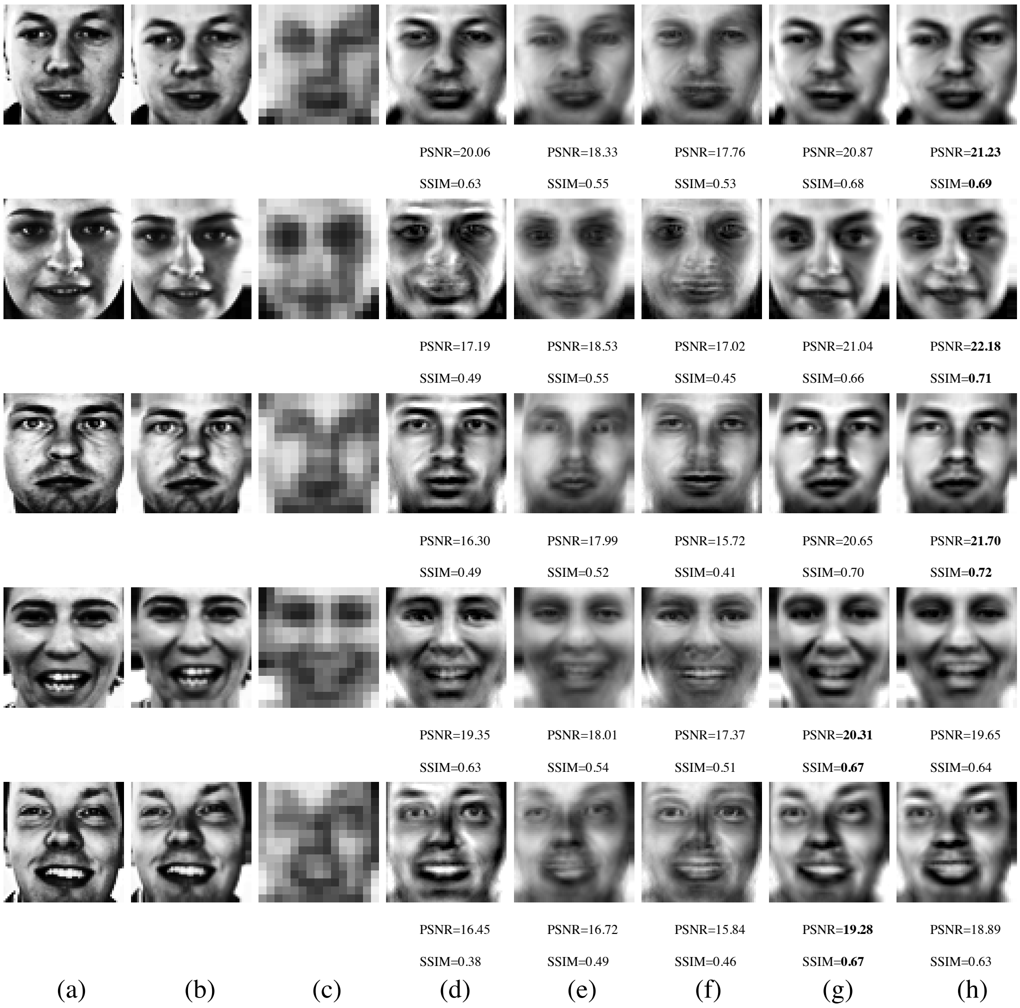

Fig. 2. Examples of face hallucination results. (a) HR faces. (b) warped HR faces. (c) LR faces obtained by blurring HR faces in (b) with Gaussian blurring kernel with $\sigma = 0.6s$, and then sub-sampling by $s = 4$ to 15×15 and adding Gaussian noise. (d) Results of FS-MAP/Soft-Constraint using the GT blurring kernel and identity transformation. (e) Results of FS-MAP/Soft-Constraint using a Gaussian blurring kernel with $\sigma = 0.2s$ and GT transformation. (f) Results of FS-MAP/Soft-Constraint using a Gaussian blurring kernel with $\sigma = 0.2s$ and identity transformation. (g) Results of proposed method. (h) Results of FS-MAP/Soft-Constraint using the GT blurring kernel and the GT transformation.

cies of face normalization using affine transformation, which is insufficient to align faces with large out-of plane rotation or large variations of facial expression and is improved by the proposed method. The performance of the proposed method in comparison with FS-MAP/Soft-Constraint using different settings can also be seen in table 1, which shows its robustness to errors in both the blurring kernel and the transformation.

5. CONCLUSIONS

The paper proposed a joint framework for face hallucination incorporating face deblurring and registration. The joint framework not only directly hallucinates LR faces without interpolating LR images, but also deblurs and aligns LR faces iteratively to improve the performance of face hallucination.

Methods	FS-MAP[1]/Soft-Constraint[2]				Proposed
Blurring Kernel σ	GT	$0.2s$	$0.2s$	GT	Estimated
Transformation	**I**	GT	**I**	GT	Estimated
PSNR (dB) mean	19.12	18.89	17.89	21.44	20.64
max	23.49	21.40	20.74	24.25	23.81
min	15.61	16.72	15.72	18.89	18.38
SSIM mean	0.56	0.57	0.51	0.68	0.65
max	0.70	0.67	0.65	0.80	0.78
min	0.37	0.48	0.40	0.55	0.52

Table 1. Performance comparison

Without the need for accurate face registration and prior knowledge of blurring kernels, it is robust to errors in face registration and blurring kernel. Experimental results showed the robust performance of the proposed method.

6. REFERENCES

[1] D. Capel and A. Zisserman, "Super-resolution from multiple views using learnt image models," in *Proc. IEEE CVPR*, 2001, pp. 627–634.

[2] C. Liu, H. Y. Shum, and W. T. Freeman, "Face hallucination: Theory and practice," *IJCV*, vol. 75, no. 1, pp. 115–134, 2007.

[3] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE TIP*, vol. 19, no. 11, pp. 2861–2873, 2010.

[4] B. K. Gunturk, A. U. Batur, Y. Altunbasak, M. H. Hayes, and R. M. Mersereau, "Eigenface-domain super-resolution for face recognition," *IEEE TIP*, vol. 12, no. 5, pp. 597–606, 2003.

[5] G. Dedeoglu, T. Kanade, and S. Baker, "The asymmetry of image registration and its application to face tracking," *IEEE TPAMI*, vol. 29, no. 5, pp. 807–823, 2007.

[6] X. Wang and X. Tang, "Hallucinating face by eigentransformation," *IEEE TSMC(C)*, vol. 35, no. 3, pp. 425–434, 2005.

[7] K. Jia and S. Gong, "Generalized face super-resolution," *IEEE TIP*, vol. 17, no. 6, pp. 873–886, 2008.

[8] W. Liu, D. Lin, and X. Tang, "Hallucinating faces: TensorPatch super-resolution and coupled residue compensation," in *Proc. IEEE CVPR*, 2005, pp. 478–484.

[9] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE TPAMI*, vol. 24, no. 9, pp. 1167–1183, 2002.

[10] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, "Learning low-level vision," *IJCV*, vol. 20, no. 1, pp. 25–47, 2000.

[11] S. Baker and I. Matthews, "Lucas-Kanade 20 years on: A unifying framework," *IJCV*, vol. 56, no. 3, pp. 221 – 255, 2004.

[12] T. F.Cootes, G. J. Edwards, and C. J.Taylor, "Active appearance models," *IEEE TPAMI*, vol. 23, no. 6, pp. 681–685, 2001.

[13] D. A. Ross, J. Lim, R. S. Lin, and M. H. Yang, "Incremental learning for robust visual tracking," *IJCV*, vol. 77, pp. 125–141, 2000.

[14] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Efficient marginal likelihood optimization in blind deconvolution," in *Proc. IEEE CVPR*, 2011, pp. 2657–2664.

[15] M. E. Tipping and C. M. Bishop, "Bayesian imge super-resolution," in *NIPS*, 2003, pp. 1279–1286.

[16] L. C. Pickup, D. P. Capel, S. J. Roberts, and A. Zisserman, "Bayesian image super-resolution, continued," in *NIPS*, 2007, pp. 1089–1096.

[17] P. J. Philips, H. Moon, P. Rauss, and S. A. Rizvi, "The FERET evaluation methodology for face-recognition algorithms," *IEEE TPAMI*, vol. 22, no. 10, pp. 1090–1104, 2000.

[18] O. Jesorsky, K. Kirchberg, and R. Frischholz, "Robust face detection using the Hausdorff distance," in *Proc. IAPR International Conference on Audio- and Video-Based Personal Authentication*, 2003, pp. 90–95.