NON-PARAMETRIC BAYESIAN ALIGNMENT AND RECOVERY OF OCCLUDED FACE USING DIRECT COMBINED MODEL

Ching-Ting Tu and Jenn-Jier James Lien
Robotics Laboratory, Department of Computer Science and Information Engineering
National Cheng Kung University, Tainan, Taiwan, R.O.C.

Keywords: Principal Component Analysis (PCA), Eigenface, Statistical Image Models.

Abstract: This paper focuses on the problem of recovering the occluded facial image automatically with the aid of domain specific prior knowledge and no manual face alignment or user-specified occlusion region is needed. The robust alignment and occlusion recovery are solved sequentially by a novel recovery scheme called the direct combined model (DCM). Local occluded facial patches are recovered by utilizing the information propagated from other non-occluded patches and is further constrained by a global facial geometry. The error residue between the recovered result and the geometric constraint is then used for updating the parameter of alignment function for the next iteration. Into this recovering framework, DCM efficiently and robustly updates the results of recovering and aligning based on a compact statistic model representing the prior updating knowledge. Our extensive experiment results demonstrate that the recovered images are quantitatively closer to the ground truth with no manual alignment and occlusion dection.

1 INTRODUCTION

Recently, with the help of a large collection of other facial images, a number of face occlusion recovery techniques have been developed. Park et al. (Park et al., 2005) and Saito et al. (Saito et al., 1999) propose methods to remove occlusion and reconstruct facial images with principal component analysis (PCA). However, the results synthesized via the original PCA process are highly affected by the appearance and location of the occlusion. Subsequently, Hwang et al. (Hwang et al., 2003) and Mo et al. (Mo et al., 2004) presented methods to recover occluded faces using two separate eigenspaces sharing the same coefficients. However, the occluded and non-occluded appearances have an entirely different character, recovery results using the same weights for two different spaces tends to be inaccurate. Furthermore, the learning-based facial recovery methods are required automatically detecting the occluded area and aligning the test image with training examples. However, pervious methods, e.g. (Hwang et al., 2003; Lin et al., 2009; Mo et al., 2004), are need a user either to specify the occlusion region or to align the occluded face images.

This paper unifies the tasks of automatic occluded face recovery, detection, and alignment in a Bayesian framework, and solves these problems sequentially by a novel particle-based recovery scheme. In this framework, we introduce a novel learning algorithm, DCM, to deterministically and robustly infer the affine parameters and the recovered facial appearance.

2 DIRECT COMBINED MODEL

The DCM algorithm assumes there are two related classes: class X and class Y, e.g. the facial appearances of occluded and non-occluded patches or the affine parameter and the corresponding facial appearance. Let the structure of the coupled training dataset be $(x_1, y_1),(x_2, y_2),\ldots,(x_N, y_N)$, where N is the total number of (x, y) feature pairs. The combined principal space $[U_X^T U_Y^T]^T$ for minimizing the energy function $E(U_X, U_Y, W)$:

$$\sum_{i=1}^{n} \| x_i - U_X w_i - \bar{x} \|_2^2 + \sum_{i=1}^{n} \| y_i - U_Y w_i - \bar{y} \|_2^2 \qquad (1)$$

can be solved by SVD process, where and w_i is the weight set calculated by projecting each feature pair, (x, y), onto the combined principal space $[U_X^T U_Y^T]^T$, and $(\bar{x}, \bar{y})$ is the mean vector pair.

Tu C. and James Lien J. (2010).
NON-PARAMETRIC BAYESIAN ALIGNMENT AND RECOVERY OF OCCLUDED FACE USING DIRECT COMBINED MODEL.
In *Proceedings of the International Conference on Computer Vision Theory and Applications*, pages 495-498
DOI: 10.5220/0002833704950498

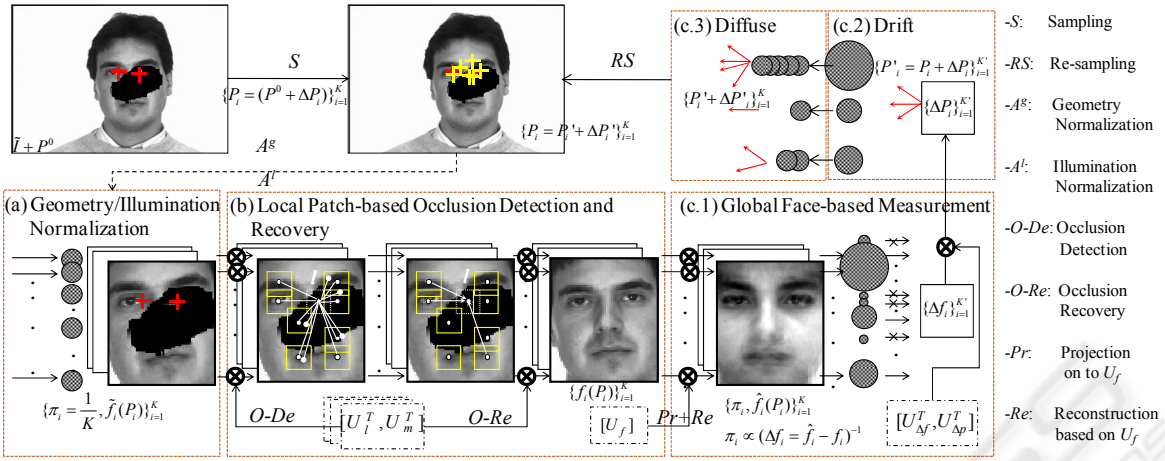

Figure 1: Workflow of testing process consists of (a) geometry/illumination normalization, (b) patch-based occlusion detection and recovery, (c) probabilistic propagation: (c.1) face-based measurement, (c.2) drift, and (c.3) diffuse.

According to (De la Torre et al., 2001), the common drawback of such combined formulation, $[U_X^T U_Y^T]^T$, is that it is not suitable for predicating from one class to another. We exploit the fundamental property of SVD in Eq.(1), the unpredictable problem is resolved via a MMSE criterion.

$$\hat{y}(x) = \arg\min_{y} \iint (y - \hat{y})^2 P(x, y) dX dY \quad (2)$$

where $P(x, y)$ is the joint probability of feature vectors x and y, and $\hat{y}(x)$ is the estimated vector y for a given x. According, it is the expected value or vector of the posterior probability of y for a given observation x, $E[Y \mid X = x]$.

Under the assumption that the joint distribution of X and Y in Eq. (1) is a multivariate normal density, Eq.(2) further is derived:

$$\hat{y}(x) = \bar{y} + U_Y U_X^{\dagger} (x - \bar{x}) \quad (3)$$

where $U_X^{\dagger}$ is the right inverse matrix of U_X that can be approximated by the SVD algorithm. Eq. (3) is defined as the DCM transformation from X to Y, where the regression matrix $U_Y U_X^{\dagger}$ can be calculated in off-line. Especially, in contrast to the standard multiple multivariate linear regression approach, the DCM transformation extracts only a few significant feature pairs to represent the relevant information, and thus, the major features of the X-Y correlation are better captured.

3 DCM-BASED BAYESIAN ALIGNMENT AND RECOVERY

We introduce a Bayesian framework to bind the tasks of occluded facial image alignment and recovery and the occlusion detection together, and address such a task by a novel particle-based scheme to model the posterior probability density function. The proposed process takes both the global and local facial appearance components of the input image into account, and sequentially recovers and propagates the particles by embedding with the DCM algorithm.

3.1 Unified Probability Model

Following the sequential recovery algorithm, the recovered facial appearance f in an image I is inferred from the previously recovered result f':

$$f^* = \arg\max_{f} \iint_{\xi,b} P(f \mid \xi, b, I) P(\xi, b \mid f', I) d\xi db \quad (4)$$

where the affine parameter ξ and the PCA weight vector b are included for aligning f with training examples and for guaranteeing the global facial geometric structure of f. It naturally decomposes into two terms:

Posterior Probability $P(f \mid \xi, b, I)$. Since the weight vector b is independent from the affine parameter ξ, the posterior density term are factored as $P(b \mid f) P(f \mid \xi, I)$ in the recovery stage (sec 3.2).

Propagation Probability $P(\xi, b \mid f', I)$. It is defined as the probabilistic propagation formulation:

$$p(\xi, b \mid f') = \iint_{\xi', b'} P(\xi, b \mid \xi', b') P(\xi', b' \mid f') d\xi' db' \quad (5)$$

where ξ' and b' are from the previous iteration.

Fig. 1 illustrates the proposed particle based solution for this Bayesian framework. In practice, particles are sampled based on two inner corner

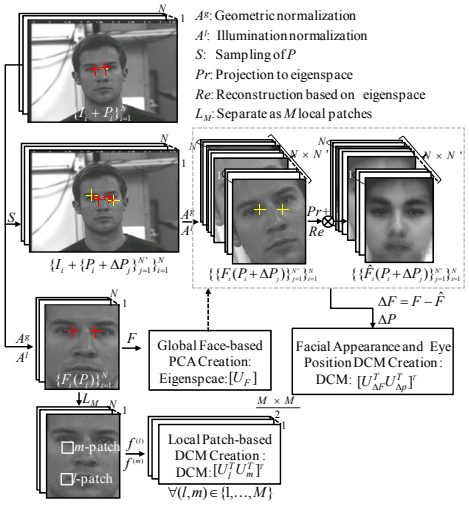

Figure 2: Training process consists of face-based PCA creation, patch-based DCM creation, and facial appearance error-to-eyes position updating DCM creation.

points of the eyes P by an affine transformation function A^g via a parameter ξ. The initial position of two inner corner points of the eyes P^0 (it is probably occluded) is roughly detected, and K initializations particles are randomly generated around P^0: $\{P = P^o + \Delta P\}$. The particle weight π is measured with its coincidence with the learned eigenspace U_f (Initial weights of all particles are equal). Further, to suppress the effects of illumination differences between different facial appearances, a illumination normalization scheme A^l is performed.

3.2 Local-based Occlusion Recovery

The posterior density, $P(f \mid \xi, b, I)$ for recovery stage is modeled by a Markov network embedded with the proposed DCM algorithm. Different from the common patch-based Markov network approaches (Freeman et al., 2000; Liu et al., 2001; Sudderth et al., 2003) that selects the recovered patches from the training database, the current approach recovers patches by other non-occluded patches via the DCM transformation.

Learning Patch Pair *l*-to-*m* DCM $[U_l^T U_m^T]^T$ (Fig. 2). For each patch $l, m \in \{1, 2, ..., M\}$ (each f is composed of M patches), the *l*-*m* patch pairs of N training facial appearances $\{f(P)\}$ are used to train the combined principal space $[U_l^T U_m^T]^T$ and the DCM transformation from l to m by Eq. (1) and Eq. (3), respectively.

Occlusion Detection (Fig. 1.b). The confidence of visibility of patch l is written as c_l which is directly proportional to the difference between the original local texture detail of patch l and the reconstructed texture by the bidirectional DCM transforms from l to m and from m back to l, where the patch m is one of neighbor patches of l.

Occlusion Recovery (Fig. 1.b). We solve the defined Markov network by the nonparametric belief propagation method (NBP) (Sudderth et al., 2003), but the recovery order is from non-occluded patches to the occluded patches sorted based on their confident values, i.e. the *c*-value.

3.3 Global-based Face Alignment

After the recovery, the higher-weighted particles are chosen to form the distribution of $P(\xi', b' \mid f')$ in the face-based measurement step of the probabilistic propagation stage (Fig. 1.c.1), where only these correctly aligned ones will be treated as the updating initializations in the following steps of re-randomization. According, the summarization of the current recovered results, $\{f(P), \pi'\}$, is the mean of these particles, $\bar{f} = E[f] = \sum_{i=1}^{N} f_i \pi_i$.

Learning the Position-facial Appearance DCM $[U_{\Delta f}^T U_{\Delta P}^T]^T$ (Fig. 2). Each training image I generates N' perturbed facial appearances, $\{f(P + \Delta P)\}$, by disturbing elements of the manually labeled position P. Subsequently, $N' \times \Delta P$ of N training images and their corresponding facial appearance difference generated by U_f are used to train the DCM combined principal space $[U_{\Delta f}^T U_{\Delta P}^T]^T$ and the DCM transformation from Δf to ΔP, respectively.

Face Alignment. The drift step (Fig. 1.c.2) updates positions $\{P' = P + \Delta P\}$ from the given facial differences, $\{\Delta f\}$ based on the combined space $[U_{\Delta f}^T U_{\Delta P}^T]^T$ in order to form the transition probability, $P(\xi, b \mid \xi', b')$. Finally, a diffuse step (Fig. 1.c.3) is done on these higher-weighted particles to generate several copies and shift them to the neighbors of the updated position, $\{P' + \Delta P'\}$. The new set of particles then forms the distribution of $P(\xi, b \mid f')$ for the following step.

4 EXPERIMENTAL RESULTS

The performance of the proposed recovery system was evaluated by performing a series of experimental trials using training and testing databases comprising 100 and 50 facial images, respectively, where specific facial feature regions of

the testing images are occluded manually. The normalized images are 28x96 pixels and the patch size and overlap size are 13 and 5, respectively.

Fig. 3 presents representative examples of the reconstruction results. Table 1 presents the average recovery and alignment errors computed over all the images in the testing database. Fig. 4 compares the recovery results obtained using the proposed DCM method with those obtained by methods in (Hwang et al., 2003 and in Park et al., 2005).

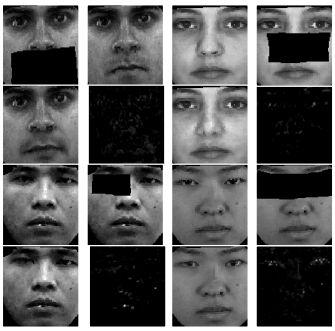

Figure 3: Reconstruction results of the DCM schemes: (1st and 3rd rows) occluded and ground truth facial images and (2nd and 4th rows) reconstructed facial images and its difference from its ground truth.

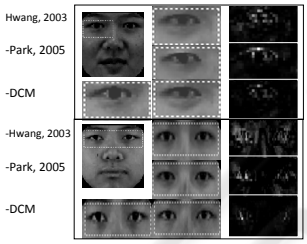

Figure 4: Two examples of reconstruction results and errors using three eigen-based method. (1st column): original occluded images and the occlude features.

5 CONCLUSIONS

This study has presented a Bayesian framework for sequential facial occlusion alignment, detection, and recovery through a DCM-based algorithm. By considering both local and global facial structures, our recovered results closely resemble the ground truth facial appearances. Overall, the proposed method is a promising way to improve the performance of existing automatic face recognition, facial expression recognition, and facial pose estimation applications.

Table 1: Average and standard deviation of the position *P* and the recovered *f* estimation errors for all images in testing database by different levels of occlusion.

Facial features	Ave. Error (Pixel/Grayvalues)		Std. Dev. (Pixel/Grayvalues)		Occlusion Area
	P	*f*	*P*	*f*	
Left Eye	0.7	6.6	0.5	1.7	10%
Right Eye	0.6	6.5	0.6	1.8	10%
Both Eye	0.7	6.6	0.6	16	24%
Nose	1.0	7.0	1.2	2.0	16%
Mouth	1.6	6.8	1.5	1.9	20%

REFERENCES

De la Torre, F. D.and Black, M. J. (2001). Dynamic Coupled Component Analysis. *CVPR*.

Freeman, W.T., Pasztor, E.C., and Carmichael, O.T. (2000). Learning Low-Level Vision. *IJCV*, 40: 25-47.

Hwang, B.W. and Lee, S.W. (2003). Reconstruction of Partially Damaged Face Images Based on a Morphable Face Model. *IEEE Trans. on PAMI*, 25: 365-372.

Lin, D. and Tang, X. (2009). Quality-Driven Face Occlusion Detection and Recovery. *ICCV*.

Liu, C., Shum, H.Y., and Zhang, C.S. (2001). A Two-Step Approach to Hallucinating Faces: Global Parametric Model and Local Nonparametric Model. *CVPR*.

Mo, Z., Lewis, J.P., and Neumann, U. (2004). Face Inpainting with Local Linear Representations. *BMVC*.

Park, J.S., Oh, Y.H., Ahn, S.C., and Lee, S.W. (2005). Glasses Removal from Facial Image Using Recursive PCA Reconstruction. *IEEE Trans. on PAMI*, 27: 805-811.

Saito, Y., Kenmochi, Y., and Kotani, K. (1999). Estimation of Eyeglassless Facial Images Using Principal Component Analysis. *ICIP*.

Sudderth, E., Ihler, A., Freeman, W., and Willsky, A. (2003). Nonparametric Belief Propagation. *ICCV*.