Recognition of Low-Resolution Faces Using Multiple Still Images and Multiple Cameras

Pablo H. Hennings-Yeomans
ECE Department
Carnegie Mellon University
phh@cmu.edu

Simon Baker
Microsoft Research
Microsoft Corporation
sbaker@microsoft.com

B.V.K. Vijaya Kumar
ECE Department
Carnegie Mellon University
kumar@ece.cmu.edu

***Abstract*— We propose a new algorithm for recognition of low-resolution faces for cases when multiple images of the same face are available at matching. Specifically, this is the case of multiple-frame, or video, face recognition, and recognition with multiple cameras.**

Face recognition degrades when probe faces are of lower resolution than those available for training. There are two paradigms to alleviate this problem, but both have clear disadvantages. One is to use super-resolution algorithms to enhance the image, but as resolution decreases, super-resolution becomes more vulnerable to environmental variations, and it introduces distortions that affect recognition. On the other hand, it is possible to match in the low-resolution domain by downsampling the training set, but this is undesirable because features important for recognition depend on high frequency details that are erased by downsampling.

We recently proposed a new framework for recognition that is different from these two paradigms, and we have shown that recognition is considerably improved for still-image face recognition. In this work, we show that the proposed method can be generalized to use a stream of frames, and produce even better recognition performance. The new algorithm incorporates the underlying assumptions of super-resolution methods with subspace distance metrics used for classification, simultaneously. We extend our previous formulation to use multiple frames, and we show that it can also be generalized to use multiple image formation processes, modeling different cameras.

In this manuscript we consider two scenarios: face recognition with multiple frames, and the case of having a camera sensor network with dual cameras in a master-slave setting. Using the Multi-PIE database, our results show an increase of rank-1 recognition accuracy of about 6% for both cases compared to the single frame settings.

I. INTRODUCTION

As multimedia applications become ubiquitous, thanks to the current capacity of storage devices and increasing processor speeds, face recognition algorithms are being extended and rewritten to take advantage of video and other sensor modalities that produce a continuous stream of frames instead of a single image.

We can see the use of video sensors almost everywhere. For example, at an airport gate entrance, video cameras are being used instead of still image digital cameras. Cellphones are equipped now with cameras capable of capturing a sequence of frames instead of a single image. Camcorders are everywhere, and the need to parse video digital libraries to extract specific content (such as faces) is soon to become a daily activity of search engines.

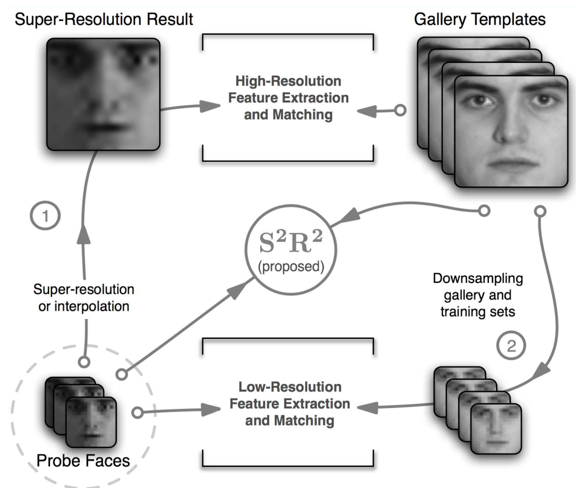

Fig. 1. Standard approaches to matching a low resolution probe to a high resolution gallery. (1) Upsampling the probe (interpolation or super-resolution) and then matching. (2) Downsampling the gallery and then matching. In this paper we propose an alternative algorithm that can outperform these two approaches.

The applications of face recognition using multiple still images are not limited to entertainment, education, or surveillance. As pointed out in a recent survey [8], critical forensics identification sometimes requires the analysis of tens of thousands of hours of video. Similarly, finding missing people using low-resolution video from traffic cameras requires dedicated experts looking at video from tens or hundreds of cameras.

Many times we don't have the option of throwing away an image with a face we may think it is of poor quality, as in some of the applications mentioned above. When the images are of very low resolution image enhancement algorithms, such as super-resolution methods, produce results that do not improve recognition by matching low-resolution faces [14]. This is the problem that motivates the study presented in this work. We are looking for new ways of performing face recognition that can be used in probes with multiple frames of low resolution faces, and at the same time use the training set available with a resolution at which we have a classifier that performs reasonable face recognition.

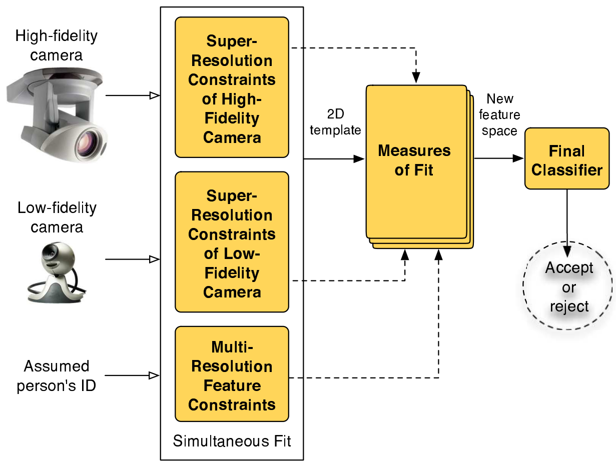

Fig. 2. Proposed procedure for recognition of low-resolution faces. A 2D template that simultaneously fits face feature constraints and super-resolution constraints is used to extract measures of fit as new features for recognition.

Figure 1 illustrates the proposed approach, as compared to the standard paradigms of face recognition.

In our work, we propose a new framework for recognition that models the properties of each frame [24], as in super-resolution methods [4][10], and we assume that a feature extraction strategy (or, in general, a discriminant function) has been selected in advance, as for any face recognition system. Our algorithm then uses these *base* face features and super-resolution priors to extract a high-resolution template that simultaneously fits both, super-resolution and face-features constraints. From this template a new set of quality-of-fit features are computed for recognition (see Fig. 2).

Additionally, we show that the modeling of image formation of multiple frames, in our framework, can be readily generalized to account for input from multiple cameras. We give preliminary results that have been obtained using a large database of more than 300 persons with illumination variations.

A. Previous Work

We are at the intersection of several fields, but our interest is in face recognition. There are two paradigms for face recognition of low resolution faces. One is to use super-resolution algorithms to enhance the image before recognition, as has been explored recently in [17], [16], [23]. Approaches that perform super-resolution in the eigenface domain have been investigated in [12], [22], while super-resolution in the face tensor-space was proposed in [15]. The problem here is that as resolution decreases, super-resolution becomes more vulnerable to environmental variations, and it introduces distortions that affect recognition [7][14], specially when target faces cannot be included in the process of basis selection (this is a real scenario that has been specified in recent face recognition protocols [20]), as assumed by most eigenface -based algorithms. On the other hand, it is possible to match in the low resolution domain by downsampling the training set, but this is undesirable because features important for recognition depend on high frequency details that are erased by downsampling. A recent study that investigates the effect of resolution in face recognition is given in [6]. In previous work [14], we showed that this approach of matching in the low-resolution domain is better than applying super-resolution when faces are of very low resolution.

An approach that does not perform super-resolution, but learns distortions in a class-by-class basis with the purpose of recognition is proposed in [2]. The authors propose to extract signatures of illumination and pose from generic training faces to represent a shape-illumination manifold [1]. Then, their algorithm learns a probabilistic model for camera downsampling artifacts from a gallery of videos at enrollment, being person-specific. Matching of a probe video to a gallery video is performed by, first, re-illuminating the face in the probe, and then fitting it to the models of pose and downsampling artifacts from the class of the gallery video. After fitting across the gallery, the algorithm outputs the class label with the highest likelihood, after some heuristic considerations for robustness against extreme pose variations. A limitation of this approach, however, is that it requires a video sequence at enrollment, making it impractical for some face recognition scenarios, such as in the FRGC protocol [20]. In fact, in the most recent face recognition competitions, which simulate realistic scenarios for current demands, when matching video probes, only a mugshot is provided per gallery class [19].

In [15], the authors propose an approach that first computes high-resolution tensor-space parameters from low-resolution ones, and then perform recognition or super-resolution. In contrast, we do not perform super-resolution in the minimum mean-squared-error sense, but use super-resolution models as constraints simultaneously with face-features constraints. Moreover, our approach is more general because it is not limited to a generative multi-factor model, but can also be discriminative.

A recent review on face recognition using multiple still images is [24]. An overview of super-resolution is given by [18]. Moving from multiple frames to multiple cameras touches on sensor networks, which we don't discuss here. A review on camera sensor networks for biometric applications can be found in [8].

B. Baseline Algorithms

To introduce the notation, we review the standard approaches to the two paradigms of low-resolution face matching mentioned above.

A first standard approach to match a low-resolution probe face (i.e., an evaluation-set face, also called query face), y_p, to a gallery image, x_g (which has a higher resolution) is to produce an estimate $\tilde{x}_p$ of the desired face, x_p, from y_p, and apply a feature extractor, $\mathbf{F}$, to compute the distance, $D(\mathbf{F}x_g, \mathbf{F}\tilde{x}_p)$. This algorithm corresponds to super-resolution or interpolation followed by classification.

We select Tikhonov regularization [13], [25] as the base super-resolution in the proposed algorithm. By using a simple super-resolution algorithm, we aim to show the effectiveness of the proposed approach. Tikhonov regularization obtains $\tilde{x}_p$ by minimizing the objective function

$$\|\mathbf{B}x - y_p\|^2 + \alpha^2\|\mathbf{L}x\|^2, \tag{1}$$

where the matrix $\mathbf{B}$ is an image formation model that transforms a high-resolution image x to its low-resolution version y. $\mathbf{B}$ may be different for every frame in a video sequence, and it encompasses a decoupled *point spread function* (PSF) of the camera, as in [4], [3]. A Gaussian kernel is used as the lens PSF and an averaging kernel for the sensor PSF. Here, $\mathbf{L}x$ is a vector with first-derivative approximations, but second derivatives and other types of constraints represented in $\mathbf{L}$ may be used. The scalar α is a regularization parameter [13].

A second approach to classify y_p, shown in Figure 1, path 2, is to downsample x_g to obtain y_g, with the same resolution as of y_p. This requires downsampling the training images and computing a new feature matrix $\mathbf{F}_L$ in the low-resolution domain. Then, the distance $D(\mathbf{F}_L y_g, \mathbf{F}_L y_p)$ can be computed and recognition is carried out in the usual manner by either comparing this distance to a threshold (verification), or comparing distances computed for each gallery image (identification). Matching with low-resolution faces, however, is undesirable since in the cases considered here, x_g and the whole training set are at a better resolution for classification, i.e., reducing the resolution will degrade the recognition performance.

II. ALGORITHMS

Recently, in [14], we proposed what we call simultaneous super-resolution for recognition ($\mathbf{S^2R^2}$). It consists of a two-step approach that uses constraints of a super-resolution algorithm and features from a classifier trained with images having the desired resolution.

Recognition is performed using the input image without obtaining a super-resolution image (in the minimum mean-squared error sense), although the first step is to compute a new template that simultaneously fits super-resolution constraints as well as feature constraints starting with the probe image and the features of the gallery image we are using for matching. The second step is to measure the fitness of the obtained template to each of these constraints and then constructing a new feature vector using these meassures. A new classifier performs recognition in this feature space.

We explain here the matching algorithm for three different scenarios: still image matching, multiple-frame matching, and matching with multiple cameras. The training algorithm is similar for these three cases. It consists of finding a set of regularization parameters and the classifier operating in the new feature space, defined below. Details are given in [14].

A. Still Image

Formally, $\mathbf{S^2R^2}$ matching is as follows. Matching a given low-resolution probe face, y_p, and a gallery image of higher resolution, x_g, from the kth class requires computing (or to look up) the features $f_g^{(k)} = \mathbf{F}x_g$. The probe image and these features are the input to the matching algorithm. $\mathbf{F}$ is a feature extractor matrix (v.g., Fisherfaces [5]), computed with training images at the desired resolution, as would be trained for any classifier.

The first step in $\mathbf{S^2R^2}$ finds a new template, $\hat{x}_p^{(k)}$, by minimizing

$$\|\mathbf{B}x - y_p\|^2 + \alpha^2\|\mathbf{L}x\|^2 + \beta^2\|\mathbf{F}x - f_g^{(k)}\|^2, \tag{2}$$

where $\mathbf{B}$, $\mathbf{L}$ and α are defined as in Eq. 1, and β is an additional regularization parameter. Each term in Eq. 2 is a set of constraints representing three different models: the image formation model, the super-resolution prior, and the feature extraction model. The optimal template, $\hat{x}_p^{(k)}$, contains information about the fitness of the low-resolution y_p to the class of x_g.

The second and last step of the process is to use the resulting $\hat{x}_p^{(k)}$ template to extract features that encode this class-specific fitness. A final classifier operating in this new feature space is trained to perform verification or identification.

For the case of face verification we make a binary decision with

$$w \cdot q(\hat{x}_p^{(k)}) \;\rightarrow\; accept/reject \tag{3}$$

where $q(\hat{x}_p^{(k)})$ is the feature vector of measures of fit, defined here as the residual norm on each set of *model* assumptions as follows:

$$q(\hat{x}_p^{(k)}) = \begin{bmatrix} \|\mathbf{B}\hat{x}_p^{(k)} - y_p\|^2 \\ \|\mathbf{L}\hat{x}_p^{(k)}\|^2 \\ \|\mathbf{F}\hat{x}_p^{(k)} - f_g^{(k)}\|^2 \end{bmatrix}. \tag{4}$$

The dot in Eq. 3 represents an inner-product, and following [14], w, here a simple projection vector, is defined as a linear discriminant, although it can be generalized to other classification schemes, since $\hat{x}_p^{(k)}$ defines a domain where other features can also be extracted.

From Eq. 4, note that the first component measures the fit between the observed low-resolution probe image and the low-resolution version of the resulting super-resolved image. The second component measures the smoothness of the super-resolution result, and the third component measures the difference between the features derived from high-resolution gallery images from class k and those obtained from $\hat{x}_p^{(k)}$.

For identification, we use Eq. 2 to compute the template $\hat{x}_p^{(k)}$ for each of the K classes in the gallery set. Then, we predict the class label for y_p by computing

$$\arg\min_k \; w \cdot q(\hat{x}_p^{(k)}) \quad k = 1, \ldots, K. \tag{5}$$

As in [14], since lower-resolution features provide some discriminative information that may be useful, in this work we also include the term $\|\mathbf{F}_L\mathbf{B}x - f_L^{(k)}\|^2$ in Eq. 2 using an additional regularization parameter. $\mathbf{F}_L$ is the feature extraction matrix computed with the low-resolution training set, and $f_L^{(k)}$ the corresponding features for class k. In general, other intermediate resolutions may be included. Here, Eq. 4 is modified accordingly to be a 4D feature vector.

B. Multiple Frames

An advantage of our algorithm is that it can be used to recognize either a single still image or a sequence of images. When multiple frames are available, as usually in video, a template $\hat{x}_p^{(k)}$ is obtained for the whole sequence of faces $y_{p_1}, y_{p_2}, \ldots$. Here, all frames in a sequence are assumed to contain the same face, i.e., the same class. Adapting $\mathbf{S^2R^2}$ from still image matching to multiple frames, needs only to redefine Eq. 2, and the new feature space, Eq. 4.

Using a similar approach as in [9], multiple-frame $\mathbf{S^2R^2}$ matching can be performed by rewriting Eq. 2 as

$$\sum_{i=1} \|\mathbf{B}_i x - y_{p_i}\|^2 + \alpha^2\|\mathbf{L}x\|^2 + \beta^2\|\mathbf{F}x - f_g^{(k)}\|^2, \quad (6)$$

where $\mathbf{B}_i$ represents the image formation process that accounts for motion and blurring corresponding to the ith frame with respect to gallery image x_g. Note that all the frames have the same regularization weight.

Once $\hat{x}_p^{(k)}$ is obtained, new features can be computed using a generalization of Eq. 4. For example, in this work, we have replaced the first feature element with

$$\sum_{i=1} \|\mathbf{B}_i \hat{x}_p^{(k)} - y_{p_i}\|^2. \quad (7)$$

Verification and identification follow as for still images, from Eq. 3 and Eq. 5, respectively.

C. Multiple Cameras

The generalization from a single camera to multiple cameras is carried out similarly as for multiple frames. In multiple cameras we model different point spread functions, and therefore, different image formation models. But, instead of giving the same weight to all the constraints from different cameras, we assign a different regularization parameter for each camera model.

For the case of two cameras, we have two still image inputs, $y_p^{(1)}$ and $y_p^{(2)}$. The effect of the cameras is accounted for with their corresponding image formation models, $\mathbf{B}^{(1)}$ and $\mathbf{B}^{(2)}$, and Eq. 2 becomes

$$\|\mathbf{B}^{(1)}x - y_p^{(1)}\|^2 + \xi\|\mathbf{B}^{(2)}x - y_p^{(2)}\|^2 + \alpha^2\|\mathbf{L}x\|^2 + \beta^2\|\mathbf{F}x - f_g^{(k)}\|^2 \quad (8)$$

where ξ is a regularization parameter.

The feature space in this case is defined similarly as in Eq. 4, but adding one feature element per camera. The generalization to use more camera inputs or multiple frames from multiple cameras is straightforward and we do not treat it here.

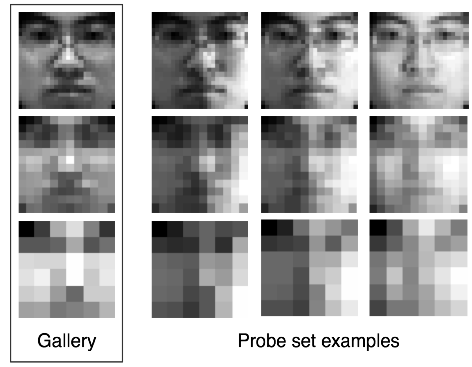

Fig. 3. Example images from the Multi-PIE database [11]. The left column shows an image at different resolutions from the gallery subset. Only one image per subject is in the gallery in all experiments. The rest of the columns show examples of probe images for the same person at different illuminations. The rows represent different resolutions, from top to bottom: 24×24 pixels, 12×12 pixels, and 6×6 pixels.

III. EXPERIMENTS AND RESULTS

To evaluate the performance of the proposed algorithm we use the Multi-PIE database [11]. The Multi-PIE database is a recent extension of the PIE database [21]. It has a total of 337 subjects (compared to 68 of PIE) that attended from one to four different recording sessions, each separated by at least a month (unlike PIE, where all images of each subject are captured on the same day in a single session). As in PIE, different face poses, expressions and illumination variations due to flashes from different angles were recorded.

Here we use all the subjects available, and we present results using frontal images of neutral expressions with different illuminations. A generic set of 73 subjects is sequestered to compute 25 Fisherfaces [5] for $\mathbf{F}$, and 40 sequestered subjects are used for learning regularization parameters and w. For evaluation, gallery and probe sets, we use 224 classes, which is the rest of the classes in the database. In our Multi-PIE experiment we have used only one image with no flash illumination as gallery, while the probe set contains all the images with flash illumination from all horizontal angles (13 images per subject). Figure 3 is an example for one subject, showing the gallery image and selected probe images at different resolutions. In total, the probe set has 2912 images, which gives 2912 true-class comparisons and 649,376 false-class comparisons.

A. Multiple Frames

We compare in this section the difference in performance of recognition of low-resolution faces using multiple frames to produce one score for the whole sequence. To evaluate the potential of this choice versus using only one still image we assume the classifier can be trained to account for the different possible shifts, and that it has perfect knowledge (possibly with help of a face detector) of the shifts in the frames. The classifier has been trained with images of 24×24 pixels and this is the desired resolution we ideally expect in the probe faces. We show results for two sets of experiments: when the probe faces are of size 12×12, and when they are 6×6 (labeled LR in the figures). That is, the magnification

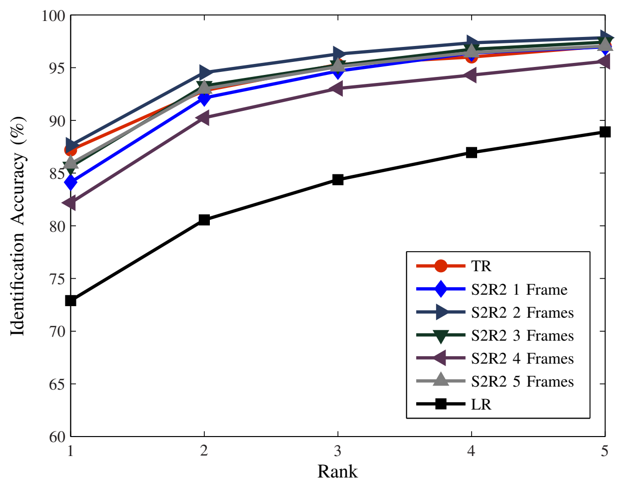

Fig. 4. Identification accuracy (%) when probe images are of size 12×12 pixels and $M = 2$, using the Multi-PIE database. CMC curve showing Rank-k IDA (%) for different number of frames.

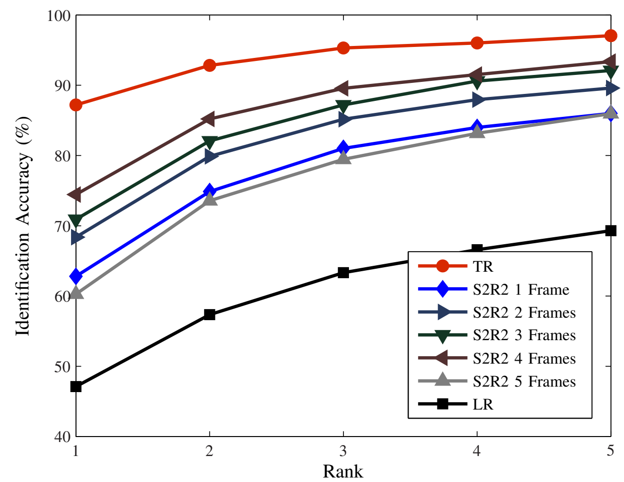

Fig. 5. Identification accuracy (%) when probe images are of size 6×6 pixels and $M = 4$, using the Multi-PIE database. CMC curve showing Rank-k IDA (%) for different number of frames.

ratio (M) is of 2 and 4, respectively. In each experiment we test for the use of 1 frame (still image) up to 5 frames (labeled under S2R2 in the figures). The implementation of the algorithm used in this section has been discussed in Section II-B.

Figure 4 shows the identification accuracy of these experiments, as a CMC curve, for the case of $M = 2$. For reference, we have included the result of matching using still probe images (single frame) for the desired training resolution (labeled TR), which in this problem settings is not available. Figure 5, similarly shows results for the case of $M = 4$.

We can observe several behaviors. First, the proposed algorithm clearly improves face recognition over matching in the low-resolution domain with a single still image. Second, for relatively small magnification factors, here for the case of $M = 2$, it is possible to outperform the hypothetical scenario of recognition at the desired training resolution. This is because in the proposed implementation we also include probe-resolution features, and the algorithm exploits this combination of high- and low-resolution domains. Finally, in both experiments, $M = 2$ and $M = 4$, it is noticeable that sequentially increasing the number of frames does not result in a monotonic increase in performance. This is probably due to the fact that in this experiment all frames are given the same importance. It is possible that favoring an instance of $\mathbf{B}_i$ over another, for example, giving more importance to smaller shifts than longer ones (using different regularization parameters) may have a positive effect. This approach is used in the following set of experiments.

B. Multiple Cameras

We have implemented the case of having two cameras with the same point of view. The typical application scenario here is a camera sensor network where every node has two cameras, one produces images at the desired resolution while the other produces low-resolution images. For example, this could be a PTZ-webcam pair in a master-slave setting, where the PTZ becomes active out of sleep mode when the webcam detects an event. An overview with examples of these scenarios is given in [8]. Our implementation is described in Section II-C.

Figure 6 shows the identification accuracy in a CMC curve, for the case of Camera 1 producing images of 24×24 pixels, and Camera 2, images of 12×12 pixels. We show as well results for the case when only Camera 1 is used for recognition, and when only Camera 2 is available.

Similarly, Figure 7 shows the identification accuracy for the case of Camera 1 producing images of 24×24 pixels, but Camera 2 producing images of size 6×6 pixels. We also include here results for the case when only Camera 1 is used for recognition, and when only Camera 2 is available.

Our results are reassuring, showing that there is an advantage in using both cameras. Moreover, comparing Fig. 6 and Fig. 7, we see that the proposed algorithm does not degrades as much in performance from recognition with ratio $M = 2$ to $M = 4$, even when recognition by Camera 2 alone does decreases considerably.

IV. CONCLUSIONS

We have proposed an approach for recognition of low-resolution faces that uses super-resolution models together with face features by including them in a regularization framework. By finding a template that fits simultaneously the super-resolution constraints and feature constraints of base super-resolution and classification algorithms, we can extract measures of fit and use them as new features for recognition. Our results show that simple linear discriminants using these features produce better recognition performance than standard approaches.

In this work, we have expanded our formulation and presented results on two generalizations of the proposed algorithm. First, we show it can achieve better recognition performance using multiple frames instead of a single image.

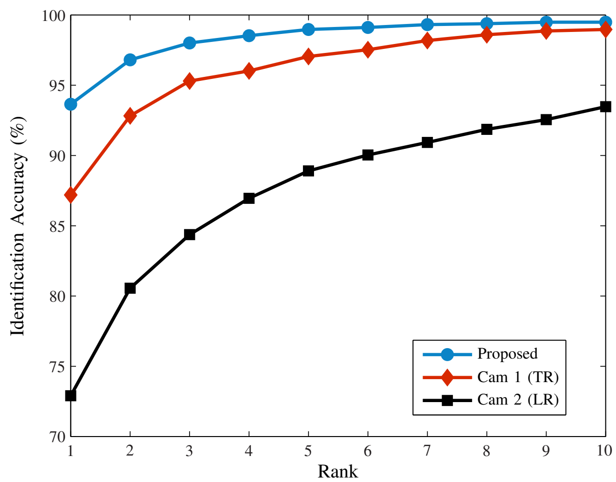

Fig. 6. Identification accuracy (%) for experiments using two cameras. CMC curve for the case of probe images being 24×24 pixels from Camera 1, and 12×12 pixels from Camera 2.

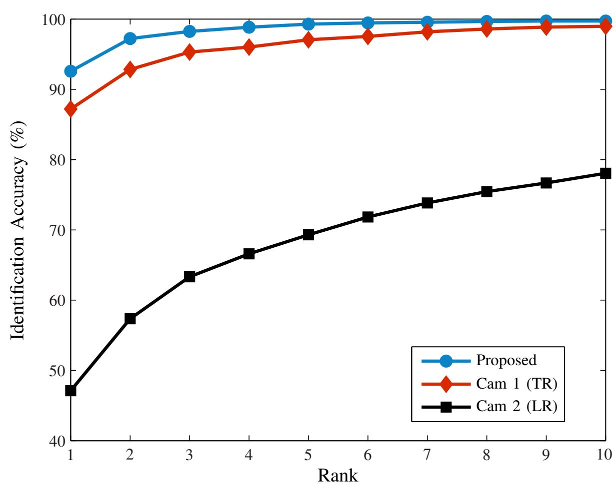

Fig. 7. Identification accuracy (%) for experiments using two cameras. CMC curve for the case of probe images being 24×24 pixels from Camera 1, and 6×6 pixels from Camera 2.

Under small magnification factors, multiple frames may be used to outperform matching under desired, but otherwise unavailable, training resolution scenarios. Second, we show that when two cameras are available, with possibly different PSFs as in camera sensor networks, our algorithm can be trained to achieve better recognition accuracy than using only the higher-resolution camera. We show, as well, that even when recognition by the low-resolution camera alone decreases considerably as resolution decreases, the proposed algorithm is robust to such variations.

Future work will include an evaluation of $\mathbf{S^2R^2}$ on recognition of faces in unconstrained video in addition to studying sensitivity to inaccuracies in the image formation model.

ACKNOWLEDGMENTS

The authors acknowledge the support of this research in part by Carnegie Mellon's ICES through a PITA grant.

REFERENCES

[1] O. Arandjelović and R. Cipolla. Face recognition from video using the generic shape-illumination manifold. In *Proc. European Conference on Computer Vision*, pages 27–40, 2006.
[2] O. Arandjelović and R. Cipolla. A manifold approach to face recognition from low quality video across illumination and pose using implicit super-resolution. *ICCV*, October 2007.
[3] S. Baker and T. Kanade. Hallucinating faces. In *Fourth Intl. Conf. on Aut. Face and Gesture Rec.*, 2000.
[4] S. Baker and T. Kanade. Limits on super-resolution and how to break them. *PAMI*, 24(9):1167 – 1183, 2002.
[5] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *PAMI*, 19(7):711–720, 1997.
[6] B. Boom, G. Beumer, L. Spreeuwers, and R. Veldhuis. The effect of image resolution on the performance of a face recognition system. In *Proc. of 9th International Conference on Control, Automation, Robotics and Vision*, pages 1–6, Dec. 2006.
[7] A. Chakrabarti, A. Rajagopalan, and R. Chellappa. Super-resolution of face images using kernel pca-based prior. *IEEE Tran. on Multimedia*, 9(4):888 – 892, June 2007.
[8] R. Cucchiara. Multimedia surveillance systems. In *Proceedings of the third ACM international workshop on Video surveillance & sensor networks*, pages 3–10, New York, NY, USA, 2005. ACM.
[9] M. Elad and A. Feuer. Super-resolution reconstruction of image sequences. *PAMI*, 21(9):817–834, 1999.
[10] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar. Advances and challenges in super-resolution. *International Journal of Imaging Systems and Technology*, 2004.
[11] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker. Guide to the CMU Multi-PIE database. Technical report, Carnegie Mellon University, 2007.
[12] B. Gunturk, A. Batur, Y. Altunbasak, I. Hayes, M.H., and R. Mersereau. Eigenface-domain super-resolution for face recognition. *Image Processing, IEEE Transactions on*, 12(5):597–606, May 2003.
[13] P. C. Hansen. *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1998.
[14] P. H. Hennings-Yeomans, S. Baker, and B. Vijaya Kumar. Simultaneous super-resolution and feature extraction for recognition of low-resolution faces. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, June 2008.
[15] K. Jia and S. Gong. Multi-modal tensor face for simultaneous super-resolution and recognition. *ICCV*, pages 1683–1690, October 2005.
[16] S.-W. Lee, J. Park, and S.-W. Lee. Face reconstruction with low resolution facial images by feature vector projection in kernel space. *Pattern Recognition, 2006. 18th International Conference on*, 3:1179–1182, 2006.
[17] J.-S. Park and S.-W. Lee. Stepwise reconstruction of high-resolution facial image based on interpolated morphable face model. In *AVBPA*, pages 102–111, 2005.
[18] S. C. Park, M. K. Park, and M. G. Kang. Super-resolution image reconstruction: a technical overview. *IEEE Signal Processing Magazine*, 20(3):21–36, May 2003.
[19] P. J. Phillips. Multiple biometric grand challenge kick-off workshop. Presentation, 18th April 2008.
[20] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek. Overview of the face recognition grand challenge. In *CVPR*, pages 947–954, 2005.
[21] T. Sim, S. Baker, and M. Bsat. The CMU pose, illumination, and expression (PIE) database of human faces. Technical report, RI, Carnegie-Mellon, 2001.
[22] X. Wang and X. Tang. Hallucinating face by eigentransformation. *Systems, Man and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 35(3):425–434, 2005.
[23] F. Wheeler, X. Liu, and P. Tu. Multi-frame super-resolution for face recognition. *BTAS*, September 2007.
[24] S. Zhou and R. Chellappa. *Face Processing: Advanced Modeling and Methods*, chapter Beyond a single still image: Face recognition from multiple still images and videos. Academic Press, 2005.
[25] A. Zomet and S. Peleg. *Super-Resolution Imaging*, chapter "Super-resolution from multiple images having arbitrary mutual motion", pages 195–209. Kluwer, 2001.