Blur Adaptive Sparse Representation of Random Patches for Face Recognition on Blurred Images

Daniel Heinsohn
Pontificia Universidad Católica de Chile
Vicuña Mackenna 4860, Santiago de Chile
diheinso@puc.cl

Domingo Mery
Pontificia Universidad Católica de Chile
Vicuña Mackenna 4860, Santiago de Chile
dmery@ing.puc.cl

1. Motivation

Face recognition in unconstraint environments is an active research area on forensics. The challenge is to deal with real world environments in order to achieve a high recognition rate under difficult conditions such as changes in ambient lightning, pose, expressions, occlusion, face size, blurriness, among others. In this research, we address one of the biggest problems: face recognition in blurred images.

There are many approaches in this field reported in the literature. On the one hand, there are inverse methods based on *deblurring*, where image restoration is performed. Among these algorithms we can find blind deconvolution [5], non-blind deconvolution [14], regularization methods on total variation [8], and Tikhonov regularization [10]. On the other hand, there are direct methods that are based on features of the image that are invariant to blurriness, such as features extracted in spatial and frequency domains (see for example [3]). In addition, algorithms based on Sparse Representation Classification (SRC) have been widely used [12]. In [11] for example, registration and illumination are simultaneously considered in the sparse representation. In [2], the dictionary is assembled by the class centroids and sample-to-centroid difference. In [9], a sparse discriminative analysis is proposed using the $l_{1,2}$-norm. In [13], a sparse representation in two phases is proposed. In [1], sparse representations of patches distributed in a grid manner are used. In [7], patches that do not give information ($e.g.$, occluded parts) are not considered in the recognition. Even though a wide variety of algorithms have been developed, face recognition in blurred images remains an open question (see for example [4]), mainly because of the difficulty to model the unknown blurriness of the images. In this work, we would like to make a contribution in this field by modeling the blurriness using a blur adaptive sparse representation.

2. Proposed method

We propose an algorithm based on ASR+, *i.e., Adaptive Sparse Representation of Random Patches* [7]. Original ASR+ consists of two stages: learning and testing. In the learning stage, for each subject of training dataset, several random patches are extracted from its face images in order to construct representative dictionaries. In the testing stage, random test patches of the query image are extracted, and for each test patch a dictionary is built concatenating the 'best' representative dictionary of each subject. Using this adapted dictionary, each test patch is classified following the Sparse Representation Classification (SRC) methodology [12]. Finally, the query image is classified by patch voting. Our proposed approach, that we call 'blur ASR+' or bASR+, consists of adding new dictionaries in the training stage that contains patches of the faces with different level of blurriness, so when a query blurred image is tested, the algorithm should recognize the face by selecting the dictionary of same level of blurriness.

In the training stage of bASR+, we create B sets of images with different levels of blurriness b, for $b = 0 \dots B-1$. Blurred training images are achieved by filtering the original sharp face image with a Gaussian mask of $\sigma = b/4$ for each blur level b. In this approach, $b = 0$ means the original image. In our experiments, we used $B = 15$. For each image of the training images, a sharpness value s is computed. This value is estimated using the single value decomposition of the gradient image (see details in [15]). Thus, for each blur level b we calculate a representative sharpness value s_b as the median of all sharpness values s of the training images that have a blur level b.

Following ASR+ methodology [7], for each image of the gallery (including the B sets), m patches are randomly extracted with a center in location (x, y). With this information a descriptor is built for each patch $\mathcal{P}$:

$$\mathbf{y} = f(\mathcal{P}) = [\,\mathbf{z}\,;\ \alpha x\ ;\ \alpha y\,] \in \mathbb{R}^{d+2}, \quad (1)$$

where $\mathbf{z}$ is a vector that contains the grayvalues of the patch

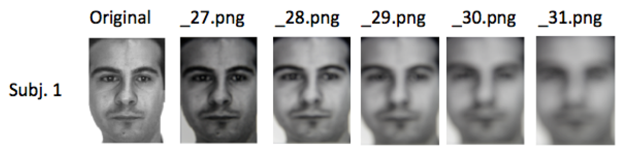

Figure 1: The image at the left is the first image of subject 1. Then, from left to right, are the ascending level of blurriness of the original photo. The top of each photo represents the index of the photo in the new dataset.

(by stacking its columns) and α is a weighting factor between the description and location. We build $\mathbf{Y}^{ib}$ that contains descriptors $\mathbf{y}$ of all patches of subject i at blur level b. Afterwards, $\mathbf{Y}^{ib}$, is clustered into Q parent clusters and then again into R child clusters in order to select a reduced number of samples. We obtain $\mathbf{c}^{ib}_{qr} \in \mathbb{R}^{d+2}$, *i.e.*, the centroid of child cluster r of parent cluster q of subject i and blur level b, for $r = 1 \ldots R$ and $q = 1 \ldots Q$. All centroids of child clusters of subject i and blur level b are stored in dictionary $\mathbf{D}^{ib}$ (see details in [7]).

Testing stage consists of identifying a subject from a query image $\mathbf{I}^t$. We obtain s^t, the sharpness value of $\mathbf{I}^t$ following the same methodology explained above based on [15]. We look for the most similar sharpness value in the training images as:

$$b^t = \underset{b}{\text{argmin}} ||s^t - s_b|| \tag{2}$$

Hence, dictionaries with this blurriness are selected, $\mathbf{D}^{ib^t}$, for $i = 1 \ldots n$. Afterwards, the algorithm ASR+ to recognize the query face is used. Thus, S patches are extracted from the query image and the Euclidean distance is measured between this patches and the parent clusters from each subject from $\mathbf{D}^{ib^t}$. We select those subjects that have a distance less than a threshold θ and for each patch an adaptive dictionary is built from the best patches of those subjects that were selected. SRC algorithm is executed using these adaptive dictionaries and the identity of the subject of the query image is chosen by majority vote from the classes selected of each patch.

3. Experiments

For these experiments we generated an extended version of AR dataset [6] by adding five new photos per subject with different levels of blurriness (see Figure 1). A blurred image was obtained in three steps: *i)* the first original (sharp) face image of an AR subject was displayed on a computer monitor, *ii)* a digital single-lens reflex camera (SLR) was set out-of-focus, *iii)* a picture of the monitor was captured by the camera. Thus, we have real out-of-focus face images of all subjects of AR database. Nevertheless, it is an out-of-focus image of a 2D object (a blurred picture of a face

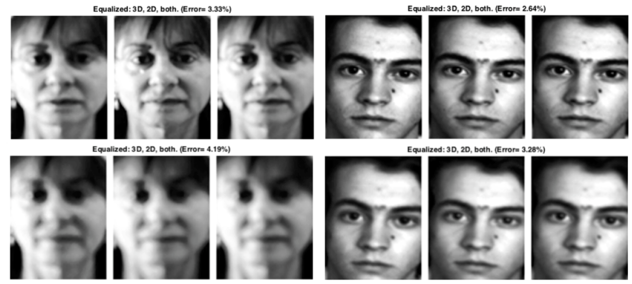

Figure 2: Each set of three images contain in the left the 3D picture (I_{3D}), in the center the 2D registered picture (I^r_{2D}) and the right picture correspond to the average of both images. The percentage of error of this technique is a 3.33% for the top-left set, 4.19% for the bottom-left set, 2.64% for the top-right set and 3.28% for the bottom-right set.

image displayed on a monitor) instead of a 3D object (a blurred picture of a real face). We call these two kind of images I_{2D} and I_{3D} respectively. In real cases all images are I_{3D}, however, in our experiments the images were I_{2D}. In order to evaluate the similarity of these two kinds of images, we conducted the following experiment: from two different subjects (relatives of the authors) we took five I_{2D} and I_{3D} images (see Fig. 2) and we measured the residual RMS after registration. The registered image computed from I_{2D} is called I^r_{2D}. In this experiment, the error between I_{3D} and I^r_{2D} was $2.64\% \sim 4.19\%$ only.

Two types of experiments were performed to evaluate the recognition rate (accuracy): 1) our approach (bASR+) was compared with other state of the art algorithms and 2) bASR+ was compared with face recognition commercial softwares. In our experiments, we randomly selected k subjects. From each selected subject, n images were randomly chosen for training and one for testing. The original image (that was used to generate the blurred images of the subject) was not considered in the training dataset. In order to obtain a better confidence level in the estimation of face recognition accuracy, the test was repeated 100 times by randomly selecting new k subjects and $n + 1$ images each time. The reported accuracy in all of our experiments is the average calculated over the 100 tests.

First, the experiments were performed with three different number of subjects (k) and training images (n) as shown in Tables 1, 2, and 3.

Second, we tested commercial softwares Picasa ©, iPhoto © for OSx 10.9.5 (Mavericks) and Photo © for OSx 10.10.5 (Yosemite). 100 subjects were tested with 13 images per subject for the training stage and 1 image for testing[1].

As shown in our results, bASR+ outperforms each

[1] In more than 50% of the images no face was detected by the software.

method when a blurred image is tested. It is clear that our algorithm can deal with overly blurred achieving an average of 85.6% of recognition.

4. Conclusions

We showed that our approach bASR+ is able to work in uncontrolled conditions, specially in images with high levels of out-of-focus blurriness. The robustness of our algorithm is due to the dictionaries learned for each subject of the gallery in the learning stage corresponded to a rich collection of blurred representations of relevant parts.

As no blur robust algorithms were found on the web, we used commercial softwares for our second experiment. These softwares achieved low recognition rates when testing with blurred images, even though they are expected to work on uncontrolled conditions. In terms of future work, we would like to extend this investigation to other types of blurriness, such as ambient interference and motion.

Acknowledgments

This work was supported by Fondecyt Grant No. 1130934 from CONICYT, Chile.

References

[1] Y. Chen, T. T. Do, and T. D. Tran. Robust face recognition using locally adaptive sparse representation. In *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pages 1657–1660. IEEE, 2010.

[2] W. Deng, J. Hu, and J. Guo. In defense of sparsity based face recognition. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 399–406. IEEE, 2013.

[3] R. Gopalan, S. Taheri, P. Turaga, and R. Chellappa. A blur-robust descriptor with applications to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(6):1220–1226, 2012.

[4] J. C. Klontz and A. Jain. A Case Study of Automated Face Recognition: The Boston Marathon Bombings Suspects. *IEEE Computer*, 2013.

[5] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Computer Science and Artificial Intelligence Laboratory Technical Report Ef &# 64257 ; cient Marginal Likelihood Optimization in Blind Deconvolution. *Optimization*, 2011.

Table 1: Comparison with different face recognition algorithms for $k = 20, n = 4$.

Method	Original	_27.png	_28.png	_29.png	_30.png	_31.png
bASR+ (ours)	100%	100%	99.8%	99.6%	98.5%	92.3%
ASR+	100%	100%	100%	98.3%	82.4%	42.4%
SRC	100%	100%	99.9%	98.5%	81.8%	45.3%
LBP	100%	100%	99.9%	98.8%	82.4%	42.3%

Table 2: Comparison with different face recognition algorithms for $k = 40, n = 9$.

Method	Original	_27.png	_28.png	_29.png	_30.png	_31.png
bASR+ (ours)	99.6%	100%	100%	100%	99.9%	93.2%
ASR+	100%	100%	100%	99.8%	86.4%	40.2%
SRC	100%	100%	100%	99.8%	86%	39.4%
LBP	100%	100%	100%	99.8%	85.4%	41.4%

Table 3: Comparison with different face recognition algorithms with $k = 100, n = 13$.

Method	Original	_27.png	_28.png	_29.png	_30.png	_31.png
bASR+ (ours)	99.8%	100%	100%	100%	99.3%	85.6%
ASR+	100%	100%	100%	99.2%	71.8%	26.3%
SRC	100%	100%	100%	99.2%	71.2%	26.3%
LBP	100%	100%	100%	99.1%	71.1%	26.6%

Table 4: Comparison with different comercial softwares for $k = 100, n = 13$.

Method	Original	_27.png	_28.png	_29.png	_30.png	_31.png
bASR+ (ours)	99.8%	100%	100%	100%	99.3%	85.6%
Picasa ©	86%	71%	76%	5%[1]	0%[1]	0%[1]
iPhoto ©	58%	21%	3%[1]	0%[1]	0%[1]	0%[1]
Photo ©	94%	87%	34%	8%[1]	0%[1]	0%[1]

[6] A. Martinez and R. Benavente. The AR face database. *CVC Technical Report 24*, June 1998.

[7] D. Mery and K. Bowyer. Face Recognition via Adaptive Sparse Representations of Random Patches. *IEEE Workshop on Information Forensics and Security (WIFS 2014), Atlanta, Dec. 3-5.*, 2014.

[8] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60:259–268, 1992.

[9] X. Shi, Y. Yang, Z. Guo, and Z. Lai. Face recognition by sparse discriminant analysis via joint l 2, 1-norm minimization. *Pattern Recognition*, 47(7):2447–2453, 2014.

[10] A. Tikhonov and V. Arsenin. *Solutions of ill-posed problems*. Vh Winston, 1977.

[11] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, H. Mobahi, and Y. Ma. Toward a practical face recognition system: Robust alignment and illumination by sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34:372–386, 2012.

[12] J. Wright and a.Y. Yang. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(March):210–227, 2009.

[13] Y. Xu, D. Zhang, J. Yang, and J.-Y. Yang. A two-phase test sample sparse representation method for use with face recognition. *Circuits and Systems for Video Technology, IEEE Transactions on*, 21(9):1255–1262, 2011.

[14] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum. Progressive inter-scale and intra-scale non-blind image deconvolution. In *ACM Transactions on Graphics (TOG)*, volume 27, page 74. ACM, 2008.

[15] X. Zhu and P. Milanfar. A no-reference sharpness metric sensitive to blur and noise. *2009 International Workshop on Quality of Multimedia Experience, QoMEx 2009*, 2(3):64–69, 2009.