

Metric Learning in the Dissimilarity Space to Improve Low-Resolution Face Recognition

Mairelys Hernández-Durán^(✉), Yenisel Plasencia-Calaña,
and Heydi Méndez-Vázquez

Advanced Technologies Application Center, 7ma A, #21406 Playa, Havana, Cuba
{mhduran,yplasencia,hmendez}@cenatav.co.cu

Abstract. Standard face recognition methods based on a feature representations are not suitable for low-resolution environments. Therefore low-resolution face recognition is still an unsolved problem where the best approaches still obtain very low recognition rates. In this paper, we propose a low-resolution face recognition method using the dissimilarity representation. In addition, we propose the use of metric learning methods to replace the standard Euclidean distance in the dissimilarity space. The effectiveness of our proposal is tested on two different data sets, one of them is the SCface database which is very challenging since the images were collected from surveillance cameras.

Keywords: Low-resolution · Metric learning · Dissimilarity space · Face recognition

1 Introduction

In real world applications such as video-surveillance, captured faces are often of low-resolution (LR). At these environments the obtained LR face image loses important details which are discriminative between persons mainly due to the distance among subjects and camera. These images can also present facial variations such as pose and expression; thus, it represents a challenge for a recognition task. Low-resolution face recognition (LRFR) methods try to cope with a classification problem between LR test images and high-resolution (HR) gallery images, causing the dimensional mismatch problem. Therefore, the dimensional mismatch between gallery/probe pairs and the lack of facial features are some of the main challenges related to LRFR [1]. Some authors have tried to find resolution-robust feature representation [2], but this is a difficult task because most of the effective features used in HR face recognition such as texture and color, may fail with LR images. The performance of traditional methods in the LR case suggests that current feature representation approaches are not suitable to cope with LRFR [3]. To improve the results, it becomes a priority to explore alternatives to the feature-based representation.

A representation based on dissimilarities between objects [4] is advantageous in situations where it is easier to define dissimilarities rather than features. The

dissimilarity space (DS) representation has been successfully used in many difficult task such as person re-identification [5]. Based on the success of previous works [4], we proposed the use of dissimilarity representations, as an alternative for LRFR. We believe that more discriminative information for classification can be obtained if the LR images are analyzed in the context of dissimilarities with other images. However, previous works assumed standard or designed dissimilarities, and not dissimilarities automatically learned for a given problem. Some researchers have shown that the classification could be greatly improved by learning a suitable distance metric. Then, we consider improving the original DS representation by using metric learning methods on top of it. Metric learning can provide a way to adapt a distance function to the given task.

In this work, the standard Euclidean distance in the DS is replaced by a learned metric, i.e., a Mahalanobis metric. We compared our proposal with some state-of-the-art representative methods based on feature vector representations. To address the dimensional mismatch, we used the best performing strategy proposed in [6], where the HR images are down-scaled and then up-scaled and the LR images are up-scaled to the same resolution. The proposal was evaluated on different face database including the SCFace [7], which is a very difficult database because it emphasizes the challenges of face recognition in surveillance environments.

The paper is organized as follows. Section 2 presents the related work on LRFR, the DS representation and the metric learning approach. Section 3 presents our proposal based on DS with automatically learned metrics to cope with the classification problem of LR facial images. Experiments and discussion are presented in Sect. 4, and concluding remarks are provided in Sect. 5.

2 Related Work

With the growing demands on surveillance applications, extensive efforts have been made on LRFR research. However, it remains an open issue due to the challenges posed by LR. Furthermore, the different resolutions between gallery and probe images lead to the so-called dimensional mismatch. To cope with this problem, different approaches have been used such as unified feature space (CLPM) [8]. This approach is used to project HR and LR images into a common space, which seems feasible to cope with the dimensional mismatch. However, it is not straightforward to find an optimal inter-resolution space and the transformations process may introduce noise. Several methods have used super-resolution (SR) techniques. However, these kind of methods mostly focus on obtaining a good visual reconstruction rather than a higher recognition rate. Current approaches mainly include feature vector representation for addressing LRFR. Resolution-robust feature representation has been considered for the LR case. Multidimensional scaling (MDS) [9] is a representative method, in which the relationships between LR and HR are explored taking into account the dimensional mismatch problem. Many authors have been working on this idea trying to find a common or inter-resolution space to project LR and their corresponding HR images on it [10].

A dissimilarity representation between objects is an alternative solution. Based on the idea proposed in [4], the dissimilarities are considered as the connection between perception and higher-level knowledge, which are key elements in the process of human recognition and categorization. By using the differences with prototypes for creating the representations we may be able to emphasize relevant information for discrimination among the classes, which, otherwise, by only analyzing the image, may be difficult to express in a feature representation. Following up on [6], they proposed the DS for LRFR but they only used the standard Euclidean distance. We believe that the use of a suitable distance metric can improve the classification accuracy. For example, in [11] they showed that it is possible to improve the K-NN classification accuracy using suitable distance metrics. The goal of metric learning algorithms is to take advantage of prior information in form of labels over standard similarity measures.

Compared with previous approaches this work is different in some aspects. We proposed the use of a dissimilarity based representation using learned metrics to achieve more discriminative distances in LRFR. In particular, our proposal is an alternative representation to feature space (FS) based on dissimilarities between objects and also introducing metric learning to replace the standard Euclidean distance in the DS.

3 Proposed Approach: Dissimilarity Space and Metric Learning for LRFR

A general scheme of the proposed strategy can be found in Fig. 1. In the following we will describe in more details the dissimilarity space construction and the metric learning approaches.



Fig. 1. General scheme of our proposal

3.1 Dissimilarity Space

Duin and Pekalska [4] proposed the DS as an Euclidean vector space in which it is possible to use several statistical classifiers. Although it has been used to solve a number of problems [12, 13] their advantages to solve the dimensional mismatch in LR case, has not been explored yet. The proximity information is intuitively more discriminative than the features or the composition of each object independently. Based on its advantages, we consider the use of the dissimilarity space to achieve a more discriminative relational representation of the LR images. Let X be the space of objects, let $R = \{r_1, r_2, \dots, r_k\}$ be the set of prototypes such that $R \in X$, and let $d : X \times X \rightarrow \mathbb{R}^+$ be a suitable dissimilarity

measure for the problem. For a training set $T = \{x_1, x_2, \dots, x_l\}$ such that $T \in X$, a mapping $\phi_R^d : X \rightarrow \mathbb{R}^k$ defines the embedding of training and test objects in the DS by the dissimilarities with the prototypes:

$$\phi_R^d(x_i) = [d(x_i, r_1) \ d(x_i, r_2) \ \dots \ d(x_i, r_k)]. \quad (1)$$

3.2 Metric Learning Approach

In this section, we introduce the general idea of metric learning for kNN classification and review some previously studied approaches: LMNN, which directly attempts to optimize k-NN classification error; another method based on the Linear Discriminant Analysis (LDA) [14]; and the KISS metric learning method.

Large Margin Nearest Neighbor (LMNN): A mapping $D : X \times X \rightarrow \mathbb{R}_0^+$ over a vector space X is defined as a metric if for all the vectors $\vec{x}_i, \vec{x}_j, \vec{x}_k \in X$ it satisfies some properties such as symmetry and triangular inequality [15]. It is possible to obtain a family of metrics on X by computing Euclidean distances after performing a linear transformation $x = Lx$. These metrics compute quadratic distances that can be expressed in terms of the square matrix $M = L'L$. Thus, any matrix M formed in this way from a real-valued matrix L is guaranteed to be positive semidefinite, refers to the Mahalanobis metric. In LMNN, the distances are viewed as generalizations of Euclidean distances, i.e., Euclidean distances are recovered by setting M to be equal to the identity matrix. The idea is based on the observation that the kNN classification could have a good performance for a sample of the data if its k-nearest neighbors share the same label. By increasing the number of training samples with this property they learned a linear transformation of the input space that precedes kNN classification using Euclidean distances. Their approach has the advantage of improving the original Euclidean distance from a classification perspective and in some cases to provide a lower-dimensional embedding of the data.

Linear Discriminant Analysis: Different ways have been proposed to estimate Mahalanobis distance metrics to compute distances in k-NN classification. One of such methods is Eigen decomposition. This approach has been used to discover informative linear transformations of the input space, which can be seen as inducing a Mahalanobis distance metric in the original space. LDA is a representative Eigenvector method. It operates in a supervised setting and uses the class labels of the inputs to derive informative linear projections. In the context of metric learning, LDA computes a linear projection L that maximizes the amount of between-class variance relative to the amount of within-class variance. The linear transformation L is chosen to maximize the ratio of between-class to within-class variance, subject to the constraint that L defines a projection matrix. The traditional LDA algorithm is still attractive compared to several recently developed metric learning algorithms [16].

Keep It Simple and Straightforward Metric Learning (KISSME):

Another strategy is to learn an optimal distance measure for genuine and impostor pairs. Koestinger et al. [17] proposed an effective method to learn the distance metric based on a likelihood-ratio test. The equivalence constraints are considered natural inputs to metric learning methods because similarity functions mainly establish a relation between pairs of points. KISSME [17] computes the covariance matrix of similar and dissimilar pairs, and uses the difference of the inverse covariance matrix as a projection matrix. It does not rely on complex iterative optimization, which is an advantage for practical applications. It applies the log likelihood ratio test of two Gaussian distributions for metric learning, and so a simplified closed-form solution can be derived.

4 Experimental Evaluation

We present the results of the proposed scheme for low-resolution face recognition. Two different database were considered for the experiments: the SCFace database [7] and the Labeled Faces in the Wild (LFW) [18]. On all of our experiments, the test images were obtained by down-scaling the original images using a bicubic interpolation at different sizes. A bicubic interpolation was also applied in the up-scaling process to obtain high resolution images. The standard Euclidean distance in the DS was replaced by a learned metric, and the linear discriminant classifier (LDC), which assumes equal covariance matrices for the classes, was used. We computed local binary patterns (LBP) on local blocks of the geometrically normalized images. Histograms were computed on each block and concatenated. The dissimilarity measure was computed on top of a feature representation. Particularly, we created the dissimilarity space using chi-square distance between LBP histograms, since it is a more discriminative measure for histograms.

4.1 Experiments and Discussion on SCFace Database

The SCface database [7] was particularly designed for simulating video-surveillance scenarios, thus it is the most suitable to evaluate the low resolution problem. It consists of 4160 images from 130 people taken in uncontrolled environment. Three different distances, namely 4.20 m (distance1), 2.60 m (distance2), and 1.00 m (distance3); each one with five cameras (cam1, cam2, cam3, cam4, cam5) were used to capture the images. The illumination was uncontrolled and the captured images were different in terms of quality, type and resolution. Example images for distances 2 and 3 appear in Fig. 2.

In order to compare our method with existing approaches we follow the protocol in [19], where the images from distance 3 were normalized to 48×48 pixels as HR images, while the corresponding LR images of 16×16 pixels were obtained from distance 2. Besides, 80 subjects were selected for training and the remaining 50 subjects were used for testing. The experiment was repeated 5 times using 200 PCA components, which provided the best results. The results in terms



Fig. 2. Some examples of SCFace database

of Recognition Rates are reported in Table 1. The standard deviation is also presented. As it can be seen in Table 1, in general the proposed scheme achieves relatively high and stable recognition rates when compared with other state-of-the-art algorithms reported in [19]. In particular, the best result is obtained using the LDA metric learning, with a significantly higher recognition rate.

Table 1. Recognition rates in the SCFace database

Method	Recognition Rate(RR)
Proposal in [19]	43.24
CLPM [8]	29.12
SDA [20]	40.08
CMFA [21]	39.56
Our proposal with KISSME (average)	47.20 ± 0.0087
Our proposal with LMNN (average)	50.80 ± 0.0056
Our proposal with LDA (average)	56.40 ± 0.0027

4.2 Experiments and Discussion on LFW Database

In order to corroborate the obtained results on another dataset and to compare the proposed use of metric learning over the dissimilarity space, we conduct experiments on LFW database [18]. It contains 13233 labeled faces from 5749 people. A subset of the database consisting of 3 832 images belonging to 178 subjects was used during the experiments, by selecting the subjects with 8 or more images. The data is challenging, as the faces are detected in the wild, taken from Yahoo! News. The images have different variations such as pose, scale, clothing, expression, focus, resolution and others. Some example images are shown in Fig. 3

All images were geometrically normalized by the center of the eyes to the LR of 16×16 pixels and to the HR of 48×48 pixels. We randomly divided the data set into two sets for training and testing of equal size five times. In this experiment we compare the standard Euclidean distance to the learned metric in the DS. The obtained results in terms of error rates are shown in Table 2. From the results in Table 2 it can be seen that learning a Mahalanobis metric to replace the Euclidean distance improves the classification in a DS by a great margin.



Fig. 3. Some examples of LFW database

Table 2. Error rates in the LFW database

Classifier	DS	LDA
LDC	0.59	0.45

5 Conclusions

In this paper we presented the use of metric learning to learn a Mahalanobis distance metric for LRFR in the dissimilarity space. This learned metric enforces objects for the same class to be closer while objects from different classes are pulled apart. Unlike current methods for LR case, which mostly consider the features space, we proposed a new representation space based on dissimilarities between objects and we improved the classification in this space with metric learning. We evaluated our proposal on two challenging datasets. Experiments showed improvements over previously reported methods. Therefore, the improvement of representations based on relational information seems to be a promising research line for future works.

References

1. Wang, Z., Miao, Z., Wu, Q.J., Wan, Y., Tang, Z.: Low-resolution face recognition: a review. *Vis. Comput.* **30**(4), 359–386 (2014)
2. Ren, C.X., Dai, D.Q., Yan, H.: Coupled kernel embedding for low-resolution face image recognition. *IEEE Trans. Image Process.* **21**(8), 3770–3783 (2012)
3. Hennings-Yeomans, P.H., Baker, S., Kumar, B.V.: Simultaneous super-resolution and feature extraction for recognition of low-resolution faces. In: *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2008*, pp. 1–8. IEEE (2008)
4. Duin, R., Pekalska, E.: *The Dissimilarity Representations for Pattern Recognition: Foundations and Applications*. World Scientific, Singapore (2005)
5. Satta, R., Fumera, G., Roli, F.: Fast person re-identification based on dissimilarity representations. *Pattern Recogn. Lett.* **33**(14), 1838–1848 (2012)
6. Hernández-Durán, M., Cheplygina, V., Plasencia-Calaña, Y.: Dissimilarity representations for low-resolution face recognition. In: Feragen, A., Pelillo, M., Loog, M. (eds.) *SIMBAD 2015*. LNCS, vol. 9370, pp. 70–83. Springer, Heidelberg (2015). doi:[10.1007/978-3-319-24261-3_6](https://doi.org/10.1007/978-3-319-24261-3_6)
7. Grgic, M., Delac, K., Grgic, S.: SCface-surveillance cameras face database. *Multimed. Tools Appl.* **51**(3), 863–879 (2011)
8. Li, B., Chang, H., Shan, S., Chen, X.: Low-resolution face recognition via coupled locality preserving mappings. *IEEE Signal Process. Lett.* **17**(1), 20–23 (2010)

9. Biswas, S., Aggarwal, G., Flynn, P.J., Bowyer, K.W.: Pose-robust recognition of low-resolution face images. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(12), 3037–3049 (2013)
10. Xing, X., Wang, K.: Couple manifold discriminant analysis with bipartite graph embedding for low-resolution face recognition. *Signal Process.* **125**, 329–335 (2016)
11. Ding, Z., Suh, S., Han, J.J., Choi, C., Fu, Y.: Discriminative low-rank metric learning for face recognition. In: *Automatic Face and Gesture Recognition (FG)*, 11th IEEE International Conference and Workshops on. vol. 1, pp. 1–6. IEEE (2015)
12. Orozco-Alzate, M., Castellanos-Domínguez, C.: Nearest feature rules and dissimilarity representations for face recognition problems. *Face Recognition; International Journal of Advanced Robotic Systems*, pp. 337–356. Vienna, Austria (2007)
13. Li, Y., Duin, R.P., Loog, M.: Combining multi-scale dissimilarities for image classification. In: *Proceedings of the 2012 21th International Conference on Pattern Recognition, ICPR 2012*. IEEE Computer Society (2012)
14. Hastie, T., Tibshirani, R.: Discriminant adaptive nearest neighbor classification. *IEEE Trans. Pattern Anal. Mach. Intell.* **18**(6), 607–616 (1996)
15. Weinberger, K.Q., Saul, L.K.: Distance metric learning for large margin nearest neighbor classification. *J. Mach. Learn. Res.* **10**, 207–244 (2009)
16. Liao, S., Lei, Z., Yi, D., Li, S.Z.: A benchmark study of large-scale unconstrained face recognition. In: *Biometrics (IJCB)*, 2014 IEEE International Joint Conference on, pp. 1–8. IEEE (2014)
17. Koestinger, M., Hirzer, M., Wohlhart, P., Roth, P.M., Bischof, H.: Large scale metric learning from equivalence constraints. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2288–2295. IEEE (2012)
18. Huang, G.B., Ramesh, M., Berg, T., Learned-Miller, E.: Labeled faces in the wild: a database for studying face recognition in unconstrained environments. Technical Report 07–49, University of Massachusetts, Amherst (2007)
19. Shi, J., Qi, C.: From local geometry to global structure: learning latent subspace for low-resolution face image recognition. *IEEE Signal Process. Lett.* **22**(5), 554–558 (2015)
20. Zhou, C., Zhang, Z., Yi, D., Lei, Z., Li, S.Z.: Low-resolution face recognition via simultaneous discriminant analysis. In: *Biometrics (IJCB)*, 2011 International Joint Conference on, pp. 1–6. IEEE (2011)
21. Siena, S., Boddeti, V.N., Vijaya Kumar, B.V.K.: Coupled marginal fisher analysis for low-resolution face recognition. In: *Fusiello, A., Murino, V., Cucchiara, R. (eds.) ECCV 2012*. LNCS, vol. 7584, pp. 240–249. Springer, Heidelberg (2012). doi:[10.1007/978-3-642-33868-7_24](https://doi.org/10.1007/978-3-642-33868-7_24)