

INSTITUTO DE COMPUTAÇÃO
UNIVERSIDADE ESTADUAL DE CAMPINAS

**Person-Specific Face Representation
for Recognition**

*Giovani Chiachia Alexandre X. Falcão
Anderson Rocha*

Technical Report - IC-11-16 - Relatório Técnico

July - 2011 - Julho

The contents of this report are the sole responsibility of the authors.
O conteúdo do presente relatório é de única responsabilidade dos autores.

Person-Specific Face Representation for Recognition

Giovani Chiachia Alexandre X. Falcão Anderson Rocha *

Abstract

Most face recognition methods rely on a common feature space to represent the faces, in which the face aspects that better distinguish among all the persons are emphasized. This strategy may be inadequate to represent more appropriate aspects of a specific person's face, since there may be some aspects that are good at distinguishing only a given person from the others. Based on this idea and supported by some findings in the human perception of faces, we propose a face recognition framework that associates a feature space to each person that we intend to recognize. Such feature spaces are conceived to underline the discriminating face aspects of the persons they represent. In order to recognize a probe, we match it to the gallery in all the feature spaces and fuse the results to establish the identity. With the help of an algorithm that we devise, the Discriminant Patch Selection, we were capable of carrying out experiments to intuitively compare the traditional approaches with the person-specific representation. In the performed experiments, the person-specific face representation always resulted in a better identification of the faces.

1 Introduction

As a consequence of being one of the most active pursuits in computer vision, the face recognition problem has been addressed from many different perspectives [2]. However, from the seminal work of Kanade [5] to the novel approaches [2], the majority of the techniques consists, at some point, on representing the faces in a common feature space in order to compare them. By doing so, a consensus about the most discriminant way to represent all the faces must be obtained, where common face aspects that better distinguish different persons are emphasized with the purpose of achieving a good representation.

By ignoring face aspects that are not good at distinguishing among all the considered persons, we may also be ignoring aspects that are actually good at discerning a particular person from the others, i.e., we may be disregarding important person-specific face aspects only because their discriminability does not generalize to all persons. This way, the representation of a person's face in a feature space conceived to underline its discriminating aspects may lead to better recognition capabilities.

Given the ability that humans have in recognizing faces, it is worth looking at some findings about the human perception of the faces. The first finding we refer to is the

*The authors are with the Department of Information Systems, Institute of Computing - Unicamp, Av. Albert Einstein 1251, 13083-852, Campinas-SP, Brazil. Phone: +55-19-3521-5881. Fax: +55-19-3521-5847. E-mail: {chiachia,afalcao,anderson.rocha}@ic.unicamp.br.

negative correlation between the triviality and the recognizability of a face [8], which states that the more unusual the appearance of an individual’s face, the easier it is to remember it. The second matter of concern is the increasing ability of humans on identifying people from faces as they become familiar with them [11]. The third and last finding that we highlight is the neural processing distinction between the invariant and the variant aspects of the face. The idea is that the invariant aspects are related to the person identity, while the variants contribute to the social interactions [8]. Throughout the paper, we are going to name these three ideas as the *particularities*, the *familiarity*, and the *invariant aspects*.

Such an employment of domain knowledge to represent the face uniqueness has been a focus of recent research [16]. Beyond being in accordance to our intuitions, the person-specific face representation seemed to us an appropriate idea to model the evidence about the human face recognition process. In order to implement this idea, we devised a method that allowed us to determine and to select discriminant patches on a class-specific basis, the *Discriminant Patch Selection* (DPS). On the considered face database, this method enabled an effective modeling of the problem toward these biological mechanisms.

In addition to the DPS and the proposed implementation of the idea, a key contribution of this work is the evaluation of the “person-specific face representation” concept itself. Six out of the seven methods that we investigate are derived from the use of different patch setups to match the faces, including the proposed one that accomplishes the framework. Given that most of the analyses consist of comparing methods that differ only by the setup of patches that they employ, the evaluation becomes clear and convincing.

A small number of works somewhat near in concept to this one is presented in Section 2. After describing the face database and dissimilarity criterion (Sections 3 and 4), the DPS algorithm is presented in Section 5. Taking the database and the DPS into account, our implementation of the person-specific representation idea is described in Section 6, and the three rounds of experiments that were conducted are presented in Section 7. The conclusions obtained are then presented in Section 8, along with the possibilities that were opened and that we intend to investigate in the future.

2 Related Work

We noticed that just a few works are related to ours in the person-specific representation sense. A close one is the work of Krishna et al. [6], in which faces are represented by Gabor wavelets and the person specificity is given by optimally finding these wavelets in order to better separate two sets corresponding to samples of the same person and of different persons. Similarity is given by comparing a query sample to the clusters of samples represented in their respective feature spaces. Despite having indirect similarities to our work, they compare their approach just with respect to PCA, and no conclusions particularly related to the person-specific issue are obtained.

Another conceptually similar work takes into account the matching of a person-specific elastic graph whose nodes are determined by a hill-climbing algorithm that considers the ability of the nodes in distinguishing among the training samples [15]. In such work, performance is compared to more methods, but once again there is no comparison regarding the

benefits of the approach over its precisely *non person-specific* counterpart. While the aim of Krishna et al. [6] and Zafeiriou et al. [15] was to advance the state-of-the-art, our objective in this paper is also to shed light on whether or not the person-specific representation is an interesting framework for face recognition.

In [7], although the authors consider person-specific SIFT features, the matching process is carried out in a feature space common to all subjects. In more restricted contexts of face recognition, Sivic et al. [12] model specific classifiers that learn the best features to discriminate among faces of characters in TV or movie material, and Yao et al. [14] address the problem of face image retrieval with the combination of an offline classifier and a person-specific online classifier that is learned from the set of query images provided for the retrieval. Finally, Jain and Park [4] have recently proposed the use of features obtained from facial marks that are inherently specific to each person (e.g., freckles, moles, and scars) to supplement the features of existing face recognition approaches.

3 Face Database

A total of 1,864 visible images from X1 collection of the UND face database were considered in this study [3], corresponding to the 54 subjects whose attendance over the weekly acquisition sessions were highest. This way, a given person was recorded at least in seven and at most in ten sessions. In each session, four face images were obtained by the combination of two lighting conditions and two facial expressions, neutral and smiling.

In order to meet the *familiarity* issue and to enable the modeling of the *invariant aspects* of the faces, a number of face images of the individuals has to be available. Therefore, two images of the earliest recording session were taken first for the training process and then for the multisample gallery. These 54 pairs of images refer to one of the illumination conditions and both expressions, and the decision of employing only two samples of each person was made by regarding that this was enough to become familiar and to model the invariances of the faces in such a controlled scenario. Given the training/gallery set, all the remaining images were made probes.

Before being used in the experiments, all the 108 images from the training/gallery set and the 1,756 images from the probe set were registered by the position of the eyes and cropped with an elliptical mask. They were made 260×300 in size in order to keep details from the face features.

4 Dissimilarity Measure

The dissimilarities between faces in this work are obtained by matching a number of histograms of *Local Binary Patterns* (LBPs) that are extracted from patches on different positions of the faces. The matching idea is similar to the work of Ahonen et al. [1], but with a different metric. As in our work the LBPs are calculated from an 8-neighborhood, the histograms from the patches have $b = \{1, 2, \dots, 256\}$ bins. Formally, let \mathcal{P}' be the set of patches considered for the matching and H_p be the histogram of the LBPs within patch

p . The patch-based dissimilarity between images \mathbf{A} and \mathbf{B} is

$$D(\mathbf{A}, \mathbf{B}, \mathcal{P}') = \sum_{\forall p \in \mathcal{P}'} \sum_{b=1}^{256} |H_{p,b}(\mathbf{A}) - H_{p,b}(\mathbf{B})|, \quad (1)$$

where $H_{p,b}$ represents the value of bin b of patch p . In simple words, the dissimilarity corresponds to the summation of the absolute difference over the bins of each patch histogram. The only method whose matching criterion is different from this one is the baseline PCA that we also compare in the experiments.

5 Discriminant Patch Selection (DPS)

The idea of DPS is to model and to select patches according to the discriminability they have in a group of training images coherent both in content and in space, i.e., from registered images representing objects of a common category.

For a given patch in a given training image, its *discriminability* is measured on an individual basis with respect to the patches in the other training samples. By interchanging such a training image, the same patch is modeled in all images (and for all classes). This is done for the whole set of patches to be modeled. At the end, all patches are associated with the discriminabilities that are taken for their *selection*.

Two important aspects of DPS are the discriminability measure and the patch selection criterion. The discriminability is based on the combination of (i) features extracted from the patches, (ii) a dissimilarity measure, and (iii) a matching strategy. For the patch selection criterion, besides the (i) discriminability of the patches, it may also be regarded their (ii) spatial information and (iii) the problem class to which the patch is associated. Given the number of possibilities that these two criteria can be defined, in our formulation they are both taken as parameters. In Section 6.1, we present how they are defined in this work.

Assume that function $F(p, \mathbf{T}, \mathcal{G})$ returns how good patch p of image \mathbf{T} discriminates its class with respect to other patches in the image set \mathcal{G} . Consider that function $C(\mathbf{T})$ retrieves the correct class $c \in \mathcal{C}$ to which image \mathbf{T} pertains. In addition, take \mathcal{P} as the set of patches considered for selection and \mathcal{T} as the set of training images. A pseudocode of the method can be defined as follows.

Algorithm 1. – DPS

INPUT: The set of images \mathcal{T} , the set of patches \mathcal{P} , the discriminability function $F(p, \mathbf{T}, \mathcal{G})$, and the *patch selection criterion*.

OUTPUT: Models M_c and patches \mathcal{P}' selected according to the provided criterion.

AUXILIARY: Function $C(\mathbf{T})$, image \mathbf{T} , and variables c and d .

1. $M_{|\mathcal{C}|, |\mathcal{P}|} \leftarrow 0 \quad \forall c \in \mathcal{C} \text{ and } \forall p \in \mathcal{P}$.
2. **For each** patch $p \in \mathcal{P}$ **do**
3. **For each** image $\mathbf{T} \in \mathcal{T}$ **do**
4. $c \leftarrow C(\mathbf{T})$.
5. $d \leftarrow F(p, \mathbf{T}, \mathcal{T} \setminus \mathbf{T})$.
6. $M_{c,p} \leftarrow M_{c,p} + d$.

7. **Select patches** from models M_c into \mathcal{P}' according to the criterion related to their discriminability.

Note that M_c is considered a class-specific model in the sense that the discriminability of patches $p \in \mathcal{P}$ modeled for class c are represented in $M_{c,p}$. The patch selection criterion may take into account the discriminability of the patches by the problem classes (i.e., by M_c), but may also fuse the classes in order to consider patch discriminabilities common to the whole training set. In the first case, class-specific *Discriminant Patch Models* are built, while in the second case, we end up with a general *Discriminant Patch Model*.

Despite this novel formulation, the concept of selecting patches to better describe object classes in images has already been studied in recent years. For example, in [13], the authors present two methods for selecting informative patches in order to detect objects. One of the methods is related to DPS in the sense that it selects patches regarding their individual discriminability, while the other is modeled as a combinatorial optimization problem. Sarfraz and Khan [10] had also explored this idea by selecting a set of patches that are used to recognize types of vehicles. In their work, they employ a probabilistic discriminant criterion and the selection of the patches is class-specific.

As mentioned in [13], it is important to recall that this approach is different from the Boosting method. DPS deals with the patches individually, with the idea that a synergistic composition of discriminant patches will produce good results at an acceptable computational burden. Boosting, in turn, is a machine learning technique based on the combination of weak classifiers, and its employment for feature selection may become very expensive in some cases. Moreover, DPS provides a natural way to determine class-specific sets of discriminant patches rather than only one.

6 Implementation of the Person-Specific Idea

The DPS was formulated to fulfill the modeling of person-specific face representations along with the regarded observations on how faces are perceived by humans. Given the constrained scenario of the employed face images and the multisample gallery, this method enabled us to capture both the *invariant aspects* and the *particularities* specific to each person's face. The training process that implements the idea is an instance of the algorithm.

6.1 DPS Training

In this section we define the four input parameters of DPS that we use in order to obtain person-specific face representations. The notations are according to Algorithm 1.

- The first parameter is the collection of training images \mathcal{T} , which refers to the 108 face images presented in Section 3.
- The second argument that we define is the set \mathcal{P} of patches considered for selection. The size of the patches was fixed in 20×20 pixels by considering that, in images with 260×300 pixels, this was appropriate to represent the face aspects. Assuming that

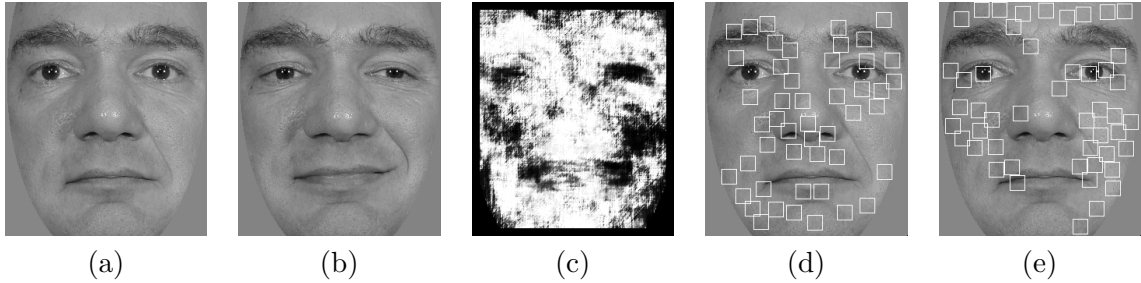


Figure 1: An illustration of the person-specific DPS process employed in this work. For a modeled person, patches from his training images (a) and (b) are alternately matched to their corresponding patches in the other training images. His *Discriminant Patch Model* is then derived (c), enabling the selection of a subset of patches by a given criterion. The higher the discriminability, the brighter the pixel that corresponds to the center of the patch. Image (d) shows the most discriminant patches and reflects the implementation of the idea, and image (e) shows the patches that would be obtained by selecting the least discriminant ones.

the candidate patches for selection must lie entirely in the image domain and that we consider all possible patches satisfying this constraint, a total of $|\mathcal{P}| = 67,721$ patches are considered for each person being modeled.

- The third DPS parameter is the criterion $F(p, \mathbf{T}, \mathcal{G})$ to measure the discriminability of a given patch p in a given pivot image \mathbf{T} with respect to the other training images. For this purpose, we use the recognition rank obtained by matching the patch according to Equation 1, where $\mathbf{A} = \mathbf{T}$, $\mathbf{B} \in \mathcal{G}$, and $\mathcal{P}' = \{p\}$. The discriminant criterion is the complement of the rank, provided that the lower the rank, the more discriminant the patch. Such a measurement of discriminability by the recognition rank was only possible because of the multisample gallery.
- The fourth and last definition necessary to obtain person-specific face representations from DPS is the *patch selection criterion*. For each individual c , the 48 most discriminant patches from the models M_c are selected into \mathcal{P}'_c . We decide to consider such a number of patches because it seemed for us that it was enough to describe the person-specific most important face aspects. In order to avoid overlapping patches, we made a constraint on the selection such that each new selected patch must have its center at a minimum distance of 28 pixels from the previously selected ones.

In Figure 1, an illustration of the person-specific DPS training is presented for one of the modeled persons. Figures 1(a) and 1(b) are his corresponding images in the training set. His person-specific *Discriminant Patch Model* is shown in Figure 1(c). The higher the discriminability, the brighter the pixel that corresponds to the center of the patch. By the end of the DPS process, patches can be selected by a given criterion. Figure 1(d) shows the most discriminant patches and reflects the implementation of the idea, and Figure 1(e) shows the patches that would be obtained by selecting the least discriminant ones.

As it is possible to observe in Figure 1(d), a greater importance was given to the invariant aspects of the face with respect to the two provided training images of this person (Figures 1(a-b)). The visual aspect of the eyebrows seemed to be, to some extent, one of his particularities, as well as the texture of his skin in the invariant parts of his face. Here it is important to remark that LBPs are good for representing texture, and this certainly reflects in the modeling (in this case, in the importance of the skin texture).

6.2 Recognition in Multiple Feature Spaces

The person-specific representation implies that we have one feature space defined by \mathcal{P}'_c for each subject c considered in the training set. In order to recognize a probe \mathbf{A} , we match it to all samples $\mathbf{B} \in \mathcal{T}$ of the gallery in all the feature spaces (see Equation 1).

This strategy leads to a number of *nearest neighbor* classifications. In order to obtain a general decision, we fuse the responses of all person-specific representation classifiers in a voting scheme, and the identity which has the most votes is the one related to the probe face.

The type of multibiometrics that we employ in this work is not accommodated by the taxonomy provided by Ross et al. [9], which states that a multibiometric system can operate considering (i) multiple sensors, (ii) multiple algorithms operating on the same data, (iii) multiple instances of the biometric trait (e.g., different fingers), (iv) multiple samples of the trait, (v) multiple traits (e.g., face and voice), and (vi) a combination of these scenarios.

Although our approach employs a multisample gallery (iv), its main characteristic is the use of multiple representations for the same face instance, obtained with the same sensor and by the same algorithm. This may be considered a new type of multibiometrics, perhaps a *Multirepresentation* one.

6.3 Running Example

In Figure 2, we provide a running example of the recognition scheme adopted in this work. From top to bottom, the diagram starts with the person-specific representation of the gallery samples $G_{c,m}$ into the feature spaces R_c , where c denotes the modeled persons in the training/gallery set, and m indicates which of the multiple samples of the person is being considered. In this example, we have two persons modeled with two gallery samples each. Thus, $c = \{1, 2\}$ and $m = \{a, b\}$. After representing the gallery samples in each person-specific feature space, we obtain samples $G_{c,m}^c$, which means $G_{c,m}$ represented in the feature space modeled for person c .

In order to recognize a probe P , we represent it in each feature space R_c , and the resulting P^c samples are correspondingly matched to the gallery. In this example, we match P^1 to the samples $G_{c,m}^1$ and P^2 to the samples $G_{c,m}^2$. The matchings are then ranked according to the dissimilarities and an identity is established by each *nearest-neighbor* classifier. Here we have two classifiers, one for R_1 and one for R_2 . Finally, a voting scheme is done by considering the decisions of the classifiers, and the person in the gallery which has the most votes is taken as the probe identity.

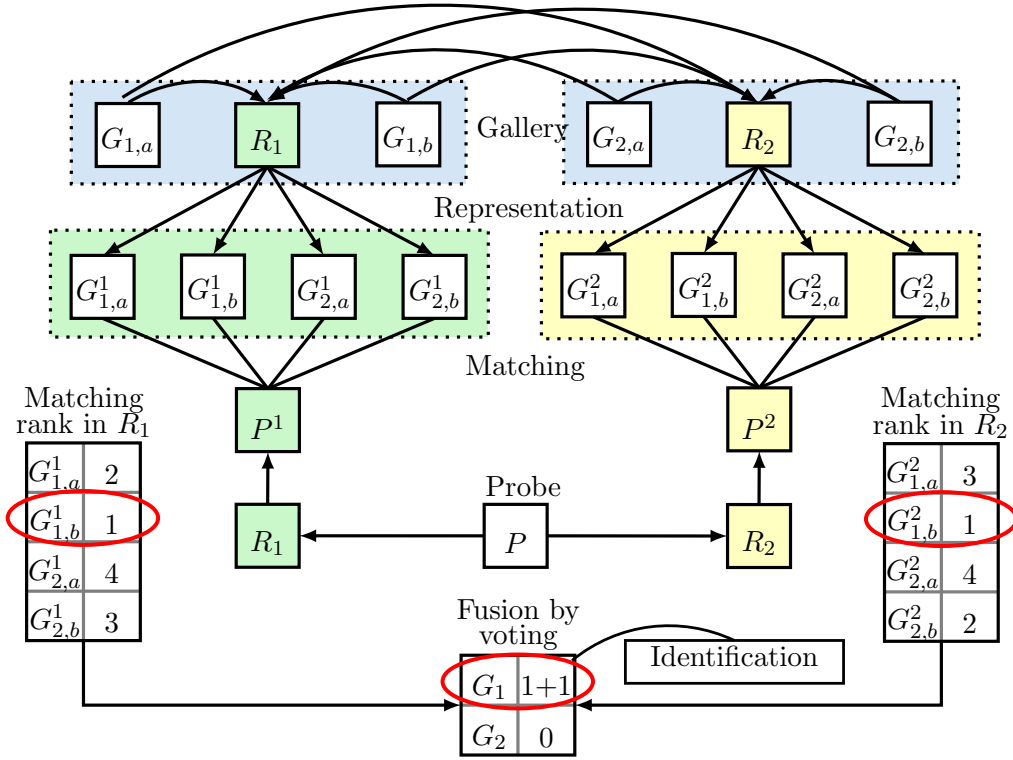


Figure 2: Illustration of the recognition scheme adopted in this work considering $c = \{1, 2\}$ persons in the training/gallery set, with $m = \{a, b\}$ samples each. From top to bottom, the diagram starts with the person-specific representation of the gallery $G_{c,m}$ into the feature spaces R_c . Such a representation results in $G_{c,m}^c$, which means $G_{c,m}$ represented in the feature space c . Given a probe P , its representations P^c are correspondingly matched to the gallery. The matchings are then ranked and an identity is established by each classifier. Finally, a voting scheme is done by considering their decisions, and the person which has the most votes is taken as the probe identity.

7 Experiments and Results

In this section, we present the three rounds of experiments that were carried out in order to assess whether or not the person-specific representation of faces has advantages. The method that reflects the idea is compared in the three rounds and, because it consists of the selection of the 48 most discriminant patches of each individual, it is named *Person-Specific Most Discriminant Patches* (PSMDP).

The results are presented with Cumulative Match Characteristic (CMC) plots. With the purpose of providing equivalent comparisons between CMC plots obtained from person-specific methods that employ the voting scheme (Section 6.2) and from traditional methods with one face representation space, only the best match of each identity was considered for

the CMC plots of the traditional methods.

7.1 Validation

We validate the person-specific representation idea by comparing the PSMDP method with two other methods that only differ by the patch setup they consider. The first method is obtained by considering the *Person-Specific Least* discriminant patches (instead of the most ones) of each individual. As it can be noticed in Figure 3, the difference in the identification rate of these two methods is considerable, and this was expected, since this experiment was aimed at assessing the effectiveness of the implementation.

The second validation is the most interesting outcome of this study. It consists of comparing the performance of PSMDP to the performance by taking into account the assembling of *All Person-Specific Most* (APSM) discriminant patches. The interesting thing to observe in this validation is the fact that the same data are employed by both methods. In the PSMDP case, we consider 54 feature spaces of 48 patches each, while in the APSM case, we consider one feature space with the same $54 \times 48 = 2,592$ patches, i.e., we are dealing with the same images, patches, etc., and the only difference relies in the feature

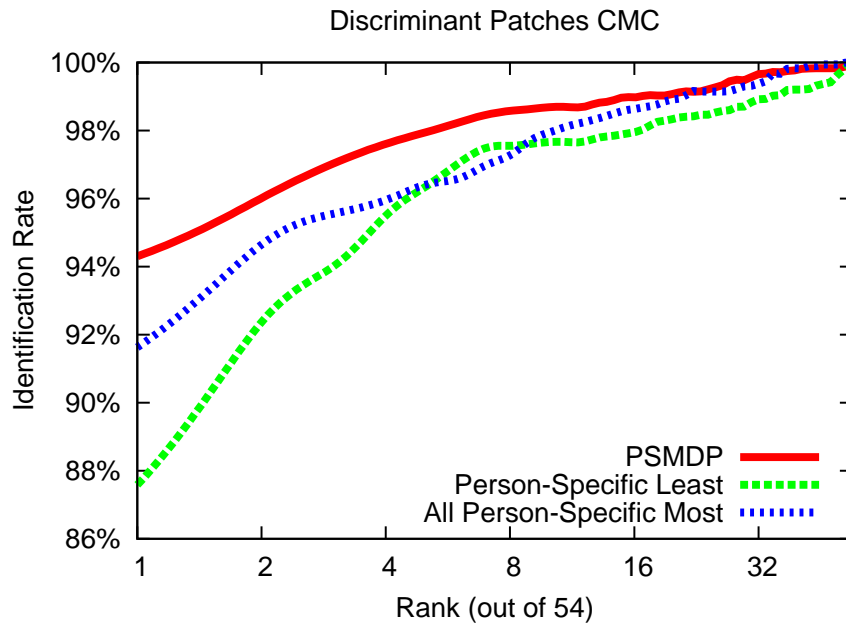


Figure 3: Validation of the idea. As expected, the performance by considering the *Person-Specific Most Discriminant Patches* (PSMDP) is considerably better than the performance of the *Person-Specific Least* discriminant patches. The most interesting outcome of this work, however, comes with the comparison between PSMDP and the *All Person-Specific Most* method. As the data employed by both are exactly the same, the idea of the person-specific face representation is strengthened here.

space “concatenation” made in APSM. This can be considered a comparison between the *person-specific* representation idea and its exactly *non-person specific* counterpart.

As it is possible to observe in Figure 3, the difference in performance between PSMDP and APSM is more than 2.5% (it dropped from 94.31% to 91.63%). Given such a difference, it becomes clear that the person-specific face representation idea is valid and that it deserves attention. It seems that the manifold of a person’s face becomes simpler in its own representation space, leading to an overall better recognition.

7.2 General and Random Representations

In this round of experiments, the aim was at comparing PSMDP with two more alternatives. The *General Discriminant* (GD) alternative refers to the selection of the 48 patches that are the most discriminant for all persons simultaneously. The difference between GD and the *All Person-Specific Most* (APSM) of the previous section is that, in GD, the patch discriminability from the persons are correspondingly merged before the selection, while, in APSM, the patches are assembled after the person-specific selections. The GD case can be viewed as an adaptation of the way that many face recognition methods represent the face features, i.e., by emphasizing the importance of the aspects that better distinguish among all the faces.

The general *Discriminant Patch Model* that was obtained is shown in Figure 4(a), and the corresponding most discriminant patches are illustrated in Figure 4(b). Note that the patch positions are the same for all individuals in this case. Some findings in the literature can be observed here [2, 11], mainly the importance of the eyebrows in distinguishing persons. The eyebrows were followed by the nose and the superior part of the eyes. Due to the presence of facial expressions in the database, the discriminant relevance of the mouth was decreased. This is in accordance to the idea of the *invariant aspects* mentioned in Section 1.

The second alternative that we compare in this section consists of the randomly and person-specifically selection of patches within the elliptical face domain. The performance of this random method was assessed ten times, and its resulting CMC is the average of these runs over the ranks. With this comparison, the person-specific discriminant face representation is further validated with respect to stochasticity in the selection of the face aspects.

In Figure 5, the comparison with these two alternatives is presented. One can see that the performance of the claimed face representation is also superior. In the *General Discriminant* case, it is worth noting that the amount of feature data is much smaller. However, if we have selected a considerably large number of patches, we would cover the entire face image. This possibility is investigated in Section 7.3.

Regarding the *Person-Specific Random* (PSR) alternative, the amount of feature data is the same as in the PSMDP case, and the difference relies only on the patch selection criterion. Despite the random patch selection, the performance of PSR was good. We believe that the random criterion, by being uniform and not allowing overlapping patches, enabled a selection of patches well distributed within and among the person-specific representations. This possible scattered representation may have led to the good performance.

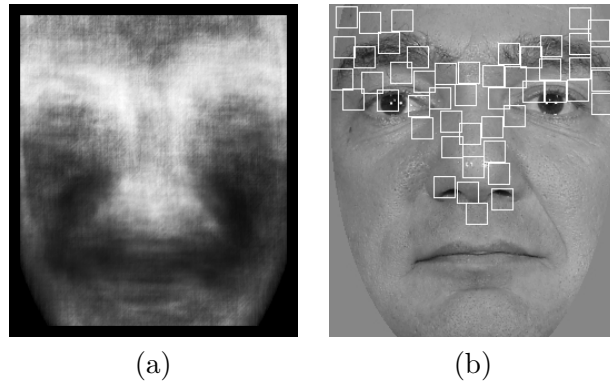


Figure 4: The general *Discriminant Patch Model* (a) and an illustration of the most discriminant patches corresponding to the model (b). In this case, the discriminant patches are the same for all individuals. The results corroborate with the literature with respect to the importance of the eyebrows in distinguishing persons [2, 11]. The eyebrows were followed by the nose and the superior part of the eyes. Due to the presence of facial expressions, the relevance of the mouth (a variant aspect) was decreased.

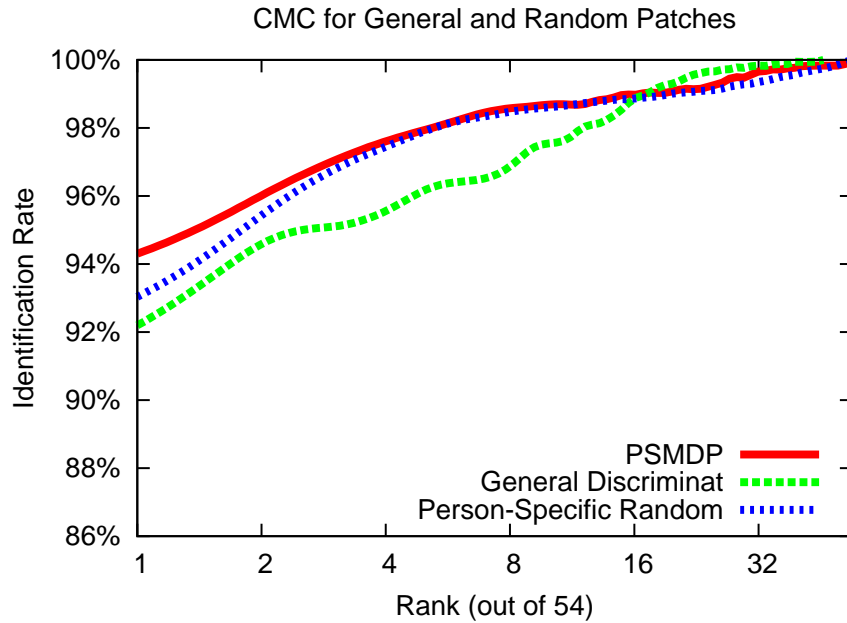


Figure 5: Comparison of the claimed face representation to the *General Discriminant* and the *Person-Specific Random* cases. Neither the common way to discriminate faces nor the random selection of the patches outperform PSMDP.

7.3 Intuitive Baselines

So far, we have compared methods directly derived from the selection of patches according to different criteria. In this last part of the experiments, we present the performance of two methods found in the literature that seemed appropriate for us to compare. The objective is to estimate how they perform in this setup of face images.

Still concerning the patch-based idea, we first compare the PSMDP method with the *Grid of Patches* approach, which stands for a regular composition of non-overlapping patches similar to the idea in [1]. Given the image and the patch size, this method employs a grid of $13 \times 15 = 195$ patches to describe the faces.

The second method of this last round of experiments is based on the use of the well-known PCA technique to achieve a good representation of the faces. The face subspace was created from the neutral image of each subject in the gallery by retaining the first eigenvectors corresponding to 95% of the total variance. Since the Mahalanobis cosine between the projected samples provided better results than the Euclidean distance, the performance is presented by considering this metric. We call this method *PCA Mahalanobis*.

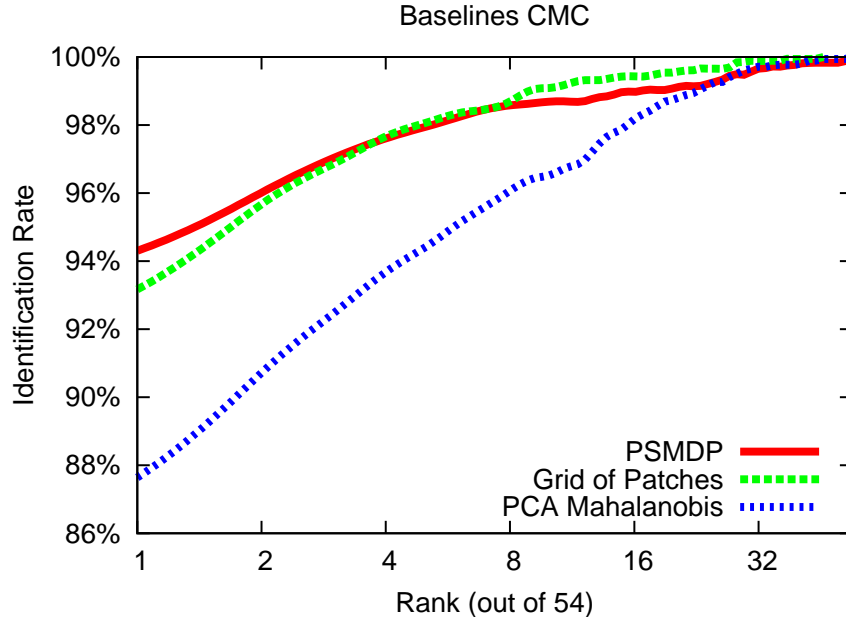


Figure 6: Intuitive baselines. The *Grid of Patches* approach stands for a regular composition of non-overlapping patches, while the *PCA Mahalanobis* method consists on measuring face dissimilarities with the Mahalanobis cosine distance in a subspace created by the well-known PCA technique.

One can observe in Figure 6 that the person-specific discriminant representation is also more effective than these two alternatives in terms of identification rate. Although the better performance of PSMDP comes at a price of matching the probe to the gallery in all

feature spaces, this should not be a matter of concern in this work, given that the person-specific face representation can be considered a new way to handle the face recognition problem and its computational burden can be readily addressed by multicore computing.

8 Conclusions and Future Work

By means of a plain and effective methodology, the results presented in this paper clearly support the person-specific face representation idea as a way to overcome recognition deficiencies related to representing the face aspects in a feature space common to all persons.

Beyond the implementation and the evaluation of the idea, another contribution of this work is the formulation of the Discriminant Patch Selection (DPS) algorithm. This method allowed us both to construct effective person-specific face models and to carry out experiments that were intuitive to compare. Combined with the considered face database, DPS also enabled us to model some findings in the human perception of faces.

As a last contribution, it is also worth remarking that the recognition scheme proposed in this study may be understood as a new type of multibiometrics, one that is based on multirepresentations [9].

The first aspect we intend to further investigate is the person-specific face representation idea under less constrained scenarios, where we can count on more challenging face samples to build the models. In this direction, we plan to assess some connections that can be made between the DPS process and a fine deformable registration of the faces. Moreover, we also intend to study heuristics to split the person-specific models based on the face pose.

Meanwhile, other subjects that are also worth attention for improving the method are related to (i) the number and the size of the patches (that here were fixed), (ii) the use of additional image descriptors other than LBPs, (iii) the improvement of the patch discriminability and selection criteria, and (iv) the employment of more sophisticated classification engines that can be built tightly coupled to the person-specific representation idea, aiming at improving both efficiency and effectiveness.

References

- [1] Timo Ahonen, Abdenour Hadid, and Matti Pietikinen. Face recognition with local binary patterns. In *ECCV*, pages 469–481. IEEE, 2004.
- [2] Rama Chellappa, Pawan Sinha, and P. Jonathon Phillips. Face recognition by computers and humans. *IEEE Computer*, 43:46–55, February 2010.
- [3] X. Chen, P. J. Flynn, and K. W. Bowyer. Visible-light and Infrared Face Recognition. *ACM Workshop on MMUA*, pages 48–55, December 2003.
- [4] Anil K. Jain and Unsang Park. Facial marks: soft biometric for face recognition. In *ICIP*, pages 37–40. IEEE, 2009.
- [5] Takeo Kanade. *Picture Processing by Computer Complex and Recognition of Human Faces*. PhD thesis, Kyoto University, 1973.

- [6] Sreekar Krishna, John Black, and Sethuraman Panchanathan. Using genetic algorithms to find person-specific gabor feature detectors for face indexing and recognition. In David Zhang and Anil Jain, editors, *Advances in Biometrics*, LNCS, pages 182–191. Springer, 2005.
- [7] Jun Luo, Y. Ma, E. Takikawa, S. Lao, M. Kawade, and Bao-Liang Lu. Person-specific sift features for face recognition. In *ICASSP*, volume 2, pages 593–596. IEEE, april 2007.
- [8] A. J. O’Toole. Psychological and Neural Perspectives on Human Face Recognition. In S. Z. Li and A. K. Jain, editors, *Handbook of Face Recognition*, chapter 15, pages 349–369. Springer, 2004.
- [9] Arun Ross, Karthik Nandakumar, and Anil Jain. Introduction to multibiometrics. In Anil K. Jain, Patrick Flynn, and Arun A. Ross, editors, *Handbook of Biometrics*, pages 271–292. Springer US, 2008.
- [10] M. Saquib Sarfraz and M.Haris Khan. A probabilistic framework for patch based vehicle type recognition. In *VISAPP*, volume 1, March 2011.
- [11] P. Sinha, B. Balas, Y. Ostrovsky, and R. Russell. Face recognition by humans: Nineteen results all computer vision researchers should know about. *Proceedings of the IEEE*, 94(11):1948 –1962, nov. 2006.
- [12] J. Sivic, M. Everingham, and A. Zisserman. “Who are you?”: Learning person specific classifiers from video. In *IEEE CVPR*, 2009.
- [13] Akshay Vashist, Zhipeng Zhao, Ahmed Elgammal, Ilya Muchnik, and Casimir Kulikowski. Discriminative patch selection using combinatorial and statistical models for patch-based object recognition. In *IEEE CVPRW*, pages 12–, 2006.
- [14] Bangpeng Yao, Haizhou AI, and Shihong Lao. Person-specific face recognition in unconstrained environments: a combination of offline and online learning. In *IEEE FG*, pages 1 –8, sept. 2008.
- [15] S. Zafeiriou, A. Tefas, and I. Pitas. Learning discriminant person-specific facial models using expandable graphs. *IEEE T. IFS*, 2(1):55–68, 2007.
- [16] W. Zhao, R. Chellapa, P. J. Phillips, and A. Rosenfeld. Face Recognition: A Literature Survey. *ACM Computing Surveys*, 35(4):399–458, December 2003.