

Partial Least Squares for Face Hashing

Cássio E. dos Santos Jr., William Robson Schwartz

Department of Computer Science - Universidade Federal de Minas Gerais
Belo Horizonte - MG - Brasil

{cass,william}@dcc.ufmg.br

Introduction

In this work, we focus on the face identification task. Specifically, the main goal is to provide a face identification approach scalable to galleries consisting of numerous subjects and on which common face identification approaches would probably fail on responding in low computational time. There are several applications for a scalable face identification method: surveillance scenarios, human-computer interaction and social media. The few aforementioned applications show the importance of performing face identification fastly and, in fact, several works in the literature have been developed in the past years motivated by these same types of applications (surveillance, forensics, human-computer interaction, and social media). However, most of the works focus on developing fast methods to evaluate one test face and a single subject enrolled in the gallery. These methods usually develop low computational cost feature descriptors for face images that are discriminative and with low memory footprint enough to process several images per second. Note that these methods still depend on evaluating all subjects in the face gallery. Therefore, if the number of subjects in the gallery increases significantly, these methods will not be able to respond fastly and new methods shall be developed to scale the face identification to this larger gallery.

Face identification methods usually consists of a face representation or description in the feature vector where mathematical models can be applied to determine the face identity. In this case, it is used one model to determine each identity in the face gallery, therefore, being necessary a number of models equal to the gallery size. Note that the parameters in each model are learned using samples for each subject in the face gallery and every model must be evaluated to correctly identify a test sample. In this work, we propose a method to reduce the number of models evaluated in the face identification by eliminating identities that are somewhat clearly not the identity in the test sample. Figure 1 illustrates the common face identification pipeline employed in practice and the main component tackled in this work.

There is an extensive literature of works regarding large-scale image retrieval that could be employed in face identification. However, most of these works focus on returning a list containing images from the dataset that are similar to the test image. Although reasonable to recover images in large datasets, such approaches are not suitable to apply directly to the face identification task. The models from subjects in the face gallery should optimally be described regarding the discriminative features related to each subject identity, which might consume less memory, specially if several samples per subject are available, and less computational time since only discriminative features are evaluated to determine the face identity.

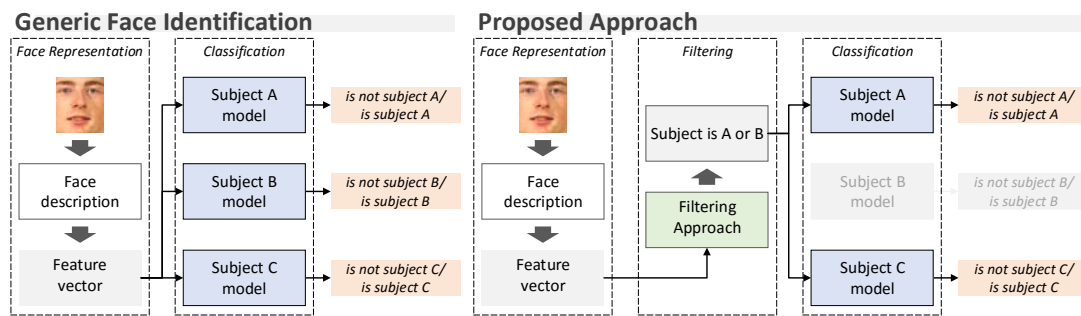


Figure 1: Common face identification pipeline and the proposed pipeline with the filtering approach which is used to reduce the number of evaluations in the classification step with low computational cost. The filtering approach is the main contribution in this work and it is tailored considering recent advances in large-scale image retrieval and face identification based on PLS.

The proposed approach is inspired by the family of methods regarded as *locality-sensitive hashing* (LSH), which are the most popular large-scale image retrieval method in the literature, and the *partial least squares* (PLS), which has been explored intensively in numerous past works regarding face recognition. We call the proposed approach *PLS for hashing*, abbreviated to PLSH and ePLSH in its extension. The main goal in LSH is to approximate the representation of samples in the high dimensional space using a small binary representation where the search can be implemented efficiently employing a hash structure to approximate near-identical binary representations. The idea in LSH is to generate random hash functions to map the feature descriptor in the high dimensional representation to bits in the binary representation.

In the PLSH approach, the random projection in the aforementioned example is replaced by a PLS regression, which provides discriminability among subjects in the face gallery and allow us to employ a combination of different feature descriptors to generate a robust description of the face image. PLSH is able to provide significant improvement over the brute-force approach (evaluating all subjects in the gallery) and compared to other approaches in the literature. Furthermore, since the evaluation of hash functions in PLSH requires a dot product between the feature and regression vectors, additional speedup can be achieved by employing feature selection methods, resulting on the extended version of PLSH (ePLSH).

The following contributions are presented in this work. (i) A fast approach for face identification that support a combination of several feature descriptors and high dimensional feature vectors. (ii) The proposed approach presents at least comparable performance with other methods in the literature and up to 58 times faster when enough samples per subject are available for train. (iii) Extensive discussion and experimentation regarding alternative implementations that may guide future development in scalable face identification methods. (iv) The proposed approach is easy to implement and to deploy in practice since only two trade-off parameters need to be estimated.

This work resulted in the following publications in which the former presents PLSH and the latter presents ePLSH. Jr, C. E. S., Kijak, E., Gravier, G., and Schwartz, W. R. (2015). Learning to hash faces using large feature vectors. In *Content-Based Multimedia Indexing (CBMI), 13th IEEE Workshop on*, pages 1–6, and C E Santos Jr, E Kijak, G. G. W. R. S. (2016). Partial least squares for face hashing. *Elsevier Neurocomputing – Special Issue on Binary Representation Learning in Computer Vision*, pages 1–44.

Experimental setup

We evaluate PLSH and ePLSH in two standard face identification datasets, FERET and FRGCv1. The facial recognition technology (FERET) dataset [Phillips et al. 2000] consists of 1,196 images, one per subject for training, and four test sets designed to evaluate the effects of lightning conditions, facial expression and aging on face identification methods. The test sets are: *fb*, consisting of 1,195 images taken with different facial expressions; *fc*, consisting of 194 images taken in different lightning conditions; *dup1*, consisting of 722 images taken between 1 minute and 1,031 days after the gallery image; *dup2*, is a subset of *dup1* and consists of 234 images taken 18 months after the gallery image. In our experiments, all images were cropped in the face region using annotated coordinates of the face, scaled to 128×128 pixels and normalized using the self-quotient image (SQI) method to remove lightning effects [Wang et al. 2004].

The face recognition grand challenge dataset (FRGC) [Phillips et al. 2005] consists of 275 subjects and samples that include 3D models of the face and 2D images taken with different illumination conditions and facial expressions. We follow the same protocol described by Yuan et al. [Yuan et al. 2005], which considers only 2D images and consists in randomly selecting different percentages of samples from each subject to compose the face gallery and using the remaining samples to test. The process is repeated five times and the mean and standard deviation of the rank-1 recognition rate and speedup (considering the brute-force approach) are reported. The samples were cropped in the facial region, resulting in size 138×160 pixels, and scaled to 128×128 pixels.

All experiments regarding parameter validation were performed on the FERET dataset, since it is the dataset with the largest number of subjects (1,196 in total). FERET consists of four test sets and we use *dup2* to validate parameters since it is considered the hardest of the dataset.

We consider four feature descriptors in this work, CLBP [Ahonen et al. 2006], Gabor filters [Randen and Husoy 1999], HOG [Dalal and Triggs 2005] and SIFT [Lowe 2004], which mainly captures information about texture and shape of the face image. This set of features was chosen because they present slightly better results in the face identification and indexing compared to the previous works [Schwartz et al. 2012, Jr et al. 2015].

The error rate of the pipeline as described in Figure 1 results from errors induced by the filter approach (fail to return identity of test sample in the candidate list) and by the face identification approach (fail to identify correctly the subject in the candidate list). Therefore, to assess the performance of the filter approach alone, we provide results considering the *maximum achievable recognition rate* (MARR) [Jr et al. 2015], which is calculated considering that a perfect face identification method is employed for different percentages of candidates visited in the list. Note that the MARR value is the upper bound for the recognition rate achieved by the filter and face identification pipeline.

Experimental results

Results regarding MARR and rank-1 recognition rate for PLSH in all test sets from the FERET dataset are presented in Figures 2(a) and 2(b). For the test sets *fb* and *fc*, about 1% of subjects in the candidates list is enough to achieve more than 95% of the rank-1 recognition rate of the brute-force approach (presented in the legend of Figure 2(b) for each test

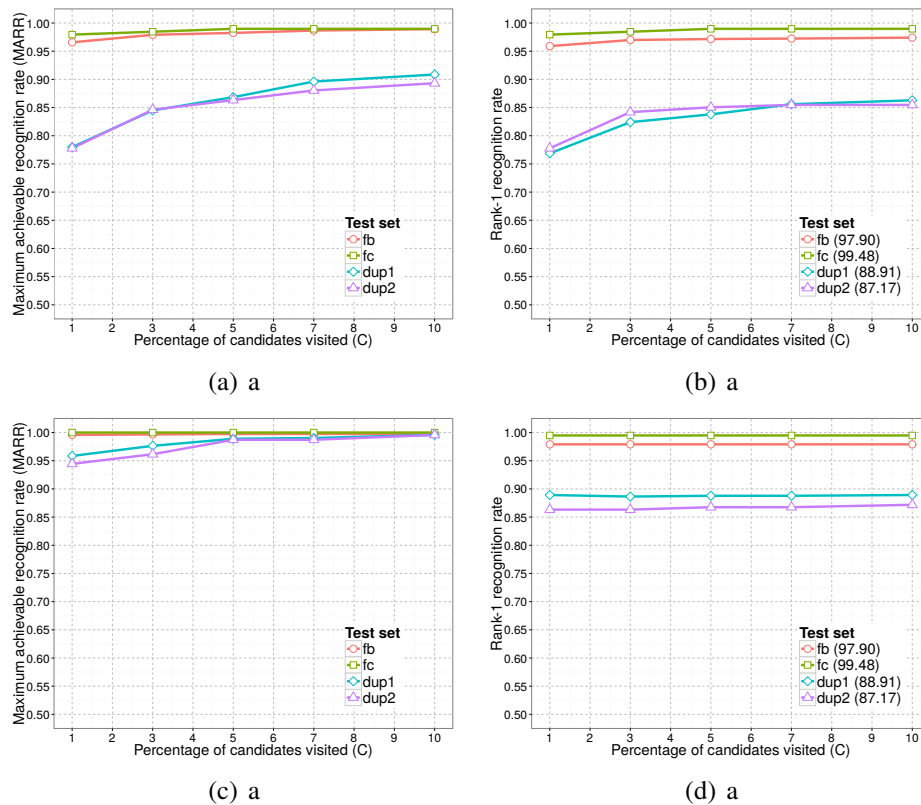


Figure 2: Results on the FERET dataset. (a) PLSH MARR curves, (b) PLSH rank-1 recognition rate, (c) ePLSH MARR curves and (d) ePLSH rank-1 recognition rate. Number in parenthesis indicate rank-1 recognition rate for the brute force approach.

set). However, for the test sets *dup1* and *dup2*, about 5% of subjects in the candidate list ensured at least 95% of the brute-force rank-1 recognition rate. The theoretical speedup in the worst case can be calculated considering the 150 PLSH hash function evaluations and the 5% of the gallery size, which consists of 60 PLS projections. In this case, the number of PLS projections would be 210 compared to the 1,196 projections necessary in the brute-force approach, which would still result in a 5.6 times speedup.

Results from ePLSH are presented in Figures 2(c) and 2(d). Using only 1% of subjects in the candidate list, it is possible to recover all subjects in the rank-1 recognition rate from brute-force approach for all four test sets. In this case, the rank-1 recognition rate from the ePLSH pipeline is the same as the brute-force approach, but with reduction to 1% of the subjects evaluated in the identification. Considering that the cost to evaluate all hash models in ePLSH is about the same as in PLSH, the theoretical speedup is 7.38 times compared to the brute-force approach in the worst case.

Results from the FRGC dataset for PLSH and ePLSH are presented in Table 1 along with results from three other methods as presented in the literature. The three methods are the cascade of rejection classifiers (CRC) from [Yuan et al. 2005], the PLS-based search tree [Schwartz et al. 2012], and our previous published work [Jr et al. 2015], which consists of PLSH with the combination of HOG, Gabor filter and LBP feature descriptors. For PLSH and ePLSH, we vary the number of hash models and the maximum percentage of subjects visited in the candidate list and we present the results with rank-1 recognition rate close to 0.95 and higher speedups. In this way, it is possible to compare

		% of samples for train				
		90%	79%	68%	57%	35%
CRC [Yuan et al. 2005]	Speedup	1.58×	1.58×	1.60×	2.38×	3.35×
	Rank-1 rec. rate	80.5%	77.7%	75.7%	71.3%	58.0%
Tree-based [Schwartz et al. 2012]	Speedup	3.68×	3.64×	3.73×	3.72×	3.80×
	Rank-1 rec. rate	94.3%	94.9%	94.3%	94.46%	94.46%
PLSH	Speedup	18.24×	8.61×	6.95×	3.96×	3.49×
	Rank-1 rec. rate	95.31%	95.31%	93.60%	94.67%	94.60%
ePLSH	Speedup	233.61×	98.93×	45.42×	22.29×	14.21×
	Rank-1 rec. rate	96.03%	95.02%	95.98%	94.67%	94.44%

Table 1: Comparison between the proposed approach and other approaches in the literature. The highest speedups are shown in bold. The full table with more information can be found in page 58 of the dissertation.

directly the maximum speedup achievable when using PLSH and ePLSH compared to the other approaches, which also provide rank-1 recognition rate close to 0.95.

According to Table 1, the speedup for PLSH and ePLSH decreases considerable as the number of samples per subject available for train reduce. The reason for that is the increase in the number of hash models and the maximum number of subjects visited in the candidate list to guarantee at least 0.95 rank-1 recognition rate. Even with reduced speedups considering 35% of samples available for train, ePLSH provides significant improvement over the speedup achieved by the tree-based approach (3.6 times faster), while PLSH provides competitive speedup.

The speedup provided by PLSH and ePLSH compared to the tree-based approach is noticed with 90% of the samples available for train, where PLSH is about 5 times faster than the tree-based approach while ePLSH is about 13 times faster than PLSH. Finally, in the worse case, ePLSH provides at least 14 times speedup considering the brute-force approach in the setup with 200 hash models and 10% of subjects in the candidate list.

Conclusions

In this work, we proposed and evaluated PLSH and its extension ePLSH for face indexing. PLSH is inspired by the well-known locality-sensitive hashing for large-scale image retrieval and PLS for face identification, which provides fast and robust results for face indexing. Additional gain in speedup was achieved with the ePLSH, a method that employs PLS-based feature selection to reduce the computational cost to evaluate hash functions, enabling a large amount of additional hash functions to be employed and raising the indexing precision. We evaluated several parameters and alternative implementations of PLSH in the hope that they will be useful for future face indexing development. The experiments were conducted on two face identification standard datasets, FERET and FRGCv1, with 1, 196 and 275 subjects, respectively. Although these datasets do not provide enough number of subjects for a proper evaluation regarding scalability to large galleries, PLSH and ePLSH still provide significant improvement in speedup compared to other scalable face identification approaches in the literature.

The conclusions and considerations regarding PLSH and ePLSH are the following: (i) they support for high dimensional feature vectors, allowing different complementary feature descriptors to be employed to increase the robustness of the face indexing; (ii) they are easy to implement and deploy in practice since the only parameters needed to be set are the number of hash models and subjects in the candidate list. (iii) they do not provide good performances when the number of samples per subject is reduced and (iv)

incremental enrollment of subjects in the framework requires re-training of the hash models, which may be prohibitive to perform in practice, specially for ePLSH which demands considerable more hash models.

Acknowledgments

The authors would like to thank the Brazilian National Research Council – CNPq (Grant #477457/2013-4), the Minas Gerais Research Foundation – FAPEMIG (Grant APQ-00567-14) and the Coordination for the Improvement of Higher Education Personnel – CAPES (DeepEyes Project).

References

- Ahonen, T., Hadid, A., and Pietikainen, M. (2006). Face description with local binary patterns: Application to face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(12):2037–2041.
- C E Santos Jr, E Kijak, G. G. W. R. S. (2016). Partial least squares for face hashing. *Elsevier Neurocomputing – Special Issue on Binary Representation Learning in Computer Vision*, pages 1–44.
- Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society Conference on*, volume 1, pages 886–893.
- Jr, C. E. S., Kijak, E., Gravier, G., and Schwartz, W. R. (2015). Learning to hash faces using large feature vectors. In *Content-Based Multimedia Indexing (CBMI), 13th IEEE Workshop on*, pages 1–6.
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110.
- Phillips, P. J., Flynn, P. J., Scruggs, T., Bowyer, K. W., Chang, J., Hoffman, K., Marques, J., Min, J., and Worek, W. (2005). Overview of the face recognition grand challenge. In *Computer vision and pattern recognition (CVPR). IEEE Conference on*, volume 1, pages 947–954.
- Phillips, P. J., Moon, H., Rizvi, S. A., and Rauss, P. J. (2000). The FERET evaluation methodology for face-recognition algorithms. *Pattern Analysis and Machine Intelligence*, 22(10):1090–1104.
- Randen, T. and Husoy, J. H. (1999). Filtering for texture classification: A comparative study. *Pattern Analysis and Machine Intelligence*, 21(4):291–310.
- Schwartz, W. R., Guo, H., Choi, J., and Davis, L. S. (2012). Face identification using large feature sets. *Image Processing, IEEE Transactions on*, 21(4):2245–2255.
- Wang, H., Li, S. Z., and Wang, Y. (2004). Face recognition under varying lighting conditions using self quotient image. In *Automatic Face and Gesture Recognition*, pages 819–824.
- Yuan, Q., Thangali, A., and Sclaroff, S. (2005). Face identification by a cascade of rejection classifiers. In *Computer Vision and Pattern Recognition (CVPR) - Workshops. IEEE Conference on*, pages 152–152.