A Binarization Scheme for Face Recognition based on Multi-scale Block Local Binary Patterns

Torsten Schlett, Christian Rathgeb and Christoph Busch[1]

Abstract: Local binary patterns (LBP) represent a well-established technique for reliable facial recognition. Multi-scale block (MB) LBP, which process average pixel values of block sub-regions instead of single pixels, have been proposed in order to achieve more robust feature vectors, which consist of histograms of (MB)-LBP values. In this work, we present a simple, yet effective, method to binarize histograms obtained from a MB-LBP face recognition system. The proposed method extracts compact binary feature vectors which allow for a rapid comparison and which are required for privacy enhanced biometric processing, when biometric template protection is enabled. On the publicly available FERET and Extended-Yale-B dataset, the presented scheme reveals only a negligible drop in biometric performance compared to the original MB-LBP system, while a substantial speed-up is achieved in the comparison stage.

Keywords: Face recognition, binarization, local binary patterns.

1 Introduction

Face recognition represents a longstanding field of research and a variety of methods have been proposed over the past three decades [Zh03, LJ11]. Generic face recognition systems comprise four major modules: face detection, face alignment, or feature extraction, and comparison, where the latter two are generally conceded as key modules. On the one hand, feature extractors are designed to generate discriminative feature vectors. On the other hand, comparators are designed to perform an efficient comparison, which is of utmost importance in identification systems. Feature descriptors such as LBP [AHP06] and its variants have been widely used in numerous face recognition schemes due to their robustness and strong discriminative power. Generally speaking, extracted LBP-values are stored in sequences of histograms, which are compared using some kind of histogram-based similarity measure, such as Chi-square distance. In contrast, a binary representation of facial feature vectors appears more appealing, as it enables a more efficient comparison based on intrinsic bit operations. In past years several approaches have been introduced to obtain binary feature vectors from facial images. Mostly, these schemes have been designed for biometric template protection schemes [RU11], where so-called *feature-type transformations* are essential, since distinct cryptographic primitives can only operate on a certain form of input [LTK15].

Diverse quantisation (or bit-allocation) schemes, which divide the feature space into segments and assign certain bits to them have been presented. Most notably, Chen *et al.*

[1] da/sec – Biometrics Research Group, Hochschule Darmstadt, Germany, torsten.schlett@stud.h-da.de, {christian.rathgeb,christoph.busch}@h-da.de

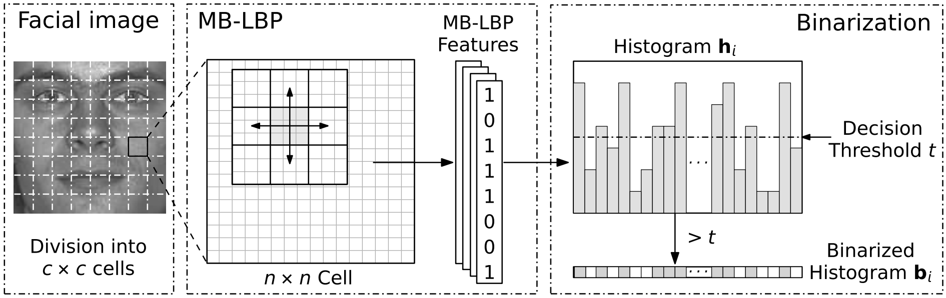


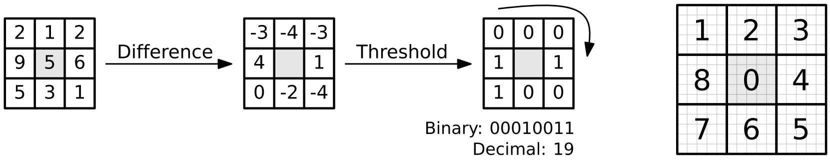
Fig. 1: System overview: MP-LBP values are extracted per cell and mapped to histograms, which are binarized based on a threshold.

[Ch09] present a quantisation system, which is referred to as detection rate optimized bit allocation (DROBA), where the detection rate is the probability of a genuine feature element staying within a distinct segment. The detection rate is maximized in order to minimize the false non-match rate (FNMR). This is achieved by an iterative segmentation of the feature space, where feature elements are weighted by the number of allocated bits. DROBA assumes negligible dependency between feature elements. In contrast, Chen and Veldhuis [CV11] propose another system in which feature elements are paired prior to the binarization process. Lim *et al.* [LTT11] present a reliability-dependent bit allocation (RDBA) system. Different from DROBA, RDBA uses a bottom-up segment-merging approach based on the reliability of binary-encoded training samples. The binary encoding is performed employing a binary reflected Gray code.

Despite these quantisation schemes, which require a (subject-specific) storage of feature space segments and their corresponding bit encoding, other direct encoding schemes have been proposed. These schemes are designed to transform integer values to a binary representation [LTK15]. In general, when there are more than two possible integer labels, the distance between feature elements in the discrete domain cannot be preserved exactly in the Hamming domain, which implies a certain loss of information. Due to such reasons, the vast majority of binarization schemes has been found to cause a significant drop in recognition accuracy.

In this paper we tackle the aforementioned issue and propose a new binarization scheme for LBP-based biometric recognition schemes. The presented scheme binarizes bins of histograms extracted using MB-LBP [Li07], which represents an extension to the basic LBP operator. By employing a static or dynamic threshold, a single bit is extracted from each histogram bin. This trivial technique is shown to obtain compact binary feature vectors, which possess enough discriminative power to almost preserve biometric performance. Moreover, the comparison stage, which is based on intrinsic bit operations, is significantly accelerated.

The remainder of this paper is organized as follows: in Sect. 2 the proposed system is described in detail. Experimental results are presented in Sect. 3. Conclusions are drawn in Sect. 4.



(a) basic LBP (b) MB-LBP

Fig. 2: The basic LBP operator and and example of the MB-LBP operator with $k = 2$.

(a) Original (b) $k = 8/512*n$ (c) $k = 10/512*n$ (d) $k = 12/512*n$

Fig. 3: Examples of MB-LBP extraction for a cropped sample face image.

2 Proposed System

In the following subsections constituting components of the proposed system, which are depicted in Fig. 1, are described in detail.

2.1 Multi-scale Block Local Binary Patterns

A quadratic cropped facial image of $n \times n$ pixels is first divided into $c \times c$ cells. For each pixel in a cell the basic LBP operator processes its 3×3 pixel neighborhood, denoted by (8,1). All eight neighboring pixels are traversed circularly, thresholded by the currently processed (center) pixel value, multiplied by powers of two and then summed up to obtain a single value, i.e. pattern. An example of this process is illustrated in Fig. 2 (a). Subsequently, LBP values are transformed to their uniform representation. Of the $2^8 = 256$ possible patterns a total number of 58 uniform patterns is obtained, which yields 59 different values. A local binary pattern is called uniform if the binary pattern contains at most two bitwise transitions from 0 to 1 or vice versa. Ojala *et al.* [OPM02] noticed in their experiments with texture images that uniform patterns account for a little less than 90% of all patterns when using the (8,1) neighborhood.

Liao *et al.* [Li07] suggested to divide cells into $(2k+1) \times (2k+1)$ blocks and extract MB-LBP values from the (8,1) neighborhood of the average pixel values of each block. An example is shown in Fig. 2 (b). Obviously, k has to be chosen according to the resolution of the facial image. Fig. 3 illustrates the extraction of MB-LBP values from a single facial image for different values of k in relation to n. As can be observed, more details are lost for increasing values of k, while extracted MB-LBP values are expected to become more robust. Depending on the scaling process, similar results might also be achieved by adapting the scale of the image to the basic LBP operator.

All patterns obtained from a cell are stored in a histogram $\mathbf{h}_i$ consisting of 59 bins, $i = 1,\ldots,c\times c$, such that the biometric template is composed of a sequence of normalized histograms, $\mathbf{H}$=$(\mathbf{h}_1,\mathbf{h}_2,\ldots,\mathbf{h}_{c\times c})$. Given a pair of histogram sequences, $\mathbf{H}$ and $\mathbf{H}'$, a dissimilarity score between them can be estimated by calculating the χ^2-distance between corresponding normalized histogram bins,

$$\chi^2(\mathbf{H},\mathbf{H}') = 1/2c^2 \sum_{i=1}^{c\times c}\sum_{j=1}^{59}(\mathbf{h}_{ij} - \mathbf{h}'_{ij})^2/(\mathbf{h}_{ij} + \mathbf{h}'_{ij}). \tag{1}$$

Obviously, the above baseline comparator, which operates on real values, is relatively demanding. Hence, in an identification scenario such a comparator could represent a bottleneck of a biometric system.

2.2 Binarization and Comparison

Given $\mathbf{H}$, a binary feature vector $\mathbf{B}$=$(\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_{c\times c})$ is obtained by successively comparing each histogram bin against a pre-defined threshold t,

$$\mathbf{b}_{ij} = \begin{cases} 1, & \text{if } \mathbf{h}_{ij} > t \\ 0, & \text{otherwise,} \end{cases} \quad \text{with } i = 1,\ldots,c\times c, \text{ and } j = 1,\ldots,59. \tag{2}$$

Hence, most dominant bin values of each histogram are encoded with 1, where an adequate value of t has to be selected in relation to other system parameters. Hence, the binarized feature vector consists of a total number of $c\times c\times 59$ bits. Alternatively, we suggest to dynamically employ the median bin value $\tilde{\mathbf{h}}_i$ of histogram $\mathbf{h}_i$ as decision threshold, i.e. $\mathbf{b}_{ij}$=1 in case $\mathbf{b}_{ij}$>$\tilde{\mathbf{h}}_i$ and $\mathbf{b}_{ij}$=0 otherwise.

Given two sequences of binarized histograms, $\mathbf{B}$ and $\mathbf{B}'$, a dissimilarity score between them is estimated as,

$$d(\mathbf{B},\mathbf{B}') = 1/c^2 \sum_{i=1}^{c\times c} 1 - \frac{2||\mathbf{b}_i \cap \mathbf{b}'_i||}{||\mathbf{b}_i|| + ||\mathbf{b}'_i||}. \tag{3}$$

That is, we calculate the number of agreeing 1s (most dominant histogram bins) in relation to the total number of 1s in corresponding binarized histograms within probe and reference feature vectors. The norm of a binary vector, denoted as $||\cdot||$, can be efficiently retrieved by intrinsic population count (`PopCnt`) functions. In case, $t = \tilde{\mathbf{h}}_i$ it might as well be possible to use a static divisor as it is expected to be equal to the number of bins, which would allow for a slightly more efficient comparison, $d'(\mathbf{B},\mathbf{B}') = 1/c^2 \sum_{i=1}^{c\times c} 1 - ||\mathbf{b}_i \cap \mathbf{b}'_i||/59$. However, the goodness of this approximation will certainly depend on the size of processed cells as well as the distribution of extracted feature values.

3 Experiments

In the following subsections we describe the experimental setup, report obtained biometric performance as well as time measurements.

Tab. 1: Overview of both employed face datasets.

Name	Subjects	Images	Resolution	Genuine comp.	Imposter comp.
FERET	993	2,707	512×768	3,620	492,528
Extended-Yale-B	38	719	168×192	6,444	703

(a) FERET

(b) Extended-Yale-B

Fig. 4: Three sample image pairs of both employed face datasets.

3.1 Experimental Setup

Experiments are conducted on the publicly available FERET and cropped Extended-Yale-B face database. An overview of both datasets is given in Table 1. For the FERET dataset images for which the face detection failed have been excluded. Biometric performance is reported in terms of genuine match rate at a false match rate of 0.1% ($\text{GMR}_{0.1}$) and equal error rate (EER). We perform all possible genuine comparisons and all impostor comparisons for the first image of each subject[3]. Sample images of both databases are depicted in Fig. 4. Time measurements are performed on a single core of an Intel Core i7-5820K CPU with 3.9 GHz (overclocked) on a standard Windows work station with sufficient RAM, where the process is executed with real-time priority.

3.2 Performance Evaluation

Obtained EERs and $\text{GMR}_{0.1}$ for reasonable configurations on the FERET and the Extended-Yale-B datasets are summarized in Table 2 and Table 3, respectively. Note that $\text{GMR}_{0.1}$ values reported on the Extended-Yale-B dataset are interpolated due to the smaller number of impostor verification attempts. Across all considered configurations, the proposed system almost maintains biometric performance causing drops in accuracy below one percent points in terms of EER. Interestingly, $t = 0$ appears to be a good default parameter choice, which might only apply in case of small cell sizes. Moreover, the median-based binarization (per histogram) turns out to be an adequate dynamic threshold across considered configurations, which is especially appealing. Receiver operation characteristic (ROC) curves

[3] For simplification we did not follow available evaluation protocols for subsets of these datasets, resulting in an even more challenging scenario.

Tab. 2: Biometric performance in terms of EER (%) and $GMR_{0.1}$ (%) on the FERET database of the original and the proposed system for best parameter configurations.

MB-LBP	Cells	Original	Proposed		
		χ^2	$t = 0$	$t = 0.1$	$t = \tilde{\mathbf{h}}_i$
$k = 8/512 * n$	16×16	13.06 / 58.24	13.34 / 56.75	13.47 / 56.54	13.36 / 56.74
$k = 8/512 * n$	32×32	13.77 / 55.08	14.05 / 54.24	14.00 / 53.96	14.05 / 54.24
$k = 10/512 * n$	16×16	13.67 / 55.44	14.07 / 54.53	14.22 / 53.91	14.05 / 54.53
$k = 10/512 * n$	32×32	14.20 / 53.19	14.44 / 52.33	14.40 / 52.12	14.44 / 52.33
$k = 12/512 * n$	16×16	14.55 / 52.87	14.81 / 53.03	15.07 / 51.88	14.81 / 53.03
$k = 12/512 * n$	32×32	14.86 / 50.53	14.99 / 50.60	15.18 / 50.13	14.99 / 50.60

Tab. 3: Biometric performance in terms of EER (%) and $GMR_{0.1}$ (%) on the Extended-Yale-B database of the original and the proposed system for best parameter configurations.

MB-LBP	Cells	Original	Proposed		
		χ^2	$t = 0$	$t = 0.1$	$t = \tilde{\mathbf{h}}_i$
$k = 8/512 * n$	16×16	16.92 / 62.30	17.10 / 60.65	16.62 / 58.99	17.10 / 60.65
$k = 8/512 * n$	32×32	13.92 / 69.19	14.79 / 67.67	14.28 / 67.20	14.79 / 67.67
$k = 10/512 * n$	16×16	19.90 / 56.55	19.73 / 58.96	19.53 / 53.06	19.73 / 58.96
$k = 10/512 * n$	32×32	17.09 / 62.29	17.91 / 60.98	17.21 / 61.64	17.91 / 60.98
$k = 12/512 * n$	16×16	23.32 / 50.82	23.02 / 49.51	23.04 / 51.41	23.04 / 51.41
$k = 12/512 * n$	32×32	19.77 / 55.17	20.19 / 54.24	20.34 / 54.18	20.19 / 54.24

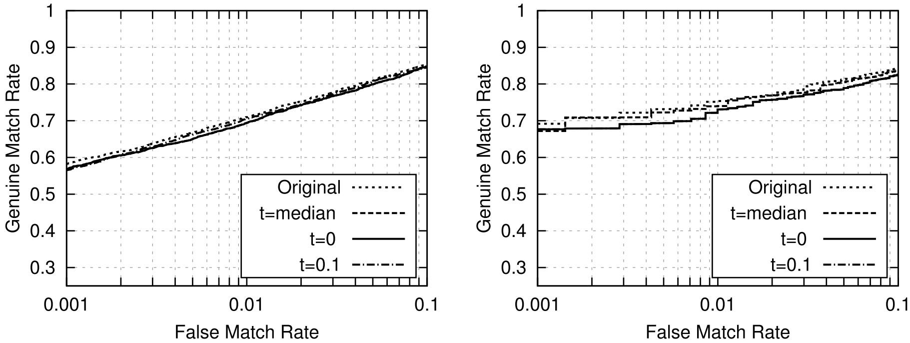
for the best configurations are plotted in Fig. 5. It can be further observed, that the binarized systems reveal almost identical behavior across relevant operation points, which further confirms the soundness of the proposed scheme.

3.3 Time Measurement

Time measurements are done by performing the above verification experiment until a total number of 200,000 comparisons is achieved. This is done in 15 loops and the median and mean times are reported for using a single CPU core in Table 4. Compared to the conventional χ^2-based system, the Hamming distance-based comparator achieves an approximately 20-fold speed-up. Note that, in the proposed system a single histogram can be stored in a single 64-bit integer. Note that, time measurements are independent of chosen block sizes.

4 Conclusions

Binary biometric feature vectors enable a compact storage and rapid comparison. In this work, we presented a trivial, yet effective, method for binarizing feature histograms extracted from a MB-LBP analysis of facial images. At a negligible drop of biometric performance, compact binary feature vectors of approximately 16,000 bits are obtained. By



(a) FERET, $c = 16$ (b) Extended-Yale-B, $c = 32$

Fig. 5: ROC curves for the best configurations on both databases using $k = 8/512 * n$.

Tab. 4: Median time measurements (seconds) for the original and the proposed schemes.

Sytem	Cells	Median	Mean/Std. Dev.
Original	16×16	14.1	14.1 $\pm$0.028
	32×32	43.9	43.9 $\pm$0.094
Proposed	16×16	0.58	0.58 $\pm$0.003
	32×32	2.22	2.22 $\pm$0.006

employing intrinsic functions of the CPU, identification systems can be efficiently operated on low-cost commodity hardware or even mobile devices.

Moreover, binarized feature vectors could serve as input of template protection systems, e.g. the fuzzy commitment scheme [JW99] or the recently proposed Bloom filter template protection scheme [RBB13]. In contrast to many proposed feature-type transformation schemes, the presented scheme enables the extraction of a binary feature vector from a single facial image.

Acknowledgements

This work was partially supported by the German Federal Ministry of Education and Research (BMBF) within the Center for Research in Security and Privacy (CRISP).

References

[AHP06] Ahonen, T.; Hadid, A.; Pietikainen, M.: Face description with local binary patterns: Application to face recognition. IEEE Trans. Pattern Anal. Mach. Intell., 28(12):2037–2041, 2006.

[Ch09] Chen, C.; Veldhuis, R.; Kevenaar, T.; Akkermans, A.: Biometric quantization through detection rate optimized bit allocation. EURASIP J. Adv. Signal Process., 2009:1–16, 2009.

[CV11] Chen, C.; Veldhuis, R.: Binary biometric representation through pairwise adaptive phase quantization. EURASIP J. Inform. Security, 2011:1–16, 2011.

[JW99] Juels, A.; Wattenberg, M.: A fuzzy commitment scheme. In: Proc. 6th ACM Conf. on Computer and Communications Security. ACM, pp. 28–36, 1999.

[Li07] Liao, S.; Zhu, X.; Lei, Z.; Zhang, L.; Li, S. Z.: Learning Multi-scale Block Local Binary Patterns for Face Recognition. In: Proc. Int'l Conf. on Biometrics (ICB'07). pp. 828–837, 2007.

[LJ11] Li, S. Z.; Jain, A. K.: Handbook of Face Recognition (2nd edition). Springer, 2011.

[LTK15] Lim, M.-H.; Teoh, A. B. J.; Kim, J.: Biometric feature-type transformation: Making templates compatible for template protection. IEEE Signal Process. Mag., 32(5):1–11, 2015.

[LTT11] Lim, M.-H.; Teoh, A. B. J.; Toh, K.-A.: An efficient dynamic reliability-dependent bit allocation for biometric discretization. Pattern Recognit., 45(5):1960–1971, 2011.

[OPM02] Ojala, T.; Pietikainen, M.; Maenpaa, M.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell., 24(7):971–987, 2002.

[RBB13] Rathgeb, C.; Breitinger, F.; Busch, C.: Alignment-free cancelable iris biometric templates based on adaptive bloom filters. In: Int'l Conf. on Biometrics (ICB'13). pp. 1–8, 2013.

[RU11] Rathgeb, C.; Uhl, A.: A Survey on Biometric Cryptosystems and Cancelable Biometrics. EURASIP J. on Inf. Sec., 2011, 2011.

[Zh03] Zhao, W.; Chellappa, R.; Phillips, P. J.; Rosenfeld, A.: Face recognition: A literature survey. ACM Comput. Surveys, 35(4):399–458, 2003.