A Complementary Local Feature Descriptor for Face Identification

Jonghyun Choi[†], William Robson Schwartz[‡], Huimin Guo[†] and Larry S. Davis[†]

[†]Institute for Advanced Computer Studies
University of Maryland, College Park
MD, USA

{jhchoi,hmguo,lsd}@umiacs.umd.edu

[‡]Institute of Computing
University of Campinas
SP, Brazil

schwartz@ic.unicamp.br

Abstract

In many descriptors, spatial intensity transforms are often packed into a histogram or encoded into binary strings to be insensitive to local misalignment and compact. Discriminative information, however, might be lost during the process as a trade-off. To capture the lost pixel-wise local information, we propose a new feature descriptor, Circular Center Symmetric-Pairs of Pixels (CCS-POP). It concatenates the symmetric pixel differences centered at a pixel position along various orientations with various radii; it is a generalized form of Local Binary Patterns, its variants and Pairs-of-Pixels (POP). Combining CCS-POP with existing descriptors achieves better face identification performance on FRGC Ver. 1.0 and FERET datasets compared to state-of-the-art approaches.

1. Introduction

In many computer vision algorithms, even ones that use sophisticated machine learning techniques, the low level description of an image is still critical to performance [23]. For this reason, it is important to design good low level feature descriptors; however, this is not a trivial task since it is not clear what visual information is important when we perceive an object or a scene [28].

The existing descriptors, such as Histograms of Oriented Gradients (HOG) [5] and Local Binary Pattern (LBP) [2] capture block-wise information, while Gabor wavelets [26] capture global shape information centered at a pixel. However, they can miss pixel-wise local information which might contain discriminative information.

In [9], a descriptor called Pairs of Pixels (POP) was proposed for detecting vehicles by a simple intensity transform along the principal axes of a vehicle. For man-made objects such as vehicles, POP is sufficient since discriminative information is mostly aligned with the principal directions of the vehicle (consider a car - along the main axes there are

distinctive changes as one moves from metallic to glass regions). For each pixel in a detection window, POP measures the integrated differences in brightness (color) between that pixel and all other pixels along the two principal directions. If the detection window contains n pixels, then POP would give rise to $n\sqrt{n}$ distinct features, approximately.

In general, however, natural objects such as faces have spatially correlated information along many directions. Thus, many of the successful descriptors for faces capture spatial information in all directions, *e.g.*, Gabor wavelets and LBP and its variants.

Generalizing the POP descriptor to face recognition, we propose a descriptor called Center Symmetric-Pairs of Pixels (CCS-POP) which efficiently compares pairs of pixels in several directions with a center symmetric scheme proposed in [7]. The definition of the CCS-POP feature vector for a pixel (x, y) is

$$V(x, y) = (d_{(x,y)}(r_1, a_1), \cdots, d_{(x,y)}(r_m, a_n)), \quad (1)$$

where $(\cdot)$ is a tuple notation and $d_{(x,y)}(r_k, a_i)$ denotes the symmetric pixel intensity difference centered at pixel (x, y) for the kth radius r_k ($k = \{1, \cdots, m\}$) along the ith angular direction a_i ($i = \{1, \cdots, n\}$) as described below.

$$
\begin{aligned}
d_{(x,y)}(r_k, a_i) = \\
I(x + r_k \cos a_i, \ y + r_k \sin a_i) \\
- I(x + r_k \cos(a_i + \pi), \ y + r_k \sin(a_i + \pi)),
\end{aligned}
\quad (2)
$$

where $I(x, y)$ denotes the intensity value of pixel (x, y) in a specific color channel. The detailed description of CCS-POP is presented with an example in Section 4.

Since computing the CCS-POP for every pixel in the face results in a large number of descriptors, we use a supervised dimensionality reduction technique called Partial Least Squares (PLS) [32] to find a much smaller number of discriminative factors without histogramming.

Experimental results show that CCS-POP further improves face identification performance over the state-of-the-

art methods in FRGC Ver. 1.0 [21] and FERET [22] dataset when combined with other descriptors.

2. Related Work

The literature on low level descriptor design contains many proposals for various applications [2, 4, 5, 14, 18, 23, 26]. Among them, we selectively review the most relevant and successful ones for face identification, the application considered in this paper.

The local binary patterns (LBP) is a well-known texture descriptor and a successful local descriptor for face recognition under local illumination variations [2, 7]. LBP descriptors are compact and easy to compare by various histogram metrics. In addition, there are many LBP variants to improve the description performance of LBP. The most popular extension is multi-scale LBP (MSLBP) which uses multiple radii [19].

The histogram of oriented gradients (HOG) has been successfully applied to tasks such as human detection [5] and face recognition [24]. Similar to LBP, edge information captured by gradients within blocks is packed into a histogram. Discarding pixel location information by block-based histogram binning, LBP and HOG gain invariance to local changes such as small facial expressions and pose variations in pedestrian images.

The Gabor wavelets are also successful face descriptors which capture global shape information centered at a pixel [26]. The convolution of multiple Gaussian-like kernels with different scales and orientations captures information insensitive to expression variation and blur at a pixel's location.

Recent research deals with parameter learning with a fixed template [4, 23] and especially learned weights on LBP information with illumination normalization (DLBP) [16]; our approach to using descriptors is similar in terms of learning discriminative weighting in a proposed template. Also, some successful LBP variants including Three-Patch based LBP (TPLBP), Four-Patch based LBP (FPLBP) [33], Patterns of Oriented Magnitude (POEM) [30] have been used recently for face recognition.

Using advanced feature weighting or selection methods such as PLS and compressive sensing, very simple features such as raw intensity values [25] or Pairs of Pixels (POP) [9] can be used, showing that the image intensity itself or simple second order intensity features can be very informative features [34].

3. Partial Least Squares for Feature Weighting

Partial Least Squares (PLS) is a supervised learning method that models linear relations between sets of independent variables and response variables via an intermediate latent space. It has been successfully applied to many

vision applications including face recognition [24, 25].

The main idea of the method is to maximize the covariance between the dependent variable Y and a weighted sum of the independent variables in X to find the most discriminative weighting vectors for each row of X in the training phase. Once it obtains the weighting vectors, w, it computes the regression vector. Using the regression vector, the response is obtained by a dot product of the regression vector and a target sample. Eq. (3) and Eq. (4) show the equations of the training phase.

Let $\mathcal{X} \subset \mathbb{R}^m$ denote an m-dimensional feature space, $\mathcal{Y} \subset \mathbb{R}$ denote a scalar space representing the response variable. PLS decomposes a $(n \times m)$ matrix $X \in \mathcal{X}$, where n denotes the number of samples, which is normalized by subtracting its average across the columns, and the $(n \times 1)$ vector $Y \in \mathcal{Y}$ into

$$
\begin{aligned}
X &= TP^T + E \\
Y &= UQ^T + F
\end{aligned}
\tag{3}
$$

where $(n \times p)$ matrices T and U are called *score* or *component* or *factor* which contain latent vectors, the $(m \times p)$ matrix P and the $(1 \times p)$ vector Q are called the *loadings* and the $(n \times m)$ matrix E and the $(n \times 1)$ vector F are called the *residuals*.

Using a greedy algorithm called NIPALS [32], we can obtain a set of weight vectors, stored in the matrix $W = (w_1, w_2, \ldots, w_p)$ iteratively, such that

$$
\text{cov}(t, u)^2 = \max_{|w|=1} \text{cov}(Xw, Y)^2
\tag{4}
$$

where t and u are the column vectors of matrices T and U in Eq. (3), respectively, and $\text{cov}(t, u)$ is the covariance matrix between latent vectors t and u. At the ith iteration, w_i is obtained in order, *e.g.*, at first iteration w is w_1 and so on. At the end of each iteration, the matrices X and Y are deflated by subtracting their rank-one approximations based on t and u and this is continued until the desired number of latent vectors is obtained, denoted by p.

4. Circular Center Symmetric Pairs of Pixels (CCS-POP)

CCS-POP captures information along various directions by computing pixel differences along angular axes with a center symmetric scheme which results in more stable angular-information by symmetry [19]. The captured information is similar to LBP but it maintains magnitude information, which was shown to be useful in [4, 9].

It concatenates the information without histogramming as is common with LBP and HOG. Since we are often able to geometrically warp face images to a canonical position by fiducial points (labeled by human or detected by off-the-shelf algorithms), the pixel-wise information in the re-

gion that is less affected by facial expression and local deformation, *e.g.*, ocular region or nose, would be complementarily useful for recognition without high sensitivity to mis-alignment. This is motivated by the fact that humans are good at recognizing people even with only subtle differences in their ocular regions. We empirically show the advantage of not using histogramming in Sec. 5.1.3 and Sec. 5.2.2.

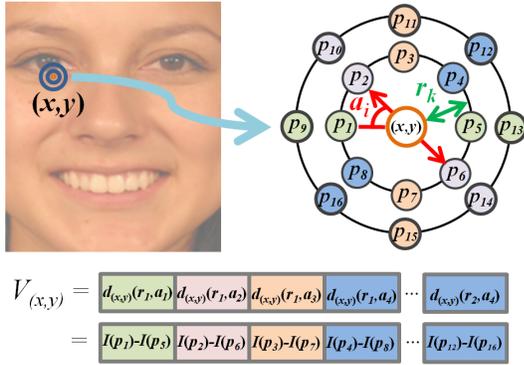Fig. 1 depicts how the CCS-POP descriptor is computed through an example.[1]

Figure 1. An example of CCS-POP descriptor for pixel (x, y).

The center point in the right diagram in Fig. 1 represents a pixel in an image. At each pixel, we compute the difference $d_{(x,y)}(r_k, a_i)$ of a pair of symmetric pixels distant r_k away from the center and oriented at the angle a_i with respect to the horizontal axis, *e.g.*, $d_{(x,y)}(1, -45°) = I(p_2) - I(p_6)$ where $I(\cdot)$ denotes the intensity value of a pixel. We then concatenate these differences over a range of distances and directions, which is used as the feature vector. In the example of Fig. 1, two radii are used, where the first radius is r_k, and the first angle a_i is $-45°$ and the remaining the angles are equally sampled by $45°$ so that the total number of CCS-POP values for pixel (x, y) is 8.

4.1. Parameters

CCS-POP can be adapted to different imaging conditions or applications by controlling its parameter configurations. According to the level of blur of the input data, we can choose the size and numbers of radii, for example, we would use larger radii for more blurred images to capture more global information as in [37].

Truncation Threshold In addition, to make CCS-POP robust to extreme local illumination variations, truncation with a threshold is used to prevent unstable information violating a smoothness assumption [29]. We denote the trun-

---
[1]Code is available in http://umiacs.umd.edu/~jhchoi/ccspop

cating threshold as T_t. When employed, the values in the vector are truncated by Eq. (5).

$$
d_{(x,y)}(r_k, a_i)
= \begin{cases} d_{(x,y)}(r_k, a_i), & |d_{(x,y)}(r_k, a_i)| < T_t, \\ \mathrm{sgn}(d_{(x,y)}(r_k, a_i)) \cdot T_t, & |d_{(x,y)}(r_k, a_i)| \geq T_t, \end{cases} \quad (5)
$$

where $\mathrm{sgn}(\cdot)$ denotes the sign function.

5. Experimental Results and Discussion

We evaluate the performance of CCS-POP with PLS weighting on two face identification datasets, FRGC Ver. 1.0 [21] and FERET [22], comparing the results to previous approaches. To obtain the PLS weighting vector, we use the one-against-all PLS regression approach which works well even when there is only a single sample per subject in the gallery [24].

We also compare the performance of CCS-POP to other existing descriptors, discuss their characteristics and show that the combination of CCS-POP with others improves face identification results in both datasets.

5.1. FRGC Ver. 1.0

In FRGC Ver. 1.0, there are three experiments for 2D images: Exp.1, 2 and 4. We follow the published protocols [21]. Experiment 1 consists of a gallery with a single controlled image and a probe with one controlled still image per subject (183 training images, 152 gallery images, and 608 probe images). Experiment 2 consists of a gallery with four controlled still images per subject (732 training images, 608 gallery images, and 2,432 probe images). Finally, experiment 4 consists of a gallery with a single controlled still image per subject and multiple uncontrolled probe images per subject (366 training images, 152 gallery images, and 608 probe images). Exp. 4 is the most challenging since it involves matching a controlled face to uncontrolled faces with extreme illumination variations, blur and expression variations.

5.1.1 Experimental Settings

Preparation The images are cropped and resized to 138 $\times$ 160 pixels and are aligned with the annotated eye positions. We do not use any pre-processing for illumination normalization.

Parameters For the parameters of CCS-POP, we use two radii, $r_1 = 2$ and $r_2 = 3$. We skip every other pixel to avoid the redundancy of information since the radii configuration is 2 and 3 so that the size of the feature vector is 44,160.

The effects of the truncation threshold are shown by rank-1 face identification performance in Table 1. As shown

in the table, a truncation threshold of $T_t = 5$ yields performance which is 10% better than without the threshold in Exp. 4. Thus we choose $T_t = 5$.

Table 1. Comparison of rank 1 face identification rate (%) of CCS-POP (weighted by PLS regression) with and without T_t on FRGC 1.0 dataset.

T_t	Exp.1	Exp.2	Exp.4
No threshold	91.6	97.2	57.7
5	96.7	99.3	67.6

The block sizes of HOG are 32×32 and 16×16 pixels with strides 8 and 4, respectively. The block size of LBP is 32×32 pixels with a stride of 16. For multi-scale LBP, we use radii 2 and 3. For Gabor wavelets, we use 8 orientations ($45°$ for each) and 5 scales with 4 pixels. These are the best performing configurations obtained from many trials.

5.1.2 Color Information

Since the FRGC Ver. 1.0 dataset contains color images, we try to use color information by applying CCS-POP on the R,G and B color channels separately. Although the color information helps to improve performance as shown in Table 2, it leads to a very high dimensional features (Refer to second row of the table). To incorporate color information without such a large increase in dimensionality, we use color channels which are less sensitive to illumination variations as described below and add them to CCS-POP descriptor for the experiments on the FRGC Ver. 1.0 dataset.

Table 2. Comparison of rank 1 face identification rate (%) of CCS-POP (weighted by PLS regression) with color information by different schemes on FRGC 1.0 dataset. Dim. refers to number of dimension of the descriptors. *CI* refers to the additional color information by taking mean on new color channel.

Color	Dim.	Exp.1	Exp.2	Exp.4
Gray	44,160	96.1	99.0	59.4
R,G,B	132,480	96.5	99.4	67.3
Gray+*CI*	60,042	96.7	99.3	67.6

To obtain a color channel which is less sensitive to illumination variations, we consider the color constancy measures which are invariant to illumination variations [6, 11]. *Chromaticity* [6] gives good color constancy under illumination variations by taking the color ratios $\frac{R}{R+G+B}$, $\frac{G}{R+G+B}$ and $\frac{B}{R+G+B}$.

Given the fact that blue channel is usually darker than other color channels so that it is noisier than others [15], to make a less noisy color constancy measure, we drop the

blue channel and extract the mean of small patches around a pixel on the remaining two channels, $\frac{R}{R+G}$, $\frac{G}{R+G}$, as well as standard gray scale intensity, $(0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B)$ since noise affects the intensity value under low light levels [3].

5.1.3 Advantage of not Using Histogram

Typically, when pixel-based features, such as LBP or gradient information, are used for face or object recognition, the features are accumulated over small regions into histograms, and recognition is then conducted based on the histogram features. There are two reasons for this. First, the compression to histograms reduces the dimensionality of the feature set. Second, the histograms are invariant to some local deformations of the underlying patterns.

However, given the availability of dimensionality reduction methods like PLS, it is worth investigating whether it is really advantageous to compress LBP-like features using histograms. Table 3 shows the comparative performance of raw LBP information (denoted by Raw LBP) without regional histogram binning and the original LBP features that are reduced to histograms, both reduced to a small number of factors by PLS for recognition on FRGC ver. 1.0. As shown in the table, raw LBP significantly outperforms the regionally histogrammed LBP, especially in the most challenging experiment scenario, Exp. 4.

Table 3. Comparison of rank 1 face identification rate (%) of Raw LBP to original (regionally histrogrammed) LBP weighted by PLS (17 factors) on FRGC 1.0 dataset. For histogram regions, 32×32 sized blocks with a stride 16 are used. Dim. refers to number of dimension of the each descriptor. Raw LBP means LBP information without encoding and histogram binning.

Descriptor	Dim.	Exp.1	Exp.2	Exp.4
LBP	16,128	85.4	96.8	13.0
Raw LBP	176,640	94.4	98.7	59.9

5.1.4 Comparison to Other Descriptors

We compare the performance of rank 1 face identification of CCS-POP to LBP, its multi-scale version (MSLBP) and POP. We also compare the proposed descriptors to Gabor wavelets and HOG, which are useful descriptors for face identification, as well as the raw image intensity. All descriptors are weighted by PLS regression for comparison. For PLS factors, we tried 4 numbers of factors (15,17,19 and 21) and choose the best performing one for each descriptor. The comparative performance is summarized in Table 4.
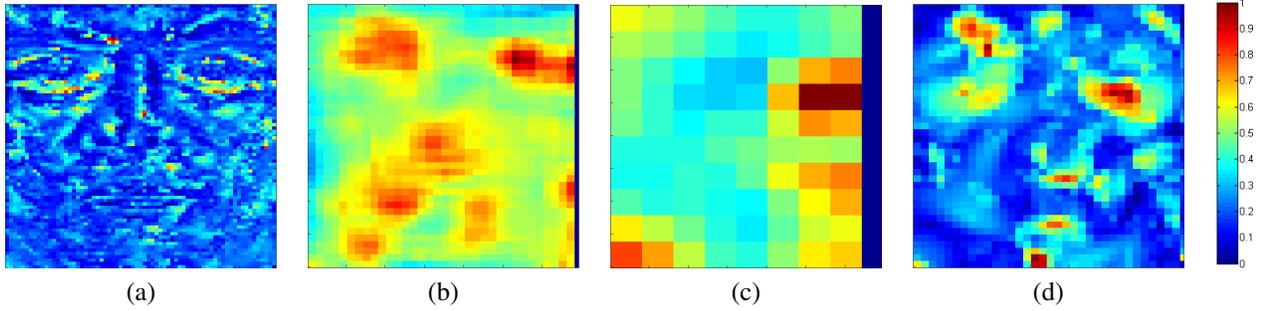
Figure 2. An example of heat maps of PLS regression for different features. (a) CCS-POP. (b) Histogram of Oriented Gradient (HOG). (c) Local Binary Pattern (LBP). (d) Gabor Wavelets. CCS-POP captures pixel-wise micro information.

Table 5. Comparison of rank 1 face identification rate (%) of our descriptor combined with other descriptors with PLS regression on FRGC 1.0 dataset.

Feature Descriptor	Dim.	Exp.1	Exp.2	Exp.4
LBP + CCS-POP	76,407	96.5	99.4	68.4
MSLBP + CCS-POP	92,535	96.5	99.4	69.2
HOG + CCS-POP	110,139	97.9	99.8	79.3
Gabor + CCS-POP	114,679	97.2	99.8	82.4
HOG + Gabor	104,260	97.2	99.6	82.9
HOG + Gabor + CCS-POP	168,690	**98.0**	**99.8**	**89.0**

Table 4. Comparison of rank 1 face identification rate (%) of our descriptor to other descriptors with PLS regression on FRGC 1.0 dataset. Dim. refers to number of dimension of the descriptors

Feature Descriptor	Dim.	Exp.1	Exp.2	Exp.4
Baseline (Intensity)	22,080	41.3	51.3	3.5
POP	607,520	83.2	94.0	29.3
LBP	16,128	85.4	96.8	13.0
MSLBP	32,256	95.2	98.9	30.9
HOG	49,860	97.5	99.4	64.5
Gabor	54,401	97.0	99.5	66.6
CCS-POP	60,042	96.7	99.3	67.6

The original POP produces very high dimensional feature, thus it is intractable to use it for all pixel pairs in a 138 × 160 image. So, we compute POP every 10 pixels, which still gives very high dimension (607,520). Even though the dimension is high, it results in poor face identification performance as shown in the result of Exp.4 in Table 4.

HOG shows fairly good performance especially in Exp.4 which has uncontrolled image acquisition conditions. This results is with the block configuration of 32×32 and 16×16 sized blocks with strides 16 and 8, which we found to be the best performing one among many trials.

Gabor wavelets capture shape information invariant to local variations and are well-known to be one of the best descriptors for face recognition [26]. As Table 4 shows, Gabor wavelets provide the best result among the existing descriptors.

CCS-POP slightly outperforms the remaining descriptors and has a similar number of features to Gabor.

5.1.5 Combinations with Other Descriptors

Since CCS-POP captures pixel-wise local edge information, after it is weighted by PLS we can visualize through heat maps the additional information it captures that is missed by HOG and Gabor which capture block-wise or image-wise information. Fig. 2 contains the heat maps of PLS weighting for the various feature channels. The fact that different parts of the face are given more weight by PLS for the different types of features suggests that combining the features for face recognition should lead to overall improved results. This is illustrated in Table 5.

As Fig. 2 shows, CCS-POP captures fine-grained pixel-wise information while HOG, Gabor and LBP capture overall information centered at pixels of each block. Thus, CCS-POP is able to to complement the information extracted by HOG and LBP by giving large weights to specific pixels, which the block-based nature prevents HOG and LBP from doing. In addition, although Gabor wavelets are also extracted for pixels, the heat map for it is smooth because of the large support of the filter (a whole image).

Since CCS-POP captures different information, combinations of other features with CCS-POP improves identifi-

cation performance especially on the most challenging experiment, Exp.4, by more than 15% comparing to the results of individual descriptors. Although combining HOG and Gabor results in good performance (Refer to fifth row of Table 5), the performance further improves by 6.1% by combining CCS-POP with HOG and Gabor.

5.1.6   Comparison to the Previous Approaches

Table 6 summarizes the performance comparison of our method to previous approaches on FRGC Ver. 1.0 dataset. CHG refers to combined descriptors of CCS-POP, HOG and Gabor with one-against-all PLS regression. CHG significantly increases the recognition rate relative to the state-of-the art by 10.8% on Exp.4.

Table 6. Comparison of rank 1 face identification rate (%) of our method and other algorithms on FRGC 1.0 dataset. **CHG** refers to the combined descriptor of CCS-POP, HOG and Gabor with PLS regression.

Method	Exp.1	Exp.2	Exp.4
PCA [17]	87.6	95.6	-
UMD [1]	94.2	99.3	-
BEE (from [27])	-	-	37.0
LC_1C_2 [27]	-	-	75.0
ROCA [10]	96.4	-	75.5
Liu [12]	-	-	78.0
Tan (from [8])	-	-	58.1
Holappa [8]	-	-	63.7
LPQ [20]	-	-	74.5
PLS [24]	97.5	99.4	78.2
CHG	**98.0**	**99.8**	**89.0**

Fig. 3 shows the cumulative matching characteristic (CMC) curve of our method on the FRGC 1.0 dataset.

5.2. FERET

The FERET dataset contains experimental protocols with partitions fa, fb, fc, $dup1$ and $dup2$. fa is a gallery set and the rest are probe sets. fb contains 1195 images taken with different expressions, fc contains 194 images taken under different lighting conditions, $dup1$ contains 722 images having aging and expression variations, and $dup2$ contains 234 images, having more serious aging and expression variations. Among the four probe sets, $dup1$ and $dup2$ are regarded as the most challenging ones since they have aging variations as well as expression and small illumination variations.
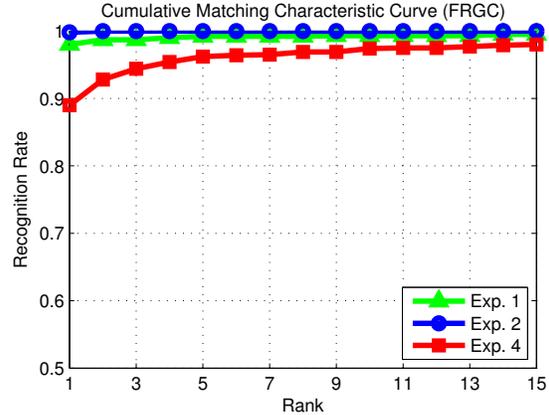

Figure 3. Cumulative matching characteristic (CMC) curve of **CHG** in Table 6 on FRGC Ver. 1.0 dataset

5.2.1   Experimental Settings

Preparation   We crop and resize the images to 110×110 and align them with the annotated eye positions. Before extracting HOG, Gabor and LBP features, we use the SQI (Self Quotient Image) pre-processing [31] as used in [24] to compare with them. We, however, use no pre-processing before computing CCS-POP.

Parameters   We use the same parameter configuration of CCS-POP as was used for FRGC Ver. 1.0 dataset; $r_1 = 2$, $r_2 = 3$, $T_t = 5$ and again skip every other pixels. Since the FERET dataset contains lots of facial expressions including wrinkles in the ocular region, to give a slight stability to expression and mis-alignment, we average the CCS-POP descriptor of a pixel with those of its immediate neighbors within a 4×4 pixel sized window centered at the pixel.

The block sizes of HOG are 32×32 and 16×16 pixels with strides 8 and 4, respectively. The block size of LBP is 32×32 pixels with a stride 16. For Gabor wavelets, we use 8 orientations ($45°$ for each) and 5 scales with stride 4 pixels.

5.2.2   Advantage of not Using Histogram

To determine if the raw LBP is better than the regionally histogrammed LBP in this dataset, we performed the same experiments of comparing raw LBP and regionally histogrammed LBP that was conducted on FRGC Ver. 1.0.

Table 7. Comparison of rank 1 face identification rate (%) of raw LBP to original LBP with PLS regression (factor 9) on FERET dataset. Convention is the same with FRGC experiments.

Descriptor	Dim.	fb	fc	$dup1$	$dup2$
LBP	6,400	57.7	26.8	18.3	10.3
Raw LBP	96,800	89.8	73.7	68.7	53.0

Table 7 shows the comparative performance of raw LBP and regionally histogrammed LBP on FERET. As shown in the table, the raw information without encoding and histogram binning in a block significantly outperforms histogrammed LBP in all protocols.

5.2.3 Comparisons to Other Descriptors

We also compare the performance of rank 1 face identification result of CCS-POP to LBP, its multi-scale version (MSLBP), POP, Gabor wavelets, HOG and intensity. All descriptors are weighted by PLS regression vector with 9 factors.

The comparative performance is summarized in Table 8. While CCS-POP outperforms LBP, MSLBP and POP in all experiments, it is less informative than HOG and Gabor on the FERET dataset.

Table 8. Comparison of rank 1 face identification rate (%) of our descriptor to other descriptors with PLS regression on FERET dataset.

Descriptor	Dim.	fb	fc	$dup1$	$dup2$
Intensity	12,100	75.1	52.6	39.5	31.2
LBP	6,400	57.7	26.8	18.3	10.3
MSLBP	12,800	92.8	29.4	56.5	49.1
POP	242,000	79.7	69.1	47.2	42.7
HOG	24,336	94.1	98.5	76.9	76.9
Gabor	29,187	97.0	97.9	79.4	76.1
CCS-POP	24,200	91.9	72.7	69.3	53.4

5.2.4 Combinations with Other Descriptors

However, combining CCS-POP with HOG and Gabor does result in better performance than the combination of HOG and Gabor. We summarize the results of combined descriptors with CCS-POP in Table 9.

5.2.5 Comparison to the Previous Approaches

Table 10 summarizes the results of the previous approaches and the combined descriptor of CCS-POP, HOG and Gabor, denoted as **CHG**. The combined descriptor increases the recognition rate more than the state-of-the-art, especially on two of the most challenging probe sets, $dup1$ and $dup2$ by 5.3% and 4.3% respectively.

Fig. 4 shows the cumulative matching characteristic (CMC) curve of CHG on FERET dataset.

6. Conclusions and Future Work

We proposed a simple but effective edge descriptor for face identification, called CCS-POP. It generalizes the Pairs

Table 10. Comparison of rank 1 face identification rate (%) of our method to other algorithms without preprocessing on FERET dataset.

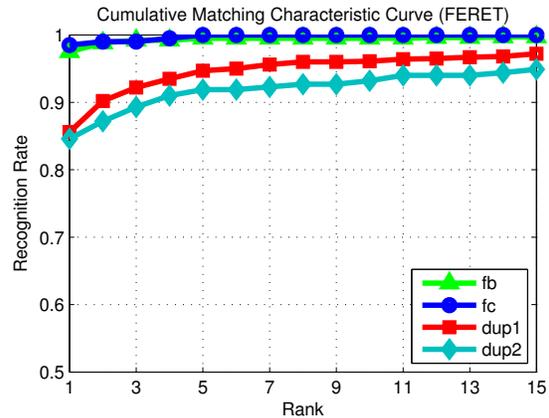| Method	fb	fc	$dup1$	$dup2$
LGBPHS [36]	98	97	74	71
HGPP [35]	**97.6**	98.9	77.7	76.1
SIS [13]	91	90	68	68
POEM [30]	**97.6**	96	77.8	76.5
PLS [24]	95.7	**99.0**	80.3	80.3
CHG	97.5	98.5	**85.6**	**84.6**

Figure 4. Cumulative matching characteristic (CMC) curve of **CHG** in Table 10 on FERET dataset

of Pixels (POP) descriptor [9], which was originally designed for vehicle detection, to be extracted over various orientations for natural objects such as face. It uses raw LBP information without encoding and block based histogram binning, which is empirically shown to preserve useful information in high dimensional feature. To weight the high dimensional information properly, we use PLS regression which gives discriminant weightings using latent spaces. In the experiments on face identification, CCS-POP outperforms both LBP and POP.

The heat maps of the PLS weighting vector showed that CCS-POP captures different information from that captured by other descriptors; its combination with other descriptors achieves state-of-the-art performance on the challenging face identification experiments of FRGC Ver. 1.0 and FERET dataset by fairly large margins: 10.8% in Exp.4 of FRGC, 5.3% in $dup1$ and 4.3% in $dup2$ of FERET.

Since CCS-POP is good at capturing pixel-wise local edge information with the PLS weighting method, it would be of interest to determine if it works well in other biometrics applications such as iris recognition and vascular

Table 9. Comparison of rank 1 face identification rate (%) of our descriptor combined with other descriptors with PLS regression on FERET dataset.

Feature Descriptor	Dim.	fb	fc	$dup1$	$dup2$
LBP + CCS-POP	30,600	94.3	82.0	73.1	58.5
HOG + CCS-POP	50,564	95.7	98.5	81.9	79.5
Gabor + CCS-POP	53,387	97.2	97.9	82.0	77.4
HOG + Gabor	53,523	97.2	99.0	84.2	83.3
HOG + Gabor + CCS-POP	59,795	97.5	98.5	85.6	84.6

References

[1] G. Aggarwal, S. Biswas, and R. Chellappa. UMD Experiments with FRGC Data. In *CVPR Workshop*, page 172, 2005. 126

[2] T. Ahonen, A. Hadid, and M. Pietikainen. Face Description with Local Binary Patterns: Application to Face Recognition. *IEEE T. PAMI*, 28(12):2037–2041, 2006. 121, 122

[3] F. Alter, Y. Matsushita, and X. Tang. An Intensity Similarity Measure in Low-Light Conditions. In *ECCV*, 2006. 124

[4] Z. Cao, Q. Yin, X. Tang, and J. Sun. Face Recognition with Learning-based descriptor. In *CVPR*, 2010. 122

[5] N. Dalal and B. Triggs. Histograms of Oriented Gradients for Human Detection. In *CVPR*, 2005. 121, 122

[6] G. D. Finlayson, S. D. Hordley, and P. M. Hubel. Color by Correlation: A Simple, Unifying Framework for Color Constancy. *IEEE T. PAMI*, 23:1209–1221, 2001. 124

[7] M. Heikkilä, M. Pietikäinen, and C. Schmid. Description of interest regions with local binary patterns. *Pattern Recognition*, 42:425–436, 2009. 121, 122

[8] J. Holappa, T. Ahonen, and M. Pietikainen. An Optimized Illumination Normalization Method for Face Recognition. In *BTAS*, pages 1–6, 2008. 126

[9] A. Kembhavi, D. Harwood, and L. S. Davis. Vehicle Detection Using Partial Least Squares. *IEEE T. PAMI*, 99, 2010. 121, 122, 127

[10] F. D. la Torre, R. Gross, S. Baker, and B. V. K. V. Kumar. Representational Oriented Component Analysis (ROCA) for Face Recognition with One Sample Image per Training Class. In *CVPR*, 2005. 126

[11] E. H. Land. Recent advances in Retinex theory. *Vision Res*, 26(1):7–21, 1986. 124

[12] C. Liu. Capitalize on dimensionality increasing techniques for improving face recognition grand challenge performance. *PAMI*, 28(5):725–737, 2006. 126

[13] J. Liu, S. Chen, Z.-H. Zhou, and X. Tan. Single Image Subspace for Face Recognition. In *AMFG'07*, pages 205–219, 2007. 127

[14] D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *IJCV*, 2(60):91–110, 2004. 122

[15] L. MacDonald, editor. *Digital Heritage: Applying Digital Imaging to Cultural Heritage*. 2006. 124

[16] D. Maturana, D. Mery, and A. Soto. Learning discriminative local binary patterns for face recognition. In *FG*, 2011. 122

[17] A. Mian, M. Bennamoun, and R. Owens. 2D and 3D multimodal hybrid face recognition. In *ECCV*, pages 344–355, 2006. 126

[18] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *IEEE T. PAMI*, 27(10):1615–1630, 2005. 122

[19] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *PAMI*, 24(7):971–987, 2002. 122

[20] V. Ojansivu, E. Rahtu, and J. Heikkila. Rotation invariant local phase quantization for blur insensitive texture analysis. In *ICPR*, 2008. 126

[21] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek. Overview of the face recognition grand challenge. In *CVPR*, volume 1, 2005. 122, 123

[22] P. J. Phillips, H. Moon, P. Rauss, and S. A. Rizvi. The FERET evaluation methodology for face-recognition algorithms. In *CVPR*, pages 137–143, 1997. 122, 123

[23] N. Pinto and D. Cox. Beyond Simple Features: A Large-Scale Feature Search Approach to Unconstrained Face Recognition. In *FG*, 2011. 121, 122

[24] W. R. Schwartz, H. Guo, and L. S. Davis. A Robust and Scalable Approach to Face Identification. In *ECCV*, 2010. 122, 123, 126, 127

[25] A. Sharma and D. Jacobs. Bypassing Synthesis: PLS for Face Recognition with Pose, Low-Resolution and Sketch. In *CVPR*, pages 593–600, 2011. 122

[26] L. Shen and L. Bai. A review on Gabor wavelets for face recognition. *Pattern Anal. Appl.*, 9:273–292, 2006. 121, 122, 125

[27] P. Shih and C. Liu. Comparative Assessment of Content-Based Face Image Retrieval in Different Color Spaces. In *AVBPA*, 2005. 126

[28] 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, 2006. 121

[29] X. Tan and B. Triggs. Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions. *IP*, 19(6):1635–1650, 2010. 123

[30] N.-S. Vu and A. Caplier. Face recognition with patterns of oriented edge magnitudes. In *ECCV*, pages 313–326, 2010. 122, 127

[31] H. Wang, S. Z. Li, and Y. Wang. Face recognition under varying lighting conditions using self quotient image. In *FG*, pages 819–824, 2004. 126

[32] H. Wold. *Partial Least Squares*, volume 6, pages 581–591. 1985. 121, 122

[33] L. Wolf, T. Hassner, and Y. Taigman. Descriptor based methods in the wild. In *ECCV Workshop*, 2008. 122

[34] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. Robust Face Recognition via Sparse Representation. *IEEE Trans. PAMI*, 31(2):210–227, 2009. 122

[35] B. Zhang, S. Shan, X. Chen, and W. Gao. Histogram of Gabor Phase Patterns (HGPP): A Novel Object Representation Approach for Face Recognition. *IEEE T. IP*, 16(1), 2007. 127

[36] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang. Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition. In *ICCV*, 2005. 127

[37] J. Zou, Q. Ji, and G. Nagy. A Comparative Study of Local Matching Approach for Face Recognition. *IEEE T. IP*, 2007. 123