Procedia Computer Science
Volume 51, 2015, Pages 1555–1564
ICCS 2015 International Conference On Computational Science

Power LBP: A novel texture operator for smiling and neutral facial display classification

Bogdan Smolka[1] and Karolina Nurzynska[2]

[1] Institute of Automatic Control, Silesian University of Technology, ul. Akademicka 16, 44-100 Gliwice, Poland,
Bogdan.Smolka@polsl.pl
[2] Institute of Informatics, Silesian University of Technology, ul. Akademicka 16, 44-100 Gliwice, Poland,
Karolina.Nurzynska@polsl.pl

Abstract
Texture operators are commonly used to describe image content for many purposes. Recently they found its application in the task of emotion recognition, especially using local binary patterns method, LBP. This paper introduces a novel texture operator called *power* LBP, which defines a new ordering schema based on absolute intensity differences. Its definition as well as interpretation are given.

The performance of suggested solution is evaluated on the problem of smiling and neutral facial display recognition. In order to evaluate the *power* LBP operator accuracy, its discriminative capacity is compared to several members of the LBP family. Moreover, the influence of applied classification approach is also considered, by presenting results for k-nearest neighbour, support vector machine, and template matching classifiers. Furthermore, results for several databases are compared.

Keywords: power LBP, local binary patterns, classification, emotion recognition

1 Introduction

There are many methods proposed in the literature [4, 12, 16] describing an image content using the information included in the texture. The most general one – first order features [20] – are based on the intensity distribution over an image histogram. They are enriched with additional parameters describing histogram mean, variance, or skewness. Other – second order features [12], [24] – concentrate on spatial relations between the luminance intensities within an image. Hence the spatial relation between pixel value distribution are also incorporated in the co-occurrence matrices, which are the starting point for many parameters calculation, for instance texture contrast, entropy, and information. Works developing this technique for color information are also presented in the literature, e.g. [19]. Different approach presents gray-tone difference matrix [5], which attempts to define a texture measures correlated to human

Selection and peer-review under responsibility of the Scientific Programme Committee of ICCS 2015

doi:10.1016/j.procs.2015.05.350

perception of texture. Next, the run-length matrix proposed in [10] is based on the assumption that textures of good quality are characterised by a small number of consecutive pixels in a given direction with similar luminance. On the contrary, the coarse textures are related to longer runs. There are many possibilities to calculate features describing various texture parameters [4],[6],[23]. More recently histograms of oriented gradients proved to well describe the texture [8] for human detection. Some smoothing filters are presented in application for snowflake recognition in [15]. Moreover, abstract templates used for face detection in [11] synthesise texture, shape and scale of the object.

Apart from the techniques previously described, the local binary patterns operator [16] is widely used. This method interprets the texture as a two-dimensional phenomenon characterized by two orthogonal properties: spatial structure (pattern) and contrast (the "amount" of local image texture). The definition starts from the joint distribution of luminance value on a circularly symmetric neighbour set of pixels in a local processing region. Then an operator invariant against any monotonic transformation of luminance is derived. Tolerance for rotation is achieved by incorporating a fixed set of rotation invariant patterns by this operator.

This work introduces a novel *power* LBP (local binary patterns) texture operator which redefines the way of LBP rank calculation. Its image description capabilities are evaluated in the task of smiling and neutral facial display recognition. Several databases are exploited in order to get better insight in the accuracy of this descriptor. The achieved results are compared with other texture operators from (LBP) family.

Section 2 shortly describes existing LBP methods. Next in Section 3 the *power* LBP approach is introduced. Since in case of emotion recognition the system set-up demands some preprocessing, all necessary details are given in Section 4. The results of performed experiments are presented in Section 5 followed by conclusions drawn in Section 6.

2 Local binary patterns

Local binary patterns [16, 17, 18] [21, Ch. 2] is an operator which characterises a small texture regions. It assumes to work on the monochromatic image I, where for each pixel, the final code is calculated considering the sign $s(z)$ of the differences between central pixel value $I(x_0)$ and its P neighbours pixel values $I(x_k)$ sampled on a circle of given radius R (see Fig. 1):

$$\mathrm{LBP}_{P,R}(x_0) = \sum_{k=1}^{P} s\{I(x_k) - I(x_0)\} \cdot 2^{k-1}, \tag{1}$$

where

$$s(z) = \begin{cases} 1, & z \geq 0, \\ 0, & z < 0. \end{cases} \tag{2}$$

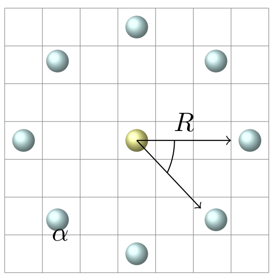
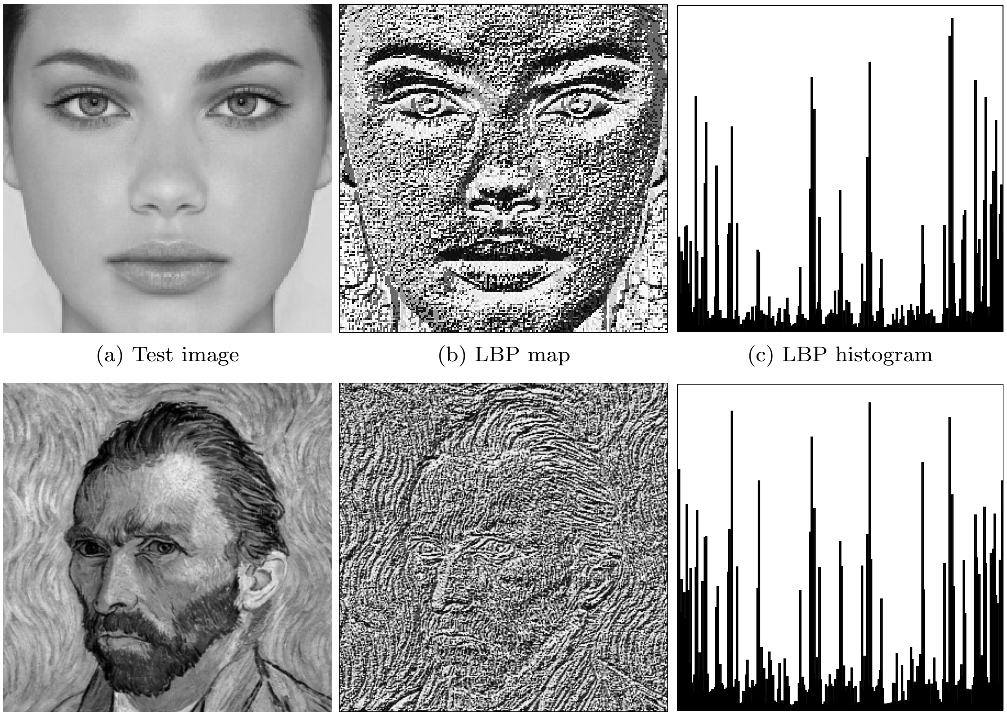


Figure 1: LBP neighbourhood: $P = 8$ sampling points on a circle of radius R.



(a) Test image (b) LBP map (c) LBP histogram

(d) test image (e) LBP map (f) LBP histogram

Figure 2: Local binary patterns visualization.

As the LBP values assigned to the image pixels for $P = 8$ are in the range $[0, 255]$, they can be visualized in form of 8-bit images. Figure 2 depicts exemplary images (a, d), corresponding maps of LBP values (b, e) and the LBP histograms (c, f).

It was noticed in [21][Ch. 2] that not all codes have the same importance. Some of them seem less relevant and therefore might fill a common bin of the histogram, whereas others need to be stored separately. When one looks at the code as a sequence of bits, those valid are characterised by up to two transitions between 0s and 1s, eg. 00111000, and are called *uniform patterns*, (uLBP). This is a natural compression of descriptive data since the histogram length is diminished.

On the other hand, when the image rotates, the distribution of the generated codes in the histogram moves correspondingly [3], [21][Ch. 2]. That disables correct classification in case of the texture rotation. Yet noticing that the points used to generate the code are sampled on the circle, one can easily find out, that rotation of the bits should remove the unwanted result. Therefore, a mapping operator ROR(m,i) was introduced, which is responsible for circular bitwise right rotation of a bit sequence m by i steps:

$$\mathrm{riLBP}_{P,R} = \min_{i} \mathrm{ROR}(\mathrm{LBP}_{P,R}, i). \tag{3}$$

One can also imagine the combination of rotation invariance applied for uniform local binary pattern, what gives a new texture operator called riuLBP.

x_1	x_2	x_3
x_8	x_0	x_4
x_7	x_6	x_5

(a) notation

5	9	11
97	15	50
86	81	72

(b) local window

0	0	0
1		1
1	1	1

(c) LBP labels

00011111

(d) LBP code

Figure 3: Standard LBP procedure.

3 Power LBP

The LBP operator can be regarded as a transformation which is based on the ranks of the gray scale values $I(x_0), I(x_1), \ldots, I(x_P)$ of the pixels $x_0, x_1, \ldots, x_P$, belonging to the local processing window W centered at x_0 as depicted in Fig. 3 a).

If we perform the ordering of the pixel intensities, an ascending sequence of values $I_{(0)}, I_{(1)}, \ldots, I_{(P)}$ is obtained. Using this notation, the pixel with lowest intensity has rank 0, and the pixel with maximum intensity has rank P, if all the pixels differ from each other. In case that two or more pixels have the same intensity we can assign them the same rank. This procedure is called dense ranking: items that compare equal, receive the same ranking number, and the next item(s) receive the immediately following ranking.

Assuming that the rank Ψ of the central pixel x_0 of W is ρ_0, $(\Psi(x_0) = \rho_0)$ we can define the basic LBP as an operation which assigns 0 to the pixels whose rank is less than ρ_0 and 1 otherwise. In this way instead of Eq. 1 we can write

$$\mathrm{LBP}_{P,R}(x_0) = \sum_{k=1}^{P} s\{\Psi(x_k) - \rho_0\} \cdot 2^{k-1}. \tag{4}$$

The ranking procedure is based only on the gray scale values of the pixels intensity and does not consider the distribution of the pixel intensities. Instead of assigning 0 or 1 using the simple ordering criterion, we propose to assign to each pixel x_i from W the cumulated distance defined as

$$D_{x_i} = \sum_{k=0}^{P} |I(x_i) - I(x_k)|^{\gamma}, \quad i = 0, \ldots, P, \tag{5}$$

where $\gamma > 0$ is the power. The ordering of pixels can be now performed with regard to the value of sum of absolute differences between pixels denoted as D raised to the power γ. The generation of the new *power* LBP code is illustrated in Fig. 4.

Then the pixels $x_0, x_1, \ldots, x_P$ can be ordered according to the corresponding values of $D_0, D_1, \ldots, D_P$. In this way a new ordering scheme based on the sum of absolute intensity differences raised to the power γ is introduced. For $\gamma = 1$, this procedure enables to calculate the median of the intensities, which is the pixel $x_{(0)}$ with minimum cumulated distance $D_{(0)}$,

381	353	343
447	331	296
370	345	318

(a) values of D

7	5	3
8	2	0
6	4	1

(b) ranks

1	1	1
1		0
1	1	0

(c) LBP labels

11100111

(d) power LBP

Figure 4: Power LBP generation for the gray scale values shown in Fig. 3.

where $D_{(k)}$ denotes the k-th smallest value of D. For $\gamma = 2$, the pixel with rank 0 in the ordered sequence is the one closest to the average of the pixel intensities.

In this way, we can determine the rank of the central pixel x_0 of the processing window and calculate the LBP value based on the newly defined ranks using Eq. (4). To avoid dense ranking procedure in the situation in which the central pixel has the same cumulative value D_{x_0} as one of its neighbors, we can simply write

$$\mathrm{LBP}_{P,R}(x_0) = \sum_{k=1}^{P} s\{D_{x_k} - D_{x_0}\} \cdot 2^{k-1}, \tag{6}$$

Figure 4 shows the calculation of the LBP code for the pixel intensities of the window in Fig. 3 b) for $\gamma = 1$. As can be observed the value 50 with rank 0 is the median of the intensities in the local window.

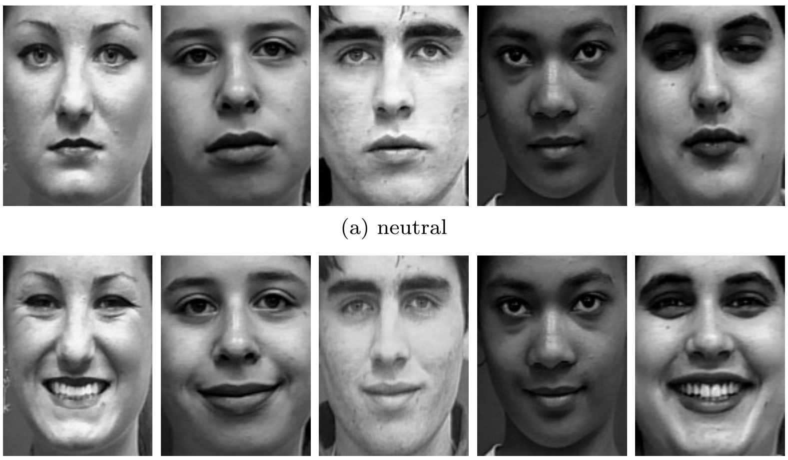
4 Experiment set-up

The performance of the novel *power* LBP operator was examined using three different databases: *Cohn-Kanade*, *Utrecht*, and *All*. The image description accuracy between smiling and neutral facial displays was compared with other approaches to local binary patterns methods: LBP, uLBP, riLBP, and riuLBP. In all presented experiments the set-ups used for calculation were similar, only the texture operator differed.

4.1 Database presentation

Cohn-Kanade AU-Coded Facial Expression Database [7], [14] consists of images with facial display representing six basic emotions: joy, surprise, anger, fear, disgust, and sadness. They are annotated with information which muscles groups were activated during the emotion. Those muscles' groups are called Action Units and were defined by Ekman [9] enabling definition of Facial Action Coding System, *FACS*.

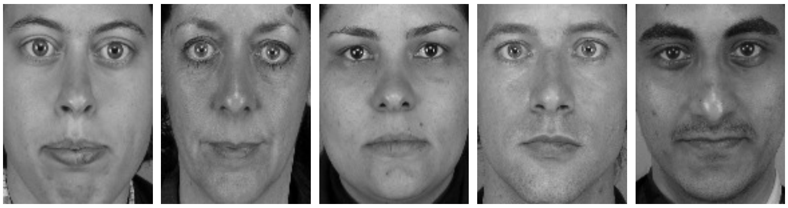
For the performed experiments images representing 82 subjects for smiling and neutral facial display were selected. The image digital resolution is 640×480 pixels with 8-bit precision for



(a) neutral

(b) smiling

Figure 5: Examples of *Cohn-Kanade* image database.



(a) neutral

(b) smiling

Figure 6: Examples of *Utrecht* image database.

a gray scale values. The subjects do not wear glasses or other covering elements, as well as do not have beards. Some examples are presented in Fig. 5.

Utrecht image database was collected at the European Conference on Visual Perception in Utrecht, 2008 [2]. The database consists of 131 colour images of 49 male and 20 female subjects usually taken for neutral and smiling expression. Each image has 900×1200 pixels resolution. In the presented experiment the smiling and neutral expressions were represented by 61 images in each group (see examples in Fig. 6).

Since many of the available databases containing emotion are rather small, it is difficult to generalize the achieved results. Therefore, we decided to use a bigger dataset for evaluation. In order to accomplish this goal, the *All* image database was prepared. It contains of 712 images equally representing two groups of emotions, and gathers images from the two already presented image datasets as well as from *Feret* [1], *Iranian*, *Nottinghams original*, and *Pain* databases [2].

4.2 System set-up overview

Preparing the data for classification is a complex process, which crucial steps are depicted in Fig. 7. In the input image, the face area is specified by the face detector described in [13] and resized to constant resolution of 112×150 pixels. Then the feature vector is obtained for a chosen texture operator. Notwithstanding, calculating global texture operator for the whole image is too general, therefore the face area is divided into 7×10 rectangular subregions, for which the LBP histograms are obtained. Next, all histograms are concatenated in order to build the final feature vector. This approach enables detailed description of each part of the face, which is necessary in order to find differences between emotion displays. Finally, the classification process takes place, during which 10-fold cross-validation schema is exploited due to small number of objects in dataset.

4.3 Classification approaches

In the literature concerning the problem of emotion recognition a plethora classification methods are exploited. Therefore, in order to better understand the accuracy of presented image description technique, the most popular classifiers are exploited:

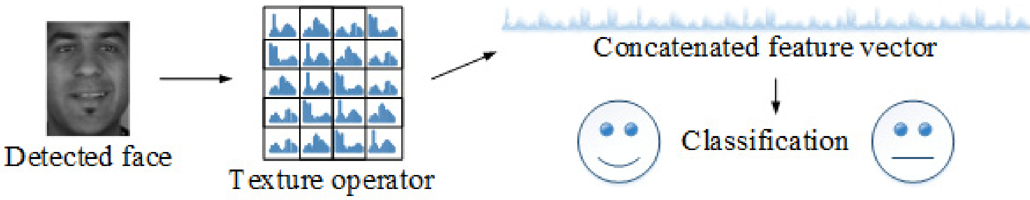


Figure 7: System data flow overview.

kNN k-nearest neighbourhood method represents the non-parametric approach, where the classification of tested element is based on the majority vote of k closest elements in the training set. For described experiments *cityblock* distance metric was used and $k = 9$.

SVM Support vector machine transforms the data into high-dimensional space in order to find a separation plane between objects from different classes. In the experiments *linear* kernel proved to give the best results.

TM Template matching requires generation of average feature vector for each class, then *kNN* with $k = 1$ votes for the one which is the most similar to the tested element. The distance is calculated using chi square metrics as detailed in [22].

5 Results

The goal of the experiment was to compare the performance of the novel *power* LBP operator with other texture operators belonging to the local binary pattern family: LBP, uLBP, riLBP, and riuLBP. All of them were calculated for $P = 8$ sampling points distributed on the circle of radius $R = 1$, the *power* parameter was set to 1. The recognition accuracy was evaluated using three databases: *Cohn-Kanade*, *Utrecht*, and *All* with different classification approaches: *kNN*, *SVM*, and *TM*.

Table 1 presents data gathered for the *kNN* classification method for all examined image sets. One can see that independently of the database the best scores are recorded for LBP, uLBP, and *power* LBP image description techniques. When the rotation invariant version of the operators is applied (riLBP and riuLBP) the efficiency diminishes by a few percent. The most beneficial classification between smiling and neutral facial displays takes place when the novel, *power LPB*, operator is applied; the score reaches: 84.92%, 94.48%, and 78.69% for *All*, *Cohn-Kanade*, and *Utrecht* image databases, respectively.

Similar pattern in described image descriptors is visible when *linear SVM* is applied (see Tab. 2). Once again the rotation invariant version of LBP deteriorates by 6-8% in comparison with other local binary pattern approaches. Here, the variety in accuracy for the best methods is minimal. Yet in case of the most general *All* dataset the *power* LBP brings the best performance of 90.75%, for *Utrecht* it is comparable with uLBP operator with score 86.07%, and for *Cohn-Kanade* it reaches 96.93% similarly as LBP, which is a little bit less accurate when compared to uLBP 97.55%.

Table 1: Classification accuracy comparison for *kNN* method.

Method	*All*	*Cohn-Kanade*	*Utrecht*
power LBP	**84.92**	**94.48**	**78.69**
uLBP	83.36	93.87	72.13
LBP	83.93	93.87	76.23
riLBP	79.37	88.34	68.85
riuLBP	80.94	87.12	68.03

Table 2: Classification accuracy comparison for *linear SVM* method.

Method	*All*	*Cohn-Kanade*	*Utrecht*
power LBP	**90.75**	96.93	**86.07**
uLBP	89.90	**97.55**	**86.07**
LBP	90.47	96.93	83.61
riLBP	81.51	90.18	71.31
riuLBP	82.93	93.25	80.33

Table 3: Classification accuracy comparison for *TM* method.

Method	*All*	*Cohn-Kanade*	*Utrecht*
power LBP	**85.35**	90.80	66.39
uLBP	80.09	**92.02**	**68.03**
LBP	85.06	87.73	65.57
riLBP	66.86	90.18	66.39
riuLBP	67.28	88.96	65.57

Results of experiments concerning the *TM* classification technique are presented in Tab. 3. Firstly, the most beneficial method for *All* database is still *power* LBP with correct classification ratio equal to 85.35%, the LBP technique is not far behind. The results follow the same pattern when the other databases are concerned. Secondly, in case of *Cohn-Kanade* image set, the best performance is 92.02% gathered for uLBP, whereas *power* LBP achieves 90.80%. More interesting is the fact, that in this test, the rotation invariant techniques brought comparable results around 90%. Finally, for the *Utrecht* database the performance is very low regardless of applied texture operator.

Looking at the results from other perspective one can easily notice, that the best performance is achieved when *linear SVM* classification is applied. Only for this solution the efficacy reaches almost 98% for *Cohn-Kanade* database and are around 90% for the other image sets. On the other hand, the weakest accuracy was recorded when the *TM* technique was used. Finally, the results differ depending on the image set. The best accuracy corresponds to the best image quality gained in the *Cohn-Kanade* database. The *Utrecht* data is also characterised by high resolution, but the subjects are wearing glasses or have beards, moreover in some pictures artefacts from the light reflections are visible. Probably all this circumstances influence the accuracy.

In order to give better overview, additional tests using other texture operators in similar experiment set-up were performed for the *Cohn-Kanade* database. Once again the best results were recorded in case of *linear SVM* and are as follows: for co-occurence matrix operator the correct classification ratio equals 88.42%, for run length matrix 88.42%, and for gray-tone difference matrix 65.48%, which proves the superiority of LBP family approach.

Apart from the classification accuracy of the *power* LBP, the influence of the power γ parameter on the performance in case of varying conditions due to illumination change and noise was investigated. In order to perform such experiment clone versions of *Cohn-Kanade* datasets were prepared, where each of them included one of the following changes: Poisson noise, salt & pepper noise (probability=0.05), and variation in illumination, which consisted of linear light changes in the image in direction from top to bottom, right to left and vice versa, combined with irregular lighting starting in each corner (see Fig. 8). Figure 9 depicts the results when recognition between smiling and neutral facial display on each of considered versions of the *Cohn-Kanade* database was examined with *linear SVM*. The plot shows that in comparison with classification performance in case of stable lighting conditions (original database), outcomes in case of varying lighting conditions do not change much, especially for higher values of the power γ. The algorithm is also resistible to Poisson noise and diminishes its accuracy by few percents when salt & pepper noise is considered.

(a) Poisson (b) salt & pepper (c) bottom-top (d) right-left (e) bottom-left

Figure 8: Examples of noise (a,b) and light changes (c,d,e) incorporated into *Cohn-Kanade* database.

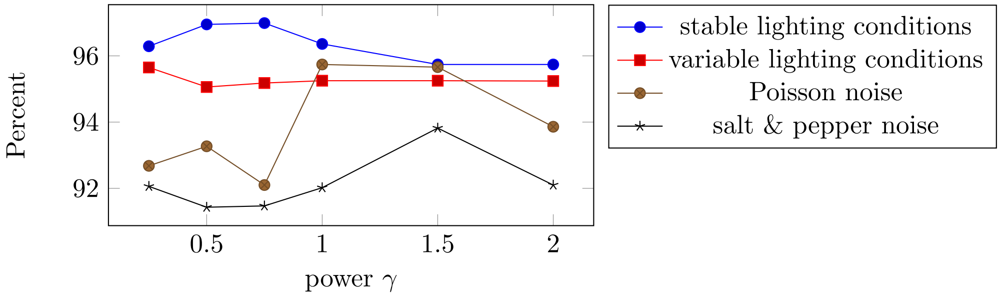


Figure 9: Influence of power γ parameter on classification accuracy of *power* LBP in case of lighting changes and noise disturbances.

6 Conclusions

This work introduces the novel *power* LBP operator, giving its definition followed by interpretation. The accuracy of data description is evaluated on the problem of smiling and neutral facial display classification with utilization of three different datasets: *All*, *Cohn-Kanade*, and *Utrecht*. The performance is compared with other well known texture operators from local binary pattern family: LBP, uLBP, riLBP, and riuLBP. According to gathered results, the accuracy in most cases were the highest among considered techniques or comparable with outcomes recorded for LBP and uLBP approaches. In the future we will focus on the effectiveness of the uniform and rotation invariant versions of the introduced *power* LBP operator.

ACKNOWLEDGEMENT

This work has been supported by the Polish National Science Centre (NCN) under the Grant: DEC-2012/07/B/ST6/01227 and was performed using the infrastructure supported by POIG.02.03.01-24-099/13 grant: GCONiI - Upper-Silesian Center for Scientific Computation.

References

[1] FERET database. http://www.itl.nist.gov/iad/humanid/ feret/feret_master.html, July 2014.

[2] Smile databases. http://pics.psych.stir.ac.uk, July 2014.

[3] J. Ahonen, T.and Matas, Ch. He, and M. Pietikäinen. Rotation invariant image description with local binary pattern histogram fourier features. In *Proceedings of the 16th Scandinavian Conference on Image Analysis*, SCIA '09, pages 61–70, Berlin, Heidelberg, 2009. Springer-Verlag.

[4] F. Albregtsen, B. Nielsen, and H.E. Danielsen. Adaptive gray level run length features from class distance matrices. In *Pattern Recognition. Proc. 15th International Conf. on*, volume 3, pages 738–741 vol.3, 2000.

[5] M. Amadasun and R. King. Textural features corresponding to textural properties. *Systems, Man and Cybernetics, IEEE Transactions on*, 19(5):1264–1274, Sep 1989.

[6] A. Chu, C.M. Sehgal, and J.F. Greenleaf. Use of gray value distribution of run lengths for texture analysis. *Pattern Recognition Letters*, 11(6):415 – 419, 1990.

[7] J.F. Cohn, A.J. Zlochower, J. Lien, and T. Kanade. Automated face analysis by feature point tracking has high concurrent validity with manual facs coding. *Psychophysiology*, 36(2):35–43, 1999.

[8] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005.*, volume 1, pages 886–893 vol. 1, June 2005.

[9] P. Ekman and W. Friesen. Facial action coding system: a technique for the measurement of facial movement. *Consulting Psychologists Press*, 1978.

[10] M.M. Galloway. Texture analysis using gray level run lengths. *Computer Graphics and Image Processing*, 4(2):172 – 179, 1975.

[11] He Guo, Yulong Yu, and Qi Jia. Face detection with abstract template. In *Image and Signal Processing (CISP), 2010 3rd International Congress on*, volume 1, pages 129–134, Oct 2010.

[12] R.M. Haralick, K. Shanmugam, and Its'Hak Dinstein. Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, SMC-3(6):610–621, Nov 1973.

[13] M. Kawulok and J. Szymanek. Precise multi-level face detector for advanced analysis of facial images. *Image Processing, IET*, 6(2):95–103, March 2012.

[14] J.-J. J. Lien, T. Kanade, J.Y. Cohn, and C. Li. Detection, tracking, and classification of action units in facial expression. *Journal of Robotics and Autonomous Systems*, 31:131–146, July 1999.

[15] K. Nurzynska, Mamoru Kubo, and Ken-ichiro Muramoto. Snow particle automatic classification with texture operators. In *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*, pages 2892–2895, July 2011.

[16] T. Ojala, M. Pietikäinen, and D. Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1):51 – 59, 1996.

[17] T. Ojala, M. Pietikäinen, and T. Mäenpää. A generalized local binary pattern operator for multiresolution gray scale and rotation invariant texture classification. In *Proc. 2nd Intern. Conf. on Advances in Pattern Recognition*, pages 397–406, London, UK, 2001. Springer.

[18] T. Ojala, M. Pietikäinen, and T. Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7):971–987, Jul 2002.

[19] C.A. Perez, C.F. Navarro, D.A. Schulz, J.A. Saravia, and C.M. Aravena. Pattern detection using a new haralick quaternion color extraction model and support vector machine classifier. In *Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on*, pages 3300–3304, Oct 2013.

[20] M. Petrou and P.G. Sevilla. *Image processing: dealing with texture.* John Wiley & Sons Inc., 2006.

[21] M. Pietikäinen, A. Hadid, G. Zhao, and T. Ahonen. *Computer Vision Using Local Binary Patterns*, volume 40 of *Computational Imaging and Vision*. Springer, 2011.

[22] C. Shan, Sh. Gong, and P.W. McOwan. Facial expression recognition based on local binary patterns: A comprehensive study. *Image Vision Comput.*, 27(6):803–816, May 2009.

[23] X. Tang. Texture information in run-length matrices. *Image Processing, IEEE Transactions on*, 7(11):1602–1609, Nov 1998.

[24] S.W. Zucker and D. Terzopoulos. Finding structure in co-occurrence matrices for texture analysis. *Computer Graphics and Image Processing*, 12(3):286 – 308, 1980.