Mouth Covered Detection For Yawn

Masrullizam Mat Ibrahim *1, John S Soroghan *2, Lykourgos Petropoulakis *3
*Centre for Excellence in Signal and Image Processing (CeSIP)
Electronic and Electrical Engineering Department, University of Strathclyde
204 George Street, Glasgow G1 1XW, Scotland, United Kingdom.
1 masrullizam.bin-mat-ibrahim@strath.ac.uk
2 j.soraghan@strath.ac.uk
3 l.petropoulakis@strath.ac.uk

***Abstract*— Yawn is one of the common fatigue sign phenomena. The common technique to detect yawn is based upon the measurement of mouth opening. However, the spontaneous human action to cover the mouth during yawn can prevent such measurements. This paper presents a new technique to detect the covered mouth by employing the Local Binary Pattern (LBP) features. Subsequently, the facial distortions during the yawn process are identified by measuring the changes of wrinkles using Sobel edges detector. In this research the Strathclyde Facial Fatigue (SFF) database that contains genuine fatigue signs is used for training, testing and evaluation of the developed algorithms. This database was created from sleep deprivation experiments that involved twenty participants.**

Keywords—Yawn, fatigue, mouth covered, destortion detection, database.

I. INTRODUCTION

Fatigue is a syndrome that is well known as tiredness, exhaustion or lethargy, and it is a common health complaint. Generally, fatigue is defined as a feeling of lack of energy and can be caused by inadequate rest, long hours of physical and mental activity, sleep disturbance, excessive stress or a combination of these factors. The symptoms of fatigue can be felt as well as having annoying effects such as impairment of hand-eye coordination, low motivation, poor concentration, slow reflexes and response, feeling overwhelmed and inability to pay attention. The fatigue causes have contributed the largest number of road accidents, leading to loss of drivers' and passengers lives. According to the UK Department of Transport[1], in 2010, 1850 people were killed, and 22,660 were seriously injured. Furthermore, fatigued drivers have contributed to a 20% of total road accidents [2]. In USA, the administration of National Highway Traffic Safety estimates that there are approximately 100,000 crashes each year caused by fatigue and drowsiness [3].

In general the fatigue detection can be recognised from physiological activities such as eye and mouth activities, brain activities, and heart activities. Furthermore, the physical activities and behavior of human also can indicate the sign of fatigue such as implemented in [4-6]. Yawn is a symptom that visually indicated by widely mouth opening. Because of this, all researches in yawn detection research focus on the technique to measure and classify the mouth opening [7-10]. In this paper, we introduce a new approach to detect the mouth covered with the distortion detection. For the mouth covered detection the LBP features and learning machine classifier are employed. In order to detect the distortion of specific face region the Sobel operator of edges detector is used. In this research, the genuine yawn signs of fatigue from Starathclyde Facial Fatigue (SFF) database are utilised for the training, testing and evaluation. This video footage database is developed explicitly to acquire signs of fatigue from the faces of subjects in various stages of sleep deprivation. The ethically approved sleep deprivation experiments were conducted at the University of Strathclyde. Twenty participants were involved in these experiments each one was sleep deprived for periods of 0, 3, 5 and 8 hours on separate occasions. During each experimental session the participants' faces were recorded while carrying out a series of cognitive tasks.

II. REGION INTEREST INITILISATION

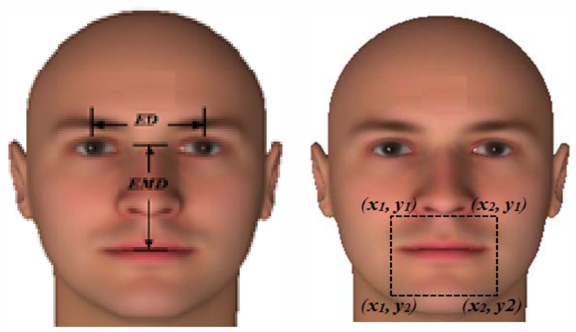
In face acquisition operation, the face, eyes and mouth are sequentially detected using a Viola Jones technique [11, 12]. In this technique Viola and Jones introduced a cascade classifier, which is a series of classifier for training that is applied to every sub-window in input image in order to increase detection performance as well as to reduce the computation time. As implemented in [13, 14], a lot of the face, eyes and mouth images are required to train the Haar-like features [12] in a cascade classifier. In this acquisition operation, the face region is firstly detected, and then the eyes and mouth regions are detected within this face region. Subsequently an anthropometric measurement is carried out to obtain the distinctive distances between these components, since each person has different facial components coordinates. This information is then used in the formation of the two regions of interest.

978-1-4799-0269-9/13/$31.00 ©2013 IEEE

A. Focused Mouth Region (FMR)

The focused mouth region (FMR) is formed based on the first detected location of the eyes and mouth. The distances of these two main face components are measured. A distance between the centre of eyes *ED*, and a distance between centre of mouth and the middle point between eyes *EMD* as shown in Fig.1 (a) are obtained. Then, the ratio of this distance *rem* is computed as follows:

$$rem = \frac{ED}{EMD} \tag{1}$$



(a) (b)

Fig.1. (a) Anthropometric measurement. (b) Focus Mouth Region (FMR)

From these two distances the coordinate of the FMR as shown in Fig.1(b) are empirically defined as follows:

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_R \\ y_R + 0.75EMD \end{bmatrix} \tag{2}$$

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_L \\ y_L + 0.75EMD + 0.8ED \end{bmatrix} \tag{3}$$

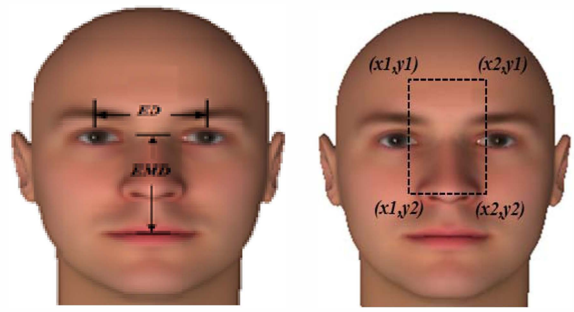
The FMR is dependent on the location and distance of the eyes, when the face moves forward the eye distance increases and so does the FMR. On the other hand, the FMR becomes smaller when the face moves backwards.

B. Focus Distortion Region (FDR)

The focused distortion region (FDR) is a region in the face that is used to measure changes of facial distortion. This region, which is most likely to undergo changes (i.e. facial distortions) during yawn, was identified based on conducted experiments using the SFF database. The coordinates of FDR, shown in 2(b), are empirically defined as:

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_R \\ y_R - 0.75EMD \end{bmatrix} \tag{4}$$

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_L \\ y_L + 0.75EMD \end{bmatrix} \tag{5}$$



(a) (b)

Fig. 2. (a) Anthropometric measurement, (b) Focused Distortion Region (FDR)

III. COVERED MOUTH DETECTION

It is not unusual to hand-cover the mouth during yawn. This is almost spontaneous human action, means that the yawn can no longer be detected from the mouth opening. Therefore, this paper introduces a new approach to detect covered mouth during yawn. In this approach the FMR is examined for whether the region is covered or not. Fig.3 shows image examples where the FMR is covered (Fig.3 (b)) and not covered (Fig.3 (a)).

(a) (b)

Fig. 3. (a) Uncovered mouth. (b) Covered mouth.

978-1-4799-0269-9/13/$31.00 ©2013 IEEE

. From these images the texture difference with the FMR region covered and not covered is clearly visible. In order to differentiate between these two regions Local Binary Pattern (LBP) features are chosen to extract the texture pattern from FMR. LBP features have been proven highly accurate descriptors of texture, robust to the monotonic gray scale changes, as well as simple to implement computationally [15].

A. *Local Binary Pattern (LBP)*

Local Binary Pattern (LBP) originally applied on texture analysis to indicate the discriminative features of texture. With reference to Fig.4, the basic LBP code of the central pixel is obtained by subtracting its value form each of its neighbours and when a negative number is obtained is substituted by a 0, else it is 1.

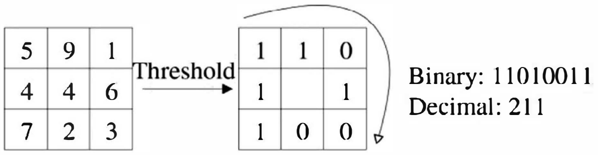


Fig.4. Example of the basic LBP operator.

The limitation of the basic LBP operator is that it represents a small scale feature structure which may be unable to capture large scale dominant features. In order to deal with the different scale of texture, Ojala *et al.* [16] presented an extended LBP operator where the radius and sampling points are increased. Fig.5 shows the extended operator where the notation (P,R) represent a neighborhood of P sampling points on a circle radius R.

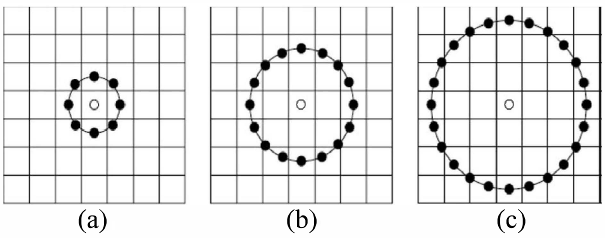


Fig.5. Example of extended LBP operator. (a) (8,1), (b) (16,2) and (c) (24,3) neighborhoods.

The result of LBP can be formulated in decimal form as follows:

$$LBP_{P,R}(x_c, y_c) = \sum_{P=0}^{P-1} s(i_P - i_c)2^P \qquad (6)$$

Where i_c and i_P denote gray level values of the central pixel and P represent surrounding pixels in the circle neighborhood with radius R. Function $s(x)$ is defined as:

$$s(x) = \begin{cases} 1, & if\ x \geq 0 \\ 0, & if\ x < 0 \end{cases} \qquad (7)$$

The operator LBP as formulated in (6) has a rotation effect if the image is rotated, then, the surrounding pixels in each neighborhood are moving accordingly along the perimeter of circle. In order to remove this effect Ojala [16] proposed rotation invariant (*ri*) LBP as follows:

$$LBP_{P,R}^{ri} = \min\{ROR(LBP_{P,R}, i) \mid \quad i = 0,1,\ldots\ldots\ldots\ldots, P-1\} \quad (8)$$

Where *ROR(x,i)* performs a circular bitwise right shift on the *P*-bit number x with i time. Ojala *et al.* also propose LBP uniform pattern (*u2*) , $LBP_{(P,R)}^{u2}$. The LBP is called uniform when it contains two bitwise transition from 0 to 1 or vice versa. For example, 1111111 (0 transition) and 00111100 (2 transition) are both uniform, for 100110001 (4 transition) and 01010011(6 transition) are not uniform.

In addition, another one LBP operator is combination of rotation invariant pattern with uniform pattern, $LBP_{(P,R)}^{riu2}$. This operator is calculated by simply counting ones in uniform pattern codes and all non-uniform pattern are labeled in a single bin. The example of LBP histogram of three types of LBP operators for covered and uncovered mouth region is shown in Fig.6.

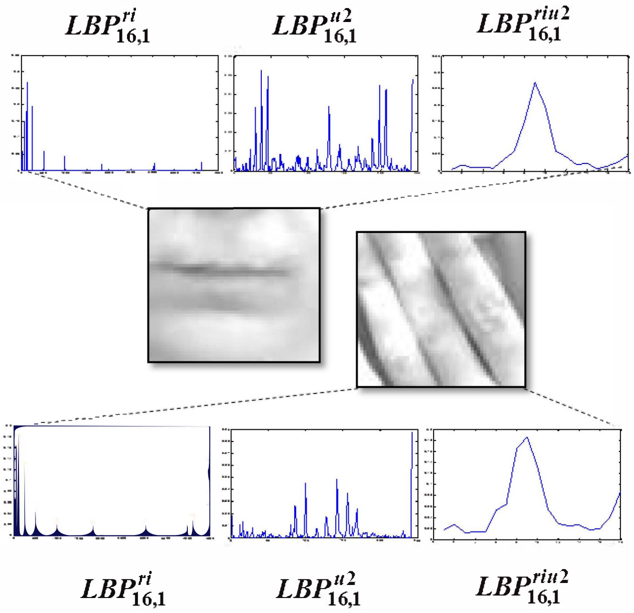


Fig. 6. LBP histogram for uncovered mouth and covered mouth region

978-1-4799-0269-9/13/$31.00 ©2013 IEEE

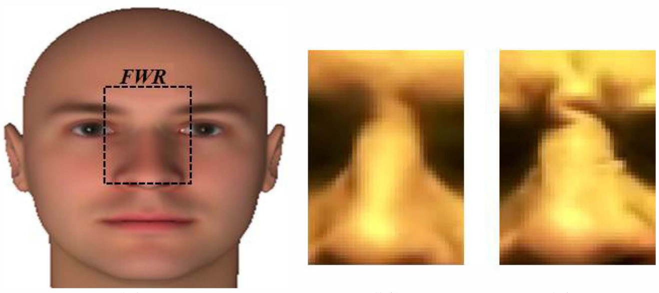
IV. DISTORTION DETECTION

The ability to detect the covered mouth is not adequate to conclude that someone is yawning. The mouth is probably accidentally or intentionally covered - not necessarily due to yawn. Thus, this paper introduces a technique to detect the wrinkles in a specific region of the face most likely to wrinkle while yawning. This region is identified based on observations and the experiments conducted using the video footage from SFF database. The most affected region that shows wrinkle changes during yawning is labeled as Focus Distortion Region (FDR) shown in Fig. 7(a). Fig 7(b) and (c) show the changes of the region due to yawn.

In order to measure the changes of FDR, the work presented here uses edge detection to detect the wrinkles. Based on the experiments carried out, the Sobel operator [17] is chosen since it is able to detect most of the required edges. The Sobel operator first calculates the intensity gradient at each point in the region of interest. Then, it provides the direction of the largest possible increase from light to dark and the rate of change in the horizontal and vertical directions. The Sobel operator represents a partial derivative of f (x, y) and of the central point of a 3x3 area of pixels.The gradients for the horizontal and vertical directions for the region of interest is then computed as follows [18] :

$$G_x = \{f(x+1,y-1)+sf(x+1,y)+f(x+1,y+1)\} \\ -\{f(x-1,y-1)+2f(x-1,y)+f(x-1,y-1)\} \quad (9)$$

$$G_y = \{f(x-1,y+1)+sf(x,y+1)+f(x+1,y+1)\} \\ -\{f(x-1,y-1)+2f(x,y-1)+f(x+1,y-1)\} \quad (10)$$



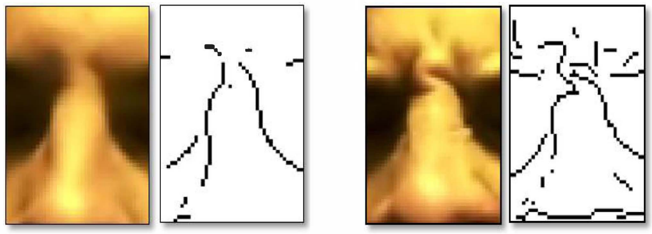
(a) (b) (c)

Fig. 7. (a) Focus Wrinkle Region (FDR). (b) normal condition in FDR. (c) yawn condition in FDR

The wrinkles in FDR are detected based on the both direction gradient and the result as shown in Fig. 8. For identifying the changes of wrinkles in FDR during yawn, sum of absolute values FDR_{SAD} (11) is applied to compute the numbers wrinkles in the region. The normalize FDR_{SAD} is calculated as in (12) where W *and* H denote the width and height of FDR respectively. During yawn the numbers of the detected edges are increased as shown in Fig.9.

$$FDR_{SAD} = \sum_{i,j} \left| I_1(i,j) - I_2(x+i, y+j) \right| \quad (11)$$

$$Normalised\ FDR_{SAD} = \frac{I_{SAD}}{255 \times W \times H} \quad (12)$$



(a) (b)

Fig. 8. FDR with input image and edges detected image. (a) normal condition in FDR. (b) yawn condition in FDR

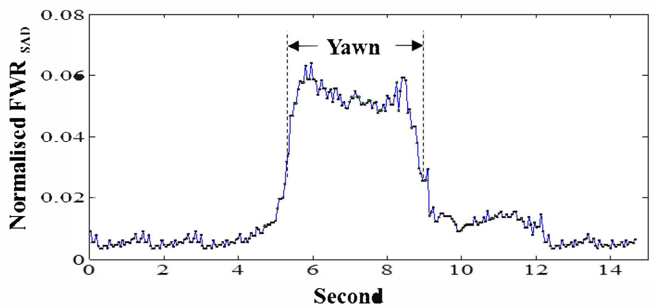


Fig. 9. The plot of Normalised FDR_{SAD} value during yawn.

V. EXPERIMENT RESULTS

An extensive experiment has been carried out using SFF database to evaluate the performance of developed algorithms. In order to detect the covered mouth, the FMR is examined in every video frame. The LBP features are extracted from the region, and then the learning machine algorithm decides the status. In this research two classifiers Support Vector Machine (SVM) and Neural Network (NN) were applied in order to determine their suitability for use in the analysis algorithm. 500 covered mouth images and 100 non covered mouth images (as shown in Fig.3) of size 50X55 pixels have been used to train the NN and SVM classifiers. For the features extraction three operators of LBP; rotational invariant (*ri*), uniform pattern (*u2*), and rotational invariant pattern with uniform pattern (*riu2*), are applied with 8, 16, and 24 radius. The results of the classification for 300 images for the SVM classifier in four different kernel functions and the NN in two different network layers are stated in Table I and Table II. Table I

978-1-4799-0269-9/13/$31.00 ©2013 IEEE

describe the percentage of detection rate and Table II shows the false positive detection rate. From the graph in Fig. 10 the detection rates are balanced between the LBP operators. However for false positive detection rate, as shown in Fig.11, the uniform pattern (*u2*) indicates the most promising results.

TABLE I
MOUTH COVERED DETECTION RATE

LBP	SVM- rbf Detect (%)	SVM-linear Detect (%)	SVM-polynomial Detect (%)	SVM-quadratic Detect (%)	NN-(5) Detect (%)	NN- (10) Detect (%)
8ri	59	97.5	86.7	80	92.5	91.2
8u2	51.6	95	57.8	83.3	83.3	90
8riu2	98.3	97.5	95	95	98.3	99.2
16ri	50.8	90	63.3	63.3	90.8	95
16u2	54.2	98.3	50.8	78.3	97.5	100
16riu2	84.1	99.2	82.5	85	92.5	92.5
24u2	55	95.8	55	69.2	96.7	98.3
24riu2	66.7	94.2	91.7	85.8	95.8	95

TABLE II
MOUTH COVERED FALSE POSITIVE DETECTION RATE

LBP	SVM-rbf False +ve (%)	SVM-linear False +ve (%)	SVM-polynomial False +ve (%)	SVM-quadratic False +ve (%)	NN-(5) False +ve (%)	NN- (10) False +ve (%)
8ri	2.2	14.4	24.4	18.3	15	8.9
8u2	0	2.2	0	1.7	1.1	2.8
8riu2	27.8	14.4	25	27.8	12.8	13.9
16ri	0	25	19.4	26.1	15	16.7
16u2	0	2.2	0	0	1.1	3.3
16riu2	17.8	12.8	22.2	26.1	16.7	11.1
24u2	0	0	0	0	1.7	1.1
4riu2	6.6	15.6	16.1	16.1	12.8	16.1

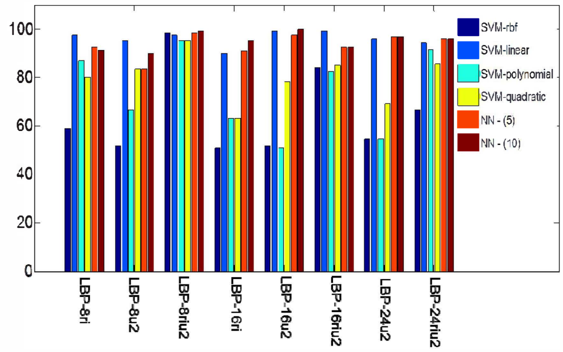


Fig. 10. Result of classification for covered mouth region detection

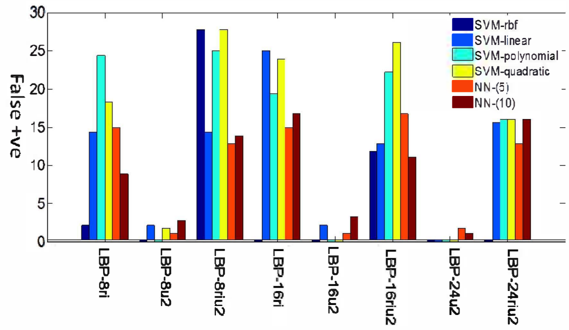


Fig.11. Result of classification for uncovered mouth region detection

VI. CONCLUSION

A new technique to detect covered mouth during yawning situations was presented. This technique is a combination of mouth covered detection, and wrinkles changes detection. In mouth covered, an LBP uniform operator is applied to extract features of the mouth region, and then the mouth covered is classified using a classifier. Based on experiment, the NN classifier (5 networks) shows the best result. In order to verify that the mouth covering is indeed due to yawning, the wrinkles of specific regions on the face are measured. Yawning is confirmed when wrinkles changes are detected.

REFERENCES

[1] D. f. T. UK, "Reported Road Casualties in Great Britain: 2010 Annual Report " Department for Transport UK, United Kingdom, Statistical Release - 29 September 2011

[2] P. Jackson, C. Hilditch, A. Holmes, N. Reed, N. Merat, and L. Smith, "Fatigue and Road Safety: A Critical Analysis of Recent Evidence," ed. London:

978-1-4799-0269-9/13/$31.00 ©2013 IEEE

Department for Transport: Department for Transport, 2011, p. 88.

[3] R. Resendes and K. H. martin, "Saving Lives Through Advanced Safety Technology," Faderal Highway Administration US, Washington DC2003.

[4] Daimler. (Accessed February 2012). *Drowsiness-detection system warns drivers to prevent them falling asleep momentarily* [online]. Available: http://www.daimler.com/dccom/0-5-1210218-1-1210332-1-0-0-1210228-0-0-135-0-0-0-0-0-0-0.html

[5] M. Michael Littner, C. A. Kushida, D. Dennis Bailey, R. B. Berry, D. G. Davila, and M. Hirshkowitz, "Practice parameters for the role of actigraphy in the study of sleep and circadian rhythms: an update for 2002," *Sleep,* vol. 26, p. 337, 2003.

[6] N. Mabbott, "ARRB Pro-active fatigue management system," *Road and Transport Research,* vol. 12, 2003.

[7] F. Xiao, Y. Bao-Cai, and S. Yan-Feng, "Yawning Detection for Monitoring Driver Fatigue," in *Machine Learning and Cybernetics, 2007 International Conference on*, 2007, pp. 664-668.

[8] J. Jiménez-Pinto and M. Torres-Torriti, "Driver alert state and fatigue detection by salient points analysis," in *Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on*, 2009, pp. 455-461.

[9] M. Omidyeganeh, A. Javadtalab, and S. Shirmohammadi, "Intelligent driver drowsiness detection through fusion of yawning and eye closure," in *Virtual Environments Human-Computer Interfaces and Measurement Systems (VECIMS), 2011 IEEE International Conference on*, 2011, pp. 1-6.

[10] Y. Lu and Z. Wang, "Detecting Driver Yawning in Successive Images," in *Bioinformatics and Biomedical Engineering, 2007. ICBBE 2007. The 1st International Conference on*, 2007, pp. 581-583.

[11] P. Viola and M. Jones, "Robust real-time face detection," in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, 2001, pp. 747-747.

[12] P. Viola and M. J. JONES, "Robust Real-Time Face Detection," *International Journal of Computer Vision,* vol. 57(2), pp. 137–154, 2004 2004.

[13] P. I. Wilson and J. Fernandez, "Facial feature detection using Haar classifiers," *Journal of Computing Sciences in Colleges,* vol. 21, pp. 127-133, 2006.

[14] M. Castrillón, O. Déniz, C. Guerra, and M. Hernández, "ENCARA2: Real-time detection of multiple faces at different resolutions in video streams," *Journal of Visual Communication and Image Representation,* vol. 18, pp. 130-140, 2007.

[15] G. Zhao and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 29, pp. 915-928, 2007.

[16] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 24, pp. 971-987, 2002.

[17] R. O. Duda and P. E. Hart, *Pattern classification and scene analysis* vol. 3: Wiley New York, 1973.

[18] J.-Y. Zhang, Y. Chen, and X.-x. Huang, "Edge detection of images based on improved Sobel operator and genetic algorithms," in *Image Analysis and Signal Processing, 2009. IASP 2009. International Conference on*, 2009, pp. 31-35.

978-1-4799-0269-9/13/$31.00 ©2013 IEEE