QUEST: Quadrilateral Senary bit Pattern for Facial Expression Recognition

(Accepted in IEEE International Conference on Systems, Man, and Cybernetics, Miyazaki, Japan, pp. 1498-1503, (2018))

By: Monu Verma

Contents

- Shortfalls of previous methods
- Proposed framework
- Proposed descriptor
- Properties of QUEST
- Qualitive analysis between proposed and existing methods
- Quantitative analysis between proposed and existing methods
- Conclusion

Shortfalls of Previous Methods

- LBP faced problem in noisy conditions .
- Directional descriptors like LDP, LDN, LDTP and LDTerP, extracted features of expressive regions by applying different compass mask as sobel, krish and robinson. Therefore, the performance of these methods fully dependent on selected predesigned compass masks.
- Most recent descriptor LDTerP mainly focuses on the extreme edge variations and ignores micro level edge information. This may lead to salient feature loss, thereby degrading the discrimination capability of the descriptor.

Procedure of proposed framework

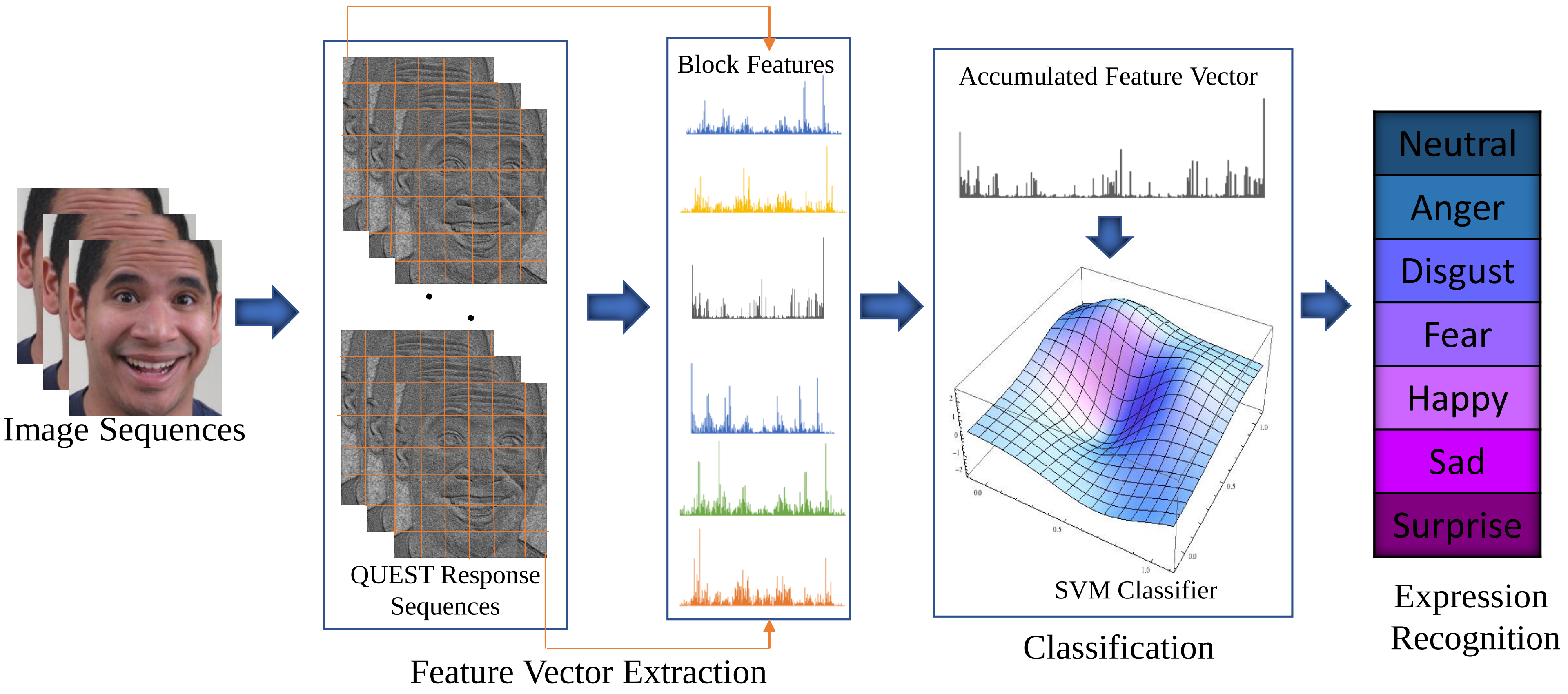


Fig. 1. Overall Process of proposed method

Proposed descriptor

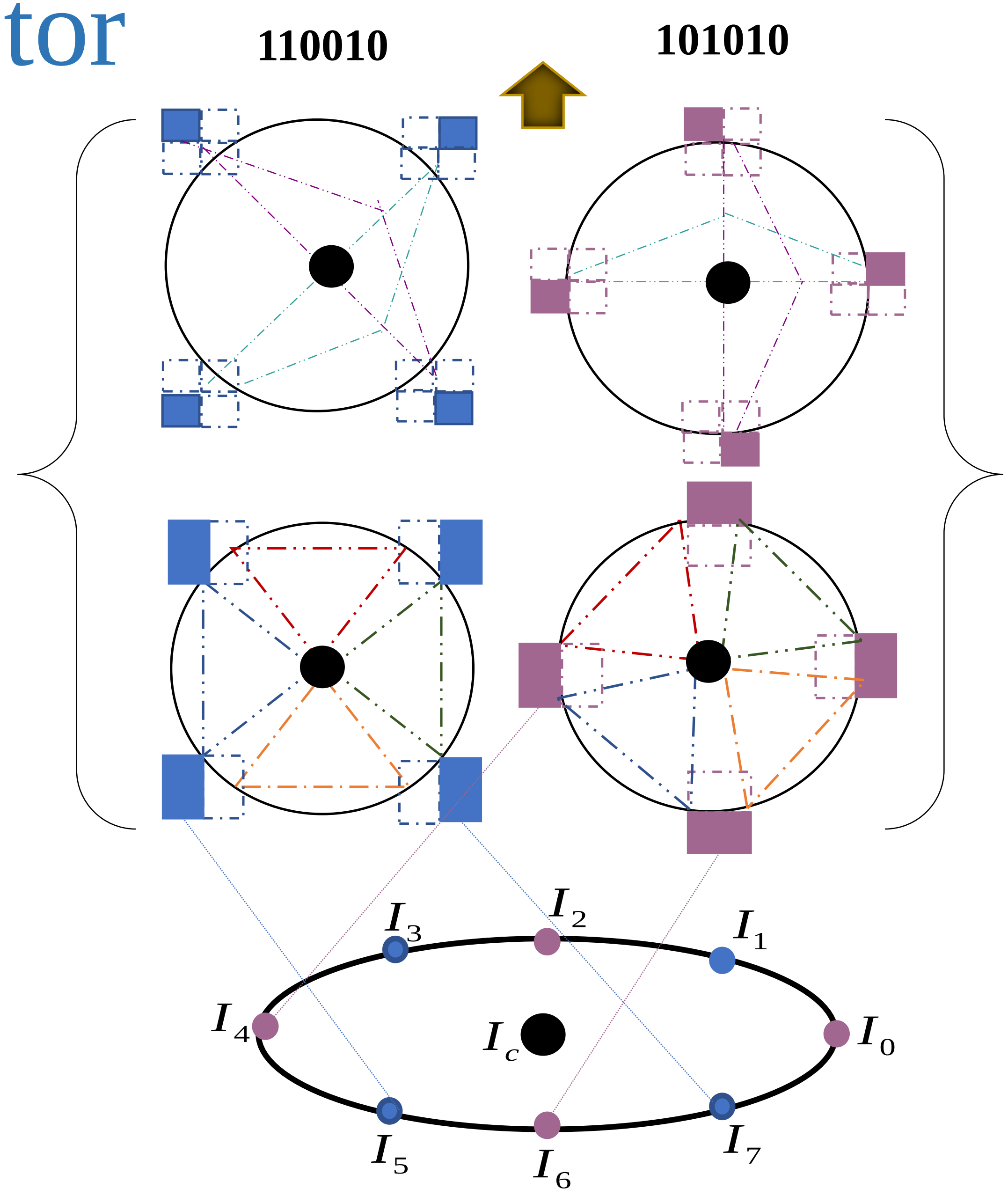


Fig. 2. Proposed Descriptor

Proposed descriptor Cont…

The detailed step wise representation of the QUEST is given in Eq (1-4).

$$L(x,y)=\sum_{\upsilon=0}^{p-3}\left\{f(Q_{\upsilon,\omega}-I_c)\times 2^{\upsilon}\right\}_{w=0}^{1} \qquad (1)$$

Where,

I_c → is the reference pixel in the image.

p → is the total number of the neighborhood.

$$Q_{\upsilon,\omega=}\frac{I_{(2\upsilon-\omega+\psi(\upsilon))\bmod p}+I_{(2(\upsilon+1)-\omega)\bmod p}}{(\psi(\upsilon)/3)+2} \qquad (2)$$

$$\psi(\eta)=(\lfloor\eta/4\rfloor)\times(p-2) \qquad (3)$$

$$f(x)=\begin{cases}1 & if\ \ x\geq 0\\ 0 & otherwise\end{cases} \qquad (4)$$

Proposed descriptor Cont…

When $\upsilon = 0, \omega = 1$

$$Q_{0,1} = \frac{I_{(2*0-1+\psi(0)) \bmod 8} + I_{(2(0+1)-1) \bmod 8}}{(\psi(0)/3)+2}$$

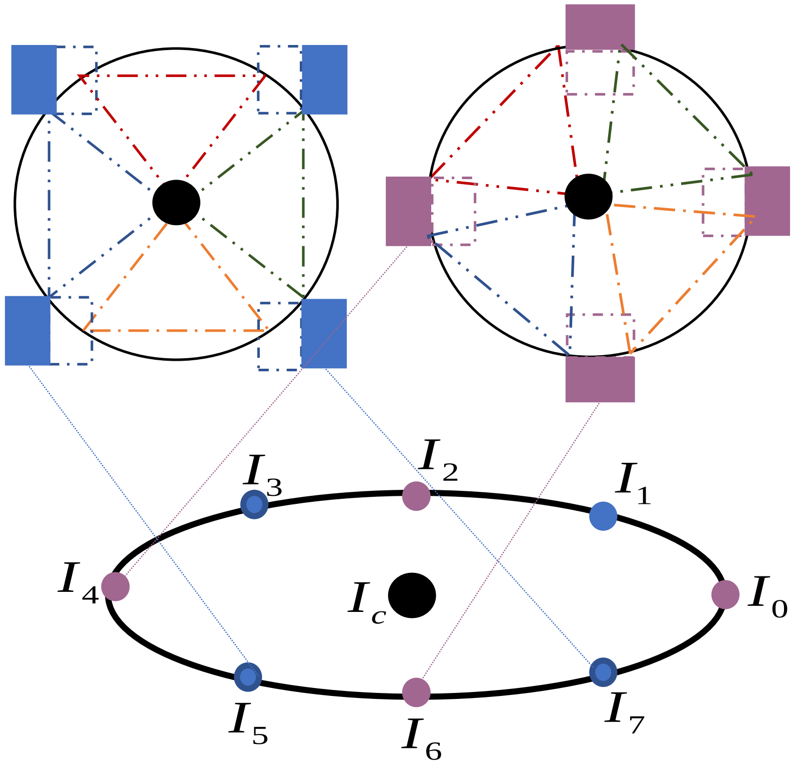
$$\psi(0) = (\lfloor 0/4 \rfloor) \times (8-2) \Rightarrow 0$$

$$Q_{0,1} = \frac{I_7 + I_1}{2}$$

$$Q_{1,1} = \frac{I_1 + I_3}{2}$$

$$Q_{2,1} = \frac{I_3 + I_5}{2}$$

$$Q_{3,1} = \frac{I_5 + I_7}{2}$$



When $\upsilon = 0, w = 0$

$$Q_{0,0} = \frac{I_{(2*0-0+\psi(0)) \bmod 8} + I_{(2(0+1)-0) \bmod 8}}{(\psi(0)/3)+2}$$

$$\psi(0) = (\lfloor 0/4 \rfloor) \times (8-2) \Rightarrow 0$$

$$Q_{0,0} = \frac{I_0 + I_2}{2}$$

$$Q_{1,0} = \frac{I_2 + I_4}{2}$$

$$Q_{2,0} = \frac{I_4 + I_6}{2}$$

$$Q_{3,0} = \frac{I_6 + I_0}{2}$$

Proposed descriptor Cont…

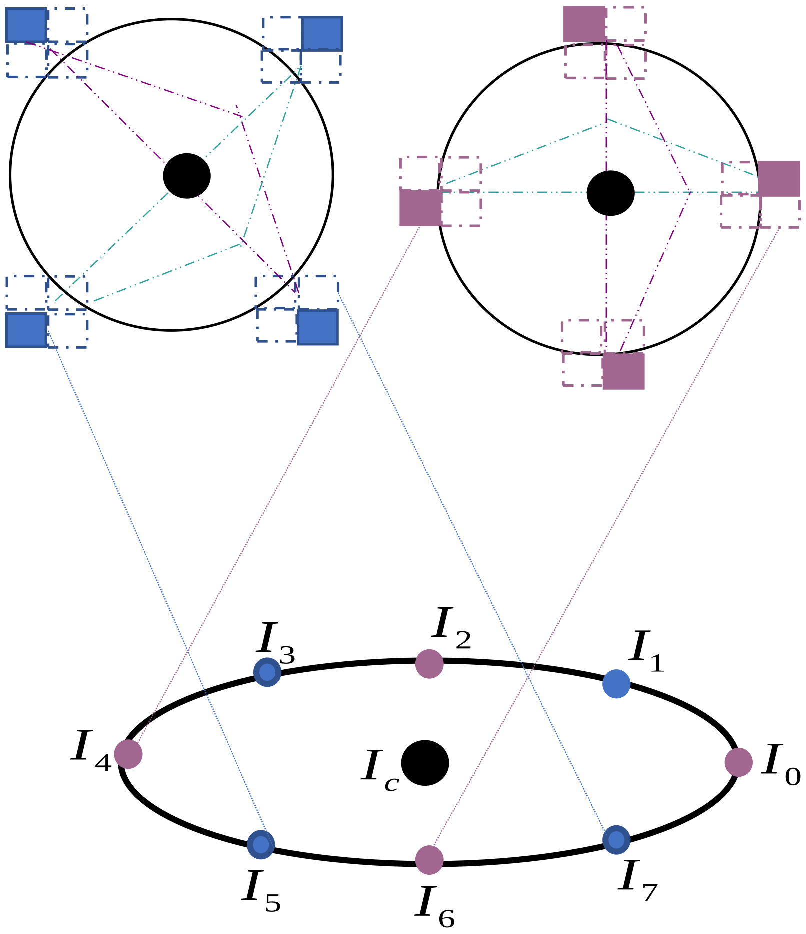
When $\upsilon = 4, \omega = 1$

$$Q_{4,1} = \frac{I_{(2*4-1+\psi(4)) \bmod 8} + I_{(2(4+1)-1) \bmod 8}}{(\psi(4)/3)+2}$$

$$\psi(4) = (\lfloor 4/4 \rfloor) \times (8-2) \Rightarrow 6$$

$$Q_{4,1} = \frac{I_5 + I_1}{4}$$

$$Q_{5,1} = \frac{I_7 + I_3}{4}$$



When $\upsilon = 4, w = 0$

$$Q_{0,0} = \frac{I_{(2*4-0+\psi(0)) \bmod 8} + I_{(2(4+1)-0) \bmod 8}}{(\psi(4)/3)+2}$$

$$\psi(4) = (\lfloor 4/4 \rfloor) \times (8-2) \Rightarrow 6$$

$$Q_{4,0} = \frac{I_6 + I_2}{4}$$

$$Q_{5,0} = \frac{I_4 + I_0}{4}$$

Properties of QUEST

The properties of QUEST are summarized as follows:

1. QUEST encoded the gradient edge information by dividing neighboring pixels into two quadratics to generate six-bit compact code. Thus, generates small feature vector.

2. QUEST extracted the gradient information by utilizing trine pixel relationship, that increases its robustness to noise, pose and lighting changes.

3. Extracted gradient information cohesively describe the disparities among the expression classes.

Comparison between proposed and existing approaches

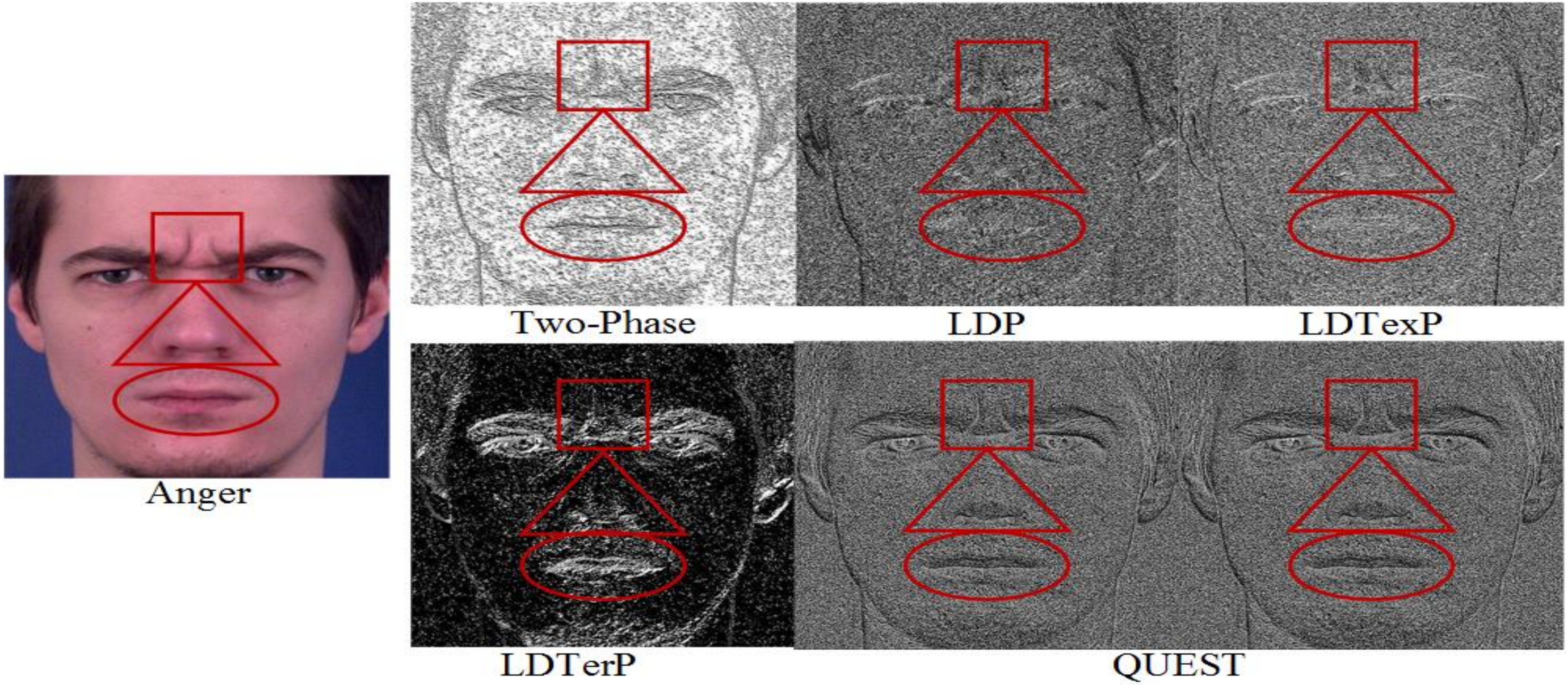


Fig. 3. Coded response results over a anger expression image of the MUG dataset.

Comparison between proposed and existing approaches

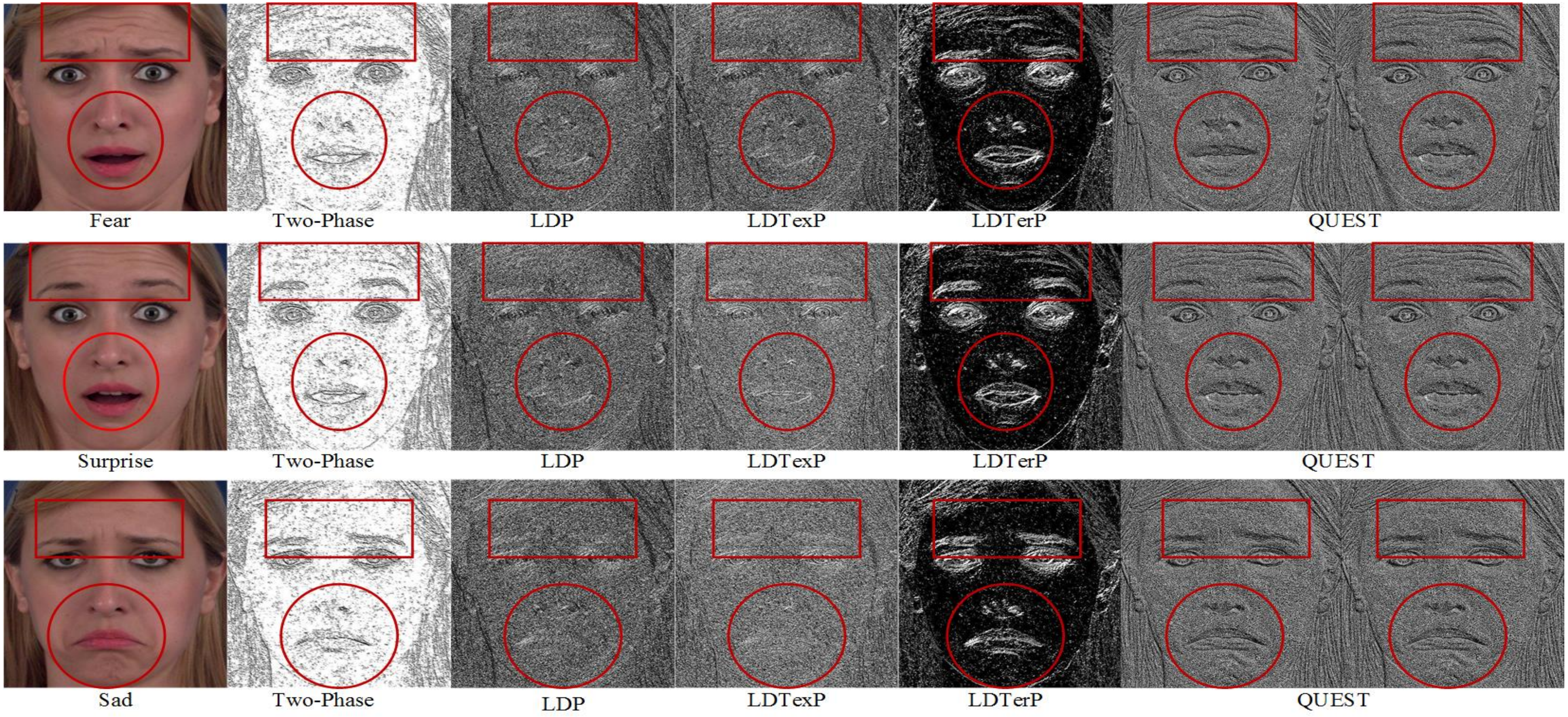


Fig. 4. Coded response results over three different expression (a) fear (b) surprise (c) sad images.

Experimental Results

TABLE I
recognition accuracy comparison on MMI dataset

Methods	6-Class Exp.	7-Class Exp.
LBP [11]	76.5	81.7
Two-Phase [13]	75.4	82.0
LDP [15]	80.5	84.0
LDN [16]	80.5	83.0
LDTexP [17]	83.4	86.0
LDTerP [18]	80.6	80.0
Spatio-Temopral* [20]	81.2	-
QUEST	**83.05**	**84.0**

TABLE II
recognition accuracy comparison on GEMEP-FERA dataset

Methods	5-Class Exp.	6-Class Exp.
LBP [11]	92.2	87.8
Two-Phase [13]	88.6	85.0
LDP [15]	94.0	90.0
LDN [16]	93.4	91.0
LDTexP [17]	94.0	91.8
QUEST	**94.3**	**91.33**

Experimental Results

TABLE III
recognition accuracy comparison on OULU_VIS 6- class expression dataset

Method	6-Class Avg. Accuracy			
	Dark	Strong	Weak	Avg.
LBP [11]	97.6	97.2	97.2	97.3
Two-Phase [13]	94.3	94.1	95.2	94.5
LDP [15]	96.6	97.5	97.9	97.3
LDN [16]	98.3	98.1	98.5	98.3
LDTexP [17]	98.1	98.0	98.2	98.1
LDTerP [18]	98.0	97.8	98.1	98.0
QUEST	**98.6**	**98.2**	**98.2**	**98.3**

TABLE IV
recognition accuracy comparison on OULU_VIS 7- class expression dataset

Method	7-Class Avg. Accuracy			
	Dark	Strong	Weak	Avg.
LBP [11]	96.4	96.9	95.9	96.4
Two-Phase [13]	93.0	92.3	91.3	92.2
LDP [15]	96.0	97.7	97.7	97.1
LDN [16]	96.7	98.1	98.0	97.6
LDTexP [17]	97.8	97.7	97.1	97.5
LDTerP [18]	97.7	96.6	98.2	98.0
QUEST	**98.3**	**98.2**	**98.2**	**98.2**

Experimental Results

TABLE V
recognition accuracy comparison on OULU_NIR 6- class expression dataset

Method	**6-Class Avg. Accuracy**			
	Dark	**Strong**	**Weak**	**Avg.**
LBP [11]	94.1	96.3	96.1	95.5
Two-Phase [13]	80.3	87.8	90.0	86.0
LDP [15]	92.7	98.4	97.2	96.1
LDN [16]	94.3	98.5	96.0	96.2
LDTexP [17]	90.3	98.5	96.6	95.1
LDTerP [18]	93.9	98.3	97.2	96.4
QUEST	**94.5**	**98.5**	**97.9**	**96.9**

TABLE VI
recognition accuracy comparison on OULU_NIR 7- class expression dataset

Method	**7- Class Avg. Accuracy**			
	Dark	**Strong**	**Weak**	**Avg.**
LBP [11]	90.1	93.3	94.1	92.5
Two-Phase [13]	86.2	87.0	89.4	87.5
LDP [15]	94.3	98.0	96.3	96.2
LDN [16]	95.3	97.8	96.7	96.6
LDTexP [17]	95.0	98.3	96.7	96.7
LDTerP [18]	92.4	98.8	96.8	96.0
QUEST	**94.9**	**99.1**	**97.2**	**97.0**

Experimental Results

	ANG	FEA	JOY	REL	SAD
ANG	0.91	0	0.05	0.03	0.02
FEA	0.05	0.93	0	0.02	0
JOY	0.03	0.01	0.94	0.02	0
REL	0.02	0.01	0.01	0.96	0
SAD	0.01	0.02	0	0	0.97

Fig. 5. Confusion matrix of QUEST for 5-class expression classification in GEMEP_FERA dataset

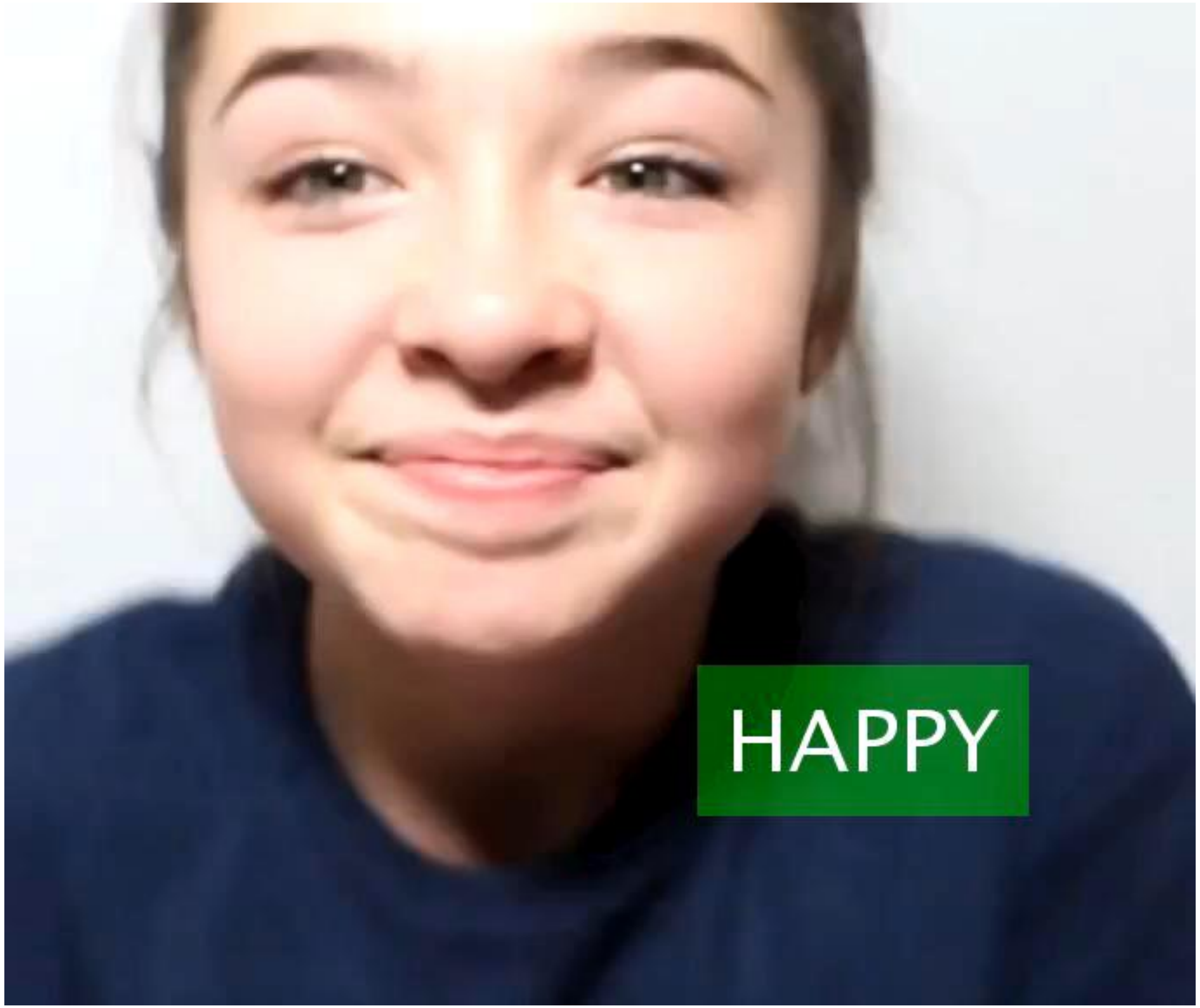
	ANG	FEA	JOY	NEU	REL	SAD
ANG	0.92	0.04	0.01	0.02	0	0.01
FEA	0.02	0.87	0.02	0.04	0.04	0.01
JOY	0.01	0.01	0.96	0.01	0.01	0
NEU	0	0.04	0	0.82	0.03	0.11
REL	0	0.02	0	0.03	0.95	0
SAD	0.04	0	0	0.04	0.02	0.90

Fig. 6. Confusion matrix of QUEST for 6-class expression classification in GEMEP_FERA dataset

Conclusion

- QUEST encoded two six-bit compact codes by thresholding neighboring pixels with reference pixel by dividing the local neighborhood into two quadrics.
- QUEST Extracts the transitional patterns by analyzing pixels located in quadrilaterals, that’s elicited edge variation patterns.
- Quadrilateral structure extracted features of expressive regions and suppress the noise to increase the robustness of the QUEST.

Results on Real time video



References

[1] P. Ekman, “Facial expression and emotion,” American psychologist, vol. 48, no.4, pp.384, 1993.

[2] P. Ekman, and W. V. Friesen, “Facial action coding system,” 1977.

[3] Y.L. Tian, T. Kanade and J.F. Cohn, “Facial expression analysis,” In Handbook of face recognit., New York; Springer-vergal, pp. 247- 275, 2005

[4] H. Hong, H. Neven and C. Von der Malsburg, “Online facial expression recognition based on personalized galleries,” In Proc. 3rd IEEE Int. Conf. Autom. Face and Gesture Recognit., pp. 354- 359,1998.

[5] I. Kotsia and I. Pitas, “Facial expression recognition in image sequences using geometric deformation features and support vector machines,” IEEE Trans. Image Process., vol. 16, no. 1, pp.172-187, 2007.

[6] J. Yang, D. Zhang, A. F. Frangi and J. Y. Yang, “Two-dimensional PCA: a new approach to appearance-based face representation and recognition,” IEEE Trans. Pattern Anal. Mach Intell., vol. 26, no. 1, pp.131-137, 2004.

[7] K. Etemad and R. Chellappa, “Discriminant analysis for recognition of human face images,” JOSA A, vol. 14, no. 8, pp.1724-1733,1997.

[8] M. S. Bartlett, G. Littlewort, I. Fasel and J. R. Movellan, “Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, vol. 5, pp. 53-53, 2003.

[9] D. Gabor, “Theory of communication,” J. Inst. Electr. Eng. III, Radio Commun. Eng., vol. 93, no. 26, pp. 429-425, 1946.

[10] L. Wiskott, J.-M. Fellous, N. Kuiger and C. von der Malsburg, “Face recognition by elstic bunch graph matching,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 7, pp. 775-779, 1997.

[11] C. Shan, S. Gong and P. W. McOwan, “Facial expression recognition based on local binary patterns: A comprehensive study,” Image Vis. Comput., vol. 27, no. 6, pp.803-816.

[12] X. Tan and B. Triggs, “Enhanced local texture feature sets for face recognition under difficult lighting conditions,” IEEE Trans. Image Process., vol. 19, no. 6, pp. 1635–1650, 2010.

[13] C. C. Lai and C. H. Ko, “Facial expression recognition based on twostage features extraction.,” Optik-Int. J. Light Electron Optics, vol. 125, no. 22, pp.6678-6680, 2014.

[14] S. Yang, and B. Bhanu, “Facial expression recognition using emotion avatar image,” In Proc. IEEE Int. Conf. Autom. Face and Gesture Recognit., pp. 866-871,2011.

References

[15] T. Jabid, M. Kabir and O. O. Chae, “Robust facial expression recognition based on local directional pattern,” J. ETRI, vol. 32, no. 5, pp.784-794,2010. [16] A. R. Rivera, J. R. Castillo and O. O. Chae, “Local directional number pattern for face analysis: Face and expression recognition,” IEEE Trans. Image Process., vol. 22, no. 5, pp.1740- 1752, 2013.

[17] A. R. Rivera, J. R. Castillo and O. O. Chae, “Local directional texture pattern image descriptor,” Pattern Recognit. Letters, vol. 51, pp.94- 100, 2015.

[18] B. Ryu, A. R. Rivera, J. Kim and O. O. Chae, “Local directional ternary pattern for facial expression recognition,” IEEE Trans. Image Process., vol. 26, no. 12, pp. 6006-6018, 2017.

[19] H. Ding, S. K. Zhou and R. Chellappa, “Facenet2expnet: Regularizing a deep face recognition net for expression recognition,” in proc. 12th IEEE Int. Conf. Automatic Face Gesture Recognit., pp. 118-126, 2017.

[20] K. Zhang, Y. Huang, Y. Du and L. Wang, “Facial expression recognition based on deep evolutional spatial-temporal networks,” IEEE Trans. Image Proc., vol. 26, no. 9, pp. 4193-4203, 2017

[21] C. Cortes and V. Vapnik, “Support-vector networks,” Machine learning, vol. 20, no. 3, pp.273-297, 1995.

[22] M. Pantic, M. Valstar, R. Rademaker and L. Maat, “Web-based database for facial expression analysis,” in Proc. IEEE Conf. Multimed. Expo., pp. 200-205, 2005.

[23] M. Valstar and M. Pantic, “Induced disgust, happiness and surprise: an addition to the mmi facial expression database,” in Proc. 3rd Int. Workshop. Emot. (Satellite of LREC), pp. 65-70, 2010.

[24] M. F. Valstar, B. Jiang, M. Mehu, M. Pantic and K. Scherer, “The first facial expression recognition and analysis challenge,” in Proc. IEEE Int. Conf. Automatic Face & Gesture Recognit. Workshops, pp. 921-926, 2011.

[25] G. Zhao, X. Huang, M. Taini, S.Z. Li and M. Pietikälnen, “Facial expression recognition from near-infrared videos,” Image Vis. Comput., vol. 29, no. 9, pp.607-619, 2011.

[26] S. L. Happy, P. Patnaik, A. Routray and R. Guha, “The Indian spontaneous expression database for emotion recognition,” IEEE Trans. Affect. Comput., vol. 8, no. 1, pp.131-142, 2017.

Thank You