Performance Evaluation of Stationary Wavelet Features with Least Squares SVM for Facial Expression Recognition

Nikunja Bihari Kar and Korra Sathya Babu

Abstract This paper presents an improved method for recognition of facial expressions. The facial expression images are first decomposed using 3-level stationary wavelet transform (SWT). The feature vectors are then derived from the coefficients of high-frequency SWT sub-bands. Moreover, these high-frequency coefficients are useful to retain the edge information from expression images. SWT overcomes the issue of translation variant that traditional discrete wavelet transform (DWT) suffers. Also, SWT performs considerably well on translated images. To generate a set of compressed and discriminant features, linear discriminant analysis + principal component analysis (LDA+PCA) is applied. Finally, least squares SVM is used for the classification. Japanese female facial expression (JAFFE) and the Extended Cohn-Kanade (CK+) datasets are used to evaluate the system proposed. The proposed method achieves an accuracy of 98.72% and 98.63% on JAFFE and CK+ datasets, respectively. Experimental results based on 5-fold cross-validation test indicate the superiority of the proposed system over state-of-the-art schemes. Besides, the effectiveness of the SWT features is compared with the DWT features.

Keywords: Discrete wavelet transform, principal component analysis, least squares support vector machine, stationary wavelet transform

Nikunja Bihari Kar
Department of Computer Science and Engineering, National Institute of Technology Rourkela, Odisha-769 008, India. e-mail: nikunjakar@gmail.com

Korra Sathya Babu
Department of Computer Science and Engineering, National Institute of Technology Rourkela, Odisha-769 008, India. e-mail: ksathyababu@nitrkl.ac.in

1 Introduction

Emotion imparts a vital role in our real-life interaction and communication. The capacity to perceive the emotional state of individuals surrounding us is an essential part of natural conversation. Affective computing is the study of systems and devices. It can translate, process, and perceive human affects. The key benefit of affective computing is stress analysis, sentiment analysis, and intelligent systems, etc. But the main drawback of it is to make out emotion from facial expressions, or body gesture, or speech, or electroencephalogram (EEG).

Facial expression analysis has become an important research area in psychology, and a large body of works have been published in the past years. However, the performance still requires potential improvement to meet real-time demand. In general, facial expression recognition (FER) system works in three stages: preprocessing, feature extraction and classification. Selection of proper features plays an important role in final decision making. In the existing system, a large number of schemes have been used to extract the features. Refer [1] to see a list of FER systems. The feature extraction techniques used by existing FER systems are discrete wavelet transform (DWT) [2], local binary pattern (LBP) [3], local ternary pattern (LTP), stationary wavelet transform (SWT) [4], ripplet transform type II (ripplet-II) [5], Gabor wavelets [6], and histograms of oriented gradients (HOG) [7]. These feature extraction techniques are utilized to extract features either from the entire face or some portion of the face with no earlier information.

The past decade has witnessed a significant amount of research on FER. For instance, Happy and Routray [8] developed a novel FER framework using appearance features of different facial patches. Siddiqi *et al.* [9] harnessed stepwise linear discriminant analysis (SWLDA) and hidden conditional random fields (HCRFs) for recognizing emotion category. In the work of Mlakar and Potocnik [10] the HOG features are calculated for neutral and peak expression images. Then the SVM classifier uses the difference vector for recognition. A novel FER approach based on HOG and PCA+LDA is proposed in [7] for feature extraction and reduction. After that backpropagation neural network (BPNN) is appended for expression classification.

In FER systems various multi-resolution analysis techniques have been proposed. Zhang and Tjondronegoro [11] considered facial movement and muscle movement features for recognition, which is obtained from patch-based 3D Gabor features. This technique demonstrates promising outcomes for face registration faults with fast processing time. Kazmi *et al.* [12] combined 3 level 2D-DWT with seven parallel SVMs to detect emotion. Wang *et al.* [13] presented a hybrid strategy to extract texture and shape features using Weber local descriptor (WLD) and HOG descriptor. Finally, k-nearest neighbors (kNN) is harnessed as the classifier. Siddiqi and Lee [14] employed Symlet wavelet for feature extraction. The use of LDA reduces the feature set. Then, HMM is harnessed for expression labeling. In the work of Uccar *et al.* [15] curvelet features are extracted from each local face regions. From each curvelet regions, textural features such as mean, standard deviation and entropy were calculated. Online sequential extreme learning machine-spherical clustering (OSELM-SC) is used to classified expression. Zhang *et al.* [16] developed a FER system using biorthogonal wavelet entropy (BWE). A SVM variant called Fuzzy SVM is trained for emotion classification. Wang *et al.* [4] harnessed stationary wavelet entropy (SWE) and Jaya learning algorithm to develop an intelligent FER system. For emotion labeling, single hidden layer feed-forward neural network (SHLFNN) is used. Qayyum *et al.* [17] used SWT and discrete cosine transform (DCT) for features and BPNN for emotion classification. In previous existing works, it has been noticed that the wavelet transform (WT) technique is an ultimate and widespread feature extraction technique for systems [4, 14, 16, 17]. But, due to some shortcoming of non-supportiveness to anisotropy and restricted directionality, the WT technique cannot extract the intrinsic and subtle features of the expression images, which is essential for emotion detection.

From the literature mentioned above, many FER techniques used DWT as a tool for feature extraction. In nature DWT is translation-variant and is used by many FER techniques to extract features, as suggested by literature. SWT can deal with the shift invariance problem. Hence in this work, SWT algorithm is used for feature extraction. SWT algorithm is a variant of DWT designed to overcome the deficiency of translation invariant property of DWT. To achieve translation-invariant property the upsamplers and downsamplers are removed from DWT coefficients. Then, the filter coefficients are upsampled by a factor of of $2(j-1)$ in the j^{th} level. Originally, the SWT is a redundant method. The dimension of the SWT coefficients at each level are same as the input sample. Thus, there are N redundant wavelet coefficients for N level decomposition.

The proposed FER system has the following characteristics,

- We introduced a new combined approach to FER.
- The proposed method used SWT as features which have three significant advantages over DWT [18].
  - The obtained SWT coefficients will not change regardless of the shift of the signal, due to translation-invariant property.
  - SWT performance is better for denoising and edge-detecting tasks.
  - SWT is applied to any arbitrary size image while DWT can only be applied to images whose size is a power of 2.
- The dimension of the SWT features is quite high. To reduce the feature dimension and improve the computational complexity PCA+LDA method is harnessed.
- Traditional SVM classifier incurred higher computational complexity. So to avoid high computation, a lighter SVM called least squares support vector machine (LS-SVM) is employed.

The remaining structure of the paper is organized in following way. In Section 2, we defined the datasets. Each step of the proposed system is described in Section 3. Section 4 presents the results and discussions. Finally, Section 5 concluded the paper.

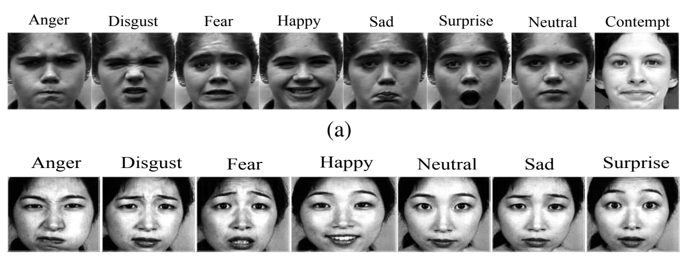
2 Dataset

In order to validate the system proposed two benchmark datasets such as CK+ [19] and JAFFE [20] are used.

The various images chosen to be used in our experiment is listed in Table 1. The CK+ dataset consists of 593 image sequences from 123 subjects. Among which 327 images are labelled with seven basic emotion categories. JAFFE dataset comprises 213 expression images collected from 10 individuals. In our experiment, we have used 213 and 450 images from JAFFE and CK+ database respectively. The sample preprocessed images for both the dataset used in the experiment are shown in Fig. 1.

Table 1: Images from each emotions for JAFFE and CK+ dataset of the experiment

	Anger	Disgust	Fear	Happy	Sad	Surprise	Neutral	Contempt	All
JAFFE	30	29	32	31	31	30	30	-	213
CK+	45	59	25	69	28	83	123	18	450



(a)

(b)

Fig. 1: Some normalized image samples resulted from (a) CK+ and (b) JAFFE dataset for the experiment

3 Proposed Methodology

The proposed methodology follows in four main phases such as (i) pre-processing of face images, (ii) feature extraction, (iii) reduction in feature dimension, and (iv) classification phase. Fig. 2 reveals the outline of the system proposed. The details of each flow of the scheme are reported in the next section.

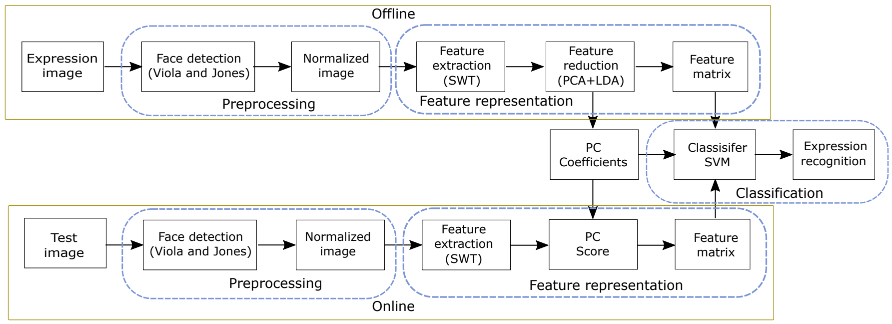


Fig. 2: Framework of the proposed scheme for recognizing emotion from expression images

3.1 Preprocessing

Initially, all the input images are changed over to grayscale images. At that point, image contrast increased by eliminating the 1% top and bottom of all pixel values. Then, Viola and Jones algorithm [21] is used for face identification. Afterwards, the identified face area is cropped and normalized to 128×128 size images.

3.2 Feature extraction using 2D SWT

In this study, we have used all the high-frequency SWT sub-bands decomposed up to 3-levels as features. SWT is a variant of DWT designed to overcome the deficiency of translation invariant property of DWT. To achieve translation-invariant property the up-samplers and down-samplers are removed from DWT coefficients. In the j^{th} level, a $2(j-1)$ factor is used to upsample the filter coefficients. The dimension of the input sample is the same at each level, as SWT is a redundant method. In our study, SWT with three level Haar wavelet is utilized to decompose the face image. The high-frequency SWT sub-bands are retained to analyze the edge and texture features of the image. The approximation sub-bands of SWT carries little energy for which it is not considered for texture analysis.The results of 2D-SWT decomposition up to three-level is shown in Fig. 3, in which 9 SWT sub-bands are stored as D1, D2, D3, H1, H2, H3, V1, V2, and V3.

3.3 Feature reduction using PCA+LDA

We observed that all 3-level SWT high-frequency coefficients are very large to be considered as desired features. Therefore, it is necessary to make the feature size

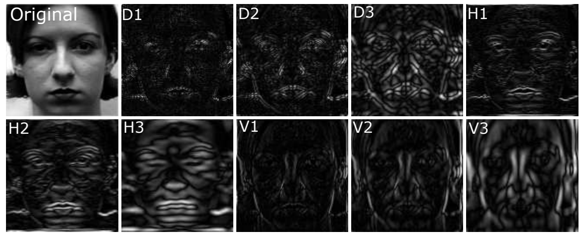


Fig. 3: 3-level SWT decomposition of a neutral face image indicating all high frequency components

small, which will result in a more convenient recognition task. In this work the dimensionality reduction of features achieved by PCA+LDA technique. Moreover, it stores the informative features from high dimensional SWT features.

Both PCA and LDA has been broadly used to reduce the dimensionality of the features. PCA does the orthogonal projection of information onto a lower dimensional linear space [22]. This linear space is called principal subspace. The aim is to maximize the variance of the projected data. LDA looks for vectors in hidden space that best separate among classes as opposed to data. In other words, data is described by numerous independent features. In our study, linear discriminant analysis (LDA) approach [23] is adopted to enhance the class separability of all the expression images. It finds the most discriminant projection vectors which map high dimensional feature space into low dimensional feature space. LDA projection vectors assist every single projected sample to form maximum between class scatter and minimum within-class scatter.

From our analyses, it is evident that if just LDA is utilized for feature reduction, at some point the within scatter matrix become singular. This issue emerges if the quantity of samples is significantly less than the feature dimension. This problem is termed as the small sample size (SSS) problem. PCA technique is used for feature reduction before LDA to eliminate this problem. PCA is used to project n-dimensional feature space into an intermediate p dimensional space. Afterward, the LDA is attached to reduce the features from p to m-dimensional $(m << p << n)$. In this work, high dimensional featuresreduced to considerably low dimension i.e. 6 features for CK+ and JAFFE dataset. An improved accuracy and lower computational complexity is achieved with these reduced features.

The number of features to be reduced has been decided according to a metric called the normalized cumulative sum of variances (NCSV). This metric can be described as follows,

$$NCSV(j) = \frac{\sum_{u=1}^{j} \alpha(u)}{\sum_{u=1}^{p} \alpha(u)} \qquad ; 1 \leq j \leq p \tag{1}$$

Where, p represents the feature dimension. It is worth noticing that the eigenvalues constructed to each feature are first arranged in decreasing order. Finally, a threshold is set to choose the reduced dimension of the feature vector.

3.4 Classification using LS-SVM

After feature reduction phase, the feature matrix is fed to the classifier for prediction of emotion. The performance of conventional SVM is low when it deals with massive datasets. Also, the computational overhead is more. To improve accuracy and computational complexity, LS-SVM [24] is utilized in this scheme, which is a more capable variant of SVM. LS-SVM supervised approach is used linear or nonlinear hyperplanes for classification of samples from two or more hyperplanes. Various experiments carried out using LS-SVM with kernels such as RBF, linear, and polynomial. Experimental results indicate that LS-SVM with RBF kernel outperforms other kernels.

Let there be N samples $\{p_i, q_i\}_{i=1}^{N}$. where, $p_i \in R^n$ is the i^{th} input data and $q_{i \in R}$ is the i^{th} output label. The LS-SVM classifier decision function can be defined as,

$$f(x) = sign\left[\sum_{i=1}^{N} \alpha_i y_i \kappa(x, x_i) + b\right] \tag{2}$$

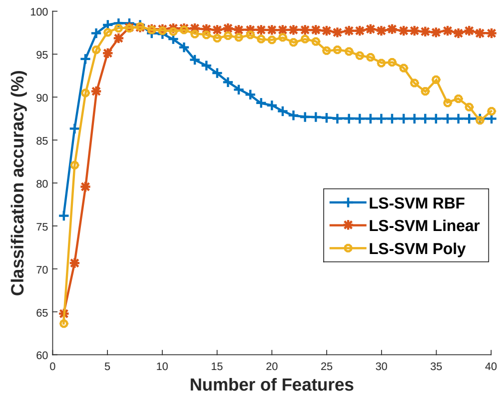
where, $\kappa(.,.)$ is the kernel function, and α_i is the Lagrange multiplier.

4 Results and Discussion

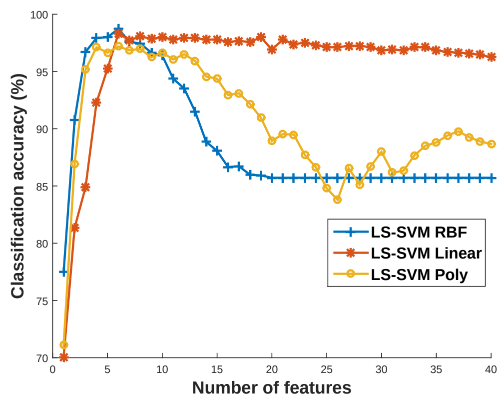
The experiments were observed on a 3.4 GHz Core i7 processor and 8GB RAM PC running under Windows OS framework. Matlab tool was used to simulate the proposed system.

4.1 Feature extraction and reduction

At first, Viola and Jones algorithm was used to recognize the face area from the original sample. Then, the detected face region was cropped, and then it is normalized. The SWT with three level haar wavelet is used to obtained features from expressive images. The results of 2D-SWT decomposition up to three-level is shown in Fig. 3, in which 9 SWT sub-bands are stored as D1, D2, D3, H1, H2, H3, V1, V2, and V3. The aggregate number of features extracted by three-level SWT decomposition with haar wavelet is $128 \times 128 \times 9 = 147456$. This large number makes the learning process difficult of any current classifier. The number of features come down to 6 with the use of PCA+LDA for CK+ and JAFFE dataset.



(a)



(b)

Fig. 4: Performance measure of (a) CK+ and (b) JAFFE dataset in terms of number of features

The results obtained from PCA+LDA on both the datasets are shown in Fig. 4. It is observed that the highest accuracy has been achieved by using only six PCs because it preserves maximum variance. Hence for performance comparison, we have used only six features in our experiment.

The classification results of LS-SVM RBF on CK+ and JAFFE datasets are presented in Tables 2 and 3. To demonstrate the performance of the LS-SVM classifier along other classifiers like KNN, RF, and BPNN have also been implemented. All the classifiers are trained and tested with the same set of data obtained from LDA.

The classifiers utilize 5-fold stratified cross-validation (SCV) to make the system more stable and generalize to other datasets. Parameters for different classifiers have been tuned experimentally. Finally, the parameters with least SCV error are chosen.

Table 2: Classification results obtained by LS-SVM RBF on CK+ 8 dataset

Expressions	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy (%)
Anger	39	397	8	6	86.00	98.00	96.88
Contempt	13	429	3	5	72.00	99.00	98.22
Disgust	58	390	1	1	98.00	99.00	99.55
Fear	25	425	0	0	100.00	100.00	100.00
Happy	68	380	1	1	98.00	99.00	99.55
Sad	26	420	2	2	92.00	99.00	99.11
Surprise	82	367	0	1	98.00	100.00	99.77
Neutral	115	317	10	8	93.00	96.00	96.00
Average performances:							98.63

Table 3: Classification results obtained by LS-SVM RBF on JAFFE 7 dataset

Expressions	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy (%)
Anger	28	183	0	2	93.00	100.00	99.06
Disgust	28	183	1	1	96.00	99.00	99.06
Fear	29	179	2	3	90.00	98.00	97.65
Happy	30	181	1	1	96.00	99.00	99.06
Neutral	30	182	1	0	100.00	99.00	99.53
Sad	27	180	2	4	87.00	98.00	97.18
Surprise	29	183	0	1	96.00	100.00	99.53
Average performances:							98.72

Table 4: Results of classification by different classifier (CK+ 8 dataset)

Classifiers	Sensitivity	Precision	F-Measure	Accuracy(%)
LS-SVM RBF	0.92	0.93	0.92	98.63
LS-SVM Linear	0.95	0.82	0.87	96.83
LS-SVM Poly	0.95	0.87	0.91	98.05
KNN	0.93	0.93	0.93	93.55
RF	0.93	0.93	0.93	93.11
BPNN	0.94	0.94	0.94	94.88

The average sensitivity, average precision, average F-measure, and the average accuracy is obtained from five-fold stratified CV procedure for the proposed system is shown in Table 4 and Table 5 on CK+ and JAFFE dataset respectively. Experimental results reported in Tables 4 and 5, indicate that the performance of LS-SVM with RBF kernel obtained higher accuracy than other classifiers on both the datasets.

Table 5: Results of classification by different classifier (JAFFE 7 dataset)

Classifiers	Sensitivity	Precision	F-Measure	Accuracy(%)
LS-SVM RBF	0.94	0.96	0.95	98.72
LS-SVM Linear	0.96	0.92	0.94	98.32
LS-SVM Poly	0.97	0.85	0.90	97.18
KNN	0.95	0.95	0.95	95.77
RF	0.93	0.93	0.93	93.89
BPNN	0.93	0.93	0.93	93.42

Table 6: Performance comparison of DWT and SWT features on CK+ and JAFFE datasets

Features	Classifiers	Accuracy(%)	
		CK+	JAFFE
DWT+PCA+LDA	LS-SVM RBF	97.36	95.03
	LS-SVM Linear	95.61	94.90
	LS-SVM Poly	96.50	93.17
SWT+PCA+LDA	LS-SVM RBF	98.63	98.72
	LS-SVM Linear	96.83	98.32
	LS-SVM Poly	98.05	97.18

Therefore we select LS-SVM RBF as the classifier for this study. Table 6 shows the superiority of SWT features over DWT features on the CK+ and JAFFE datasets. It also demonstrates the classification results of LS-SVM classifier with different kernels. Finally, we compare our proposed approach with other six state-of-the-art methods on both CK+ and JAFFE datasets. It is observed from Table 7 that our method achieves highest classification accuracy on JAFFE dataset and achieves second best result on CK+ dataset.

Table 7: Performance comparison between proposed and reported systems on CK+ and JAFFE dataset

Reference	**Facial feature**	**Classifier**	**Accuracy**(%)	
			CK+	JAFFE
Zhang and Tjondronegoro [11]	Patch based Gabor	SVM	94.48	99.23
Wang *et al.* [13]	HOG+WLD	KNN	93.97	95.86
Mlakar and potocinik [10]	HOG difference vector	SVM	95.64	87.82
Happy *et al.* [8]	Salient facial patches	RBF SVM	94.09	91.79
Siddiqi *et al.* [9]	SWLDA	HCRF	96.83	96.33
Uccar *et al.* [15]	Local Curvlet transform	OSELM-SC	95.17	94.65
Proposed method	**SWT+PCA+LDA**	**LS-SVM RBF**	**98.63**	**98.72**

5 Conclusion

This work suggested a more practical approach for recognition of facial expressions. SWT was applied to decompose the preprocessed image into several sub-bands. The high-frequency elements of SWT were then collected to form the feature vector. PCA+LDA method was utilized to derive a discriminant feature set from high dimensional SWT coefficients. Finally, the resultant elements were subjected to the LS-SVM classifier for detecting emotions. The experimental results on benchmark datasets showed that the proposed scheme achieved higher detection accuracy as compared to the state-of-the-art methods. The performance of SWT features showed the superiority in comparison to DWT features.

In the proposed method, 3-level Haar wavelet was used for decomposition of images using SWT; however, choosing the optimal decomposition level and mother wavelet, remains a concern. The performance of the proposed method needs to be verified over a larger and diverse dataset. Various efficient schemes for feature extraction and machine learning can be applied to the existing proposed system, to enhance the performance.

References

1. Pantic, M., Rothkrantz, L.J.: Automatic analysis of facial expressions: The state of the art. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(12), 1424–1445 (2000)
2. Dongcheng, S., Jieqing, J.: The method of facial expression recognition based on DWT-PCA/LDA. In: International Congress on Image and Signal Processing (CISP). vol. 4, pp. 1970–1974. IEEE (2010), doi:10.1109/CISP.2010.5648166
3. Gritti, T., Shan, C., Jeanne, V., Braspenning, R.: Local features based facial expression recognition with face registration errors. In: 8th IEEE International Conference on Automatic Face & Gesture Recognition, FG'08. pp. 1–8. IEEE (2008), doi:10.1109/AFGR.2008.4813379
4. Wang, S.H., Phillips, P., Dong, Z.C., Zhang, Y.D.: Intelligent facial emotion recognition based on stationary wavelet entropy and jaya algorithm. Neurocomputing 272, 668–676 (2018), doi:10.1016/j.neucom.2017.08.015
5. Kar, N.B., Babu, K.S., Sangaiah, A.K., Bakshi, S.: Face expression recognition system based on ripplet transform type ii and least square svm. Multimedia Tools and Applications pp. 1–24 (2017), doi:10.1007/s11042-017-5485-0
6. Bartlett, M.S., Littlewort, G., Frank, M., Lainscsek, C., Fasel, I., Movellan, J.: Fully automatic facial action recognition in spontaneous behavior. In: 7th International Conference on Automatic Face and Gesture Recognition. pp. 223–230. IEEE (2006), doi:10.1109/FGR.2006.55
7. Kar, N.B., Babu, K.S., Jena, S.K.: Face expression recognition using histograms of oriented gradients with reduced features. In: Proceedings of International Conference on Computer Vision and Image Processing. pp. 209–219. Springer (2017), doi:10.1007/978-981-10-2107-7_19
8. Happy, S., Routray, A.: Automatic facial expression recognition using features of salient facial patches. IEEE transactions on Affective Computing 6(1), 1–12 (2015), doi:10.1109/TAFFC.2014.2386334
9. Siddiqi, M.H., Ali, R., Khan, A.M., Park, Y.T., Lee, S.: Human facial expression recognition using stepwise linear discriminant analysis and hidden conditional random fields. IEEE

Transactions on Image Processing 24(4), 1386–1398 (2015), doi:10.1109/TIP.2015.2405346
10. Mlakar, U., Potočnik, B.: Automated facial expression recognition based on histograms of oriented gradient feature vector differences. Signal, Image and Video Processing pp. 1–9 (2015), doi:10.1007/s11760-015-0810-4
11. Zhang, L., Tjondronegoro, D.: Facial expression recognition using facial movement features. IEEE Transactions on Affective Computing 2(4), 219–229 (2011), doi:10.1109/T-AFFC.2011.13
12. Kazmi, S.B., Jaffar, M.A., et al.: Wavelets-based facial expression recognition using a bank of support vector machines. Soft Computing 16(3), 369–379 (2012), doi:10.1007/s00500-011-0721-4
13. Wang, X., Jin, C., Liu, W., Hu, M., Xu, L., Ren, F.: Feature fusion of HOG and WLD for facial expression recognition. In: International Symposium on System Integration (SII). pp. 227–232. IEEE (2013), doi:10.1109/SII.2013.6776664
14. Siddiqi, M.H., Lee, S.: Human facial expression recognition using wavelet transform and hidden markov model. In: International Workshop on Ambient Assisted Living. pp. 112–119. Springer (2013), doi:10.1007/978-3-319-03092-0
15. Uçar, A., Demir, Y., Güzeliş, C.: A new facial expression recognition based on curvelet transform and online sequential extreme learning machine initialized with spherical clustering. Neural Computing and Applications 27(1), 131–142 (2016), doi:10.1007/s00521-014-1569-1
16. Zhang, Y.D., Yang, Z.J., Lu, H.M., Zhou, X.X., Phillips, P., Liu, Q.M., Wang, S.H.: Facial emotion recognition based on biorthogonal wavelet entropy, fuzzy support vector machine, and stratified cross validation. IEEE Access 4, 8375–8385 (2016), doi:10.1109/ACCESS.2016.2628407
17. Qayyum, H., Majid, M., Anwar, S.M., Khan, B.: Facial expression recognition using stationary wavelet transform features. Mathematical Problems in Engineering 2017 (2017), doi:10.1155/2017/9854050
18. Gnanadurai, D., Sadasivam, V.: Undecimated wavelet based speckle reduction for SAR images. Pattern Recognition Letters 26(6), 793–800 (2005)
19. Lucey, P., Cohn, J.F., Kanade, T., Saragih, J., Ambadar, Z., Matthews, I.: The Extended Cohn-Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). pp. 94–101. IEEE (2010), doi:10.1109/CVPRW.2010.5543262
20. Lyons, M., Akamatsu, S., Kamachi, M., Gyoba, J.: Coding facial expressions with Gabor wavelets. In: Third IEEE International Conference on Automatic Face and Gesture Recognition. pp. 200–205. IEEE (1998), doi:10.1109/AFGR.1998.670949
21. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR. vol. 1, pp. I–511. IEEE (2001)
22. Bishop, C.M.: Pattern recognition and machine learning. springer (2006)
23. Yu, H., Yang, J.: A direct LDA algorithm for high-dimensional datawith application to face recognition. Pattern recognition 34(10), 2067–2070 (2001)
24. Suykens, J., Vandewalle, J.: Least squares support vector machine classifiers. Neural Processing Letters 9(3), 293–300 (1999), doi:10.1023/A:1018628609742