Multiple Local Curvature Gabor Binary Patterns for Facial Action Recognition

Anıl Yüce, Nuri Murat Arar and Jean-Philippe Thiran

Signal Processing Laboratory (LTS5),
École Polytechnique Fédérale de Lausanne, Switzerland
{anil.yuce,murat.arar,jean-philippe.thiran}@epfl.ch

Abstract. Curvature Gabor features have recently been shown to be powerful facial texture descriptors with applications on face recognition. In this paper we introduce their use in facial action unit (AU) detection within a novel framework that combines multiple Local Curvature Gabor Binary Patterns (LCGBP) on different filter sizes and curvature degrees. The proposed system uses the distances of LCGBP histograms between neutral faces and AU containing faces combined with an AU-specific feature selection and classification process. We achieve 98.6% overall accuracy in our tests with the extended Cohn-Kanade database, which is higher than achieved previously by any state-of-the-art method.

1 Introduction

Being the most important non-verbal means of human communication, facial expressions have gained extreme importance for computer vision especially with the ease of access of recent technological developments. Besides the various applications of basic emotion recognition, recognizing facial actions now serves much more in areas such as intelligent human computer interaction or diagnosis and treatment of certain pathological conditions.

The recent interesting work, such as [11], which investigates the detection of expressions of pain, or [17], which presents a method for the detection of asymmetric lip movements, for instance, all point out to the need for increased facial action recognition accuracy to capture more subtle muscle movements in both intensity and time. To this end, we introduce a framework with a novel set of features for action unit (AU) detection that achieves much higher accuracy than previously presented state-of-the-art methods.

The AUs are the basic units of facial movement, that are defined by the Facial Action Coding System (FACS) [7]. FACS serves as a method to objectively define every independent motion on the face. For an overview of recent advances and the state-of-the-art in AU and facial expression detection, the reader is referred to [19], which is the meta-analysis of the first facial expression recognition challenge and includes the summary of up to date work using shape and appearance based methods.

In this work we propose the variation among frames of a combination of Local Curvature Gabor Binary Patterns (LCGBP) as descriptors of facial action. LCGBP is an extension to the Local Gabor Binary Patterns (LGBP) which have been used extensively for face recognition and AU detection (e.g. [22], [16]), since they have proven

to be quite robust against variations of conditions such as illumination. By adding the effect of curved formations, which exist commonly in the facial texture, the curvature Gabors provide a more efficient way of representing the facial components [8] and it was previously shown in our recent work that they are successful in recognizing facial identity [2]. Here, we apply this idea by using the change in LCGBP histograms between neutral and expressive images for detecting the AUs. Using this variation of histograms between frames has shown to be more efficient than using the histograms themselves directly ([16], [21]). The main contribution of our work is introducing a unique way of extracting Gabor features, which includes the curvature information and proves by the very high accuracy results to be very powerful descriptors for facial actions, combined with a feature selection and classification phase that was proven in our previous work [21] to be efficient with such features.

The rest of the paper is formed as follows: In Section 2 we explain the formulation of LCGBP, in Section 3 we describe the framework that we propose for AU detection and detail the parameter selection. Section 4 presents the test settings and results obtained by several experiments on the CK+ database and comparisons with other types of features and recent existing methods in the literature. Finally, we report our conclusions and possible future directions for further improving the system in Section 5.

2   Local Curvature Gabor Binary Patterns

2.1   Curvature Gabor (CG) Wavelets

Gabor wavelets have been recognized as one of the most successful feature extraction methods for face representation. They form a well-established image decomposition because of their spatial locality and orientation selectivity characteristics. Therefore, they are optimally localized in the space and frequency domains, and can be used successfully in facial image processing for face and facial expression recognition and analysis.

The conventional Gabor wavelet definition is as follows:

$$\psi(\boldsymbol{x};\nu,\mu) = \frac{k_{\nu,\mu}^2}{\sigma^2} e^{\left(-\frac{k_{\nu,\mu}^2 ||\boldsymbol{x}||^2}{2\sigma^2}\right)} [e^{(ik_{\nu,\mu}\boldsymbol{x})} - e^{(-\frac{\sigma^2}{2})}] \tag{1}$$

where $\boldsymbol{x} = \begin{pmatrix} \acute{x} \\ \acute{y} \end{pmatrix} = \begin{pmatrix} xcos(\frac{\mu\pi}{M}) + ysin(\frac{\mu\pi}{M}) \\ -xsin(\frac{\mu\pi}{M}) + ycos(\frac{\mu\pi}{M}) \end{pmatrix}$ and $e^{(ik_{\nu,\mu}\boldsymbol{x})}$ is the oscillatory wave function, whose real and imaginary parts are respectively the cosine and sine functions. μ controls the orientation of the filters, with M being the total number of different orientations, and ν scales the center of the filter in the frequency domain [6].

A typical neutral face image contains curve-like features because it contains permanent facial components (e.g. eyebrows, lips) as well as straight features. Since facial expressions are generated by the movement of groups of muscles in any orientation and transient features like wrinkles and furrows, images with expressions contain even more curvature characteristics than straight ones. Therefore, to model these curve-like features, we include CG wavelets for face representation in addition to the conventional Gabor wavelets.

Peters et al. [14] obtained CG wavelets by adding a curvature parameter to the conventional Gabor formulation as follows:

$$\psi(\boldsymbol{x}; \nu, \mu) = \frac{k_{\nu,\mu}^2}{\sigma^2} e^{(-\frac{k_{\nu,\mu}^2 ||\boldsymbol{x}||^2}{2\sigma^2})} [e^{(ik_{\nu,\mu}\acute{x})} - e^{(-\frac{\sigma^2}{2})}]$$ (2)

$$\boldsymbol{x} = \begin{pmatrix} \acute{x} \\ \acute{y} \end{pmatrix} = \begin{pmatrix} xcos(\frac{\mu\pi}{M}) + ysin(\frac{\mu\pi}{M}) + c(-xsin(\frac{\mu\pi}{M}) + ycos(\frac{\mu\pi}{M}))^2) \\ -xsin(\frac{\mu\pi}{M}) + ycos(\frac{\mu\pi}{M}) \end{pmatrix}$$ (3)

where c corresponds to the curvature ratio.

CG wavelets do not have the orientation symmetry as in conventional Gabor wavelet as shown in Fig. 1 ([2]). For the conventional Gabor wavelet setting, it is usually sufficient to have 8 orientations ($M = 8$). However, this number should be increased to 16 to obtain the same orientation utilization in case of CG wavelets.
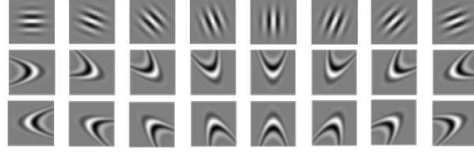

Fig. 1: Illustration of orientation asymmetry in CG wavelets with $c = 0.1$ (middle and bottom row) in comparison with the conventional Gabor wavelet (top row) (Image courtesy of [2]).

In CG wavelets, one can use different curvature degrees, i.e., $c = \{0.05, 0.1, 0.2\}$, and Gaussian sizes, i.e., $\sigma \in \{0.5\pi, \pi, 2\pi\}$, for multi-curvature utilization as well as scale space utilization. In this way, a more powerful representation of facial structures is obtained by extracting both fine and coarse features with straight and curved filters.

2.2 Local Binary Patterns

The local binary pattern (LBP) transformation has been proposed as a texture description method [12] and has proven to be very effective in representing facial texture and been widely used for both face and facial action recognition [1][18]. It maps the texture variation around each pixel to a binary pattern and the histogram of these patterns in a local window can be used directly as a descriptor for that certain region of interest. The computation of the pattern for a pixel at position x of an image I is as follows:

$$LBP_P(x) = \sum_{p=0}^{P-1} t(I(x_p) - I(x)).2^p$$ (4)

In this representation, each $I(x_p)$ is a neighboring pixel of the center pixel $I(x)$ on a neighborhood defined by the number of pixels P as well as the shape (e.g. rectangular or circular) and the distance to the central pixel which determines the resolution of the transformation. The function $t(x)$ is the simple thresholding function which returns 1

if the input pixel difference is positive and 0 if it is negative. In this way we obtain a $P - bit$ binary value, or an integer between 0 and $2^P - 1$ to represent each pixel.

In this work we use an 8-pixel circular neighborhood with the radius 1, giving 256 possible patterns. It has been shown, however, that only 58 of these patterns, called the uniform patterns, are sufficient to describe the majority of the texture information [13]. So we can reduce the size of the descriptor to 59 bins by assigning all the non-uniform patterns into a single bin.

Applying the LBP on top of Gabor magnitude images with various scales and orientations results in obtaining a richer representation and finer description of the facial texture [22]. In our work we extend this variation of descriptors by also including multiple curvature degrees and Gaussian sizes, obtaining the Local Curvature Gabor Binary Patterns (LCGBP) representation. Of course, this extension substantially increases the number of features obtained, and introduces more redundancy between features and possibly noise for the final classification task. Therefore, whether using directly the LCGBP histogram bins as features or, as we perform in this particular work, using a dissimilarity measure for the histograms between frames, a feature selection or dimension reduction technique is essential to be able to perform a meaningful classification using these features. The details on how we compute the histogram dissimilarity as well as the feature selection technique and the types of selected features are explained in more detail in the following sections.

3 Facial Action Recognition Framework

This section describes in detail each step in our automated facial action unit detection system using LCGBP as seen in Fig. 2.

3.1 Face Localization

To be able to perform an effective feature extraction among all images in the dataset, we first need to locate our region of interest, which is the face, as accurately and consistently as possible. Face detection systems which output a rectangular region around the face are generally not reliable enough to extract appearance features because of the variety across subjects, expressions and head poses. Therefore, we choose to use a facial point tracking system instead, which provides more stable boundaries for the face region.

In this paper, we localize 66 facial landmarks as seen in Fig. 2, using a publicly available automatic face tracking system proposed by Saragih et al. [15]. The face tracker is based on constrained local models (CLM) [5] with regularized landmark mean-shift as the fitting strategy [15]. The CLM, similar to the Active Appearance Model (AAM) [4], uses a combined model for the shape and texture, but the model in CLM consists of templates of appearance around each facial landmark point which allows accurately tracking facial points even under extreme head poses, intensive facial expressions and presence of occlusions.

Once we locate the facial landmarks using the face tracker, we crop the image using the most extreme landmarks on the horizontal and vertical directions of the facial

FACIAL POINT TRACKING   FACE LOCALIZATION AND SCALING   GABOR TRANSFORM USING 504 FILTERS   LBP TRANSFORM & WINDOW SELECTION

GENTLEBOOST + SVM AU1

GENTLEBOOST + SVM AU2

GENTLEBOOST + SVM AU4

GENTLEBOOST + SVM AU27

AU1   AU4   AU6

AU12   AU20   AU25

FINAL DECISION OF DETECTED AU's

FEATURE SELECTION AND CLASSIFICATION FOR 17 AU's

HISTOGRAM EXTRACTION & DISTANCE COMPUTATION FOR 504 x 400 (201600) REGIONS
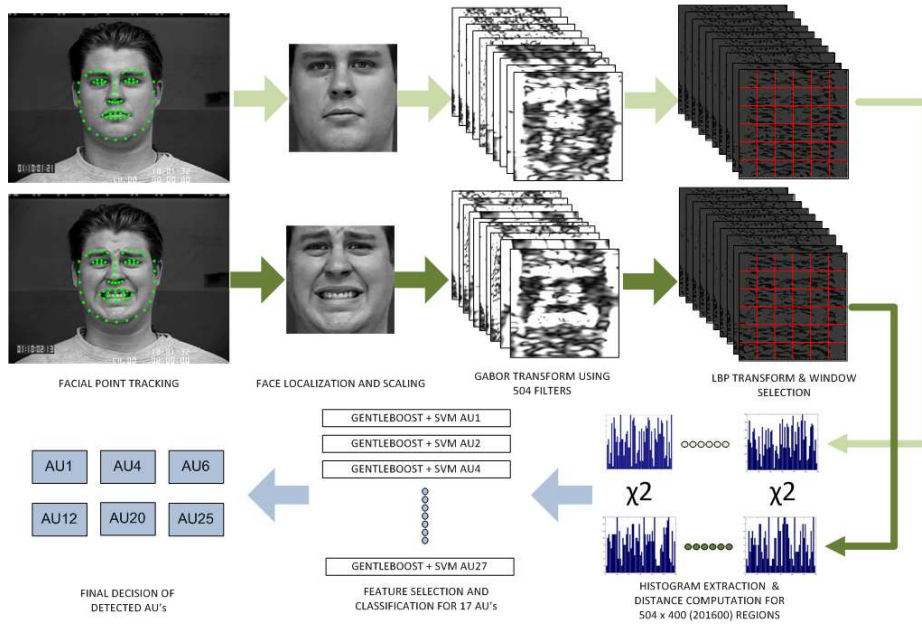
Fig. 2: Complete flowchart of the proposed framework for an input video

mask obtained, with a certain safety margin (Fig. 2). No rotation or texture warping is performed, since the databases that we use to train and test our system were recorded in quite constrained situations with respect to head pose and since the types of features we use have proven to be robust against misalignments. In this paper we only aim to show the strength of LCGBP as features for facial action recognition compared to other types of features. As future work registration of facial texture will be added to the system to gain more robustness against cases of unconstrained head pose. In our system we only scale each detected face region to a fixed size of 120 by 120 pixels.

3.2 Feature Extraction

After locating and scaling the face region we extract the appearance features using a combination of LCGBP transforms, which is the LBP transform applied on top of the image filtered by various curvature Gabor wavelets, as explained previously in Section 2 and as represented in Fig. 2. For our training and testing purposes we apply this filtering to the frame with a neutral expression and the frame with the peak of the posed expression separately for each sample video, since we utilize the comparison between those frames. For the CK+ database [10] these frames correspond to the first and last frames respectively. At this point the system requires that a frame is marked as neutral expression, then the method can be applied to any other frame of the same subject to detect action units at different intensities. This automatization problem can be solved by projecting the subject face with any expression to the PCA space created by examples

of expressionless faces, as proposed in [16]. However, this method was not tested in the scope of this paper.

The first step of feature extraction is applying the Gabor transforms to the input images. The classic method for generating Gabor representations of images is to apply wavelets in different scales and orientations with a fixed Gaussian size. In addition to adding the curvature component in various degrees we also include wavelets with different Gaussian sizes, similar to [2]. This is expected to result in a richer representation of finer details of facial texture components, which are crucial for high accuracy action recognition, compared to a single Gaussian size, and so is proven with our test results (presented in the following section). To be more precise we use Gabor wavelets of 3 different scales ($\nu \in \{0, 1, 2\}$), 8 (or 16 in case of curvature, see Fig. 1) orientations ($\mu \in \{0, \cdots, 7\}$), 3 Gaussian sizes ($\sigma \in \{\pi/2, \pi, 2\pi\}$) and 4 curvature degrees ($c \in \{0, 0.05, 0.1, 0.2\}$). This results in a total of 504 separate filters ($1 \times 3 \times 3 \times 8 + 3 \times 3 \times 3 \times 16$).

Next we apply the uniform LBP transform on each of the magnitude images of the outputs of these 504 filters for both the neutral and peak expression frame. Then to obtain the local texture information we calculate the histograms on 400 overlapping windows of sizes 20 by 20, 20 by 40, 40 by 20 and 40 by 40 with an overlap size of 10, all units in pixels. The conventional tendency in the literature for LBP histogram extraction has been to use non-overlapping windows of a fixed size, but as shown recently in our previous work [21], varying the size and performing a more extensive search using overlaps, combined with a powerful feature selection step, results in a more informative feature set. Then we compute for each of these windows the χ^2 distance of corresponding histograms in the neutral and peak expression frames, and obtain our full set of features of size 201600 (400×504). Using these alterations from the neutral face as features not only eliminates the variation caused by identity ([16], [21]) but also allows tracking the relative intensity of the movement between frames.

3.3   Relevant Feature Selection and AU detection

The extensive representation and search strategy chosen in the feature extraction technique results in a huge number of features which causes two main problems. First problem is that most of these features are correlated with each other so using them in combination in a classification task introduces an unnecessary computational burden. Secondly, only a portion of them are relevant to the task, i.e. detecting a specific action unit. The irrelevant features cause only noise and a decrease in accuracy in classification. Therefore we need to use a feature selection method that addresses both of these problems and that is specific to each action unit. Boosting techniques allow both reducing the dimensionality of the feature vector and eliminating the irrelevant features, since they are trained in a manner that maximizes the classification rate.

We adopt in this work the GentleBoost technique, since it has already been shown in the literature to be effective when used in combination with Support Vector Machines (SVM) ([20],[21]), which is the classification method that we utilize. For 17 AUs, which have a reasonable number of examples in our training database (CK+), we select 1000 features out of 201600 using GentleBoost separately, so we obtain the most relevant features in terms of Gabor parameters and the location in the 2D space. Then we train,

once again for each AU, an SVM, for which the two output classes are whether the AU is present or not. The details of the various tests and results are presented in the following section.

4 Action Unit Detection Results

In this section we report the results of our experiments performed on the Extended Cohn-Kanade database of facial expressions (CK+) [10]. The database consists of 593 videos of 123 subjects performing a facial expression formed by a single or multiple AUs, manually coded by trained FACS coders. Each sample video starts with a neutral expression and ends with the peak of the expression. We train and test our system using only this final frame of each sequence. All presented results are those obtained by a leave-one-subject-out test, i.e. training the SVM classifier on samples of 122 subjects and testing it on the remaining subject. We perform the tests for each AU using 100, 200, 300, 400, 500, 750 and 1000 features in the SVM and at each case choose the number of features giving the highest overall accuracy rate. Using the publicly available LibSVM implementation [3] we have tested both linear-SVMs and RBF kernels (parameters optimized using a 5-fold cross validation). Here, however, we only report results using the RBF kernels, since they result in better accuracy compared to the linear SVM in every AU, but there is no substantial difference when comparing different types of features.

4.1 Comparing types and combinations of Gabor features

We first compare the test results obtained by various parameter settings for the LCGBP and also using only LBP as a baseline comparison method. All settings are kept the same for this comparison, except only for the LBP the maximum number of features tested in the training phase of SVM is kept at the highest possible, i.e. 400.

We have tested 14 configurations in addition to the standard LBP features; namely 12 settings for LCGBP with 3 scales and 8 (or 16) orientations and a fixed Gaussian size (σ) chosen from 0.5π, π or 2π and fixed curvature degree (c) from $0, 0.05, 0.1, 0.2$ (0 meaning standard LGBP with 9600 total features, each of the rest yields 19200), one setting combining all proposed σ choices with $c = 0$ (28800 features) and one setting combining all possible σ and c choices (201600 features), which is the setting for the main proposed system. The comparison in three types of accuracy measures (overall accuracy, F1 and area under Receiver-Operator-Characteristics (ROC) curve (AUC)) averaged over 17 AUs (Upper face AUs 1, 2, 4, 5, 6, 7, 9 and lower face AUs 11, 12, 15, 17, 20, 23, 24, 25, 26, 27) can be seen in Fig. 3. The ROC curve was obtained by alternating the SVM decision threshold. The first observation, other than the definite superiority of LGBP to standard LBP, is that for all the single σ settings the curvature Gabors perform significantly better than the non-curvature standard Gabor setting, which is the first indication of the effectiveness of curvature features for facial action recognition. Another important comparison is the one between the 4 non-curvature LGBP settings. Using different sizes of Gaussians in the Gabor formulation in combination with each

(a) Overall Accuracy(%)
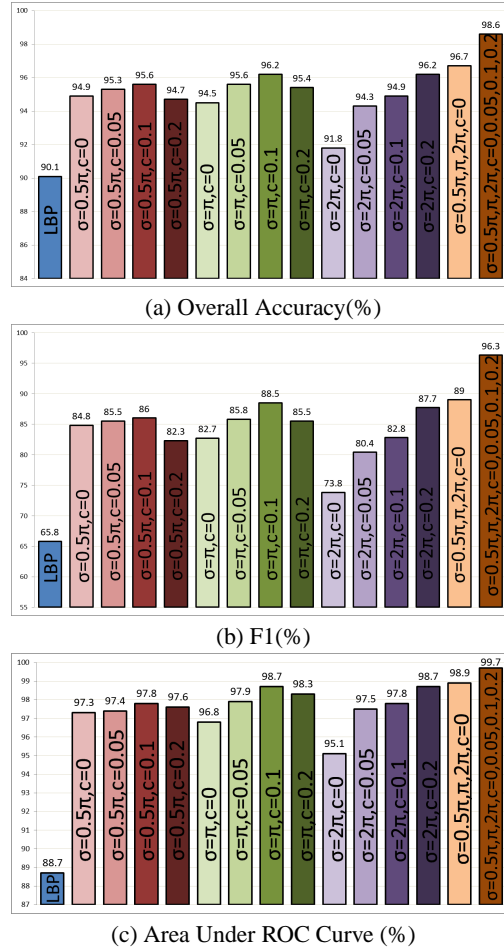

(b) F1(%)

(c) Area Under ROC Curve (%)

Fig. 3: Comparison of three different accuracy measures for different LCGBP feature settings & LBP

other results in a substantial increase in accuracy with respect to any fixed σ configuration. This indicates the necessity of alternating the Gaussian size along with the scale and orientation in any Gabor setting, which contradicts with the usual tendency in the literature for selecting Gabor wavelets for facial expression or AU detection.

The proposed setting, which is combining 3 different σ values and 4 different curvature degrees gives the highest classification accuracy for all action units, as expected. The results for each AU tested can be seen in Table 1 in comparison with the non-curvature case combining different σ's. The superiority is clearly not because of the greater number of features extracted (201600 vs. 28000), but because the various curvature degrees and filter sizes allow extracting those that are relevant to each action unit. We observe that for some AUs the difference between the two cases is less sig-

Table 1: Number of features used (No Feat.), Overall accuracy (OA), F1 and area under ROC curve (AUC) values for combinations of LCGBP(Curv.) and LGBP (No curv.) for 17 AUs

AU	No Feat.		OA		F1		AUC	
	Curv.	No curv.	Curv.	No curv.	Curv.	No curv.	Curv.	No curv.
AU1	750	750	0.976	0.958	0.959	0.928	0.995	0.983
AU2	1000	1000	0.992	0.987	0.978	0.965	0.997	0.998
AU4	750	750	0.963	0.935	0.942	0.897	0.994	0.976
AU5	750	1000	0.985	0.965	0.956	0.895	0.997	0.992
AU6	1000	1000	0.985	0.955	0.963	0.884	0.998	0.991
AU7	750	750	0.968	0.936	0.917	0.835	0. 996	0.969
AU9	750	300	1	0.995	1	0.979	0.998	0.994
AU11	1000	300	0.997	0.979	0.969	0.786	0.999	0.984
AU12	1000	1000	0.988	0.968	0.973	0.923	0.998	0.994
AU15	1000	750	0.988	0.969	0.962	0.897	0.999	0.993
AU17	1000	1000	0.975	0.956	0.963	0.935	0.993	0.989
AU20	500	1000	0.983	0.975	0.937	0.905	0.996	0.991
AU23	750	500	0.993	0.971	0.967	0.838	0.999	0.993
AU24	750	1000	0.993	0.965	0.964	0.796	0.999	0.989
AU25	500	1000	0.979	0.966	0.982	0.969	0.994	0.994
AU26	1000	1000	0.989	0.959	0.938	0.721	0.999	0.987
AU27	200	1000	0.998	0.995	0.994	0.981	0.999	0.999
Avg.			**0.986**	0.967	**0.963**	0.89	**0.997**	0.989

nificant than others, and this can be explained by the variation of amount of curvature that shapes the deviation from the resting state for each action unit. However, observing Table 2, which shows the ratio of features chosen by the GentleBoost with respect to σ and c values and the deviation among action units, we can say that none of the types of features show a too powerful dominance over others in none of the AUs, although the non-curvature features are selected significantly less frequently than the rest. This suggests that every type of feature chosen is of similar importance to the detection task and their combination is essential for such a high classification accuracy.

Table 2: Mean and standard deviation of percentage of features chosen from different Gaussian sizes defined by σ and curvature values (c).

c \ σ	0.5π	π	2π	Total
0	3.7 ± 0.6	4.8 ± 0.9	5.4 ± 1.4	13.9 ± 1.6
0.05	7.6 ± 1	9.2 ± 1.2	10.6 ± 1.4	27.3 ± 2.2
0.1	7.6 ± 1.2	8.8 ± 1	11.3 ± 1.1	27.7 ± 1.1
0.2	8.2 ± 1	9.7 ± 0.8	13.3 ± 2.6	31.1 ± 2.3
Total	27 ± 2.6	32.5 ± 1.8	40.5 ± 2.7	

4.2   Comparison with existing work

We compare our results, as shown in Table 3, with 3 recently conducted works ([20], [16] and [21]) which have reported results on the Cohn-Kanade database and have used similar techniques either in the feature extraction or the classification phase, in addition to the baseline system proposed in [10]. Valstar et al. [20] have used the evolution of certain facial landmarks throughout the video sequence as features and utilized the Gentleboost and SVM as the feature selection and classification methods. In our recently published work [21], we have also used Gentleboost and SVM with a combination of shape features similar to [20] and LBP features that are improved with the help of three filters. The work in [16] uses as features directly the bins of histogram difference of LGBP magnitude images extracted from 16 non-overlapping windows with a fixed Gaussian size and no curvature, and as classification adopts SVM with a specially trained kernel. In [10] the database was validated using 68 geometrically normalized facial point locations and canonical normalized appearance vectors as features for an SVM classification.

Table 3: Accuracy comparison with 4 other methods; No of AUs represents the number of common AUs taken into consideration

Type of acc.	F1(%)		AUC(%)		AUC(%)		AUC(%)	
No of AUs	14		17		16		14	
Method	[20]	Our Met.	[10]	Our Met.	[16]	Our Met.	[21]	Our Met.
	61.86	**96.09**	94.5	**99.7**	96.45	**99.69**	96.9	**99.65**

As seen in Table 3, our method certainly outperforms all the other state-of-the-art methods on the CK+ database in AU detection accuracy. The comparison with the two methods ([20] and [21]) using the same type of feature selection and classification, and the database validation system [10] which also uses SVM, shows the efficiency of the type of features utilized in our system. Although it is only fair to say that the comparison with [20] is not exactly straightforward since the authors have used many frames from each sequence, instead of using only the peak expression one, which naturally causes a decrease in the classification accuracy. Also, they have used the CK database [9] instead of the CK+, which is a previous version that includes less subjects and sequences. The comparison with [16], which uses a rather complicated classification scheme, also proves the utility of using curvature based features in addition to combining different sizes of Gabor wavelets.

These initial results obtained on the CK+ database demonstrates a great potential of the proposed features, but additional tests certainly need to be performed on larger databases to show the generalizability of the system, which remains as principal future work.

5 Conclusions

We have presented a novel framework for facial action unit detection in videos. The proposed system consists of extracting a combination of curvature Gabor features at different filter sizes, applying the LBP on top and computing the difference in histograms for neutral and peak frames. Then the obtained features are used in an AU specific feature selection and classification process to detect the present AUs. We achieve 98.6% accuracy, 96.3% F1 and 99.7% AUC scores in average for the leave-one-out test performed on the CK+ database, which is to our knowledge the highest reported to date. To assess the generalizability of the system, further tests should be performed with a dataset containing a larger variability among expressions. However, the extremely high accuracy presented in this work already shows the representation and discriminative power of the proposed features, which we believe will constitute an important position in future facial action recognition and expression analysis research.

References

[1] Ahonen, T., Hadid, A., Pietikainen, M.: Face description with local binary patterns: Application to face recognition. Pattern Analysis and Machine Intelligence, IEEE Transactions on 28(12), 2037–2041 (2006)

[2] Arar, N.M., Gao, H., Ekenel, H.K., Akarun, L.: Selection and combination of local gabor classifiers for robust face verification. Biometrics: Theory, Applications and Systems (BTAS),IEEE 5th International Conference on pp. 297–302 (2012)

[3] Chang, C.C., Lin, C.J.: Libsvm: A library for support vector machines. Intelligent Systems and Technology, ACM Transactions on 2(3), 1–27 (2011)

[4] Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. Pattern Analysis and Machine Intelligence, IEEE Transactions on 23(6), 681–685 (2001)

[5] Cristinacce, D., Cootes, T.: Feature detection and tracking with constrained local models. In: Proceedings of the British Machine Vision Conference 2006. pp. 929–938 (2006)

[6] Daugman, J.: Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two-dimensional visual cortex filters. Journal of Opt. Soc. Amer. 2(7), 1160–1169 (1985)

[7] Ekman, P., Friesen, W.: The facial action coding system: A technique for the measurement of facial movement. San Francisco CA: Consulting Psychologists Press, Inc. (1978)

[8] Hwang, W., Huang, X., Noh, K., Kim, J.: Face recognition system using extended curvature gabor classifier bunch for low-resolution face image. Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE International Conference on pp. 15–22 (2011)

[9] Kanade, T., Cohn, J., Tian, Y.: Comprehensive database for facial expression analysis. In: Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on. pp. 46–53 (2000)

[10] 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. Computer Vision and Pattern Recognition Workshops (CVPRW),IEEE International Conference on pp. 94–101 (2010)

[11] Lucey, P., Cohn, J., Matthews, I., Lucey, S., Sridharan, S., Howlett, J., Prkachin, K.: Automatically detecting pain in video through facial action units. Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on 41(3), 664–674 (2011)

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

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

[14] Peters, G., Kruger, N., von der Malsburg, C.: Learning object representations by clustering banana wavelet responses. 1st International Workshop on Statistical Techniques in Pattern Recognition pp. 113–118 (1997)

[15] Saragih, J.M., Lucey, S., Cohn, J.F.: Deformable model fitting by regularized landmark mean-shift. Int. J. Comput. Vision 91(2) (Jan 2011)

[16] Senechal, T., Bailly, K., Prevost, L.: Automatic facial action detection using histogram variation between emotional states. Pattern Recognition (ICPR), 20th International Conference on pp. 3752–3755 (2010)

[17] Senechal, T., Turcot, J., el Kaliouby, R.: Smile or smirk? automatic detection of spontaneous asymmetric smiles to understand viewer experience. Automatic Face and Gesture Recognition,IEEE International Conference on (2013)

[18] Shan, C., Gong, S., McOwan, P.W.: Facial expression recognition based on local binary patterns: A comprehensive study (2009)

[19] Valstar, M.F., Mehu, M., Jiang, B., Pantic, M., Scherer, K.: Meta analysis of the first facial expression recognition challenge. Systems, Man, Cybernetics, Part B: Cybernetics, IEEE Transactions on 42(4), 966–979 (2012)

[20] Valstar, M.F., Pantic, M.: Fully automatic recognition of the temporal phases of facial actions. Systems, Man, Cybernetics, Part B: Cybernetics, IEEE Transactions on 42(1), 28–43 (2012)

[21] Yuce, A., Sorci, M., Thiran, J.P.: Improved local binary pattern based action unit detection using morphological and bilateral filters. Automatic Face and Gesture Recognition(FG),IEEE International Conference on (2013)

[22] Zhang, W., Shan, S., Gao, W., Chen, X., Zhang, H.: Local gabor binary pattern histogram sequence (lgbphs): a novel non-statistical model for face representation and recognition. Computer Vision, 10th IEEE International Conference on 1, 786–791 (2005)