

Binary Pattern Flavored Feature Extractors for Facial Expression Recognition: An Overview

Rasmus Lyngby Kristensen*, Zheng-Hua Tan[†], Zhanyu Ma[‡] and Jun Guo[‡]

*Section of Image Analysis and Computer Graphics

DTU Compute, Technical University of Denmark, Kgs. Lyngby, Denmark

Email: raly@dtu.dk

[†]Signal and Information Processing section (SIP)

Department of Electronic Systems, Aalborg University, Aalborg, Denmark

Email: zt@es.aau.dk

[‡]Pattern Recognition and Intelligent System Lab.

Beijing University of Posts and Telecommunications, Beijing, China

Email: mazhanyu@bupt.edu.cn

Abstract—This paper conducts a survey of modern binary pattern flavored feature extractors applied to the *Facial Expression Recognition* (FER) problem. In total, 26 different feature extractors are included, of which six are selected for in depth description. In addition, the paper unifies important FER terminology, describes open challenges, and provides recommendations to scientific evaluation of FER systems. Lastly, it studies the facial expression recognition accuracy and blur invariance of the Local Frequency Descriptor. The paper seeks to bring together disjointed studies, and the main contribution is to provide a solid overview for future research.

I. INTRODUCTION

When a scientific field sees a burst of published articles it often results in a clutter where an overview of state-of-the-art is lost. Today, this is what has happened to the field of *Facial Expression Recognition* (FER). Many of the recent articles focus on feature extraction and description. Especially extractors that are somewhat related to *Local Binary Patterns* (LBP) are popular. We term this kind of extractors: *Binary Pattern Flavored* feature extractors. The main contribution of this paper is to provide an overview of how this feature extractor type has been applied to FER. From our literature survey it is evident, that a substantial amount of researchers still base their FER systems on the relatively outdated LBP feature. Often it is not explicitly stated why LBP was chosen, but better performance and general robustness could have been achieved if a different and newer feature had been used. Thus, a clear overview of the current state-of-the-art will allow for the construction of better systems as well as extending them.

In order to provide a thorough overview, we make the following contributions:

- Unification of the FER terminology
- Elucidation of unresolved challenges in feature extraction for FER
- Considerations for choosing a feature extractor
- Overview of binary pattern flavored features currently used for FER

- Recommendations for good evaluation practice of feature extractors for FER

The remainder of the paper is divided according to the list above. Section II defines the FER terminology, Section III elucidates unresolved challenges, Section IV defines feature extractor considerations, Section V provides an overview of current binary pattern flavored features, Section VI provides evaluation recommendations, Section VII performs an example study of how the *Local Frequency Descriptor* (LFD) feature extractor performs for FER, and Section VIII concludes the paper.

II. TERMINOLOGY

By looking through the literature, one can see that different researchers use a slightly different FER terminology. To encourage consistency in the future, we seek to define a set of pivot terms based on the general FER surveys [1], [2], [3], and [4].

Spontaneous expressions are based on a true, underlying felt emotion. They are typically what FER systems aim at recognizing as they are the ones carrying true emotion information.

Posed expressions are acted expressions. According to [5], posed and spontaneous expressions look differently. Spontaneous expressions are hard to obtain, which is why most facial expression databases contain posed expressions.

Micro expressions are expressions which never reach full apex because the executing person suppresses the expressions before they are fully formed. Micro expressions can be very informative in some applications, such as clinical psychology.

Basic facial expressions were defined by [6]. They are six expressions which has been proved universal and are as follows: Happiness, Sadness, Anger, Surprise, Disgust and Fear. Sometimes Neutral is included as well. The basic expressions have been extensively used in FER research, but [7] and [8] argue that they are inadequate.

Entity based FER recognizes expressions by considering the entire face as one entity. This is by far the most popular

and most researched form of FER [9].

Action Unit based FER recognizes muscle contractions and relaxations in the face and use them to recognize the expression. These activations of face muscles are known as *Action Units* (AU) as described by the *Facial Action Coding System* (FACS) [10].

Permanent facial features are features being permanent in the face, such as the eyes and mouth [11].

Transient facial features are volatile face features, meaning that they occur and disappear as the face moves, such as wrinkles and furrows [11].

Geometric features are image features which describe some form of physical parameters of the face, such as positions, shapes, sizes, or similar. Usually, a set of easily recognizable salient points are defined on the face and their positions and mutual distances are used to describe the face.

Appearance features describe image properties based directly on pixel intensities. Often, they are formed by filtering images by a bank of image filters and use their outputs as features. Binary pattern flavored features fall in this category.

Global feature descriptors describe the features over the entire face. Thus, a very small amount of feature location information is kept.

Local features descriptors contain information about the spatial position of features. Many ways of retaining locality information exist, but it is often done by dividing faces into smaller regions, each with its own descriptor.

Engineered features are features which have been explicitly designed and therefore convey some form of meaningful information about the objects in the image.

Learned features are features which are learned from a dataset of observations by a learning algorithm. As a result, the features do usually not convey any meaningful information.

Binary pattern flavored features describe objects in an image by binary patterns extracted from small local image patches. The standard example is LBP. Some authors use the term "local features" but obviously this term is too easily confused with the term *local feature descriptors*, as defined above.

III. RESEARCH CHALLENGES

In 2011, the first *Facial Expression Recognition and Analysis Challenge* (FERA2011) [12] was held as a way to uncover the then current status of the FER field. A meta analysis of the submitted papers was published in 2012 [9]. Based on this meta analysis, and our thorough literature survey, there seems to be a recent overweight of papers that focus on entity based FER from high-resolution faces viewed front on. The meta analysis concluded that this problem can be considered largely solved for the person-specific case.

We believe that the continued substantial amount of research going into this problem is due to missing awareness of the state-of-the-art. Clarifying the current general challenges of the FER field also results in clarification of the feature

specific challenges, as they must be correlated.

Be aware that research has already been done on some of the challenges in the following list, but that more is needed. Also note that most papers only tackle few challenges at a time. Thus, there is unexplored territory to conquer for researchers who are interested in creating multi-invariant systems.

Blurred faces can result from disturbances in the atmosphere or out-of-focus images. *Local Phase Quantization*, developed by [13] for face recognition in blurred images and applied to FER by [14] and [15], is particularly interesting. Note that neither [14] nor [15] investigates FER in blurred images. They use LPQ solely for its high descriptiveness.

Illumination variations results from changes in direction or intensity of light used to illuminate faces. This is often seen in face images obtained in a real-world setting. A substantial amount of research has been done on intensity invariance, and it is a natural property of most binary pattern flavored features. In our survey, we only encountered research on direction invariance as a subparts of larger studies, such as [16].

Transformation of face position and orientation happens if a person is allowed to move freely, as would be the case in a real-world setting. Quite a lot of research has gone into rotation and scaling invariance, such as rotation and scale invariant LBP [17]. Translations aligned with the image plane are often handled by the segmentation algorithm. Changes to face orientation have been studied, such as [18], but it remains an open problem.

Face occlusion happens if there is not a direct line of sight between the camera and the face. Occlusion is often due to mustaches, glasses or hair. It has seen a relatively large amount of research, such as the comparative study done by [4]. However, the topic still remains active and unsolved.

Variations in faces between different people make facial expressions person dependent. [19] proposed to use the difference between the expressive face and the neutral face to eliminate the person specific contributions. However, this requires the setting to be person specific, thus with known identities.

Multiple faces in one image could be a problem. We are aware that this challenge has been studied for face recognition, but so far we have not encountered any such research related to FER.

Variations in facial expression intensity happen because humans can form both subtle and explicit versions of facial expressions. A smile might be small and subtle, it can be large and explicit, or somewhere in between. So far, we have not encountered any research which addresses this challenge.

Recognition of AUs has been done, but as noted in the meta analysis of FERA2011 there is room for a lot more. Reliable recognition of AUs is still a challenge.

Processing time is an important factor in instant feedback systems, which is usually the case for social robots [56] and the like. Some researchers provide notes about the processing time of their systems, but often just as a minor detail. A comparative study would be worthwhile.

TABLE I. FEATURE ABBREVIATIONS

Abbr.	Phrase
LBP	Local Binary Patterns [26]
$LBP_{R,P}^{riu2}$	Rotation Invariant LBP [27]
FLBP	Fuzzy LBP [28]
PLBP	Pyramid of LBP [29]
MBP	Median Binary Patterns [30]
LTP	Local Transitional Patterns [31]
MTP	Median Ternary Patterns [32]
VLBP	Volume LBP [33]
LBP-TOP	LBP from Three Orthogonal Planes [33]
SLMBP	Spatiotemporal Local Monogenic Binary Patterns [34]
LGBP	Local Gabor Binary Patterns [35]
LGBP-TOP	LGBP from Three Orthogonal Planes [36]
LGDP	Local Gabor Directional Patterns [21]
RGJ	Radial encoded Gabor Jets [37]
LGTP	Local Gabor Transitional Patterns [38]
LDP	Local Directional Patterns [39]
LDPv	LDP variance [40]
GDP	Gradient Directional Patterns [41]
WLD	Weber Local Descriptor [42]
LPQ	Local Phase Quantization [14]
RI-LPQ	Rotation Invariant LPQ [43]
VLQP	Volume LPQ [44]
LPQ-TOP	LPQ from Three Orthogonal Planes [15]
LCVBP	Local Color Vector Binary Patterns [45]
TPCF	Tensor Perceptual Color Framework [16]
CLBP	Curvature Local Binary Patterns [46]
PRICoLBP	Pairwise Rotation Invariant Co-occurrence Local Binary Pattern [47]
Ferns	[48], [49]
LEP	Local Energy Patterns [50]
LFD	Local Frequency Descriptor [51]

IV. FEATURE EXTRACTOR CONSIDERATIONS

A set of considerations for selecting a feature extractor for FER has to be made. A flowchart of the choices can be seen in Fig. 1. Five binary primary choices, all based on ones research focus, has to be made, some with sub-choices. Note that two opposite choices can be combined for potential higher recognition rates as reported by [20].

Most researchers tend to make these choices implicit without arguments. We believe that making choices explicit with solid argumentation will lead to better understanding and therefore better science. Therefore, we encourage future researchers to provide argumentation for the choices they make.

V. BINARY FLAVORED FEATURES

This section provides an overview of modern binary pattern flavored features which have previously been applied to the FER problem. Fig. 2 shows the relationships between these features, including four new, promising features. These four features were selected because they have provided promising results for other applications. We strived to make the list of previously used features as comprehensive as possible, but please note that some could be missing by accident.

From the figure it is evident, that there is a large overweight of feature extractors which use still images containing 2D intensity data as input. In fact, only a single extractor relies on 3D depth data. Only a handful of extractors relies on image sequences. The rest use still images. Based on our survey, this tendency seems to be a good indicator for the status of the FER field in general. We hope that this survey will help to change that. The used abbreviations are explained in Table I.

We encourage readers to use this overview to select a proper feature extractor for their own research, and to minimize the development of redundant features. In 2011, [18] proposed to use *Local Gabor Binary Patterns* (LGBP) for making a face-orientation invariant FER system. Later for the same year, a system using LGBP for AU based FER won the AU recognition sub-challenge of FERA2011 [35]. As recent as 2014, a new paper on entity based FER using LGBP was published by [52]. The authors seems to be unaware of the previous use of LGBP, as they make no notes or citations to the two articles listed above. Thus, redundant work could have been avoided by a clear overview of the state-of-the-art.

Below we have selected six of the most interesting recently used feature extractors for a more detailed description. We hope that the descriptions will help readers choose a good, modern feature extractor, and thereby enable them to develop highly accurate and robust FER systems for practical use. Note that the novel *Local Frequency Descriptor* (LFD) will be studied in Section VII. As proposed by [53], all of the described features use local descriptors formed by dividing the input images into a set of non-overlapping blocks, each with its own descriptor. The descriptors are then concatenated into one, large descriptor, thus keeping some locality information. This is informally known as the *grid division approach*. Also note that all the referenced methods use entity based FER.

A. Features from static image data

Pyramid of Local Binary Patterns (PLBP) was proposed in 2013 by [29]. PLBP is interesting due to their high level of descriptiveness and low complexity compared to other pyramid methods. The key idea is to extract $LBP_{R,P}^{u2}$ features from a given rectangular region of the image. That region is then sub-divided into four smaller regions, each of which can be further sub-divided and so on. LBP features are extracted from all regions, and their resulting descriptors are concatenated into one. The authors show that two subdivisions are enough, thus creating a total of five histograms for each region. They use the region covering the mouth and the region covering the eyes. PLBP requires less memory and provides faster computation than previous pyramid features. Further, PLBP provides a stronger discriminative ability than previous non-pyramid methods. However, it has not been compared against LPQ or similar modern features.

Median Ternary Patterns (MTP) was proposed in 2013 by [32]. MTP is interesting due to their robustness and high level of descriptiveness. MTP combines the strengths of MBP and LTP to form a highly illumination and random noise invariant feature. The idea is to quantize each pixel into a ternary pattern. For a given pixel, the median of the intensities of its eight directly neighboring pixels is calculated. The neighbors are quantized into three states depending on their relation to the median. The ternary pattern is split into two binary patterns and their two corresponding descriptors are concatenated together. MTP has a higher recognition rate and better illumination and noise robustness than pixel intensity *Principal Component Analysis* (PCA), LBP, MBP, and LTP. No comparison against LPQ or similar has been done.

Local Color Vector Binary Patterns (LCVBP) was proposed in 2013 by [45]. LCVBP is interesting because they

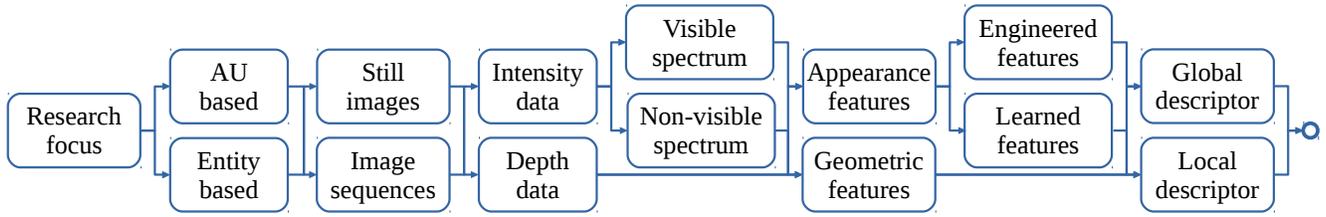


Fig. 1. Flowchart of considerations for selecting a feature extractor. Making these choices should always be based on what is best in relation to the research focus. Opposite choices can be combined, e.g. a system could use both intensity data and depth data to achieve a higher descriptiveness.

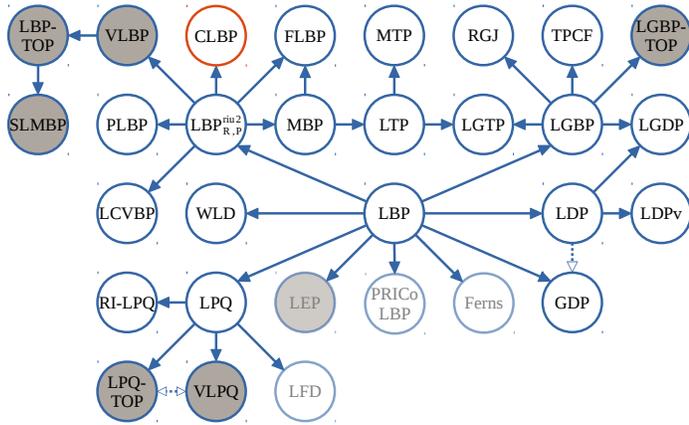


Fig. 2. Overview of 26 binary flavored feature extractors which has been applied to FER, along with four new features. Each node is a feature. Coloring codes: White node = still image data, Gray node = image sequence data, Blue rim = 2D intensity image data, Orange rim: 3D depth data. Solid arrows indicate a direct extension of a previous method. Dotted arrows indicate an inspirational link to a previous method. Transparent nodes indicate new features. The abbreviations are defined in Table I.

are among the few binary flavored features which incorporates color information. The key idea is to treat the color of each pixel as a three dimensional vector of which the magnitude can be calculated by $r = \sqrt{R^2 + G^2 + B^2}$. Binary patterns are created from the magnitude response similarly to the process used by $LBP_{R,P}$. In addition to these patterns, also magnitude difference patterns, color directional patterns, and color directional arc patterns are extracted. The descriptors from the four pattern types are concatenated to form the final descriptor. LCVBP provides a higher recognition accuracy than a large set of previous features, including $LBP_{8,2}^{u2}$, LGBP, and TPCF, but excluding LPQ. Further, it is shown that LCVBP allows for sparse representation, is person-independent, and robust against illumination changes.

Local Gabor Directional Patterns (LGDP) was proposed in 2012 by [21]. LGDP is interesting because it revives the use of Gabor wavelet filters in modern binary pattern features, and attain a high recognition rate in doing so. LGDP combines Gabor wavelet filtering and *Local Directional Patterns* (LDP). This is done by exchanging the eight Kirsch compass kernel filters normally used in LDP by eight Gabor wavelet filters corresponding to eight different angles. The eight filter responses are encoded using the same approach as LDP. This is done at five different scales, creating five descriptors which are concatenated. LGDP is showed to provide higher recognition accuracy than LDP and LBP (among others), but it is not compared to LPQ. It is further showed that it is invariant to

illumination changes and random noise.

B. Features from image sequence data

Local Phase Quantization from Three Orthogonal Planes (LPQ-TOP) was proposed in 2011 by [15]. LPQ and its derivatives are interesting due to its well established theoretical background, invariance to image blurring and illumination changes, and high level of descriptiveness. It has been showed, that the spatial phase angles of a centrally symmetric *Point Spread Function* (PSF), such as atmosphere and out-of-focus blur, is 0 at frequencies with positive real valued frequency responses. LPQ encodes the phase angles, which makes it blur invariant. The Three Orthogonal Planes framework was developed as an LBP extension by [33]. A given spatial 2D image sequence is considered as a 3D space by letting the movement in time represent the third dimension. From the 3D space, three orthogonal 2D planes are extracted. LPQ descriptors are extracted from each plane and concatenated together to form the sequence descriptor. LPQ-TOP provides higher recognition rates than LBP-TOP and is invariant to image blurring as well as intensity changes. The TOP framework has also been applied to LGBP, forming the LGBP-TOP descriptor by [36]. Unfortunately, LGBP-TOP and LPQ-TOP was not compared.

C. Features from 3D depth data

Curvature Local Binary Patterns (CLBP) was proposed in 2013 by [46]. CLBP is interesting because it is the only binary pattern flavored feature we have encountered which uses 3D depth data. CLBP is invariant to changes in face rotation, face orientation (pose), and changes in illumination. Four types of curvature measures are calculated, namely the two principal curvatures k_1 and k_2 , the mean curvature, and the shape index. The four curvature measures are encoded by LBP, and the four resulting feature descriptors are concatenated. It is showed that the recognition accuracy of CLBP is higher than that of two state-of-the-art 3D FER methods, which does not use binary pattern flavored features. It would be interesting to compare this 3D feature to state-of-the-art 2D features with respect to invariance and recognition accuracy. Such a study would provide arguments for choosing one type of input data over another.

VI. RECOMMENDATIONS FOR EVALUATION

To extend the state-of-the-art, one has to benchmark results against those previously reported. Benchmarking can easily be done if all researchers use the same well defined benchmarking datasets along with the same well defined benchmarking procedures. As things are now, quite a few different facial

expression databases and evaluation procedures are used. This includes different databases containing static and sequential images, static and sequential 3D depth data, labeled with basic facial expressions for entity based FER, and/or AUs for AU based FER.

Some 3D depth databases of facial expressions exist, but as was shown in Section V, research on 3D methods is relatively sparse. Therefore, we have not been able to identify a standard benchmarking database and procedure for 3D methods.

For 2D methods in intensity images, the *Cohn-Kanade* (CK) and the *Extended Cohn-Kanade* (CK+) [22] databases are popular. Both databases were primarily created for research in AUs and AU based FER. Therefore, they contain AU labels for all images, but only CK+ contains facial expression labels. As a consequence, researchers who use the original CK database have to manually apply labels to the images, thus potentially creating different labels. To fairly compare results, procedures such as the number of cross-validation groups and the specific identity of samples in each cross-validation groups have to be precisely the same. This process also tends to vary from researcher to researcher.

A good candidate for a unified benchmarking dataset is the well defined dataset and procedure which was used for FERA2011 called GEMEP-FERA2011 [12]. It is composed of two subsets of the original GEMEP dataset, one annotated by FACS for AU based FER and one annotated by emotions for entity based FER. The used emotions are: Anger, Fear, Joy, Relief, and Sadness. Though it has been little used since the competition, it is still maintained and available.

Naturally one usually wants to know how well their FER system recognizes real facial expressions. As mentioned in Section II, true spontaneous expressions are difficult to obtain. Some databases containing spontaneous expressions have been compiled, such as the Rochester/UCSD FacialActionCodingSystem Database 1 (RUFACS1) [23] for 2D intensity data and the BP4D-Spontaneous [24] database for 3D depth data. However, when using only a single database there is the risk of training ones system to recognize some underlying property of the images, other than the actual facial expressions. If so, artificially high recognition rates will generally be obtained when testing on images from the same dataset. To avoid this, [25] showed that cross-database evaluation on multiple facial expression databases will provide a closer-to-real-world recognition accuracy.

VII. AN EXAMPLE STUDY

This section performs an experiment which studies how well LFD recognizes blurred facial expressions. We chose to compare LFD against $LBP_{8,1}^{u,2}$ due to its continued popularity within the field and against LPQ due to its link to LFD as well as its high blur invariance and level of descriptiveness. LFD was developed by [51] as an extension of LPQ which includes magnitude information. They showed that LFD provides higher recognition rates than LPQ for face recognition in blurred images. Thus, our hypothesis was that LFD would also provide higher recognition for FER in blurred images.

We chose to use the CK and the *Karolinska Directed Emotional Faces* (KDEF) [54] databases because of the larger number of images contained in both databases compared to the GEMEP-FERA2011 dataset. Previous research on LPQ have focused on entity based FER. As we wished to benchmark LFD against the state-of-the-art, we also focused on entity based FER. For the same reason we chose still images as our input data.

We chose to base our system on the baseline system defined for FERA2011 because of its simplicity. Our system started by extracting LBP, LPQ, and LFD features using the grid division approach. Then, the dimensionality of each of the feature descriptors was reduced to 140 dimensions by PCA. Lastly, One-Versus-One Support Vector Machines (SVM) [55] were used for classification. Five fold cross-validation was used, and the experiment was executed a total of 10 times with 10 different partitions for the five cross-validation groups. This was done to lower the risk of obtaining bad results due to a bad selection of cross-validation groups.

Every time the experiment was run, the clear images in the four training cross-validation groups were used. The images in the remaining cross-validation group were blurred and used for testing. Gaussian blur was used, and a total of 17 different kernel variances were tried, ranging from 0 to 4 in steps of 0.25. The results of the experiment can be seen in Fig. 3.

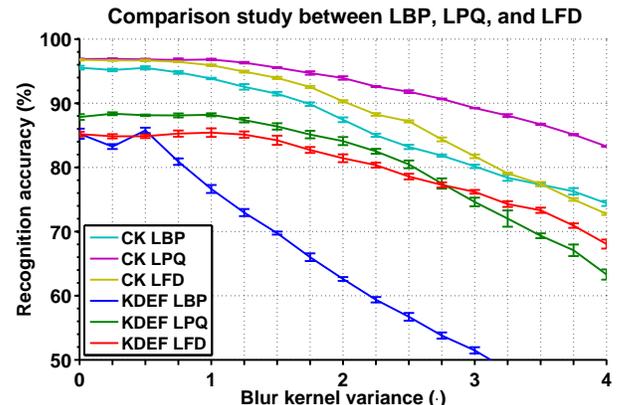


Fig. 3. FER accuracy of LBP, LPQ, and LFD features extracted from blurred images from the CK and KDEF databases. Note that LPQ performs generally better than LFD which again performs generally better than LBP for both databases. Error bars are used to indicate the variance of the recognition accuracy calculated based on 10 runs of five fold cross-validation.

We observe that LPQ generally performs better than LFD, which generally performs better than LBP. A steeper decline in the recognition accuracy of LPQ than LFD is observed for KDEF, which makes LFD better than LPQ at blur-variances of 3 or higher. However, this effect is not seen for CK. A reason might be, that the CK dataset contains more explicit expressions compared to the KDEF set (which is why the accuracy for CK is generally 10%-points higher than for KDEF). Thus, LFD might have an advantage over LPQ for recognizing subtle expressions in very blurred images. Note that this correlation is highly speculative, and more studies on other datasets containing subtle expressions would have to be done to state anything conclusive.

VIII. CONCLUSION

This paper compiled the most comprehensive overview of the use of binary pattern flavored features for FER to date. We chose to focus on this type of features due to their high popularity and thus large number of recent publications. In addition to the overview, we unified the fundamental FER terminology, elucidated unresolved FER challenges, provided considerations for selecting an FER feature extractor as well as recommendations for scientific evaluation of FER systems, and studied the descriptiveness of the novel LFD feature by comparing it against LBP and LPQ.

Our study of LFD showed that it provides a recognition rate and blur invariance which is higher than LBP, but lower than LPQ.

Despite the large quantity of research in the field, we showed that there are still challenges to overcome. In particular, we showed that multi-robust systems, the use of 3D depth data, the use of image sequences, and FER based on AUs are all more or less open topics. In particular, we see AU based FER as an interesting topic. Recognizing AUs would open up for the possibility of recognizing micro expressions as well as indicating the intensity of recognized expressions. Those properties would be of value to clinical psychology as well as advanced human-machine interfaces. Further, de-identification would be a natural part of an AU system if only the recognized AUs are stored. This could make the system easier applicable to public spaces as no personal identity information would be stored.

IX. ACKNOWLEDGMENTS

We would like to extend our gratitude to ?, ?, ?, and ? for providing travel funds which made our corporation between ? and ?.

REFERENCES

- [1] V. Bettadapura, "Face Expression Recognition and Analysis: The State of the Art", *arXiv.org*, abs/1203.6722, 2012.
- [2] N.U. Khan, "A comparative analysis of facial expression recognition techniques", *IEEE Third International Advance Computing Conference (IACC)*, Ghaziabad, 2013, pp. 1262-1268.
- [3] C. Caeanu, "Face expression recognition: A brief overview of the last decade", *IEEE 8th International Symposium on Applied Computational Intelligence and Informatics (SACI)*, Timisoara, 2013, pp. 157-161.
- [4] C.F. Liew and T. Yairi, "A comparison study of feature spaces and classification methods for facial expression recognition", *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Shenzhen, 2013, pp. 1294-1299.
- [5] P. Ekman and E.L. Rosenberg, "What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)", *Oxford University Press*, vol. 1, 1997.
- [6] P. Ekman and W.V. Friesen, "Constants across cultures in the face and emotion", *Journal of Personality and Social Psychology*, vol. 17(2), 1971, pp. 124-129.
- [7] W.G. Parrott, "Emotions in Social Psychology: Key Readings", *Psychology Press*, vol. 1, 2000.
- [8] S. Du, Y. Tao, A.M. Martinez, "Compound facial expressions of emotion", *Proceedings of the National Academy of Sciences*, vol. 111(15), 2014, pp. 1454-1462.
- [9] M.F. Valstar, M. Mehu, B. Jiang, M. Pantic, and K. Scherer, "Meta-Analysis of the First Facial Expression Recognition Challenge", *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 42(4), 2012, pp. 966-979.
- [10] P. Ekman and W.V. Friesen, "Manual for the Facial Action Coding System", *Consulting Psychologists Press*, vol. 3, 1978.
- [11] Y. Tian, T. Kanade, and J.F. Cohn, "Recognizing action units for facial expression analysis", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23(2), 2001, pp. 97-115.
- [12] M.F. Valstar, B. Jiang, M. Mhu, M. Pantic, and K. Scherer, "The First Facial Expression Recognition and Analysis Challenge", *IEEE International Conference on Automatic Face and Gesture Recognition*, Santa Barbara, 2011, pp. 921-926.
- [13] V. Ojansivu and J. Heikkil, "Blur Insensitive Texture Classification Using Local Phase Quantization", *Image and Signal Processing*, vol. 5099, 2008, pp. 236-243.
- [14] S. Yang and B. Bhanu, "Facial expression recognition using emotion avatar image", *IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG 2011)*, Santa Barbara, 2011, pp. 866-871.
- [15] A. Dhall, A. Asthana, R. Goecke, and T. Gedeon, "Emotion recognition using PHOG and LPQ features", *IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG 2011)*, Santa Barbara, 2011, pp. 878-883.
- [16] S.M. Lajevardi and H.R. Wu, "Facial Expression Recognition in Perceptual Color Space", *IEEE Transactions on Image Processing*, vol. 21(8), 2012, pp. 3721-3733.
- [17] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24(7), 2002, pp. 971-987.
- [18] S. Moore and R. Bowden, "Local binary patterns for multi-view facial expression recognition", *Computer Vision and Image Understanding*, vol. 115(4), 2011, pp. 541-558.
- [19] S. Zafeiriou and M. Petrou, "Sparse representations for facial expressions recognition via l_1 optimization", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, San Francisco, 2010, pp. 32-39.
- [20] K. Yu, Z. Wang, M. Hagenbuchner, and D.D. Feng, "Spectral embedding based facial expression recognition with multiple features", *Neurocomputing*, vol. 129, 2014, pp. 136-145.
- [21] S.M.Z. Ishraque, A.K.M.H. Banna, and O. Chae, "Local Gabor directional pattern for facial expression recognition", *15th International Conference on Computer and Information Technology (ICCIT)*, Chittagong, 2012, pp. 164-167.
- [22] P. Lucey, J.F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, San Francisco, 2010, pp. 94-101.
- [23] M. Frank, J. Movellan, M. Bartlett, and G. Littleworth, "Rochester/UCSD FacialActionCodingSystem Database1", Machine Perception Laboratory, U.C. San Diego
- [24] X. Zhang, L. Yin, J.F. Cohn, S. Canavan, M. Reale, A. Horowitz, P. Liu, and J.M. Girard, "BP4D-Spontaneous: a high-resolution spontaneous 3D dynamic facial expression database", *Image and Vision Computing*, vol. 32(10), 2014, pp. 692-706.
- [25] C. Mayer, M. Eggers, and B. Radig, "Cross-database evaluation for facial expression recognition", *Pattern Recognition and Image Analysis*, vol. 24(1), 2014, pp. 124-132.
- [26] X. Feng, "Facial expression recognition based on local binary patterns and coarse-to-fine classification", *The Fourth International Conference on Computer and Information Technology (CIT)*, 2004, pp. 178-183.
- [27] C. Shan, S. Gong, and P.W. McOwan, "Facial expression recognition based on Local Binary Patterns: A comprehensive study", *Image and Vision Computing*, vol. 27(6), 2009, pp. 803816.
- [28] M.R. Mohammadi and E. Fatemizadeh, "Fuzzy local binary patterns: A comparison between Min-Max and Dot-Sum operators in the application of facial expression recognition", *8th Iranian Conference on Machine Vision and Image Processing (MVIP)*, Zanjan, 2013, pp. 315-319.

- [29] R.A. Khan, A. Meyer, H. Konik, and S. Bouakaz, "Framework for reliable, real-time facial expression recognition for low resolution images", *Pattern Recognition Letters*, vol. 34(10), 2013, pp. 1159-1168.
- [30] A. Hafiane, G. Seetharaman, and B. Zavidovique, "Median Binary Pattern for Textures Classification", *4th International Conference on Image Analysis and Recognition*, Montreal, 2007, pp. 387-398.
- [31] T. Jabid, and O.S. Chae, "Local Transitional Pattern: A Robust Facial Image Descriptor for Automatic Facial Expression Recognition", *Proceedings of the International Conference on Computer Convergence Technology*, Seoul, 2011, pp. 3333-3344.
- [32] F. Bashar, A. Khan, F. Ahmed, and M.H. Kabir, "Robust facial expression recognition based on median ternary pattern (MTP)", *International Conference on Electrical Information and Communication Technology (EICT)*, Khulna, 2014, pp. 1-5.
- [33] G. Zhao and M. Pietikainen, "Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29(6), 2007, pp. 915-928.
- [34] X. Huang, G. Zhao, W. Zheng, and M. Pietikainen, "Spatiotemporal Local Monogenic Binary Patterns for Facial Expression Recognition", *IEEE Signal Processing Letters*, vol. 19(5), 2012, pp. 243-246.
- [35] T. Senechal, V. Rapp, H. Salam, R. Segulier, K. Bailly, and L. Prevost, "Facial Action Recognition Combining Heterogeneous Features via Multikernel Learning", *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 42(4), 2012, pp. 993-1005.
- [36] T.R. Almaev and M.F. Valstar, "Local Gabor Binary Patterns from Three Orthogonal Planes for Automatic Facial Expression Recognition", *Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII)*, Geneva, 2013, pp. 356-361.
- [37] W. Gu, C. Xiang, Y.V. Venkatesh, D. Huang, and H. Lin, "Facial expression recognition using radial encoding of local Gabor features and classifier synthesis", *Pattern Recognition*, vol. 45(1), 2012, pp. 80-91.
- [38] T. Ahsan, T. Jabid, U. Chong, "Facial Expression Recognition Using Local Transitional Pattern on Gabor Filtered Facial Images", *IETE Technical Review*, vol. 30(1), 2013, pp. 47-52.
- [39] T. Jabid, H. Kabir, and O. Chae, "Robust Facial Expression Recognition Based on Local Directional Pattern", *ETRI Journal*, vol. 32(5), 2010, pp. 7847-794.
- [40] H. Kabir, T. Jabid, and O. Chae, "A Local Directional Pattern Variance (LDPv) Based Face Descriptor for Human Facial Expression Recognition", *Seventh IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Boston, 2010, pp. 526-532.
- [41] F. Ahmed, "Gradient directional pattern: A robust feature descriptor for facial expression recognition", *Electronics Letters*, vol. 48(19), 2012, pp. 1203-1204.
- [42] X. Wang, C. Jin, W. Liu, M. Hu, L. Xu, and F. Ren, "Feature fusion of HOG and WLD for facial expression recognition", *IEEE/SICE International Symposium on System Integration (SII)*, Kobe, 2013, pp. 227-232.
- [43] J. Li and Z. Ying, "Facial Expression Recognition Based on Rotation Invariant Local Phase Quantization and Sparse Representation", *Second International Conference on Instrumentation, Measurement, Computer, Communication and Control (IMCCC)*, Harbin, 2012, pp. 1313-1317.
- [44] J. Pivrinta, E. Rahtu, and J. Heikkil, "Volume Local Phase Quantization for Blur-Insensitive Dynamic Texture Classification", *Proceedings of the 17th Scandinavian Conference on Image Analysis (SCIA)*, Ystad, 2011, pp. 360-369.
- [45] S.H. Lee, H. Kim, Y.M. Ro, and K.N. Plataniotis, "Using color texture sparsity for facial expression recognition", *10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, Shanghai, 2013, pp. 1-6.
- [46] Y. Wang and M. Meng, "3D Facial Expression Recognition Based on Curvature Local Binary Patterns", *5th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, vol. 2, Hangzhou, 2013, pp. 123-126.
- [47] X. Qi, R. Xiao, C. Li, Y. Qiao, J. Guo, and Xiaou Tang, "Pairwise Rotation Invariant Co-Occurrence Local Binary Pattern", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36(11), 2014, pp. 2199-2213.
- [48] M. Ozuysal, M. Calonder, V. Lepetit, and P. Fua, "Fast Keypoint Recognition Using Random Ferns", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32(3), 2010, pp. 448-461.
- [49] M. Calonder, V. Lepetit, M. Ozuysal, T. Trzcinski, C. Strecha, and P. Fua, "BRIEF: Computing a Local Binary Descriptor Very Fast", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34(7), 2012, pp. 1281-1298.
- [50] J. Zhang, J. Liang, and H. Zhao, "Local Energy Pattern for Texture Classification Using Self-Adaptive Quantization Thresholds", *IEEE Transactions on Image Processing*, vol. 22(1), 2012, pp. 31-42.
- [51] Z. Lei, T. Ahonen, M. Pietikainen, and S.Z. Li, "Local frequency descriptor for low-resolution face recognition", *IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG2011)*, Santa Barbara, 2011, pp. 161-166.
- [52] M. Abdulrahman, T.R. Gwadabe, F.J. Abdu, and A. Eleyan, "Gabor wavelet transform based facial expression recognition using PCA and LBP", *EEE 22nd Signal Processing and Communications Applications Conference*, Trabzon, 2014, pp. 2265-2268.
- [53] T. Ahonen, M. Pietikainen, A. Hadid, and T. Maenpaa, "Face recognition based on the appearance of local regions", *Proceedings of the 17th International Conference on Pattern Recognition (ICPR)*, vol. 3, Cambridge, 2004, pp. 153-156.
- [54] D. Lundqvist, A. Flykt, and A. hman, "The Karolinska Directed Emotional Faces - KDEF", *CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet*, Solna, 1998.
- [55] C. Chang and C. Lin, "LIBSVM: A library for support vector machines", *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2(3), 2011, pp. 27:1-27:27.
- [56] Z.-H. Tan, N.B. Thomsen and X. Duan, "Designing and Implementing an Interactive Social Robot from Off-the-shelf Components", *3rd IFToMM Symposium on Mechanism Design for Robotics (MEDER 2015)*, Aalborg, 2015.