HAL
open science

Customer satisfaction measuring based on the most significant facial emotion

Mariem Slim, Rostom Kachouri, Ahmed Ben Atitallah

To cite this version:

Mariem Slim, Rostom Kachouri, Ahmed Ben Atitallah. Customer satisfaction measuring based on the most significant facial emotion. SSD 2018, Mar 2018, Hammamet, Tunisia. pp.502-507, 10.1109/SSD.2018.8570588 . hal-01790317

HAL Id: hal-01790317

https://hal.science/hal-01790317v1

Submitted on 11 May 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Customer satisfaction measuring based on the most significant facial emotion

Mariem Slim
University of Sfax, LETI laboratory
ENIS, BP W
3038 Sfax, Tunisia
mariem.slim@esiee.com

Rostom Kachouri
University of Paris-Est,
Gaspard-Monge Computer Science Laboratory
A3SI, ESIEE Paris, CNRS, France
rostom.kachouri@esiee.fr

Ahmed Ben Atitallah
University of Sfax, LETI laboratory
ENIS, BP W
3038 Sfax, Tunisia
ahmed.benatitallah@enetcom.usf.tn

***Abstract*—Customer satisfaction (CS) measuring has become one of the strategic tools for companies. Many methods exist to measure customer's satisfaction. However, generally results could be wrong. Otherwise, CS can be deduced from customers emotion state. In this paper, we propose a new end-to-end method for facial emotion detection. For that, six new characteristic features are proposed. We characterize especially the most significant emotions namely "Happy", "Surprised" and "Neutral". The proposed method is invariant to camera position. A very challenging datasets as Radboud Faces and Cohn-Kanade (CK+) are used for study and evaluation. Obtained results show that our method reaches a high recognition accuracy and outperforms Action Unit (AU) features based Support Vector Machine (SVM) and K-Nearest Neighbors (KNN).**

***Index Terms*—Facial expression, emotion recognition, most significant emotions, feature extraction, action unit, SVM, KNN.**

I. Introduction

In an increasingly competitive world, negative reviews of customers can have an important impact on a business. Indeed, customer satisfaction is actually measured in many applications such as shops, banks, mobile apps and others. Many methods and metrics are designed. In instance, Customer Satisfaction (CSAT) is asked using emails, phones, face-to-face interview, tablets and touchpads. As well, Net Promoter Score (NPS) is asking about the intention of product or service recommendation and consumer respond by assigning a rating or a number of stars. However, collected information with these methods could be wrong.

In another way, human interaction is recognized essentially through facial expressions which provide a natural and compact way for humans to convey their emotional state to another party. In addition, nowadays, supervision cameras are everywhere. Thus, we can consider that facial expressions are accessible and we can deduce whether the customer is satisfied or not based on his emotional state.

Across different cultures, we found seven universally recognizable emotions [1]: anger, disgust, fear, happy, sad, surprised and neutral. However, in this work, we note that no need to recognize all emotions. In fact, for customer satisfaction, only "Happy", "Surprised" and "Neutral" are significant. In this paper, we propose an original and performing method for customer satisfaction measuring based on the most significant facial emotion.

The organization of the paper is as follows: In section II, related works are presented. Then, we propose a novel facial emotion recognition in Section III. Experimental results are discussed in Section IV and finally conclusions are drawn in Section V.

II. Related works

For the first time, Charles Darwin demonstrated that despite their differences and countries, human beings have innate and common emotion [2]. Historically, there are seven types of human emotions: anger, disgust, fear, happy, sad, surprised and neutral (Fig. 1). Paul Ekman later conducted several experiments and observations on a large number of individuals and confirmed that emotions are universal recognizable expression [2].

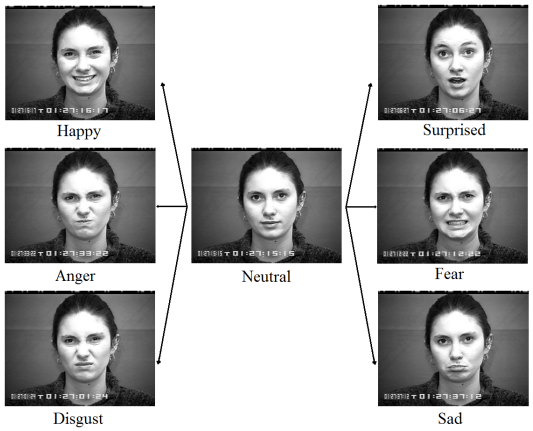


Fig. 1: The basic emotions.

Many approaches are proposed to recognize facial expressions. They can be split into two broad categories: appearance-based approaches and features-based approaches. Appearance-based approaches detect emotion via a model which illustrate the general facial shape and texture ([3], [4]). On the other hand, features-based approaches extract person's characteristics. Then, extracted features can be combined and classified

[5]. Generally, appearance-based approaches are very time-consuming and the human face appearance could be affected due to variations of lighting conditions. Therefore, we consider in this paper features-based approach.

Fig. 2: The three facial recognition modules.

Facial emotion recognition systems comprised essentially three steps [6] (Fig. 2). The first step is face and landmarks detection. For that, several open source facial behavior analysis systems exist. In instance, CHEHRA[1] [7], TCDCN[2] face alignement tool [8], CLandmark[3] [9], OpenFace[4] [10]... etc. We note that OpenFace is the most recent fully real time and open source system [10].

Feature extraction is the second step of facial emotion recognition system. P. Ekman and W. Friesen developed the Facial Action Coding System (FACS) to describe movements (action units) caused by expansion and contraction of facial muscles [11]. Using FACS, human coders can manually code almost any anatomically possible facial expression [12]. More recently, Hammal et al propose the use of just 5 characteristic distances as features [5]. They supposed that all the necessary information is contained in the castration of the eyes, mouth and eyebrows.

The final step is emotion recognition. Many features based emotion classification methods exist. Two main categories can be identified: supervised and unsupervised classification techniques [13]. Unsupervised learning or clustring don't require human intervention and don't need a prior knowledge or predefined categories [14]. In the other side, supervised learning requires an expert labeled database. The process is performed in two phases. During the learning phase, a model is determined from the labeled data. Using the learned model, the test phase called consist on predicting the label of a new non labeled data [15]. We note that, supervised learning is more efficient and used to resolve linear and non-linear classification problems [16]. The most popular supervised machine learning algorithms [17] are: Support vector machine (SVM) [18], k-nearest neighbors (KNN) [19], Convolutional Neural Networks (CNN) [20]. However, generally these methods could require a large number of attempts to move towards the best possible recognition performance [21].

Up to our knowledge, no facial emotion recognition model has yet proposed an affordable and performing system for customer satisfaction. Otherwise, our method deals with emotion recognition from images, and focus on feature extraction and analyze. To measure customer satisfaction, we note that in the rest of this paper, we consider only three significant emotions: "Happy", "Surprised" and "Neutral".

[1]https://sites.google.com/site/chehrahome/
[2]https://github.com/zhzhanp/TCDCN-face-alignment
[3]https://github.com/uricamic/clandmark
[4]https://github.com/TadasBaltrusaitis/OpenFace

III. Proposed method

Our goal is to propose a feasible and especially performing method. We use OpenFace for face and landmarks detection (see subsection III-A). Then, we present our extracted features in subsection III-B. To recognize the most significant emotions, we propose in subsection III-C a new approach to classify features. To develop a working model, we use Radboud Faces Database [22]. It is a highly standardized set of pictures. It contains 67 subjects displaying 8 emotional expressions.

A. Face and landmarks detection

OpenFace is an open source real-time facial behavior analysis system [10]. As shown in Fig. 3, it is able to estimate head pose and detect facial landmarks. It uses for that the recently proposed Conditional Local Neural Fields (CLNF) [10]. OpenFace can also recognize facial action unit and estimate eye-gaze. It uses a modern C ++ toolkit called Dlib library [26]. It can be used free of charge in any application because Dlib's license is free.

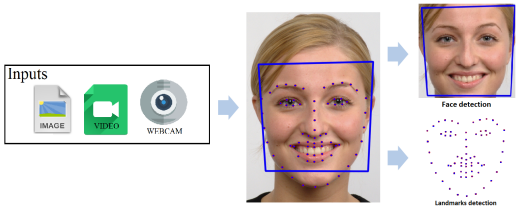


Fig. 3: OpenFace: face and landmarks detection.

The provided landmarks are 68 points on the face such as the outline of the mouth, the eyes, the eyebrows, the noise and the facial contour. The facial landmark detection aptitude was evaluated on four sub-dataset Annotated Faces in the Wild (AFW) [27], IBUG [28], LFPW [29], and Helen [30]. It was also compared to other facial landmark detection algorithms. Implementations' approaches are available online and have been trained to detect the same facial landmarks (or their subsets). The baselines were: Discriminative Response Map Fitting (DRMF) [31], tree based deformable models [32], extended version of Constrained Local Models [33], GaussNewton Deformable Parts Model (GNDPM) [34], and Supervised Descent Method (SDM) [35]. The train and test datasets are different for all of the methods. The obtained results affirmed that OpenFace is state-of-the-art performance.

B. Feature extraction

Our method is based on feature-based approaches. We proposed seven features that we extract from the previously presented face and landmarks. First of all, as shown in Fig. 4, we compute the frame's diagonal (D) from the detected face. This feature corresponds to the customer's position from camera. Using the feature D makes our analysis invariant to the position of the face with respect to the camera.

In addition to the diagonal (D), as illustrated in Fig. 5, new characteristic features named F_1 to F_6 are defined as

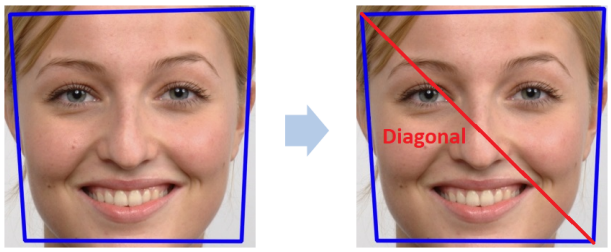


Fig. 4: The frame's diagonal (D).

a variation of some facial landmarks provided by OpenFace. Where:

- F_1 = (F_{1L} + F_{1R}) / 2: the opening of the eyes,
- F_2 = (F_{2L} + F_{2R}) / 2: the distance between the eyebrow and the corner of the eye,
- F_3 : the width of the mouth,
- F_4 : the opening of the mouth,
- F_5 = (F_{5L} + F_{5R}) / 2: the distance between the nose and the tip of the mouth,
- F_6 : the length of the lower inner lip.

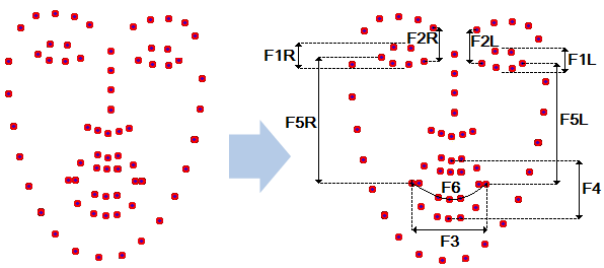


Fig. 5: Our feature extraction (F_1 to F_6).

C. Emotion recognition

Our aim is to characterize each "Happy", "Surprised" and "Neutral" emotion by a specific hyper-planes which allow to recognize the corresponding emotion depending on the associated characteristic features. The conducted hyper-planes have to be invariant to the frame's diagonal (D) of each face. Radboud facial database[5] [22] is used. First, we evaluate each extracted features according to the number of images (see Fig. 8(a)-(f)). According to this feature study and based on predicted behavior of each extracted feature for the most significant emotions, we present in the following our proposed hyper-planes for each emotion.

1) Happy emotion: We note that in this case, F_1 have to be small because eyes are little closed. In addition, generally, smiling increases the values of F_3 and F_6 and decreases F_5's value. Equation 1 presents our proposed feature combination F_h to well distinguish the "Happy" emotion:

$$F_h = F_3 + F_6 - F_5 - F_1 \tag{1}$$

[5]http://www.socsci.ru.nl:8180/RaFD2/RaFD?p=main

Fig. 6 illustrates the evaluation of the new F_h representation of the most significant emotions versus the frame's diagonal D. We can see that the "Happy" emotion curve can be entirely separated from the two other emotions according to the affine hyper-plane Y_h (equation 2).

$$Y_h = A_h * D + B_h \tag{2}$$

Where Ah = 0.21 is the slope and B_h = -1 is the Y_h-intercept.

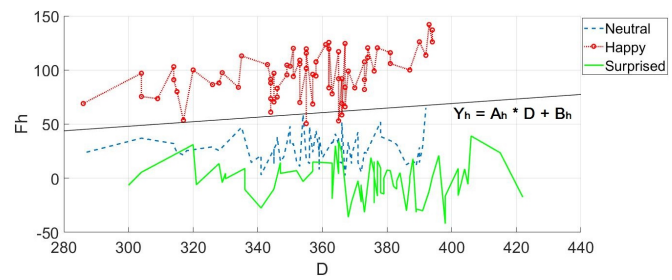


Fig. 6: F_h representation of the most significant emotions versus the frame's diagonal D within Radboud Face database.

Finally, the recognition of the "Happy" emotion is ensured through the equation 3:

$$F_h > Y_h ==> \text{“}Happy\text{”} emotion \tag{3}$$

2) Surprised emotion: In this case, we note that F_2 value have to be high because eyebrows are raised. In addition, the mouth is like the letter "O" so it increases the value of F_4 and decreases F_3 value. Thus, the "Surprised" emotion is distinguished by the feature combination F_s (equation 4):

$$F_s = 2 * F_2 + 3 * (F_4 - F_3) \tag{4}$$

The evaluation of the new F_s representation of the most significant emotions versus the frame's diagonal D is illustrated in Fig. 7.

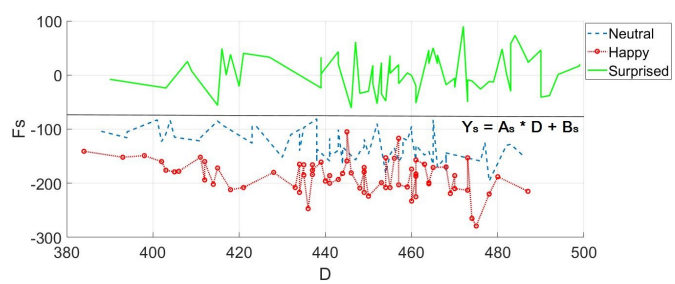
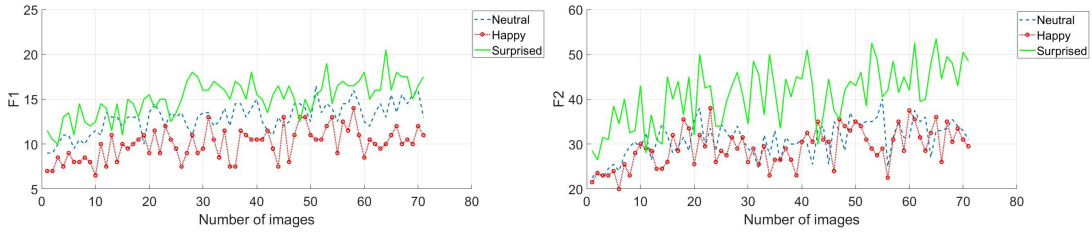


Fig. 7: F_s representation of the most significant emotions versus the frame's diagonal D within Radboud Face database.

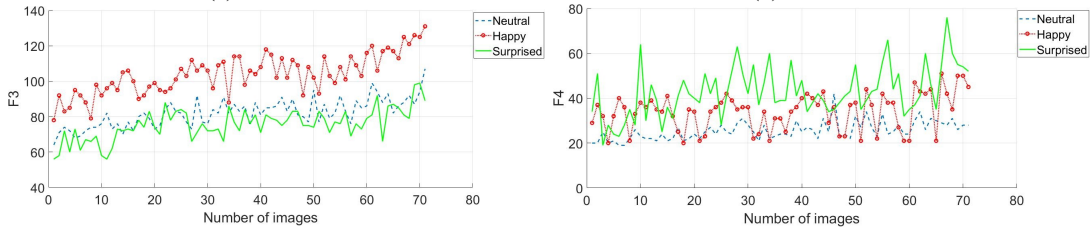
The "Surprised" emotion curve can be entirely separated from the two other emotions (Fig. 7) according to the affine hyper-plane Y_s (equation 5).

$$Y_s = A_s * D + B_s \tag{5}$$



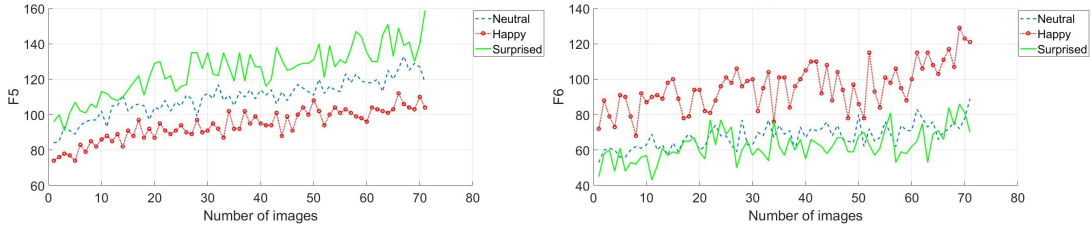
(a) F_1 curves.

(b) F_2 curves.



(c) F_3 curves.

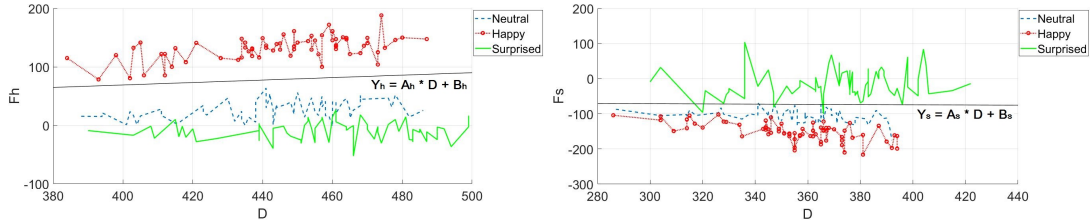
(d) F_4 curves.



(e) F_5 curves.

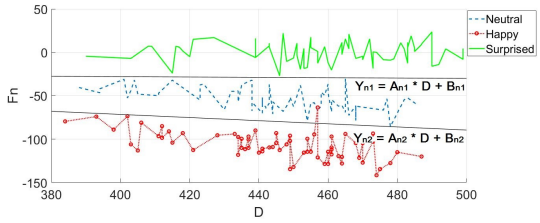
(f) F_6 curves.

Fig. 8: Behaviour feature study: (a) F_1, (b) F_2, (c) F_3, (d) F_4, (e) F_5, (f) F_6, curves of the most significant emotions according to 3*67 images from Radboud Face database.



(a) “Happy” emotion identification.

(b) “Surprised” emotion identification.



(c) “Neutral” emotion identification.

Fig. 9: The three proposed feature representation of the most significant emotions versus the frame’s diagonal D within CK+ database: (a) F_h for “Happy” emotion (b) F_s for “Surprised” emotion and (c) F_n for “Neutral” emotion.

Where As = -0.03 is the slope and B_s = -62.2 is the Y_s-intercept.

The recognition of the "Surprised" emotion is ensured through the equation 6:

$$F_s > Y_s ==> \text{“}Surprised\text{”} emotion \tag{6}$$

3) Neutral emotion: To differentiate "Neutral" emotion, F_2 and F_6 have to be smaller compared to their respectively "Surprised" and "Happy" values. In addition, we note that the closed mouth in this case decreases F_4 value. Then, we consider that feature combination F_n (equation 7) allows to well distinguish the "Neutral" emotion:

$$F_n = F_4 + F_2 - 1.5 * F_6 \tag{7}$$

Fig. 10 illustrates the evaluation of the new F_n representation of the most significant emotions versus the frame's diagonal D. We can see that the "Neutral" emotion curve can be entirely separated from the two other emotions according to the two affine hyper-planes Y_{n1} and Y_{n2} (equations 8 and 9).

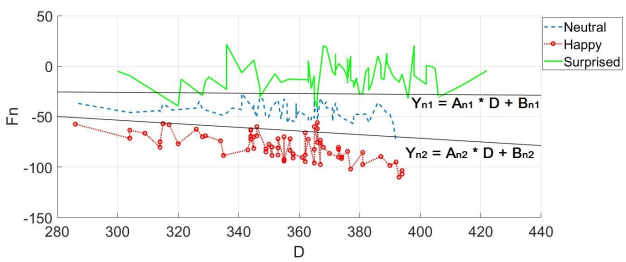


Fig. 10: F_n representation of the most significant emotions versus the frame's diagonal D within Radboud Face database.

$$Y_{n1} = A_{n1} * D + B_{n1} \tag{8}$$

$$Y_{n2} = A_{n2} * D + B_{n2} \tag{9}$$

Where A_{n1} = -0.02, B_{n1} = -20, A_{An2} = -0.18 and B_{n2} = 0.5.

The recognition of the "Neutral" emotion is finally ensured through the equation 10.

$$Y_{n1} > F_n > Y_{n2} ==> \text{“}Neutral\text{”} emotion \tag{10}$$

IV. Experimental results

Many facial expression databases (e.g. The Cohn-Kanade database [23], Radboud Faces Database [22], the MMI-Facial Expression Database [24], and the JAFFE database [25]) exist. In our experiments, we use a well known and popular facial expression dataset. Indeed, the extended Cohn-Kanade database (CK+)[6] [23] is used in this work to evaluate the performance of our method. This database have a large number of images with expressions' labels. The CK+ database contains 327 image sequences for "Happy", "Surprised" and "Neutral" emotions. To develop our method, we use Microsoft Visual Studio 2015 to employ OpenFace and MATLAB R2017a for drawing curves and illustrate results.

Fig. 11 and 12 shows the performance of our method on both Radboud Faces and CK+ databases. We note that our reached accuracy is higher then 94% for all the three used most significant emotions.

Fig. 9(a)-(c) confirm on CK+ database the previously conducted study using Radboud Faces Database. That validates our original feature presentation F_h, F_s and F_n and proposed affine hyper-planes Y_h, Y_s, Y_{n1} and Y_{n2}.

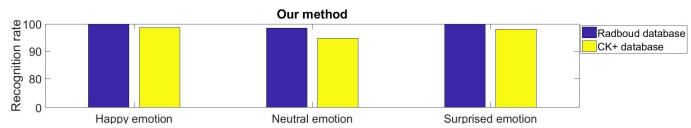


Fig. 11: The recognition rate of our proposed method within respectively the Radboud Face and the CK+ database for the 3 most significant emotions.

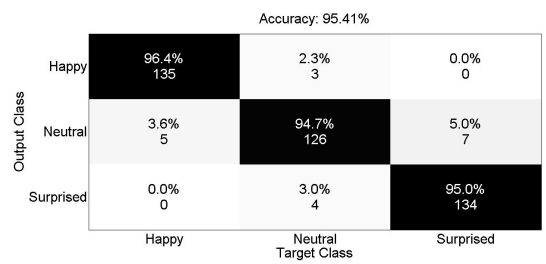
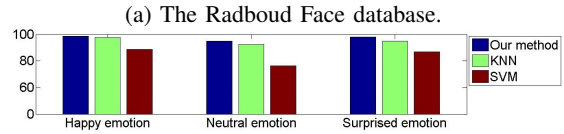


Fig. 12: The confusion matrix of our proposed method within the CK+ database for the 3 most significant emotions.

For comparison reasons, we employ two machine learning methods SVM [18] and KNN [19]. Facial Action Coding System (FACS) [11] are used as features with these two classifiers. For fair comparison, we use Radboud dataset for training and CK+ dataset for prediction. Obtained results illustrated in Fig. 13(a)-(b) and presented in table I prove that our proposed method outperform SVM and KNN classifiers.

(a) The Radboud Face database.



(b) The CK+ database.

Fig. 13: The obtained recognition rates of action unit based on SVM and KNN classifiers within (a) the Radboud Face and (b) the CK+ databases for the 3 most significant emotions.

[6]http://www.consortium.ri.cmu.edu/ckagree/

Table I shows the obtained average accuracy (Av Acc) of the three compared methods. We can see that our method reach 97.02% and outperforms SVM (84.04%) and KNN (95.32%) classifiers.

TABLE I: Our proposed method, SVM, KNN recognition rate comparison within CK+ database

Method \ Emotion	Av Acc	Happy	Surprised	Neutral
Our method	**97.02%**	**98.59%**	**94.83%**	**98.12%**
KNN + AU	95.32%	97.65%	92.49%	94.84%
SVM + AU	84.04%	88.73%	76.53%	86.85 %

V. Conclusion

We propose in this paper a new and performing method for customer satisfaction measuring based on the most significant facial emotion. Six new characteristic features are used and original feature representation and hyper-planes emotion separation are proposed. The proposed method is invariant to camera position. A very challenging datasets as Radboud Faces and Cohn-Kanade (CK+) are used for study and evaluation. Obtained results show that our method reaches a high recognition accuracy and outperforms Action Unit (AU) features based Support Vector Machine (SVM) and K-Nearest Neighbors (KNN).

References

[1] Mistry, Kamlesh Zhang, Li Neoh, Siew Peng Lim, Chee Fielding, Ben. (2016). A Micro-GA Embedded PSO Feature Selection Approach to Intelligent Facial Emotion Recognition. IEEE Transactions on Cybernetics.

[2] Darwin Charles, Ekman Paul, Prodger Phillip (1998) The Expression of the Emotions in Man and Animals, 3rd edn, London: Harper Collins.

[3] Naveen Kumar H N, Jagadeesha S and Amith K Jain, Human Facial Expression Recognition from static images using shape and appearance feature, Applied and Theoretical Computing and Communication Technology (iCATccT), 2016 2nd International Conference

[4] K. Mistry, L. Zhang, S.C. Neoh, M. Jiang, A.r Hossain and B. Lafon, Intelligent Appearance and shape based facial emotion recognition for a humanoid robot, Software, Knowledge, Information Management and Applications (SKIMA), 2014 8th International Conference

[5] Z. Hammal. Segmentation des Traits du Visage, Analyse et Reconnaissance des Expressions Faciales par les Modeles de Croyance Transferable ; Traitement du signal et de l'image, Universite Joseph-Fourier - Grenoble I, 2006. Francais.

[6] L. Greche, N. Hamaoui and N. ES-Sbai, Facial expression recognition on Android, Mediterranean Telecommunication Journal, Vol. 5, N 2, June 2015y

[7] A. Asthana, S. Zafeiriou, S. Cheng and M. Pantic. Incremental Face Alignment in the Wild. In CVPR 2014.

[8] Zhanpeng Zhang, Ping Luo, Chen Change Loy, Xiaoou Tang. Facial Landmark Detection by Deep Multi-task Learning, in Proceedings of European Conference on Computer Vision (ECCV), 2014

[9] M. Uricar, V. Franc, D. Thomas, A. Sugimoto, and V. Hlavac, Multi-view facial landmark detector learned by the Structured Output SVM, Image and Vision Computing, 2016.

[10] P. Robinson, and L. P. Morency, OpenFace: an open source facial behavior analysis toolkit Tadas Baltruaitis, in IEEE Winter Conference on Applications of Computer Vision, 2016.

[11] Mousavi, Nima Siqueira, Henrique Barros, Pablo Fernandes, Bruno Wermter, Stefan. (2016). Understanding how deep neural networks learn face expressions. Conference: 2016 International Joint Conference on Neural Networks (IJCNN)

[12] Freitas-Magalhes, A. (2012). Microexpression and macroexpression. In V. S. Ramachandran (Ed.), Encyclopedia of Human Behavior (Vol. 2, pp. 173183). Oxford: Elsevier/Academic Press. ISBN 978-0-12-375000-6

[13] M. Cord, P.Cunningham, M. Cord, P. Cunningham, Machine Learning Techniques for Multimedia, Cunningham, P. (Eds.), 2008, XVI, 289p., Hardciver,

[14] Zoubin Ghahramani, Unsupervised learning, Advanced Lectures on Machine Learning 3176, 72–112

[15] S. B. Kotsiantis. 2007. Supervised Machine Learning: A Review of Classification Techniques. In Proceedings of the 2007 conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in eHealth, HCI, Amsterdam, The Netherlands, The Netherlands, 3-24.

[16] R. Sathya and Annamma Abraham, Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification International Journal of Advanced Research in Artificial Intelligence(IJARAI), 2013.

[17] P. Liu and L. Yin, Spontaneous facial expression analysis based on temperature changes and head motions, in FG, 2015.

[18] Auria, Laura and Moro, R. A., Support Vector Machines (SVM) as a Technique for Solvency Analysis (August 1, 2008). DIW Berlin Discussion Paper No. 811.

[19] W. Gu, C. Xiang, Y. Venkatesh, D. Huang, and H. Lin, Facial expression recognition using radial encoding of local gabor features and classifier synthesis, PR, vol. 45, no. 1, pp. 8091, 2012.

[20] Nikhil Churamani, Matthias Kerzel, Erik Strahl, Pablo Barros, Stefan Wermter.: Teaching Emotion Expressions to a Human Companion Robot using Deep Neural Architectures. International Joint Conference on Neural Networks (IJCNN), pages 627–634 - May 2017. To appear on IEEE Xplore

[21] Ruiz-Garcia, Ariel Elshaw, Mark Altahhan, Abdulrahman Palade, Vasile. (2017). Stacked deep convolutional auto-encoders for emotion recognition from facial expressions. 1586-1593.

[22] Langner, Oliver , Dotsch, Ron , Bijlstra, Gijsbert , Wigboldus, Daniel H. J. , Hawk, Skyler T. and van Knippenberg, Ad(2010) 'Presentation and validation of the Radboud Faces Database', Cognition & Emotion, 24: 8, 1377 1388

[23] Kanade T, Cohn J, Tian Y: Comprehensive Database for Facial Expression Analysis. 4th IEEE International Conference on Automatic Face and Gesture Recognition, France 2000.

[24] M. Valstar and M. Pantic. Induced disgust, happiness and surprise: an addition to the mmi facial expression database. Proc. 3rd Intern. Workshop on EMOTION (satellite of LREC): Corpora for Research on Emotion and Affect, page 65, 2010.

[25] M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba. Coding facial expressions with gabor wavelets. Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on, pages 200205, 1998.

[26] Davis E. King, Dlib-ml: A Machine Learning Toolkit, Journal of Machine Learning Research 10 (2009) 1755-1758

[27] X. Zhu and D. Ramanan. Face detection, pose estimation, and landmark localization in the wild. In CVPR, 2012.

[28] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic. 300 faces in-the-wild challenge: The first facial landmark localization challenge. In ICCV, 2013.

[29] P. N. Belhumeur, D. W. Jacobs, D. J. Kriegman, and N. Kumar. Localizing parts of faces using a consensus of exemplars. In CVPR, 2011.

[30] V. Le, J. Brandt, Z. Lin, L. Bourdev, and T. S. Huang. Interactive facial feature localization. In ECCV, 2012.

[31] A. Asthana, S. Zafeiriou, S. Cheng, and M. Pantic. Robust discriminative response map fitting with constrained local models. In CVPR, 2013.

[32] X. Zhu and D. Ramanan. Face detection, pose estimation, and landmark localization in the wild. In CVPR, 2012.

[33] T. Baltrusaitis, N. Banda, and P. Robinson. Dimensional affect recognition using continuous conditional random fields. In FG, 2013

[34] G. Tzimiropoulos and M. Pantic. Gauss-Newton Deformable Part Models for Face Alignment In-the-Wild. Computer Vision and Pattern Recognition, (c):18511858, 2014.

[35] X. Xiong and F. De la Torre. Supervised descent method and its applications to face alignment. In CVPR, 2013