International Journal of Natural and Engineering Sciences
E-ISSN: 2146-0086 13 (1): 01-06, 2019

Geometry Based Three Dimensional Face Expression Recognition

Neşe GÜNEŞ[1*]

[1]Department of Computer Engineering, Istanbul Technical University, Istanbul, TURKEY

*Corresponding Author
E-posta:nese.gunes@itu.edu.tr

Received: 30 August 2019
Accepted: 07 October 2019

Abstract

This study discusses the problem of three-dimensional face expression recognition and presents a novel approach for recognizing facial expressions in the presence of three-dimensional geometry of face models. As a fundamental step in mesh processing pipeline using 3D scanned data, we first perform data alignment between different meshes. After registration step, we apply mesh smoothing algorithms which filter out high frequency noise and Taubin smoothing is used for regularization. Proposed approach is based on modeling surface geometry with completely regular triangle meshes called geometry images. Results demonstrating face expression recognition rates using end-to-end learning are presented and a three-dimensional facial expression database is used to provide a series of experiments.

Keywords: Mesh Processing, Conformal Mapping, Geometry Images, Biometrics

INTRODUCTION

Face expression analysis has attracted numerous scientists because of its several applications and purposes. It takes part in a significant role in expression analysis therefore plays a part in human-computer interaction systems' development. In addition, it can strengthen facial recognition systems through supplying preliminary understanding on face movements and facial feature distortions. This is especially fascinating regarding as the cheek area involves remarkable number of discriminatory details, further it is where the majority of the facial distortions occur. Various utilizations contain however are not restricted to psychology related analysis, facial animation, fatigue detection, virtual reality as well as robotics. Face expressions are produced when face muscles contract as a result of that short-term facial distortion in both facial texture and geometry occur.

Previously, the center of attention of expression recognition has been the two-dimensional domain because of the widespread presence of data such as videos and images. Although 2D facial expression recognition (FER) systems have carried out exceptional achievement, even now we confront challenges such as illumination and pose variations in 2D expression recognition. However, three-dimensional data, are not varying according to these changes and are providing a great deal of information through its nature.

In 2006, the BU-3DFE database [3] was released and since then 3D FER has attracted significant attention.

In pursuit of giving an answer to the following question, we study on 3D face expression recognition task: "Can we employ geometry images to have a geometry based solution to the task of 3D face expression recognition using the power of neural networks as well?" The article has been put in writing to share obtained results.

The article studies the recognition of 3D facial expressions when a subject is attempting to show her/his emotions. These expressions make the detection and recognition problem difficult for inexperienced people. A motivation of 3D face expression recognition in computer vision is an automatic recognition of human behavior. A large variety of disciplines may benefit from revealing the phenomenon, e.g. human robot interaction, security services, psychologists, teachers, etc.

We propose a pipeline to recognize 3D facial expressions based on geometry images of 3D face models. Geometry images are obtained opening up a 3D mesh onto a square domain. The expression phenomenon results in a change in geometry based approach. The proposed method uses VGG deep neural network model pre-trained on ImageNet to make accurate expression classification.

The article is structured as follows. We first introduce the task of 3D facial expression. Then we give detailed

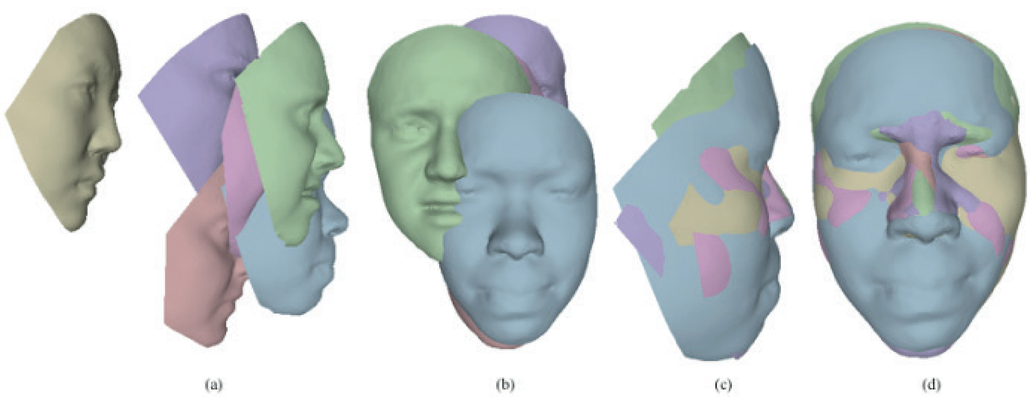

Figure 1. Depiction of 3D registration. (a) The original face scans with different poses, side view. (b) The original face scans with different poses, frontal view. (c) The resulting transformation after alignment, side view. (d) The resulting transformation after alignment, frontal view.

information about mesh processing pipeline. Moreover, we describe the proposed method using deep neural networks. We furthermore present the proposed pipeline for 3D facial expression recognition with implementation details. Finally, we conclude the study.

A 3D Face Expression Recognition Framework

In the past, a list of typical properties of a face expression recognition framework [4] has been made. The framework of interest will profit from solution methods to several computer vision research problems such as landmark localization, illumination normalization and face detection. Furthermore, based on the modality of the data, in addition to corresponding challenges, the methods to accomplish these objectives can be diverse. For example, using 3D data it is achievable to handle a large scale of rigid head movements, additionally to extract uniform geometric features which bypasses the pose estimation problem. By its nature, three-dimensional data are invariant to lighting changes. Thus, a unit appointed to handle changes in illumination is not desired anymore. Nonetheless, when a completely automated FER framework is required, we shall select the proper unit or supply arrangement to handle these tasks based on the modality of the data for instance 3D data with associated texture. Various existing studies utterly address the key problems of 3D FER for instance expression classification and feature computation instead of constructing a general automatic framework. For evaluation, we use the BU-3DFE database which provides cropped 3D models therefore face detection is not a necessary step anymore. Furthermore, we skip one extra step due to manually annotated landmark set from which the feature extraction will be done.

The initial research on this area has existed since the seventies accompanied by the front runner study established previously. Within this work, this is showed that several principal facial emotions occur that might be classified into six categories, that is to say, sadness, happiness, disgust, anger, surprise, fear and as well as the neutral appearance. This classification of face expressions has been also demonstrated to be compatible over several nationalities and societies; therefore, these emotions are in certain perception in all cases recognizable.

At a recent time, there has been an increasing shifting from 2-Dimensional to 3-Dimensional in face recognition viewpoints, mostly inspired by the strength of the 3-Dimensional face shape model to lighting variations, posture, and size changes. In spite of the fact that numerous research has performed to accomplish 3-D facial analysis, a small number of have used benefit of the 3D surface geometry data to conduct facial emotion analysis. Previously, the initial results to accomplish facial expression analysis in an automatic way using 3D facial scans were suggested employing extremely small-scale datasets and classifying just a certain number of facial emotions.

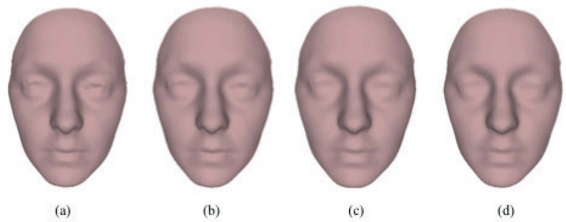

Figure 2. Mesh smoothing. (a) The original mesh. (b) Taubin smoothing is performed for 10 iterations. (c) 50 iterations. (d) 100 iterations.

The accessibility of novel facial expression datasets, such as those established at the Bosphorus dataset (Boğaziçi University) moreover at the BU-3DFE dataset have currently moved the study on the concept forwards. Specifically, the BU-3DFE dataset has begun to be the existing level of quality for measuring the similarity of facial expression analysis methods. For the reason that different from variants of 3D face databases, the BU-3DFE database supplies an exact classification of 3D facial models as specified via six simple face emotions and neutral expression, as well as making available distinct amount of emotion strengths.

3D Registration

Our approach on face expression recognition shows that 3D registration (alignment) is a fundamental step especially in geometric approaches. Actually, an error in alignment might not be fixed in the next steps of our method or other different approaches. Eventually, we represent a registration technique which provides accurate and robust alignment even with the facial expressions being present. The traditional idea is that we apply pose correction to each mesh using the same face model as a reference and as primary step in mesh processing pipeline.

The alignment evaluates a rigid transformation

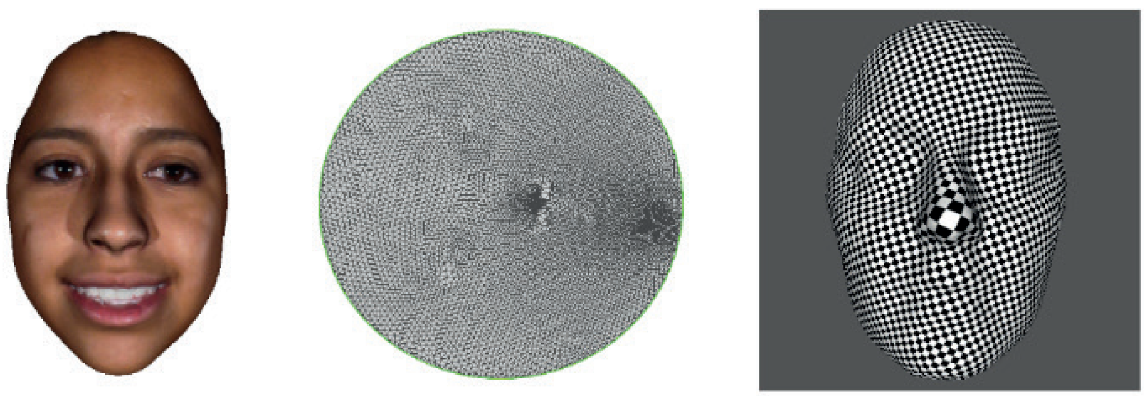

Figure 3. Fixed Boundary Conformal Mapping on a circle. (L) Original mesh displayed with G3dOGL.exe. (M) The mesh is unfolded in UV space, also known as parameter space. (R) Parameterized mesh, textured mesh with checkers.

which involves rotation and translation. Our alignment method employs the Iterative Closest Point algorithm. The algorithm deals with the alignment task through reducing the distance between two different oriented triangle meshes. The resulting transformation is shown in Figure 1 and next step uses the output of this step as an input.

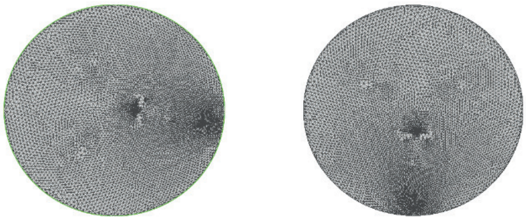

Figure 4. Alignment to the ground in parameter space. (L) Conformal map wire framed and displayed in Meshlab. The boundary edges are shown in green. (R) Aligned planar mesh.

Taubin Smoothing

The smoothing algorithm employed in this study performs a back and forward Gaussian smoothing without shrinking of a triangulated mesh as described by Gabriel Taubin [5]. Firstly, using a positive scale factor λ, Gaussian smoothing is applied through all the vertices of the mesh. Afterwards, a consecutive smoothing step is also applied using a negative scale factor μ, regarding λ is less than μ in magnitude ($0 < \lambda < -\mu$). These two steps must be applied a number of times to produce a significant smoothing effect. Figure 2 shows some examples of applying Taubin smoothing algorithm to facial expression models from BU-3DFE data.

Conformal Mapping

Mesh parameterization describes the procedure for mapping a three-dimensional triangular mesh over a two-dimensional (flattened) domain, the majority of mesh parameterization algorithms use the concept differential geometry as a basis. A conformal parameterization represents a three-dimensional surface on to a planar two-dimensional domain, in such a way that the parameterization is angle-preserving, or identically, conformal mapping directly maps extremely small circles on the 3D surface to the extremely small circles on the 2D plane. It is shown in the subsequent Figure 3, the three-dimensional human face surface is mapped on to the unit planar disk using a conformal mapping algorithm. In Figure 3, we place the checker board texture on the flattened disk, and draw it back on to the 3D face surface, so that we preserve all the right corner angles of the checkers. In a similar fashion, we place a circle packing texture on the planar disk, afterwards draw it back on to the 3D face surface, eventually we preserve all the small circles. As specified by uniformization theory, every surface in actual life could be parameterized conformally to one of three recognized shapes, the hyperbolic space, the plane, the sphere.

Alignment in Parameter Space

In UV space, we first compute the boundary positions of the conformal maps. In Figure 4, the boundary is shown in green. Then we rotate the coordinates of all the vertices along with the boundary positions. The rotation is not relative to each other but relative to the ground and the transformation performs alignment to the ground in parameter space.

Color Interpolation using Barycentric Coordinates

In 3D Modelling, mostly triangles are used. Triangles are stored as a sequence of three vertices. We generally know information about the vertices, for instance color. We should note that Barycentric coordinates can be thought as an internal coordinate system for a triangle. Hence, we use Barycentric coordinates to define the color of an arbitrary point inside the triangle. Barycentric coordinates allow us to interpolate over the whole triangle. For each triangle, we first compute the mean value coordinates, and we use them to interpolate the data from the boundary of the unit disk to its interior.

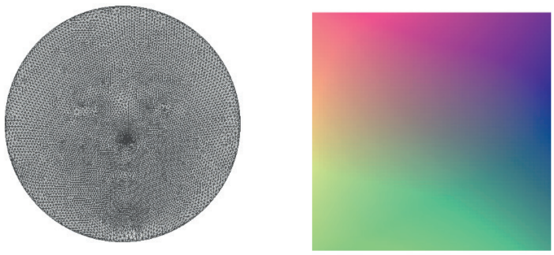

Figure 5. Interpolating inside a planar mesh. (L) Flattened mesh, output of the previous steps. (R) The interpolated data: Geometry pixels.

Geometry Image Representation

The standard geometric representation used in today's graphics hardware is the irregular mesh. A mesh consists of an array of triangles, and an array of vertices, where each triangle

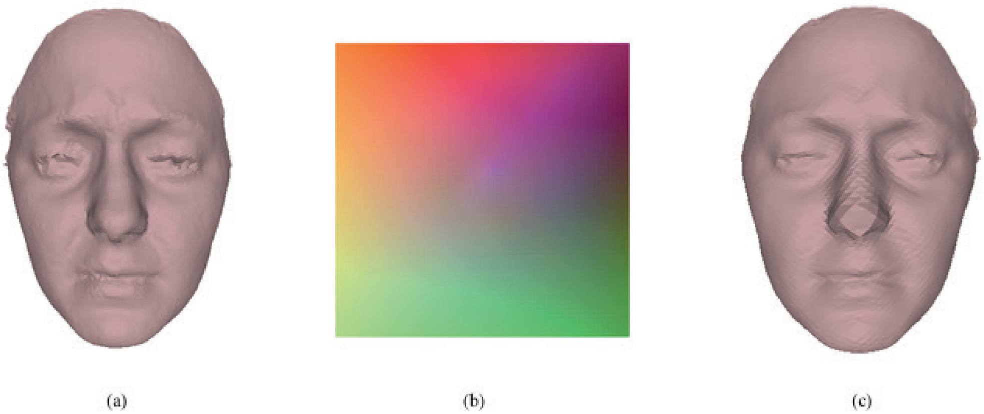

Figure 6. Rendering of a geometry image. (a) The initial face scan. (b) Geometry image is visualized as [r, g, b], 12 bits/channel. (c) Reconstruction of geometry, obtained completely from (b).

refers to three vertex indices. To represent more detail over meshes, it is common to use texture mapping. The mesh vertices are assigned texture coordinates, defining a parametrization of the mesh onto an UV domain. An image is placed on this domain, which is then mapped back to the surface. In this case, the texture image represents surface normals, which can be used in per-pixel shading. Previous work in remeshing has gone part of the way in creating more regular geometric representations. But, such methods still use an irregular base mesh, which leads to a sampling that is only semi-regular. Our approach is to employ an arbitrary surface using a completely regular grid of samples on a square domain which is called a geometry image. In Figure 3, as you can see, it looks just like an ordinary image. It has 257 by 257 pixels. The only difference is that it has 12 bits per channel instead of the usual 8. It is called a geometry image because its RGB colors encode XYZ positions, and in fact entirely describe the 3D face model shown in part (c) in Figure 6.

The outline of the fundamental steps involved in creating a geometry image is as follows: First, we open up the mesh through a proper cut path sets, thus we create a surface having the topology of a disk. Second, parametrization is performed to this disk surface to map it into the square domain of the geometry image. Next, we overspread a regular 2D sampling grid over the domain and map these samples through the parametrization back onto the surface. Noting that it is important to create a good cut and a good parametrization, so that the samples are evenly distributed over the surface. Then, we compute the XYZ positions of the surface samples, as well as other surface attributes and store these in the geometry image. Therefore, we visualize the XYZ data as RGB colors. Given a geometry image, it is straightforward to render it as shaded 3D geometry.

Figure 7. Seven Expressions with a variety of ethnic ancestries: The BU-3DFE Database.

Consequently, geometry images are a powerful, completely regular representation of triangular meshes, describing each vertex and its attributes as a pixel in one or more rectangular images and thus defining an implicit connectivity through neighbouring pixels. It can be understood as a next step after color and normal mapping, representing the whole model as a set of images

.

Performance Evaluation

In this part, we concisely discuss the database we employed to assess our approach for expression recognition. Then we present the outcomes of our experiments utilizing this database and make comparisons with other experimental set-ups.

Until now, for expression analysis, there exists three unrestricted 3D databases: the Bosphorus, the BU3DFE and the BU-4DFE databases. We should indicate that we think of a database addressed to expression recognition as if it includes datasets showing six typical expressions or various action units of the facial model. Various facial databases including GavabDB [6] and FRGC v2 [7] and are hardly utilized for the purpose of emotion recognition even though they consist of expression variations, because of an insufficient set of expressions or a non-uniform distribution of different expressions.

Figure 8. BU-3DFE dataset: textured 3D facial scan of a participant performing happy face expression at each stage of different intensities (low to highest) also neutral expression.

The BU-3DFE Database

Due to being freely accessible by the research communities, the BU-3DFE database [3] is used to assess the majority of the current 3D FER frameworks. Additionally, the manually annotated facial landmarks are supplied along with the database release. There exist 100 subjects in the database, 56% of them is female and 44% of them is female. Each subject performed six expressions and a neutral expression. The six typical expressions are captured at different intensities from ground to peak at four levels. The database includes the original face scans along with corresponding texture images and it also involves the cropped face models from these raw scans. Therefore, for each subject, there exists 25 three-dimensional face expression models. As a result, the BU-3DFE database contains 2,500 three-dimensional face expression models.

Experiments

At a recent time, deep learning has become one of the most popular research topics and has established cutting-edge results for various applications.

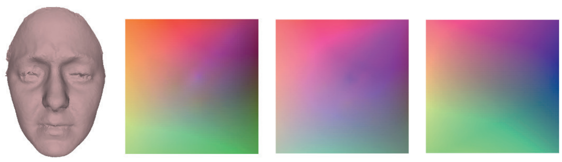

Figure 9. Different geometry images of the same face model. (L) No alignment [v1]. (M) Alignment in Cartesian coordinate system [v2]. (R) Alignment in parameter space [v3].

Deep learning aims to represent high-level abstractions by hierarchal architectures of numerous non-linear representations and transformations. We concisely describe the deep neural network model which have been used for 3D FER in our experiments. Deep networks can perform FER in an end-to-end way. They perform not only the feature extraction step, but also the feature classification step consecutively unlike the traditional methods. Thus, the deep features are extracted and learned features are classified into one of the seven expression categories. In addition to the end-to-end learning way, a different possibility is to extract features using the deep neural network and afterwards employ other independent classifiers, for instance support vector machine or random forest, to the extracted features.

Table 1. Training and test time losses and accuracies for performance evaluation of using different GIM databases.

Database	Training Time		Test Time	
	Loss	Accuracy	Loss	Accuracy
GIM v1	0.899	0.659	3.142	0.259
GIM v2	1.244	0.513	2.295	0.281
GIM v3	1.833	0.234	1.953	0.210
GIM5K	1.279	0.496	2.318	0.272

During the experiments, we employ a pre-trained VGG16 model, which is previously trained on ImageNet, and add three fully-connected layers along with a loss layer at the end of the model architecture to adjust the backpropagation error. Hence, the prediction results of each model can be immediately output by the network. We employ softmax loss which is a commonly used function to minimize the cross-entropy between the predicted class probabilities and the groundtruth labels.

CONCLUSION

In the article, we studied 3D facial expressions. The importance of experimenting in 3D domain is that 3D modeling improves the 2D drawbacks such as illumination, pose variations, etc. Therefore, having a good recognition method might be useful for human computer interaction, criminal investigation, airport security or psychological examination. A geometry based recognition method is proposed. The method was designed to spot 3D expressions from geometry images obtained using 3D face models. The method is based on opening up a 3D mesh onto a square domain using specific cut-paths. The BU-3DFE database is one of the most widely used databases in 3D domain. We use our own geometry image database which is created using 3D expression models from BU-3DFE. There exist 2500 facial models in the original database and associated 2500 geometry images in the GIM database.

The proposed method was evaluated on our GIM databases. VGG deep neural network model was used to obtain classification results. We observed that the expressions tend to give similar scores to other recognition frameworks. We keep experimenting on geometry images and it could be possible to design a more sophisticated classifier with higher recognition accuracy to improve state-of-the art results.

In this paper, we conducted experiments on three-dimensional geometry of facial expression models using end-to-end learning. One kind of image inputs namely geometry images were employed. Effects of important hyper-parameters including optimizer, learning rate, batch size and number of epochs on the classification accuracy were studied. Optimizers including Adam, SGD, RMSProp, and Nadam, learning rate ranging from 0.01 to 0.0001, four batch sizes (64, 128, 256, 512) and number of epochs ranging from 50 to 1000 were tested on the model. Our experimental results show that VGG16 model has a best classification accuracy of 28.1% on geometry image database. In conclusion, the results reveal that a pre-trained VGG16 network is capable of handling complicated information from 3D geometry pixels of expression models. Our results also produce precious intuition into the application of neural networks on 3D domain along with the facial expression recognition task.

REFERENCES

[1] Fang, T., Zhao, X., Ocegueda, O., Shah, S.K., Kakadiaris, I.A. (2012). 3D/4D facial expression analysis: An advanced annotated face model approach. *Image Vision Comput.*, 30, 738-749.

[2] Gu, X., Gortler, S.J., Hoppe, H. (2002). Geometry images. *ACM Trans. Graph.*, 21, 355-361.

[3] Yin, L., Wei, X., Sun, Y., Wang, J., Rosato, M.J. (2006). A 3D facial expression database for facial behavior research. *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, 211-216.

[4] M. Pantic and L. Rothkrantz. Automatic analysis of facial expressions: The state of the art. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1424–1445, December 2000.

[5] G. Taubin. Curve and surface smoothing without shrinkage. *In Proceedings, Fifth International Conference on Computer Vision*, pages 852–857, June 1995.

[6] A. B. Moreno and A. Sanchez. GavabDB: A 3D Face Database. *Proceedings 2nd COST Workshop on Biometrics on the Internet: Fundamentals, Advances and Applications*, Vigo, 25-26 March 2004, pp. 77-82.

[7] Phillips, P. J., Flynn, P. J., Scruggs, T., Bowyer, K. W., Worek, W. (n.d.). Preliminary Face Recognition Grand Challenge Results. *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*. https://doi.org/10.1109/fgr.2006.87

[8] Alyüz, N., Gökberk, B., Dibeklioglu, H., Savran, A., Salah, A. A., Akarun, L., & Sankur, B. (2008). 3D face recognition Benchmarks on the bosphorus database with focus on facial expressions. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5372 LNCS, 57–66. https://doi.org/10.1007/978-3-540-89991-4_7

[9] Amin, D., & Sinha, K. (n.d.). Touchy Feely?: *An Emotion Recognition Challenge*, 92.

[10] Byeon, Y. (2014). Facial Expression Recognition Using 3D Convolutional Neural Network, 5(12), 107–112.

[11] Convolutional, B., Autoencoder, C. M., & Convolutional, N. (2018). Convolutional Mesh Autoencoders for 3D Face Representation, 1–11.

[12] Fang, T., Zhao, X., Ocegueda, O., Shah, S. K., & Kakadiaris, I. A. (2012). 3D/4D facial expression analysis: An advanced annotated face model approach. *Image and Vision Computing*, 30(10), 738–749. https://doi.org/10.1016/j.imavis.2012.02.004

[13] Gu, X., Gortler, S. J., & Hoppe, H. (n.d.). Geometry Images, (Figure 1).

[14] Hasani, B., & Mahoor, M. H. (2017). Facial Expression Recognition Using Enhanced Deep 3D Convolutional Neural Networks. Retrieved from http://arxiv.org/abs/1705.07871

[15] Hussain, N., Ujir, H., Hipiny, I., & Minoi, J. (2017). 3D Facial Action Units Recognition for Emotional Expression. *Advanced Science Letters*, 24, 1–6.

[16] Huynh, X., Tran, T., & Kim, Y. (n.d.). Convolutional Neural Network Models for Facial Expression Recognition Using BU-3DFE Database. https://doi.org/10.1007/978-981-10-0557-2

[17] Jan, A., Ding, H., Meng, H., Chen, L., & Li, H. (n.d.). Accurate Facial Parts Localization and Deep Learning for 3D Facial Expression Recognition.

[18] Kim, D., & Choi, J. (n.d.). Deep 3D Face Identification.

[19] Losasso, F., Hoppe, H., Schaefer, S., & Warren, J. (2003). Smooth Geometry Images.

[20] Mayya, V., Pai, R. M., & M, M. P. M. (2016). Automatic Facial Expression Recognition Using DCNN. *Procedia - Procedia Computer Science*, 93(September),

453–461. https://doi.org/10.1016/j.procs.2016.07.233

[21] Mollahosseini, A., Hassani, B., Salvador, M. J., Abdollahi, H., Chan, D., Mahoor, M. H., & Jan, C. V. (n.d.). Facial Expression Recognition from World Wild Web.

[22] Ou, J. (2012). Classification Algorithms Research on Facial Expression Recognition. *Physics Procedia*, 25, 1241–1244. https://doi.org/10.1016/j.phpro.2012.03.227

[23] Oyedotun, O. K., Demisse, G., El, A., & Shabayek, R. (2017). Facial Expression Recognition via Joint Deep Learning of RGB-Depth Map Latent Representations, (December). https://doi.org/10.1109/ICCVW.2017.374

[24] Paine, T. Le, & Huang, T. S. (n.d.). Do Deep Neural Networks Learn Facial Action Units.

[25] Pantic, M. (n.d.). 4DFAB: A Large Scale 4D Facial Expression Database for Biometric Applications.

[26] Phillips, P. J., Flynn, P. J., Scruggs, T., Bowyer, K. W., Chang, J., Hoffman, K., … Worek, W. (2005). Overview of the Face Recognition Grand Challenge * 2 . *Design of Data Set and Challenge*, 1–8.

[27] Revina, I. M., & Emmanuel, W. R. S. (2018). A Survey on Human Face Expression Recognition Techniques. *Journal of King Saud University - Computer and Information Sciences*. https://doi.org/10.1016/j.jksuci.2018.09.002

[28] Sander, P. V, Wood, Z. J., Gortler, S. J., Snyder, J., & Hoppe, H. (2003). Multi-Chart Geometry Images.

[29] Savran, A. (2008). Non-Rigid Registration of 3D Surfaces by Deformable 2D Triangular Meshes. *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 1–6. https://doi.org/10.1109/CVPRW.2008.4563083

[30] Savran, A., Alyüz, N., Dibeklioğlu, H., & Çeliktutan, O. (n.d.). Bosphorus Database for 3D Face Analysis, 1–10.

[31] Sumathi, C. P., & Santhanam, T. (2012). Automatic Facial Expression Recognition Using 3D Faces. *International Journal of Computer Science & Engineering Survey*, 3(6), 47–59. Retrieved from https://doi.org/10.1145/1631272.1631358

[32] Wei, X., B, H. L., & Gu, X. D. (2017). Three Dimensional Face Recognition via Surface Harmonic Mapping and Deep Learning, 1, 66–76. https://doi.org/10.1007/978-3-319-69923-3

[33] Yin, L., Wang, J., & Rosato, M. J. (2006). A 3D Facial Expression Database For Facial Behavior Research.

[34] Zhang, Z., Girard, J. M., Wu, Y., Zhang, X., Liu, P., Ciftci, U., … Yin, L. (2016). Multimodal Spontaneous Emotion Corpus for Human Behavior Analysis. https://doi.org/10.1109/CVPR.2016.374

[35] Zulqarnain, S., Ajmal, G., & Engineering, S. (2018). Learning from Millions of 3D Scans for Large-scale 3D Face Recognition *(This the preprint of the paper published in CVPR 2018)*.