Automatic Selection of Key Points for 3D-Face Deformation

Huynh Cao Tuan[1], Do Nang Toan[2], and Lam Thanh Hien[1]
[1] Lac Hong University, Dong Nai, Vietnam
[2] Institute of Information Technology, VAST, Vietnam
Email: {caotuan, lthien}@lhu.edu.vn, donangtoan@gmail.com

***Abstract*—Human face has been an interesting research topic in several fields in terms of both theoretical research and deployment applications. Such real-world problems have contributed to promoting virtual reality studies, computer graphics in general, and 3D face modeling studies in particular. Several studies have shown that different control points are selected based on different applications which need different 3D face models. This study proposes an innovative approach to automatically identify the set based on facial deformations by clustering similar deformations. Its good performance is successfully validated through practical experiments with JAFFE database.**

***Index Terms*—control point set, face deformation, 3D face model**

I. INTRODUCTION

Human face has been an interesting research topic in several fields in terms of both theoretical research and deployment applications, such as detecting deception through facial expressions [1], virtual reality, computer graphics, among others. Thus, facial recognition and its techniques have well attracted the special interest of several researchers worldwide as presented in [2], [3]. For example, in fiction movies, we often make videos that transform from one character to another or create changes in expressions as well as abnormal and deformed transformations on the characters' faces. There are also cases, such as in criminal science, where people often have to create images of human faces from descriptions of others or even need to recover human faces from deformed bodies or even the remains of skulls only. Or it has been successfully used in beauty industries [4]-[6]. These are cases that require a high level of detail, but there are also types of applications that only need a sufficient level of detail; for instance, in game applications, it is important to have a certain balance between aesthetics and the performance of the program during its operations whose lag may be affected by the internet latency and computer configuration. These are just a few examples of the role of facial studies that can be easily visualized. Such real-world problems have contributed to promoting virtual reality studies, computer graphics in general, and 3D face modeling studies in particular.

One of the popular approaches used in transforming the 3D surface model is to use the control points [7]. A 3D face model is featured and controlled by a set of selected 3D points. The deformations of the 3D surface model is then parameterized according to the deformations of the 3D point set [5], [8]. Hence, to get the deformation of the 3D surface model, instead of having to control each point separately, we only need to make a few transforms and let other points to be accordingly transformed. For example, when we laugh, we only need to change a few points around the mouth, other points nearby will be shifted accordingly depending on the influence degree. This approach not only simplifies the manipulation process to perform 3D face model transformation but also closely relates to human knowledge of face deformation. Before transforming a 3D face model, we must define the control point set and parameterize the relationships between other points to this set of control points.

Practically, several studies have shown that different control points are selected based on different applications which need different 3D face models. A popular approach is based on knowledge of facial anatomy [9]. In term of biology perspective, the changes in the human face are caused by the movement of facial muscles on the base of the skull; hence, the control point set for the 3D face model can be selected based on related studies of anatomical experts so that we can identify which positions are particularly influencing the changes of the face. In some other cases, the basis for the selection may be simpler than those that clearly distinguish from others but it is also consistent across different facial deformations. For example, if we concern a specific region of a mouth, appropriate control points include corners of the mouth and some between the lips while other ones may not be of interest.

Fig. 1 are some examples of the human face model in the realm of virtual reality.

In the field of animation, a set of control points used in the construction of face animations is the set of FAPs in MPEG-4 format [9]. As shown in Fig. 2, FAPs construct face specification using 84 points along with a set of activation parameters corresponding to facial movements to create deformation of the face model. This sequence of parameters is assigned a timestamp that produces a specific facial animation. The control points of FAPs are

---

Manuscript received December 15, 2021; revised May 22, 2022.

doi: 10.12720/jait.13.4.332-337

mounted on a number of quite intuitive positions of the human face such as the position of two eyes, mouth, etc. In addition, these points are also classified by subgroups such as nose, mouth, etc.

Figure 1. Some examples of the human face model in the realm of virtual reality.

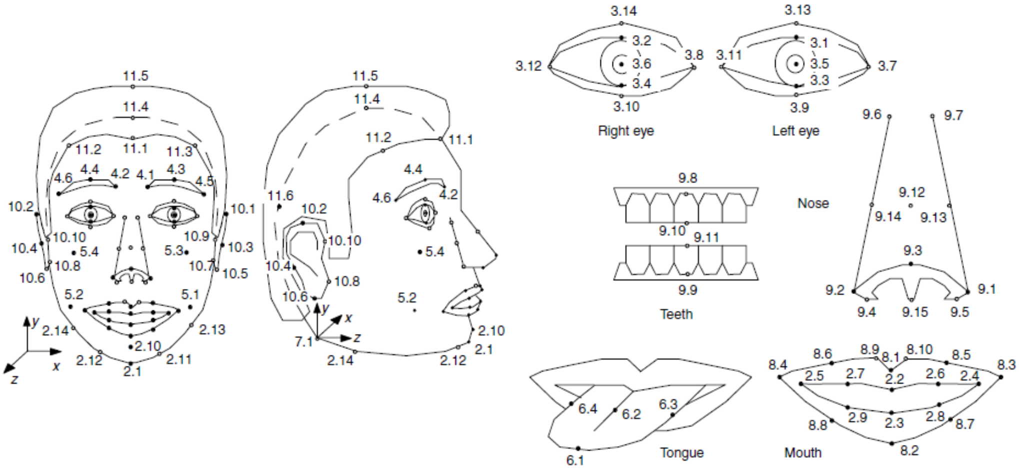


Figure 2. Control point set of FAPs.

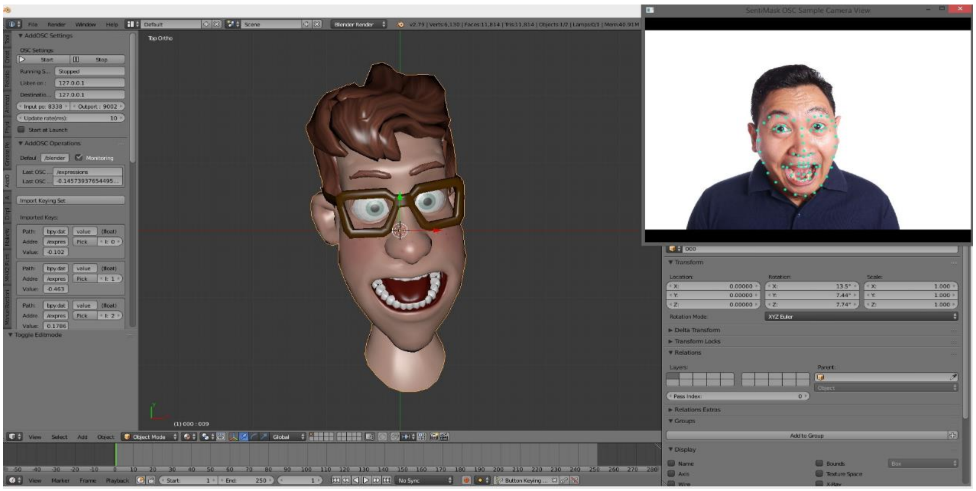


Figure 3. Control point set used in the product of NeuroTechnology.

Another control point set used in the NeuroTechnology company's product includes many features of facial image analysis, shown in Fig. 3. This set of points consists of 68 points assigned with coordinates of facial components such as eyes, eyebrows, eye contours, mouth, and nose.

Such built-in sets have achieved great success in both real-world applications and theoretical studies related to facial transformation in the realm of virtual reality. Literally, facial control point set is manually identified based on knowledge of biological face because 3D faces have many different types of distortions based on expressions, verbal gestures to changes from person to person. However, this study proposes another approach to automatically identify the set based on facial deformations by clustering similar deformations. Moreover, we also describe how to assess the quality of such deformations when employing the set of control points in a 3D model deformation technique.

II. PROPOSED APPROACH

A. Deformation Characteristics of a Point

Let consider a face model A consisting of N points expressed as $V(A)=\left\{v_0, v_1, ..., v_{N-1}\right\}$, where v_i is the i^{th} point of the face A, $v_i \in R^3$.

Suppose we have an input data set consisting of M variations of the face model A, i.e. we have a set S including M different 3D face models with the same of number and order of points as well as the edge-to-surface relation among them but the models are different in their coordinates. Thus, the set S can be described as $V(S)=\left\{v_{ij} \mid i=\overline{0, N-1}, j=\overline{0, M-1}\right\}$, where v_{ij} is the i^{th} point of the j^{th} deformation of 3D face model A, $v_{ij} \in R^3$.

At the i^{th} point, we have M presentations corresponding to M deformations. Let $T_i=\left\{v_{i0}, v_{i1}, ..., v_{i,M-1}\right\}$ be the deformation characteristics of i^{th} point in set S, i.e. T_i represents different coordinates of i^{th} point in the set S of the investigated 3D face model.

B. Identification of Clusters and Selection of Control Points

With a deformation set M, we compute deformation characteristics of each facial point and obtain a set of features; each feature respectively represents different coordinates of a point on 3D face model. We then perform clustering on this feature set to find groups of points with similar deformation in the input data set M. The algorithm is designed based on the data of the deformation characteristics of the groups and the center point of each group is selected as a control point of the model. To cluster, we need a distance function between two features:

```
Function: deformationFeatureMetric
Input: 2 deformation features $T_1$, $T_2$
Output: distance d
Process:
   1. m = 0
   2. for i in range(0, |$T_1$|):
   3.   m = m + distance($T_1$[i], $T_2$[i])
   5. m = m/|$T_1$|
Clustering algorithm:
Function: clusterDeformationFeature
Input: Set of features T, point groups k
Output: Set of control points centers, set of point groups clusters
Variables: Temporary variable to store point groups tmpClss, Temporary variable to store set of control points tmpCenters
Process:
   1. centers = random_select(T, K)
   2. resize(tmpClss, K)
   3. for l in range(0, MAX_LOOP):
   4.   for k in range(0, K):
   5.     tmpClss[k] = ∅
   6.   for i in range(0, |T|):
   7.     cid = findNearest(T, i, centers)
   8.     tmpClss[cid] = tmpClss[k] ∪ {i}
   9.   tmpCenters = calcCenters(T, tmpClss)
   10.  eps = 0
   11.  for k in range(0, K):
   12.    eps += deformationFeatureMetric(centers[k], tmpCenters[k])
   13.  eps = eps/K
   14.  if eps<MIN_EPS:
   15.    break
   16.  centers = tmpCenters
   17.  cluster = tmpCluster
```

Each iteration of the loop specified in line 3 recalculates the centers of the point groups. To ensure that the center is a point of the face model, from the calculated center of each group, one point is selected as the center point provided that it has the closest distance identified by *deformationFeatureMetric* to the point groups. Moreover, the *clusterDeformationFeature* algorithm results in clusters of similar deformations. We then choose clusters with the most and least deformed points to put into control point set. The deformed degree is evaluated based on a reference model. Thus, each point of the reference model corresponds to a deformation feature. A deviation vector is determined by the distance of each point in the deformation feature to its corresponding point on the reference model. The evaluation value of deformed degree of each corresponding point is calculated based on the deviation vector.

C. Analysis of Efficiency in Transforming Model

The proposed algorithm aims at building control points to serve the transformation of the 3D face model. Thus, a good control point set is considered good once it is effectively used to transform 3D surface model. In this subsection, we propose an algorithm to analyze the efficiency of the results obtained in the previous step. This algorithm is deployed in a 3D model transformation technique based on control point set. The algorithm is described as follows:

```
Function: errorAnalysis
Input: Reference model R, Deformation feature set T,
Evaluation model set M
A set of K corresponding values KC
Output: Control point set P
Variables:
        Cluster temporary variables clusters
        Center temporary variables centers
Process:
   1. minErr = MAX_VALUE
   2. for K in KC:
   3.   clusterDeformationFeature(T, K, centers, clusters)
   4.   tmpP = selectControlPoints($T_1$, clusters, R)
   5.   err = 0
   6.   for i in range(0, |M|):
   7.     err += errMorph(M[i], R, tmpP)
   8.   err = err/|M|
   9.   if err < miner:
   10.    minerr = err
   11.    P = tmpP
```

In line 7, *errMorph* function evaluates the error between the evaluation model *M[i]* and the face model transformed from the reference model R based on the given set of control points.

III. EXPERIMENTAL RESULTS

This study conducts an experiment with the following inputs: a 3D model transformation technique based on the control point, one reference 3D object, one set of 3D varied objects to calculate control points and a set of 3D varied objects to evaluate the quality of the selected control point set. Each experiment works with a different K value to evaluate its errors. The rectification technique used in the experiment is the 3D object manipulation technique using the Radial Basic Function (RBF) [10].

The 3D object used in this study includes a face model with 3448 vertices and 6736 surfaces created corresponding to patterns in the JAFFE database which contains 213 images of 7 different facial expressions (6 basic facial expressions including happy, sad, surprised, angry, angry, scared, and one neutral status) captured on 10 Japanese female models, as shown in Fig. 4. The database was built by Ref. [11]. The images in the JAFFE database were evaluated with the proportions of the six emotional states. These assessments were made by the perception of 60 Japanese students with the highest value of 5 and the lowest value of 1. Based on the selected data set, our experiments were conducted under three different transformation scenarios, including among different people with different expressions, among different people with the same neutral state and of a person with different expressions.

Figure 4. Example of 3D face model.

In the case of the variation among different people with different expressions, two different sets of variations for the experiment were obtained by randomly dividing the sample data set into two. With the variation set for selecting control point set, the deformation feature set of the points is calculated as shown in Fig. 5. As shown in Fig. 6, obtained feature sets are used in clustering with different K values to determine each region with similar deformation.

The obtained cluster of each K results in a set of control points which is then used to transform the reference 3D model according to each sample in the variation set for output quality assessment. Each model transformed with RBF technique is compared with the reference model to determine errors. Since the data nature in the 3D model and the deformation features of the point are both sets of 3D points, the error between the transformed model and the reference model is also calculated through the distance function between two variation features. The error value of the control point set on the variation set for quality assessment is calculated by the average error of each sample investigated. Fig. 7 shows the relationship between K values and obtained errors.

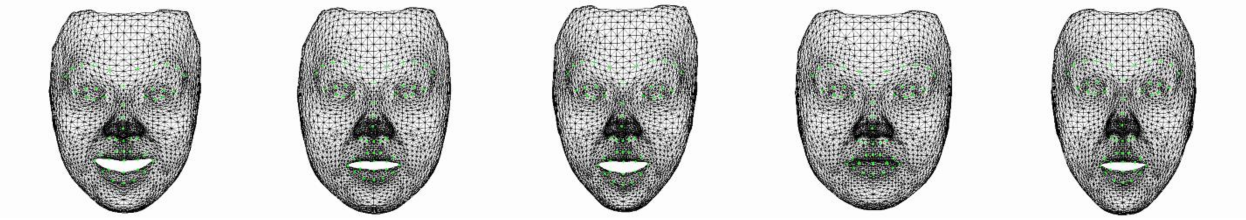


Figure 5. Several sample variations of a 3D object.

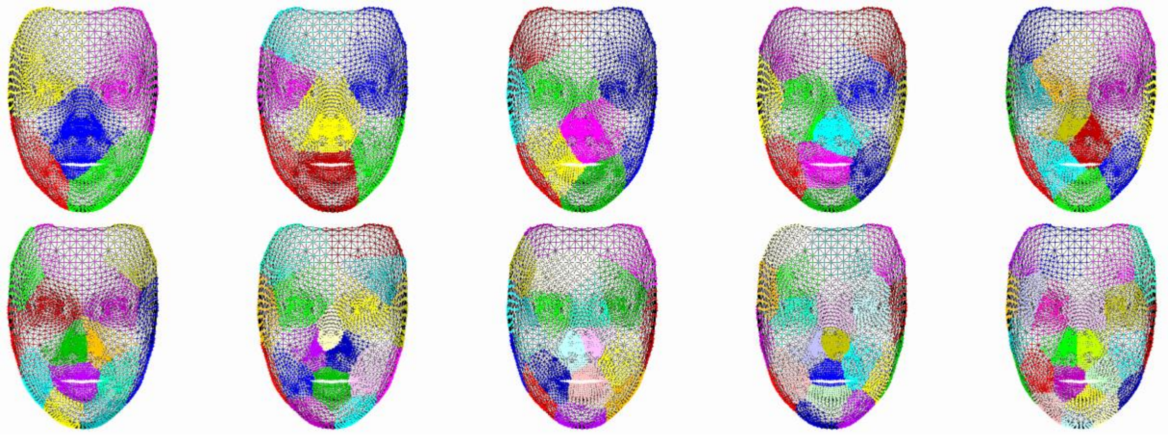


Figure 6. Some typical clustering results with different K levels.

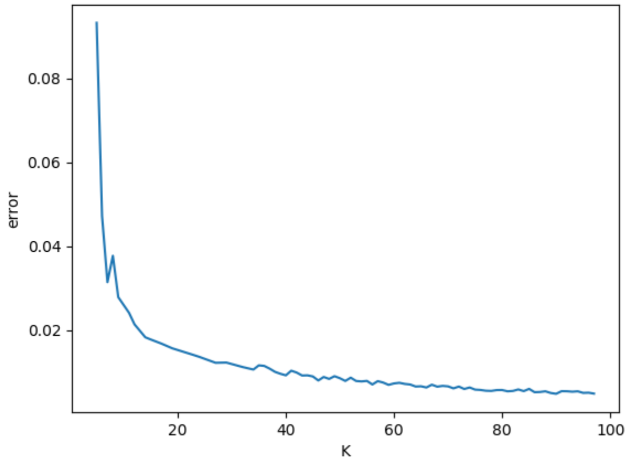
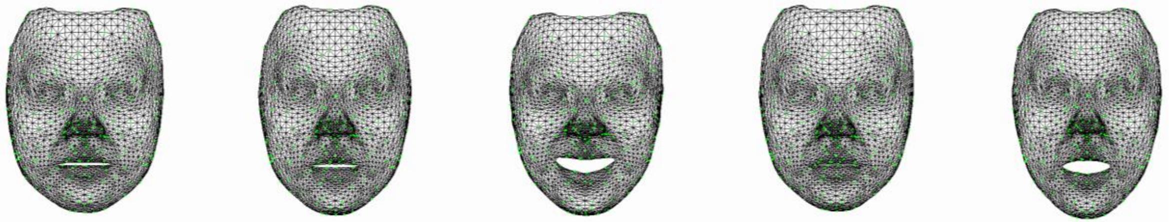


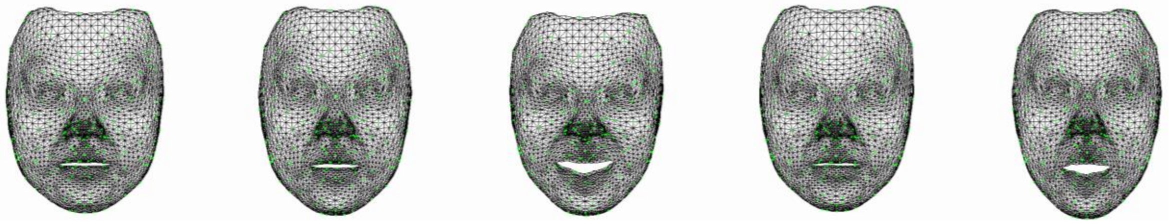
Figure 7. Relationship between obtained errors and K

Particularly, larger K values, i.e., increase in the size of the control point set results in better accuracy of the transformation. Hence, in our experiments, we choose a reasonable size set of control points to meet allowed error range. Fig. 8 shows our typical experimental results.

For the other two cases, namely the transformation among different people on the same neutral expression and the transformation of a person with different expressions, the computation steps are performed similarly but with a small difference in selecting two sets of variations for calculating control points and assessing their quality. In the case of judging among different people on the same neutral expression, relevant models corresponding to the neutral expression are selected. Two sets of variations are generated by linear aggregation among several models in the randomly selected set. The analysis in this case aims to transform the model from person to person. Similar to the transformation of a person with different expressions, for a particular person, the chosen set is the model corresponding to that person to analyze the transformation of his/her expression states. The obtained results in these two cases are respectively shown in Fig. 9 and Fig. 10.



Reference 3D faces



Transformed 3D faces

Figure 8. Transformation results.

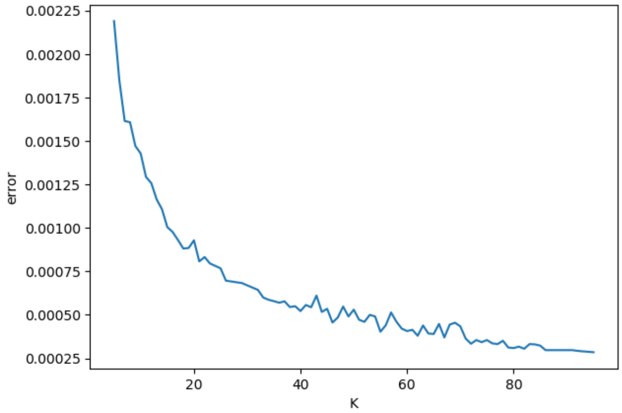


Figure 9. Relationship between errors and K of different people with the neutral expression.

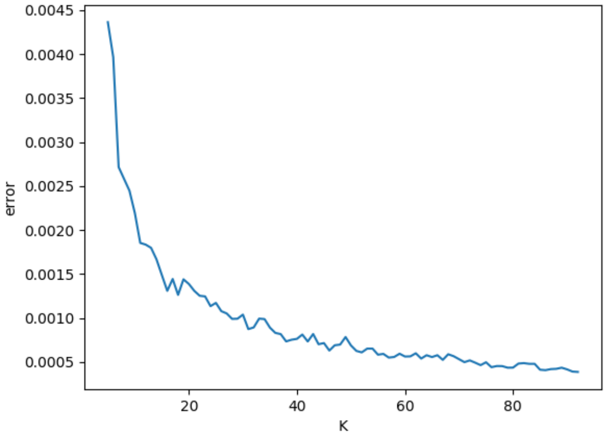


Figure 10. Relationship between obtained errors and K of a particular Model KA.

The relationship of obtained errors and K in the three scenarios shown in Fig. 8 – Fig. 10 indicates that when the increase in K value gradually lessens the transformation errors, which is especially true with our proposed approach in automatically selecting control point set and using RBF technique. However, it is also shown that the errors in the case of transformation among different people on the same neutral expression is much smaller than in other two cases, which also reflects the nature of the RBF interpolation as a global transformation which is more suitable than other ones when we consider the transformation of different expressions of a human face because the observed changes are usually localized on certain facial components such as the mouth and eyes only.

IV. CONCLUSION

Due to the various applications of 3D human face modelling in different fields, fully capturing of facial deformation plays critical role. Thus, it is mandatory to effectively identify key points for controlling the deformation. However, the traditional approaches select the control points manually based on prior knowledge. To overcome the shortcoming, this study proposes an innovative algorithm to automatically identify key points and cluster similar deformations with radial basic function (RBF) technique to improve the deformation of 3D faces. From our experiments with images in JAFFE database, it is found that our proposed approach results in a better performance in terms of lower errors in the transformation and deformation.

It is suggested to further investigate our proposed approach in dealing with other databases to firmly validate its performance in theoretical and practical applications.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

H. C. Tuan conducted the research; D. N. Toan and L. T. Hien analyzed the data. H. C. Tuan and L. T. Hien wrote the paper. All authors had approved the final version.

ACKNOWLEDGMENT

The authors wish to thank Lac Hong University for the financial support.

REFERENCES

[1] M. Monaro, S. Maldera, C. Scarpazza, G. Sartori, and N. Navarin, "Detecting deception through facial expressions in a dataset of videotaped interviews: A comparison between human judges and machine learning models," *Computers in Human Behavior*, vol. 127, p. 107063, 2021.

[2] J. A. C. Moreano and N. B. L. S. Palomino, "Global facial recognition using gabor wavelet, support vector machines and 3D face models," *J. Advances in Information Technology*, vol. 11, no. 3, pp. 143-148, August 2020.

[3] F. Z. Canal, *et al.*, "A survey on facial emotion recognition techniques: A state-of-the-art literature review," *Information Sciences*, vol. 582, pp. 593-617, 2022.

[4] Q. Xiao, Y. Wu, D. Wang, Y. L. Yang, and X. Jin, Beauty3DFaceNet: Deep geometry and texture fusion for 3D facial attractiveness prediction," *Computers & Graphics*, vol. 98, pp. 11-18, 2021.

[5] E. R. Manal, Z. Arsalane, M. Aicha, and O. A. Allah, "3D facial attractiveness enhancement using free form deformation," *Journal of King Saud University - Computer and Information Sciences*, 2020.

[6] A. Laurentini and A. Bottino, "Computer analysis of face beauty: A survey," *Comput. Vis. Image Underst.*, vol. 125, pp. 184-199, 2014.

[7] V. Blanz and T. Vetter, "Face recognition based on fitting a 3d morphable model," *IEEE Trans. Pattern Analysis Machine Intel.* vol. 25, no. 9, pp. 1063-1074, 2003.

[8] J. Zwirner, B. Ondruschka, M. Scholze, and N. Hammer, "Passive load-deformation properties of human temporal muscle," *Journal of Biomechanics*, vol. 106, p. 109829, 2020.

[9] I. S. Pandzic and R. Forchheimer, *MPEG-4 Facial Animation: The Standard, Implementation and Applications*, John Wiley & Sons, 2003.

[10] M. D. Buhmann, *Radial Basis Functions: Theory and Implementations*, Cambridge: Cambridge Univ. Pr., 2003.

[11] M. J. Lyons, S. Akemastu, M. Kamachi, and J. Gyoba, "Coding facial expressions with gabor wavelets," in *Proc. 3rd IEEE International Conference on Automatic Face and Gesture Recognition*, 1998, pp. 200-205.

Huynh Cao Tuan obtained his Master degree in Information Technology 2012. He is now working the Director of the Center of Information and Resources of Lac Hong University. His research interest includes: robotic technology, data mining, artificial intelligence, machine learning, image processing, etc.

Do Nang Toan is an Associate professor in Computer Science. He received BSc. Degree in Applied Mathematics and Informatics in 1990 from Hanoi University and PhD in Computer Science in 2001 from Vietnam Academy of Science and Technology. He is currently working as Associate Professor in Computer Science at Institute of Information Technology, VAST, Vietnam. His main research interests are Pattern recognition, Image processing and Virtual reality.

Lam Thanh Hien received his MSc. Degree in Applied Informatics Technology in 2004 from INNOTECH Institute, France; and PhD Degree in Computer Science in 2016 from Vietnam Academy of Science and Technology. He is currently working as the Rector of Lac Hong University. His main research interests are Information system and image processing.