International Journal of Electrical and Computer Engineering (IJECE)
Vol. 14, No. 2, April 2024, pp. 1928~1936
ISSN: 2088-8708, DOI: 10.11591/ijece.v14i2.pp1928-1936

An algorithm for decomposing variations of 3D model

Tran Thanh Phuong[1], Lam Thanh Hien[1], Ngo Duc Vinh[2], Ha Manh Toan[3], Do Nang Toan[3]
[1]Faculty of Information Technology, Lac Hong University, Bien Hoa, Vietnam
[2]Faculty of Information Technology, Hanoi University of Industry, Hanoi, Vietnam
[3]Institute of Information Technology, Vietnam Academy of Science and Technology, Hanoi, Vietnam

Article Info

Article history:

Received Jul 14, 2023
Revised Oct 27, 2023
Accepted Nov 29, 2023

Keywords:

3D deformation
3D face models
Conditional decomposition
Linear synthesis
Virtual reality

ABSTRACT

In recent times, there has been an increasing number of people who are concerned about the virtual reality field. Parameterization of deformations of 3D models is a meaningful problem in theoretical research and application development of virtual reality. This paper proposes a technique for conditional decomposition of 3D model variations based on a given set of 3D observations of an object, along with a set of input strain weights. The proposed algorithm is conducted through an optimal iterative process with solving the non-negative least squares problem. The output of the technique is a set of base models corresponding to different types of strain. The result of the proposed technique allows the creation of a new 3D model variant of the object in a simple and visually observable way. The algorithm has been tested and proven effective on data that are 3D face models created from the Japanese Female Facial Expression (JAFFE) dataset with labeled expression weights.

Corresponding Author:

Do Nang Toan
Institute of Information Technology, Vietnam Academy of Science and Technology
18 Hoang Quoc Viet, Cau Giay District, Hanoi 10072, Vietnam
Email: dntoan@ioit.ac.vn

1. INTRODUCTION

The 3D model deformation is an important problem in the research and development of applications in the field of virtual reality. Since objects in the real world always move and change shape, objects in the virtual world also need to exhibit corresponding distortions. We have seen that water surface objects, grass, or rice fields always have the rules of motion under the influence of wind. We also see examples of the rapid growth of a tree, a leaf from sprouting, expanding, and extending. Even the distortions we can see every day like the stretching of instant noodles when soaked in hot water contain mathematical models of itself as shown in Figure 1.

In essence, the 3D model can be understood as consisting of a set of points in three-dimensional space with the relations of edges as well as surfaces between those points. In addition, depending on the specific research and application case, the 3D model can add different features. Given a shape is represented with a meshed, the 3D model deformation problem is defined as finding the best position of the vertices so that the new shape fits the target while preserving the local geometric details [1]. The problem of 3D model deformation has been very interesting in recent times. Many techniques have been proposed, mainly divided into 3 groups:

a. The first group is called the free-form approach [1]–[4] the methods in this group use voxel grids of the volume enclosing the surface as the control point and define a smooth, strain function that interpolates the weights from the points control to the vertices of the mesh.
b. The second group is the nest-based approaches [5]–[8] which, unlike the first, derive control points not from the voxel grid but from a rough scaffold grid surrounding the input.

c. The last group is vertex-based approaches [9]–[13]. In this approach, the mesh vertex locations describe properties that need to be preserved, such as Laplacian grids [10], [13] or local rigidity [9], [11], [12], so that the objective function for optimization is directly defined by the mesh vertex locations.

Accordingly, the 3D model deformation is essentially the rule to change that information on a model. In the majority of research and application cases, the 3D model deformation is usually considered in terms of the relative variability of the points together in the 3D model. For example, when we consider the transformation of a 3D face model from a neutral face model to a smiley face model.

It is clear that we always imagine that the number of points as well as the relation of edges and surfaces between the points is fixed, just that there is a change in the coordinates of the points together, such as when the mouth smiles, the points at the corners of the mouth relax and move up. Or consider the deformation of the leaf when swayed with the wind or bent, the same changes as shown in Figure 2. When studying deformation on a certain type of model, it is often attempted to parameterize these warping according to certain rules. Such rules allow the authors to manipulate the distortions in order to achieve the desired goals.

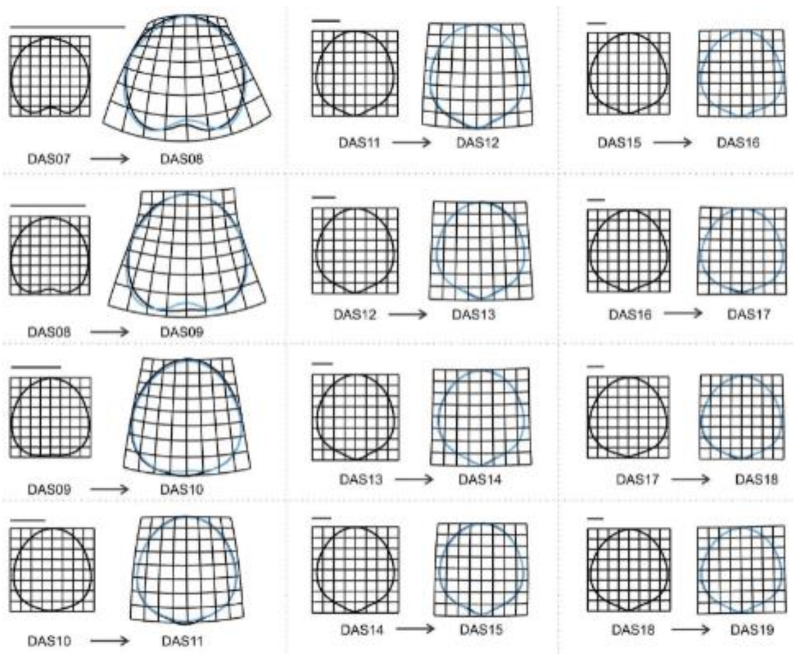

Figure 1. Deformed samples of leaf tissue averaged a day in the study [14]

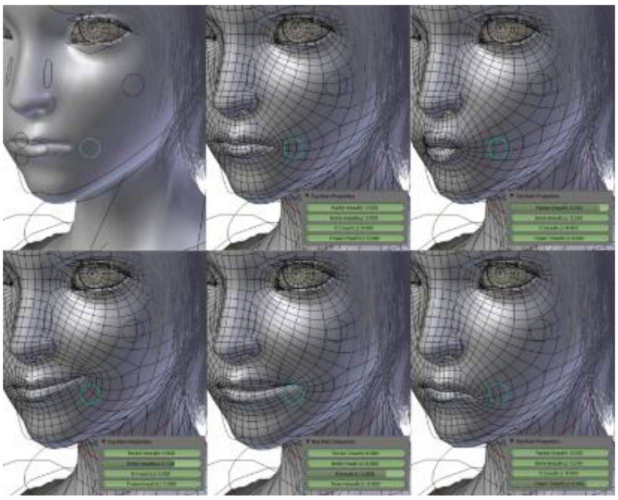

Figure 2. Deformed the face model of the protagonist in the open source Sintel project [15]

Blanz *et al.* [16] studied a statistical model of the face constructed from 3D laser scanning data on a Cyberware 3030PS machine. Group results allow us to perform model transformations to best fit the face on an input image. Thus, after that process, the model parameters will respectively represent that face. This research allows for in-depth analysis of facial data on 3D models such as emotional analysis as well as facial recognition as shown in Figure 3.

Figure 3. Some examples of model matching with facial images in group Blanz's study [16]

Wang *et al.* [17] proposed a technique to generate a variation of a 3D model. According to a target model (represented as a 2D or 3D image), they deform the source grid so that it resembles the target model as closely as possible. In this strain model, they keep the triangular topology of the source grid fixed and update only the vertex positions as shown in Figure 4.

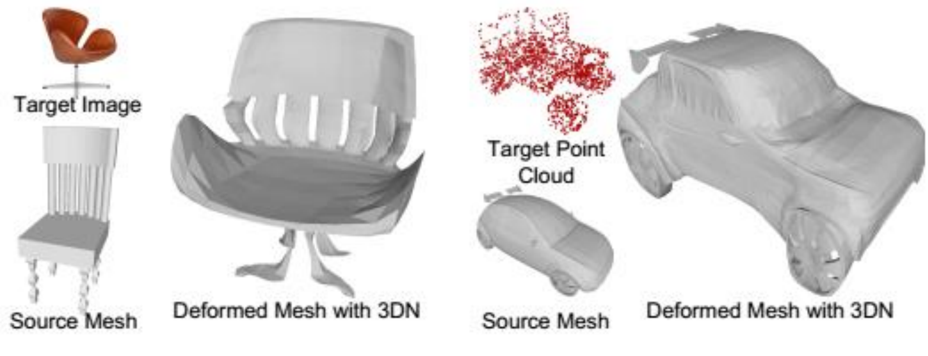

Figure 4. Deformed mesh with 3DN [17]

Groueix *et al.* [18] propose shape deformation networks, a comprehensive, all-in-one solution to template-driven shape matching. A shape deformation network learns to deform a template shape to align with an input observed shape. Given two input shapes, they align the template to both inputs and obtain the final map between the inputs by reading off the correspondences from the template. Kurenkov *et al.* [19] extended the spatial transform network (STN) model to DeformNet using the free-form deformation (FFD) module, whereby DeformNet takes image input, searches for the closest shape pattern from the database and deforms the sample to fit the query image. Pontes *et al.* [20] and Jack *et al.* [21] introduce methods to learn FFD. Yang *et al.* [22] propose FoldingNet which deforms a 2D grid into a 3D point cloud while preserving locality information. Tuan *et al.* [23] proposed an innovative algorithm to automatically identify key points and cluster similar deformations with the radial basic function (RBF) technique to improve the deformation of 3D faces. Another study is concerned with the problem of synthesizing a face model from a set of basic surface models [24]. In this study, to perform interpolation to represent gestures and expressive states of human faces, the research team used a set of pattern models, which included a neutral face model and the model shows the face state according to specific expressions. The resulting model is generated by linear synthesis of those underlying models as shown in Figure 5.

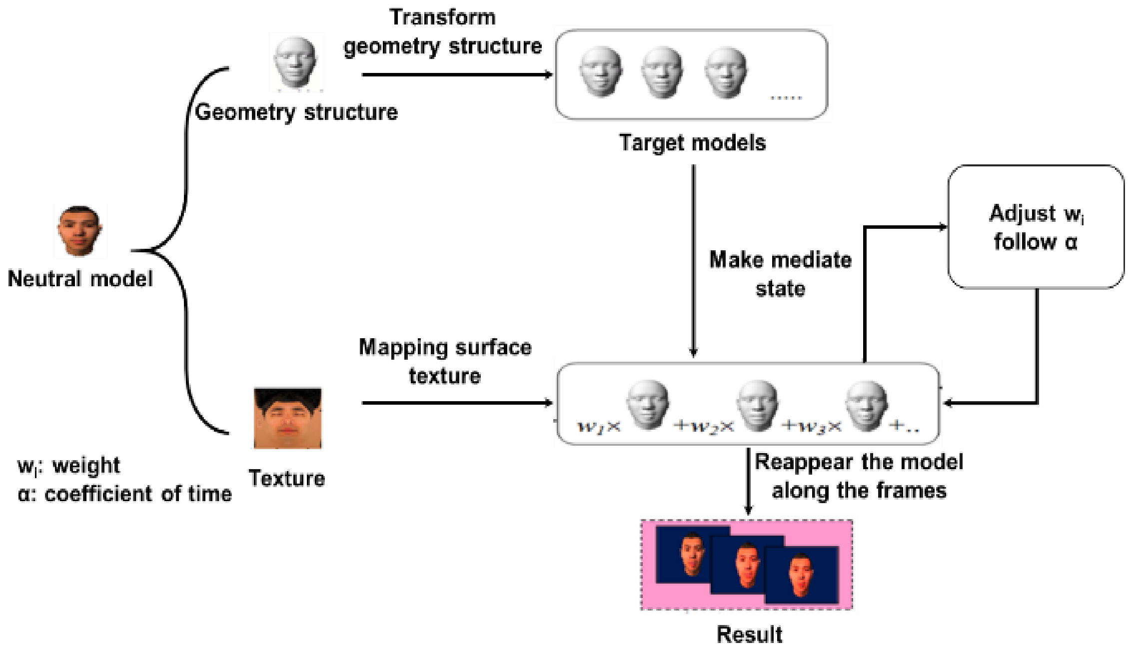

Figure 5. The process of synthesizing of the face model [24]

The idea of studying the model deformation of the paper is posed in the opposite way of the study [24]. Accordingly, the paper proposes a technique for conditional decomposition of 3D model variations based on a given set of 3D observations of an object, according to a set of input strain weights. The results of the proposed technique allow the creation of a new 3D model variant, of the object in a simple and visually observable way.

Specifically, the proposed technique will analyze a 3D model data set weighted labels for each type of deformation given. The output of the technique is a set of basic models corresponding to different types of strain. In the next content, section 2 of the paper will detail the proposed algorithm, section 3 will be the experimental results, and section 4 contains conclusion.

2. METHOD

2.1. 3D model synthesis

Variations of the 3D model are assumed to be represented in (1):

$$s = m + \sum_{i=0}^{n} \alpha_i * e_i \tag{1}$$

where s is the model variant to be aggregated; m is a model pattern, can be understood as a model of balance between distortions; e_i is the offset model corresponding to the deformation i; α_i is the weight corresponding to deformation i; and n is the number of types of deformation of the model.

The equation (1) is assumed to represent any variation of the model. Thus, to synthesize a model, we need to have data on the right side of the formula (1). In which the set of basic models including m and e_i are predetermined data, the α_i are the deformation weights.

2.2. The problem non-negative least squares

In this problem, the goal is to estimate a set of basic models, that is to compute m and e_i. We need the input data set of variants s and the weight values α_i corresponding to each variant. The input data set is modeled by the above formula as a matrix as (2):

$$S = M + AE \tag{2}$$

where S is a set of input samples, each of which is considered a vector 1×L, set S consists of N samples, S is a matrix $N \times L$; M is a matrix $N \times L$, in this model m is replicated N times and is lined up on each line; A is a matrix $N \times n$ deformation weights; and E is a matrix $n \times L$ are the compensation models. If M is assumption, yes, the formula (2) will return:

$$S - M = AE \tag{3}$$

Because, the deformation weight values α_i are non-negative values, in order to find E, the solution of (3) is converted to non-negative least squares [25].

2.3. Deformation decay algorithm

The deformation decay algorithm is presented with the input S, A, and the output m, E. The estimation of the set of the base model is done through an optimal iterative process as presented in Figure 6. Thus, the estimation of the set of the base model is done through an optimal iterative process as shown in Figure 6. At the initial step, the sample model m is initialized by averaging the variant models in the input data set S, and the complement models are initialized by 0. At each iteration, the recalculation is performed with two values in turn.

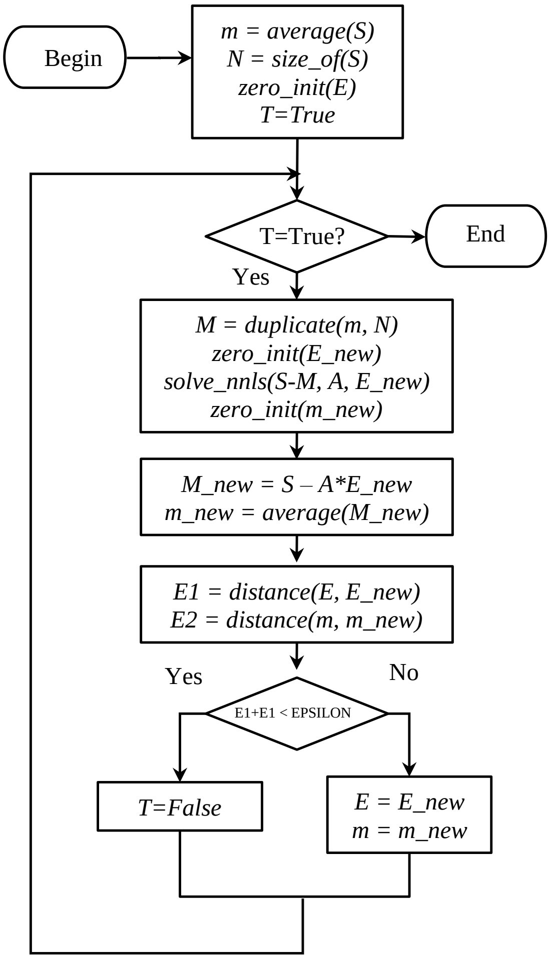

Figure 6. Diagram deformation decay algorithm

First, the model *m* models are duplicated to obtain a matrix *M* and we solve using the non-negative least squares method to get the new set of complementary models *E_new*. Then, the new complement model set *E_new* is used to recalculate the new model *m_new* by calculating the matrix *M_new* and averaging it to obtain *m_new*. After calculating the value of *E_new* and *m_new* at each iteration step, we perform the calculation of the deviation of each update step on 2 values, namely the deviation error E1 of the compensation model and the deviation error E2 of the sample model. If the sum of the deviations is small enough, that is, the iteration process no longer updates the complement model and the pattern model, the algorithm ends.

3. RESULTS AND DISCUSSION

In the experimental program, the non-negative least squares method is taken from the Eigen library, which is specialized in calculating linear algebra. The technique is tested on the Japanese Female Facial Expression (JAFFE) data set in which the 3D object is a face model consisting of 6,736 surfaces and 3,448 vertices generated corresponding to each sample in the database. JAFFE database contains 213 images, each corresponding to 7 different emotional expressions (6 basic emotions, including happy, surprised, angry, scared, sad, angry, plus neutral status) recruited from 10 Japanese female models. This database was built by Michael Lyons, Miyuki Kamachi, and Jiro Gyoba with research assistant Reiko Kubota. The photos were taken at the Department of Psychology at Kyushu University. The images in the JAFFE database were evaluated with the proportions of the six emotional states. These assessments were made by the sentiments of 60 Japanese students with the highest being 5 and the lowest being 1. Figure 7 shows a 3D face model from an example image of the JAFFE database.

Therefore, the algorithm is applied to decompose a 3D face model according to the emotional parameters. In this case, each sample offset corresponds to a specific expression in the input data set. To reduce the influence of another factor, testing was conducted on a particular person's data. Thus, in the test, the input is a set of 3D models of faces with the weights corresponding to the levels of each expression type and the output is the human face model and the sample compensation for each particular expression.

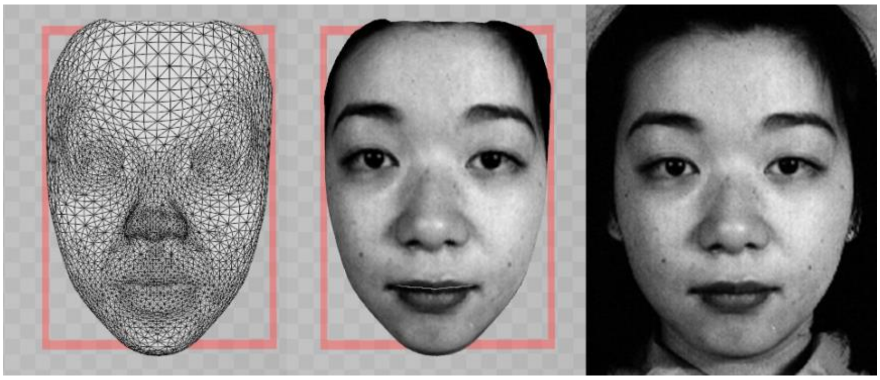

Figure 7. Example of 3D face model data

During the experiment, both the sample model and the complement model converge and are shown with deviation errors asymptotic to zero after more than 200 loops as shown in Figure 8. The results are then applied to randomly synthesize the face model as shown in Figure 9. At the synthesis step, each value in the expressive weight set is also taken at random in the segment [0,1].

The experimental results show the ability to generate new variant instances of a 3D object, particularly a 3D face in this case. In recent times, there has been a trend to apply a deep learning approach to synthesize a 3D object by modeling the 3D deformation. For example, in 2019, Wang *et al.* [17] presented a 3D deformation network that can deform the base model to be similar to the target as a point cloud acquired as a depth scan, 3D model, or 2D image. Another example is the study of Jack *et al.* [21]. In their work, learning free-form deformations was conducted to reconstruct a 3D object to generate dense point clouds or meshes. In comparison to the high-complexity computation of deep learning networks, our study focuses on a simpler method to generate 3D objects. Our technique outputs a set of base models corresponding to different types of strain, that allows the creation of a new 3D model variant of the object in a low-complexity computation and visually observable way.

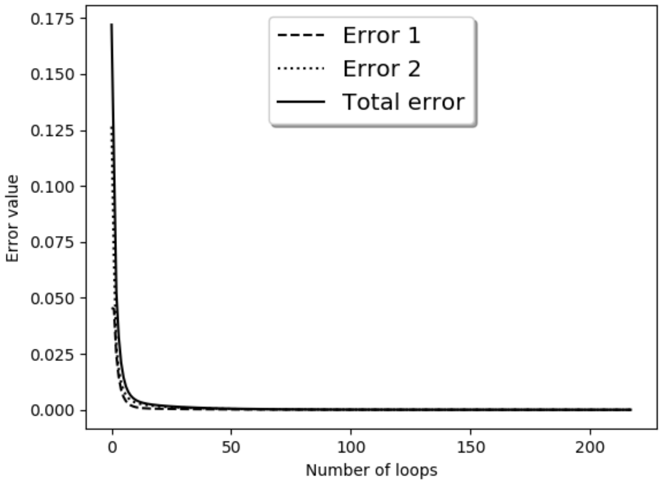

Figure 8. The process of updating deviation error

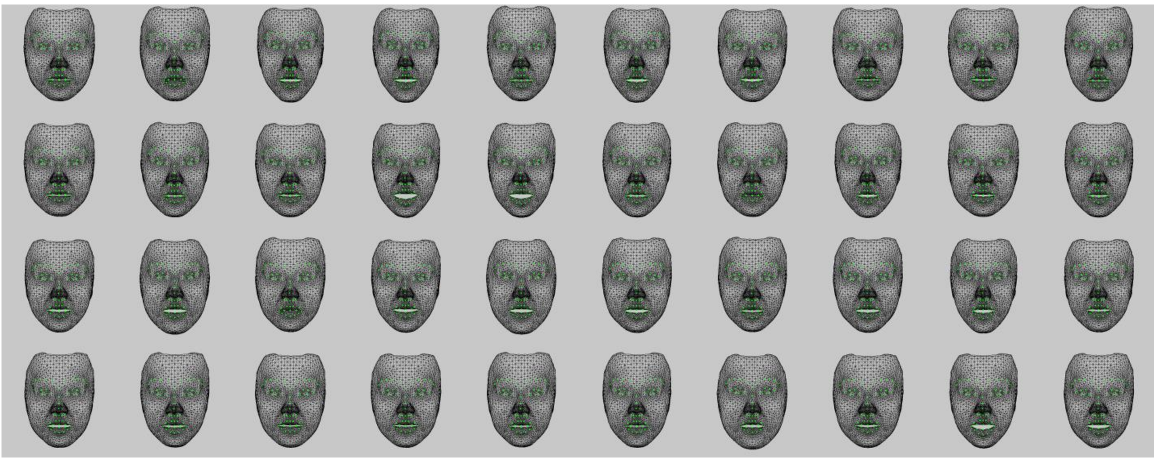

Figure 9. Some faces are randomly synthesized

4. CONCLUSION

This paper mentions the virtual reality field from the perspective of the deformation problem. Research on 3D model deformation is an important issue in both theory and practical implementation. The paper presented a technique of conditional decomposition of 3D model variations based on a given set of 3D observations of an object. The results of the paper allow to creation of a new 3D model variant, of the object in a simple and visually observable way that is linear synthesis. So, results of this study toward some perspective of the 3D modeling problem, specifically 3D face models in the experiment here. The technique was tested on 3D facial models generated from the JAFFE dataset with labeled expressive weights. The convergence algorithm and the obtained results have also proven effective when applied to perform facial reconstruction. This study is the basis for our research on the 3D synthesis problem next time, especially the 3D face modeling for building the virtual newsreader or the virtual vocal music teacher.

ACKNOWLEDGEMENTS

This research is funded by the project CS23.04 "Research on some techniques for processing spatial information of objects in medical images".

REFERENCES

[1] V. Kraevoy, A. Sheffer, A. Shamir, and D. Cohen-Or, "Non-homogeneous resizing of complex models," *ACM Transactions on Graphics*, vol. 27, no. 5, pp. 1–9, Dec. 2008, doi: 10.1145/1409060.1409064.

[2] T. W. Sederberg and S. R. Parry, "Free-form deformation of solid geometric models," in *Proceedings of the 13th Annual conference on Computer graphics and interactive techniques*, Aug. 1986, pp. 151–160, doi: 10.1145/15922.15903.
[3] M. E. Yumer and N. J. Mitra, "Learning semantic deformation flows with 3D convolutional networks," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9910, Springer International Publishing, 2016, pp. 294–311.
[4] H. C. Tuan, D. N. Toan, L. T. Hien, and T.-L. Nguyen, "An innovative approach to automatically identify control point set for model deformation rectification," *International Journal of ADVANCED AND APPLIED SCIENCES*, vol. 6, no. 8, pp. 45–52, Aug. 2019, doi: 10.21833/ijaas.2019.08.007.
[5] P. Joshi, M. Meyer, T. Derose, B. Green, and T. Sanocki, "Harmonic coordinates for character articulation," *ACM Transactions on Graphics*, vol. 26, no. 3, Jul. 2007, doi: 10.1145/1276377.1276466.
[6] T. Ju, S. Schaefer, and J. Warren, "Mean value coordinates for closed triangular meshes," in *ACM SIGGRAPH 2005 Papers*, Jul. 2005, pp. 561–566, doi: 10.1145/1186822.1073229.
[7] Y. Lipman, D. Levin, and D. Cohen-Or, "Green coordinates," *ACM Transactions on Graphics*, vol. 27, no. 3, pp. 1–10, Aug. 2008, doi: 10.1145/1360612.1360677.
[8] O. Weber, M. Ben-Chen, and C. Gotsman, "Complex barycentric coordinates with applications to planar shape deformation," *Computer Graphics Forum*, vol. 28, no. 2, pp. 587–597, Apr. 2009, doi: 10.1111/j.1467-8659.2009.01399.x.
[9] T. Igarashi, T. Moscovich, and J. F. Hughes, "As-rigid-as-possible shape manipulation," *ACM Transactions on Graphics*, vol. 24, no. 3, pp. 1134–1141, Jul. 2005, doi: 10.1145/1073204.1073323.
[10] Y. Lipman, O. Sorkine, D. Cohen-Or, D. Levin, C. Rossl, and H.-P. Seidel, "Differential coordinates for interactive mesh editing," in *Proceedings Shape Modeling Applications*, 2004, pp. 181–190, doi: 10.1109/SMI.2004.1314505.
[11] Y. Lipman, O. Sorkine, D. Levin, and D. Cohen-Or, "Linear rotation-invariant coordinates for meshes," *ACM Transactions on Graphics*, vol. 24, no. 3, pp. 479–487, Jul. 2005, doi: 10.1145/1073204.1073217.
[12] O. Sorkine and M. Alexa, "As-rigid-as-possible surface modeling," *Eurographics Symposium on Geometry Processing (SGP)*, pp. 109–116, 2007.
[13] O. Sorkine, D. Cohen-Or, Y. Lipman, M. Alexa, C. Rössl, and H. P. Seidel, "Laplacian surface editing," in *ACM International Conference Proceeding Series*, Jul. 2004, vol. 71, pp. 175–184, doi: 10.1145/1057432.1057456.
[14] A. G. Rolland-Lagan, L. Remmler, and C. Girard-Bock, "Quantifying shape changes and tissue deformation in leaf development," *Plant Physiology*, vol. 165, no. 2, pp. 496–505, Apr. 2014, doi: 10.1104/pp.113.231258.
[15] Sintel, "Sintel 'lite' – character file for testing/learning," *SINTEL*, 2011. https://durian.blender.org/news/sintel-lite/ (accessed Oct. 24, 2023.
[16] V. Blanz, S. Romdhani, and T. Vetter, "Face identification across different poses and illuminations with a 3D morphable model," in *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition*, 2002, pp. 202–207, doi: 10.1109/AFGR.2002.1004155.
[17] W. Wang, D. Ceylan, R. Mech, and U. Neumann, "3DN: 3D deformation network," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2019, pp. 1038–1046, doi: 10.1109/CVPR.2019.00113.
[18] T. Groueix, M. Fisher, V. G. Kim, B. C. Russell, and M. Aubry, "3D-CODED: 3D correspondences by deep deformation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11206, Springer International Publishing, 2018, pp. 235–251.
[19] A. Kurenkov *et al.*, "DeformNet: free-form deformation network for 3D shape reconstruction from a single image," in *Proceedings - 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018*, Mar. 2018, pp. 858–866, doi: 10.1109/WACV.2018.00099.
[20] J. K. Pontes, C. Kong, S. Sridharan, S. Lucey, A. Eriksson, and C. Fookes, "Image2Mesh: a learning framework for single image 3D reconstruction," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11361, Springer International Publishing, 2019, pp. 365–381.
[21] D. Jack *et al.*, "Learning free-form deformations for 3D object reconstruction," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11362, Springer International Publishing, 2019, pp. 317–333.
[22] Y. Yang, C. Feng, Y. Shen, and D. Tian, "FoldingNet: point cloud auto-encoder via deep grid deformation," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun. 2018, pp. 206–215, doi: 10.1109/CVPR.2018.00029.
[23] H. C. Tuan, D. N. Toan, and L. T. Hien, "Automatic selection of key points for 3D-face deformation," *Journal of Advances in Information Technology*, vol. 13, no. 4, pp. 332–337, 2022, doi: 10.12720/jait.13.4.332-337.
[24] T. T. Phuong, N. D. Vinh, H. M. Toan, D. N. Toan, and N. M. Ngoc, "A technique to simulate the movement of a 3D human head model according to Vietnamese dialogue," *TNU Journal of Science and Technology*, vol. 228, no. 15, pp. 20–28, 2023, doi: 10.34238/tnu-jst.8297.
[25] C. L. Lawson and R. J. Hanson, "Solving least squares problems," *Solving Least Squares Problems*, 1995, doi: 10.1137/1.9781611971217.

BIOGRAPHIES OF AUTHORS

Tran Thanh Phuong received his master of IT in 2008 at Lac Hong University. Currently, he is a lecturer at Information Technology of Lac Hong University. His research interests involve computer vision, image processing and machine learning. He can be contacted at email: thanhphuong@lhu.edu.vn.

Lam Thanh Hien received the M.Sc. degree in applied informatics technology in 2004 from INNOTECH Institute, France, and his Ph.D. in 2017 at Vietnam Academy of Science and Technology. Currently, he is a rector of Lac Hong University. His research interests involve computer vision, image processing and machine learning. He can be contacted at email: lthien@lhu.edu.vn.

Ngo DucVinh received his master of IT in 2006 from the Military Technical Academy, and his Ph.D. in 2016 at the Graduate University of Science and Technology, Vietnam Academy of Science and Technology. Currently, he is a lecturer at Hanoi University of Industry. His research interests involve computer vision, image processing and machine learning. He can be contacted at email: ngoducvinh@haui.edu.vn.

Ha Manh Toan received the B.Sc. degree in applied mathematics and informatics in 2009 from College of Science, Vietnam National University, Hanoi. He is currently working as a researcher at Institute of Information Technology, Vietnamese Academy of Science and Technology. His main research interests are image processing, computer vision. He can be contacted at email: hmtoan@ioit.ac.vn.

Do Nang Toan is an associate professor at the institute of information technology, Vietnam Academy of Science and Technology. He received B.Sc. degree in applied mathematics and informatics in 1990 from Hanoi University, and Ph.D. in computer science in 2001 from Vietnam Academy of Science and Technology. His main research interests are pattern recognition, image processing and virtual reality. He can be contacted at email: dntoan@ioit.ac.vn.