Journal of Theoretical and Applied Information Technology
30th November 2020. Vol.98. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DEVELOPMENT OF INDONESIAN TALKING-HEAD ANIMATIONS ON HUMANOID AND NON-HUMANOID CHARACTERS USING THE RETARGETING PROCESS OF FACIAL MOTION CAPTURE DATA

ARIPIN[1], MUHAMMAD NASRULLOH[2]

[1]Lecturer, Faculty of Engineering, Department of Biomedical Engineering, Universitas Dian Nuswantoro, Semarang, Indonesia

[2]Student, Faculty of Electrical Technology, Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

E-mail: [1]arifin@dsn.dinus.ac.id, [2]nazh.rull@gmail.com

ABSTRACT

The difference in the facial structure of the human and the 3D virtual model (Non-Humanoid) is a challenge in the talking-head animation development. This difference resulted in the talking-head animation of the 3D virtual model that does not match with the talking-head animation of the human models. One example is the difference in the mouth width of a human model and the 3D virtual model. This research aims to develop the talking-head animation on several 3D virtual characters based on the facial motion capture (MoCap) data. The facial MoCap data are recorded using MoCap technology. A radial basis function (RBF) method is used to process retargeting of facial MoCap data on several 3D virtual characters. This method is used to map the feature points from the face source to the target face. The experimental results showed that the talking-head animation on several 3D virtual models can imitate the mouth movement of an actor. The result of the space transformation on tortoise face model has a standard deviation of 0.028261. The value of a relatively small standard deviation shows that the talking-head animation on a tortoise face model corresponds to the mouth movement of an actor. We also evaluate the talking-head animation quality on 3D virtual characters using MOS (Mean Opinion Score) method. The result of MOS calculation shows that the talking-head animation on several 3D virtual characters is 4.133. It means that the 3D virtual characters can imitate mouth movements of an actor.

Keywords: *Talking-Head Animation, Retargeting Process, 3D Virtual Characters, Facial MoCap Data, Radial Basis Function*

1. INTRODUCTION

Facial animation is an important aspect in the 3D virtual models that bring characters to life, either on human or other characters. The use of this technology includes 3D games, interactive software, and the 3D animated films. An animator takes a long time to build realistic facial animation. It is due to the complexity of facial animation. Up until now, research products are still a lot generated of facial animation, especially two important aspects, namely the facial rigging and the retargeting process of facial animation on several 3D virtual characters.

The process of animating a human face has traditionally relied on the animator to form a series of movements in the formation of the facial expression animation. The problem is when the use of the rig of human facial movements is applied in several different models, it will be time consuming for the animator [1]. Some research on the facial animation have been conducted, especially on two important aspects, they are: facial rigging process and retargeting of human facial expressions in the 3D virtual characters. In this research, we use the second aspect to build talking-head animation on several virtual characters, both humanoid and non-humanoid, as shown in Figure 1.

One of the ways used by animators to create facial animation is by using MoCap technology to

Actor Human Tortoise Shrew Serpent

Figure 1: The Talking-Head in Several Characters

record facial movement. Today, the use of motion capture is very broad, such as animated characters in the film industry, the gaming industry and others. Figure 2 shows one example of the use of MoCap in the film industry. The use of MoCap technology will provide motion information and accurate timelines, which is why MoCap technology is widely used in the quality film production. Realistic facial animation can be created using MoCap technology by placing the markers in the area of face following the facial structure [2]. Another approach commonly used by animators is the blendshapes approach [3].

Figure 2: The use of MoCap Technology in AVATAR film Production [4]

The retargeting process of facial motion from the source face to the target face can save time significantly. Motion of source face can have a variety of formats, including 2D face video, 3D MoCap face and animated facial mesh. While, models of target face are 3D facial mesh or blendshape face model. Transfer of the facial motion between two mesh of 3D face can be done through geometric deformation. The cloning techniques of facial animation can be used to transfer the vertex position changes from a single source of 3D face model into a target 3D face model that enables us to have the proportional difference of geometry and mesh structure [5]. The basic idea is to build a mapping of the vertices motion among models through changes in radial basis function (RBF). General framework written by [6] has tried to move the geometric changes between the two triangular meshes automatically. It can be directly applied retargeting process the face motion from the mesh of source face into target face.

Facial modeling is the process of making a 3D face model in the computer. Models should be made in detail and match the size and scale of the predetermined sketches so that the model looks ideal and proportional. On the other hand, facial rigging is the formation 3D facial bone structure so that it can move the 3D model [7]. Generally, the facial rigging process aims to perform planning and setting of deformation mechanism and control interface to animate characters. The process is required to manipulate every corner of the 3D model, especially for the complex face model. It requires thousands detailed vertices for deformation. Depending on the complexity of the model's face and the level of control desired, the character's face can have dozens of rig control. For example, Woody, the main character in Toy Story has 712 controls [8].

The facial structure of the human character is different to the facial structure of non-humanoid character. The example of non-humanoid characters such as monsters, animals, and other fantasy characters like aliens. Figure 3 shows an example of a non-humanoid character. The differences are in the shape of its big eyes, the wider mouth and the small lip. Other differences can also be found in the bone structure of the jaw and the skull. Adjustment of motion between a human face character and a non-humanoid face characters is usually conducted in the eyes, nose and lips. In the development of the talking-head animation, the lips become the focus in the adjustment process of motion.

Figure 3: Example of the Non-Humanoid Character Face Model

The animation retargeting is the process of the animation skeletal control at one model to another model. In general, there are differences between the structure of the existing skeleton on the facial animation including the naming of each joint and bone in the skeletal structure models. This retargeting process is done by providing a marker for each joint and bone in the animation and model that serves as the target link. Animation skeleton and the model serve as source, and marked as a target face that can be connected to the face source [9]. Figure 4 shows one of the frameworks of retargeting method written by Tamás Umenhoffer et al [10].

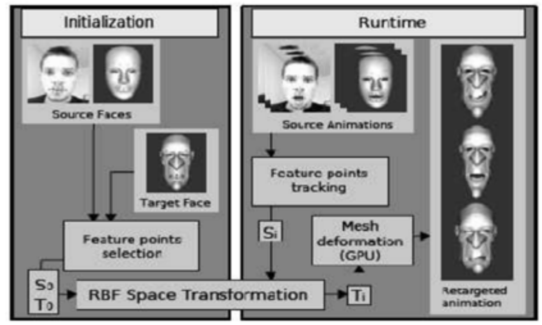


Figure 4: Retargeting Method [10]

2. RELATED WORKS

Research on the facial 3D animation on both human and avatar characters has been conducted widely. Animation techniques that are commonly used including the skeleton animation, motion capturing and retargeting. MoCap is a technology used to record the movement of an actor into a digital model. The MoCap data are mapped into a three-dimensional model so that the model shows the same action as an actor. Whereas the motion retargeting, motion sequences of a model saved and derived to be reused by other models. The motion retargeting is

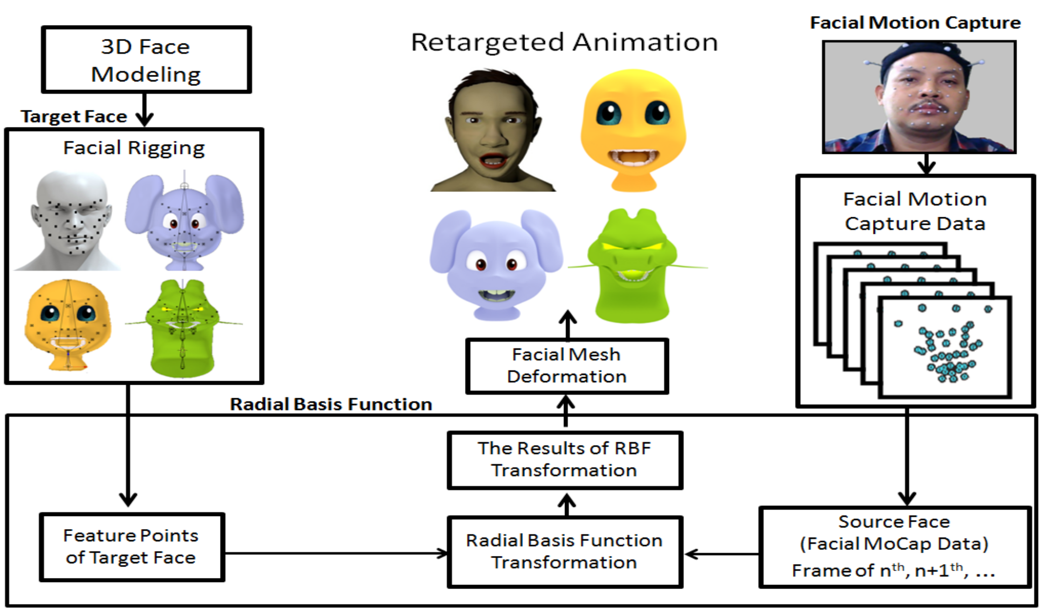


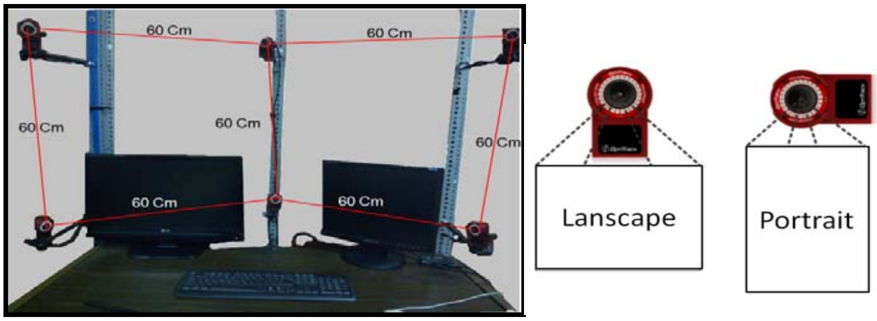
Figure 5: Framework of the approach

Journal of Theoretical and Applied Information Technology
30th November 2020. Vol.98. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

used to make the facial motion of a model becomes easier, unlike the traditional method.

process using parameter-parallel layer method to transform MoCap data for facial animation is explained in [17]. The research illustrates the result of the retargeting process for characters of facial animation that has different structure and facial configuration. The characters are human and turtle face. The turtle face has a unique structure and configuration, which is very different with the human face. The result showed that the model animation is able to visualize various emotions like blinking and speaking effectively.



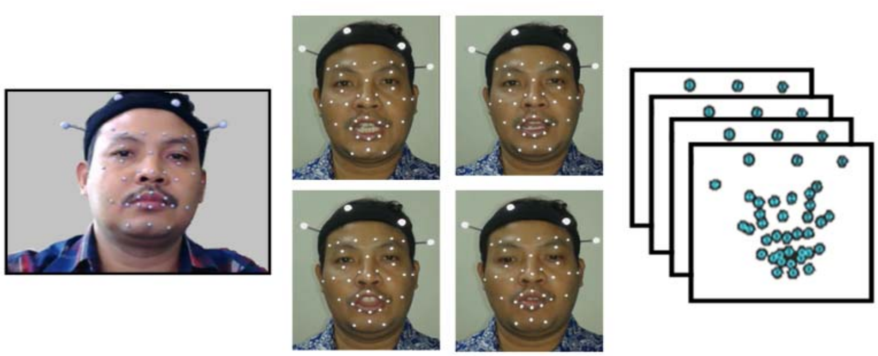
(a) (b)

Figure 6: (a) The Formation of the OptiTrack Cameras, (b) The OptiTrack Camera Type VR100:R2

An automatic rigging method is used to attach a generic rigging skeleton to 3D character animation [11]. However, this method requires complex mathematical computation when embedding rigging skeleton to 3D character animation. Gaurav et al [12] proposed a method to animate the 3D character automatically from a rigging skeleton of multi-component characters. The mesh model and rigging skeleton is used as input. A paper written by [13] presents a method to create a human animated skeleton automatically from a geometric of 3D human model. A Kinect device is used to capture the human motion. The shape of the skeleton animation generated by the Kinect device is similar to skeleton of the animator design. It can also provide information on the motion of joints.

The research [1], [15] describes in detail on the human-face motion capture and the retargeting process. Meanwhile, differences about the MoCap data formats are discussed in [16]. Retargeting

Referring to the related works, we developed the talking-head animation on humanoid and non-humanoid characters based on the retargeting process of facial MoCap data. In this research, non-humanoid characters are fantasy characters who have a different facial structure from a human facial structure. The success of this retargeting process is highly dependent on the proposed RBF method.



(a) (b)

Figure 7: (a) Markers Installation on The Actor's Face, (b) Facial MoCap Data

Journal of Theoretical and Applied Information Technology
30th November 2020. Vol.98. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3. PROPOSED METHOD

Several steps are needed to produce the talking-head animation on several 3D virtual characters. The steps are forming the facial MoCap data through the recording process of an actor's face by using MoCap technology, modeling of 3D virtual characters, making the facial rigging on each 3D facial character, conducting the targeting process on several 3D face characters and skinning mesh deformation to adapt the location of the feature points. In the retargeting process, facial MoCap data source (frame of n^{th}, $n+1^{th}$, ...) used in the process of space transformation on the target faces based on the location of the feature points. Furthermore, the deformation process is used to adapt the location of the feature points so the talking-head animation of several 3D virtual characters is formed. Figure 5 displays the framework of our approach. We use several the 3D virtual characters such as human faces, tortoise, shrew and serpent as a model to test the results of the RBF process. Each Character's face has a different facial structure.

3.1. Data Acquisition

One of the necessary equipment in the process of recording facial movements is 6 OptiTrack cameras types VR100:R2. The camera is arranged to resemble the circular arc with 120^{O} angle. Each camera can be mounted in portrait and landscape mode. Three cameras are arranged on top of the head and three cameras are arranged at chest height. Distance of the camera and the model is 60 cm as shown in figure 6.

We also use the WebCam to record sound and video. The voice data are recorded into .wav file format and Audacity software is used to process voice signals. Once the MoCap equipment is installed, the next step is to prepare the OptiTrack ARENA MoCap software that can be used to record facial movements. The process of synchronization and calibration of the OptiTrack Camera MoCap is necessary so equipment can be used to record facial movements. The synchronization process aims to set the focus of the 6 OptiTrack Camera. Meanwhile, the calibration process aims to regulate the quality and coverage area captured by the OptiTrack Camera [18].

Data recording was performed on an actor's face movement. In this research, we installed as many as 37 markers on the actor's face, where 33 markers are placed in face area and 4 markers are placed above the head. The number and the marker formation refer to the approach of FACS (Facial Action Coding System). Selection of this formation is expected to result in realistic face and mouth movements. Figure 7 illustrates the installation of markers on the actor's face and facial MoCap data resulted from the recording process.

3.2. Facial MoCap Data

The facial MoCap data is the data of facial movements of an actor which is mounted and captured using MoCap technology. While the action undertaken by the actor is to say 1529 sentences in Indonesian. The recording process produces facial MoCap data that is stored in a C3D file format. We also use a webcam to record voice and video scenes of actors. Video can be used as a reference for checking the actor's mouth movement. Meanwhile, voice data is used as one of the input data in the development of this system. We use Audacity Software to process the voice to suit the needs of the development of this system.

3.3. The 3D Face Modeling

The 3D face modeling is the process of forming a 3D face object using a computer so it looks like life. The 3D modeling process consists of several steps, namely the formation of the base object, the method of 3D object modeling, lighting and animate the movement of objects in the process sequence to be performed.

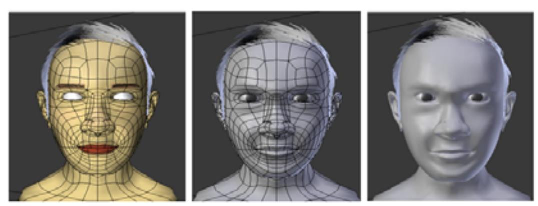


Figure 8: The 3D Human Face Modeling

There are several methods used for 3D modeling, namely NURBS (Non-Uniform Rational Bezier Spline), polygon or subdivision. Polygon modeling is a triangle and rectangle shapes that determine the surface area of a character as shown in Figure 8. Each polygon determines a flat surface by putting a polygon area so that it can create forms of surface. To obtain a smooth surface, it takes a lot of areas of polygons. Meanwhile, NURBS Modeling is the most popular method to build a model of organic. It is due to the NURBS curves can be formed by just three points. Compared with a polygon curve that requires a lot of points (vertices) that it makes is easier to be controlled. One point of CV (Control Vertex) can control one area of the texture process.

Journal of Theoretical and Applied Information Technology
30th November 2020. Vol.98. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The next step in the modeling process is rendering. In the rendering process, all data entered in the modeling process like animation, texture, lighting is transformed into a form of output. In order for the 3D models look more realistic, the lighting is required. There are several lighting techniques such as key light, fill light, back light, kickers light and bounce light.

3.4. Facial Rigging on Several 3D Face Models

Facial Rigging is the process of making the control of face to move facial muscles. Facial rigging aims to provide markers as the feature points when moved using MoCap data. The facial rigging process is performed on each of the 3D face model manually. Each the 3D face model is given the feature points that refers to the FACS approach used in the OptiTrack MoCap system. Figure 9 shows the facial rigging process on two 3D virtual characters. They have different facial structures than the structure of the human face. They have a wider mouth shape, larger eyes, smaller nose and shorter the distance between the eyes and mouth. Therefore, determination of rigging at two 3D virtual characters adapts to the structure of their faces.

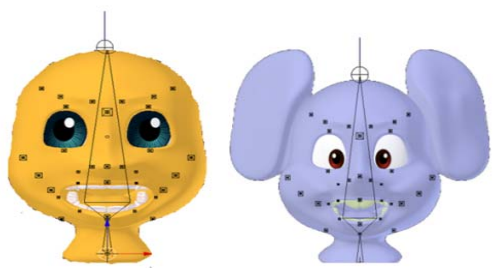


Figure 9: Facial Rigging on Two 3D Virtual Characters

The 3D character models equipped with bones at the point features that have been determined as blendshape control of the character's face. 3D character models are actuated in accordance with the facial MoCap data. The next step is connecting the facial MoCap data with bones at each feature point on the character's face.

3.5. Radial Basis Functions (RBF)

Radial Basis Functions (RBF) is commonly used in computer graphics applications interpolated surface [19] and the problem of scattered data interpolation [20]. In this research, RBF is used as a space transformation from data of the source face into the target face. The initial step of the transformation process of space is defining two spaces with two sets of feature points. Each function can be interpolated with a degree of accuracy that varies with RBF formula as in Equation (1).

$$y(x) = \sum_{i=1}^{N} \omega_i \phi(\|x - x_i\|) \qquad (1)$$

where N is the number of points and ϕ is a radial function, its value depends on the distance from the source file (in this equation is x_i). There are several radial functions commonly is used. In this research the multi quadratic function as shown in Equation (2) is used.

$$\phi(r) = \sqrt{1 + (\varepsilon r)^2} \qquad (2)$$

Each component of the target parameter defines the set of the RBF, with a weighted value of ω_i should be calculated.

The number of a basis functions (N) is the number of samples that have been determined, x_i is the source configuration of the sample of *i*. The target configuration of $y_i = y(x_i)$ defines the control points of the set of the interpolation function. The weighting value of x can be computed by completing linear equations such as shown in Equation (3).

$$T = H.W \qquad (3)$$

where $H_{ij} = \phi(\|x_j - x_i\|)$, $W_i = \omega_i$, and $T_i = y_i$.

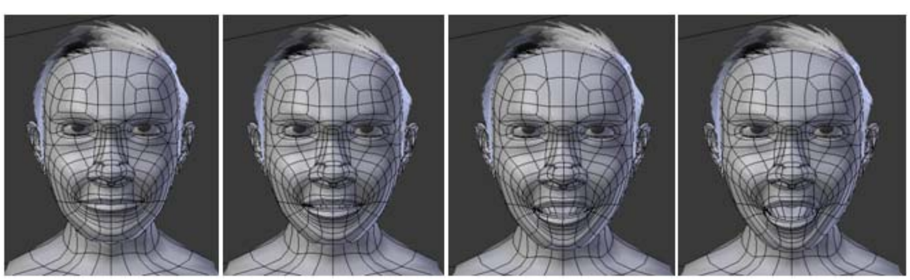


Figure 10: Visualization of changes in the mouth shape

3.6. Facial Mesh Deformation

Deformation is an image processing technique that converts the image by using a reference point in making changes to the image [21]. On the other hand, mesh modeling is a basic technique of modeling that is widely used in most 3D software. The mesh modeling is a vertex-based modeling or face by face, which means it is only possible to be conducted on vertex level. Facial mesh deformation is an attempt to change the form of facial mesh to produce a quality facial animation as shown in figure 10. Mesh deformation aims to smooth the mouth movement of the talking-head animation. The result is the 3D face model that is able to animate the face animate the facial movements of human actors.

One of the deformation methods commonly used is FFD (Free Form Deformation). This method can deform the 3D object surface so that results in a curve shape freely. When a control point is moved, then the vertices effected control point of an object will move to follow. Deformation by this method produce smoother object surface.

3.7. Retargeting Process

The 3D face models be moved adjusting face MoCap data in C3D format. Retargeting process served to connect the data C3D with bone in each point of 3D face models. It produces mouth movements on the 3D face models which followed a face MoCap. Retargeting process can be done through the mapping process by using RBF method as described previously.

The results of the retargeting process continued with the facial mesh deformation using linear blend skinning (LBS). It is used to adapt to changes in the location of the feature points by calculating the position of the feature points as a point of joint motion. Weights for each of the feature point is defined as $\sum W_i = 1$ and bone weights i^{th} is $W_i \in [0,1]$. Weight with a value of 0 means that the feature points are not influencing the vertices in the mesh point. If the weight value is 1, it means that the vertex point is entered. It also means that the vertex point in the mesh is affected only by the feature point. The vertex $\vec{v}$ in frame f is defined by Equation (4).

$$\overrightarrow{v_f} = \overrightarrow{v_0} + \sum_{i=1}^{N} W_i . \overrightarrow{d_{if}} \qquad (4)$$

where N is the number of points and $\vec{d}_{if}$ is changes in the location of feature points of i^{th} in f frame.

4. EXPERIMENTAL RESULTS

In this research, we have done several steps including: acquisitions of facial MoCap data, modelling and facial rigging process of the 3D virtual characters, retargeting process on several 3D virtual characters. The markers placement in several areas such as eyebrow, orbital upper, eyelids, ear, orbital lower, nostril base & bulge, puffer, lip & mouth, and jaw end, markers are placed on the face part of the left and right symmetrically.

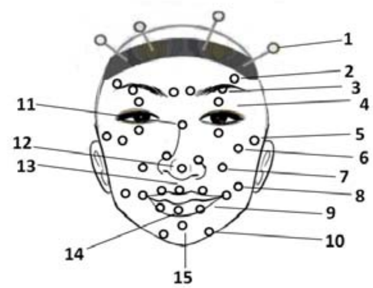


Figure 11: Placement area of markers on an actor's face

Table 1: The number of markers each area

No	Face Area	The Number of Markers
1	Head	4
2	Eyebrow	6
3	Orbital upper	2
4	Eyelids	4
5	Ear	2
6	Orbital Lower	2
7	Nostril Base & bulge	2
8	Puffer	2
9	Lip and Mouth	6
10	Jaw end	2
11	Nose bridge	1
12	Nose tip	1
13	Upper tip	1
14	Lower tip	1
15	Chin	1

The 6 OptiTrack cameras (0.3 MP, 100 FPS) is used on MoCap technology. The data recording is based on placing markers on the actor who refers to the approach of Facial Action Coding System (FACS). Markers placement position on an actor's face refers to facial feature points as shown in Figure 11 and Table 1. The placement of markers at the corresponding position is very helpful in the process of recording to obtain a facial motion data.

To determine the displacement of the feature points on several 3D virtual characters, calculation of the movement of the feature points result of retargeting process is needed. We used 13 feature

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

points representing each marker position in the area of the mouth of the 3D virtual face. Each of the feature points is considered to represent the animated talking formation of each 3D virtual character. Figure 12 shows the feature points of each 3D virtual character. The feature points on human models used as reference to the placement the feature points of other 3D virtual characters.

Retargeting process was tested on several 3D virtual characters from facial MoCap data. The results of the retargeting process show that the 3D virtual characters can visualize the talking-head animation similar to those exhibited by the actor. Figure 13 displays the results of visualization of 3D virtual characters that imitate the mouth movement by an actor. It's shown that the process of data retargeting mocap into several 3D virtual characters can generate realistic talking-head animation.

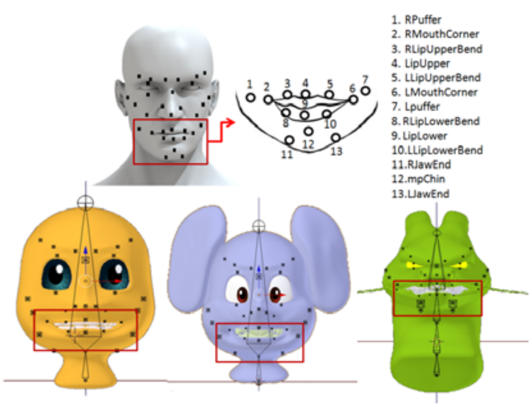


Figure 12. Position of feature points from each 3D virtual character

To find the displacement value of the feature points of 3D virtual characters (in this experiment we focused on tortoise), the calculation of the feature points of the RBF result is required. We calculate 13 feature points of the tortoise face model that represent the talking-head animation. In this experiment, we use seven frames of facial MoCap data and data from the RBF process. Facial MoCap data is used as reference to find displacement of the feature points. When the feature points of facial MoCap data are changed, then the feature points on the target face move in accordance with the data source. Table 2 displays displacement the feature points of the RBF process of the tortoise face models.

Table 2: The Result of Displacement of the feature points on the Tortoise Face Models

Feature Points	Frame# 0	1	2	3	4	5	6
FP#1	0	0.03280	0.01802	0.02002	0.02211	0.05924	0.04118
FP#2	0	0.04272	0.05779	0.05979	0.05758	0.08291	0.04578
FP#3	0	0.04050	0.06330	0.06530	0.06274	0.09174	0.06750
FP#4	0	0.00043	0.02840	0.03040	0.01781	0.05430	0.06513
FP#5	0	0.03977	0.00819	0.01019	0.01085	0.07348	0.03860
FP#6	0	0.02760	0.01664	0.01864	0.02070	0.07058	0.05456
FP#7	0	0.05022	0.05479	0.05679	0.05004	0.08394	0.02491
FP#8	0	0.03332	0.05128	0.05328	0.05462	0.08032	0.05359
FP#9	0	0.06005	0.07867	0.08067	0.06761	0.06653	0.03725
FP#10	0	0.07162	0.08527	0.08727	0.06401	0.05020	0.02566
FP#11	0	0.04978	0.07737	0.07937	0.07511	0.06909	0.04591
FP#12	0	0.07993	0.08839	0.09039	0.08011	0.05988	0.03526
FP#13	0	0.05601	0.07623	0.07823	0.06710	0.04001	0.02210

In the frame of 0th, value of the displacement is 0. Frame of 0th is the initial frame the movement of feature points. Meanwhile, in the next frame (frame # 1 to 6) occurs the frame displacement which indicates the movement of the feature points.

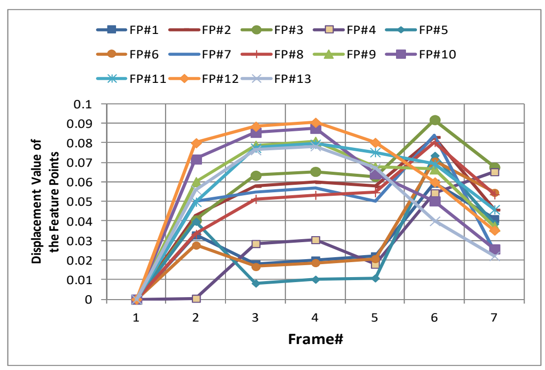


Figure 13: Graph of Displacement the feature points on the Tortoise Face Models

Figure 13 shows that occurs displacement the feature points of the tortoise face model. The displacement process of the facial feature points occurs in a linear manner and depend on the morphological form of the models. Morphological form of human or similar to humans will be illustrated by similar graph. Meanwhile, the models have different morphologies such as tortoise, shrew, and the serpent had a different displacement graph.

The correspondence degree of the RBF process result can be computed using the standard deviation. Standard deviation illustrates the differences between sample value and the average. The calculation result of the standard deviation of the feature points as shown in Table 2 is 0.028261. The value of a relatively small standard deviation shows that the mouth movement of the talking-head animation (tortoise) and an actor are nearly similar. The mouth movement of the tortoise model can exactly mimic the mouth movement of an actor.

Journal of Theoretical and Applied Information Technology
30th November 2020. Vol.98. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 3: Indonesian texts spoken by each 3D virtual character

No	The Indonesian Texts
1	saya suka baju yang berwarna merah tua
2	boneka beruang di toko itu lucu sekali
3	sepatuku kotor belum aku cuci dari kemarin
4	jalan itu ramai sekali setiap pagi hari
5	lusa aku akan pergi ke rumah paman
6	besok saya pergi ke toko baju bersama ibu
7	toko itu buka mulai jam enam pagi
8	ibu menyirami bunga di halaman setiap sore
9	buanglah sampah di tempat yang telah disediakan
10	masakan nenek paling enak di rumah kami

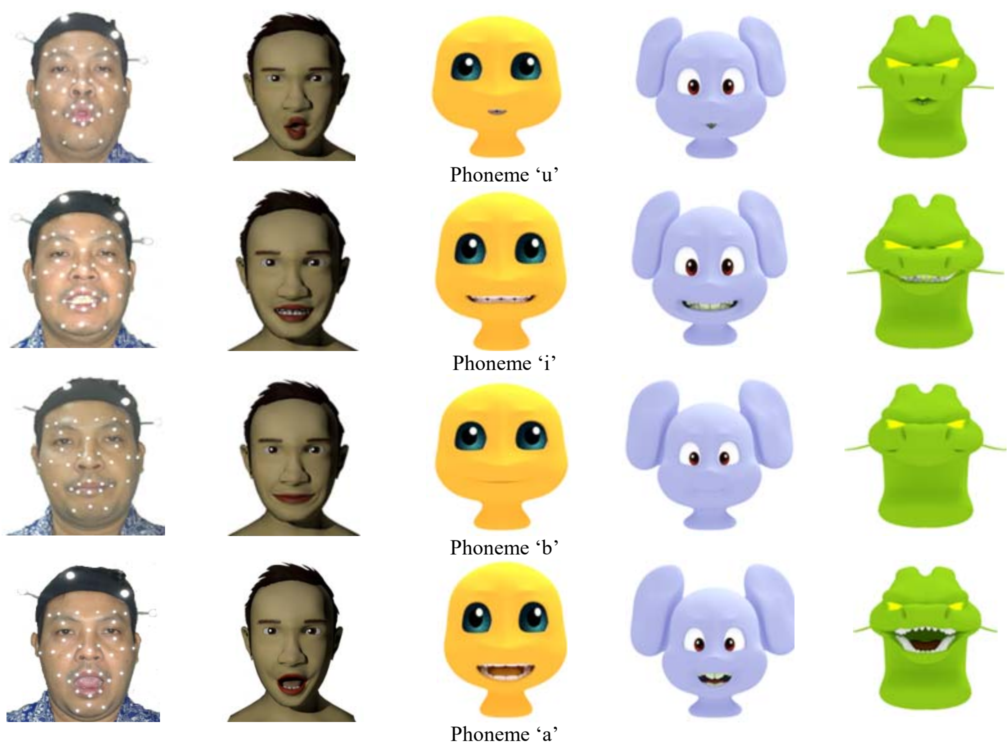


Figure 14: The results of retargeting process on several 3D Virtual Characters

$$MOS = \sum_{i=1}^{n} \frac{y(i).w}{J} \quad (5)$$

Where $y(i)$ is the sample value of i^{th}, w is the number of weight and J is the number of respondents.

Table 4: The MOS assessment criteria

MOS value	Quality	Description
5	Very Good	Quality of mouth movement against Indonesian spoken text is very good
4	Good	Quality of mouth movement against Indonesian spoken text is good
3	Adequate	Quality of mouth movement against Indonesian spoken text is adequate
2	Bad	Quality of mouth movement against Indonesian spoken text is bad
1	Very Bad	Quality of mouth movement against Indonesian spoken text is very bad

Figure 14 displays The results of retargeting process on several 3D Virtual Characters. The talking-head animation resulting from this research is presented in Figure 15. From this talking-head animation the phoneme pronunciation scene in Figure 14 is obtained. To measure the talking-head quality of each 3D virtual character, we conduct a subjective test to 30 respondents. Each respondent provides an assessment about the Correspondence degree of the mouth shape of each virtual 3D character for pronouncing 10 Indonesian texts as shown in Table 3. The method used to measure it is

MOS (mean opinion score). Each respondent gives descriptive assessment with 5 levels of correspondence criteria as shown in Table 4. Furthermore, MOS can be calculated using Equation (5).

Table 5: Recapitulation of respondents' assessment for each 3D virtual character

3D Virtual Characters	The talking-head quality of each character				
	Very bad	Bad	Adequate	Good	Very Good
Human	0	0	5	10	15
Tortoise	0	0	5	12	13
Shrew	0	1	8	9	12
Serpent	0	3	7	11	9

Results of the assessment of all respondents were calculated for each of the 3D virtual characters as seen in Table 5. Next, we calculate MOS using Equation (6) and the result obtained is 4.133. This value indicates the average assessment of all respondents to each talking-head of 3D virtual characters. The MOS value indicates that the level of conformity of each talking-head of 3D virtual characters against the Indonesian text is good.

Figure 15: Talking-Head Animation Resulting From This Research

5. CONCLUSION AND FUTURE WORK

The retargeting process has been implemented in several 3D virtual characters. The retargeting process by using RBF can create talking-head animation of several 3D virtual characters similar to an actor's mouth movements. The 3D virtual characters are able to imitate the mouth movements of an actor. It is based on the results of the standard deviation calculation and subjective testing by the respondents using MOS method. It shows that the level of conformity of the talking-head animation of RBF space transformation is a linear approach that does not require complex computational processes. The use of the RBF space transformation approach can provide a solution for the development of realistic and flexible talking-head that is in harmony with the mouth movement of the actor.

There are important points that can be emphasized from the results of this research, namely : 1) facial MoCap data can be used in the retargeting process to several virtual characters, both humanoid and non-humanoid, 2) the use of 37 markers on an actor's face can produce a talking-head animation that can mimic the actor's facial movements include facial expression, 3) The RBF space transformation method can be used to overcome the difference in dimensional space between the coordinates of the source face shape and the coordinates of the animated target face shape. If the position of the marker points on the actor is not symmetrical, it will not cause deviation in the retargeting process on the animated target face.

Based on the experimental results and descriptions above, it can be concluded that the retargeting process of facial MoCap data using the proposed RBF method in this research can produce realistic and natural Indonesian talking-head animation mimicking the actor's speaking.

The use of MoCap technology in this research has many advantages such as high accuracy, actor's freedom of movement, high recording speed, face expressions recording, and most popularly used [22],[23]. The disadvantages of using MoCap technology are requirement of special programs and long time to manipulate MoCap data after it is retrieved and processed. If MoCap data input is wrong, data recording process must be repeated. The other disadvantages are the susceptible recording process against light and the process would become more complicated if the actor is more than one person [24]. Some of the efforts made are to arrange the recording process scenario and perform the retargeting process on some virtual character models. To produce an animation with realistic facial expression then more markers should be placed on the actor's face.

6. ACKNOWLEDGEMENTS

The author would like to appreciate the '*Program Penelitian Terapan Unggulan Perguruan Tinggi (PTUPT) Direktorat Riset dan Pengabdian Masyarakat*' of the Ministry of Research and Technology Republic of Indonesia then this reserach to be completed.

REFERENCES:

[1] Aleix Martinez, Shichuan Du, "A Model of the Perception of Facial Expressions of Emotion by Humans: Research Overview and Perspectives", Journal of Machine Learning Research, May 2012, pp. 1589-1608.
[2] Akinobu Maejima, Hiroyuki Kubo, and Shigeo Morishima, "Realistic Facial Animation by Automatic Individual Head Modeling and Facial Muscle Adjustment", Virtual and Mixed Reality - Systems and Applications, pp 260-269, 2011.
[3] Clément Reverdy, Sylvie Gibet, Caroline Larboulette, "Optimal Marker Set for Motion Capture of Dynamical Facial Expressions", MIG2015 8th ACM SIGGRAPH Conference on Motion in Games, Paris, France, Nov 2015, pp.31-36.
[4] Deepali Aneja, Bindita Chaudhuri, Alex Colburn, Gary Faigin, Linda Shapiro and Barbara Mones, "Learning to Generate 3D Stylized Character Expressions from Humans", 2018 IEEE Winter Conference on Applications of Computer Vision, DOI 10.1109/ WACV.2018.00024, 2018, pp. 160-169.
[5] Qiang Xu*, Yaping Yang, Qun Tan and Lin Zhang, "Facial Expressions in Context: Electrophysiolo-gical Correlates of the Emotional Congruency of Facial Expressions and Background Scenes", Department of Psychology, Ningbo University, Ningbo, China, 12 December 2017.
[6] Shaojun Bian , Anzong Zheng, Lin Gao, Greg Maguire, Willem Kokke, Jon Macey, Lihua You, and Jian J. Zhang, "Fully Automatic Facial Deformation Transfer", Symmetry - Open Access Journal, Symmetry 2020, 12, 27; doi:10.3390/ sym12010027, Published: 21 December 2019.
[7] Orvalho, V., Bastos, P., Oliveira, B., and Alvarez, X., "A Facial Rigging Survey", In Proc. of the 33rd Annual Conference of the European Association for Computer Graphics-EUROGRAPHICS, May 13-18, Cagliari, Italy, Volume 32, 2012, pp. 10-32.
[8] Samuel Gandang Gunanto, Mochamad Hariadi, and Eko Mulyanto Yuniarno, "Computer facial animation with synthesize marker on 3D faces surface", 2016 2nd International Conference of Industrial, Mechanical, Electrical, and Chemical Engineering (ICIMECE), DOI: 10.1109/ICIMECE.2016.7910452, 27 April 2017.
[9] Ning Kang, Junxuan Bai, Junjun Pan, Hong Qin, "Real-time Animation and Motion Retargeting of Virtual Characters Based on Single RGB-D Camera", 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), DOI: 10.1109/VR.2019. 8797856, 2019.
[10] Tamás Umenhoffer and Balázs Tóth, "Facial Animation Retargeting Framework Using Radial Basis Function", Sixth Hungarian Conference on Computer Graphics and Geometry, Budapest, 2012.
[11] Matahari Bhakti Nendya, Eko Mulyanto Yuniarno, and Samuel Gandang Gunanto, "Facial Rigging for 3D Character", International Journal of Computer Graphics & Animation (IJCGA) Vol. 4, No. 3, July 2014, pp. 21-29.
[12] B. Gaurav, T. Thormählen, H. P. Seidel, and C. Theobalt, "Automatically rigging multi-component characters", Computer Graphics Forum, vol. 31, Blackwell Publishing Ltd, 2012, pp. 755-764.
[13] Abdul Razzaq, Zhongke Wu, Mingquan Zhou, Sajid Ali, and Khalid Iqbal, "Automatic Conversion of Human Mesh into Skeleton Animation by Using Kinect Motion", International Journal of Computer Theory and Engineering, Vol. 7, No. 6, December 2015, pp. 482-488.
[14] L.-Y. Yu, Q. Han, and X. Niu, "An improved contraction-base method for mesh skeleton

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

extraction", Multimedia Tools and Applications, 2013, pp. 1-14.

[15] Kfir Aberman, Rundi Wu, Dani Lischinski, Baoquan Chen, Daniel Cohen-Or, "Learning Character-Agnostic Motion for Motion Retargeting in 2D", Computer Vision and Pattern Recognition SIGGRAPH 2019, DOI: 10.1145/ 3306346.3322999, May 2019.

[16] Eline van der Kruk and Marco M. Reijne, "Accuracy of human motion capture systems for sport applications; state-of-the-art review", European Journal of Sport Science, Volume 18, Issue 6, 2018.

[17] Mekides Assefa Abebe and Jon Yngve Hardeberg, "Application of Radial Basis Function Interpolation for Content Aware Image Retargeting", 2018 14th International Conference on Signal-Image Technology & Internet-Based Systems, DOI:10.1109/ SITIS.2018.00035, 2018.

[18] KyungHun Cho, Xi Chen, "Classifying and Visualizing Motion Capture Sequences Using Deep Neural Networks", Department of Infomration and Computer Science Aalto University School of Science, Finland, 2013.

[19] Yihjia Tsai, Hwei Jen Lin, and Fu Wen Yang, "Facial Expression Synthesis Based on Imitation", International Journal of Advanced Robotic Systems, Vol. 9, 148: 2012, First Published May 15, 2017, pp. 1-6.

[20] Troy, Pranowo, and Samuel Gandang Gunanto, "2D to 3D space transformation for facial animation based on marker data", 2016 6th International Annual Engineering Seminar (InAES), DOI: 10.1109/INAES. 2016.7821896, 2016.

[21] Shuang Li, Xiang Li, Xiaohan Li, and Xiang Zhou, "Global Deformation Model for 3D Facial Combination", Twelfth International Conference on Digital Image Processing, 2020, Osaka, Japan, https://doi.org/10.1117/ 12.2573892, June 2020.

[22] Tiago Henrique Ribeiro and Milton Luiz Horn Vieira, "Motion Capture Technology; Benefits and Challenges", International Journal of Innovative Research in Technology & Science (IJIRTS), ISSN : 2321-1156, Vol. 4, Number 1, January 2016, pp. 48-51.

[23] Slava Milanova Yordanova, Nikolay Todorov Kostov and Yasen Dimchev Kalchev, "Motion Capture System Overview and Accelerometer MoCaps System", Faculty of Computing and Automation, Technical University of Varna, Bulgaria, ISBN: 978-1-61804-247-7, 2014, pp. 138-141.

[24] Ashish Sharma, MukeshAgarwal, Anima Sharma, and Pankhuri Dhuria, "Motion Capture Process, Techniques and Applications", International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC), ISSN 2321-8169, Vol. 1 Issue : 4, April 2013, pp. 251-257.