

Face recognition techniques using artificial intelligence for audio-visual animations

Daniel Izario
Bruno Izario
Diego Castro
Yuzo Iano

Cite this article:

Izario, Daniel, Izario, Bruno, Castro, Diego, Iano, Yuzo; 2017. Face recognition techniques using artificial intelligence for audio-visual animations. SET INTERNATIONAL JOURNAL OF BROADCAST ENGINEERING. ISSN Print: 2446-9246 ISSN Online: 2446-9432. doi: 10.18580/setijbe.2017.11. Web Link: <http://dx.doi.org/10.18580/setijbe.2017.11>

Face recognition techniques using artificial intelligence for audio-visual animations

¹Daniel Izario, ²Bruno Izario, ³Diego Castro and ⁴Yuzo Iano

Abstract—This work developed an artificial intelligence based web tool that uses face recognition for computer animation applications. It enables that the producers can simulate the face movements of the actors in a digital environment. All the inherent applications of the tool favor the area of computer-animated films and television series.

Index Terms—Artificial Intelligence, Computer Animation, and Expressive Facial Animation.

I. INTRODUCTION

COMPUTER facial animation based on artificial intelligence is a very challenging topic and has become one of the most used techniques for the development of cartoons and film productions. Among the most well-known computer graphic techniques, the stop motion technique was one of the first technique used in filmmaking, animations and live-actions. This technology consisted in the reproduction of objects in a scene thought of a series of aligned static images. Due to the advent of new technologies, 3D animated graphics allowed to the creators have access to different perspectives and layers of an object, however the time spent used to create this type of animations increases because of its high complexity. [1][2]

For this reason, new computer facial animation techniques should include artificial intelligence and machine learning based methods. For instance, it is possible to extract information using face recognition algorithms based on the movements of the interpreter's own face. In base on this premise, this work will explain the developed web tool in a sequential order of steps. The first process is to record the movements and expressions of interpreter's face. These samples are the input of the training process and learn how each interpreter facial movement is performed. Then, the respective weights extracted from the past process are stored in the database for later use. The next step is to create a 3D pre-modeling of the face to generate the computer animation of the character.

Item II explains the stages involved in the development of the application and its functionalities. Item III, presents the final results. Finally, Item IV presents the conclusions and future works.

All Author are with the Laboratory of Visual Communications from University of Campinas/SP, Brazil. For ¹Daniel Izario the e-mail is daniel_izario@hotmail.com; ²Bruno Izario is bruno_izario@hotmail.com; ³Diego Castro is diego.pajuelo.castro@gmail.com; and ⁴Yuzo Iano is yuzo@decom.fee.unicamp.br.

II. DEVELOPMENT OF THE APPLICATION

This tool enables to create animations based on the author's face interpretation in real time. As a part of the development, the HTML5 (HyperText Markup Language) [3], CSS3 (Cascading Style Sheets) [4] and JavaScript [5] programming languages were used.

The first stage is the dubbing process of the animated character; this process consists of giving life to animated characters. Actors interpret scenes in front of screens and microphones for many hours in order to recreate the more real expression of character's feelings required on the scene. The application detects and captures the facial movement and stores the feature points as is shown in Fig. 1.

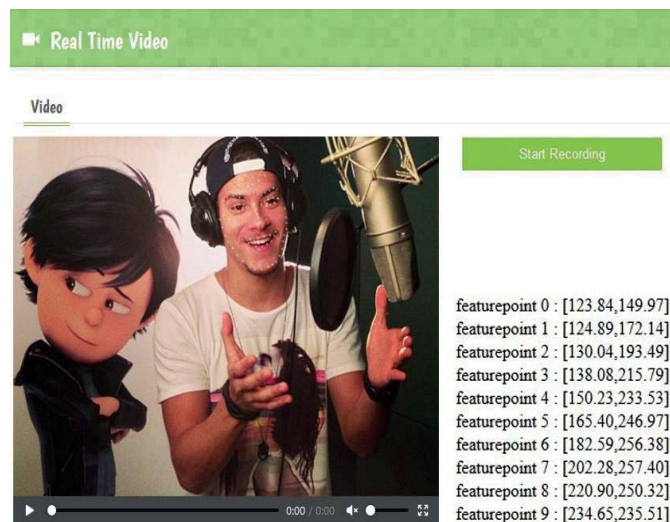


Fig. 1. Character dubbing and capturing facial points.

Despite these feature points are used to pre-model the animated character. This work used this information to train the neural network, it means that a computer could create an animated character as a combination of these features. For this application, is constructed after 7 steps, as seen in Fig. 2. The developed tool uses spreadsheets of models, which are precisely groups of images created by points that show all the possible expressions that a character can have, and all the different positions that they can adopt according to the recorded images. [6] Thus, the tool can generate a modeled face with the expressions of the actor when talking in stage 5 of the development process, reducing the time spent of audio-visual productions.

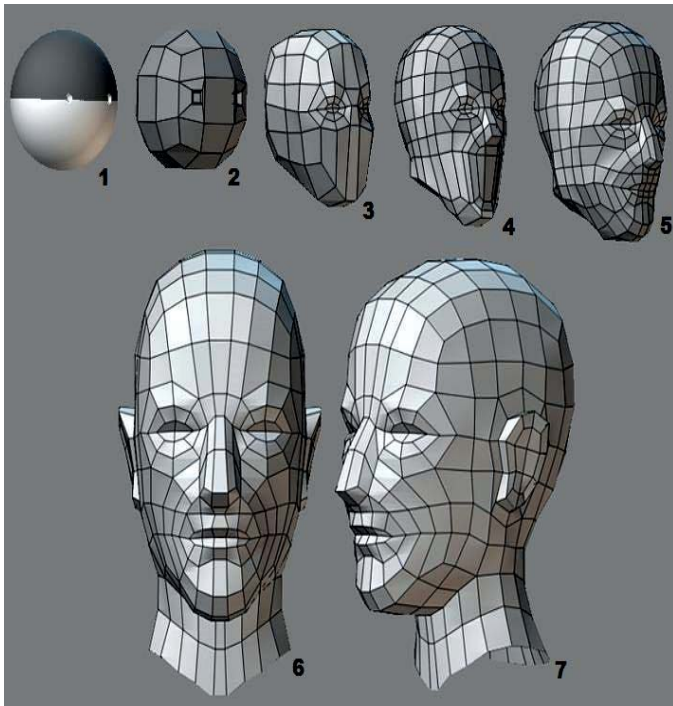


Fig. 2. Standard modeling of a character in 7 steps.

The spreadsheets are created to remain the character's details and the drawing be uniform. The captured expressions are categorized and stored as specific frames, as is shown in Fig. 3. Thus, animators, along the course of the editions, can choose a state of the character's face for a given scene, simplifying the entire creation stage.

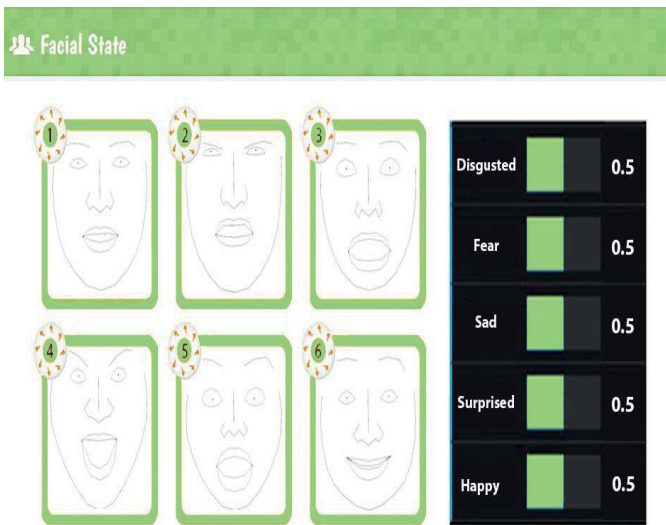


Fig. 3. Standard facial modeling with adjustment panel.

Once the modeling stage is finished, the creation of layer-upon-layer color using the RGB color model is performed. The tool enables to the user to insert the 3 layers created by the color histogram, also known as color frequency distribution, which is a graphical representation in columns with the data sets previously tabulated and divided into classes automatically to check the possible discrepancies of the image, as shown in Fig. 4.

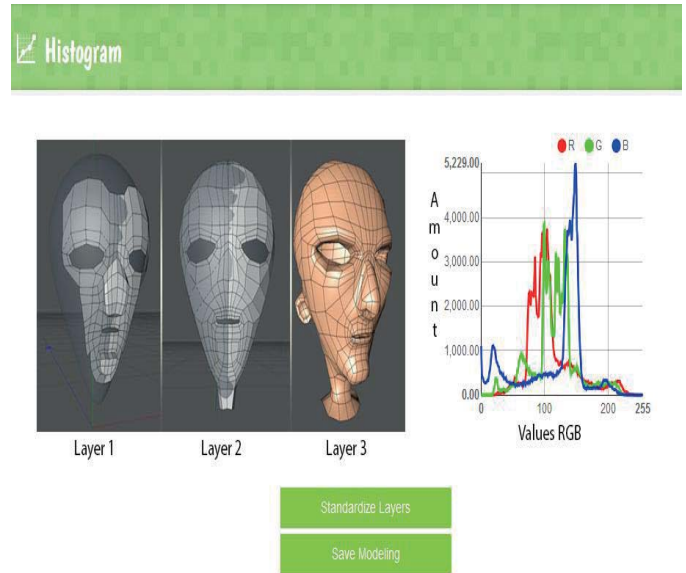


Fig. 4. Histogram of color for each layer.

If there is any discrepancy, the system acknowledges and prevents the user to select a wrong model. After the correction, the model can be saved and the next step is applied.

The modeling in this stage is raw, only the layer was made with the colors by frequencies, and it is now necessary to add the textures, which are created to fit the artistic concepts planned with the plot created by the scripter of the film. These textures are created in the form of maps and associated with the model, as is shown in Fig. 5.



Fig. 5. Texture created in map form.

Finally, the application allows uploading the created face with texture, unifying theme, and forming the final file, as is shown in Fig. 6.



Fig. 6. Examples of facial modeling, with mapped texture applied.

The entire system has restoring points at each stage of creation of the animations to recover valuable information if were the case of errors during the playing.

A. Artificial Intelligence based method

An Artificial Intelligence (AI) system is not only capable of storing and manipulating data, but of acquiring, representing, and manipulating knowledge. This manipulation includes the ability to deduce or infer new knowledge (new relationships about facts and concepts) from prior knowledge, to solve complex problems. [7] Thus, the main goal of the tool is the application of machine learning techniques. By doing so, the system can learn the movements made by the author at the time of the interpretation of the character using the feature points previously stored.

In total, there are 312 captured and stored facial points represented by the vector (x, y, z) in the recorded environment. These values are binaries to apply machine instructions among the samples [8].

Once the tool can gain more knowledge, the system will be increasingly accurate and efficient, since it is a slow learning process, and requires a high number of data to perform actions close to reality. Due to this, this tool stores the dots and movements to train the neural network, but it is not able to generate an entire computer animation with all the molded movements. However, the diagram with all the steps covered by the developed tool is shown in Fig. 7.

The current architecture of the neural network consists of a finite amount of data for the training process. In mathematical

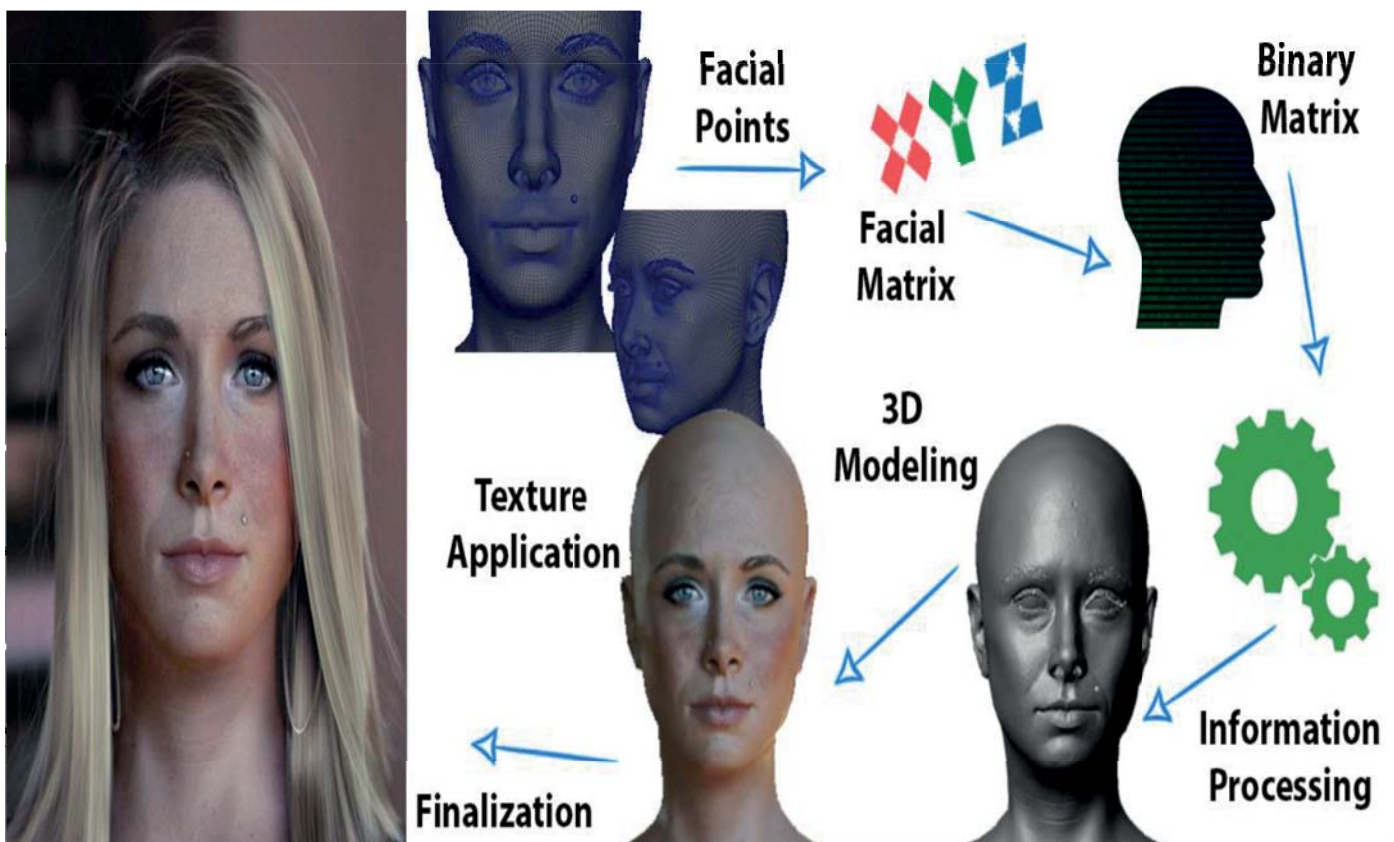


Fig. 7. Diagram of the complete technique.

function $f(x)$ (which generated the data), this function would have many errors, as shown in Fig. 8.

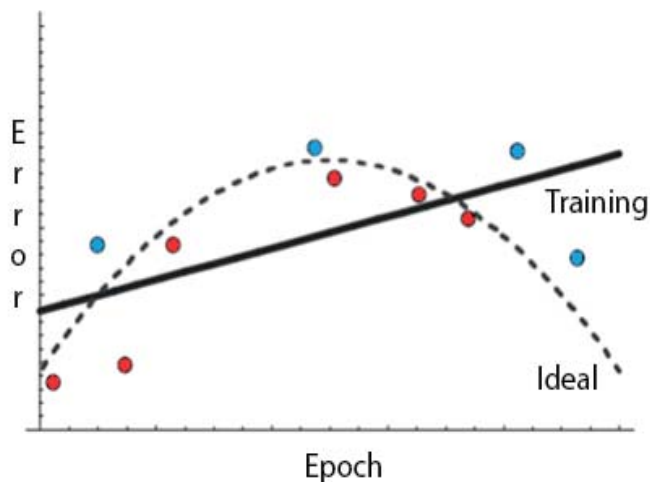


Fig. 8. Example of learning with many cataloged errors.

By the use of the developed tool, the application will be able to be adequately trained to an ideal level. Consequently, the errors will be lesser during the entire process and those would be imperceptible, as shown in Fig. 9.

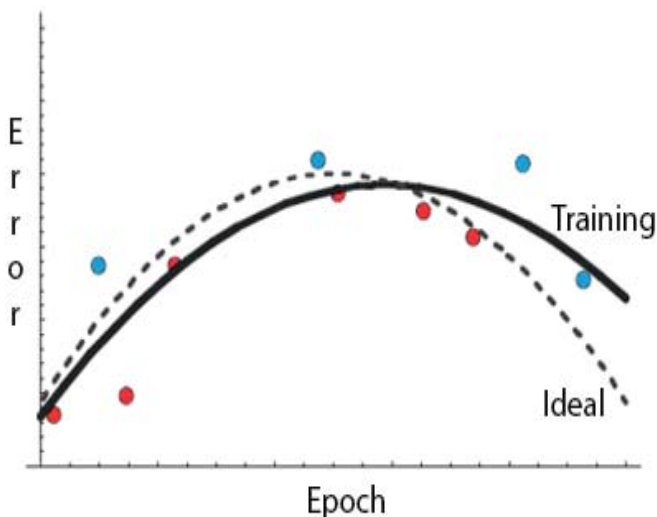


Fig. 9. Example of ideal learning for the application.

B. Neural Network based method

Artificial neural networks are computational techniques that present a mathematical model inspired by the neural structure of intelligent organisms and that acquire knowledge through experience. They are used to solve AI problems where a human brain is simulated, including its behavior, which is, learning, making mistakes and making discoveries. [9]

This work used the Hopfield model as neural network model. This network serves as a content memory system that can be addressed with binary limit nodes. It is ensured that the animation is converted to a local minimum, but sometimes converge to a false standard through mathematical analysis tests, the latter case appears in only 12% of all content

generated, without distortion of the final results [10].

The overall network consists of feeding its output into its own inputs, a feedback process. For instance, the units assume a binary state (active or inactive) and these units are connected together by symmetrical edges with different weights. Whether the edge has positive weight, it indicates that the two units tend to activate by itself. By the contrary, negative values indicates that an active unit can deactivate another unit. [11] The mathematical model for this representation is (1):

$$s_i \leftarrow \begin{cases} +1 \therefore \sum_j \omega_{ij}s_j \geq \theta_i \\ -1 \therefore \sum_j \omega_{ij}s_j < \theta_i \end{cases} \quad (1)$$

Where:

ω_{ij} is the weight of the unit connection weight j the unity i ;

s_j is the state of unity j ;

θ_i is the limit of unity i ;

ρ is otherwise;

The last analysis is the energy configuration of the network, highlighting the current state of the network (above the valley). This work looks for an attractive state to which it would converge. The minimum energy level and a shaded attraction area is shown in Fig. 10.

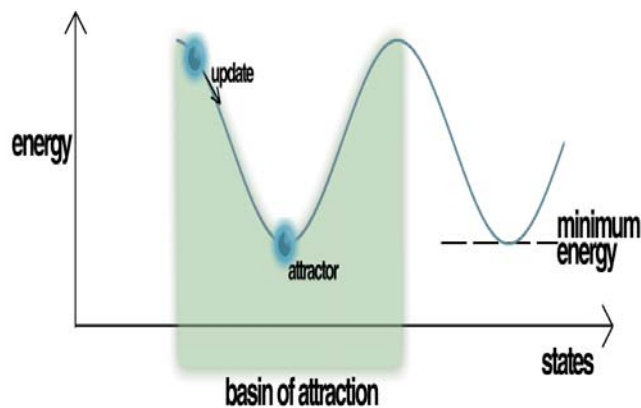


Fig. 10. Example of power configuration in the Hopfield network.

The energy (E) represented in (2) guarantees that, when the units to be updated are randomly chosen, the energy will decrease in value or remain the same. [11]

$$E = -\frac{1}{2} \sum_{i,j} \omega_{ij}s_i s_j + \sum_i \theta_i s_i \quad (2)$$

III. RESULTS

The developed tool allows to the user approximate the contour of a face. This can be adjusted regarding to the perspective of the animation being used in the scene. After the facial mapping process generated by a computer, it is possible to change the points to the desired pattern, creating a tabulation, as shown in Fig. 11.

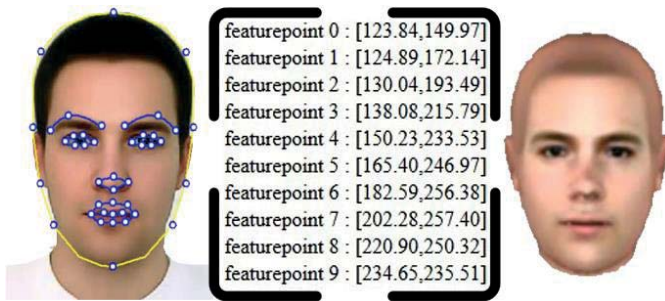


Fig. 11. Expected result.

IV. CONCLUSION

This article discusses a simple tool, easy to implement and of high importance in the cinematographic industries. The incorporation of an artificial intelligence based proposal is emerging and is expected to be part of the next-generation of animation techniques. As a future work, the application would create pre-models of faces and movements using the generated patterns after applying artificial intelligence based methods, making the computer animation process a more simple task for producers.

REFERENCES

- [1] P. Gulati, "Step-by-Step: How to Make an Animated Movie", Envato Tuts+. Available in: <<https://cgi.tutsplus.com/articles/step-by-step-how-to-make-an-animated-movie--cg-3257>>. Access in: 22 July 2017.
- [2] N. M. Thalmann; D. Thalmann, "Complex Models for Animating Synthetic Actors", Journal IEEE Computer Graphics and Applications archive, vol. 1, pages 32-44, 1991.
- [3] J. Keith; R. Andrew, "HTML5 for Web Designers", 2nd ed., A Book Apart, 2016.
- [4] D. Cederholm, "CSS3 for Web Designers", 2nd ed., A Book Apart, 2015.
- [5] D. Flanagan, "JavaScript: The Definitive Guide", 6th ed., O'Reilly Media, 2011.
- [6] S. Bouaziz; Y. Wang; M. Pauly, "Online Modeling For Realtime Facial Animation", Journal ACM Transactions on Graphics, vol. 32, 2013.
- [7] J. Harlan; J. Struthers; C. Baker, "Artificial Intelligence", 1st ed., Thames and Hudson, 2009.
- [8] Rein-Lien Hsu; M. Abdel-Mottaleb; A.K. Jain, "Face detection in color images", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 290, pages 696 - 706, 2002.
- [9] M. T. Hagan; H. B. Demuth; M. H. Beale, "Neural Network Design", 2nd ed., Martin Hagan, 2014.
- [10] H. Akça; R. Alassar; V. Covachev; Z. Covacheva; E. Al-Zahrani, "Continuous-time additive Hopfield-type neural networks with impulses", Journal of Mathematical Analysis and Applications, vol. 290, pages 436-451, 2004.
- [11] H.A. Rowley; S. Baluja; T. Kanade, "Neural network-based face detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, pages 23 - 38, 1998.



Izario, D. Bachelor's at National Institute of Telecommunications/MG, Brazil - Inatel in Computer Engineering (2017), master's degree student at Computer Engineering from State University of Campinas. He is currently a freelancer in the development and planning of websites for stores and businesses and personal websites. His research interests are digital transmission and image/video/data processing, javascript, facial recognition, discrete cosine transform, data transmission and storage, defense systems and smart cities.



Izario, B. Bachelor's at National Institute of Telecommunications/MG, Brazil - Inatel in Electrical Engineering (2011), master's at Electrical Engineering from Mackenzie Presbyterian University (2015) and is working towards his doctoral degree in Sciences and Telecommunications at the State University of Campinas (Unicamp). He works as Systems Engineer at SAVIS Embraer Defense and Security. His research interests are digital transmission and processing of images/video/data, networks 4G/5G, antennas, transmission, defense systems and smart/digital cities.



Castro, D. Graduate in Electrical Engineering from the Peruvian University of Applied Sciences (UPC), Lima, Peru in 2012. He is currently working towards his doctoral degree in Sciences and Telecommunications at the State University of Campinas (Unicamp). His research interests are video and audio coding, hdr, image processing, digital television and satellite communications.



Iano, Y. is the head and founder of the Laboratory of Visual Communications since 1972. He obtained his BSc (1972), MSc (1974) and PhD (1986) in Electrical Engineering at University of Campinas, SP-Brazil. Research interests: digital signal processing (images/audio/video), digital TV, 4G (LTE) and 5G cellular networks, pattern recognition, smart cities, smart grid, internet of things.