Title	Appearance capturing and animation techniques for realistic, true-identity gaming avatars
Author(s)	Ionita, Mircea C.; Corcoran, Peter
Publication Date	2009-10-23
Publication Information	Ionita, M. C., & Corcoran, P. (2009). Appearance capturing and animation techniques for realistic, true-identity gaming avatars. Paper presented at the Games Innovations Conference, 2009. ICE-GIC 2009.
Publisher	IEEE
Link to publisher's version	http://dx.doi.org/10.1109/ICEGIC.2009.5293604
Item record	http://hdl.handle.net/10379/1435

Appearance Capturing and Animation Techniques for Realistic, True-Identity Gaming Avatars

Mircea C. Ionita and Peter Corcoran, *Senior Member*, IEEE

College of Engineering & Informatics, National University of Ireland Galway, Galway, Ireland

Abstract—**Making use of a class of 2D statistical face models known as *active appearance models* (AAMs), we investigate a technique for generating real-time animated avatars suitable for video gaming. We consider here a number of improvements of these models in order to make them more robust to varying environmental conditions, as well as to reduce the computational load such that to become adapted for real-time implementations.**

Keywords – AAM; face models; gaming avatars.

I. INTRODUCTION

The player's representation in a game, known as avatar, is most of the time fixed throughout the game. This representation is typically an artificial one, a certain character being picked from a template data set, with or without the possibility to further accessorize/personalize it.

We address in this paper the possibility to generate avatars that are realistic replicas of user's facial appearance, preserving thus his/her true identity.

In our approach we combine several computer vision/ image processing techniques in order to obtain accurate results based on versatile animated models. An image-based system is used to generate offline a statistical model of facial appearance. A set of parameters to animate the facial model are extracted from annotated image sequences of real faces. The true identity is obtained by extracting the facial appearance, based on an advanced segmentation technique.

The outline of this paper is as follows. In section II we do an overview of the facial appearance modeling system. In section III we describe our combined approach for capturing the facial appearance and for performing real-time animation. A series of tests are presented in section IV in order to evaluate the overall performance of the proposed approach. Finally, section V summarizes the conclusions of this work.

II. OVERVIEW OF THE IMAGE-BASED APPEARANCE MODELING SYSTEM

A particular useful characteristic of these statistical models is that they separate between the photometric and geometric aspects of the modeled object. For face modeling, the pixel values across the facial region are extracted subsequent to a geometrical projection that normalizes the face from the point of view of geometry; the extracted photometric information is also known as *texture*. The geometry, also known as *shape*, is represented as a vector of coordinates of a set of landmark points which sparsely define the basic facial features.

A. Model Training

The face model is statistically built using annotated sets of image sequences as shown in Fig. 1.

We employ here an improved model, which uses full color information, and which we demonstrated to increase the model fitting accuracy [3]. By means of an appropriate conversion to I1I2I3 color space [9], we are able to apply a more efficient photometric normalization, by considering the data on each color channel individually:

Fig 1: *Training image sequence.*

AAMs model both texture and geometry by means of principal component analysis (PCA). Due to the fact that PCA cannot differentiate between the various sources of facial variability, e.g. identity, illumination, expression etc., we already focused in our previous research on separating the sources of variability inside the texture space [6]. We proposed there a method for splitting the texture space of an AAM into two orthogonal texture subspaces accounting respectively for identity and illumination variation. The directional lighting model is as well incorporated to increase the robustness to lighting variations, as well as to allow for casting arbitrary shading and shadowing.

We further developed in the current work a method for identifying and separating the intrinsic identity-related shape parameters from 3D pose and expression parameters.

B. Model Fitting

The statistical model is linear both in shape and texture or appearance. However, fitting the model to a new image example represents a non-linear optimization problem. The fitting algorithm is based on minimizing the error between the query image and the model-synthesized image. It makes use of an estimate for the Jacobian of the error both with respect to the intrinsic model parameters and the so-called pose parameters – which position the model inside the image; the Jacobian matrix is estimated offline from a number of training images.

III. APPEARANCE CAPTURING AND ANIMATION

A. Capturing Appearance

Using a basic webcam, a frontal image of the player is initially captured. The face region is then roughly estimated by employing the common Viola-Jones real-time face detector [7]; our implementation is based on the OpenCV library developed by Intel [8].

The exact head region is further extracted from the background by means of an advanced iterative image segmentation method, which is also optimized to reduce the computational load. Once the head region is extracted, the player's appearance is used to generate the video-realistic avatar.

B. Template Animation

We are now able to animate the captured facial appearance by altering the appropriate shape parameters, while keeping the texture fixed. By altering these parameters one can generate different facial expressions, as well as 3D head movements within a certain range.

IV. EXPERIMENTAL RESULTS

The I1I2I3 color space enables the application of more suitable color texture normalization and, as a consequence, the convergence accuracy of the model is significantly improved. Table I shows a comparison between fitting the common RGB-GN model and our improved normalization model I1I2I3-SChN. The comparison is realized in terms of shape boundary errors (Pt-Crv) and texture errors (PTE). The tests have been realized on two different databases.

TABLE I
CONVERGENCE RESULTS ON *UNSEEN* DATABASES

Model	Success Rate [%]	Pt-Crv (Mean/Std)	PTE (Mean/Std)
(PIE) Grayscale*	92.17	5.10 / 1.66	4.28 / 1.03 *
(PIE) RGB-GN	98.26	4.98 / 1.44	7.49 / 1.98
(PIE) I1I2I3-SChN	**99.13**	**3.60 / 1.32**	**5.10 / 1.01**
(IMM) Grayscale*	63.89	4.85 / 2.12	4.90 / 3.44 *
(IMM) RGB-GN	65.28	4.55 / 2.03	9.68 / 2.81
(IMM) I1I2I3- SChN	**86.81**	**3.53 / 1.49**	**6.04 / 2.56**

*This implementation uses a grayscale model that is applied on the grayscale variant of the image data set.

From our tests we also found that the AAM-like model, enhanced with a dedicated texture subspace for directional lighting, yields a more accurate convergence on general images. This means that the initial requirement of training/fitting the model on images acquired in a controlled environment can now be extended to a broader range of common face images with no strict constraints.

As the texture space is represented as two orthogonal texture subspaces of ID variations and lighting variation, respectively, allows for the manipulation of the two types of variations independently.

V. CONCLUSIONS

In this paper we summarized initial work towards real-time generation and animation of realistic avatars for video gaming, capable of capturing the true identity of the player.

Our approach is based on a combination of several efficient and accurate techniques, such as face detection, advanced foreground extraction, and statistical modeling of appearance.

Several improvements in the use of a particular class of 2D affine models are shown to increase the robustness of the appearance modeling system.

REFERENCES

[1] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models", *Lecture Notes in Computer Science*, vol. 1407, pp. 484–, 1998

[2] G. J. Edwards, T. F. Cootes, and C. J. Taylor, "Advances in active appearance models," in *International Conference on Computer Vision (ICCV'99)*, 1999, pp. 137–142.

[3] M. Ionita, "Advances in the design of statistical face modeling techniques for face recognition", *Ph.D. Thesis*, NUI Galway, 2008.

[4] A. Bigdeli, C. Sim, M. Biglari-Abhari and B. C. Lovell, "Face Detection on Embedded Systems, Proceedings of the 3rd international conference on Embedded Software and Systems", *Springer Lecture Notes In Computer Science*; Vol. 4523, p295-308, May 2007.

[5] P. Corcoran, M.C. Ionita, I. Barcivarov, "Next generation face tracking technology using AAM techniques", *Signals, Circuits and Systems, 2007 (ISSCS 2007). International Symposium on*, Volume 1, p 1-4, 13-14 July 2007.

[6] M. Ionita, I. Bacivarov, and P. Corcoran, "Separating directional lighting variability in statistical face modelling based on texture space decompo-sition," in *Proceedings of the 15th International Conference on Digital Signal Processing (DSP'07)*, Cardiff, Wales, UK, July 2007, pp. 252–255.

[7] P. A. Viola and M. J. Jones, "Robust real-time face detection", *Int. Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.

[8] Open Source Computer Vision Library; http://www.intel.com/technology/computing/opencv/

[9] Y. Ohta, T. Kanade, and T. Sakai, "Color Information for Region Segmentation",*Ó Comput. Graphics Image Process.*, vol. 13, pp. 222-240, 1980.

[10] M. B. Stegmann and R. Larsen, "Multi-band modelling of appearance," *Image and Vision Computing*, vol. 21, no. 1, pp. 61–67, jan 2003. [Online]. Available: http://www2.imm.dtu.dk/pubdb/p.php?1421

[11] I. Bacivarov, M. Ionita, P. Corcoran, "Statistical Models of Appearance for Eye Tracking and Eye-Blink Detection and Measurement". IEEE *Transactions on Consumer Electronics*, August 2008.

[12] I. Bacivarov, M.C. Ionita, and P. Corcoran, "A Combined Approach to Feature Extraction for Mouth Characterization and Tracking", in *Signals and Systems Conference, 208. (ISSC 2008). IET Irish*, , Volume 1, p 156-161, Galway, Ireland 18-19 June 2008.

[13] D. Gonzalez-Jimenez and J. Alba-Castro, "Toward pose-invariant 2-d face recognition through point distribution models and facial symmetry," *Information Forensics and Security, IEEE Transactions on*, vol. 2, no. 3, pp. 413–429, Sept. 2007.

[14] Theobald, B.J.; Bangham, J.A.; Matthews, I.A.; Cawley, G.C., "Towards video realistic synthetic visual speech," *Acoustics, Speech, and Signal Processing, 2002. Proceedings. (ICASSP '02). IEEE International Conference on* , vol.4, no., pp. IV-3892-IV-3895 vol.4, 2002.

[15] Theobald, B.J.; Cawley, G.C.; Matthews, I.A.; Bangham, J.A., "Near-videorealistic synthetic visual speech using non-rigid appearance models," *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on* , vol.5, no., pp. V-800-3 vol.5, 6-10 April 2003.

[16] B. Theobald, Iain Matthews, Jeffrey Cohn, and S. Boker, "Real-time expression cloning using active appearance models," *Proceedings of the ACM International Conference on Multimodal Interfaces (ICMI'07)*, 2007, pp. 134 - 139.

Demos

AZ66

(Demo submission)

S. Froment, M. Ginibre, S. Mader, A. Sarafian,
A. Schwartz, D. Soriano

CNAM - ENJMIN – 121, rue de Bordeaux
16000 Angoulème, France

Jerome Dupire, Alexandre Topol

CNAM - CEDRIC – 292, rue St Martin
75003 Paris, France
{topol, dupire}@cnam.fr

Abstract — **Many researches were conducted in the affective computing field, since the precursor work of Picard in the 90's. Affective computing devices are slowly appearing with some particular digital entertainment products. We present in this paper a prototype, resulting from the association of a physiological sensing device (EKG, GSR and temperature) with an original virtual reality game.**

Keywords-component; video game, interactive experience, affective interaction, physiological sensors.

I. Introduction

Many researches were conducted in the affective computing field, since the precursor work of Picard in the 90's [1]. The physiological aspects of the emotion were studied, in order to determine which variables are interesting to be monitored. New hardware were built or improved: devices became smaller, more accurate and even wireless [2]. The sensing parts of these devices were also affected. They became more discrete, lighter, going as far as being fully integrated in the clothing [3,4]. Signal processing also contribute to this enhancement. New filters were applied, resulting in a better material to study. Analysis techniques and recognition mechanism are now able (under certain condition) to recognize a basic set of emotion from a human user and all of this led to numerous fields of application, from the health care [5] to the entertainment computing.

We present in this paper a prototype, resulting from the association of a physiological sensing device with an original virtual reality game. This article is structured as follow: the monitoring device is presented and the design of the game (or interactive experience) is described, also including implementation aspects. We finally conclude on this experience, also considering our future work.

II. Affective Computing and Video Games

Bridging physiological sensors and video game is not a novel work. One of the most famous (and older?) one is the "Relax to win" game [6]. It was developed for medical purposes, only using a GSR sensor. The aim of the game is to win a race between two dragons: the more the player relaxes, the faster the dragon runs. Many examples can also be found in very different contexts, like pervasive gaming or tangible interfaces [7].

Using the physiological information to disturb the game, adapt its contents and even manipulate the player is quite interesting, in a game validation process. For example, this previous work could evolve in a validation tool, in order to test and verify that a particular gameplay interests and/or affects the player at a level it is supposed (expected) to do. It would be very useful for a game development team to be able to test some particular aspect of the gameplay, from the early stages.
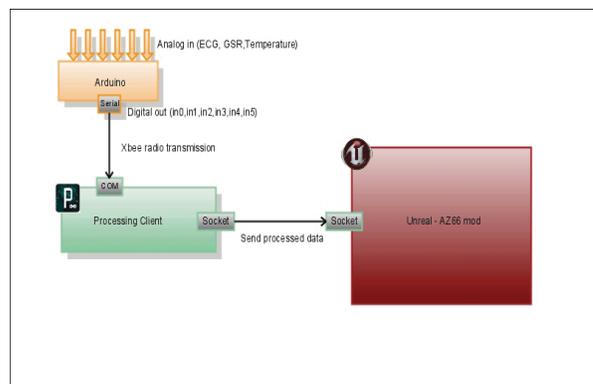

Fig.1 : Hardware and software architecture

III. Hardware and Software Architecture

The physiological monitoring apparatus was built by us [8]. Inspired from different schematics available on the web, we designed a hardware which is able to monitor the heart signal (EKG), the galvanic skin response (GSR) and the temperature (TEMP). It is declined into two versions, both wireless: one is embeddable and the other not. In this experiment, we use the second one. Both are built around an Arduino card [9], which provide a microcontroller and thus, the intelligence of the apparatus. A global view of the hardware and software architecture.

Wireless communication is based on the ZigBee protocol. We use homemade ZigBee-USB dongle, which embed a XBee module from Digi [10]. Whereas the wireless communication

is not necessary in the standalone version, it still allows the system to be isolated from the main power line and thus, to be electrically safe for the user (the sensors are battery powered). The temperature sensor is placed on one user's wrist, maintained with a wristband. The two electrodes of the EKG are also located on each wrist, while the ground one is placed on the ankle. Finally, the two electrodes of the GSR sensor are placed on two user's fingers, on the same hand.

All the data are transmitted wirelessly and integrated into the game engine (Ureal Engine 3 [11]), through a socket communication. This game engine is dedicated to create mods of the original FPS game "Unreal Tournament". AZ66 is a FPS like game, real time 3D virtual world. The interactions remains typical (through the keyboard and the mouse) and only the monitoring hardware differs from a usual playing environment.

IV. INTERACTIVE EXPERIENCE OR VIDEO GAME?

"In a close future, the player embodies a citizen, registered as AZ66. He lives in a quiet society, without any pain or violence. Responsible for the daily cleaning of his bosses' offices, he is a discrete and well disciplined employee, who never makes any trouble. One day, he is called for a medical check-up. Without any explanation, he's told to wear physiological sensors and asked to pass three tests…"

The story brings the player to the building in which the three tests will occur. Along these three tests, the player will be exposed to stressful challenges (like to disarm a bomb in a short and limited time). His/her physiological states, monitored through the specific hardware, will be taken into account. For example, an increasing stress level will penalize the player whereas a good self control will give him/her some extra time or bonuses.

The originality of this work resides in the fact that the player thinks that he/she has the control on the game, whereas not (always). Indeed, depending on the test, the player can change or not the result by his/her actions. The game makes him/her believe that he/she can succeed, with fake feedbacks for example. The ratio of the player control is variable, from none to full.

At the end of the game, the player is given a physiological report, showing how his/her body reacted during the game, linking the events to the different challenges.

V. CONCLUSION

We have presented AZ66, an interactive application influenced by the physiological states of the player. Beyond the rich interaction provided by the dynamic adaptation of the virtual environment in this game, we expect to design a game validation tool. It will allow the game development team to compare the effective effects of a specific gameplay (or a narrative aspect) with the expected ones.

REFERENCES

[1] Picard, Rosalind W. Affective computing. Ma, USA, The MIT Press (1997).

[2] Sung M. , Pentland A. Minimally-Invasive Physiological Sensing for Human-Aware Interfaces, HCI International, Las Vegas, Nevada, USA (2005).

[3] Loriga G. , Taccini N. , De Rossi D. and Paradiso R. Textile Sensing Interfaces for Cardiopulmonary Signs Monitoring. IEEE Engineering in Medicine and Biology Conference, Shanghai, China (2005).

[4] Weber W. , Glaser R. , Jung S. , Lauterbach C. , Stromberg G., and Sturm T. Electronics in Textiles: The Next Stage in Man Machine. 2nd CREST Workshop on Advanced Computing and Communicating Techniques for Wearable Information Playing Interaction. Nara, Japan, (2003).

[5] Lisetti C. , Lerouge C. Affective Computing in Tele-home Health. 37th Hawaii International Conference on System Sciences (2004).

[6] Matt J. S. Relax to Win - Treating children with anxiety problems with a biofeedback videogame. http://medialabeurope.org/research/library/Sharry_Relax_2003.pdf (2003).

[7] Magerkurth, C., Cheok, A. D., Mandryk, R. L., and Nilsen, T. Pervasive games: bringing computer entertainment back to the real world. Comput. Entertain.(2005).

[8] Dupire J. , Gal V. and Topol A. Physiological Player Sensing: New Interaction Devices for Video Games . International Conference on Entertainment Computing, Pittsburgh (2008).

[9] The Arduino project. http://www.arduino.cc/

[10] The XBee module. http://www.digi.com/products/wireless/point-multipoint/xbee-series1-module.jsp

[11] The Unreal Engine. http://www.unrealtechnology.com/