R. Amini, and C. Lisetti. HapFACS: an Open Source API/Software to Generate FACS-Based Expressions for ECA Animation and for Corpus Generation. In Proceedings of the Affective Computing and Intelligent Interaction International Conference, 2013.

HapFACS: an Open Source API/Software to Generate FACS-Based Expressions for ECAs Animation and for Corpus Generation

Reza Amini
School of Computing and Information Sciences
Florida International University
Miami, Florida, USA
Email: ramin001@fiu.edu

Christine Lisetti
School of Computing and Information Sciences
Florida International University
Miami, Florida, USA
Email: lisetti@cis.fiu.edu

***Abstract*—We present HapFACS (ver. *beta*), a new open source software and API for generating FACS-based facial expressions on 3D virtual characters that have accompanying lip-synchronized animation abilities. HapFACS has two main usage scenarios: *First*, with the HapFACS software, users can generate repertoires of realistic FACS-validated facial expressions, either as static images or as videos; *Second*, with the accessible HapFACS API, users can animate speaking virtual characters with real-time realistic facial expressions, and embed these expressive characters in their own application(s) without any prior experience in computer graphics and modeling. We describe how HapFACS (1) provides control over 49 FACS Action Units at all levels of intensity; (2) enables the animation of faces with a single AU or a composition of AUs, activated unilaterally or bilaterally; and (3) can be applied to any supported character in the underlying 3D-character system[1]. Finally, we provide details of evaluation experiments we conducted with FACS-certified scorers to validate the facial expressions generated by HapFACS.**

I. Introduction

Since their first appearance, simulated human characters have become increasingly common elements of user interfaces for a wide range of applications such as interactive learning environments [1], e-commerce [2], virtual patients [3], virtual health coaches [4], and for entertainment applications such as video games [5] and virtual worlds [6].

Contrary to simulated characters in virtual environments and video games which represent and are controlled by the user, Embodied Conversational Agents (ECA) are virtual entities of their own. ECAs are digital systems created with a physical anthropomorphic representation, and capable of having a conversation with a human counterpart, using some artificial intelligence broadly referred to as an "agent". ECAs provide computer users a natural interface through which they interact with the humans using humans' natural and innate communication modalities such as facial expressions, body language, speech, and natural language understanding. In that sense they are intended to provide an interface to computer artifacts that require quasi no computer literacy nor technical knowledge, and to thereby render computing systems more approachable and accessible to even non-technical people.

[1]http://www.haptek.com

One of the most important media in human social communication is the face [7], and the integration of its signals with other non-verbal and verbal messages is crucial for successful social exchanges [8]. However, while much is known about the appearance and human perception of facial expressions, specially emotional facial expressions, researchers and animation professionals experience difficulties when attempting to create believable animated characters that have realistic expressions.

II. Related Work

A. Facial Action Coding System (FACS)

The Facial Action Coding System (FACS) [9], [10] is a widely used system for measuring all visible facial movements. FACS describes facial activities in terms of muscle Action Units (AU). Each AU is associated with the underlying muscles that cause the movement. The AUs in FACS are based on their location on the face and the type of facial action involved. The upper-face AUs include the eyebrows, forehead, and eyelids muscles; and the lower-face AUS include muscles around the mouth and lips. For example, the inner brow raiser muscle corresponds to AU1.

The muscle groups underlying all facial movements involve 44 AUs for facial expressions and 12 AUs for gaze and head directions. The AUs can further be broken down into left and right areas of the face, which can be activated asymmetrically. Trained FACS coders can identify specific AUs present in an expression, as well as their intensities. FACS score of an expression consists of the list of AUs that produce it.

Among other uses, FACS has been used to define emotional facial expressions. Emotional FACS (EmFACS) [11] is introduced for generating emotions based on AUs. EmFACS provides subsets of AUs in FACS that are relevant to facial expressions of the universal emotions identified by Ekman [12], namely fear, anger, surprise, disgust, sadness, and happiness.

Ekman and Friesen used FACS to study people's emotions around the world. They found that people have an inherited understanding of how to create and read facial expression among all cultures. This cross-cultural universality in physical emotional reactions is highly significant to animators as well as to psychologists. Animators can use well-rigged characters capable of expressing the facial actions based on FACS to express cross-culturally readable expressions.

Since FACS is a widely accepted standard, and its learning process is sometimes a difficult job, researchers who are interested in controlling facial expressions in AU-level can use HapFACS **software** as a learning tool. In addition, researchers who are not interested in controlling facial action in AU-level can use HapFACS to generate EmFACS emotions. Also, the HapFACS **API** provides methods for developers to manipulate AUs in real-time.

B. Datasets for Facial Recognition and Generation Research

Researchers interested in human facial expressions – whether they are interested in correlating the movements of the face to the expression of emotions [9], [10], [12], [13], [14], or in considering facial displays as social signals of intent which have meaning only in social context [15], [16] – typically base their research on large databases of human facial expression images and/or videos.

Facial databases are used to enable humans to view facial displays and code them in a variety of meaningful ways, e.g. what muscles move, which muscles move together, how long are they activated, are the movements sequential or occur in parallel, what emotional or social processes are they associated with?

Several databases of human facial expressions have been developed. These databases provide standard sets of facial expression images and videos, including different emotional facial expressions and faces with specific activated AUs (e.g., Karolinska Directed Emotional Faces [17]; Pictures of Facial Affect[2]; UC Davis Set of Emotion Expressions [18]; Montreal Set of Facial Displays of Emotion[3]; Amsterdam Dynamic Facial Expression Set [19]; SmartKom video database [20]; and Belfast Naturalistic video database[4]).

These databases commonly provide data on six to nine human facial expressions of emotion. Although the databases have been used successfully for facial expression recognition and synthesis, they have common limitations such as: (1) limited number of facial movements are provided; (2) not all the possible intensities of different expressions are provided; (3) not all possible combinations of the AUs activations with different intensities are provided; (4) facial expressions generally differ between different posers in intensity and underlying facial actions, but the datasets can not provide all possible expression on one single face; (5) facial actions are not provided on faces of different ages, ethnicities, and genders; and (6) most of the provided emotional expressions are static (images).

In this article, we posit that realistic facial expressions generated and validated based on FACS, can provide valuable additional data on facial expression generation and a platform for researchers to experience with. As explained later, HapFACS aims at addressing some limitations of current repertoires available to researchers by increasing the number of facial movements that can be activated and manipulated according to FACS on a virtual character's face.

[2]http://face.paulekman.com/face/productdetail.aspx?pid=1

[3]http://www.er.uqam.ca/nobel/r24700/Labo/Labo/MSEFE.html

[4]http://sspnet.eu/2010/02/belfast-naturalistic/

C. Virtual Character Animation

Most of the current studies into facial expressions consider the FACS to be the first major step towards a scientific understanding of facial expressions. For example animation and computing projects such as MPEG-4 standard [21], the Maya-based Facial Animation Toolset[5], and iFACE [5] are informed by the FACS. However, these products either need high prior graphics knowledge to work with or do not provide AU-level manipulation of faces.

Similarly, the virtual character animation system developed by Shapiro [22] in the Institute for Creative Technologies (ICT) provides important aspects of realistic character modeling such as locomotion, facial animation, speech synthesis, reaching/grabbing, and various automated non-verbal behaviors, however, graphics prior knowledge is needed to use this software and it requires a time intensive work for creating animations.

Recently, a few software have started to become available with similar goals to the HapFACS, i.e. to generate FACS-based facial animations on virtual characters. FACSGen [23] is a software very similar to the HapFACS. The FACSGen is developed on top of the FaceGen[6] virtual character and simulates 35 AUs. However, some aspects of this software can limit its extended use: (1) it only implements 35 out of 44 effective AUs; (2) it cannot activate the bilateral AUs asymmetrically; (3) characters do not have lip-synchronization abilities (which limits its appeal for researchers interested in speaking characters); (4) strong graphics expertise is needed to embed their virtual characters in other applications, or to perform lip-synchronization; and (5) the system is not freely available and downloadable to the research community.

In a research by Helmut and Leon [24], they implemented 13 AUs on faces with soft-looking skins which can simulate wrinkles on skin. however, in addition to the same limitations as the FACSGen, their system in limited in exporting the generated facial expressions as images or videos and it is not possible to embed the character in other applications.

Villagrasa and Sanchez [25] present a 3D facial animation system named FACe!, which is able to generate different facial expressions throughout punctual and combined activation of AUs. This system is implemented on 3DStudio Max platform for building bones system, the full-rigged and skinning. The resulting head is able to activate single or combined AUs, express emotions, and display phonemes on lips. Although their system addresses some of the fore-mentioned limitations such as low number of implemented AUs and bilateral activations of AUs, still, making changes in their character's skin, age, gender, ethnicity, lighting, and enabling lip-synchronized speech requires graphics expertise. Moreover, it is not freely available to the research community either.

D. Haptek Avatar System

In research fields which employ the ECAs as the user interface, expressive and accessible 3D virtual character systems are needed that enable easy integration with low graphics knowledge. The Haptek software is an avatar system widely used in

[5]http://research.animationsinstitut.de/44.0.html

[6]http://www.facegen.com/

research projects over the years [4], [26], [27], [28]. The accessibility of the Haptek software makes it popular in projects which do not focus on graphics design but need to experiment the systems with an anthropomorphic interface. The character can be integrated easily to applications and enables to have a real-time talking character with lip synchronization based on speech synthesizers or pre-recorded sound files. However, it suffers from not having an accessible programming interface for facial expression generation which is the crucial part of emulating face-to-face interaction with nonverbal behaviors. The HapFACS addresses this issue by providing an API to control facial expressions in real-time based on the FACS and control emotions based on the EmFACS.

III. HapFACS

HapFACS is an open-source[7] stand-alone software and API implemented in the C# language. It uses the characters created in the commercial software PeoplePutty[8] from the Haptek company. HapFACS is able to map the FACS AUs to the facial, head and eye variables of the Haptek characters. The HapFACS currently includes more that 120 characters and hair styles, however, to generate more faces users can use the PeoplePutty software.

For mapping, we match each AU introduced in the FACS manual [10] to a combination of the Haptek registers and switches. The Haptek original registers and switches are not based on the FACS, therefore, to generate facial expressions based on the FACS, we followed four steps:

1) we explored all the Haptek facial, head, and eye registers and switches which manipulate the facial, head, and eye movements. The Haptek registers/switches used in our implementation include 6 for head movements, 4 for eye movements, 8 for upper-face movements, and 21 for lower-face.
2) for each AU, we found a set of the registers/switches whose activation simulates the same movements in the face as actual activation of that AU. For example, for AU4, the used registers are *MidBrowUD*, *LBrowUD*, *RBrowUD*, and *eyes_sad*.
3) the FACS manual [10] introduces 6 intensity levels for each AU: (0) not active, (A) trace, (B) slight, (C) marked or pronounced, (D) sever or extreme, and (E) maximum. Assuming that *E* is activating the AU with 100% intensity, we assign 85% to *D*, 55% to *C*, 30% to *B*, 15% to *A*, and 0% to *0*.
4) for each AU intensity level, we apply the same percentages to the intensity range of the Haptek registers. Although the intensity values in the FACS are represented as ranges, in HapFACS we represented them as discrete values based on our empirical approximations. For example the maximum register values for AU4 are selected as *MidBrowUD = 2.00*, *LBrowUD = 0.60*, *RBrowUD = 0.60*, and *eyes_sad = 1.25*.

When a HapFACS user activates an AU with a specific intensity, the set of registers found in step 2 are activated with the intensities found in step 4 for that AU. A snapshot of the HapFACS software interface is shown in Figure 1.

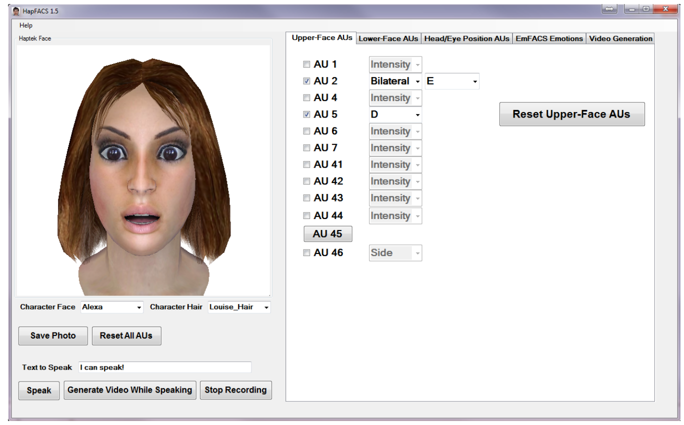


Fig. 1: HapFACS interface for generating static outputs of AUs.

When the AUs are activated, user can take a photo of the generated face using a provided "Save Photo" button. In addition to the AU and its intensity, for 19 AUs, the user can also select to activate the AU unilaterally. For example, user can activate the AU2-left (left outer brow raiser) unilaterally with intensity *E* and AU2-right with intensity *B*. For video generation, users need to provide the AU, side (i.e., left or right), starting intensity, ending intensity, starting time, and ending time of the AU activation. The intensity will change in a linear way. Users can activate different AUs in parallel during the video generation. In addition, users can generate EmFACS-based emotional facial expressions with different intensities for 9 emotions namely happiness, sadness, surprise, anger, fear, disgust, contempt, embarrassment, and pride. Figure 2 shows sample activated AUs. The users can also change the avatar and its hair style. HapFACS can import characters of different ages, ethnicities, and genders generated in PeoplePutty software.

The HapFACS enables the users to re-produce the same expressions again, or to activate the same AUs in another Haptek-based software by the HapFACS API. The API returns the *hypertext* which can be imported to any Haptek character and activate the same AUs on any face.

HapFACS provides various possibilities and controls over the characters' facial expressions such as: (1) controlling 49 AUs (12 upper-face, 21 lower-face, and 16 head/eye position); (2) activating individual and composition of AUs with different intensities; (3) activating the AUs bilaterally and unilaterally; (4) generating faces with different lightings, backgrounds, and observer's vantage points; (5) importing new faces with different facial skin textures, ages, genders, and ethnicities generated in the PeoplePutty; (6) generating reproducible, realistic, 3D, static and dynamic (video) outputs; (7) generating the Haptek hyper-texts provided in a C# API to enable reproduction of the HapFACS facial expressions in other applications with embedded Haptek avatars; (8) providing an easy-to-use interface which requires no prior computer or FACS expertise to enable researchers in various disciplines to easily take advantage of the HapFACS.

The expressiveness, accessibility, easy integration with

[7] [reference to website is removed for blind review.]
[8] http://www.haptek.com/products/peopleputty/

(a) AU1 (b) AU2-Left (c) AU22 (d) AU61

Fig. 2: Individual AUs activated in HapFACS.

other application, and lip-synchronized characters of the Haptek avatar system makes it a popular one among others. The HapFACS adds the possibility of creating and recording valid facial expressions to the Haptek avatar system. Therefore, HapFACS is useful in different research fields such as simulation of facial expressions and generation of facial repertoire in **psychology and emotion theory**, animation of ECAs in the **affective computing**, avatar-based **Human-Computer Interaction**, **animation**, **teleconference**, as well as generation of facial analysis data for the **FACS training**.

IV. Validation

We performed four experiments to validate the AUs and emotional facial expressions generated by the HapFACS.

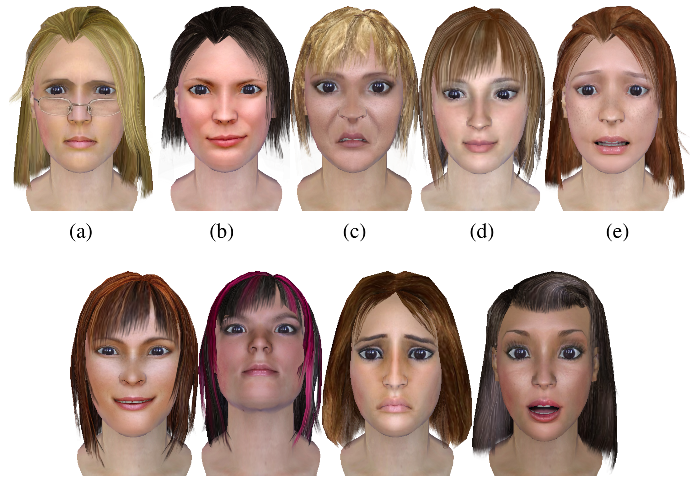
A. Experiment 1: Validation of AUs

In this experiment, we examine whether the AUs generated by the HapFACS correspond to the same descriptions provided in FACS manual [10]. We generated 49 two-second videos each of which shows a single AU activation. In each video, the AU intensity is changed from *0* (i.e., neutral) to *E* (i.e., highly intense) linearly in one second and returned back from *E* to *0* in one second. Videos are generated with different male and female faces in different ages and ethnicities (e.g., White, Black, Asian), with frame rate of 30 frames per second, and size of 480×480 pixels.

Videos were named with numbers and presented to three FACS-certified coders in a random order. Each coder coded the presence/absence of the AUs in the videos. FACS coders were not asked to score the activation intensity of the AUs in videos, because it is shown that judgments over the intensity show poor inter-rater agreement [29]. Results show that **39** AUs are recognized correctly by **all** three coders. The 10 AUs that are listed in Table I had recognition rates of less than 100%. The average recognition rate of all the AUs was **90.48%**.

TABLE I: AUs with recognition rate of less than 100%.

AU	Recognition Rate	AU	Recognition Rate
10	66.67%	**16**	33.33%
11	66.67%	**20**	66.67%
12	66.67%	**23**	33.33%
13	66.67%	**25**	33.33%
14	66.67%	**28**	33.33%



(a) (b) (c) (d) (e) (f) (g) (h) (i)

Fig. 3: HapFACS emotions (a) Angry (4E+5B+7B+23A+24B), (b) Contempt (12A+14RE), (c) Disgust (10E+15E+16E), (d) Embarrassment (12B+54C+62B+64C), (e) Fear (1E+2C+4D+5B+20C+26D), (f) Happiness (6B+12E+25B), (g) Pride (12B+53C+58B+64E), (h) Sadness (1E+4C+15E), (i) Surprise (1E+2E+5B+26E)

B. Experiment 2: Validation of Emotion Inferences Drawn from HapFACS Expressions (Color Version)

This experiment examines whether the images of the emotional facial expressions generated by HapFACS convey the same emotional meaning as that of human expressions. Nine emotions namely angry, contempt, embarrassment, disgust, happiness, sadness, pride, sadness, and surprise are portrayed in this experiment based on the EmFACS. Emotions are displayed at the high intensity, with 9 Caucasian female faces. Images are of size 485×480 pixels. The AU combination used to generate each emotion, and faces generated by HapFACS are shown in Figure 3. In this experiment we asked regular people (i.e., not FACS-coders) to recognize the emotions shown in the images of the emotional facial expressions.

Subjects were recruited in Amazon Mechanical Turk (AMT). From the total of 123 subject we filtered 50 of them. These 50 subjects used a web-based interface to tag each image with an emotion from a list of 9 emotions. AMT is a crowd-sourcing Internet marketplace in which ”Requesters” post tasks (e.g., choosing the best emotion label for a face image), and ”Workers” complete them for a monetary payment set by the Requester. However, not all of the Workers put enough time into the assigned jobs and try to finish them as soon as possible to get their money (i.e., possibility of fake answer). So, we use some rules for filtering their jobs to prevent random and fake tagging (e.g., they may select one tag for all of the images). From the three rules we defined below, if a subject dissatisfy two out of three rules, it shows that the answers are selected randomly, so we can exclude that job.

1) asking verifiable questions from the subjects to verify the legitimacy of their other answers (e.g., we excluded a subject if he/she gives multiple obvious conflicting answers).

2) obtaining multiple responses from a user to verify his/her response validity (e.g., we excluded a subject if he/she tags all/multiple images with a same tag).
3) excluding subjects who did not answer any of the questions correctly to prevent random tagging of the images.

Table II shows the emotion recognitions rates by the subjects with the average of **86.22%** for color-images.

TABLE II: Results of experiment 2.

Emotions	Angry	Contempt	Disgust	Embarrass	Fear	Happiness	Pride	Sadness	Surprise
Angry	84%	8%	4%	0%	2%	0%	2%	0%	0%
Contempt	2%	68%	2%	8%	2%	0%	18%	0%	0%
Disgust	6%	6%	84%	2%	0%	0%	0%	0%	2%
Embarrass	0%	8%	0%	80%	4%	4%	4%	0%	0%
Fear	0%	2%	4%	4%	88%	0%	0%	2%	0%
Happiness	0%	0%	0%	2%	0%	98%	0%	0%	0%
Pride	6%	8%	6%	0%	0%	0%	78%	2%	0%
Sadness	0%	0%	0%	0%	4%	0%	0%	96%	0%
Surprise	0%	0%	0%	0%	0%	0%	0%	0%	100%

C. Experiment 3: Validation of Emotion Inferences Drawn from HapFACS Expressions (Gray-Scale Version)

This experiment aims to determine whether the image quality and image color biases the participants' recognition of the portrayed emotions. We replicated the second experiment with gray-scale lower quality images. Each face in Experiment 2 was converted to gray-scale and resized to a 340×336 pixels image. We used AMT to recruit the subjects. Totally, 162 subjects participated in this experiment who where filtered using the same techniques mentioned in Section IV-B to 50 subjects. These 50 subjects used the same web-based interface to tag each image with an emotion from a list of 9 emotions. Table III shows the emotion recognition ratings with the average of **87.11%** for gray-scale images.

TABLE III: Results of experiment 3.

Emotions	Angry	Contempt	Disgust	Embarrass	Fear	Happiness	Pride	Sadness	Surprise
Angry	80%	4%	0%	12%	0%	0%	4%	0%	0%
Contempt	2%	68%	2%	10%	2%	2%	14%	0%	0%
Disgust	4%	2%	94%	0%	0%	0%	0%	0%	0%
Embarrass	6%	18%	0%	70%	0%	2%	4%	0%	0%
Fear	0%	0%	2%	0%	98%	0%	0%	0%	0%
Happiness	0%	2%	0%	2%	0%	96%	0%	0%	0%
Pride	8%	6%	2%	6%	0%	0%	78%	0%	0%
Sadness	0%	0%	0%	0%	0%	0%	0%	100%	0%
Surprise	0%	0%	0%	0%	0%	0%	0%	0%	100%

D. Experiment 4: Validation of Emotion Inferences Drawn from HapFACS Videos

As discussed in Section II-D, Haptek characters are capable of speaking with lip-synchronization. So, in many applications (e.g., On-Demand Virtual Health Counselor [4]) Haptek characters are used to deliver information both verbally and non-verbally. This experiment evaluates the same emotional expressions as in previous experiments while the character is speaking with an emotionally neutral voice. Neutral voice is selected to prevent biasing the subjects.

We generated the same 9 EmFACS emotions on characters' faces. Then, we made the character read a sentence with a neutral voice (e.g., "Can you recognize my emotional facial expression while I am talking with a neutral voice?"). We recorded the videos of the speaking character with frame rate of 30 frames per second, and size of 480×480 pixels. We recruited 122 subjects from the AMT and filtered 50 subjects using the techniques mentioned in Section IV-B. The subjects watched the videos in a web-based interface and selected the expressed emotion in each video from a list of 9 emotions. Table IV shows the recognition ratings with the average recognition rate of **77.11%**.

TABLE IV: Results of experiment 4.

Emotions	Angry	Contempt	Disgust	Embarrass	Fear	Happiness	Pride	Sadness	Surprise
Angry	64%	22%	8%	0%	4%	0%	0%	2%	0%
Contempt	2%	50%	2%	2%	0%	8%	38%	0%	0%
Disgust	4%	2%	94%	0%	0%	0%	0%	0%	0%
Embarrass	0%	6%	0%	84%	0%	6%	4%	0%	0%
Fear	0%	2%	2%	12%	80%	0%	0%	4%	0%
Happiness	0%	4%	0%	0%	2%	86%	8%	0%	0%
Pride	24%	24%	0%	4%	0%	0%	44%	0%	4%
Sadness	0%	0%	0%	0%	6%	0%	0%	94%	0%
Surprise	0%	0%	0%	0%	0%	2%	0%	0%	98%

V. DISCUSSION

The 49 generated AUs in the HapFACS were recognized correctly with the average recognition rate of **90.48%** which validates the correctness of the AU generation and expression by HapFACS. The AUs that are not perfectly recognized are mostly lower-face AUs. Since lower-face AUs are all located in a close proximity around the mouth, it is hard to depict the subtle lower-face AUs and differentiate between them with the limited number of Haptek registers in this area. Results of experiment 1 support this hypothesis. However, improvement is needed in future versions of Haptek for lower-face AUs.

Results of the second and the third experiments show that, images of emotional facial expression generated by HapFACS are validated with high average recognition rated of **86.22%** for the color-images and **87.11%** for the gray-scale ones, which proves that quality and colorfulness of the images do not bias the subject emotion recognition. As shown in Table II, *contempt* emotion is sometimes perceived as *pride*, which can be because of the connection between these two emotions, i.e., people who feel *contempt* may feel *pride* at the same time too, and vise versa. This can cause subjects to tag these two emotions interchangeably.

Results of forth experiment indicate that generated dynamic emotional expressions are validated with the average recognition rate of **77.11%**. While the character is speaking and expressing an emotion, many of the lower-face AUs may be activated for speaking. For emotions such as *contempt* that is expressed only by lower-face AUs, speaking can decrease the recognition rates. Also, the discussion about the related emotions such as *contempt* and *pride* is still valid in this experiment.

We need to perform more validation experiment in which different combinations of the AUs are also validated. We will ask more FACS-coders to score some common and some difficult combinations of the AUs. Therefore, at the moment, we introduce the HapFACS as a *beta* version until we perform more validations.

VI. Conclusion and Future Work

We presented HapFACS, a new open source API and software for generating FACS-based facial expressions on 3D virtual characters that have lip-synchronized speech abilities. Using the HapFACS **software**, users can generate repertoires of realistic FACS-validated facial expressions, either as images or as videos. HapFACS (1) provides control over 49 AUs at all levels of intensity; (2) enables the animation of faces with a single AU or a composition of AUs, activated unilaterally or bilaterally; and (3) can be applied to any supported character in the Haptek 3D-character system with different ethnicities, genders, ages, and skin textures. We conducted four evaluation experiments to validate the facial expressions generated by HapFACS. Results show that AUs generated in HapFACS are highly validated by certified FACS coders. Moreover, the images and videos of emotional facial expressions generated by HapFACS are validated by the subjects.

In future versions of the HapFACS, we aim to (1) improve the expressiveness of the AUs that were reported as not very expressive (see Table I); (2) ask more FACS-coders to score some common and some difficult combinations of the AUs, therefore, we introduce the HapFACS as a beta version until we perform more validations; (3) provide non-linear changing of the intensity for video generation; (4) use 0-100% continuous intensity instead of 6-level discrete intensity; and (5) evaluate the lip-synchronization performance of the characters.

References

[1] A. Sarrafzadeh, S. Alexander, F. Dadgostar, C. Fan, and A. Bigdeli, "See Me, Teach Me: Facial Expression and Gesture Recognition for Intelligent Tutoring Systems," *2006 Innovations in Information Technology*, pp. 1–5, Nov. 2006.

[2] T. Bosse, G. F. Siddiqui, and J. Treur, "An Intelligent Virtual Agent to Increase Involvement in Financial Services," in *IVA'10 Proc. of the 10th int'l conf. on Intelligent Airtual Agents*. Springer-Verlag Berlin, Heidelberg, 2010, pp. 378–384.

[3] P. Kenny, T. Parsons, J. Gratch, A. Leuski, and A. Rizzo, "Virtual patients for clinical therapist skills training," in *Intelligent Virtual Agents (IVA'07)*, C. P. et Al., Ed. Springer-Verlag Berlin Heidelberg, 2007, pp. 197–210.

[4] C. L. Lisetti, U. Yasavur, C. D. Leon, R. Amini, and N. Rishe, "Building an On-demand Avatar-based Health Intervention for Behavior Change," in *Proc. of FLAIRS'2012 Association for the Advancement of Artificial Intelligence (www.aaai.org)*, no. Mi, Miami, FL, US, 2012.

[5] A. Arya, S. DiPaola, L. Jefferies, and J. T. Enns, "Socially communicative characters for interactive applications," in *14th Int'l Conf. on Computer Graphics, Visualization and Computer Vision (WSCG2006)*. Plzen, Czech Republic: UNION Agency Science Press, 2006.

[6] S. C. Marsella, J. Gratch, and J. Rickel, "Expressive Behaviors for Virtual Worlds," in *Life-Like Characters, Tools, Affective Functions, and Applications*, M. Prendinger, H., Ishizuka, Ed. Springer, Heidelberg, 2003.

[7] N. Ambady and R. Rosenthal, "Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis." *Psychological Bulletin*, vol. 111, no. 2, pp. 256–274, 1992.

[8] J. Allwood, "Bodily communication dimensions of expression and content," *Multimodality in language and speech systems*, vol. 7, no. 26, pp. 1–15, 2002.

[9] P. Ekman and W. V. Freisen, *Facial Action Coding System: A Technique for the Measurement of Facial Movement*, C. P. Press, Ed. Consulting Psychologists Press, 1978.

[10] P. Ekman, W. V. Freisen, and J. C. Hager, *Facial Action Coding System*, 2nd ed. Salt Lake City, UT: Research Nexus eBook, 2002, vol. 160, no. 4.

[11] W. V. Friesen and P. Ekman, "EMFACS-7: Emotional Facial Action Coding System," *Unpublished manuscript, University of California at San Francisco*, 1983.

[12] P. Ekman, R. W. Levenson, and W. V. Freisen, "Autonomic Nervous System Activity Distinguishes among Emotions," *Science*, vol. 221, no. 4616, pp. 1208–1210, 1983.

[13] P. Ekman and W. V. Freisen, "Detecting Deception From The Body Or Face," *Jrnl. of Personality and Social Psychology*, vol. 29, no. 3, pp. 288–298, 1974.

[14] K. R. Scherer, *Appraisal processes in emotion: Theory, methods, research*, T. Scherer, Klaus R. and Schorr, Angela and Johnstone, Ed. New York, NY, US: Oxford University Press, 2001.

[15] N. Chovil, "Discourse-Oriented Facial Displays in Conversation," *Research on Language and Social Interaction*, vol. 25, pp. 163–194, 1991.

[16] A. Fridlund, *Human Facial Expression: An Evolutionary View*. San Diego: Academic Press, 1994.

[17] D. Lundqvist, A. Flykt, and A. Öhman, "The Karolinska Directed Emotional Faces - KDEF," 1998.

[18] J. L. Tracy, R. W. Robins, and R. A. Schriber, "Development of a FACS-verified set of basic and self-conscious emotion expressions." *Emotion Washington Dc*, vol. 9, no. 4, pp. 554–559, 2009.

[19] J. Van Der Schalk, S. T. Hawk, A. H. Fischer, and B. Doosje, "Moving faces, looking places: Validation of the Amsterdam Dynamic Facial Expression Set (ADFES)." *Emotion Washington Dc*, vol. 11, no. 4, pp. 907–920, 2011.

[20] F. Schiel, S. Steininger, and U. Türk, "The SmartKom Multimodal Corpus at BAS," in *Proc. of the 3rd conf. on language resources and evaluation LREC02*, no. 34, 2002, pp. 200–206.

[21] S. Battista, F. Casalino, and C. Lande, "MPEG-4: a multimedia standard for the third millennium, Part 1." in *IEEE Multimedia*. IEEE, 2000, vol. 7, no. 1, pp. 74–83.

[22] A. Shapiro, "Building a character animation system," *Motion in Games(MIG'11)*, vol. LNCS, no. 7060, pp. 98–109, 2011.

[23] E. G. Krumhuber, L. Tamarit, E. B. Roesch, and K. R. Scherer, "FACSGen 2.0 animation software: Generating three-dimensional FACS-valid facial expressions for emotion research." *Emotion*, vol. 12, no. 2, pp. 351–363, Jan. 2012.

[24] L. Beutl, "A simulation for the creation of soft-looking, realistic facial expressions," Master Thesis, University of Wien, 2011.

[25] S. Villagrasa and A. S. Sánchez, "Face! 3d facial animation system based on facs," in *IV Iberoamerican Symp. in Computer Graphics (SIACG 2009)*, O. Rodríguez, F. Serón, R. Joan-Arinyo, J. Madeiras, J. Rodríguez, and E. Coto, Eds., Isla Margarita, 2009, pp. 203–209.

[26] A. N. Fellner, G. Matthews, K. D. Shockley, J. S. Warm, M. Zeidner, L. Karlov, and R. D. Roberts, "Using emotional cues in a discrimination learning task: Effects of trait emotional intelligence and affective state," *Jrnl. of Research in Personality*, vol. 46, no. 3, pp. 239–247, Jun. 2012.

[27] J. Szymaski and W. Duch, "Information retrieval with semantic memory model," *Cognitive Systems Research*, vol. 14, no. 1, pp. 84–100, Apr. 2012.

[28] M. Turunen, J. Hakulinen, O. Stå hl, B. Gambäck, P. Hansen, M. C. Rodríguez Gancedo, R. S. de la Cámara, C. Smith, D. Charlton, and M. Cavazza, "Multimodal and mobile conversational Health and Fitness Companions," *Computer Speech & Language*, vol. 25, no. 2, pp. 192–209, Apr. 2011.

[29] M. Sayette, J. Cohn, and J. Wertz, "A psychometric evaluation of the facial action coding system for assessing spontaneous expression," *Jrnl. of Nonverbal Behavior*, vol. 25, no. 3, pp. 167–185, 2001.