HapFACS 3.0: FACS-Based Facial Expression Generator for 3D Speaking Virtual Characters

Reza Amini, Christine Lisetti, and Guido Ruiz

Abstract—With the growing number of researchers interested in modeling the inner workings of affective social intelligence, the need for tools to easily model its associated expressions has emerged. The goal of this article is two-fold: 1) we describe HapFACS, a free software and API that we developed to provide the affective computing community with a resource that produces static and dynamic facial expressions for three-dimensional speaking characters; and 2) we discuss results of multiple experiments that we conducted in order to scientifically validate our facial expressions and head animations in terms of the widely accepted Facial Action Coding System (FACS) standard, and its Action Units (AU). The result is that users, without any 3D-modeling nor computer graphics expertise, can animate speaking virtual characters with FACS-based realistic facial expression animations, and embed these expressive characters in their own application(s). The HapFACS software and API can also be used for generating repertoires of realistic FACS-validated facial expressions, useful for testing emotion expression generation theories.

Index Terms—Facial Action Coding System (FACS), FACS-based facial expression generation, 3D facial animation

1 Introduction

One of the most important media in human social communication is the face [1], and the integration of its signals with other non-verbal and verbal messages is crucial for successful social exchanges [2]. However, while much is known about the appearance and human perception of facial expressions, especially emotional facial expressions, emotion researchers still have open questions about the *dynamics* of human facial expression generation and their perception [3].

There are therefore advantages to have software such as HapFACS that can emulate facial expression generation via 3D animated characters: 1) to animate intelligent virtual agents (IVA) that can naturally portray human-like expressions when they interact with humans; and 2) to develop and test emotion theories of (human) dynamic facial expression generation.

Indeed, IVAs—which simulate humans' innate communication modalities such as facial expressions, body language, speech, and natural language understanding to engage their human counterparts—have emerged as a new type of computer interfaces for a wide range of applications, e.g., interactive learning [4], e-commerce [5], virtual patients [6], virtual health coaches [7], video games [8] and virtual worlds [9].

IVAs therefore need to be able to portray appropriate levels of social realism, e.g. portray believable facial expressions, and gestures. However, many IVA researchers interested in modeling social intelligence do not have the graphics expertise for the difficult task of generating and animating 3D models in order to showcase the embodiment of their models of social intelligence. Similarly, psychologists working on facial expression generation rely on the analysis of large corpora of videos and images of human expressions, but do not have the means to test their theories in a systematic fashion.

Because it is difficult to animate 3D characters, many researchers buy and use third-party software. To date, there are a handful of systems that provide the ability to manipulate facial expressions on 3D characters in terms of Action Units (AU), most of them are not freely available to the research community and/or require computer graphics expertise, with characters often difficult to integrate with one's system.

We developed HapFACS, a free software and API, based on the Haptek 3D-character platform running on the free Haptek Player[1], to address the needs of IVA researchers working on 3D speaking characters, and the needs of psychologists working on facial expression generation. HapFACS provides the ability to manipulate the activation—in parallel or sequentially—of combinations of the smallest groups of facial muscles capable of moving independently (referred to as Action Units), which can help the development of facial expression generation theories, as well as the creation of believable speaking IVAs.

Our first version of HapFACS was very well received by affective computing researchers [7], [10], [11], [12], [13], [14], [15], [16], so in this article, we describe the expanded version of our software functionalities, along with the extensive evaluation of images and videos of single and combinations of AUs that we conducted to validate their fidelity to FACS.

• The authors are with the School of Computing and Information Sciences, Florida International University, Miami, FL 33199.
E-mail: {ramin001, gruiz044}@fiu.edu, lisetti@cis.fiu.edu.

Manuscript received 18 Sept. 2014; revised 21 Apr. 2015; accepted 1 May 2015. Date of publication 12 May 2015; date of current version 1 Dec. 2015.
Recommended for acceptance by C. Pelachaud.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TAFFC.2015.2432794

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

2 Related Work

2.1 Facial Action Coding System (FACS)

The Facial Action Coding System [17] is a widely used facial coding system for measuring all visible facial movements in terms of the smallest possible muscle Action Units causing movement. As shown in Fig. 1, AUs are grouped 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, e.g., the inner brow raiser muscle corresponds to AU1; the "lower-face" AUs include muscles around the mouth and lips; and the "head and eye movement" AUs include the neck muscles which move the head, whereas the eye muscles move the gaze direction.

AUs act as multi-level "switches", which can create custom expressions depending on which AUs are activated/deactivated at a given time. Since not all expressions require the farthest reach of a muscle, intensity levels are used to discuss subtle (i.e., less intense) facial movements. Intensities are annotated from "0" to "E", with "0" for a neutral face without any activated AUs, "A" for the weakest trace of an AU, and "E" for the maximum intensity.

The muscle groups underlying all facial movements form 44 AUs for facial expressions and 12 AUs for head and gaze directions.

Emotional FACS (EmFACS) [18] is a subset of FACS, focused on facial expressions of emotions. EmFACS provides subsets of AUs used to generate Ekman's [19] 6 universal emotions: fear, anger, surprise, disgust, sadness, and happiness; other emotional facial expressions such as contempt, pride, and embarrassment can also be depicted by AU combinations.

2.2 Facial Expression Datasets

Researchers interested in human facial expressions typically base their research on large databases of human facial expression images and/or videos. These researchers include both (1) the ones interested in correlating facial movements to the expression of emotions [3], [17], [19], [20], [21], and (2) the ones considering facial displays as social signals [22], [23].

Several databases of human facial expressions have been developed: [24], [25], [26], [27], [28], [29], Pictures of Facial Affect[2], Montreal Set of Facial Displays of Emotion[3], and Belfast Naturalistic video database[4]. These databases provide standard sets of facial expression images and videos, including different emotional facial expressions and faces with specific activated AUs.

Although these databases have been used successfully for facial expression recognition and synthesis, they have common limitations such as: (1) only a limited number of facial movements are provided; (2) all the possible intensities of different expressions are not provided; (3) all combinations of AUs activation with different intensities are not provided for each face (i.e. it is difficult for a human actor to generate all combinations of 44 AUs); (4) because it is difficult for human posers to display exactly the same AU activation intensity, datasets are not always consistent across subjects [30], [31] (furthermore, since emotions and their expressions can be triggered unconsciously, some combinations of AU activations may be physiologically possible but difficult to generate on-demand); (5) most of the provided emotional expressions are static (images) which can limit their usefulness to study gradual changes of facial muscle movements; and/or (6) although a few databases provide images and videos of individual AUs and of combination of AUs, most of the databases provide data on six to nine human facial expressions of emotions only.

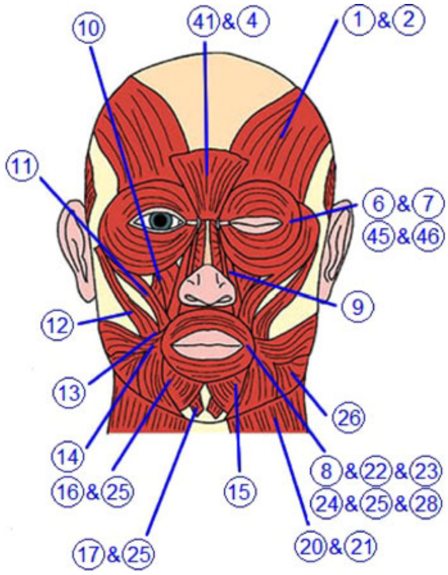

Fig. 1. Sample AUs of the FACS.

2.3 Virtual Character Animation

FACS has been used for animating virtual characters' facial expressions. Smartbody [32], [33], for example, provides realistic character modeling such as locomotion, facial animation (11 AUs), speech synthesis, reaching/grabbing, and various automated non-verbal behaviors. It is a powerful tool, however some prior knowledge of graphics is needed to use, create, and integrate animations in one's application.

FACSGen [30] is a software developed on top of the FACSGen[5] third party software and simulates 35 AUs on 3D virtual characters. Whereas FACSGen faces are quite realistic, characters do not have lip-synchronization abilities (which limits FACSGen appeal for IVA researchers interested in speaking characters), nor do they have bodies (so graphics expertise is needed to combine faces with bodies). Because it is the software closest to HapFACS, we compare it with HapFACS on some aspects in Section 4.3.2.

Villagrasa and Sánchez [34] presented a 3D facial animation system named FACe!, which is able to generate different facial expressions throughout punctual and combined activation of AUs. The virtual model is able to activate single or combined AUs, express emotions, and display phonemes on lips. FACe! simulates a total of 66 movements and AUs, and can express four emotions (happy, fear, sad, anger). However, endowing these characters with lip-synchronized speech requires graphics expertise, and it is not free.

2. www.paulekman.com/product/pictures-of-facial-affect-pofa
3. www.er.uqam.ca/nobel/r24700/Labo/Labo/MSEFE.html
4. http://sspnet.eu/2010/02/belfast-naturalistic/
5. http://www.facegen.com/

Alfred [35] is a facial animation system which uses a slider-based GUI, a game-pad, and a data glove with which the user can manipulate facial expressions. Alfred uses the 23 FACS AUs for the description and creation of facial expressions, and supports lip synchronization.

Greta [36], [37] is a virtual character which displays and animates facial expressions based on Facial Animation Parameters (FAPs) and MPEG-4 standard [38]. Greta can also simulate skin wrinkles. However, since it is not based on FACS, it does not provide simulations of AUs and asymmetric animations.

Some systems have started to use motion capture techniques to map marked points on an actress' face to the vertices of her 3D face model. Wojdel and Rothkrantz [39] presented a parametric approach to the generation of FACS-based facial expressions. They used 38 control markers on one side of a human subject's face as well as positions of facial features such as mouth-contour, eye-contour and eye-brows and took frontal pictures of a face when single AUs were activated. The authors found mathematical functions of the marker movements for each AU activation, and implemented 32 symmetric AUs. They evaluated the AU generation accuracy with 25 subjects on Ekman's 6 universal expressions mentioned above and reported a 64 percent recognition rate. However, this software does not have lip-synchronization abilities.

Digital Emily [40], a photo-real 3D character, was generated by filming an actress while she spoke and by capturing the motion of the actress' facial expressions. The actress posed for thirty-three different facial expressions based loosely on FACS. A semi-automatic video-based facial animation system was then used to animate the 3D face rig. However, Digital Emily was rendered offline, and only involved the front of the face. Alexander et al. generated a new virtual character called Digital Ira [41] using a similar approach to [40]. Digital Ira is a real-time photo-real digital human character which can be seen from any viewpoint, in any lighting, and can perform realistically from video performance capture even in a tight closeup. However, this virtual character is not available to the research community, and integration of the head model to a real application needs graphics expertise.

2.4 Haptek Avatar System

The Haptek[6] software, developed by Chris Shaw, is a light weight avatar system popular for research groups working on speaking IVAs [5], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], among others. Unlike high-resolution 3D characters seen in video games and digital animation movies (which require prescripting movements and many costly graphic artists and animators to render each facial expression), Haptek offers free programmable 3D characters that can speak with synchronized lip movements.

Haptek characters come with lip-synch, with head-only, torso, and full-body, and can be animated with gestures. They can easily be integrated in real-time applications with speech synthesizers or pre-recorded sound files. However, the default Haptek functionality does not enable FACS-based facial expression generation.

HapFACS, discussed next, provides an API to control FACS-validated AUs on Haptek characters' faces in real-time, and to generate facial expressions of emotions based on EmFACS [18].

3 HapFACS

HapFACS[7] is a free stand-alone software and API implemented in the C# language. It uses virtual characters rendered in the free HaptekPlayer software, and simulates FACS AUs on these characters. Currently, HapFACS includes more than 165 characters and hair styles, and users can use Haptek PeoplePutty software to create new characters and import them to HapFACS.

3.1 HapFACS Functionalities

HapFACS enables users to: (1) control 49 AUs (12 upper-face, 21 lower-face, and 16 head/eye position) of characters' faces; (2) activate individual AUs and AU combinations with different intensities; (3) activate AUs bilaterally and unilaterally; (4) generate EmFACS emotions with different intensities; (5) generate reproducible, realistic, 3D, static and dynamic (video) outputs; (6) generate Haptek hypertexts provided by a C# API to enable reproduction of the HapFACS facial expressions in other applications with embedded Haptek avatars[8].

For image generation, when a HapFACS user activates an AU with a specific intensity, its corresponding set of registers (described in step 2 earlier) is activated with the intensities described in step 4. In addition, users can select to activate the AU unilaterally for 19 AUs. When the AUs are activated, users can take and save a photo of the generated face.

For video generation, users need to provide: the AU, side (i.e., left/right or bilateral), starting intensity, ending intensity, starting time, and ending time of the AU activation. HapFACS currently changes the intensity linearly from the start intensity to the end intensity and generates a video of the resulting expression (non-linear activation is planned for a future version). Users can activate different AUs in parallel during the video generation and select overlapping activation times for the AUs.

In addition, users can generate EmFACS-based emotional facial expressions with different intensities for nine emotions: happiness, sadness, surprise, anger, fear, disgust, contempt, embarrassment, and pride [18].

3.2 Registers, Switches and Action Units

For simulating the AUs on the characters' faces, we match each AU introduced in the FACS manual [17] to a

6. http://www.haptek.com

7. We have an agreement with Haptek to be able to distribute HapFACS source code under a free non-exclusive license only for academic, research or non-profit centers and only for personal and non-commercial purposes as per the license terms provided at http://ascl.cis.fiu.edu/hapfacs.html. HapFACS requires Haptek Player available at www.haptek.com which is also free for non-commercial use, per its license agreement.

8. Features such as modifiable background, lighting, and skin texture are provided by the Haptek API, and HapFACS user-friendly interface enables non-experts to utilize all these functionalities without having to learn Haptek C++ or JavaScript APIs.

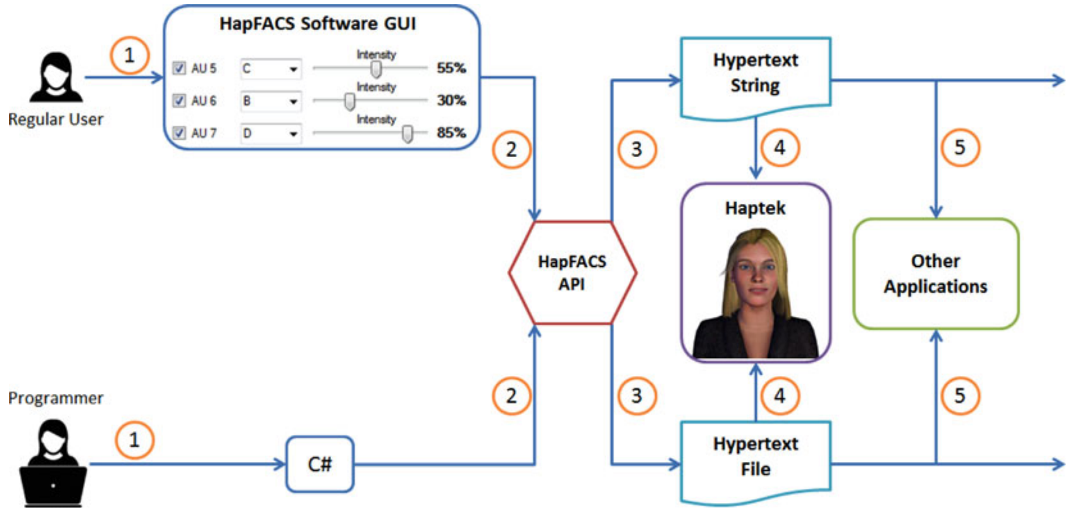

Fig. 2. HapFACS work flow.

combination of the Haptek registers and switches. If we think of a Haptek character as an object in an object-oriented programming language, registers would be the character's (i.e., object's) properties which can be adjusted to change the facial appearance of the character, and switches would be the methods (or functions) that simulate some gestures on the character (e.g. head nod).

Haptek provides a total of 71 registers for head characters (152 and 136 registers for full-body and torso characters respectively), out of which we use 41 registers. Other registers are used for changing the morph for generating different genders and ethnicities. That includes changing neck, head and face proportions, such as the neck thickness, the nose shape, the chin shape, and the distance between eyes. However expertise is needed to learn how registers work[9].

Haptek original registers and switches are not based on FACS, therefore, to generate facial expressions based on the FACS, we followed five steps:

(1) We explored all the Haptek facial, head, and eye registers and switches which manipulate the facial, head, and eye movements and gestures. Haptek registers/switches used in our implementation include six for head movements, four for eye movements, eight for upper-face movements, and 21 for lower-face.

(2) For each AU, we found a subset of $<r, v>$ tuples, where r is a register/switch whose activation in a combination simulates the same movements in the face as an actual AU activation; and v is the maximum intensity value of this specific register/switch (note that registers are character-independent, all characters have the 3D mesh and the same registers, and our results apply to all the characters). The HapFACS designer, who is a FACS-certified coder, found the maximum intensity of the registers experimentally based on the maximum possible activation of the AUs on a human's face. For example, for AU4, the set of tuples used is *{<MidBrowUD, 2>, <LBrowUD, 0.6>, <RBrowUD, 0.6>, <eyes_sad, 1.25>}*.

(3) The FACS manual [17] introduces six intensity levels for each AU: (0) not active, (A) trace, (B) slight, (C) marked or pronounced, (D) severe or extreme, and (E) maximum. Assuming that *E* is activating the AU with 100 percent intensity, based on the FACS manual, we assigned 85 percent to *D*, 55 percent to *C*, 30 percent to *B*, 15 percent to *A*, and 0 percent to *0*.

(4) For each AU intensity level, we applied 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, in AU4, the maximum value of the *MidBrowUD* register is 2.00, so its value for different intensities are: $A = 0.15 \times 2.0 = 0.3, B = 0.6, C = 1.1, D = 1.7, E = 2$.

(5) In addition to the discrete intensity levels (i.e., 0, A, B, C, D, and E), we enabled users to change AU intensities continuously from neutral to maximum intensity, by mapping the $[0\%, 100\%]$ intensity range to the $[0, v]$ of each register/switch r.

When we combine two or more AUs that share a common register, we accumulate their shared register intensities. However, we limit the intensity to the minimum of the maximum intensity values of the common register in combined AUs, so that, the summation of intensities does not deform the face beyond its physiologically possible appearance. For example, let's say for AUs x and y, a single register r is in common with maximum possible intensities of v_x and v_y respectively. So, if we generate an expression with combining AUs x and y, with intensities *a* and *b*, the final intensity of the common register r is calculated using Equation (1):

$$Int_r = \begin{cases} a + b, & \text{if } (a+b) \leq min\{v_x, v_y\} \\ min\{v_x, v_y\}. & \text{if } (a+b) > min\{v_x, v_y\}. \end{cases} \quad (1)$$

3.3 Integration with Other Software

As shown in Fig. 2 (with circled numbers indicating the order of events), HapFACS work flow is simple. HapFACS users can re-produce expressions generated using HapFACS

9. For users interested in registers involved for each AU, HapFACS source code and documentation are freely available upon request at http://ascl.cis.fiu.edu.

on any Haptek characters: when a facial expression animation is generated using HapFACS, a hypertext file with the *.hap* extension (the animation file type for Haptek characters including the relevant registers for the animation, its intensities and activation time) is exported automatically, as well as the video file.

Users only need to load the exported hypertext *.hap* file in their own Haptek character, who will then portray the same previously generated expression(s). Animation files can be loaded to the characters by dragging and dropping the files to the character, or passing the following hypertext to the Haptek character: $\backslash load[file = [fileName.hap]]$[10].

In order to embed HapFACS in another application and use the complete set of HapFACS functionalities, users can either: (1) include the Dynamic-Link Library (DLL) file provided for HapFACS API in their software, and call HapFACS methods; or (2) modify and include HapFACS C# code in their system.

3.4 HapFACS 3.0

In comparison with our previous version and evaluations of HapFACS 1.0 [59], [60], HapFACS 3.0 provides improved functionalities, including improved AU accuracy, and the ability to change intensity continuously and to generate parallel and sequential AU activations (as well as an easy to use interface to change the scene background, light, characters, and hair models).

4 Validation

We performed experiments to validate the HapFACS-generated AUs expressions, and emotional facial expressions. We designed our experiments based on the experiment design constructed to evaluate FACSGen, a software discussed earlier with similar goals to HapFACS [30]. The only differences in the design are that (1) we evaluated 49 AUs (versus 35 in FACSGen); (2) we used a 5-point Likert scale (versus 7-point in FACSGen); and (3) we also evaluated emotion expressions believability and lip-synchronization perception while the characters are speaking in order to inform IVA researchers working with speaking characters[11].

One set of experiments in Section 4.1 (experiments 1 and 2, described below in 4.1.1 and 4.1.2) were conducted with *FACS-certified coders* in order to attempt to *objectively assess the validity of AU activations* (singly or in combinations). Another set of experiments (experiments 3 and 4, described in Sections 4.2.1 and 4.3.1) were conducted with lay participants to evaluate their *subjective perceptions* of HapFACS-generated AUs and EmFACS expressions.

In experiments 1 and 2, we used accuracy, precision, recall, and F1-measure as our objective evaluation measures. Definitions of these terms follow:

$$Accuracy = \frac{t_p + t_n}{t_p + t_n + f_p + f_n} \tag{2}$$

10. We have created introductory and tutorial videos on HapFACS functionalities with animation demos, available at http://ascl.cis.fiu.edu.

11. Although we validated both color and black-and-white stimuli as in the FACSGen study [30], we do not report them here since results show that neither color nor quality affected the recognition tasks.

Fig. 3. Models used for evaluation (happy, disgust, fear, pride, sad, embarrassed, surprised, angry).

$$Precision = \frac{t_p}{t_p + f_p}, \tag{3}$$

$$Recall = \frac{t_p}{t_p + f_n}, \tag{4}$$

$$F1 - Measure = 2 \times \frac{Precision \times Recall}{Precision + Recall}, \tag{5}$$

where t_p, f_p, t_n, and f_n are defined as follows:

- true positive (t_p): number of AUs correctly coded to be present in a given combination.
- false positive (f_p): number of AUs incorrectly coded to be present in a given combination.
- true negative (t_n): number of AUs correctly coded to be absent in a given combination.
- false negative (f_n): number of AUs incorrectly coded to be absent in a given combination.

We also used the Cronbach's alpha as a reliability measure which is a value between 0 and 1 showing the agreement between the raters. The higher the value of the α, the higher the agreement between the coders. Cronbach α values higher than 0.7 are considered as reliable. Equation (6) is used to calculate the alpha value:

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k} \sigma^2_{Y_i}}{\sigma^2_X}\right), \tag{6}$$

where k is the number of coders (i.e., subjects), σ^2_X is the variance of the observed total AUs, and $\sigma^2_{Y_i}$ is the variance of coder i's ratings. It is important to note that inter-rater agreement does not necessarily show the correctness of the raters ratings, but only shows how close different their ratings are.

4.1 FACS-Certified Coders Evaluations

Participants. For experiments 1 and 2 described below in Sections 4.1.1 and 4.1.2, we asked three FACS-certified coders (36 years old female; 26 years old female; and 31 years old male) to rate our individual and combinations of AUs.

Stimuli and design. One 3-second video was generated for each of the 49 individual AUs, and each of the 54 AU combinations (the most common ones identified in the FACS manual). Each of the videos were performed by one of the eight created character models (two white males, two black males, two white females, and two black females) shown in Fig. 3.

TABLE 1
AU Recognition Results in Experiment 1

Action Unit	Accuracy	Precision	Recall	F1-Measure
AU11, AU13, AU14, AU16, AU20, AU41, AU42, AU44	0.99	0.67	0.67	0.67
Other AUs	1.00	1.00	1.00	1.00

TABLE 2
Inter-Rater Correlation (Cronbach α) for Individual AU Recognitions

AU	α	AUs Mistaken for	AU	α	AUs Mistaken for
4	0.922	-	**25**	0.750	-
11	0.750	38	**38**	0.899	-
12	0.922	-	**41**	0.750	43
13	0.750	14	**42**	0.750	4
14	0.750	12	**43**	0.899	-
16	0.750	25	**44**	0.750	4
20	0.750	24	**All Other AUs**	1.0	-
24	0.899	-			

Almost equal number of videos were created using each model. Each video displayed a linear change of the AU intensity starting from 0 percent (i.e., 0 intensity in the FACS manual) to 100 percent (i.e, E intensity in the FACS manual) in 1 second (i.e., onset duration = 1,000 ms), constant at peak for 1 second (i.e., apex duration = 1,000 ms), and lowered from peak to 0 percent intensity in 1 second. All videos have a frame rate of 30 frames/sec and a constant size of 485×480 pixels.

As in [30], videos of the *individual* AUs were accompanied by three still 485×480 pixel images of the AU in three different intensities (30, 60, 90 percent) performed by the same model as in the video. These images are intended to help the coder to see different frames from the same video in different intensities and help him/her in recognition. Videos of AU *combinations* were accompanied by one still 485×480 pixel image of the same AU combination in 70 percent intensity performed by the same model as in the video to help the coders see the activated combination in a single frame.

Procedure. Videos and images were both named randomly to prevent subjects' tagging bias. The videos were hosted on YouTube[12], while the images were hosted on the survey website. In Experiment 1, we asked the FACS-certified coders to identify which AU was activated in each video (each accompanied with three images). In Experiment 2, FACS-certified coders were asked to identify all active AUs in each combination video. As in [30], the coders were not asked to rate the intensity, because judgments over the intensity have been shown to have poor inter-rater agreement [61].

4.1.1 Experiment 1: Individual AUs Validation Results

In this experiment, we examined the accuracy of the AUs generated by HapFACS in terms of AUs described in the FACS manual [17].

12. http://www.youtube.com

TABLE 3
Individual AU Recognition Results in AU Combinations of Experiment 2

Action Unit	Accuracy	Precision	Recall	F1-Measure
AU1	0.99	1.00	0.95	0.98
AU2	0.98	1.00	0.8	0.89
AU4	0.96	0.80	0.8	0.8
AU5	0.98	1.00	0.78	0.88
AU6	0.90	0.53	0.94	0.68
AU7	0.91	0.52	0.93	0.67
AU9	0.98	0.71	0.83	0.77
AU10	0.90	0.76	0.77	0.76
AU12	0.90	0.83	0.84	0.83
AU14	0.98	0.90	0.75	0.82
AU15	0.96	0.88	0.85	0.87
AU16	0.91	0.53	0.6	0.56
AU17	0.93	0.83	0.83	0.83
AU18	0.99	0.75	1.00	0.86
AU20	0.97	1.00	0.67	0.80
AU22	0.99	0.67	1.00	0.80
AU23	0.88	0.79	0.61	0.69
AU24	0.95	0.56	0.56	0.56
AU25	0.96	0.92	0.96	0.94
AU26	0.93	0.62	0.53	0.57
AU27	0.98	0.69	1.00	0.82
AU43	0.99	1.00	0.67	0.8
AU53	1.00	1.00	1.00	1.00
AU54	1.00	1.00	1.00	1.00
AU62	1.00	1.00	1.00	1.00
AU64	0.98	1.00	0.33	0.50

Results and discussion. As shown in Table 1, 41 out of the 49 simulated AUs were recognized by all the three coders with 100 percent recognition rate. The other eight AUs, namely AU11 (Nasolabial Deepener), AU13 (Sharp Lip Puller), AU14 (Dimpler), AU16 (Lower Lip Depressor), AU20 (Lip Stretcher), AU41 (Glabella Lowerer), AU42 (Inner Eyebrow Lowerer), and AU44 (Eyebrow Gatherer), were recognized correctly by only two of the coders (i.e., 66.67 percent precision). The main reason for not being recognized perfectly is that these AUs are very similar to other AUs (listed in Table 2) and are mistakenly identified as other AUs. The *average recognition rate of all 49 AUs was 94.6 percent*.

The Cronbach α values calculated in Table 2 show that, for all the AUs, inter-rater correlation of the recognized AUs are greater than 0.75, which means that the ratings are statistically reliable. For AUs 4, 12, 24, 25, 38, and 43 the α score is less that one, because other AUs were mistakenly recognized as these AUs.

4.1.2 Experiment 2: AU-Combination Validation Results

This experiment was designed to examine the accuracy of AU combinations (documented in FACS as the most common ones) generated by HapFACS.

Results and discussion. Table 3 shows the accuracy, precision, recall, and F1-measure, defined in Equations (2) to (5), of each individual AU recognition in all the AU combinations. Individual AUs used in the combinations are recognized with an average accuracy of *0.96*, average precision of *0.81*, average recall of *0.83*, and average F1-measure of *0.81*.

TABLE 4
AU Combination Recognition Results in Experiment 2

HapFACS-Generated AU Combination	Recognized AUs by FACS-Certified Coders	Accuracy	Precision	Recall	F1-Measure
1 + 2	1 + 2	1.00	1.00	1.00	1.00
1 + 4	1 + 4	1.00	1.00	1.00	1.00
1 + 2 + 4	1 + 2 + 4	0.99	1.00	0.78	0.88
1 + 2 + 5	1 + 2 + 5	1.00	1.00	1.00	1.00
4 + 5	4 + 5 + (41 + 42)	0.98	0.71	0.83	0.77
5 + 7	5 + 7	0.99	1.00	0.83	0.91
6 + 43	6 + 43	0.97	0.63	0.83	0.71
6 + 7 + 12	6 + 7 + 12	0.99	1.00	0.89	0.94
6 + 12 + 15	6 + 12 + 15 + (13 + 14)	0.9	0.78	0.78	0.78
6 + 12 + 15 + 17	6 + 12 + 15 + 17 + (7 + 10)	0.96	0.75	0.75	0.75
6 + 12 + 17 + 23	6 + 12 + 17 + 23 + (7 + 10)	0.97	0.82	0.75	0.78
7 + 12	7 + 12 + (6)	0.99	0.86	1.00	0.92
7 + 43	7 + 43	0.99	1.00	0.83	0.91
9 + 17	9 + 17 + (10 + 13 + 24)	0.98	0.67	1.00	0.80
9 + 16 + 25	9 + 16 + 25 + (22 + 41)	0.98	0.80	0.89	0.84
10 + 14	10 + 14 + (12 + 25)	0.97	0.63	0.83	0.71
10 + 15	10 + 15 + (25)	0.99	0.86	1.00	0.92
10 + 17	10 + 17 + (11 + 24 + 25)	0.97	0.63	0.83	0.71
10 + 12 + 25	10 + 12 + 25 + (6 + 7 + 9)	0.97	0.70	0.78	0.74
10 + 15 + 17	10 + 15 + 17 + (25 + 38)	0.97	0.78	0.78	0.78
10 + 16 + 25	10 + 16 + 25 + (26)	0.97	0.78	0.78	0.78
10 + 17 + 23	10 + 17 + 23 + (9 + 15 + 24)	0.97	0.70	0.78	0.74
10 + 20 + 25	10 + 20 + 25 + (16)	0.98	0.88	0.78	0.82
10 + 23 + 25	10 + 23 + 25 + (11 + 16 + 26)	0.97	0.70	0.78	0.74
10 + 12 + 16 + 25	10 + 12 + 16 + 25 + (6 + 7 + 26)	0.95	0.69	0.75	0.72
12 + 15	12 + 15 + (6 + 17 + 23)	0.97	0.56	0.83	0.67
12 + 17	12 + 17 + (6 + 7 + 23)	0.97	0.63	0.83	0.71
12 + 23	12 + 23 + (6 + 13)	0.98	0.71	0.83	0.77
12 + 24	12 + 24 + (6 + 23)	0.97	0.56	0.83	0.67
12 + 25 + 26	12 + 25 + 26 + (6 + 7 + 10 + 27)	0.95	0.58	0.78	0.67
12 + 25 + 27	12 + 25 + 27 + (7 + 10 + 16)	0.96	0.64	0.78	0.70
12 + 15 + 17	12 + 15 + 17 + (6 + 7)	0.97	0.78	0.78	0.78
12 + 16 + 25	12 + 16 + 25 + (6)	0.98	0.80	0.89	0.84
12 + 17 + 23	12 + 17 + 23 + (7 + 28)	0.97	0.78	0.78	0.78
20 + 23 + 25	20 + 23 + 25 + (12 + 15)	0.97	0.75	0.67	0.71
22 + 23 + 25	22 + 23 + 25 + (16 + 38)	0.97	0.78	0.78	0.78
23 + 25 + 26	23 + 25 + 26	0.98	1.00	0.67	0.80
14 + 17	14 + 17 + (6 + 7 + 12 + 23)	0.97	0.60	1.00	0.75
14 + 23	14 + 23 (12 + 18)	0.97	0.67	0.67	0.67
15 + 17	15 + 17	0.99	1.00	0.83	0.91
15 + 23	15 + 23 + (17)	0.99	0.83	0.83	0.83
17 + 23	17 + 23 + (24)	0.98	0.80	0.67	0.73
17 + 24	17 + 24 + (10)	0.98	0.80	0.67	0.73
18 + 23	18 + 23 + (26)	0.98	0.80	0.67	0.73
20 + 25 + 26	20 + 25 + 26 + (10 + 16)	0.97	0.75	0.67	0.71
20 + 25 + 27	20 + 25 + 27 + (12 + 16)	0.98	0.80	0.89	0.84
4 + 5 + 7 + 24	4 + 5 + 7 + 24 + (23 + 41 + 42)	0.96	0.75	0.75	0.75
10 + 16 + 25 + 26	10 + 16 + 25 + 26 + (15 + 27)	0.95	0.73	0.67	0.70
14 + 54 + 62 + 64	14 + 54 + 62 + 64 + (12)	0.97	0.89	0.67	0.76
1 + 2 + 4 + 5 + 20 + 25 + 26	1 + 2 + 4 + 5 + 20 + 25 + 26 + (13 + 16)	0.94	0.83	0.71	0.77
6 + 12	6 + 12	1.00	1.00	1.00	1.00
12 + 53 + 64	12 + 53 + 64	0.99	1.00	0.78	0.88
1 + 4 + 15	1 + 4 + 15 + (44)	0.99	0.89	0.89	0.89
1 + 2 + 5 + 25 + 27	1 + 2 + 5 + 25 + 27	0.99	1.00	0.93	0.97

Parentheses show false-positive recognitions.

Table 4 shows the recognition accuracy, precision, recall, and F1-measure for different combinations. The generated AU combinations are recognized with an average accuracy of *0.98*, average precision of *0.80*, average recall of *0.81*, and average F1-measure of *0.80*.

4.2 Validation of Emotion Inferences

4.2.1 Experiment 3: EmFACS Inference Evaluation

We evaluated how well HapFACS characters portray EmFACS10 standard emotions [18] in static images and dynamic videos: namely *neutral*, *happiness*, *sadness*, *surprise*,

TABLE 5
AUs Involved in Emotional Expressions

Emotion	AU	Emotion	AU
Happiness	6, 12, 25	*Disgust*	9, 15, 16
Sadness	1, 4, 15	*Contempt*	12, 14R
Surprise	1, 2, 5, 26	*Embarrassment*	12, 52, 62, 64
Anger	5, 7, 9, 10, 15, 17, 42	*Pride*	12, 53, 58, 64
Fear	1, 2, 4, 5, 20, 26	*Neutral*	0

anger, *fear*, *disgust*, *contempt*, *embarrassment*, and *pride*. The AUs involved in each emotion are shown in Table 5.

Stimuli and design. We used two types of stimuli: colored videos, and colored images. Each stimuli type was used with two intensities of 50 and 100 percent, hence a total of 20 videos and 20 images were used.

Lengthwise, videos were 3 seconds and their size was 485×485 pixels. Videos showed the activation of the emotion from 0 to 100 percent-intensity (or 50 percent for lower intensity version) in one second (i.e., onset duration $=$ 1,000 ms), constant at peak intensity for one second (i.e., apex duration $=$ 1,000 ms), and decreasing from peak to 0 percent intensity in one second.

The same eight models as in the previous experiments were used for expressing the emotions (see Fig. 3).

Procedure. Subjects were asked to choose which emotion was portrayed in each video/image, as well as to rate how believable each emotion was being portrayed by the character on a 5-level Likert scale (0: not believable at all, 5: very believable). Participants could select the same emotion for multiple videos/images (e.g., they could label four videos/images with happiness), or select 'None' if they believed some other emotion (not included in the list of emotions provided) was portrayed, or select 'Neutral' when they perceived no emotional expressions.

Participants. For each set of stimuli, we recruited a group of subjects from students on Florida International University (FIU) campus as well as participants on Amazon Mechanical Turk (AMT). Subjects were compensated with $5 vendor gift cards. Table 6 shows the demographic information of the subjects assigned to each of the stimuli groups.

Results and discussion. Tables 7, 8, 9, and 10 show the recognition rates of the emotions when we used different types of stimuli with different intensities. These results show that both intense (100 percent intensity) and subtle (i.e., 50 percent intensity) emotional facial expression animations (either dynamic videos or static images) are perceived with high recognition rates.

Because we wanted to test whether there were major differences between the annotations of the AMT subjects and

TABLE 7
Emotion Recognition Rates for 100 Percent Intensity Video Stimuli

		Recognized Emotions										
		Anger	*Contempt*	*Disgust*	*Embarrass.*	*Fear*	*Happiness*	*Pride*	*Sadness*	*Surprise*	*Neutral*	*None*
Videos	*Anger*	79.3	2.4	6.1	0.0	1.2	1.2	0.0	7.3	0.0	1.2	1.2
	Contempt	0	68.3	0.0	3.7	0.0	18.3	2.4	0.0	1.2	2.4	3.7
	Disgust	7.3	2.4	75.6	0.0	7.3	0.0	0.0	4.9	0.0	1.2	1.2
	Embarrass	0.0	6.1	0.0	75.6	0.0	3.7	1.2	1.2	0.0	3.7	8.5
	Fear	0.0	1.2	2.4	6.1	86.6	0.0	0.0	1.2	2.4	0.0	0.0
	Happiness	0.0	1.2	0.0	4.9	0.0	91.5	1.2	0.0	1.2	0.0	0.0
	Pride	1.2	9.8	1.2	1.2	1.2	2.4	74.4	0.0	2.4	1.2	4.9
	Sadness	0.0	1.2	3.7	1.2	0.0	0.0	0.0	87.8	1.2	0.0	4.9
	Surprise	0.0	0.0	0.0	0.0	1.2	0.0	0.0	0.0	98.8	0.0	0.0
	Neutral	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100	0.0

Average rate is 83.8 percent.

TABLE 8
Emotion Recognition Rates for 50 Percent Intensity Video Stimuli

		Recognized Emotions										
		Anger	*Contempt*	*Disgust*	*Embarrass.*	*Fear*	*Happiness*	*Pride*	*Sadness*	*Surprise*	*Neutral*	*None*
Videos	*Anger*	54.7	0.0	6.3	0.0	6.3	0.0	3.1	26.6	0.0	1.6	1.6
	Contempt	0.0	65.6	0.0	0.0	1.6	6.3	6.3	0.0	0.0	7.8	12.5
	Disgust	10.9	4.7	62.5	3.1	0.0	0.0	0.0	7.8	0.0	4.7	6.3
	Embarrass	1.6	4.7	0.0	67.2	0.0	4.7	10.9	0.0	0.0	3.1	7.8
	Fear	1.6	0.0	9.4	3.1	76.6	1.6	0.0	1.6	1.6	0.0	4.7
	Happiness	0.0	4.7	0.0	0.0	0.0	84.4	4.7	0.0	1.6	1.6	3.1
	Pride	1.6	7.8	3.1	1.6	4.7	4.7	56.3	0.0	6.3	4.7	9.4
	Sadness	0.0	4.7	0.0	4.7	3.1	1.6	1.6	78.1	3.1	1.6	1.6
	Surprise	0.0	0.0	0.0	1.6	0.0	1.6	1.6	1.6	92.2	1.6	0.0
	Neutral	1.6	3.1	0.0	0.0	0.0	3.1	0.0	3.1	0.0	85.9	3.1

Average rate is 72.3 percent.

TABLE 9
Emotion Recognition Rates for 100 Percent Intensity Image Stimuli

		Recognized Emotions										
		Anger	*Contempt*	*Disgust*	*Embarrass.*	*Fear*	*Happiness*	*Pride*	*Sadness*	*Surprise*	*Neutral*	*None*
Images	*Anger*	82.9	1.4	2.9	0.0	2.9	0.0	0.0	7.1	0.0	1.4	1.4
	Contempt	0.0	65.7	0.0	5.7	0.0	15.7	8.6	0.0	1.4	0.0	2.9
	Disgust	10.0	2.9	78.6	0.0	2.9	0.0	0.0	5.7	0.0	0.0	0.0
	Embarrass	0.0	7.1	0.0	77.1	0.0	5.7	0.0	0.0	0.0	2.9	7.1
	Fear	0.0	0.0	4.3	0.0	82.9	0.0	0.0	5.7	4.3	0.0	2.9
	Happiness	0.0	4.3	0.0	2.9	0.0	91.4	0.0	0.0	1.4	0.0	0.0
	Pride	0.0	5.7	0.0	0.0	1.4	2.9	70.0	0.0	2.9	7.1	10.0
	Sadness	1.4	0.0	2.9	0.0	4.3	0.0	0.0	88.6	0.0	1.4	1.4
	Surprise	0.0	0.0	0.0	1.4	4.3	1.4	0.0	0.0	91.4	1.4	0.0
	Neutral	0.0	1.4	0.0	0.0	1.4	0.0	1.4	0.0	0.0	92.9	2.9

Average rate is 82.1 percent.

the local subjects, we performed two T-tests between these two populations for each individual emotion (one for the 50 percent intensity and one for the 100 percent intensity). Our goal was to know if we can combine all the AMT and local users in our analyses. Having the null hypothesis as "two ratings are coming from the same population", the T-test results show that with $\alpha = 0.05, df = 20$, for all emotions, the t value obtained is less than the critical value of 2.085. Therefore, we fail to reject the null hypothesis, or in

TABLE 6
Subjects' Demographic Data in Each Stimuli Group

Stimuli	Intensity	# of Subjects	Female (Age Avg.)	Male (Age Avg.)	White	Black	Asian	Hispanic
Video	**100 percent**	82	42.7 percent (31.3)	57.3 percent (28.3)	57.3 percent	7.3 percent	11 percent	24.4 percent
	50 percent	64	45.3 percent (32.1)	54.7 percent (28.1)	54.7 percent	9.4 percent	14.1 percent	21.9 percent
Image	**100 percent**	70	42.9 percent (33.5)	57.1 percent (27.7)	52.9 percent	8.6 percent	15.7 percent	22.9 percent
	50 percent	76	43.4 percent (30.6)	56.6 percent (28)	61.8 percent	6.6 percent	10.5 percent	21.1 percent

TABLE 10
Emotion Recognition Rates for 100 Percent Intensity Image Stimuli

		Recognized Emotions										
		Anger	*Contempt*	*Disgust*	*Embarrass.*	*Fear*	*Happiness*	*Pride*	*Sadness*	*Surprise*	*Neutral*	*None*
Images	*Anger*	68.4	1.3	11.8	0.0	2.6	0.0	2.6	2.6	0.0	6.6	3.9
	Contempt	0.0	56.6	0.0	2.6	0.0	11.8	7.9	0.0	1.3	13.2	6.6
	Disgust	6.6	2.6	64.5	2.6	5.3	0.0	0.0	0.0	3.9	5.3	9.2
	Embarrass	0.0	3.9	0.0	63.2	0.0	7.9	0.0	0.0	0.0	15.8	9.2
	Fear	0.0	0.0	1.3	5.3	72.4	0.0	0.0	10.5	3.9	0.0	6.6
	Happiness	0.0	7.9	1.3	1.3	0.0	81.6	1.3	0.0	1.3	1.3	3.9
	Pride	1.3	0.0	0.0	3.9	2.6	0.0	53.9	0.0	3.9	18.4	15.8
	Sadness	2.6	0.0	1.3	3.9	1.3	0.0	0.0	86.8	0.0	0.0	3.9
	Surprise	0.0	1.3	0.0	2.6	5.3	1.3	0.0	0.0	89.5	0.0	0.0
	Neutral	1.3	0.0	0.0	3.9	2.6	1.3	0.0	1.3	0.0	88.2	1.3

Average rate is 72.5 percent.

other words, we can not say that the two samples are coming from two different populations.

Table 11 shows the reported believability of the emotional facial expressions by the subjects (i.e., how natural and believable is the character when expressing the emotion), as well as the reliability measure (Cronbach α) for each of the stimuli groups. Cronbach α values were computed for each of the expressions, using participants' ratings as columns (items) and the 10 videos as rows (cases). Results shown in Table 11 indicate that for all expressions $\alpha > 0.7$, which means that the faces were rated reliably.

For each of the stimuli types, we calculated a 10×2 (i.e., Emotion $\times$ Intensity) ANOVA ($df = 1$). For all stimuli types, $p > 0.05$, which shows that there is not enough statistical evidence to show that intensity of facial expressions has a significant effect on recognition rates.

Also, for each stimuli type, we performed a 10×2 (i.e., Emotion $\times$ Intensity) ANOVA analysis ($df = 1$) to find out the effects of intensity on believability. For both video stimuli types, $p > 0.05$, showing that there is not enough statistical evidence to show that intensity of the facial expressions in videos has a significant effect on believability of expressions. However, for image stimuli types, ANOVA analysis shows that there is a significant effect of intensity on believability of images ($F = 9.55, p < 0.05$). Increasing the intensity increases the salience of the emotion in images, which causes an increase in believability. We did not have the similar effect of intensity on the believability in video expressions.

TABLE 11
Believability (belv.) and Its Standard Deviation (sd); and Cronbach α Reliability Measure for Different Stimuli Types and Intensities

Emotion	Measure	Video		Image	
		100%	*50%*	*100%*	*50%*
Anger	*Belv. (sd)*	3.9 (0.97)	3.84 (0.86)	4.3 (0.84)	3.8 (0.85)
	Cronbach α	0.992	0.966	0.993	0.984
Contempt	*Belv. (sd)*	3.5 (1.00)	4.22 (0.7)	4.1 (0.64)	3.8 (0.78)
	Cronbach α	0.989	0.982	0.981	0.972
Disgust	*Belv. (sd)*	3.83 (0.93)	3.38 (1.05)	4 (0.76)	3.3 (0.89)
	Cronbach α	0.990	0.975	0.990	0.982
Embarrass.	*Belv. (sd)*	3.72 (0.93)	3.88 (0.77)	3.7 (0.94)	3.8 (0.84)
	Cronbach α	0.992	0.982	0.991	0.982
Fear	*Belv. (sd)*	3.73 (0.9)	3.78 (0.79)	3.9 (0.79)	3.5 (0.89)
	Cronbach α	0.995	0.988	0.993	0.989
Happiness	*Belv. (sd)*	3.1 (1.10)	3.52 (1.11)	3.9 (0.79)	3.5 (0.99)
	Cronbach α	0.997	0.994	0.997	0.993
Pride	*Belv. (sd)*	3.7 (0.86)	3.66 (0.98)	3.4 (0.79)	3.5 (0.97)
	Cronbach α	0.990	0.965	0.986	0.974
Sadness	*Belv. (sd)*	3.7 (0.93)	3.88 (0.98)	4.1 (0.89)	3.8 (0.9)
	Cronbach α	0.997	0.989	0.996	0.996
Surprise	*Belv. (sd)*	4.13 (0.77)	4.27 (0.7)	3.9 (0.92)	3.7 (0.91)
	Cronbach α	1.00	0.997	0.997	0.996
Neutral	*Belv. (sd)*	4.3 (0.82)	4.02 (0.77)	4.1 (0.80)	3.7 (0.91)
	Cronbach α	1.00	0.995	0.998	0.996
Average Believability		3.8 (0.97)	3.84 (0.92)	3.9 (0.80)	3.7 (0.91)

TABLE 12
Recognition Rates for Speaking Characters

		Recognized Emotions										
		Anger	*Contempt*	*Disgust*	*Embarrass.*	*Fear*	*Happiness*	*Pride*	*Sadness*	*Surprise*	*Neutral*	*None*
Videos	*Anger*	66.7	0.0	5.0	0.0	3.3	0.0	0.0	16.7	0.0	5.0	3.3
	Contempt	0.0	45.0	0.0	5.0	0.0	23.3	13.3	0.0	0.0	6.7	6.7
	Disgust	11.7	3.3	80.0	0.0	1.7	0.0	1.7	1.7	0.0	0.0	0.0
	Embarrass	0.0	6.7	1.7	70.0	0.0	5.0	0.0	3.3	0.0	6.7	6.7
	Fear	3.3	1.7	1.7	1.7	76.7	0.0	0.0	15.0	0.0	0.0	0.0
	Happiness	0.0	6.7	0.0	0.0	0.0	85.0	0.0	1.7	0.0	3.3	3.3
	Pride	0.0	1.7	0.0	1.7	0.0	1.7	48.3	0.0	25.0	1.7	3.3
	Sadness	1.7	1.7	3.3	0.0	0.0	0.0	1.7	88.3	0.0	0.0	3.3
	Surprise	0.0	0.0	0.0	0.0	1.7	0.0	6.7	0.0	90.0	0.0	1.7
	Neutral	0.0	1.7	0.0	1.7	0.0	0.0	0.0	0.0	0.0	93.3	3.3

4.3 Evaluating Speaking Characters

4.3.1 Experiment 4: Validation of Emotional Speech

For IVA researchers interested in speaking characters, we validated how well HapFACS simulated characters can show emotions while they speak.

Participants. We recruited 20 FIU students and 40 AMT workers for this experiment. They included 61.7 percent females (avg. age $= 30.7$) and 38.3 percent males (avg. age $= 28.6$). Subjects were from different ethnicities including White (63.3 percent), Black (6.7 percent), Asian (8.3 percent), Hispanic (20 percent), and Caucasian (1.7 percent).

Stimuli and design. Ten videos were generated (8.3 seconds long in average) with random sentences. The emotions portrayed in the videos were the ones shown in Table 5 with 100 percent intensity. Eight models were randomly selected (two white males, two black males, two white females, and two black females) to portray each video. In all videos, a neutral voice and utterance was used in order to focus the study on emotional facial expressions (rather than voice or utterance).

Procedure. Each subject was asked to recognize the expressed emotion while the character was speaking. Also, participants were asked to rate the believability of the character while speaking on a 5-level Likert scale (0: not believable at all, 5: very believable).

Results and discussion. Table 12 shows the participants' *emotion recognition rates of speaking characters with an average rate of 74.3 percent*.

While the character is speaking and expressing an emotion, many of the lower-face AUs are activated for speaking. For emotions such as *contempt* that are expressed only with lower-face AUs, speaking can lower the recognition rates. Also, as confirmed by other studies [62], [63], recognition rates for contempt, sadness, disgust, and fear may be lower than happiness, surprise, and anger. Taken together, the low recognition rates on these emotions may be a general feature of the emotion expressions, and not of the presented images/videos.

On average, the believability of the characters in the speaking videos was rated as 3.3 (stdev $= 1.01$). Cronbach α values were computed for each of the expressions, using participants' ratings as columns (items) and the 10 videos as rows (cases). Table 13 shows the believability and reliability

TABLE 13
Believability (Belv.) and Its Standard Deviation (sd); and Cronbach α Reliability Measure for Speaking Characters

	Belv. (sd)	Cronbach α
Anger	3.3 (0.91)	0.979
Contempt	3.4 (0.91)	0.944
Disgust	2.9 (0.95)	0.989
Embarrassment	3.3 (0.94)	0.983
Fear	3.2 (1.08)	0.987
Happiness	3.6 (0.98)	0.994
Pride	3.2 (1.16)	0.949
Sadness	3.2 (0.95)	0.996
Surprise	3.1 (1.07)	0.996
Neutral	3.6 (0.98)	0.998

measure for each individual video. Cronbach alpha values indicate that the faces were reliably rated.

4.3.2 HapFACS and FACSGen

Finally, we compared HapFACS with FACSGen, which is the most similar software to HapFACS [30]. Although FACSGen characters do not have lip-synchronization nor bodies (i.e., only heads) and therefore currently have limited appeal for researchers interested in speaking IVAs, FACSGen is powerful to generate realistic 3D faces. As discussed earlier, we matched our experiment settings (including stimuli and process) to those performed with FACSGen in order to be able to compare it with HapFACS.

Results. HapFACS expresses 49 individual AUs with average recognition rate of 94.6 percent, whereas FACSGen expresses 35 individual AUs with average recognition rate of 98.6 percent. HapFACS is evaluated as 98 percent accurate in expressing 54 AU combinations while FACSGen is evaluated as 80.1 percent accurate in expressing the same 54 combinations.

Table 14 compares the recognition rates and expression believability of HapFACS and FACSGen for static emotional expressions.

Discussion. One interesting aspect of generating software such as FACSGen or HapFACS is that they can help

TABLE 14
Comparison between Recognition Rates and Believability of the Static Emotional Expressions Generated in FACSGen and HapFACS

Emotion	HapFACS		FACSGen	
	100 percent (Bel.)	*50 percent (Bel.)*	*100 percent (Bel.)*	*50 percent (Bel.)*
Angry	75 (4.0)	67.5 (3.7)	87.82 (3.6)	71.79 (3.3)
Contempt	62.5 (3.9)	55 (3.5)	56.41 (3.1)	48.08 (3.1)
Disgust	75 (3.4)	65 (3.8)	68.59 (3.0)	61.54 (2.9)
Embarrassment	80 (3.9)	80 (3.5)	69.23 (3.4)	60.26 (3.1)
Fear	87.5 (3.8)	82.5 (3.7)	72.44 (3.2)	67.31 (3.1)
Happiness	92.5 (3.7)	90 (3.4)	88.46 (3.7)	77.56 (3.4)
Pride	67.5 (3.6)	62.5 (3.5)	74.36 (3.7)	71.79 (3.5)
Sadness	90 (3.9)	87.5 (3.8)	83.97 (3.4)	76.28 (3.3)
Surprise	97.5 (3.9)	87.5 (3.7)	87.82 (3.7)	87.82 (3.5)
Neutral	97.5 (3.8)	100 (3.7)	-	-

FACSGen believability rates are scaled from 7 to 5-scale for comparison.

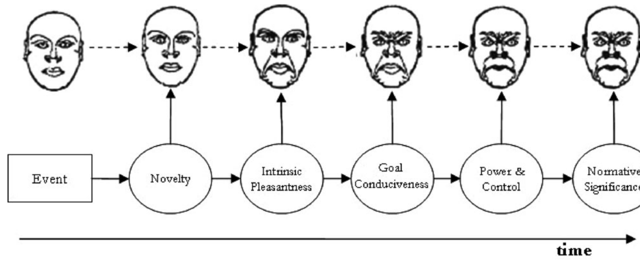

Fig. 4. Representation of emotional facial process.

the generation and testing of human emotion expression generation.

According to Scherer's Component Process Theory (CPT) [3], for example, humans evaluate each emotion-related event with respect to five Sequential Evaluation Checks (SEC), and each SEC is associated with the sequential activation of AU(s). When an event occurs, one sequentially evaluates the event *novelty* in term of its suddenness as well as its familiarity and predictability. One then evaluates the *pleasantness* or unpleasantness of the event, followed by how much the event is *conducive* to one's goals or aversions, and one's *coping potential* (i.e. whether the event is controllable, followed by how much *power* one has to modify the event outcome). The last component is *normative significance*: one evaluates the event with respect to one's external and internal standards, as social and individual norms.

The CPT process is represented in Fig. 4, and the corresponding AU activations suggested for each SEC are shown in Table 15: for each component, an AU-based emotional sub-expression is generated sequentially, and the dynamic activation sequence of all these emotional sub-expressions represents the individual's emotional facial expression for that event. CPT also suggests specific combinations for a specific set of 12 known emotions.

Using HapFACS to simulate CPT [3], we were able to confirm that the open research questions identified earlier [64] need to be answered before emotional expressions generated based on CPT reach the desired level of believability:

- Are all the SEC related sub-expressions equally important? If they are not, should some of them last longer or be more intense?

TABLE 15
SEC to AUs Predictions [3]

SEC	AUs
Novelty	1, 2 & 5
	4, 7, 26 & 38
Pleasantness	5, 26 & 38
	12 & 25
Unpleasantness	4, 7, 9, 10, 15, 17, 24 & 39
	16, 19, 25 & 26
Goal-Need Conduciveness (discrepant)	4, 7, 17 & 23
No Control	15, 25, 26, 41 & 43
Control & High Power	4 & 5
	7, 23 & 25
	23, 24 & 38
Control & Low Power	1, 2, 5, 20, 26 & 38

Fig. 5. Timing and intensities issues.

- How should different sub-expressions fuse? (see Fig. 5 for examples)?
- Are all AUs related to the same sub-expression equally important?
- How should we choose among the different possible SEC to AU predictions shown in Fig. 6?

The development of theories such as CPT that study the dynamic generation of facial expressions during an emotional event, and the ability to test these theories with tools such as HapFACs or FACSGen are likely to lead to scientific answers about emotion generation and their expressions.

5 General Discussion

We studied HapFACS validity for creating AUs defined in FACS, and the emotional meaning conveyed by HapFACS expressions. Experiment 1 reported validation data for 49 single AUs. Experiment 2 reported validation data for 54 AU combinations and validation of individual AUs used in those combinations. All the expressions were implemented in faces of different sexes and ethnicity.

The recognition rates of the AUs were high and the AUs interacted predictably in combination with each other (to generate facial expressions of emotions). For all AUs, validity of the AU appearance was scored satisfactorily by FACS-certified coders. Based on the reported high recognition rates for combinations and for individual AUs, obtained with the good inter-rater reliability scores, results suggest that the AUs synthesized by HapFACS 3.0 are valid with respect to the FACS.

Overall, when performed with high intensity, *surprise*, *fear*, *happiness*, *sadness*, and *neutral* were the most easily recognizable emotions, whereas *contempt* and *pride* were the most difficult to detect. When performed with low intensity, *surprise*, *happiness*, and *neutral* were the easiest to recognize, while *anger*, *contempt*, *pride*, and *disgust* were more difficult to recognize.

The *contempt* expression is sometimes perceived as *pride* (or even *happiness*), which we hypothesize is due to the subtle asymmetric smile that can be displayed in pride (but in happiness as well). The low recognition rate of *contempt* also confirms the findings by Langner et al. [62] and van der Schalk et al. [63], who discuss the reason as a general feature of the contempt expression, which is not as expressive nor visual as other expressions.

Experiments 3 showed that participants recognized the expected affective meanings conveyed by emotional expressions generated with HapFACS. The reported recognition rates were high and comparable to previous research [62], [63], [65], [66], [67].

6 Conclusion and Future Work

We presented HapFACS 3.0, a new free 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 lay participants.

Future versions of HapFACS will have improved expressiveness of the imperfect AUs (see Section 4.1.1); and will provide non-linear changing of the intensity for video generation since AU activation can be non-linear from a geometric point of view.

We hope that HapFACS will prove useful for the affective community, particularly psychologists working on facial expression of emotion generation, and intelligent virtual agents who need to be able to speak.

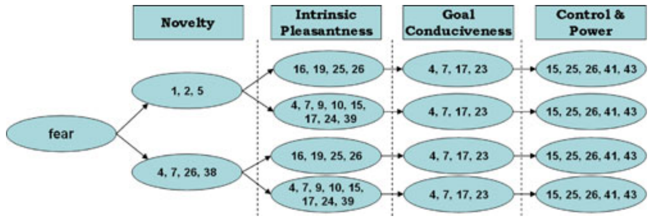

Fig. 6. The multiple prediction issue for *fear* [64].

Acknowledgments

We would like to acknowledge Drs. Erika Rosenberg and Daniel Messinger for their expertise on the Facial Action Coding System and help identifying certified FACS scorers. Part of the research described in this article was developed with funding from the National Science Foundation, grant numbers HRD-0833093, IIP-1338922, IIP-1237818, and IIS-1423260.

References

[1] N. Ambady and R. Rosenthal, "Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis," *Psychol. Bull.*, vol. 111, no. 2, pp. 256–274, 1992.
[2] J. Allwood, "Bodily communication dimensions of expression and content," *Multimodality Lang. Speech Syst.*, vol. 7, no. 26, pp. 1–15, 2002.
[3] K. R. Scherer, A. Schorr, and T. Johnstone, *Appraisal Processes in Emotion: Theory, Methods, Research*. New York, NY, USA: Oxford Univ. Press, 2001.
[4] A. Sarrafzadeh, S. Alexander, F. Dadgostar, C. Fan, and A. Bigdeli, "See Me, Teach Me: Facial expression and gesture recognition for intelligent tutoring systems," in *Proc. Innovations Inf. Technol.*, Nov. 2006, pp. 1–5.
[5] T. Bosse, G. F. Siddiqui, and J. Treur, "An intelligent virtual agent to increase involvement in financial services," in *Proc. 10th Int. Conf. Intell. Virtual Agents*, 2010, pp. 378–384.
[6] P. Kenny, T. Parsons, J. Gratch, A. Leuski, and A. Rizzo, "Virtual patients for clinical therapist skills training," in *Proc. Intell. Virtual Agents*, 2007, pp. 197–210.

[7] C. Lisetti, R. Amini, U. Yasavur, and N. Rishe, "I can help you change! An empathic virtual agent delivers behavior," *ACM Trans. Manag. Inf. Syst.*, vol. 4, no. 4, pp. 1–28, 2013.

[8] A. Arya, S. DiPaola, L. Jefferies, and J. T. Enns, "Socially communicative characters for interactive applications," in *Proc. 14th Int. Conf. Comput. Graph., Visual. Comput. Vision*, Plzen, Czech Rep.: Union Agency Science Press, 2006.

[9] S. C. Marsella, J. Gratch, and J. Rickel, "Expressive behaviors for virtual worlds," in *Life-Like Characters, Tools, Affective Functions, and Applications*, M. Prendinger and H., Ishizuka, Eds. New York, NY, USA: Springer, 2003.

[10] M. S. Hussain, S. K. D'Mello, and R. A. Calvo, "Research and development tools in affective computing," in *The Oxford Handbook of Affective Computing*, 1st ed. London, U.K.: Oxford Univ. Press, 2015, ch. 25, pp. 349–357.

[11] M. Alencar and J. F. Netto, "TUtor collaborator using multi-agent system," *Collab. Technol. Social Comput.*, vol. 460, pp. 153–159, 2014.

[12] R. Amini, C. Lisetti, and U. Yasavur, "Emotionally responsive virtual counselor for behavior-change health interventions," in *Proc. 9th Int. Conf. Design Sci. Res. Inf. Syst. Technol.*, 2014, pp. 433–437.

[13] U. Yasavur, C. Lisetti, and N. Rishe, "Let's talk! Speaking virtual counselor offers you a brief intervention," *J. Multimodal User Interfaces*, vol. 8, pp. 381–398, 2014.

[14] R. Mattheij, M. Nilsenova, and E. Postma, "Vocal and facial imitation of humans interacting with virtual agents," in *Proc. Humaine Assoc. Conf. Affective Comput. Intell. Interaction*, Sep. 2013, pp. 815–820.

[15] R. Mattheij, M. Postma-Nilsenová, and E. Postma, "Mirror, mirror in the wall: Is there mimicry in you all?" *J. Ambient Intell. Smart Environ.*, vol. 1, pp. 1–5, 2013.

[16] R. Amini, C. Lisetti, U. Yasavur, and N. Rishe, "On-demand virtual health counselor for delivering behavior-change health interventions," in *Proc. IEEE Int. Conf. Healthcare Informat.*, 2013, pp. 46–55.

[17] P. Ekman, W. V. Freisen, and J. C. Hager, *Facial Action Coding System*, vol. 160, 2nd ed. Salt Lake City, UT, USA: Research Nexus eBook, 2002.

[18] W. V. Friesen and P. Ekman, "EMFACS-7: Emotional facial action coding system," Univ. California, San Francisco, CA, USA, unpublished manuscript, 1983.

[19] 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.

[20] P. Ekman and W. V. Freisen, *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Mountain View, CA, USA: Consulting Psychol. Press, 1978.

[21] P. Ekman and W. V. Freisen, "Detecting deception from the body or face," *J. Personality Social Psychol.*, vol. 29, no. 3, pp. 288–298, 1974.

[22] N. Chovil, "Discourse-oriented facial displays in conversation," *Res. Lang. Social Interaction*, vol. 25, pp. 163–194, 1991.

[23] A. Fridlund, *Human Facial Expression: An Evolutionary View*. San Diego, CA, USA: Academic, 1994.

[24] D. Lundqvist, A. Flykt, and A. Öhman, *The Karolinska Directed Emotional Faces—KDEF*. (CD ROM). Solna, Sweden: Dept. Clinical Neurosci., Psychol. Sec., Karolinska Inst., 1998.

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

[26] 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*, vol. 11, no. 4, pp. 907–920, 2011.

[27] F. Schiel, S. Steininger, and U. Türk, "The smartkom multimodal corpus at BAS," in *Proc. 3rd Conf. Lang. Resources Evaluation*, 2002, pp. 200–206.

[28] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, I. Matthews, and F. Ave, "The extended Cohn-Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression," in *Proc. Comput. Vision Pattern Recog. Workshops*, 2010, pp. 94–101.

[29] M. Valstar and M. Pantic, "Induced disgust, happiness and surprise: An addition to the MMI facial expression database," in *Proc. Int. Conf. Lang. Resources Evaluation, Workshop on Emotion*, 2010, pp. 65–70.

[30] 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.

[31] K. R. Scherer and H. Ellgring, "Multimodal expression of emotion: Affect programs or componential appraisal patterns?" *Emotion*, vol. 7, no. 1, pp. 158–71, Feb. 2007.

[32] M. Thiebaux, M. Rey, A. N. Marshall, S. Marsella, and M. Kallmann, "SmartBody: Behavior realization for embodied conversational agents," in *Proc. 7th Int. Conf. Auton. Agents Multiagent Syst.*, 2008, pp. 12–16.

[33] A. Shapiro, "Building a character animation system," in *Proc. 4th Int. Conf. Motion Games*, 2011, pp. 98–109.

[34] S. Villagrasa and A. S. Sánchez, "Face! 3d facial animation system based on FACS," in *Proc. IV Ibero-Amer. Symp. Comput. Graphics*, 2009, pp. 203–209.

[35] N. Bee, B. Falk, and E. Andre, "Simplified facial animation control utilizing novel input devices: A comparative study," in *Proc. 14th Int. Conf. Intell. User Interfaces*, 2009, pp. 197–206.

[36] S. Pasquariello and C. Pelachaud, "Greta: A simple facial animation engine," in *Proc. 6th Online World Conf. Soft Comput. Ind. Appl. Session Soft Comput. Intell. 3D Agents*, 2001, pp. 511–525.

[37] E. Bevacqua, K. Prepin, R. Niewiadomski, E. de Sevin, and C. Pelachaud, "GRETA: Towards an Interactive Conversational Virtual Companion," in *Close Engagement with Artificial Companions: Key Social, Psychological, Ethical and Design Issues*, Y. Wilks, Ed. Amsterdam, The Netherlands: John Benjamins, 2010, pp. 143–156.

[38] S. Battista, F. Casalino, and C. Lande, "MPEG-4: A multimedia standard for the third millennium, Part 1," *IEEE Multimedia*, vol. 6, no. 4, pp. 74–83, Oct.–Dec. 1999.

[39] A. Wojde and L. Rothkrantz, "Parametric generation of facial expressions based on FACS," *Comput. Graphics Forum*, vol. 24, no. 4, pp. 743–757, 2005.

[40] O. Alexander, M. Rogers, W. Lambeth, M. Chiang, and P. Debevec, "Creating a photoreal digital actor: The digital emily project," in *Proc. Conf. Visual Media Production*, Nov. 2009, pp. 176–187.

[41] O. Alexander, G. Fyffe, J. Busch, X. Yu, R. Ichikari, A. Jones, P. Debevec, J. Jimenez, E. Danvoye, B. Antionazzi, M. Eheler, Z. Kysela, and J. Pahlen, "Digital Ira: Creating a real-time photoreal digital actor," in *Proc. ACM SIGGRAPH*, 2013, p. 4.

[42] A. T. Hirsh, M. P. Jensen, and M. E. Robinson, "Evaluation of nurses' self-insight into their pain assessment and treatment decisions," *J. Pain*, vol. 11, no. 5, pp. 454–61, May 2010.

[43] C. Arrington, D.-M. Wilson, and L. Lehmann, "Improving performance and retention in computer science courses using a virtual game show," in *Proc. 49th Annu. Southeast Regional Conf.*, 2011, p. 320.

[44] T. S. Behrend and L. F. Thompson, "Similarity effects in online training: Effects with computerized trainer agents," *Comput. Human Behavior*, vol. 27, no. 3, pp. 1201–1206, May 2011.

[45] T. S. Behrend, S. Toaddy, L. F. Thompson, and D. J. Sharek, "The effects of avatar appearance on interviewer ratings in virtual employment interviews," *Comput. Human Behavior*, vol. 28, no. 6, pp. 2128–2133, Jul. 2012.

[46] K. Bransky and D. Richards, "Users' expectations of IVA recall and forgetting," in *Proc. 10th Int. Conf. Intell. Virtual Agents*, 2011, pp. 433–434.

[47] B. De Carolis, I. Mazzotta, N. Novielli, and S. Pizzutilo, "Social robots and ECAs for accessing smart environments services," in *Proc. Int. Conf. Adv. Visual Interfaces*, 2010, pp. 275–278.

[48] 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," *J. Res. Personality*, vol. 46, no. 3, pp. 239–247, Jun. 2012.

[49] R. Ishii and I. N. Yukiko, "An empirical study of eye-gaze behaviors: Towards the estimation of conversational engagement in human-agent communication," in *Proc. Workshop Eye Gaze Intell. Human Mach. Interaction*, 2010, pp. 33–40.

[50] Y. Liu, O. Sourina, and M. K. Nguyen, "Real-time EEG-based human emotion recognition and visualization," in *Proc. Int. Conf. Cyberworlds*, Oct. 2010, pp. 262–269.

[51] I. Nakano and R. Ishii, "Estimating user's engagement from eye-gaze behaviors in human-agent conversations," in *Proc. 15th Int. Conf. Intell. User Interfaces*, 2010, pp. 139–148.

[52] N. Novielli, F. de Rosis, and I. Mazzotta, "User attitude towards an embodied conversational agent: Effects of the interaction mode," *J. Pragmatics*, vol. 42, no. 9, pp. 2385–2397, Sep. 2010.

[53] C. Smith, N. Crook, J. Boye, D. Charlton, S. Dobnik, D. Pizzi, M. Cavazza, and S. Pulman, "Interaction strategies for an affective conversational agent," in *Proc. 10th Int. Conf. Intell. Virtual Agents*, 2010, pp. 301–314.

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

[55] M. Turunen, J. Hakulinen, O. Ståhl, 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," *Comput. Speech Lang.*, vol. 25, no. 2, pp. 192–209, Apr. 2011.

[56] C. E. Vargas and D. Field, "'How was your day?' An architecture for multimodal ECA systems," in *Proc. 11th Annu. Meeting Special Interest Group Discourse Dialogue*, 2010, pp. 47–50.

[57] M. Cavazza, C. E. Vargas, J. R. n. Gil, I. D. Telefónica, N. Crook, D. Field, and S. Sheffield, "'How was your day?' An affective companion ECA prototype," in *Proc. 11th Annu. Meeting Special Interest Group Discourse Dialogue*, 2010, pp. 277–280.

[58] O. Sourina, Y. Liu, and M. K. Nguyen, "Emotion-enabled EEG-based interaction," in *Proc. SIGGRAPH Asia 2011 Posters*, 2011, p. 1.

[59] R. Amini and C. Lisetti, "HapFACS: An open source API/software to generate FACS-based expressions for ECAs animation and for corpus generation," in *Proc. 5th Biannu. Humaine Assoc. Conf. Affective Comput. Intell. Interaction*, 2013, pp. 270–275.

[60] R. Amini, U. Yasavur, and C. L. Lisetti, "HapFACS 1.0: Software/API for generating FACS-based facial expressions," in *Proc. 3rd Int. Symp. Facial Anal. Animation*, 2012, p. 17.

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

[62] O. Langner, R. Dotsch, G. Bijlstra, D. H. J. Wigboldus, S. T. Hawk, and A. van Knippenberg, "Presentation and validation of the radboud faces database," *Cognition Emotion*, vol. 24, pp. 1377–1388, 2010.

[63] J. van der Schalk, S. T. Hawk, A. H. Fischer, and B. J. Doosje, "Validation of the Amsterdam dynamic facial expression set (ADFES)," *Emotion*, vol. 11, pp. 907–920, 2011.

[64] M. Paleari, A. Grizard, C. L. Lisetti, and S. Antipolis, "Adapting psychologically grounded facial emotional expressions to different anthropomorphic embodiment platforms," in *Proc. 20th FLorida Artif. Intell. Res. Soc. Annu. Conf. Artif. Intell.*, 2007, pp. 7–9.

[65] M. G. Beaupre and U. Hess, "Cross-cultural emotion recognition among Canadian ethnic groups," *J. Cross-Cultural Psychol.*, vol. 36, pp. 355–370, 2005.

[66] E. Goeleven, R. de Raedt, L. Leyman, and B. Verschuere, "The Karolinski directed emotional faces: A validation study," *Cognition Emotion*, vol. 22, pp. 1094–1118, 2008.

[67] J. L. Tracy, R. W. Robins, and R. A. Schriber, "Development of a FACS-verified set of basic and self-conscious emotion expressions," *Emotion*, vol. 9, pp. 554–559, 2009.

Reza Amini received the bachelor's degree in biomedical engineering from the University of Isfahan, Iran, in 2006, the master's degree in electrical engineering from the Isfahan University of Technology, Iran, in 2009, and the PhD and master's degree in computer science from Florida International University, in 2012 and 2015, respectively. He is a senior scientist at IPsoft. His research interests include artificial intelligence, machine learning, and affective computing.

Christine Lisetti is an associate professor at the School of Computing and Information Sciences, Florida International University, where she directs the Affective Social Computing Laboratory (http://ascl.cis.fiu.edu). She is one of the founders of the field of Affective Computing, and a member of the Editorial Board of the *IEEE Transactions on Affective Computing*. Her long-term research goal is to create virtual engaging socially intelligent agents that can interact naturally with humans via expressive multimodal communication in a variety of contexts involving socioemotional content (e.g., health coach, social companion, educational games).

Guido Ruiz is currently working toward the bachelor's degree in computer science at Florida International University. His research interests include artificial intelligence.