

RESEARCH ARTICLE

Laughter and smiling facial expression modelling for the generation of virtual affective behavior

Miquel Mascaró¹, Francisco J. Serón², Francisco J. Perales¹, Javier Varona^{1*}, Ramon Mas¹

1 Department of Mathematics and Computer Science, University of the Balearic Islands, Palma de Mallorca, Spain, **2** Department of Computer Science, Zaragoza University, Zaragoza, Spain

* xavi.varona@uib.es



OPEN ACCESS

Citation: Mascaró M, Serón FJ, Perales FJ, Varona J, Mas R (2021) Laughter and smiling facial expression modelling for the generation of virtual affective behavior. PLoS ONE 16(5): e0251057. <https://doi.org/10.1371/journal.pone.0251057>

Editor: Enkelejda Kasneci, University of Tübingen, GERMANY

Received: February 14, 2020

Accepted: April 20, 2021

Published: May 12, 2021

Copyright: © 2021 Mascaró et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The MASEPEsmile data set (freely accessible from <http://ugivia.uib.es/MASEPEsmile/>) consists of 91 videos of seven subjects, with acting experience.

Funding: The authors acknowledge the Ministerio de Economía, Industria y Competitividad (MINECO), the Agencia Estatal de Investigación (AEI), and the European Regional Development Funds (ERDF) for its support by the Project EXPLainable Artificial INtelligence systems for health and well-beING (EXPLAINING), PID2019-104829RA-I00/AEI/10.13039/501100011033, and

Abstract

Laughter and smiling are significant facial expressions used in human to human communication. We present a computational model for the generation of facial expressions associated with laughter and smiling in order to facilitate the synthesis of such facial expressions in virtual characters. In addition, a new method to reproduce these types of laughter is proposed and validated using databases of generic and specific facial smile expressions. In particular, a proprietary database of laugh and smile expressions is also presented. This database lists the different types of classified and generated laughs presented in this work. The generated expressions are validated through a user study with 71 subjects, which concluded that the virtual character expressions built using the presented model are perceptually acceptable in quality and facial expression fidelity. Finally, for generalization purposes, an additional analysis shows that the results are independent of the type of virtual character's appearance.

Introduction

Facial expression modelling of virtual characters presents many difficulties, including time constraints, cost and complexity. On the one hand, from the field of psychology, approaches range from the pioneering work of Duchenne [1] and Darwin [2] to more recent contributions such as those of Provine [3], Gruner [4], Morreall [5] and Ruch [6]. These works address facial expressions and their relationship to human emotions. Ekman and Friesen published what they called Facial Action Coding System (FACS) [7], which has been used as a standard to categorize the facial expression of emotions. FACS has been employed by both psychologists and animators. On the other hand, from the field of artistic representation, most authors, such as Bridgman [8], Loomis [9], Hogarth [10] and Faigin [11], refer to the importance of knowledge of anatomy for proper representation of facial expressions.

In particular, laugh and smile are complex facial expressions. Both offers a great number of both emotional meanings and visual information. For example, the smile can have elements of other expressions such as sadness and anger, creating interesting effects of ambiguity and

by the previous projects: TIN2016-81143-R, TIN2011-24660, TIN2015-67149-C3-1-R, TIN2015-67149-C3-2-R, TIN2016-80347-R. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

complexity. Therefore, the representation of laughter and smiling can be approached from different perspectives. In a previous work on laughter synthesis, DiLorenzo et al. [12] propose a physically based parametric model of the human chest that can be automatically driven from pre-recorded audio laugh samples. This model is anatomically inspired and synthesizes the movements of the torso muscle activated by the air flow within the body. The model is restricted to the respiration during the laughter act and it does not involve any facial motion. Cosker and Edge [13] propose another data-driven model for non-speech related articulations such as laughs, cries, yawns and sneezes. The model is based on a Hidden Markov Model (HMM) trained from motion-capture data and audio segments. Motion capture obtains the data using 30 markers placed on the face and normalized to a facial model. During training, this model learns the correlations between the recorded audio and the visual data. Griffin et al. [14] investigate how the laughter is perceived from body movements. But, they can only identify 5 different types of laughter.

From the virtual character modelling point of view, Niewiadomski and Pelachaud [15] consider how laughter intensity modulates facial motion. They focus on modelling both laughter and respiration. In their work, facial motions depends only on the laughter intensity but, different types of laughter are not considered. The same authors [16] also study the factors that influence the perception of AUs in virtual characters: Stimulus intensity, presentation (static or animated) and presence of wrinkles. In a first study, they evaluate the AUs of laughter, in terms of identification, naturalness and realism. A second study evaluates the expressions of laughter in terms of measuring the quality of the animation and its meaning. Also Niewiadomski et al. [17] do an experiment on how the intensity of the incongruity between audio and animation is perceived in an avatar with a laughing animation. The test assesses, among other things, its naturalness, plausibility and credibility. Urbain et al. [18] propose to compare the similarity of a sample with recorded laughter audio information and then select the corresponding sequence of facial expressions. The selection is done by measuring the acoustic similarities between the input laughter and the output one. Ding et al. [19] model hilarious laughter. They have developed a generator for face and body motions that takes as input the sequence of pseudo-phonemes of laughter and the time duration of each pseudo-phoneme. The generator learns the relationship between input signals and human motions. Then, it can be used to automatically produce laughter animation in real time. The same author [20] presents an animation controller of the upper body of a virtual character from audio input. Ochs et al. [21] identify the morphology and dynamic characteristics of different smiles in a virtual agent. They claim that there are three different types of smiles: funny, polite, and embarrassed. They create an algorithm and a web application to generate smiles with a virtual character. Mancini et al. [22] provide a virtual character with models of synthesis of laughter based on a paradigm of “expressivity-copying” and check how the presence of the character affects the perception of music and the mood of the user. Most of these works have a common approach: the integration of laughter into virtual characters (avatars) as a fundamental task for machine-human communication. This enables the design of sociable conversational agents using natural-looking and natural-sounding laughter. Unlike these previous works, the presented approach is only centred in the visual information (i.e., it does not depend on audio data).

Previous works focus on the generation of standalone valid models of laughter and smiles whilst our proposal is to develop a model within a general animation framework. To our knowledge, no references are found on laughter synthesis that focus on improving or creating specific tools for the character animation.

The structure of the paper is as follows. First, we present a computational model for the generation of facial expressions associated with laughter and smiling in order to facilitate the synthesis of such facial expressions in virtual characters. This model is based on the learning of

real expressions by means of the facial feature tracking in video sequences. Next, by using the learnt animation parameters, we present the procedural animation system to reproduce different types of smiles, which are validated by conducting a user study. Finally, we finish by presenting the conclusions and discussing future work.

Laughter and smiling modeling

In this section, we present a computational model for the generation of facial expressions associated with laughter and smiling. First, we define a taxonomy of laugh and smile. Next, we present the character's facial model and how it is connected with a facial tracker in order to obtain the data for learning the human facial features' motions. Finally, based on the obtained data set we describe a procedure to synthesize such facial expressions in virtual characters.

A taxonomy

First, we present a taxonomy to classify the different types of laugh and smile. Earlier laugh and smile classifications come from the field of psychology, and their main interest is to ascertain the expression's authenticity. In this category are the works of Ekman [6, 23–25]. From these works, smiling and laughter are claimed to be universal indicators of joy [26]. From the expression synthesis point of view, we refer to the work of Faigin [11], where it can be found an exhaustive taxonomy of different types of expressions of joy (see Fig 1).

This taxonomy is based on the observation of separate key elements in laughter and smiles and how they work. First, it present a classification according to the joy intensities, going from *laughter* to *smile* (Fig 1, from 1 to 4). To this classification, it is possible to add the results of combining the actions of the zygomatic and orbicularis muscles with different positions of the eyebrows, corresponding to the forms defined in the universal expressions. Thus, the *sly smile* combines a smile with eyebrows of the anger expression (Fig 1.9); *avid laugh* with surprise expression (Fig 1.7); the *ingratiating smile* with the expression of fear (Fig 1.8) and the *melancholy smile* with the sadness expression (Fig 1.6). Other categories involve voluntary expressions: the *stifled smile* (Fig 1.5) is produced to suppress the spontaneous laughter, and the *abashed smile* (Fig 1.11) is produced to repress satisfaction. The *debauched smile* (Fig 1.10) is the combination of the smile with a position of the eyelids in which the pupils are partly covered; it is the position that can be associated with states of drowsiness and intoxication. Finally, the taxonomy is completed by the *false smile* and *false laughter* (Fig 1.12 and 1.13).

Facial model

A virtual character's facial model must allow the representation of diverse faces in both realistic and cartoon aesthetics while retaining its anthropomorphic properties. Facial rigging is the process of defining the animation controls for a facial model. In this regard, the virtual character's rig should provide a method to faithfully reproduce all facial muscle activity. For this reason, the character's rig has been designed using the edge-loop technique [27]. This technique optimizes the facial deformations distributing the geometry as muscle lines. There is no standard for defining the rig system interface; in our case, we have chosen an interface based on a 3D view with 2D handlers, similar to the solutions proposed by Alexander et al. [28] and Digital tutors [29]. With these approximations, we have an exhaustive control of the geometric deformation that performs the mesh simulation of facial muscles as well as an intuitive tool to easily perform all the facial actions. In conclusion, we use a general-purpose rig that, in addition to helping in the design of any facial expression, can be used specifically to describe laughter and smiles (see Fig 2). To define the deformations of the character meshes, our rig uses a combination of techniques for bone and blendshape interpolation. Thus, the user is able to

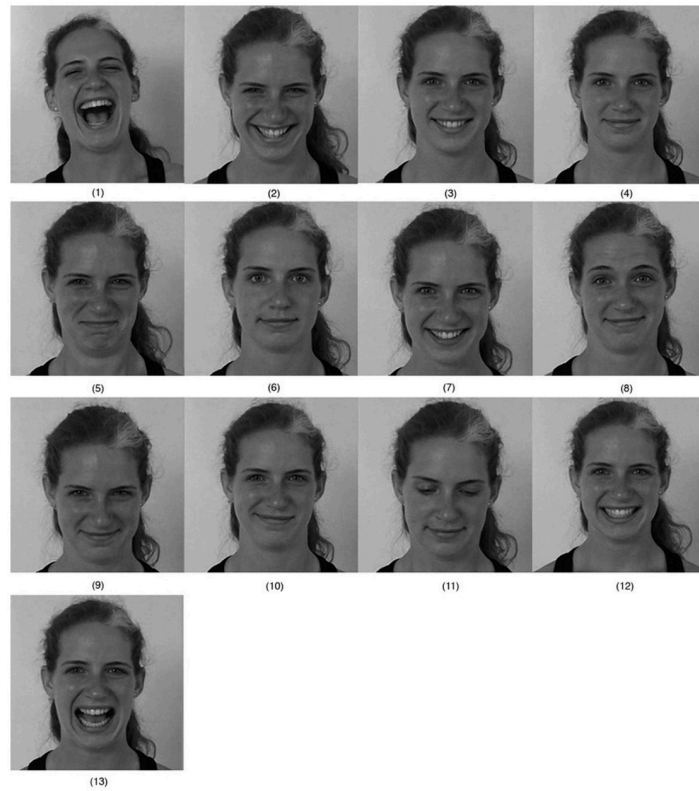


Fig 1. Laugh and smile taxonomy based on the different expressions of joy.

<https://doi.org/10.1371/journal.pone.0251057.g001>

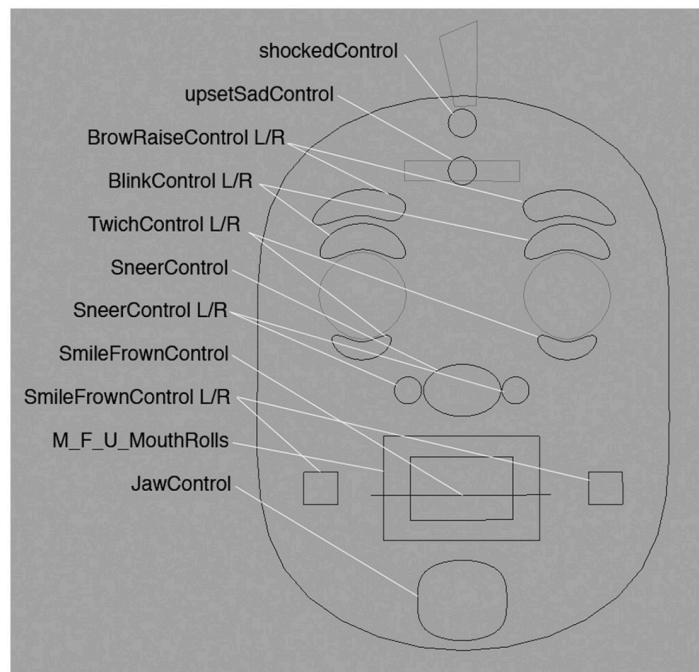


Fig 2. Rig controllers.

<https://doi.org/10.1371/journal.pone.0251057.g002>

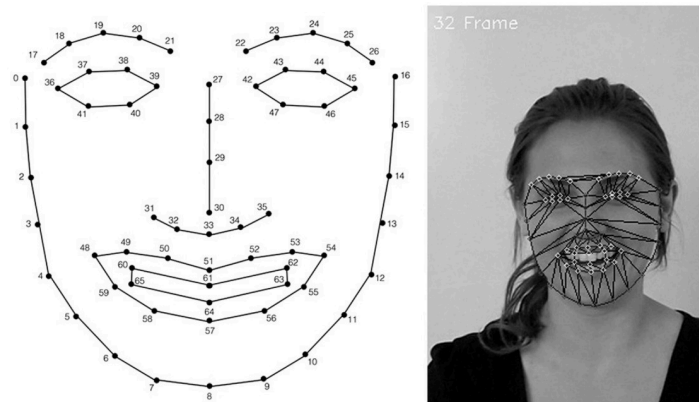


Fig 3. Visual tracking of facial landmarks.

<https://doi.org/10.1371/journal.pone.0251057.g003>

control hundreds of polygons to generate all the facial expressions with the simplified use of animation curves. These animation curves automatically drive all the rig controllers.

In order to model the animation curves for each laugh and smile facial expression, we want to learn the real facial features’ motions associated with these expressions. To avoid the use of special hardware, learning is based on the computer vision algorithm of Saragih et al. [30]. By means of this algorithm, a 66-point mesh of facial landmarks is obtained for each sequence frame, which is tracked following the motion of the user’s motions (see Fig 3).

A temporal rule-based method [31] is used to set the facial rig animation curves from the tracked facial landmarks. First, in Table 1 we define the model facial parameters $FP = \{FP_1, FP_2, \dots, FP_{15}\}$. Next, correspondences from facial parameters to facial landmarks (P_i , where $0 \leq i \leq 65$) and controller actions (C_i , where $1 \leq i \leq 17$) are defined in Table 2.

Therefore, we have defined a direct correlation between the facial points’ motions and the values that the rig controllers must take to generate the desired facial expression. Thus, for a frame i , the controller action value, C_i , is given by Eq (1)

$$C_i = (FP_{x_i} - FP_{x_0}) \cdot \varepsilon, \tag{1}$$

Table 1. Facial parameter definition.

FPs	Action
FP_1	Half opening of lower lid (right)
FP_2	Internal distance between eyebrow and eyelid (right)
FP_3	Half opening of lower lid (left)
FP_4	Internal distance between eyebrow and eyelid (right)
FP_5	Distance of eye opening (right)
FP_6	Distance of eye opening (left)
FP_7	Width of mouth
FP_8	Distance corner of mouth to nose (right)
FP_9	Distance corner of mouth to nose (left)
FP_{10}	Height of mouth
FP_{11}	Vertical between mouth and nose
FP_{12}	Distance between corner of mouth and nose-wing (right)
FP_{13}	Distance between corner of mouth and nose-wing (left)
FP_{14}	Vertical between chin and nose
FP_{15}	Distance between oval of face

<https://doi.org/10.1371/journal.pone.0251057.t001>

Table 2. Facial landmarks and rig controllers' correspondences.

FPs	Landmarks	Rig controller
FP ₁	46–47	rightTwich
FP ₂	22–27	upsetSad, shocked, rightBrownRaise
FP ₃	40–41	leftTwich
FP ₄	21–27	upsetSad, shocked, leftBrownRaise
FP ₅	22–42	rightBlink
FP ₆	21–39	leftBlink
FP ₇	48–54	smileFrown
FP ₈	33–54	rightSmileFrown
FP ₉	33–48	leftSmileFrown
FP ₁₀	51–57	jaw
FP ₁₁	51–53	sneer
FP ₁₂	35–54	rightSneer
FP ₁₃	31–48	leftSneer
FP ₁₄	8–33	jaw
FP ₁₅	0–16	neck

<https://doi.org/10.1371/journal.pone.0251057.t002>

where FP_x is the specific facial parameter defined in Table 1, the subscript 0 refers to the neutral expression, and ε is the controller adjustment factor. This factor depends on the animator, who is responsible for choosing the hardness or softness of an expression (affecting on the degree of realism he wants to obtain). Empirically, we found that visually correct values of ε may range from 0.1 to 0.3 for a 640x490 face image resolution corresponding to a subject's frontal plane. For instance, Fig 4 shows the animation curve for the *SmileFrown* controller of the *Smiling Open-Mouthed* expression. The x-axis represents time (frame) and the y-axis represents the controller's values. The scaled curve at a value of $\varepsilon = 0.1$ is shown at the top. At the bottom is represented the curve adjusted for $\varepsilon = 0.3$. It is possible to observe that this parameter controls the expression smoothing between realistic values.

By means of these transformations, our system generates a curve for each controller action from a video sequence containing the desired facial expression of laugh and smile.

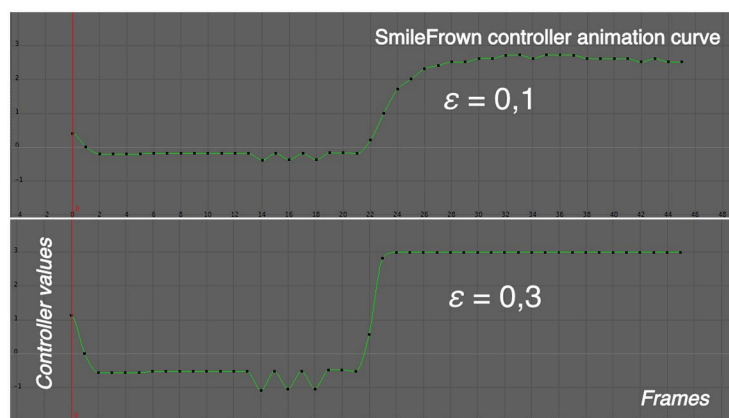


Fig 4. Animation curve for the *SmileFrown* controller at different ε values.

<https://doi.org/10.1371/journal.pone.0251057.g004>

Laugh and smile learning

In order to learn the animation curves to create realistic expressions that involve laugh and smile, we employ different facial expression data sets with the types of laughter and smiles defined in the previous taxonomy.

Classical databases of facial expressions such as **CK+** [32] and **MMI** [33] include expressions that are all conveniently annotated and include the main emotions. In particular, there are data sets for detection and classification of types of laughter as the **BBC Smile Dataset** [34], the **Uva-Nemo** database [35], which was generated to study the dynamics of a spontaneous smile and a voluntarily generated smile, and the **MAHNOB Laughter database** [36], which has been used for the differentiation and detection of laughter and smiling during speech sessions, Haddad et al. [37] present the AmuS database which is a database for the synthesis and recognition of joy in speech (English and French). It is primarily a study of speech acoustics, distinguishing between smiled-speech and speech-laugh. Jansen [38] has created MULAI which is a database that classifies types of social laughter based on their social function. It includes recordings of dyadic interactions (two subjects) and contains body movement data, ECG (electrocardiogram) and GSR (galvanic skin response). However, these databases do not address other types of laughter than the associated with expressions of joy. Therefore, regarding the used taxonomy, in previous databases there are representative examples of *uproarious laughter*, *laughter*, *open-mouth smile*, *closed-mouth smile*, *false smile* and *false laugh*, but not of others. For this reason, we have built a new facial expression data set, the **MASE-PESmile** database, which provides exactly the 13 types of laughter described in the taxonomy.

The **MASEPESmile** data set (freely accessible from <http://ugivia.uib.es/MASEPESmile/>) consists of 91 videos of seven subjects, with acting experience. The subjects has given written informed consent (as outlined in PLOS consent form) and they know that their participation is voluntary, and that their images will be used for research purposes, assuring that their privacy, anonymity, and confidentiality will be protected.

The data set recording was done in a room with controlled lighting. Each actor entered the recording set individually in order to minimize the influence of the other actors participating in the recording. Before each recording, the director described in detail the expression to perform. For example, for the expression *Sly Smile*: “It is the laughter of the evil one, of the astute one, of the paid one of itself, the eyebrows rise from the outside and wrinkle in the middle, the eyelids swell with pressure at the bottom and the lips are thin and stretch upwards pressing along the skull.” In addition, in front of the actor was projected an image with a picture of the expression taken from Faigin’s work [11]. Once ensured the actor understood what he has to perform, the expression was recorded. Finally, the expression was validated by a human expert and the recording is repeated as many times as necessary to meet the smile requirements. Each subject was asked to perform the 13 types of laugh and smile, starting at the neutral expression. Fig 5 shows 7 subjects of the data set performing four of the 13 categorized expressions. Specifically and ordered by columns we can see: *uproarious laughter*, *smiling open mouthed*, *melancholy smile* and *ingratiating smile*.

In previous section is detailed how to process one frame to compute one controller curve’s point given the landmarks (see Eq 1 and Tables 1 and 2). Then, it is possible to compute the animation curve for each sequence of the MASEPESmile data set. For instance, in Fig 6 is shown a virtual character animation key frame, the controllers’ animation curves, and the original video frame of expression *open-mouth smile*.

Once computed the animation curves for all the data set sequences, following the previously explained procedure for each sequence, it is possible to establish the most representative curves for the animation of all the considered types of laugh and smile. These most representative



Fig 5. Multiple subjects with different laughter and smiling types of the MASEPESmile data set.

<https://doi.org/10.1371/journal.pone.0251057.g005>

curves will allow the generation of a procedural animation system for reproducing different types of laugh and smile of the data set. In order to learn the most representative animation curves, first, we apply dynamic time warping to find the optimal alignment between each controller values. Thus, we can determine the similarities between different facial expression performances. For instance, for the expression *open-mouth smile*, recorded by the seven subjects of our experiment (Fig 7), in Table 3 are shown the alignment values for the *smileFrow* controller.

This process is extended to all controllers and all the laugh and smile facial expressions in order to learn the representative curves. For each subject, we apply the K-means algorithm

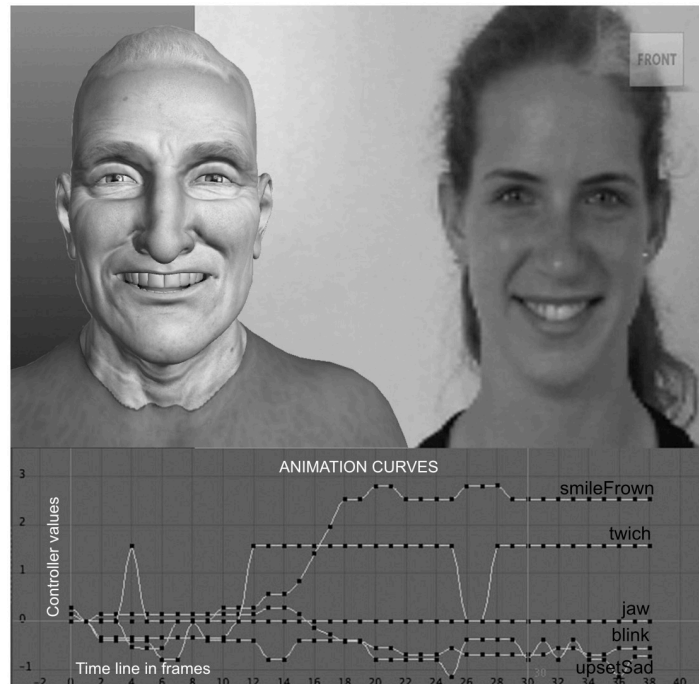


Fig 6. Mapping data for expression *open-mouth smile* to a virtual character.

<https://doi.org/10.1371/journal.pone.0251057.g006>

using the alignment value as distance between each data set sample. The resulting performance prototypes are selected as the most representative animation curves. Finally, the desired facial expression animation is obtained by applying the most representative curves to the facial model. Thus, newly generated expressions take into account the different performances included in the data set to generalize the virtual character animation curves for the synthesis of new laugh and smile facial expressions. Examples of the results are shown in the videos included in the supporting information.

In addition, by analyzing the alignment tables for different expressions and controllers, it is possible to explain the obtained expression models. For instance, in the case of the expression of *uproarious laughter*, the minimum alignment distance for the *smileFrown* controller has a



Fig 7. Open-mouth smile performed by different subjects.

<https://doi.org/10.1371/journal.pone.0251057.g007>

Table 3. Distance alignment values for the *smileFrown* controller for the *open-mouth smile* expression.

	<i>s01</i>	<i>s02</i>	<i>s03</i>	<i>s04</i>	<i>s05</i>	<i>s06</i>	<i>s07</i>
<i>s01</i>		12.85	12.88	27.82	16.68	36.82	11.31
<i>s02</i>	12.85		42.15	10.90	8.66	16.17	9.16
<i>s03</i>	12.88	42.15		48.25	35.20	38.21	18.76
<i>s04</i>	27.82	10.90	48.25		4.78	4.29	13.42
<i>s05</i>	16.68	8.66	35.20	4.78		9.56	12.43
<i>s06</i>	36.82	16.17	38.21	4.29	9.56		13.66
<i>s07</i>	11.31	9.16	18.16	13.42	12.43	13.66	
<i>mean</i>	19.73	16.65	32.57	18.24	14.55	19.79	13.12
<i>standard deviation</i>	10.32	12.79	13.82	17.00	10.87	14.32	3.21

<https://doi.org/10.1371/journal.pone.0251057.t003>

huge value (42.38), which indicates that several data set expressions are less generalizable due to a subject's performance dependency. But, in these cases, the cycle of laughter is more prevalent. This fact, is demonstrated in the work of Ruch [6], which describes how the cycle basically depends on the lung volume, which obviously depends on the subject.

Results and discussion

User study: Expression description

For the data set recording, prior to perform an expression, the actor receive a description of a context where such an expression would occur and its corresponding classification. A user study was conducted in order to validate these descriptions used to build the MAPESmile data set. In addition, with this user study, we can compare the perception of the real data set expressions with the generated virtual expressions.

Fifty-one unpaid students of both sexes aged between twenty and twenty-eight years old were recruited. All the participants has given written informed consent (as outlined in PLOS consent form) prior to testing to ensure that they know that their participation is voluntary, that they will incur no physical or psychological harm, that they can withdraw at any time, and that their privacy, anonymity, and confidentiality will be protected.

In order to conduct the experiment, each participant rank six or seven expressions randomly chosen from the taxonomy (approximately the 50% of the expressions). For each chosen expression, first, the participant sees the description of the expression employed to built the data set. Next, the participant sees the two videos of the real actor and the virtual character of the selected expression, and answers the question: *Does the expression corresponds to the description?*. This question is scored on a scale ranging from 0 (very little) to 5 (very much).

As expected, it can be observed in Fig 8 that real expression are better perceived than the generated virtual expression. The grand mean for the real expression was 4.27 (SD = 0.32). For the generated virtual expression, the grand mean was 3.4 (SD = 1.09). Then, the answers of the question show that the generated virtual character expressions corresponds to their descriptions.

User study: Movement quality and representativeness

To validate our laugh and smile model, we design a user study for a human evaluation of the movement quality and representativeness of the generated expressions in one avatar. In addition, we are interested in studying whether the avatar used could influence the user perceptions. We used two different avatars: a realistic avatar, and a cartoon avatar (see Fig 9). The

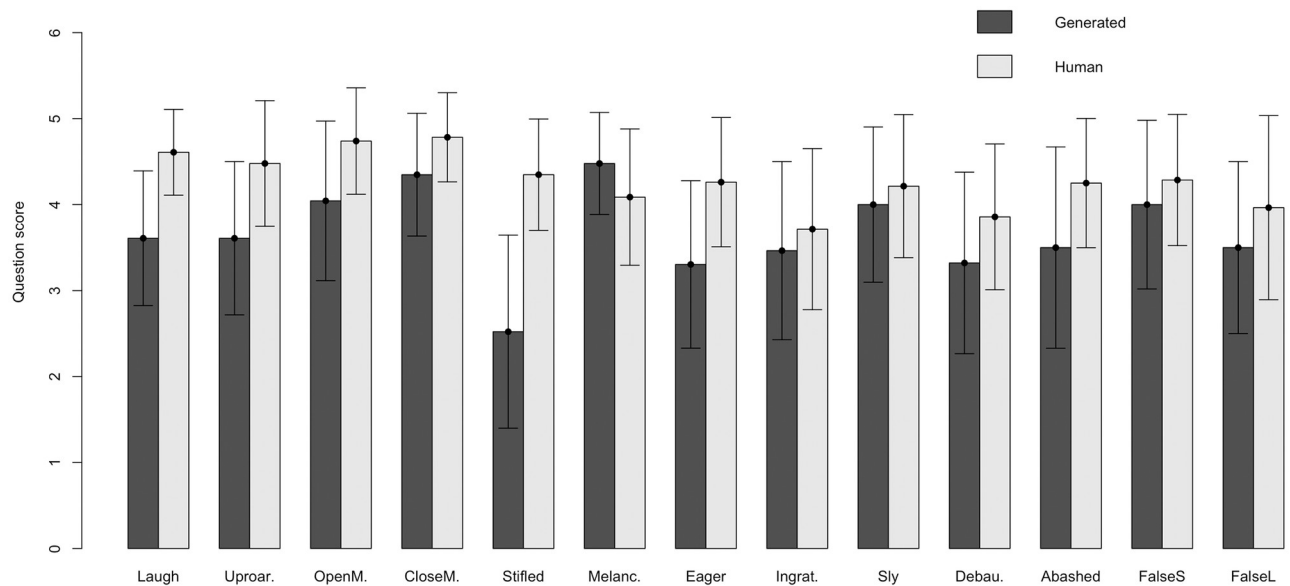


Fig 8. Comparison of the question scores for each expression performance (real vs generated).

<https://doi.org/10.1371/journal.pone.0251057.g008>

expressions of both avatars are made with the same animation curves, synthesized by applying our previously described method.

In order to study the perception of the synthesized virtual character expressions, seventy-one unpaid university students of both sexes aged between nineteen and twenty-one years old were recruited (a within-subject design was used). All the participants has given written informed consent (as outlined in PLOS consent form) prior to testing to ensure that they know that their participation is voluntary, that they will incur no physical or psychological harm, that they can withdraw at any time, and that their privacy, anonymity, and confidentiality will be protected. Therefore, each participant seeing thirteen videos of virtual character facial expressions, one for each expression in our study. After seeing each video, they answered two questions:

- Q1: Does the motion feel right?
- Q2: Does the motion represent the description?

The objective of Q1 is to test the expression realism; meanwhile, the goal of Q2 is to validate the perceived expression.

Both questions are answered on a sheet where a brief description of the expression will be found. Each question is scored on a scale ranging from 0 (very little) to 5 (very much).

The grand mean for Q1 was 3.32 (SD = 0.45). For Q2, the grand mean was 3.34 (SD = 0.51). Thus, the test answers show that the generated expressions were realistic and that they were positively perceived by the users. Nevertheless, it can be observed in Fig 10 that the users did not perceive all the expressions in the same manner. The best-scoring expression is sly smile (M = 4.13 and SD = 1.14 for movement; and M = 4.30 and SD = 1.11 for representativeness), followed by the smile with mouth closed (M = 4.11 and SD = 0.91 for movement; and M = 4.19 and SD = 0.90 for representativeness). The worst-scoring expression is ingratiating smile (M = 2.49 and SD = 1.28 for movement; and M = 2.45 and SD = 1.30 for representativeness), followed by eager smile (M = 2.87 and SD = 1.04 for movement; and M = 2.82 and SD = 1.21 for representativeness).

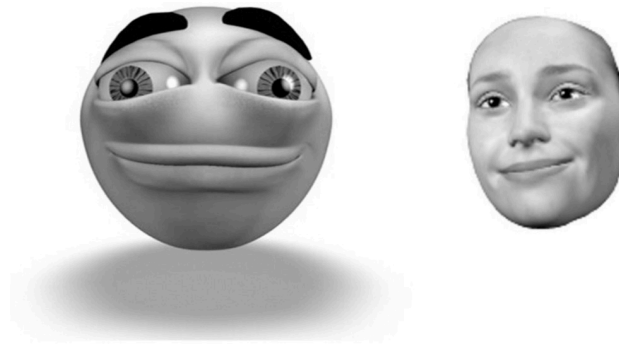


Fig 9. Character avatar synthesis for the melancholy smile.

<https://doi.org/10.1371/journal.pone.0251057.g009>

There is a distinction between the highest-scoring expressions and those rated worse. We think that among the latter are those whose meaning can be more confusing, while the meanings of the top-rated expressions are less ambiguous. Thus, *ingratiating smile* in the test has been described as *suggestive* or *to get sympathy*, and *eager smile* has been described as

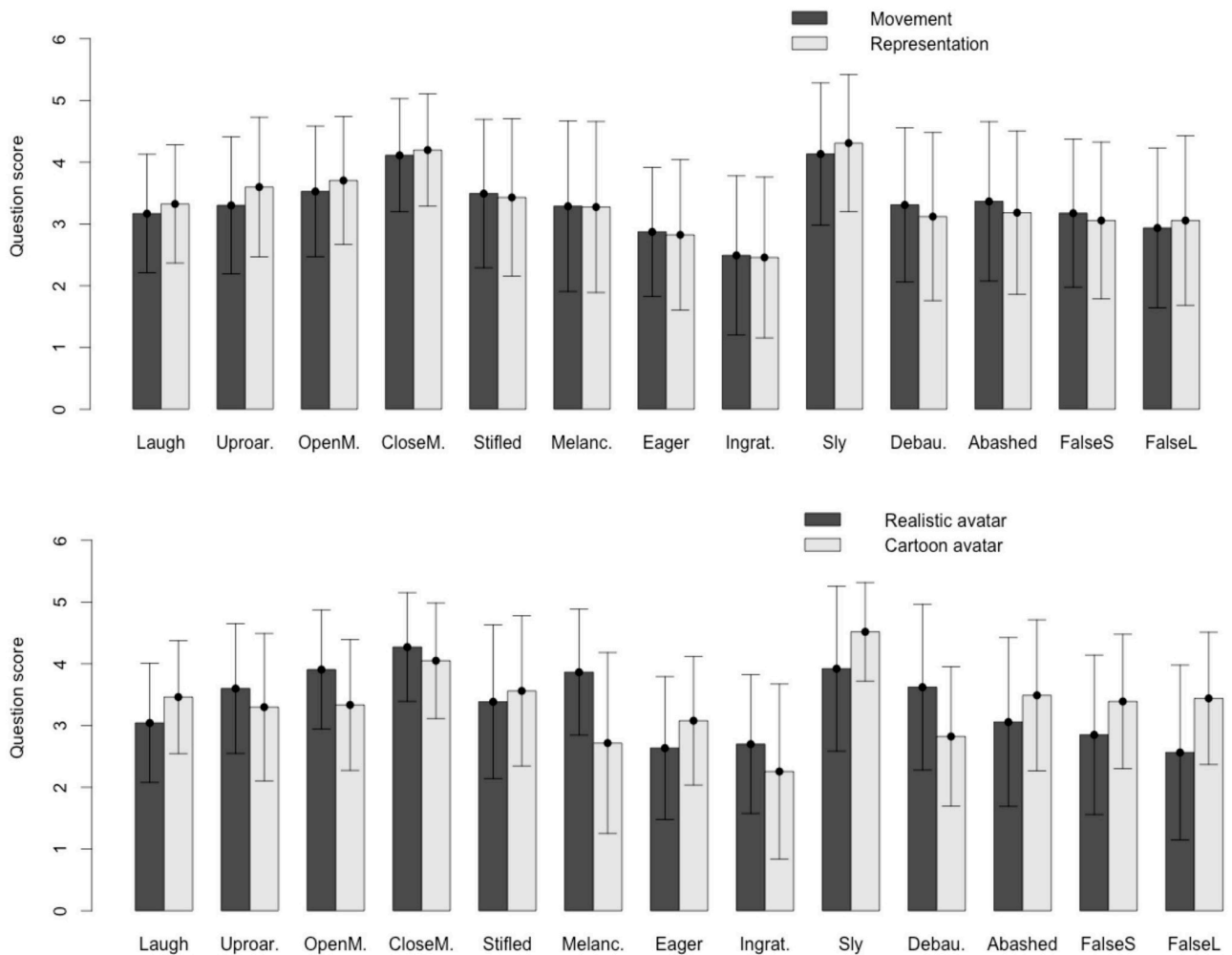


Fig 10. Mean and standard deviation (+/- SD) of the scores for both questions for all the evaluated expressions (71x13x2 = 1846 scores and 71x13x4 = 3692 scores).

<https://doi.org/10.1371/journal.pone.0251057.g010>

anxious or *euphoric*, which received the worst scores, coinciding with the ambiguity in their descriptions.

By analyzing the scores of both questions by the kind of avatar used to represent the expression, we found that the grand mean for the realistic avatar was 3.34 (SD = 0.56) and that for the cartoon avatar was 3.33 (SD = 0.56).

With two-way ANOVA, we found that the main effect of the avatar on the score was not statistically significant ($F(1,3640) = 2.825, p > 0.05$). The results by avatar and expression are shown in Fig 10.

There was a significant expression interaction effect ($F(12,3640) = 5.167, p < 0.001$), which was due to the difference between the two avatar expressions of *melancholy smile* and *false laughter*, as determined by a Scheffé post hoc analysis. In the *melancholy smile* case, the realistic avatar has better scores, while for the *false laughter* expression, the cartoon avatar is better perceived. This fact is explained if we consider that *melancholy smile* is the combination of facial action units corresponding to two opposite emotions such as joy and sadness, so, this expression is one of the most complex to interpret, especially when out of context, as in our test. It is reasonable that the representation of this expression by the realistic avatar has a better acceptance because the realism of the avatar facilitates the understanding and validation of the movement. In the case of *false laughter*, we should realize that this expression is the one with the largest interpretative load, not only because it is consciously performed but also because it has a voluntary component of exaggeration. This is what is expected from all the expressions of cartoon characters and what explains the result obtained.

Summarizing, from the analysis of the user study, we can conclude that the expressions generated using the presented method are acceptable in movement quality and in fidelity of representation of expressions, and this result is independent of the type of avatar used; that is, the expressions generated by the presented system are accepted independently of whether they are performed by a realistic or by an unrealistic character, as it follows the anthropomorphic rules.

Limitations

Our model is versatile and adapts to most facial animation rigs, as long as the isolated animation of the movement of the eyebrows, eyelids and mouth accepts the animation keys that our system generates. Of course, is a requirement for the animation environment to accept the .anim format of the Autodesk Maya software. Another constraint of our model is that it is not currently integrated as a plugin within the animation environment, this can be a drawback as the user needs to manually import the curves into their scene.

Conclusions and future work

In this work, we defined a framework capable of producing realistic representations of the different types of laughter and smile in virtual characters. The presented model is systematic, automatic, and it is generalizable to different characters.

A video data set has been generated with thirteen expressions involving laughter or smile, which follows the presented taxonomy.

For the generation of the automatic animation of the facial rig controllers, we have provided a facial motion capture-based system, which converts these data into information for the animation rig. The integration of this information into the animation environment allows it to be further refined and modified by the animator. We have presented a rigging system for control of geometric interpolation of the twenty-six polygonal surfaces, which correspond to the muscle actions and are capable of being combined to form AUs, as described by Ekman and Friesen in the work on FACS [7].

The method followed to generate facial animation has been extensively applied to our data set, to choose the most representative controller's animation curve for each expression. This choice has been made using dynamic time warping and the K-means algorithm. Therefore, our study provides a library of representative animation curves to animate expressions of laughter and smile on virtual characters, improving the efficiency of the animation process.

To validate the resulting virtual character animations, we have conducted a test on rating of the quality of the motion and of the expression representativeness. The obtained results validate our proposal.

From the presented framework, different lines of investigation have been opened, which could define various extensions of the work. First, in order to improve the obtained results, the data set could be extended by recording and analysing the laughter and smile expressions of a greater number of individuals. In addition, the presented framework could be employed to model other facial expressions.

Supporting information

S1 Video. Uproarious laughter.

(M4V)

S2 Video. Laughter.

(M4V)

S3 Video. Smiling open-mouthed.

(M4V)

S4 Video. Smiling closed-mouth.

(M4V)

S5 Video. Stifled smile.

(M4V)

S6 Video. Melancholy smile.

(M4V)

S7 Video. Eager smile.

(M4V)

S8 Video. Ingratiating smile.

(M4V)

S9 Video. Sly smile.

(M4V)

S10 Video. Debauched smile.

(M4V)

S11 Video. Closed-eye (abashed) smile.

(M4V)

S12 Video. False smile.

(M4V)

S13 Video. False laughter.

(M4V)

Author Contributions

Conceptualization: Miquel Mascaró, Francisco J. Serón, Francisco J. Perales, Ramon Mas.

Formal analysis: Miquel Mascaró, Francisco J. Serón, Javier Varona.

Funding acquisition: Francisco J. Serón, Francisco J. Perales, Javier Varona.

Investigation: Miquel Mascaró, Francisco J. Serón, Francisco J. Perales, Javier Varona.

Methodology: Miquel Mascaró, Javier Varona.

Software: Miquel Mascaró, Ramon Mas.

Supervision: Francisco J. Serón, Francisco J. Perales, Ramon Mas.

Validation: Miquel Mascaró, Javier Varona.

Visualization: Miquel Mascaró, Ramon Mas.

Writing – original draft: Miquel Mascaró.

Writing – review & editing: Francisco J. Serón, Francisco J. Perales, Javier Varona, Ramon Mas.

References

1. Duchenne GB. The mechanism of human facial expression. Cambridge university press.; 1990.
2. Darwin C. The expression of the emotions in man and animals. Oxford University Press.; 1988.
3. Provine RR. Laughing, Tickling, and the Evolution of Speech and Self. *Current Directions in Psychological Science*. 2004 Dec; 13(6):215–218. <https://doi.org/10.1111/j.0963-7214.2004.00311.x>
4. Gruner CR. The game of humor: A comprehensive theory of why we laugh. Transaction publishers.; 2000.
5. Morreall J. Taking laughter seriously. Suny Press.; 1983.
6. Ruch W, Ekman P. The expressive pattern of laughter. *Emotion, qualia, and consciousness*. 2001; p.426–443. https://doi.org/10.1142/9789812810687_0033
7. Ekman P, Friesen W. Facial action coding system: a technique for the measurement of facial movement. Palo Alto: Consulting Psychologists.; 1978.
8. Bridgman GB. Bridgman's complete guide to drawing from life. Sterling Publishing Company, Inc.; 2009.
9. Loomis A. Drawing the Head and Hands. London: Titan Books.; 2011.
10. Hogarth B. Drawing the human head. Watson-Guption.; 1989.
11. Faigin G. The artist's complete guide to facial expression. Watson-Guption.; 2012.
12. DiLorenzo PC, Zordan VB, Sanders BL. Laughing out loud: control for modeling anatomically inspired laughter using audio. *ACM Transactions on Graphics (TOG)*. 2008; 27(5):125. <https://doi.org/10.1145/1409060.1409078>
13. Cosker D, Edge J. Laughing, crying, sneezing and yawning: Automatic voice driven animation of non-speech articulations. *Proceedings of Computer Animation and Social Agents, CASA*. 2009; p. 225–234.
14. Griffin HJ, Aung MSH, Romera-Paredes B, McLoughlin C, McKeown G, Curran W, et al. Perception and Automatic Recognition of Laughter from Whole-Body Motion: Continuous and Categorical Perspectives *IEEE Transactions on Affective Computing*, 2015; v 6(2):165–178. <https://doi.org/10.1109/TAFFC.2015.2390627>
15. Niewiadomski R, Pelachaud C. Towards Multimodal Expression of Laughter. In: Nakano Y, Neff M, Paiva A, Walker M, editors. *Intelligent Virtual Agents: 12th International Conference, IVA 2012, Santa Cruz, CA, USA, September, 12-14, 2012*. Proceedings. Berlin, Heidelberg: Springer Berlin Heidelberg. 2012. p. 231–244.
16. Niewiadomski R, Pelachaud C. The effect of wrinkles, presentation mode, and intensity on the perception of facial actions and full-face expressions of laughter *ACM Transactions on Applied Perception*. 2015; 12(1):1–21. <https://doi.org/10.1145/2699255>

17. Niewiadomski R, Ding Y, Mancini M, Pelachaud C, Volpe G, Camurri A. Perception of intensity incongruence in synthesized multimodal expressions of laughter In: 2015 International Conference on Affective Computing and Intelligent Interaction (ACII), 2015, pp. 684–690.
18. Urbain J, Niewiadomski R, Bevacqua E, Dutoit T, Moinet A, Pelachaud C, et al. AVLaughterCycle. *J Multimodal User Interfaces*. 2010; 4(1):47–58. <https://doi.org/10.1007/s12193-010-0053-1>
19. Ding Y, Prepin K, Huang J, Pelachaud C, Artières T. Laughter animation synthesis. In: Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems. International Foundation for Autonomous Agents and Multiagent Systems. 2014; p. 773–780.
20. Ding Y, Huang J, and Pelachaud C. Audio-driven laughter behavior controller *IEEE Transactions on Affective Computing*. 2017; 8(4):546–558. <https://doi.org/10.1109/TAFFC.2017.2754365>
21. Ochs M, Niewiadomski R, Brunet P, Pelachaud C. Smiling virtual agent in social context *Cognitive Processing*. 2012; 13(2):519–532. <https://doi.org/10.1007/s10339-011-0424-x>
22. Mancini M, Biancardi B, Pecune F, Varni G, Ding Y, Pelachaud C, et al. Implementing and evaluating a laughing virtual character *ACM Transactions on Internet Technology*. 2017; 17(1):1–22. <https://doi.org/10.1145/2998571>
23. Ekman P, Friesen WV. Felt, False, and Miserable Smiles. *Journal of Nonverbal Behavior*. 1982; 6(4):252. <https://doi.org/10.1007/BF00987191>
24. Ekman P, Friesen WV, O'Sullivan M. Smiles When Lying. *J Pers Soc Psychol*. 1988 Mar; 54(3):414–20. <https://doi.org/10.1037/0022-3514.54.3.414>
25. Kawakami K, Takai-Kawakami K, Tomonaga M, Suzuki J, Kusaka T, Okai T. Origins of smile and laughter: a preliminary study. *Early Human Development*. 2006; 82(1):6.
26. Hofmann J, Platt T, and Ruch W. Laughter and Smiling in 16 Positive Emotions *IEEE Transactions on Affective Computing*. 2017; 8(4):495–507. <https://doi.org/10.1109/TAFFC.2017.2737000>
27. Unay J, Grossman R. Hyper-real advanced facial blendshape techniques and tools for production. In: *ACM SIGGRAPH*; 2005. p. 113–120.
28. Alexander O, Rogers M, Lambeth W, Chiang M, Debevec P. The Digital Emily project: photoreal facial modeling and animation. In: *ACM SIGGRAPH 2009 Courses*. ACM; 2009.
29. Digital-Tutors. Facial Rigging in Maya; 2009. Available from: <http://www.digitaltutors.com>.
30. Saragih J, Lucey S, Cohn J. Deformable Model Fitting by Regularized Landmark Mean-Shift. *International Journal of Computer Vision*. 2011; 91(2):200–215. <https://doi.org/10.1007/s11263-010-0380-4>
31. Pantic M, Patras I. Detecting facial actions and their temporal segments in nearly frontal-view face image sequences. In: *IEEE International Conference on Systems, Man and Cybernetics*, 2005;4:3358–3363.
32. Lucey P, Cohn JF, Kanade T, Saragih J, Ambadar Z, Matthews I. The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*; 2010. p. 94 –101.
33. Pantic M, Valstar M, Rademaker R, Maat L. Web-based database for facial expression analysis. In: *IEEE International Conference on Multimedia and Expo*, 2005.
34. BBC—Science & Nature—Human Body and Mind—Spot The Fake Smile; 2012. Available from: <http://www.bbc.co.uk/science/humanbody/mind/surveys/smiles/> [cited 2012-11-23].
35. Dibeklioğlu H, Salah A, Gevers T. Are You Really Smiling at Me? Spontaneous versus Posed Enjoyment Smiles. In: Fitzgibbon A, Lazebnik S, Perona P, Sato Y, Schmid C, editors. *Computer Vision—ECCV 2012*. vol. 7574 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg; 2012. p. 525–538.
36. Petridis S, Martinez B, Pantic M. The MAHNOB Laughter Database. *Image and Vision Computing Journal*. 2013; 31(2):186–202. <https://doi.org/10.1016/j.imavis.2012.08.014>
37. El Haddad K, Torre I, Gilmartin E, Çakmak H, Dupont S, Dutoit T, et al. Introducing amus: The amused speech database In: *International Conference on Statistical Language and Speech Processing*, 2017, pp. 229–240.
38. Jansen MP, Truong KP, Heylen DK, Nazareth DS. Introducing MULAI: A Multimodal Database of Laughter during Dyadic Interactions In: *Proceedings of The 12th Language Resources and Evaluation Conference*, 2020, pp. 4333–4342.