Spatial and Temporal Linearities in Posed and Spontaneous Smiles

LAURA C. TRUTOIU, ELIZABETH J. CARTER, and NANCY POLLARD, Carnegie Mellon University
JEFFREY F. COHN, University of Pittsburgh
JESSICA K. HODGINS, Carnegie Mellon University

Creating facial animations that convey an animator's intent is a difficult task because animation techniques are necessarily an approximation of the subtle motion of the face. Some animation techniques may result in linearization of the motion of vertices in space (blendshapes, for example), and other, simpler techniques may result in linearization of the motion in time. In this article, we consider the problem of animating smiles and explore how these simplifications in space and time affect the perceived genuineness of smiles. We create realistic animations of spontaneous and posed smiles from high-resolution motion capture data for two computer-generated characters. The motion capture data is processed to linearize the spatial or temporal properties of the original animation. Through perceptual experiments, we evaluate the genuineness of the resulting smiles. Both space and time impact the perceived genuineness. We also investigate the effect of head motion in the perception of smiles and show similar results for the impact of linearization on animations with and without head motion. Our results indicate that spontaneous smiles are more heavily affected by linearizing the spatial and temporal properties than posed smiles. Moreover, the spontaneous smiles were more affected by temporal linearization than spatial linearization. Our results are in accordance with previous research on linearities in facial animation and allow us to conclude that a model of smiles must include a nonlinear model of velocities.

Categories and Subject Descriptors: I.3.7 [**Computer Graphics**]: Three-Dimensional Graphics and Realism

General Terms: Facial Animation, Realism

Additional Key Words and Phrases: Smile animation

ACM Reference Format:

Laura C. Trutoiu, Elizabeth J. Carter, Nancy Pollard, Jeffrey F. Cohn, and Jessica K. Hodgins. 2014. Spatial and temporal linearities in posed and spontaneous smiles. ACM Trans. Appl. Percept. 11, 3, Article 12 (August 2014), 15 pages.
DOI: http://dx.doi.org/10.1145/2641569

1. INTRODUCTION

Smiles are likely the most common and easily recognized facial expressions. Yet, depending on the context or emotion conveyed, smiles vary dramatically in terms of geometric appearance and dynamics. This variation is used to convey subtle nuances of emotion and expression. Over 18 labels, including

This study was funded by the NSF and Disney Research, Pittsburgh.
Author's address: L. C. Trutoiu, E. J. Carter, N. Pollard, J. F. Cohn, and J. K. Hodgins, Robotics Institute, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213; emails: ltrutoiu@cs.cmu.edu; ejcarter@andrew.cmu.edu; nsp@cs.cmu.edu; jeffcohn@cs.cmu.edu; jkh@cs.cmu.edu.

DOI: http://dx.doi.org/10.1145/2641569

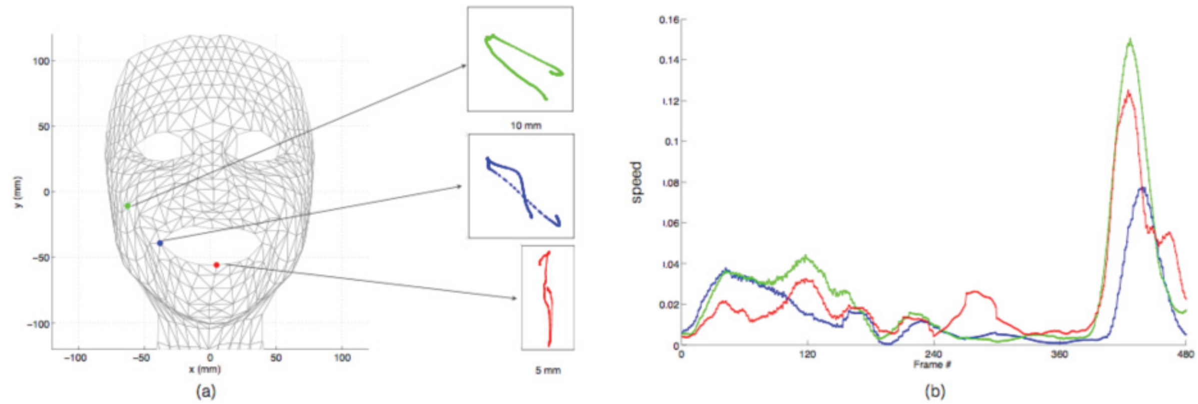

Fig. 1. Spatial and temporal nonlinearities during a spontaneous smile. (a) Spatial nonlinearities represented by the nonlinear geometric paths of three vertices during the smile. (b) Temporal nonlinearities illustrated by the speed of the three vertices during a 4-second smile. The smile is not symmetric: the speed profiles at the start (frames 1 to 180) and end (frames 300 to 480) of the smile are different.

polite, amused, embarrassed, and fearful, have been used to describe smiles and how viewers perceive and interpret them [Ekman 2001; Cohn and Schmidt 2004; Ambadar et al. 2009]. One important characteristic of smiles is genuineness. Genuine smiles are recognized as expressing positive emotions across cultures [Ekman and Friesen 1982]. In contrast, nongenuine or posed smiles are intended to mask true emotions. In this article, we investigate how animation techniques affect the perceived genuineness of smiles. In particular, we explore the impact of linearization in time and space.

In animation, facial expressions, including smiles, can be created by specifying the deformations of the face, represented as a vertex mesh, over time. Thus, a smile has two components: (i) the spatial, or the geometric path of the vertices; and (ii) the temporal, or the rate of change in the vertex position (vertex speed). High-resolution motion capture data shows that these deformations are complex and nonlinear in both space and time. The geometric path of the vertices is nonlinear and the vertex speed is not constant (Figure 1).

Common animation techniques, which rely on keyframing, approximate the temporal or spatial properties of a smile. Craft books often describe smiles in terms of combinations of basic blendshapes [Kalwick 2006]. The simplest model for a smile animation consists of two blendshapes (a neutral pose and a peak smile pose) and an interpolation function. The interpolation between two blendshapes results in a linear geometric path of the vertices, while the speed of vertices is determined by the interpolation function. The choice for the interpolation function has important consequences. If a linear interpolation function is used, the motion may look mechanical and unrealistic because of constant vertex speed. As in the case of eye blinks [Trutoiu et al. 2011], the two-blendshape model with a linear interpolation function may be perceived differently than animations with data-driven, or nonlinear, interpolation functions.

We evaluate the perceptual benefit of preserving data-derived motion characteristics (geometric path, interpolation function) for realistic smile animation. Our perceptual results show how approximations in the temporal or spatial characteristics of the data affect the genuineness of a smile expression (Figure 2). We compare smiles animated with motion capture with smiles in which the geometric path of the vertices is linear or the interpolation function is linear. We find that linearizations lead to smiles being perceived as less genuine. Additionally, we find similar results for animations with or without head motion. We contribute to previous knowledge by disentangling the effects of spatial

(a)

(b)

Fig. 2. (a) A spontaneous and (b) a posed smile animated from motion capture data. The posed smile is rated as significantly less genuine than the spontaneous smile. Linearizations in time and space reduce the perceived genuineness of the spontaneous smile to the level of the posed smile.

properties (geometric path) and temporal properties (interpolation function). Furthermore, animators will benefit from knowing how to avoid creating fake smiles, or, conversely, knowing exactly what parameters cause a smile to look posed.

2. RELATED WORK

Our study focuses on examples of smile animations because smiles are complex yet common expressions. In this section, we discuss research related to smile perception and linearities in facial animation.

Smile genuineness is often associated with a slight wrinkling on the outer corner of the eyes known as the *Duchenne marker* [Ekman et al. 1990]. In a more recent study, however, Krumhuber and Manstead [2009] discovered that over 80% of participants could pose in photographs with smiles that included the Duchenne marker. Furthermore, when viewing static pictures of smiles, volunteers similarly rated both posed and spontaneous smiles as genuine. Conversely, participants recognized posed smiles more often in videos, which may indicate that the timing of different facial actions is relevant.

Multiple studies have analyzed the temporal properties of spontaneous and posed smiles. For example, spontaneous smiles have been found to have smaller amplitude and slower onset than posed smiles [Cohn and Schmidt 2004; Schmidt et al. 2006]. When examining computer-generated smiles, Krumhuber and Kappas [2005] found that perceived smile genuineness increased as a function of onset and offset durations and decreased as a function of apex duration.

Temporal and geometric cues also affect the perceived meaning of smiles in more subtle ways. Ambadar et al. [2009] analyzed and annotated short movies of smile sequences using the Facial Action Coding System [Ekman and Friesen 1978]. The authors characterized them as amused, embarrassed, nervous, polite, or other. Smile categories were differentiated by temporal cues, including duration, onset velocity, offset velocity, asymmetry of velocity, and head movement. For example, amused smiles had larger maximum velocities and longer durations than polite smiles.

Several studies have begun to establish the importance of nonlinear motion for facial animation. For example, nonlinear temporal and geometric motion can affect the accuracy of emotion recognition [Wallraven et al. 2008]. Wallraven [2008] created animations for seven posed expressions, including happiness, using either ground-truth data or linear interpolation between two blendshapes (peak and

Fig. 3. Setup for recording smiles: facial motion was recorded with a commercial motion capture system that tracks the position of 250 3mm markers on the face.

neutral). Linear interpolation in this case created animations with both a linear geometric path and linear timing (constant speed). Viewers were better able to recognize emotions conveyed in the animations created with the ground-truth data than in those with linear interpolation.

Cosker et al. [2010] similarly showed that the originally recorded motion of short facial movements is preferred to linearly interpolated motions. The researchers captured posed expressions using dynamic 3D scanning. Then, participants viewed animations made from the nongeometric recorded data as well as animations created using linear geometric movement between blendshapes. Viewers generally preferred the nonlinear geometric movement and rated it as more natural than the linear movement; however, this was not true for posed smiles [Cosker et al. 2010].

Liu et al. [2011] considered spatial nonlinearities for a discrete set of points on the face (chin and eyelids). They compared linear and nonlinear geometric paths for these points and determined that nonlinear geometric paths were rated as more realistic.

Our study contributes to the existing work in several ways. We consider the effect of linearization for both posed and spontaneous smiles, whereas previous research has focused on posed expressions. We expect that spontaneous expressions have greater nonlinear motion and will, therefore, be more heavily impacted by linearization. Furthermore, animations with a nonlinear geometric path while the timing information is linear have not been previously investigated. We investigate linearizations in both space and time, and we hypothesize that these linearizations will decrease the perceived smile genuineness.

3. APPROACH

In this section, we discuss our approach to capturing and processing our high-resolution dataset. We also describe the animation process.

3.1 Performance Capture

We recorded over 100 smile sequences from two participants (one male, one female) during 3-hour recording sessions. To elicit smiles, we asked participants to (a) view amusing videos, (b) rate one-panel comic captions, and (c) perform smiles according to the experimenter's instructions.

Facial expressions were captured with an 18-camera Vicon system by recording the 3D positions of markers at 120 frames per second (fps); torso and head motion were also recorded. Participants wore 250 reflective markers spaced approximately 1cm apart on the face, as shown in Figure 3. We applied the reflective markers in a similar pattern for both participants. However, due to differences in their facial geometry, the marker positions were not identical or in direct correspondence. In the following section, we describe how the raw motion capture data was processed to obtain 3D meshes deforming over time, with the vertices in correspondence, for the two participants.

3.2 Data Processing

The motion capture system records the 3D position of markers on a frame-by-frame basis. Because of marker density, the motion capture system does not generate consistent marker labels throughout a sequence. Furthermore, physical marker positions and distribution vary across participants. The goal of the data processing step is to obtain smile sequences with the markers for each participant in direct correspondence. To achieve this goal, we first cleaned all the motion capture data so that all markers were present in each frame. We then standardized the meshes for all sequences.

To clean the high-resolution motion capture data, we used the method and semiautomatic tool developed by Akhter et al. [2012]. Their algorithm uses a bilinear spatiotemporal data representation and Expectation Maximization to simultaneously label, denoise, and compute missing points in motion capture data.

The configuration of facial markers differed significantly between participants. Additionally, during the session, some markers were displaced from their positions at the start of recording. Following the approach of Tena et al. [2011], we fit a dense 3D generic mesh template with more than 8,000 vertices to our entire motion capture database. The mesh was subsampled to a limited number of vertices (approximately 400), which resulted in 3D meshes in full direct dense correspondence. Our goal was to analyze and animate facial expressions independently from head motion. For each sequence, rigid-body transformations, such as head motion, were removed by aligning each motion capture frame to the subsampled generic mesh template using ordinary procrustes analysis [Dryden and Mardia 2002].

3.3 Original Smile Sequences

In this article, we used 12 basic smile expressions from two participants (six from each participant). To select the smile samples, we first ordered the smile videos for each participant according to duration. Each video contained at least one smile as determined by visual inspection. Expressions of fewer than 10 seconds were selected for further annotation. From these short smile videos, we picked three spontaneous smile sequences and three posed smile sequences such that the smiles started and ended in a neutral expression. The start, end, and peak frames of the smile expression were identified based on the velocity of the vertices on the cheeks.

We evaluated the genuineness of the 12 selected smiles in a brief experiment on *Amazon Mechanical Turk*. Thirty participants rated each smile video twice on a scale from 0 (not genuine) to 100 (genuine). The videos were recorded during the motion capture process and thus the actors KB and SD were wearing motion capture markers similar to the actor shown in Figure 3.

As expected, posed smiles (average rating =27) were rated as less genuine than spontaneous smiles (average rating =70). The ratings for the two characters are shown in Figure 4. We suspect that the ratings for spontaneous smiles do not reach an average closer to 100 because the video sequences selected are relatively simple. Based on preliminary testing of other video sequences, we posit that higher ratings of genuineness are generally associated with sequences closer resembling laughter.

3.4 Animation

A professional artist created virtual replicas (CG) of the male (KB) and female (SD) actors from the motion capture sessions (Figure 5). Using a series of photographs for reference, the artist matched the geometric shape of the actors' faces to the CG characters. The photographs also provided a high-resolution texture for the characters.

In our framework, a smile sequence is parametrized as a matrix S of m (markers)$\times$ $3F$ (frames: for each motion axis x, y, and z). Using the 3D modeling and animation software *Autodesk Maya*, we created animations based on the motion of the markers, quantified by the matrix S. Spheres

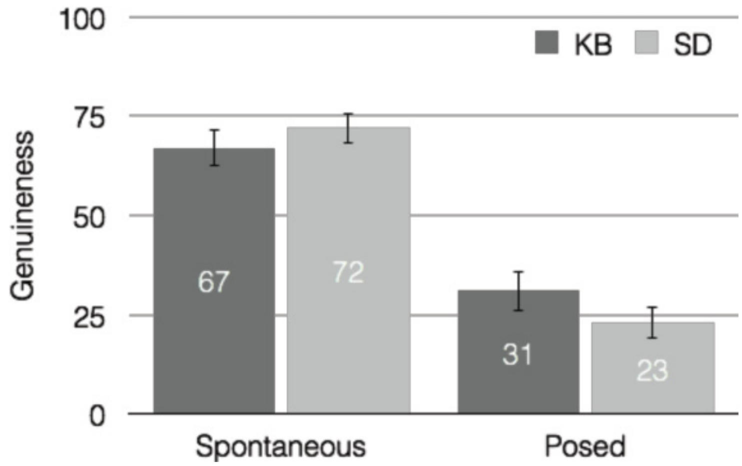

Fig. 4. Average genuineness ratings for the smile videos selected for animation. Three smiles for each actor for each category (posed or spontaneous) were rated by 30 participants. KB is the male actor and SD the female actor. The values are plotted with standard error bars.

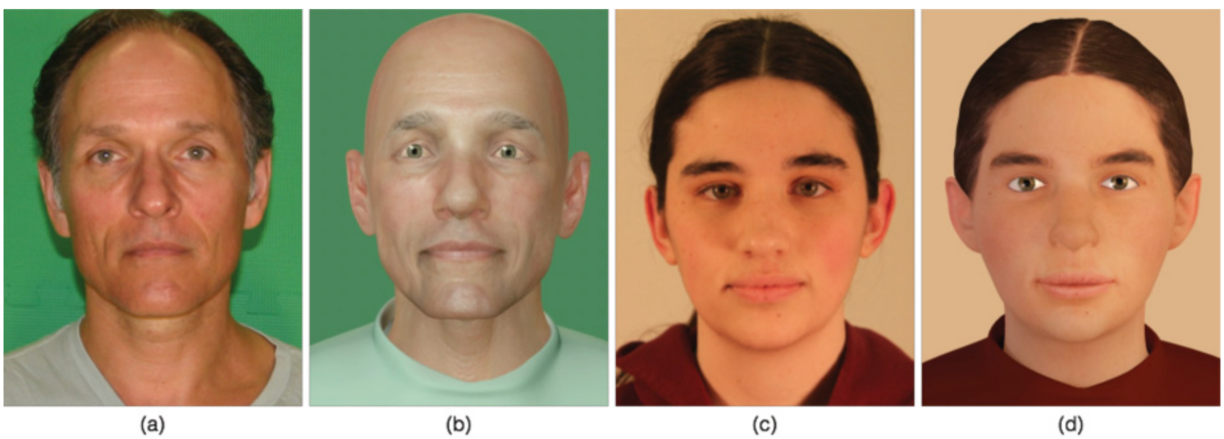

Fig. 5. Actors whose smiles were recorded and their CG character counterparts used for the perceptual experiments: (a) KB, the male actor, (b) KB's CG character, (c) SD, the female actor, and (d) SD's CG character.

corresponding to the marker positions in the neutral expression are used as influence binds on the 3D mesh of the character to be animated. The spheres deform the skin surface by influencing position attributes (translation) of nearby mesh vertices. Influence objects deform smooth skin objects in the same manner that joints can influence smooth skin objects. Virtual markers, controlled by the matrix S, are parented to the spheres such that the position of the markers over time deforms the mesh accordingly. We linearized space and time to create the matrix S for the different animation conditions described in Section 4.1.

4. EXPERIMENT 1: LINEARIZED ANIMATIONS WITH HEAD MOTION

The goal of this first experiment was to determine how linearization impacts the perceived genuineness of smiles. We chose genuineness as the dependent variable because of its usefulness in animation.

We examined the perception of four animation conditions resulting from linearizing time or space. We used 12 recorded smiles for two CG characters with an equal number of posed and spontaneous smiles (three of each type per each character). Animations were presented with the originally recorded head and torso motion. We used a within-subject repeated-measures experimental design.

We collected genuineness ratings of animated smiles through controlled experiments on Amazon Mechanical Turk. Fifty-seven viewers successfully rated 48 animations of the two CG characters. Viewers were at least 18 years old and located in the United States. The animations were displayed in a randomized order to control for possible order effects. After viewing each animation, participants were asked to rate the smile on a scale from 0 (not genuine) to 100 (genuine). Similar to the method used by Krumhuber and Kappas [2005], we described a genuine smile as a smile that someone shows when she/he is joyful, happy, or amused. Though the characters in the clips are computer-generated and, therefore, not truly happy, we explained that the animations reflect certain aspects of smiles that differ in real people. The experiment took no longer than 20 minutes.

We focus on two independent variables that can be linearized: the geometric path of the vertices and the interpolation function. Interactions between these variables result in the animation conditions described in the following text. Additional independent variables are the CG character used (female or male), the type of smile (posed or spontaneous), and the smile sample. We used a total of 12 smile sequences: three smiles for each type for each character.

4.1 Animation Technique

The animation conditions result from the combination of two independent variables:

—Spatial, determined by the geometric path of the vertices as data-derived (SN) or linear (SL)
—Temporal, determined by the interpolation function as data-derived (TN) or linear (TL).

For a vertex i, we define its position at time t as $V_i(t)$ where t ranges from 1, the first frame of the smile, to p, the peak displacement of the smile relative to frame 1, and n, the end of the smile. For a data-derived geometric path (SN), V_i is directly recorded from the motion capture data. For the linear geometric path (SL), we define a piece-wise linear function composed of the linear path of vertex V_i between the position at frame 1 and the peak frame p and the linear path of vertex V_i between the position at frame p and the end frame n.

We next describe each of the four animation conditions, focusing on deriving the position of vertices for the first part of the smile, from frames 1 to p. Similar computations are defined for frames $p+1$ to n. Visual representations of the vertex path and the interpolation functions used are shown in Figures 6 and 7.

SL-TL: A linear geometric path for vertices $V_i t$ between 1 and p moving with constant speed based on *linear interpolation* is computed as

$$V_i(t) = V_i(1) + \frac{t-1}{p-1} * [V_i(p) - V_i(1)]. \tag{1}$$

This condition is equivalent to using two blendshapes (the neutral and the peak frame) with a linear interpolation between them.

SL-TN: For a data-derived speed and linear geometric path for the vertices, the position of vertices V_i at time t is computed as

$$V_i(t) = V_i(1) + rc(t) * [V_i(p) - V_i(1)], \tag{2}$$

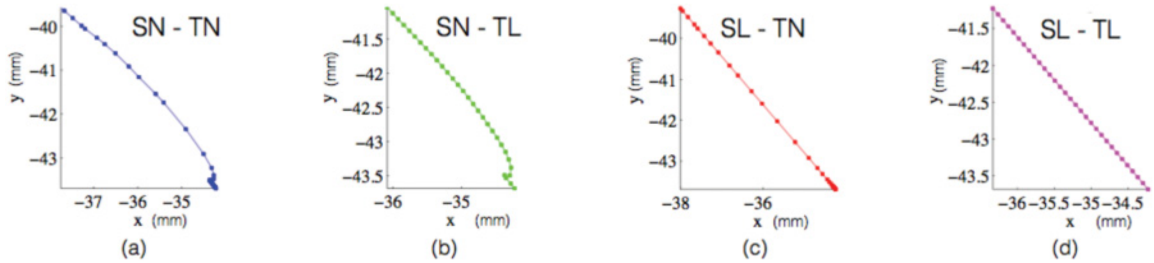

Fig. 6. Short sequence for the geometric path of the right mouth corner vertex for all animation conditions. The frequency of the dots reflects the interpolation function used. (a) SN-TN: data-derived spatial path with data-derived interpolation function. (b) SN-TL: data-derived spatial path with linear interpolation function. (c) SL-TN: linear spatial path with data-derived interpolation function. (d) SL-TL: linear spatial path with linear interpolation function.

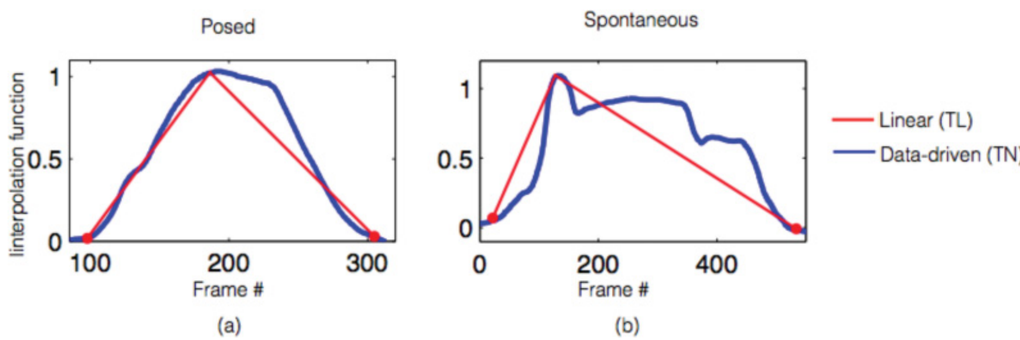

Fig. 7. Data-driven (TN condition) and linear (TL condition) interpolation functions for (a) a posed smile and (b) a spontaneous smile at 120fps.

where $rc(t)$ is the reconstruction coefficient computed as the least-squares solution to minimize

$$||V(t) - V(1) + rc(t) * [V(p) - V(1)]||. \tag{3}$$

Note that at every frame t, the positions of vertices are computed based on frames 1 and p and the sequence follows a linear path between 1 and p. This condition corresponds to using two blendshapes (the neutral and the peak frame) with a data-based interpolation function. The data-derived reconstruction coefficient may have values outside of the [0 : 1] interval, resulting in the point moving forward and then backward along the same path. This effect can be seen in smiles when the smile is released and a small lip adjustment occurs: the lips are pressed together during the release and then relax in a natural position.

SN-TL: The vertices V_i move with constant speed across a data-derived geometric path. We first compute *PathLength*, the length of the path traversed by each vertex V_i from frame 1 to p. The position of each vertex V_i at time t is computed iteratively such that

$$V_i(t+1) - V_i(t) = \frac{PathLength}{p-1}. \tag{4}$$

SN-TN: The vertices have a data-derived geometric path and speed directly based on the recorded motion capture data. This condition is considered to be the ground-truth animation, closest in naturalness to the video.

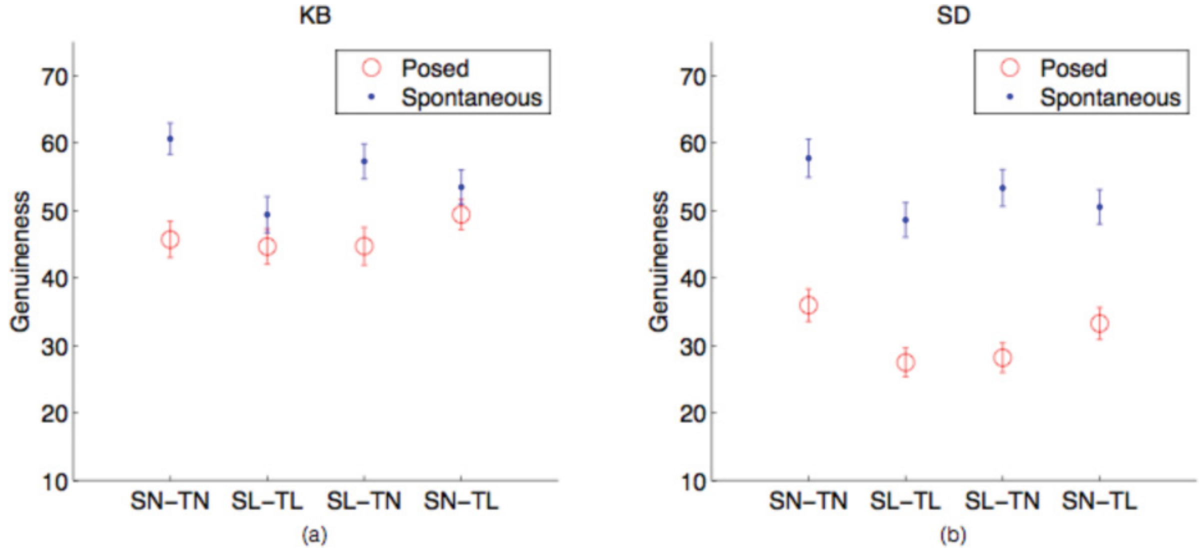

Fig. 8. (a) Genuineness ratings for the KB character smiles. (b) Genuineness rating for the SD character smiles. The values are plotted with standard error bars.

Trutoiu et al. [2013] showed that smiles have a characteristic blink placement relative to the smile start and end. We, therefore, considered it important to add eye blinks to the animations at the same frame locations as in the video sequence.

4.2 Results

We conducted a repeated-measures ANOVA to analyze the possible effects of linearization on genuineness ratings. We found that linearizations in either time or space influence the perceived genuineness of smiles.

Our hypothesis that using a linear interpolation function reduces the perceived genuineness of smiles was supported. Animations where the vertices move in a linear geometric path ($SL = 44.21$) are rated significantly lower than animations where the vertices follow the original path ($SN = 48.34$), $F(1{,}840) = 21.96$, and $p < 0.0001$. Similarly, animations where the velocity of vertices is constant ($TL = 44.60$) are rated lower than animations in which the vertices follow the temporal profile of the original data ($TN = 47.95$), $F(1{,}840) = 14.43$, and $p = 0.0002$.

Significant interactions occur between the type of smile and the temporal conditions and between the type of smile and the CG character. We investigated these interactions in more detail with post hoc contrast tests. A constant velocity for the vertices (TL) impacts spontaneous ($p < 0.0001$) but not posed smiles.

We also explored how the space and time parameters interact and impact genuineness for spontaneous smiles. We analyzed the differences in spontaneous smiles and found that the original animations are not significantly different than the animations in which space is linearized while temporal information is preserved. In contrast, the original animations are significantly more genuine than animations in which the geometric path of the vertices is preserved while the speed of the deformation is constant. The overall interactions between space and time variables are shown in Figure 8.

As expected, we also found that spontaneous smiles are rated as more genuine than posed smiles, with averages of 53.87 and 38.63, respectively, $F(1{,}840) = 297.58$, $p < 0.0001$. The two CG characters were rated in significantly different ways: KB averaged ratings of 50.66, while SD averaged 41.89, $F(1{,}840) = 99.06$, and $p < 0.0001$. It is worth noting that the ratings for the two CG characters do not differ for spontaneous smiles, while for posed smiles they are significantly different, with SD posed smiles being rated as the lowest, average of 31.23.

It is interesting to contrast the ratings for animated smiles with the original data (condition SN-TN in Figure 8) with their corresponding video ratings (Figure 4). The animated spontaneous smiles had lower ratings than their video counterparts, while the reverse was true for posed smiles. We suspect that presenting the smiles in animation makes it more difficult for participants to distinguish between posed and spontaneous smiles. Regardless, the animations preserved the ordering between posed and spontaneous smiles.

One potential confound of this study is that modifications in the geometric path of vertices or their speed may be desynchronizing the facial expressions from the head motions. To overcome this confound, we conducted a second experiment in which animations were displayed without head motion.

5. EXPERIMENT 2: LINEARIZED ANIMATIONS WITHOUT HEAD MOTION

In a second experiment, we explored the effect of linearization on genuineness in clips without head motion. The procedures and measures were in all other respects identical to those used in Experiment 1. Sixty-one participants successfully took part in this study.

5.1 Results

The results of this experiment are in most respects similar to those from Experiment 1. Linearizing either the space or interpolation function resulted in significantly lower genuineness ratings. Animations with a linear interpolation function ($TL = 42.86$) were rated lower than animations with nonlinear interpolation functions ($TN = 47.65$), $F(1{,}859.9) = 19.85$, and $p < 0.0001$. Animations with a linear geometric path ($SL = 43.38$) were rated lower than animations with nonlinear geometric path ($SN = 47.13$), $F(1{,}859.9) = 32.44$, and $p < 0.0001$.

For spontaneous smiles using a data-driven interpolation function and a linear geometric path (SL-TN = 51.35), ratings were not significantly different than the original animations (SN-TN = 53.71), $F(1{,}859.9) = 2.11$, and $p = 0.146$. The interactions between the spatial and temporal independent variables are shown in Figure 9. Interestingly, all of the SD animations without head motion, including those of spontaneous smiles, were rated as less genuine than even the posed KB animations without head motion, which was not the case in Experiment 1.

Overall, the ratings of animations without head motions were lower (mean = 45.6) than the ratings with head motions (mean = 46.2) in Experiment 1. A further study is needed to statistically compare the animations with and without head motions.

6. QUANTITATIVE DIFFERENCES BETWEEN POSED AND SPONTANEOUS SMILES

As shown in the previous experiments, linearizing space or time decreases genuineness more for spontaneous smiles than for posed smiles. In this section, we quantify the differences between posed and spontaneous smiles for the following characteristics: duration, spatial nonlinearity, and mouth corner vertex speed. The quantified variables are shown in Table I.

In our study, the six spontaneous smiles were longer (average duration of 4.9 seconds) than posed smiles (average duration of 2.9 seconds). These differences in duration are consistent with previous research. Ambadar et al. [2009] found that, on average, perceived amused smiles lasted about 4 seconds, whereas perceived polite or embarrassed/nervous smiles lasted for 2 seconds and 2.9 seconds, respectively. In our case, perceived amused smiles equate with spontaneous smiles (which were rated as highly genuine); posed smiles, which lack genuineness, likely include both polite and embarrassed

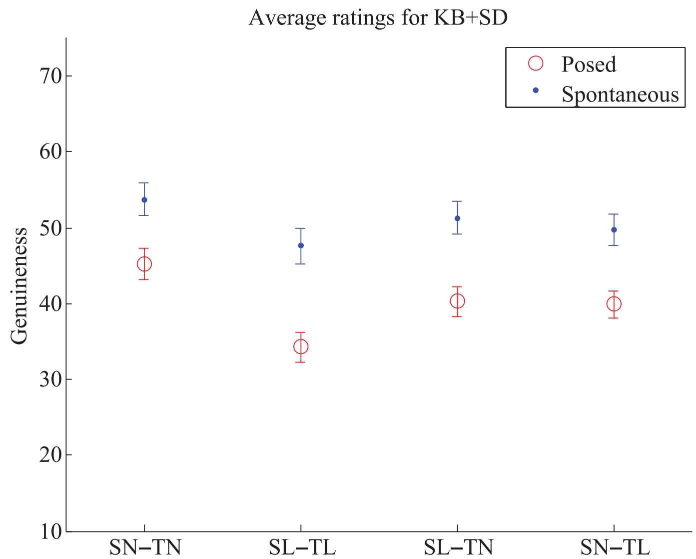

Fig. 9. Average genuineness ratings for animations without head motion. The values are plotted with standard error bars.

Table I. Differences between Posed and Spontaneous Smiles Quantified in Terms of Duration, Nonlinearity, and Mouth Corner Vertex Speed

	Sample	Duration (sec)	Spatial nonlinearity (mm)	Right mouth corner vertex speed (mm/sec)	Left mouth corner vertex speed (mm/sec)	Average mouth corner vertex speed (mm/sec)	Difference in mouth corner Speed (L-R) (mm/sec)
KB posed	KB_p1	3.68	0.62	6.89	7.99	7.44	1.10
	KB_p2	3.45	0.60	7.38	9.33	8.36	1.96
	KB_p3	3.96	0.62	6.98	8.13	7.55	1.15
	average	3.70	0.61	7.08	8.49	7.78	1.40
KB spont	KB_s1	4.71	0.85	4.61	6.72	5.66	2.10
	KB_s2	4.09	0.65	3.68	5.26	4.47	1.59
	KB_s3	5.70	0.64	2.44	3.83	3.14	1.40
	average	4.83	0.71	3.57	5.27	4.42	1.70
SD posed	SD_p1	2.10	0.46	4.03	3.73	3.88	−0.30
	SD_p2	1.70	0.08	0.78	1.06	0.92	0.28
	SD_p3	2.10	0.44	3.85	5.23	4.54	1.38
	average	1.97	0.33	2.89	3.34	3.11	0.46
SD spont	SD_s1	5.00	0.49	3.89	4.41	4.15	0.52
	SD_s2	5.25	0.58	3.22	4.50	3.86	1.28
	SD_s3	5.17	0.32	3.13	3.94	3.53	0.81
	average	5.14	0.46	3.41	4.28	3.85	0.87

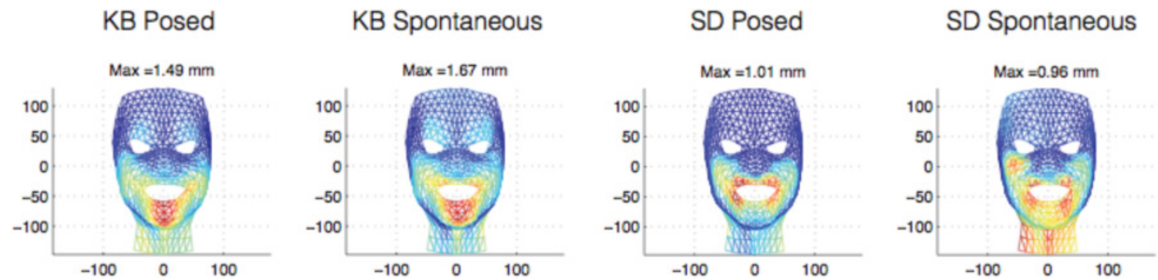

Fig. 10. Visual representation of the difference between the original nonlinear animation and its spatially linearized counterpart. The differences were averaged across the three smile samples. In each figure, the vertex with the maximum Euclidean distance is represented in dark red and noted in the title.

smiles. The durations for the 12 smiles used in the perceptual experiments are reported in the first column of Table I.

Linearizing space impacted more spontaneous smiles than posed smiles. We hypothesized that spontaneous expressions are more complex and, therefore, more nonlinear. To quantify spatial nonlinearity, we computed the differences between the original and the linearized geometric path. We used a Euclidean distance-based measure as proposed by Cosker et al. [2010]. For each frame, we computed the Euclidean distance between the vertices in the original animation and their linear counterparts. The second column of Table I shows the average nonlinearity measure per vertex normalized by smile duration. As expected, spontaneous smiles had more nonlinear motion: the Euclidean distance is larger for spontaneous (average of 0.59mm) than for posed (average of 0.47mm) smiles. Note that though these values are small, they are averaged over 430 vertices. In our perceptual experiments, spatial linearization did not impact posed smiles as strongly as it impacted spontaneous smiles.

In Figure 10, the nonlinear motion of spontaneous smiles appears to be more diffuse on the face. Different patterns are visible for the two actors. For example, KB, the male actor, shows nonlinear motion in the eyebrow and lower jaw region. In contrast, SD, the female actor, shows more nonlinear motion in the mouth corner region.

We considered the differences in the vertex speed for the two mouth corner vertices. There is evidence that posed and spontaneous smiles differ in vertex speed and symmetry. For posed expressions, Schmidt et al. [2006] showed that movement asymmetry (measured by change in pixel values over time) was significant for expressions of joy, including smiles, with more movement on the left side of the face. Our results similarly show that the left mouth corner speed was consistently higher than the right right mouth corner speed (Table I, last column).

Ekman and Friesen [1982] first posited that spontaneous smiles are more symmetric than posed smiles. However, their study did not quantitatively assess the movement asymmetry. In our analysis, for the KB actor, posed smiles had larger mouth corner vertex speeds than spontaneous smiles. However, the opposite was true for the SD actor. In future work, we intend to analyze a larger dataset of posed and spontaneous expressions across a broader pool of subjects to quantify motion asymmetry.

A representative example of vertex speed over time for posed and spontaneous smiles in shown in Figure 11. In previous research, viewers associated irregularities in the offset of smile, as measured by the frequency of phasic change, with posed smiles [Hess and Kleck 1994]. However, in our examples, spontaneous smiles showed more changes in velocity.

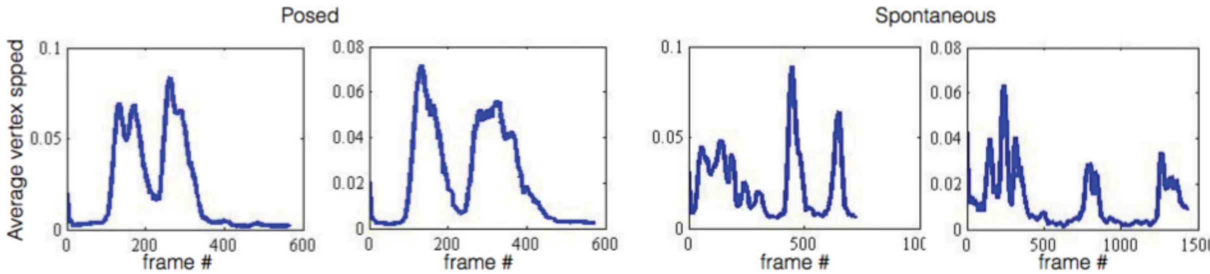

Fig. 11. Average vertex speed for four smiles computed at 120fps. Spontaneous smiles have more peaks compared to posed smiles. The first and last peaks correspond to the onset and offset of the smile.

7. DISCUSSION

Our experiments reveal that linearizing either space or time decreases the perceived genuineness of smiles. However, spontaneous smiles with a linearized geometric path (space) and a data-driven interpolation function (time) are rated as genuine as the original data. Based on these results, we will investigate a parsimonious model for smiles consisting of data-driven interpolation functions that capture the dynamics of the facial expressions.

The dependent variable in our studies was the perceived genuineness of smiles. Genuineness is a relevant metric in the animation domain and it is useful in distinguishing between posed and spontaneous smiles. Naturalness, another perceptual metric we considered, though valuable in evaluating realistic animations, does not have the same discriminatory quality. That is, posed smiles may be rated as natural because they are plausible expressions that occur naturally.

There are several limitations to our study. We recorded and animated smiles for two CG characters. Genuineness ratings for the two characters differed in their respective animations, with KB's smiles rated more genuine than SD's. Several explanations may account for these differences: intrinsic differences in the smile expressions between the two participants, differences in the quality or rendering style of the CG character, and perceptual differences related to age and gender. For example, the male participant, KB, is a professional actor while the female participant, SD, had no acting experience. This difference in acting experience could potentially explain why KB's posed smiles were rated as more genuine that SD's posed smiles even though their spontaneous smiles had similar values.

The main limiting factor in conducting this kind of study is determined by the availability of high-resolution CG characters. We aimed for our CG characters to be of similar quality. However, previous research has shown that small differences in rendering styles can influence perceptual judgments of CG characters [McDonnell et al. 2012; Hyde et al. 2013]. SD, the female character, had fewer wrinkles, leading to a smoother face appearance. In future work, we will consider conducting a larger study using more smiles from more actors and counterbalancing for age and gender.

We used two types of interpolation functions (data-driven and linear) and found that data-driven interpolation functions are needed for animating genuine smiles. However, many animation techniques use ease-in/ease-out interpolation functions that mimic the effects of acceleration and deceleration seen in physical systems. It may be the case that ease-in/ease-out interpolation functions are sufficient to create genuine smiles. Though more complex than posed smiles, the spontaneous smiles chosen started and ended in neutral expressions and were relatively short. It is possible that with more complex spontaneous smiles, preserving temporal information would be insufficient.

In experiments 1 and 2, participants rated animations with and without head motion, respectively. Our results show that the effect of linearization on smile genuineness is similar in both experiments. However, previous research has shown a correlation between head motion and the dynamics of smiles

[Cohn and Schmidt 2004]. Furthermore, for animations without head motion, SD's animations of spontaneous smiles were rated lower than KB's posed smile animations. This result indicates that the contribution of head motion to the perception of smile genuineness may vary across individuals. Further analysis is required to determine the relationship between head motion and smiles.

Our experiments indicate that some simplifications that occur in traditional blendshape animation may lead to smiles being perceived as posed and less genuine. These results suggest that if animators want to create genuine smiles, they should use nonlinear, preferably data-driven, interpolation functions. Furthermore, our study underlines the importance of using spontaneous rather than posed expressions in studies that quantify facial dynamics.

ACKNOWLEDGMENTS

We thank Moshe Mahler, Rafael Tena, Jim McCann, Jenn Hyde, our actors, and the Graphics Lab interns for their help with this study.

REFERENCES

Ijaz Akhter, Tomas Simon, Sohaib Khan, Iain Matthews, and Yaser Sheikh. 2012. Bilinear spatiotemporal basis models. *ACM Transactions on Graphics* 31, 2 (April 2012), 17:1–17:12.

Zara Ambadar, Jeffrey F. Cohn, and L. I. Reed. 2009. All smiles are not created equal: Morphology and timing of smiles perceived as amused, polite, and embarrassed/nervous. *Journal of Nonverbal Behavior* 33, 1 (2009), 17–34.

Jeffrey F. Cohn and Karen L. Schmidt. 2004. The timing of facial motion in posed and spontaneous smiles. *Journal of Wavelets, Multi-resolution and Information Processing* 2 (2004), 1–12.

Darren Cosker, Eva Krumhuber, and Adrian Hilton. 2010. Perception of linear and nonlinear motion properties using a FACS validated 3D facial model. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*. ACM, 101–108.

Ian L. Dryden and Kanti V. Mardia. 2002. *Statistical Shape Analysis*. John Wiley and Sons.

Paul Ekman. 2001. *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage (Revised and Updated Edition)* (2 rev sub ed.). W. W. Norton & Company.

Paul Ekman, R. J. Davidson, and Wallace Friesen. 1990. The duchenne smile: Emotional expression and brain physiology. II. *Journal of Personality and Social Psychology* 58, 2 (Feb. 1990), 342–353.

Paul Ekman and Wallace Friesen. 1978. *Facial Action Coding System: A Technique for the Measurement of Facial Movement.* Consulting Psychologists Press, Palo Alto, CA.

Paul Ekman and Wallace Friesen. 1982. Felt, false, and miserable smiles. *Journal of Nonverbal Behavior* 6, 4 (1982), 238–252.

Ursula Hess and Robert E. Kleck. 1994. The cues decoders use in attempting to differentiate emotion-elicited and posed facial expressions. *European Journal of Social Psychology* 24, 3 (1994), 367–381. DOI:http://dx.doi.org/10.1002/ejsp.2420240306

J. Hyde, E. J. Carter, S. Kiesler, and J. K. Hodgins. 2013. Perceptual effects of damped and exaggerated facial motion in animated characters. In *Proceedings of the 10th IEEE International Conference on Automatic Face and Gesture Recognition*. DOI:http://dx.doi.org/10.1109/FG.2013.6553775

David J. Kalwick. 2006. *Animating Facial Features and Expressions*. Charles River Media.

Eva Krumhuber and Arvid Kappas. 2005. Moving smiles: The role of dynamic components for the perception of the genuineness of smiles. *Journal of Nonverbal Behavior* 29, 1 (April 2005), 3–24. DOI:http://dx.doi.org/10.1007/s10919-004-0887-x

Eva G. Krumhuber and Antony S. R. Manstead. 2009. Can duchenne smiles be feigned? New evidence on felt and false smiles. *Emotion* 9, 6 (2009), 807–820.

Xuecheng Liu, Shihong Xia, Yiwen Fan, and Zhaoqi Wang. 2011. Exploring non-linear relationship of blendshape facial animation. In *Computer Graphics Forum*, Vol. 30. Wiley Online Library, 1655–1666.

Rachel McDonnell, Martin Breidt, and Heinrich H. Bülthoff. 2012. Render me real?: Investigating the effect of render style on the perception of animated virtual humans. *ACM Transactions on Graphics* 31, 4, Article 91 (July 2012), 11 pages. DOI:http://dx.doi.org/10.1145/2185520.2185587

Karen L. Schmidt, Yanxi Liu, and Jeffrey F. Cohn. 2006. The role of structural facial asymmetry in asymmetry of peak facial expressions. *Laterality* 11, 6 (Nov. 2006), 540–61. DOI:http://dx.doi.org/10.1080/13576500600832758

J. Rafael Tena, Fernando De la Torre, and Iain Matthews. 2011. Interactive region-based linear 3d face models. In *ACM Transactions on Graphics (TOG)*, Vol. 30. ACM, 76.

Laura C. Trutoiu, Elizabeth J. Carter, Iain Matthews, and Jessica K. Hodgins. 2011. Modeling and animating eye blinks. *ACM Transactions on Applied Perception* 8, 3 (Aug. 2011), 1–17.

Laura C. Trutoiu, J. K. Hodgins, and Jeffrey F. Cohn. 2013. The temporal connection between smiles and blinks. In *Proceedings of the 10th IEEE International Conference on Automatic Face and Gesture Recognition*.

Christian Wallraven, Martin Breidt, Douglas W. Cunningham, and Heinrich H. Bülthoff. 2008. Evaluating the perceptual realism of animated facial expressions. *ACM Transactions on Applied Perception* 4, 4 (Feb. 2008), Article 4, 20 pages. DOI:http://dx.doi.org/10.1145/1278760.1278764

Received June 2014; accepted June 2014