

Clinical photography vs digital video clips for the assessment of smile esthetics

Brian J. Schabel^a; Tiziano Baccetti^b; Lorenzo Franchi^b; James A. McNamara, Jr.^c

ABSTRACT

Objective: To test the null hypothesis that there is no clinically significant difference between the post-orthodontic treatment images of smiles of subjects captured by clinical photography and the smiles of the same subjects obtained from digital video clips.

Materials and Methods: Clinical photographs and digital video captures were obtained from 48 orthodontically treated patients. An updated version of the Smile Mesh™ program was used to quantify and compare smile characteristics obtained with the two methods. A paired-samples *t*-test was performed to test for mean differences in Smile Mesh measurements generated from both smile images. The relationship between the various Smile Mesh measurements obtained from both smile images was examined by way of Pearson product-moment correlation.

Results: A significant difference was found between 7 of the 14 mean Smile Mesh measurements. The absolute values of all these differences, however, were smaller than 1 mm and therefore were not clinically significant. With the exception of lower lip to maxillary incisor, all measurements showed a moderate to strong relation with each other (*P* values ranging from .47 to .82; *P* < .001).

Conclusions: The hypothesis cannot be rejected. A significant positive correlation was noted between Smile Mesh measurements obtained from smiles captured by clinical photography and those captured with digital video clips. This supports the conclusion that a standard digital photograph appears to be a valid tool for analysis of the posttreatment smile. (*Angle Orthod.* 2010;80:678–684.)

KEY WORDS: Smile; Digital photograph; Video clip; Orthodontic treatment; Esthetics

INTRODUCTION

In an attempt to meet the ever-increasing esthetic demands of patients, orthodontic researchers have

been prolific in recent years with articles that have examined various aspects of dentofacial esthetics.

Unfortunately, these reports often are contradictory and misleading, in part because of the subjective nature of beauty and the lack of a standardized scale by which to measure it. In addition, the reliability of static photographs for evaluating the smile has been questioned, and digital videography has been advocated for use in capturing the dynamic nature of facial animation with special emphasis on the smile.^{1,2}

Ackerman et al.¹ among others have defined two main types of smiles: social smiles and enjoyment smiles. A social smile is “the voluntary smile a person uses in social settings or when posing for a photograph.” The social smile is “posed,” which means that it is not elicited or accompanied by emotion. This type of smile can be sustained as a static facial expression and does not appear strained.² On the other hand, enjoyment smiles are involuntary and are elicited by laughter. The enjoyment smile is unposed and reflects the emotion that one is experiencing at that moment. This smile appears strained because the mouth bursts

^a Graduate Orthodontic Program, The University of Michigan, Ann Arbor, Mich. Private practice, Santa Cruz, Calif.

^b Assistant Professor, Department of Orthodontics, The University of Florence, Florence, Italy; Thomas M. Graber Visiting Scholar, Department of Orthodontics and Pediatric Dentistry, School of Dentistry, The University of Michigan, Ann Arbor, Mich.

^c Thomas M. and Doris Graber Endowed Professor of Dentistry, Department of Orthodontics and Pediatric Dentistry, School of Dentistry; Professor of Cell and Developmental Biology, School of Medicine; and Research Professor, Center for Human Growth and Development, The University of Michigan, Ann Arbor, Mich. Private practice of orthodontics, Ann Arbor, Mich.

Corresponding author: Dr Tiziano Baccetti, Department of Orthodontics, Università degli Studi di Firenze, Via del Ponte di Mezzo, 46-48, 50127, Firenze, Italy
(e-mail: t.baccetti@odonto.unifi.it)

Accepted: June 2007. Submitted: May 2007.

© 2010 by The EH Angle Education and Research Foundation, Inc.

forward to reveal the maximal expansion of the lips, but it cannot be sustained.

The unstrained social smile has been referred to as a reliable reference for measurement and characterization of the smile.³ Orthodontic records play an essential role in capturing the unstrained social smile to be used for objective analysis. These records must allow us to observe each patient frontally, vertically, obliquely, and from profile, both statically and dynamically, to obtain a true smile representation.^{4,5}

Static records used to capture the smile include study models, radiographs, and film or digital photographs.⁴ The *American Academy of Cosmetic Dentistry Photographic Accreditation Review* in 1995 recommended that facial photographs for esthetic treatment planning should include full face smiling, full face with lips relaxed, profile full smile, and right and left lateral views of full smile.⁴ It is interesting to note that this proposed sequence is advocated for appropriate visualization of even a single restorative unit (tooth), yet the universal orthodontic standard for facial images includes frontal at rest, frontal smile,⁶ and profile at rest.

Digital videography has become an adjunct tool for orthodontic and orthognathic surgery evaluation.^{3,5,7} Video clips taken before, during, and after treatment enable the clinician to observe the dynamic display zone in the frontal view during facial animation; such clips can be used as a means of comparison to assess the effects of treatment and facial change over time. In addition to diagnostic information acquired from dynamic visualization of the smile, video imaging has the potential to affect communication at consultations and at staff meetings, as well as interactions with other offices, and in other areas not yet realized.⁷

Tarantili et al.⁸ have described a progression of the smile using digital video that consists of an initial *attack* period, a sustaining period, and a fade-out or *decay* period. If a clinical photograph is taken during the attack or the decay phase, the resulting smile will not be a reliable reference. For this reason, it is postulated that video may have a distinct advantage over clinical photographs for accurately capturing a true representation of the smile.^{3,8}

To quantify the reliability and reproducibility of the posed smile, Ackerman et al.¹ developed the Smile Mesh™ (TDG Computing, Philadelphia, Pa) program. They reported high interrater and intrarater reliability of the Smile Mesh program and a high correlation coefficient ($r = 0.78$ to 0.99) between repeated measures. They also found smiles in their study to be reproducible.

The aim of the present study was to compare the smiles of subjects after orthodontic treatment when captured by clinical photography vs smiles obtained

from digital video clips. These smiles were quantified with the Smile Mesh program to determine whether these two methods of smile capture differed significantly.

MATERIALS AND METHODS

Patient Selection

Subjects enrolled in this study were recruited from the University of Michigan Graduate Orthodontic Clinic during a routine posttreatment appointment (ie, final records or retention check). Potential subjects were given a brief introduction to the study and were asked if they would be willing to participate. None of the subjects received compensation for their participation.

Each adult subject (ie, 18 years of age or older) reviewed and signed a consent form created in accordance with the rules and regulations of the University of Michigan Health Sciences Institutional Review Board. Each subject younger than 18 years of age reviewed and signed a child's assent form, and a legal guardian reviewed and signed a consent form, in accordance with the Institutional Review Board. Each subject also reviewed and completed a consent form created by the University of Michigan in accordance with the Health Insurance Portability and Accountability Act for the use and disclosure of protected health information.

To be included in the study, subjects had to present with the following characteristics: (1) age ranging from 12 to 20 years; (2) white ancestry; (3) orthodontic treatment completed within the last 6 months; (4) absence of missing or malformed teeth; and (5) a complete set of diagnostic posttreatment records, including intraoral/extraoral photographic series and a good quality video clip of the smile. The protocol proposed for the study required that 48 subjects be recruited to satisfy the design of the Q-sort method. A test was performed to determine the power of this sample size with respect to correlation tests (Type I error = .05). For a bivariate normal distribution and a sample size of 48, a test of $H_0: P = 0$ (ie, the correlation coefficient under the null hypothesis) was found to have a power of 0.80 to detect a linear correlation of $r = 0.38$. Thus, the default sample size for the Q-sort procedure was deemed adequate for purposes of testing for correlation.

Image Capture

Clinical photography. The extraoral photographic series included photographs of the subject in repose, during smiling, and in profile. For the purpose of the current study, only the extraoral smiling photographs

were used. A Canon® EF 35 mm SLR camera (Canon U.S.A., Inc., Lake Success, NY) was mounted to a frame set at a fixed distance of 36 inches between the lens and the subject. The camera was connected to a two-strobe lighting source that illuminated the subject indirectly from a flash that reflected off of a photographic umbrella. All photographs were taken by one of two dental school staff photographers.

Before taking the smiling image, the photographer instructed the subject to “smile.” The reproducibility of the posed smile derived from the static photograph has been demonstrated by Ackerman et al.¹ Each image was captured on Kodak® EV-100 slide film (Eastman Kodak Company, Rochester, NY). The film was developed, and the 2" × 2" slides were used. The slides were scanned using the Nikon® Super Coolscan 4000 ED (Nikon Inc., Melville, NY) and were imported directly into a commercially available image editing software program (Adobe® Photoshop® 7.0, Microsoft Corporation, Redmond, Wash). Each slide was scanned at maximum dpi (dots-per-inch) to enhance image quality.

Digital videography. A digital video camera was used to record the dynamic range of each subject's smile, with slight modifications to the protocol reported by Ackerman and Ackerman.³ To standardize the technique, a Panasonic® PV-GS200 digital video camera (Panasonic, Secaucus, NJ) was used in the same location under standard fluorescent lighting. The camcorder was mounted on an adjustable microphone stand and was set at a fixed distance of 60 inches from the subject. Each subject was seated and had his or her head positioned such that an imaginary line between the top of the ear and the midpoint between the upper eyelash and eyebrow paralleled the floor. The video camera was adjusted vertically to be directly in line with the subject's mouth, and the zoom feature was used to focus only on the mouth and adjacent soft tissues to protect the anonymity of the subject. Each video clip was obtained by the senior author.

Before the video clip was recorded, subjects were given the following instructions:

1. You will be asked to smile and then relax three separate times.
2. When you are asked to relax, please touch your lips lightly together.
3. When you are asked to smile, please smile until you are told to relax again.

Once the instructions were understood, the recording began. Each video clip lasted approximately 10 to 15 seconds.

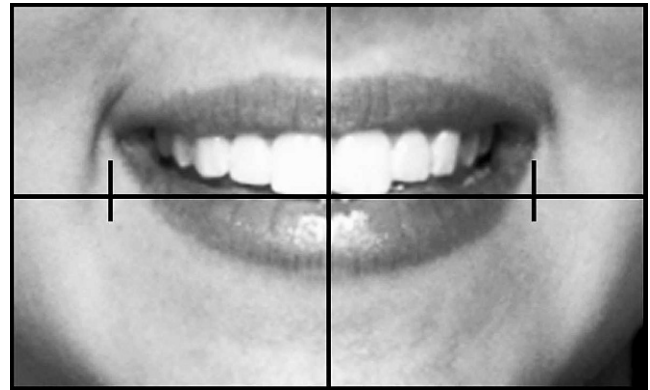


Figure 1. A standardized smile image obtained using the 3" × 5" template.

Image Editing

A 3" × 5" template was created to standardize the size and location of each image. Images were opened in Photoshop® (Microsoft Corporation), and the template was superimposed on top of the image. Smile images were enlarged until the outer commissures of the lips matched the vertical tickmarks inset three-quarters of an inch from the border of the template. The smiling images then were positioned so that the maxillary incisal edges coincided with the horizontal line of the template (Figure 1).

After the images were enlarged and positioned correctly, the portion of the image outside of the template was cropped. Resulting images were edited further in Photoshop by using the *healing brush* tool to remove blemishes, skin irregularities, and other extraneous marks that could influence the rater when evaluating the image. Images were labeled with a four-digit number unique to each subject that was obtained from a random number generator. Following the number, photos obtained from still photography were denoted with a “p” and photos obtained from digital video clips were denoted with a “v.” Once the editing was complete, each image was compressed to approximately 150 kb and was saved as a JPEG file.

Video Editing

Raw digital video clips of each subject were transferred to a computer using a commercially available video editing software package (Adobe® Premiere® 6.0, Microsoft Corporation). This program allowed the streaming video to be converted into individual photographic frames at the rate of approximately 30 frames per second. Thus, a 10 second video resulted in roughly 300 individual frames. The frame representing the subject's posed unstrained social smile was selected, as advocated by Ackerman et al.^{1,3} This frame, identified by the examiner as the

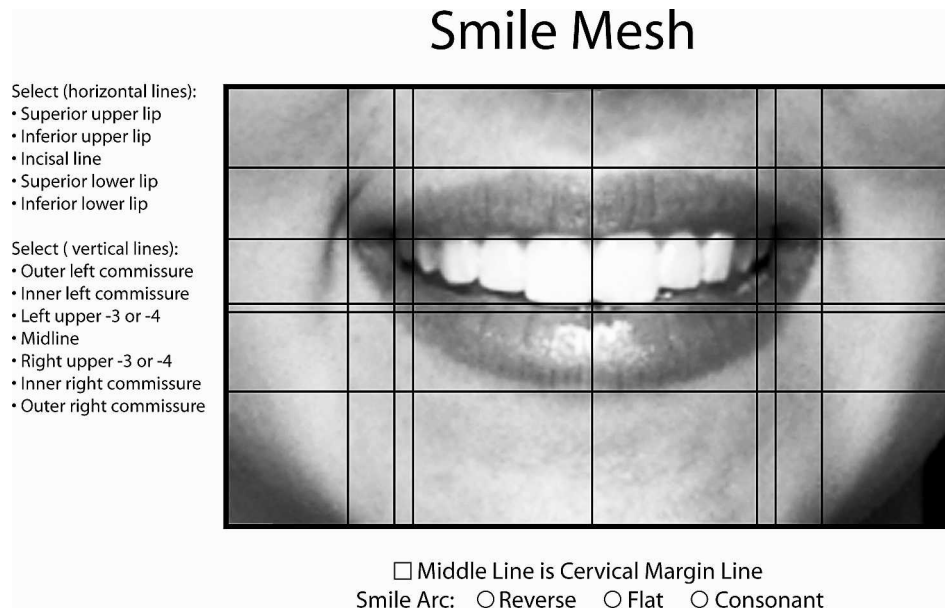


Figure 2. The Smile Mesh program used to measure various lip-tooth relationships associated with anterior tooth display.

“held smile,” was one of 15 consecutive frames in which the smile did not change. This unedited image was saved as a JPEG file.

Smile Mesh Assessment

An updated version of the Smile Mesh program was used in the current study to quantify and compare the characteristics of anterior tooth display found in “attractive” vs “unattractive” smiles. Edited smile images captured by clinical photography and obtained from digital clips of each of the 48 subjects used in this study were scanned into the Smile Mesh program. The height and width of the right maxillary central incisor for each corresponding image were entered into the program before starting. Two adjustable vertical lines, superimposed on the smile image, were moved to correspond with the mesial and distal border of the right central incisor. This enabled a computer-generated algorithm to calibrate the smile measurements to actual life size.^{3,5} The Smile Mesh consisted of an adjustable grid system that comprised seven vertical lines and five horizontal lines superimposed on the smile image. These grid lines were adjusted to correspond with specific hard and soft tissue landmarks (Figure 2). The Smile Mesh then generated 15 lip-tooth characteristics associated with anterior tooth display (Table 1).

Statistical Analysis

Standard descriptive statistics (means, standard deviations, and ranges) were calculated for the Smile Mesh measurements. A Shapiro-Wilks test for normal-

ity performed on the data revealed that these variables were distributed normally. Therefore, parametric statistics were used for inferential tests.

To test the hypothesis that an individual’s smile captured by clinical photography is the same as that obtained from a digital video clip, a paired-samples *t*-test was performed to test for mean differences in Smile Mesh measurements generated from both smile images. The relationship between the various Smile Mesh measurements obtained from smiles captured by clinical photography and from smiles obtained from digital video clips was examined by way of Pearson product-moment correlation. The correlation coefficient estimated the strength of the relationship between these two methods of smile capture.

The type I error rate for all statistical tests was set at .05. All statistical tests were performed with the aid of a statistical software program (Statistical Package for the Social Sciences for Windows, version 12.0, Chicago, Ill).

RESULTS

Standard descriptive statistics were calculated for Smile Mesh measurements taken from smile images obtained from clinical photographs and digital video clips. The significance levels (*P* values) of the paired differences between all measurements are summarized in Table 2. A significant difference was found between 7 of the 14 mean Smile Mesh measurements.

Pearson correlation coefficients were calculated to examine the relationship between Smile Mesh measurements of individual subjects obtained by the two methods of smile capture (Table 3). Other than lower

Table 1. Characteristics of Anterior Tooth Display Obtained from the Smile Mesh Program

Smile Attribute	Description
Maximum incisor exposure	Amount of vertical display of the maxillary central incisors
Upper lip drape	Amount of vertical coverage of the maxillary central incisors by the upper lip (or amount of gingival display)
Lower lip to upper incisor	Vertical distance from the incisal edge of the maxillary right central incisor to the deepest midline point on the superior margin of the lower lip
Interlabial gap	Distance between the most inferior portion of the tubercle of the upper lip and the deepest midline point on the superior margin of the lower lip
Visible posterior teeth width	Distance from the most lateral aspect of the most visible maxillary posterior tooth on the right and left sides
Smile width	Distance from the right outer commissure to the left outer commissure
Smile index	Ratio of smile width divided by interlabial gap
Commissure corridor left	Horizontal distance from the left inner commissure to the left outer commissure
Commissure corridor right	Horizontal distance from the right inner commissure to the right outer commissure
Buccal corridor left	Horizontal distance from the most lateral aspect of the left most posterior visible tooth to the left inner commissure
Buccal corridor right	Horizontal distance from the most lateral aspect of the right most posterior visible tooth to the right inner commissure
Buccal corridor ratio	Distance between the most visible maxillary right and left teeth/Distance between the right and left inner commissures
Smile arc	Curvature of the incisal edges of the maxillary incisors, canines, and first premolar relative to the curvature of the lower lip
Upper lip thickness	Vertical distance from the most superior margin of the upper lip to the most inferior portion of the tubercle of the upper lip
Lower lip thickness	Vertical distance from the deepest midline portion of the superior margin of the lower lip to the most inferior portion of the lower lip

lip to maxillary incisor, all measurements showed a moderate to strong relation with each other (P values ranging from .47 to .82; $P < .001$).

DISCUSSION

The aim of the present study that focused on the esthetics of the smile was to evaluate the relationship between smiles captured by clinical photography and smile images obtained from digital video clips. Because esthetics concerns have become more critical in orthodontic diagnosis and treatment planning, a

fundamental question arises: Are standard static records obtained routinely by orthodontists capable of capturing the smile accurately?

Ackerman et al.¹ introduced the Smile Mesh program to quantify characteristics of anterior tooth display from photographs. They reported that this morphometric tool could measure lip-tooth relationships of the posed social smile accurately and reliably in a clinical setting. The Smile Mesh program was used in the present study to quantify and compare 14 characteristics of smiles captured by clinical photography and digital videography.

Table 2. Descriptive Statistics and Paired-Samples t -Test of Smile Mesh Measurements Obtained from Images of Smiles Captured by Clinical Photographs and Digital Video Clips

Smile Mesh Measurements	Smiles Captured by Clinical Photography		Smiles Obtained from Digital Video Clips		P Value
	Mean	SD	Mean	SD	
Maximum incisor exposure, mm	8.5	1.5	8.9	1.3	.02*
Upper lip drape, mm	0.9	1.9	0.1	2.3	<.01**
Lower lip to maxillary incisor, mm	3.0	1.9	2.8	1.5	.60
Interlabial gap, mm	11.9	2.9	12.5	3.2	.12
Visible posterior teeth width, mm	48.5	4.2	47.7	4.4	.89
Smile width, mm	59.0	5.0	59.1	5.3	.88
Smile index	5.3	1.6	5.0	1.1	.12
Commissure corridor left, mm	5.3	1.8	4.8	1.6	.06
Commissure corridor right, mm	5.0	1.6	4.6	1.7	.07
Buccal corridor left, mm	5.0	1.6	5.9	1.6	<.01**
Buccal corridor right, mm	4.5	1.3	5.5	1.7	<.01**
Buccal corridor ratio	0.8	0.1	0.8	0.1	<.01**
Upper lip height, mm	7.4	1.8	8.0	1.8	<.01**
Lower lip height, mm	10.1	1.8	11.1	1.4	<.01**

* $P < .05$

** $P < .01$

Table 3. Correlation Between Smile Mesh Measurements Captured by Clinical Photography and Smile Mesh Measurements Obtained from Digital Video Clips

Smile Mesh Measurements of Smiles Captured by Clinical Photography	Smile Mesh Measurements of Smiles Obtained from Digital Video Clips	
	Correlation, <i>r</i>	<i>P</i> Value
Maximum incisor exposure	0.78	<.001
Upper lip drape	0.80	<.001
Lower lip to maxillary incisor	0.36	<.01
Interlabial gap	0.56	<.001
Visible posterior teeth width	0.71	<.001
Smile width	0.81	<.001
Smile index	0.47	<.001
Commissure corridor left	0.65	<.001
Commissure corridor right	0.60	<.001
Buccal corridor left	0.64	<.001
Buccal corridor right	0.51	<.001
Buccal corridor ratio	0.53	<.001
Upper lip height	0.82	<.001
Lower lip height	0.74	<.001

A paired samples *t*-test was conducted to evaluate mean differences between Smile Mesh measurements obtained from clinical photographs and digital video clips of the 48 participants. Significant differences ($P < .001$) were found between 7 of the 14 mean Smile Mesh measurements. However, examination of the descriptive statistics, namely, the mean measurement values, revealed some interesting trends. Smiles obtained from digital video clips had larger mean Smile Mesh measurements with respect to three direct measurements of the buccal corridor (buccal corridor right, buccal corridor left, and buccal corridor ratio). These three measurements could have varied because of methodologic differences in smile capture (ie, use of ambient lighting when obtaining smiles from digital video clips, as opposed to use of a supplemental flash when capturing smiles with clinical photography) rather than anatomic differences in the smiles.

More to the point, capturing a smile with ambient light could have created an illusion of increased buccal corridor space and decreased visible posterior teeth width as seen in smiles obtained from digital video clips. Other investigators have reported that the buccal corridor (which also affects the width of visible posterior teeth) appears more pronounced when no supplemental light is added, and that these dark spaces can be eliminated simply by using a flash on the camera.^{2,9,10}

An important consideration with regard to the remaining statistically significant paired Smile Mesh measurements (eg, upper lip drape, upper lip height, lower lip height) is clinical significance. Mean differences of 1 mm or less generally are regarded as clinically insignificant. Therefore, it should be pointed out that none of these average measurements differed by more than 1 mm.

Pearson product-moment correlation was used to examine the relationship between individual Smile Mesh measurements among smiles captured by photographs and digital video clips. Each Smile Mesh measurement of the 48 subjects was correlated significantly (correlation coefficients ranging from 0.47 to 0.82; $P < .001$), with the exception of the measurement of lower lip to maxillary incisor ($P < .01$). Of particular interest, correlations between the statistically significant differences measured with the paired samples *t*-test (other than those associated with the buccal corridor) ranged from 0.74 to 0.82. The strength of these correlation coefficients suggests that anterior tooth display is similar in a smile captured by clinical photography and a digital video clip.

As a technical aside, selecting the specific frame that represented the posed social smile from the video clip, as advocated by Ackerman and Ackerman,³ seemed as arbitrary as capturing the smile at a single time point with clinical photography. As mentioned previously, Tarantili et al.⁸ noted a progression of the smile that consisted of an initial attack period, a sustaining period, and a fade-out or decay period, when the smile is captured by digital video. This progression also was observed in the present study; however, these differences were slight, especially when still images of the smile captured at 30 frames per second were evaluated. Undeniably, error was associated with selecting the appropriate still frame that represented the posed social smile; similarly, a photograph taken of the smile has error associated with it.

Results of the present investigation suggest that a clinical photograph is adequate for analyzing the smile of subjects after orthodontic treatment. The accessibility of digital photography, in particular, should allow us to capture the posed social smile more accurately and reliably because we have instant access to the image. Regardless of whether static or dynamic records are used to capture the smile, the resultant image is only as good as the clinician's ability to capture it accurately.

It should be noted that these results in no way discount the use of digital video as a diagnostic tool for treatment planning. Streaming video allows the clinician to observe the dynamic character of the smile that cannot be seen with a static photograph. Reemphasis on the clinical examination of the patient supplemented by static and dynamic records simply enhances our ability to define specific esthetic goals before providing treatment.

CONCLUSIONS

- A significant positive correlation was noted between Smile Mesh measurements obtained from smiles

captured by clinical photography and digital video clips.

- Digital video clips offer a tremendous amount of information for analyzing the dynamic character of the smile, but a standard digital photograph allows for immediate viewing, and is a valid tool for analysis of the posttreatment smile.

ACKNOWLEDGMENTS

Funds for this research were derived in part from the LeGro Fund, as well as from sources made available through the Thomas M. and Doris Graber Endowed Professorship of the University of Michigan.

REFERENCES

1. Ackerman JL, Ackerman MB, Brensinger CM, Landis JR. A morphometric analysis of the posed smile. *Clin Orthod Res*. 1998;1:2–11.
2. Ackerman JL, Ackerman MB. Smile analysis and design in the digital era. *J Clin Orthod*. 2002;36:221–236.
3. Ackerman JL, Ackerman MB. Digital video as a clinical tool in orthodontics: dynamic smile analysis and design in diagnosis and treatment planning. In: McNamara JA Jr, ed. *Information Technology and Orthodontic Treatment*. Ann Arbor, Mich: Monograph 40, Craniofacial Growth Series, Department of Orthodontics and Pediatric Dentistry and Center for Human Growth and Development, The University of Michigan; 2003:195–203.
4. Sarver DM. The face as the determinant of treatment choice. In: McNamara JA Jr, ed. *Frontiers of Dental and Facial Esthetics*. Ann Arbor, Mich: Monograph 38, Craniofacial Growth Series, Department of Orthodontics and Pediatric Dentistry and Center for Human Growth and Development, The University of Michigan; 2001:19–54.
5. Sarver DM, Ackerman MB. Dynamic smile visualization and quantification: Part 1. Evolution of the concept and dynamic records for smile capture. *Am J Orthod Dentofacial Orthop*. 2003;124:4–12.
6. Janzen EK. A balanced smile—a most important treatment objective. *Am J Orthod*. 1977;72:359–372.
7. Sarver DM. Videoimaging: the pros and cons. *Angle Orthod*. 1993;63:167–170.
8. Tarantili VV, Halazonetis DJ, Spyropoulos MN. The spontaneous smile in dynamic motion. *Am J Orthod Dentofacial Orthop*. 2005;128:8–15.
9. Ackerman MB, Brensinger C, Landis JR. An evaluation of dynamic lip-tooth characteristics during speech and smile in adolescents. *Angle Orthod*. 2004;74:43–50.
10. Dougherty HL. Clubs, quips, phrases, and hype: musing for the new millennium. *Am J Orthod Dentofacial Orthop*. 2000; 117:586–588.