# Tongue in Cheek

George Nagy[1] and Naomi Nagy[2(✉)]

[1] Rensselaer Polytechnic Institute, ECSE, Troy, NY, USA
nagy@ecse.rpi.edu
[2] University of Toronto, Linguistics, Toronto, ON, Canada
naomi.nagy@utoronto.ca

**Abstract.** Differences between image processing and other disciplines in the definition and the role of shape are explored. Diverse approaches to quantifying shape are reviewed. Attention is drawn to the need for close coupling between image acquisition and shape analysis. An example of the effect of the means of observation on dynamic shape extraction is drawn from linguistics. Current instrumentation for measuring shape changes in the vocal tract are described. Advanced image processing based on emerging imaging technologies is proposed for linguistic and therapeutic applications of articulatory phonetics.

**Keywords:** Morphology · Morphometrics · Image processing · Computer graphics · Linguistics · Articulatory phonetics

*Hamlet*
Do you see yonder cloud that's almost in shape of a camel?
*Polonius*
By the mass, and 'tis like a camel, indeed.
*Hamlet*
Methinks it is like a weasel.
*Polonius*
It is backed like a weasel.
*Hamlet*
Or like a whale?
*Polonius*
Very like a whale.

## 1 Introduction

In biology, geography, geology, mineralogy and the fine arts shape is usually considered as an indication of the *origin, development* or *functionality* of the object of study. In contrast, the usual goal of quantifying shape in image processing, computer graphics and computer vision projects is to *recognize* or *label* objects or persons, to *track* targets, or to *render* natural or computer-created scenes. Another difference is that in other fields, researchers spend much effort to optimize image capture, while we often conduct experiments on large image databases with limited information about their original

V. Murino and E. Puppo (Eds.): ICIAP 2015, Part I, LNCS 9279, pp. 332–342, 2015.
DOI: 10.1007/978-3-319-23231-7_30

purpose and acquisition. We will elaborate these distinctions by reviewing approaches to shape within and outside our community, then focus on examples of a familiar shape in linguistics. This shape, that of the tongue, has been studied by many methods, but none provide real-time feedback during naturalistic speech. We outline the necessary parameters for a method or product to fill this lacuna, an endeavor perhaps suited to the expertise of some participants in this conference.

Consider the first sentences of a stimulating 2014 essay on *What is a Pattern,* by Eva and Godfried Toussaint, as justification of this digression from the mainline problems of image processing. Substitute *shape* for *pattern*: *The use of the word 'pattern'* ['shape'] *is ubiquitous in written and spoken discourse. Furthermore, most authors assume the intended reader knows what a pattern* [shape] *is, and hence do not define the term* [1]. Toussaint and Toussaint suggest that a *pattern* must be a sequence exhibiting some regularity, and that its opposite is *chaos* or *randomness*. It seems, however, impossible to conceive an object without any shape at all.

According to the celebrated perceptual psychologist J.J. Gibson (the second author's bio-academic great-grandfather!), *shape, figure, structure, pattern, order, arrangement, configuration, plan, outline, contour are similar terms without distinct meanings* that must be defined precisely for scientific study [2].

Dictionary definitions barely suggest the key constraint of *invariance* under certain operations or transformations, and are too abstract to serve the needs of image processing. It would also be appropriate to add the *method of observation* to any definition, because the characteristics of shape that matter depend so much on the sphere of discourse.

Shapes are often described in terms of familiar items: *egg, pear, banana, almond, peanut, pancake, gugelhupf* (an edible Teutonic truncated right cone with a central vertical cylindrical through-hole), *heart* (usually stylized), *whale, snail, butterfly, tulip, tear drop, bone, corkscrew, bottle* (wine, beer, or Klein), and *boomerang*. In engineering, *boss, chamfer, clevis, crown, driftpin, fillet, fin, flange, kerf, key, lug, camber, dowel, I-beam, hip, lintel,* and *web* have shape associations. A comprehensive survey of mathematical models of shape in geometric and topological terms is [3].

We will extend some reflections suggested by studies of the shapes of Roman and Chinese characters in the 1960s and by subsequent encounters with shape in neurophysiology, physiognomy, remote sensing and computational geometry [4]. Several notions presented in this essay are motivated by studies of variation in natural language.

In the next section, we review approaches to shape in image-processing. In Section 3, we give examples of the role of shape in some other disciplines, including mathematics. In Section 4, as a more detailed example of shape-oriented research that seems to us easily distinguishable from a computer-scientific and engineering approach, we introduce the challenging problem of dynamic characterization of the shape of the tongue in speech production.

## 2 Shape in Image Processing and Computer Graphics

Shape-based recognition has always played a prime role in image processing and computer vision. Although color, texture and shading all provide a direct path to image segmentation, they also offer an indirect path to shape recognition by way of *shape-from-x* [5].

Shape is equally important in computer graphics, where progress is often measured by the realism of rendering stationary or moving persons, animals and objects [6].

Some shape extraction methods mimic known psychophysical mechanisms, including gestalt perception, visual interpolation, and the interpretation of plane curves as 3-D surfaces. Other methods make use of geometry, topology and functional analysis. Physics-based modeling of deformable objects builds on optics, statics, kinematics, elasticity and plasticity. Chain codes, line-adjacency graphs, and delta-hinges preserve contour properties. Popular shape-preserving mappings include the Hu, Mellon and Zernike Moments, Medial Axis, Fourier, Haar, Radon/Hough, Hadamard/Walsh/Radamacher, Daubechies Wavelet, and Euclidian Distance transforms. Vector-space shape-feature representations can be derived from any transform coefficients, or more directly from window operations like the Histogram of Gradients (HOG) or Local Binary Patterns (LBP).

Matheron and Serra formulated *mathematical morphology* as a foundational theory for the manipulation of general geometrical structures in both continuous and discrete spaces [7]. The theory quickly found applications in image processing, first to bi-level images, then to gray-scale images, and eventually to 3-D image arrays. The expressive names of the many useful lattice operations are part and parcel of the image processing lexicon.

Tomographic volume imaging and reconstruction are based on filtered back projection (rooted in the inverse Radon transform) of a stack of 2-D sections. Direct recording methods use 3-D laser scanners (scanning range finders), ultrasound, stylus profilometers or coordinate measuring machines.

Image processing methods are often developed, tested and evaluated competitively on heterogeneous or homogenous collections of images (e.g. outdoor scenes or mug shots) without any need to derive technical or scientific information about the pictured subjects or objects. As we will see, in other fields the study of shape is generally motivated by what it reveals about the nature of the object under study.

## 3 Shape in Other Fields

The eighteenth century scientist Johann von Goethe coined *morphology* for the study of shape [8]. He conducted far-ranging studies of geology and mineralogy in addition to botany. He accumulated a collection of over seventeen thousand rock samples. (In his spare time he wrote novels, plays and poems, advised princes and governments, rebuilt a castle, and begat a large progeny.) *Morphometry*, the measurement and quantification of shape, is an important tool of paleontology that draws increasingly on image processing [9].

*Phyllotaxis* is the study, and also the configuration, of the non-structural components of plants. Some, like pine cones and sunflowers, exhibit repetitive and cylindrically symmetric spirals and logarithmic patterns. The lengths of consecutive sub-sequences of these plant organs have been linked to Fibonacci series. The extraordinary variety of beautiful yet purposeful shapes in Nature has inspired a large literature from Carl von Linné (Linnaeus) through Alexander von Humboldt, Charles Darwin and William

Jackson Hooker [10], to D'Arcy Wentworth Thompson [11] and Oliver Sacks [12]. They might have welcomed some computer help in their laborious visual quantification of organic growth and form.

*Geomorphology* is the study of the evolution of landforms, including the genesis and motion of the continents. *Physiography* (physical geography) tends to focus more on systematic measurement and classification of existing topographic and bathymetric features than on their formation. Most landforms are synonymous with their shape: *dome, spire, crater, escarpment, crevasse, cirque, ridge, mound, butte, knob, hummock, horn, kettle, moulin, col, saddle, mesa, berm, monadnock, cliff, canyon, cape, fjord, promontory…* However, unlike landforms, the corresponding shapes cannot be qualified by location, size or orientation.

*Homology*, the study of evolutionary change in organisms, was at first based exclusively on shape. However, *sequence homology*, the contemporary approach to evolutionary developmental biology, is based on the similarity of DNA sequences due to common ancestry.

Homology has a more precise meaning in mathematics. Indeed, it has several precise meanings (in different branches of mathematics), all related to homeomorphism (continuous transformations). Although *isomorphic* objects must have some kind of similarity, they need not have the same shape. Axis-parallel constraints on shape are called *isothetic*. The mathematical definition includes gridlines that meet at either of two arbitrary points.

Although symmetry in lay language is a relatively simple concept related to shape, in mathematics it is defined as invariance under some specified transformation over lattices. It is formalized in algebra via *symmetry groups*. In image processing it is usually sufficient to consider *translational, scaling, reflection* and *rotational symmetries*. *Helical symmetry* is invariance under a combination of rotation and translation along the axis of rotation. Humans and higher order animals exhibit external bilateral symmetry. Plants and micro-organisms often have rotational symmetries. At the atomic level, symmetries account for many crystallographic properties.

All five Platonic, thirteen Archimedean and four Kepler-Poinsot solids were christened long ago (e.g. *Rhombicosidodecahedron*). For an authoritative discussion of polyhedra, see Coxeter's *Regular Polytopes* [13]. Crystallography adopted the geometrical nomenclature for the shapes of crystals and their symmetries, but Gemology (perhaps more a craft than a science) has its own terminology. The overall shape of the finished stone is called the *cut*. Although the relative sizes and angles of the facets for maximum brilliance and dispersion have been known for over a century, the arbitrary shape of the raw stones often necessitates non-ideal cuts. The corresponding shapes are defined by the proportions of the *table* (flat top*), crown, pavilion, girdle* and *culet*. Cuts, although often polyhedral, have less forbidding names than the 3-polytopes: *marquise, pear, brilliant, trilliant, radiant, princess, emerald, Mazarin, Peruzzi. Cabochon* cuts are rounded, without facets, for opaque stones. Part of a shape design spec (from the American Gem Society) reads as follows:

> *OVAL 6 Main Pavilion Length to Width 1.8:1, Table 55%, Lower Girdle Height 80%, Upper Girdle Height 64%, and 3% Girdle Thickness at the Mains Girdle Must be Faceted*

The perception of shape depends on the point of view of the observer relative to the object. In anatomy and more generally in biology and medicine, there are three principal views (and associated planes and sections). The *sagittal* (medial) plane divides the organism into left and right halves. The *coronal* plane (and frontal view) is face-on. The *transverse* (axial or horizontal) plane is perpendicular to the other two and, in humans, perpendicular to the spine. *Sinister* and *dexter, posterior* and *anterior*, and *inferior* and *superior*, refer to the two possible orientations of each plane or point of view.

Interestingly, there is a school of sculpture where the whole point is to conceal the shape of the sculpture in a specific 2-D projection thereof. These sculptures appear to be randomly twisted rods or wires which, when projected by one or more lamps onto the white wall behind the sculpture, show a familiar shape like a bicycle or a high-heeled shoe or a youth dribbling a basketball [14].

# 4 The Shape of the Tongue

Phonetics is the study of the sounds of speech. Its applications include language (re-)learning after brain, mouth or tongue injury, helping the deaf to speak, acquiring a new language or a new (perhaps socially privileged) dialect, and improving the articulation of opera singers. Tongue-gesture recognition has been proposed as a computer interface [15]. All of these benefit from real-time feedback showing the position of the tongue relative to the palate. In this section, we sketch the role of the tongue in speech production, describe relevant data-collection methodologies, and review an exemplary study that sets the stage for speculating about the applicability of emerging imaging modalities to this problem.

Articulatory phonetics could be an exciting opportunity for image-processing research directly related to shape extraction. The role of the tongue in speech production has long intrigued physiologists, psychologists and linguists (Latin *lingua* means 'tongue' or 'speech'), but appears to have attracted scant attention from the image processing and computer vision community. Research on the subject is more likely to appear in the *Journal of Phonetics, The Journal of the Acoustical Society of America, Clinical Linguistics & Phonetics, the International Congress of Phonetic Sciences* (ICPhS) and the *Acoustics, Speech and Signal Processing* (ICASSP) conference than at *ICPR, CVPR* or *ICIAP* (but see [16] and[17]).

Some of the obstacles to precise observation of the tongue are the following. The tongue is a highly deformable, muscular, mucous-coated organ undergoing rapid motion in multiple dimensions. Its spatial configuration must be determined with respect to the cavity which surrounds it. This cavity has moving walls and only an intermittent opening to the exterior. The internal surfaces of the cavity vary in material and optical properties. The size and shape of the tongue and the cavity vary from subject to subject, and with age and sex. The motion of the tongue relative to the other articulatory components must be synchronized with the audio stream that they generate. Data collection should not interfere with normal speech production in the course of reading, reciting or conversing. Finally, small-scale differences in target locations along multiple dimensions distinguish meaningfully different sounds.

Figure 1a illustrates the fine-grained differences among s-like sounds within languages (e.g., ‘sip’ vs. ‘ship’ in English) as well as sounds distinguishing languages (e.g., Polish has a three-way contrast between “s” as in *kasa* ‘cash’, post-alveolar-s as in *kasza* ‘buckwheat, and palatal-s as in *Kasia* ‘Katie’, but no English-like “sh,” while English lacks the latter two Polish sounds). Figure 1b shows slight differences in tongue position (on the left) that distinguish English vowels (plotted on the right). Figure 1c shows a tongue making a non-linguistic gesture.

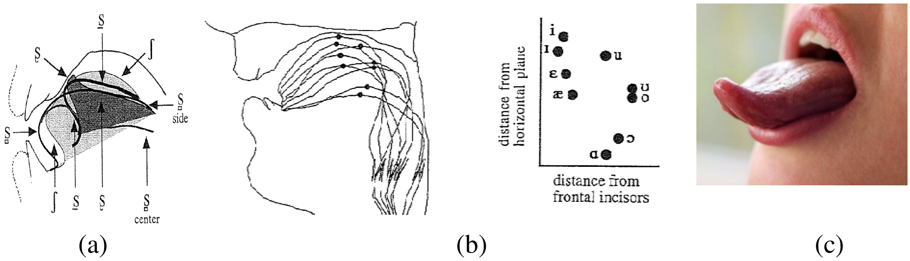
ʂ
ʂ
ʃ
s̺
side
s̺
ʃ
s̪
ʂ
s̺
center
i
ɪ
u
ɛ
ʊ
æ
o
ɔ
ɑ
distance from horizontal plane
distance from frontal incisors
(a)
(b)
(c)

**Fig. 1.** (a) Tongue position for sibilant consonants (all found in the Dravidian language Toda) [17], (b) Tongue positions distinguishing American English vowels [17], (c) an ambiguous lingual gesture. Copyright 1996 by Peter Ladefoged and Ian Maddieson. Reproduced with permission.

## 4.1 Speech Production

Systematic study of articulatory phonetics by means of visual observation, palatography and high-speed cinematography began nearly one hundred years ago. Approximations to the position of various parts of the tongue (and therefore of its shape) have been embedded in the International Phonetic Alphabet devised in 1888.

Complete modeling systems must represent the vocal folds (two thin sheets of tissue in the larynx) that vibrate under pressure from the lungs, the velum (soft palate) that opens and closes the nasal cavity and the tongue, considered the most important speech organ, that changes the shape and size of the oral cavity (which are also affected by the lips and the uvula). Increasingly complex and accurate models of the vocal tract have been developed over the last century. Mechanical, electrical circuit, and simulation models typically consist of a fixed-frequency switchable pulse train source and a series of coupled cavities with variable resonance and attenuation.

## 4.2 Instrumentation

Static X-rays and X-ray cinematography are no longer acceptable for recording speech production, but a 55-minute lateral X-ray film (at 50 fps) of four subjects taken in a Quebec City hospital in 1974 has been transcribed to videodisc and is still used.

Palatography requires putting a dye (or cocoa powder) on the tongue or palate to determine the point of contact during utterance of an isolated sound. Electropalatography (EPG), a long-established but still used technique, substitutes for the dye an artificial palate with several dozen embedded electrodes. EPG provides timing information in addition to the point(s) of contact between tongue and palate.

Other non-imaging modalities are Electromagnetic Articulometry (EMA) and X-ray Microbeam (XMB). EMA records the position and orientation, relative to external transmitters, of receiver coils glued to the articulators (as a function of time). The number of transmitters required depends on whether 2-D (usually in the mid-sagittal plane) or 3-D data is wanted. XMB uses 3-4 radio-opaque pellets glued to the tongue. The tracked locations are called *fleshpoints* in both media.

Besides its widespread clinical applications, ultrasound provides non-invasive real-time images of the upper surface of the tongue (the transducer is below the chin). A frame rate of 30 fps is marginal for some tongue movement and 60fps is the new standard. The tissue-air interface at the top of the tongue reflects 99% of the sound energy, therefore the top of the tongue is clearly visible in sagittal recordings (except for its tip and root). With interactive contour tracing, spatio-temporal surfaces can be generated, but only for a very limited volume of data. Methods to improve the generally noisy images (Fig. 2) by means of head and transducer stabilization, wetting the tongue, and customized instruments are described in Maureen Stone's 46-page guide [18].

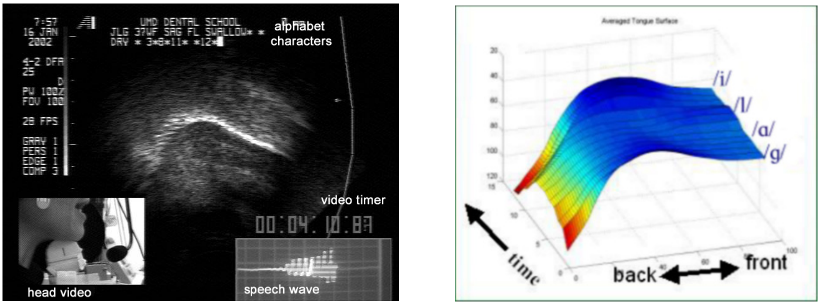

**Fig. 2.** Ultrasound image and generated trajectory of Subject uttering "golly" [18].

Magnetic Resonance Imaging (MRI) is even more expensive and cumbersome than ultrasound, but it provides useable images of soft tissue [19] that is invisible to ultrasound. Real-time MRI (rtMRI) captures images at up to 25 fps (e.g. Fig. 3). These images are sometimes overlaid on crisper static MRI depictions of muscles and edges in the vocal tract. Fleshpoints can be tracked with Tagged Cine-Magnetic Resonance Images (tMRI). The experimental protocol must accommodate the subject in supine confinement in a noisy cylinder, and able to communicate with the experimenter only via microphone and headset.

The speech production and articulation knowledge group (SPAN) maintains the ISC-TIMIT database of 3-D rtMRI and EMA with the same set of ten male and female speakers. The subjects read 460 sentences with statistically representative sounds and sound-transitions that were compiled fifteen years ago. The vocal tract data is interpolated from 12.5 fps 68 x 68 pixel image sequences. The SPAN also distributes free software for visualization [21].

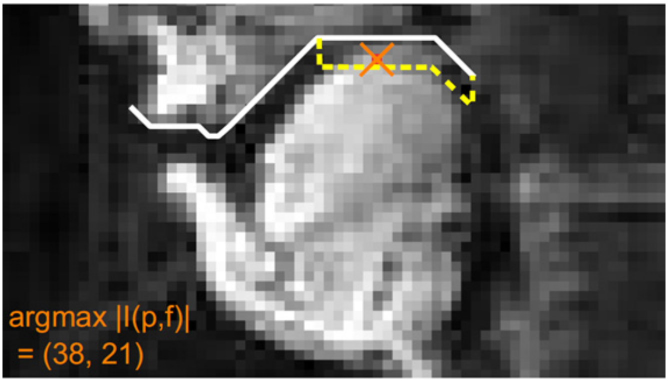

**Fig. 3.** Automatic location of dorsal constriction target: sagittal section of intervocalic stop [ak:o] produced by adult male Italian speaker. Cross indicates center of maximum change in locally-correlated pixel intensity over a surrounding 20-frame sequence [20].

## 4.3 Patterns of Tongue Movement

A fascinating study by K. Iskarous, based on the ancient X-ray movies mentioned at the beginning of the previous section, reveals a startling aspect of tongue kinematics [22]. It punctures the conjecture that the initial and final positions fix a linear trajectory of the corresponding surface points. Careful analysis of B-splines interactively superimposed on the superior edge of the tongue visible in the sagittal recordings shows that 86% of the 600 observed partitions (150 for each subject) fall into just two types: *arch* and *pivot.* Arch transitions merely raise the tongue until the tip touches the palate.

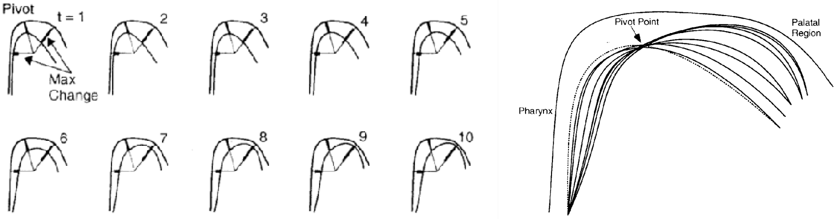

**Fig. 4.** Ten consecutive tongue splines for the transition [ai]. The palatal, uvular and pharyngeal locations of the vocal tract are shown on each frame. The splines are superimposed on the right to indicate the pivot point clearly [21].

The counterintuitive patterns are those of the pivots. The location of the pivot point depends on the vocal task. In Fig. 4, the middle section of the tongue is raised to narrow the distance between the tongue and the palate. On successive frames, the pivot point appears stationary, as shown by the ten superimposed frames of Fig. 4. Measurement of the squared distance (roughly proportional to the area) appears to show a stationary point on the tongue. Nevertheless, the pivot point is not a stationary point of the tongue! The whole back of the tongue passes through the uvular region. The part of the tongue that appears stationary actually moves parallel to the palate.

This finding supports a hypothesized distinction between high-level *sound-production* and low-level *articulatory* control functions. Automating the detection of shape trajectories by means of image processing could extend this type of study to many other puzzling aspects of how the dynamics of the vocal tract affect speech production.

### 4.4 A Modest Proposal for Imaging and Image Processing of Tongue Motion

Analysis of tongue kinematics and other articulatory phonetic studies could be improved and accelerated by wireless in-the-mouth cameras positioned on a molar and an incisor providing simultaneous sagittal and front views of both the tongue and the palate. They would have higher resolution than ultrasound and MRI, show rate of movement, record longer utterances, and provide real-time visual feedback without head restraints. They could also be combined with EPG or ultrasound recording via post-recording synchronization.

Current miniature cameras are still too large. CCD endoscopes and borescopes with a ½" diameter head have VGA resolution, 45°-67° field of view, video frame rate, and LED or fiberoptic illumination [23]. Fiberoptic endoscopes have a smaller head but a thicker cable. Either tether would interfere only moderately with speech. Intra-oral cameras are already used in orthodontics, but they require keeping the mouth open. Perhaps capsule endoscopes can be modified for this purpose. Another development on the horizon is lensless ultra-miniature CMOS computational imagers [24]. These tiny (~100μm) devices don't produce images, but images, or relevant features of images, can be reproduced, at least hypothetically, from their spectral output.

The proposed Tongue in Cheek (TIC) system has two video-cameras and a microphone, and provides real-time feedback via audio and graphics screen output. During a CAVIAR-like [25] training phase with many speakers and diverse utterances, the parameters of a kinetic articulatory model are adjusted to synthesize each speaker's mouth and tongue motion to match the audio input. In therapeutic operation, the screen displays a measure of the spatiotemporal differences between the patient's and the target articulation. In a replay mode, TIC can display a composite split-screen video of the patient's and the desired articulatory motions synchronized with either the patient's or the target audio.

## 5 Conclusion

The study of shape in image processing appears to be less closely tied to function than in many older disciplines. Practitioners in these disciplines tend to develop their own version of well-known image processing algorithms for specialized tasks. Articulatory phonetics is an example of opportunities for exciting multidisciplinary research using unconventional imaging technologies. In bocca al lupo!

**Acknowledgments.** We extend our gratitude to Alexei Kochetov and Craig Diegel, both at the University of Toronto, for sharing their wide-ranging knowledge of phonetics and photography, respectively.

## References

1. Toussaint, E.R., Toussaint, G.T.: What is a pattern? In: Proceedings of Bridges 2014: Mathematics, Music, Art, Architecture, Culture, Seoul, Korea (2014). http://archive.bridgesmahart.org//brid-ges-293.pdf
2. Gibson, J.J.: What is a form? Psychological Review **58**(1951), 403–412 (2014). Cited in Toussaint & Toussaint [1]
3. Biasotti, S., De Floriani, L., Falcidieno, B., Frosini, P., Giorgi, D., Landi, C., Papaleo, L., Spagnuolo, M.: Describing shapes by geometrical-topological properties of real functions. ACM Comput. Surv. **40**(4), October 2008
4. Nagy, G.: The dimensions of shape and form. In: Arcelli, C., Cordella, L., Sanniti di Baja, G. (eds.) Visual Form. Plenum Press, New York (1992)
5. Horn, B.K.P.: Robot Vision. McGraw-Hill, NY (1986)
6. Ferwerda, J.A.: Three varieties of realism in computer graphics. In: Rogowitz, B.E., Pappas, T.N. (eds.) Human Vision and Electronic Imaging VIII, Santa Clara, CA, vol. 5007, January 20, 2003
7. Serra, J.: Image Analysis and Mathematical Morphology. Academic Press (1982)
8. Bloch, R.: Goethe, Idealistic Morphology and Science. American Scientist **40**(2), 313–322 (1952)
9. Zollikofer, C.P., Ponce de Leon, M.: Virtual Reconstruction: A Primer in Computer-Assisted Paleontology and Biomedicine. Wiley, NY (2005)
10. Jackson Hooker, W.: Exotic Flora, vol. 3, pp. 1822–1827. Bibliobazaar, Charleston (2008)
11. D'Arcy Wentworth Thompson, On Growth and Form. MacMillan, New York (1945)
12. Sacks, O.: The Island of the Colorblind. Random House Vantage Books, New York (1998)
13. Coxeter, H.S.M.: Regular Polytopes, Dover (1963)
14. Kagan, L.: Object/Shadow, Installations of Steel and Light. The Butler Institute of America, Youngstown (2009)
15. Saponas, T.S., Kelly, D., Parviz, B.A., Tan, D.S.: Optically sensing tongue gestures for computer input. In: Procs. ACM Symposium on User Interface Technology (2009)
16. Yang, Y., Guo, X.X.: Tongue visualization for a specified speech task. In: SIGGRAPH (2012)
17. Farrar, E., Balasubramanian, A., Coleman Eubanks, J.: Real-time motion capture of the human tongue. In: SIGGRAPH (2009)
18. Stone, M.: A Guide to Analyzing Tongue Motion from Ultrasound Images. Clinical linguistics & phonetics **19**(6-7), 455–501 (2005)
19. Kim, Y.-C., Proctor, M.I., Narayanan, S.S., Nayak, K.: Improved Imaging of Lingual Articulation Using Real-Time Multislice MRI. Journal of Magnetic Resonance Imaging **35**(4), 943–948 (2012)
20. Proctor, M.I., Lammert, A., Katsamanis, A., Goldstein, L., Hagedorn, C., Narayanan, S.S.: Direct Estimation of Articulatory Kinematics from Real-time Magnetic Resonance Image Sequences. Interspeech, Florence (2011)
21. Narayanan, S., Toutios, A., Ramanarayanan, V., Lammert, A., Kim, J., Lee, S., Nayak, K., Kim, Y.-C., Zhu, Y., Goldstein, L., Byrd, D., Bresch, E., Ghosh, P., Katsamanis, A., Proctor, M.: Real-time magnetic resonance imaging and electromagnetic articulography database for speech production research (TC). The Journal of the Acoustical Society of America **136**(3), 1307–1311 (2014)

22. Iskarous, K.: Patterns of tongue movement. J. Phonetics **33**, 363–391 (2005)
23. Schlegel, S., Blase, B., Brüggemann, D., Bühs, F., Dreyer, R., Kelp, M., Lehr, H., Oginski, S.: Endoscope with flexible tip and chip-on-the-tip camera. In: Long, M. (ed.) World Congress on Medical Physics and Biomedical Engineering May 26-31, 2012 Beijing, IFMBE Proceedings, vol. 39, pp. 2111–2114. Springer, Heidelberg (2013)
24. Gill, P.R., Stork, D.G.: Lensless Ultra-miniature imagers using odd-symmetry spiral phase gratings. In: Proceedings of the Computational Optical Sensing and Imaging (2013)
25. Zou, J., Nagy, G.: Visible models for interactive pattern recognition. Pattern Recognition Letters **28**, 2335–2342 (2007)