

A New Technique for Quantitative Measurement of Laryngeal Videostroboscopic Images

Joel A. Sercarz, MD; Gerald S. Berke, MD; David Arnstein, MD;
*Bruce Gerratt, PhD; Manuel Natividad

• The objective measurement of laryngeal function and pathophysiology is one of the goals of current laryngeal research. We describe a new computerized tool for voice analysis systems that allows the quantitative analysis of individual videostroboscopic images. We describe this new technique compared with previous methods of videostroboscopic image analysis and discuss its clinical and research applications.

(*Arch Otolaryngol Head Neck Surg.* 1991;117:871-875)

Laryngologists, speech pathologists, and voice researchers have long sought to define objective measures of vocal function. Despite advances in the study of laryngeal function, surgeons typically rely on their visual examination of the larynx and their perception of the patient's vocal

quality in judging the results of therapy. Objective measures avoid reliance on the subjective interpretation of the listener, which is prone to inconsistency, bias, and error. As phonosurgery becomes more popular, objective analysis of the results may help to establish the benefits of new procedures compared with those of standard therapies.

Herein, we describe an addition to our office-based system for vocal analysis. The new system allows the quantitative analysis of individual videostroboscopic frames. A central feature is a multipurpose software package also used to analyze satellite photographs. The system holds promise both for clinical laryngology and basic laryngeal research.

Objective analysis of vocal function may include imaging and physiologic aerodynamic and acoustic measurements.¹ The information derived from glottography can be further used to generate indexes representing vocal cord motion.² Videostroboscopy is an imaging technique that provides reliable information about vocal fold

movement and overcomes the technical difficulties involved in high-speed motion picture photography. Stroboscopy, however, has been limited by the lack of an accurate and readily available quantitative method of image analysis.

Laryngeal stroboscopy dates back to 1878 and to Oertel, who shone a bright light through a perforated rotating disk and used the alternating light produced to reflect off of a laryngeal mirror. Because it was cumbersome, the device was received unenthusiastically.³ Kallen was the first to use the modern flash tube type of stroboscope in 1932.⁴ Recent advances in stroboscopy units and video cameras, particularly in fiberoptics, have allowed the widespread use of videostroboscopy for the analysis of laryngeal disorders.⁵

Stroboscopy circumvents the inability of the eye to visualize very brief events occurring in laryngeal vibration. An image on the human retina persists for 0.2 second, and sequential events lasting milliseconds are not perceptible. In electron stroboscopy, light flashes can be synchronized with the

Accepted for publication January 8, 1991.

From the Divisions of Head and Neck Surgery, UCLA Medical Center and West Los Angeles Veterans Administration Medical Center, and the Division of Head and Neck Surgery, Department of Surgery, Stanford University Medical Center, Palo Alto, Calif.

Reprint requests to the Division of Head and Neck Surgery, UCLA Medical Center, Los Angeles, CA 90024 (Dr Berke).

fundamental frequency of the subject's laryngeal vibration. If synchronized precisely, a still picture is produced. A moving image is simulated by producing flashes slightly out of phase with the fundamental frequency of phonation. The movement is created from a montage of several different glottic cycles photographed at slightly different points within the cycle.

Some knowledge of the physiology of laryngeal vibration is necessary to correctly interpret videostroboscopic images. Based on anatomic and vibratory studies, Hirano and Kakita⁶ described the body/cover theory of vocal cord vibration, which emphasizes the importance of a layered structure in vocal fold vibration. The stiff underlying body consists of the muscular portion of the vocal cords and the deep lamina propria. The pliable cover includes the overlying mucosa as well as the superficial lamina propria and accounts for the mucosal wave visualized in laryngostroboscopy.⁶ The mucosal wave is best visualized at loud, midrange pitch phonation and diminishes at the falsetto register.³

In 1961, vonLeden⁷ described his initial clinical experience with stroboscopy, emphasizing its use in the early detection of true vocal cord carcinomas and in the management of vocal nodules. Mucosal lesions fixed to the underlying muscle, or body in the model of Hirano and Kakita, abolish the mucosal wave.^{3,6} Fex and Elmqvist⁸ and Fex⁹ subsequently noted the disappearance of the traveling wave in patients with recurrent paralysis and its early reappearance with return of neural function.

Laryngostroboscopic images have been analyzed perceptually with criteria previously described by Kitzing¹⁰ and Bless et al.¹¹ The most frequently used variables are listed in Table 1 and will be discussed briefly. Symmetry not only includes the anatomic structures of the larynx but also reflects symmetry of glottal opening and closing during each cycle. Glottic closure has been divided into seven different categories by Bless et al; these labels identify adequacy of closure, gap location (if present), and the presence of bowing. Periodicity is assessed by synchronizing the strobe flashes with the

Table 1.—Variables of a Perceptually Based Analysis of Videostroboscopy*

Symmetry
Regularity
Glottic closure (complete vs incomplete)
Amplitude of vibration
Presence of mucosal wave
Nonvibrating portion
Presence of ventricular fold hyperadduction

*Data are from Bless et al.¹¹

frequency of phonation. If the larynx does not appear to be a still image, aperiodicity is present. This aperiodicity frequently is manifested clinically as hoarseness. The above measures are not possible with indirect laryngoscopy, because in real time the free margins of the vocal fold form a blurred, vibrating mass.

The measurement of the glottic area has been a primary goal of prior objective studies of laryngeal imaging. High-speed laryngeal photography was developed by the Bell Laboratories in the 1930s. Early efforts to calculate glottic area involved the use of tracing paper, on which laryngeal images from high-speed photography were projected. Several laboratories have developed computerized methods of analyzing projected laryngeal images.^{12,13} Tanabe et al¹³ employed a computer that analyzed a filmed strobe image projected onto a data reduction system, a type of optical-to-digital transformation device. Their system allowed measurement of the excursion curves of the vocal folds and the calculation of glottal width, length, and area.

A substantial drawback of such methods is the expense and cumbersome nature of high-speed photography compared with the current generation of laryngeal videostroboscopes. For example, typical filming of the larynx would involve 10 seconds of phonation or 50 000 frames, which would take weeks to analyze.

More recently, Koike and Imaizumi¹⁴ used a data reduction-based system similar to that of Tanabe et al¹³ to study stroboscopic images on 16-mm motion picture film. Because of a lack of availability and ease, however, motion picture-based methods to calculate the glottic area have not been

Table 2.—Cost of System for Voice Laboratories Already Equipped With Videostroboscopy

Equipment	1990 Cost in \$US
Computer	2000
Data translation (analogue to digital image board)	1500
Image Pro software	1000
Mavigraph video printer	3000

widely used in clinical or research laryngology. Herein, we summarize our early experience with a new method of quantitative analysis of videostroboscopic images, outlining its possible research and clinical applications.

MATERIALS AND METHODS

Our office-based system for objective voice evaluation has been previously described.¹⁵ Briefly, the system involves a personal computer to digitize photoglottographic and electroglottographic measures and acoustic input. Videostroboscopic recording of vocal cord vibration has been used, but the video images were formerly analyzed based on subjective criteria only.

The system used for objective measurement of videostroboscopic images is shown in Fig 1. The vocal cords were visualized on a video monitor through a telescope (Nagashima, Tokyo, Japan) (70°) that was connected to a charge-coupled device video camera (IK-C30A, Toshiba, Buffalo Grove, Ill). A 50-mm lens (Briel & Kjaer/Computar, Orange, Calif) and a diopter lens (Vivitar No. 1, Santa Barbara, Calif) were used to magnify and focus the image of the glottis. Illumination was provided by a stroboscope (4914-4, Bruel & Kjaer) that employs a hand-held stethoscope for synchronization of the strobe pulse with the vocal frequency. The videostroboscopic images were then recorded with a three-fourths inch videotape recorder (VO 9850, Sony) equipped with a time-code generator. A video time-date generator (WJ-810, Panasonic) was also used to mark each video field.

Recorded vocalization of the vowel *i* was played back with the job and shuttle search mode of the videotape recorder. Desired video fields were identified for digitization. Before digitization, the video field was stabilized through the use of a video printer (Mavigraph UKP-5000, Sony, Teaneck, NJ). The image was then digitized with a video Frame Grabber (2853-SQ, Data Translation, Marlboro, Mass) and stored in the computer. Thereafter, the images could be retrieved and observed on either the video monitor or a printed photograph from the video printer.

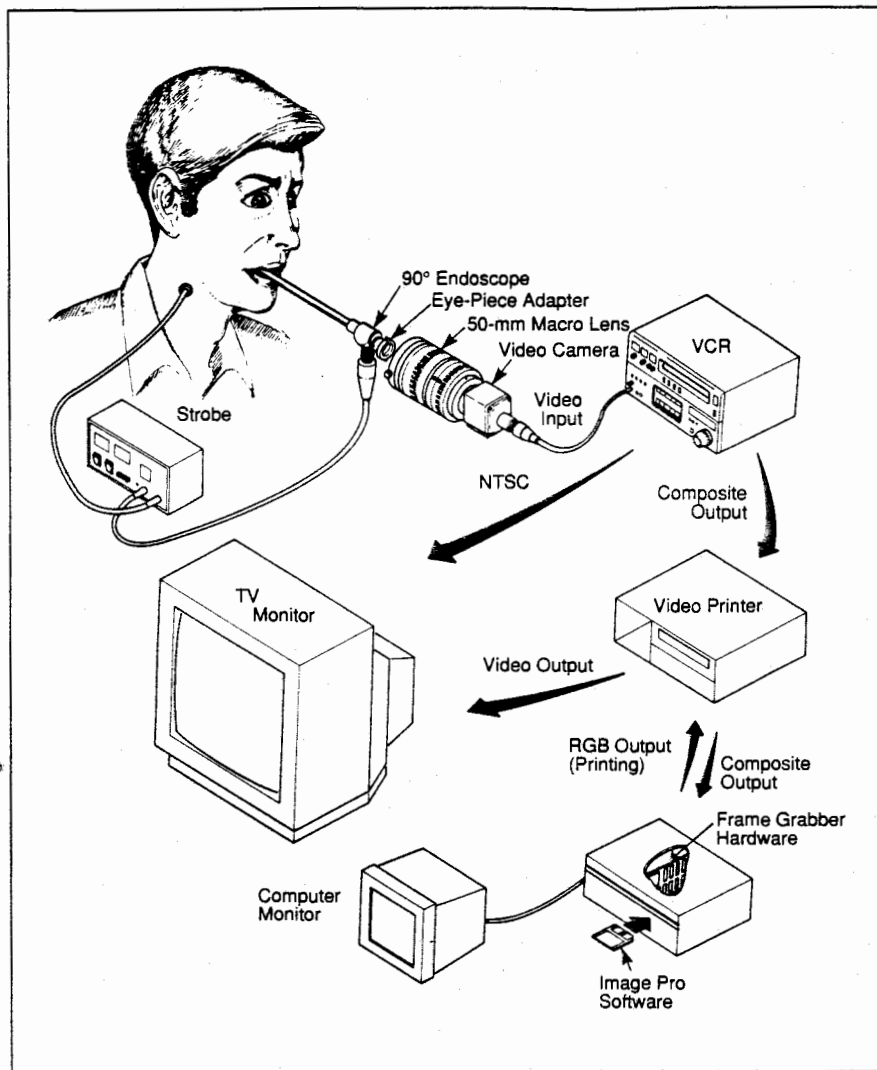


Fig 1.—Diagram of system for objective analysis of videostroboscopic images. RGB indicates red, green, and blue; VCR, videotape recorder; and NTSC, National Television Standard Commission.

The area of the glottis was measured with mouse-driven software (Image Pro, Media Cybernetics, Silver Spring, Md). The area of the measured trace was expressed in pixel units or calibrated to express the area in other units. Figure 2 illustrates the method of calibration used for an in vivo canine experiment. A standard centimeter ruler was lowered to the level of the glottis and measured with the software, which gave a pixel value equivalent to 5 mm in length. To compare measurements from one subject, it is critical that the camera is positioned at a constant distance from the larynx.

The mouse-driven software can measure the length or area of any portion of an image as directed by the mouse. It also allows the user to label images, rotate an image, measure angles, combine two images, change the scale of an image, or place multiple video frames on a single screen or photograph.

CLINICAL APPLICATIONS

Some uses of this system are presented in the following examples. The first subject was a volunteer with a normal larynx; the second and third subjects were patients with unilateral vocal cord paralysis. Patient 1 underwent laryngoplasty, and patient 2 underwent Isshiki type I thyroplasty. The surgical patients were examined before and after the operation.

A still video image from the normal volunteer is illustrated in Fig 3, top. This image was digitized and stored in the computer. Figure 3, middle, shows the digitized image after being recalled from the computer hard drive. As an example of the use of the mouse-driven

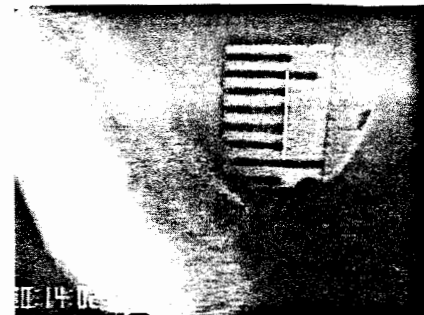


Fig 2.—Calibration ruler placed at the level of the vocal cord for use in canine studies. The software provides a pixel equivalent for a 5-mm length.

software, we measured the relative area of the vocal folds. The examination of the volunteer yielded normal findings, and there was no history of laryngeal disease. Figure 3, bottom, shows the outlining of the vocal fold with the mouse. Five measurements were made on five separate video frames and averaged. The area of the right vocal cord averaged 11 735 pixels, and that of the left vocal cord averaged 11 872 pixels. This result was expected, given the symmetry of the larynx and lack of disease. This technique has potential value for monitoring vocal fold atrophy or reinnervation in the months following a laryngeal paralysis.

Figure 4 demonstrates the computerized analysis technique comparing relative glottic area in patient 1 before and after laryngoplasty for left-sided vocal cord paralysis, during the most closed and most open segments of the cycle. We have found that within a montage of stroboscopic images, the most closed and most open portions of the cycle are easily identifiable. Figure 4, top, illustrates the outlining of the transglottic area with use of the mouse-driven software. Area measurements were made in patient 1 from digitized stroboscopic images after normalization of the anterior to posterior lengths. To normalize an image, the glottis is framed from its anterior and posterior extent and placed into a standard-sized box with use of the software. The images in Fig 4, bottom, have been normalized, allowing the comparison of images obtained at dif-



Fig 3.—Top, A videostroboscopic image obtained from a normal human volunteer. Middle, Same image after digitization by the video Frame Grabber. Bottom, Same image showing outlining of the vocal fold with use of the mouse-driven software.

ferent times with the camera at slightly different distances from the larynx.

For the patient illustrated in Fig 4, bottom, the preoperative ratio of the most closed to most open areas was 8327/9375 pixels, or 0.89, indicating minimal glottic closure. Postoperatively, the ratio of the most closed to the most open area was 815/4057 pixels, or 0.20. For comparison, the ratio is zero in the normal case in which there is no glottal gap during the most closed portion of the cycle. The ratio of the postoperative to preoperative most closed area is 815/8357 pixels, or 0.076,

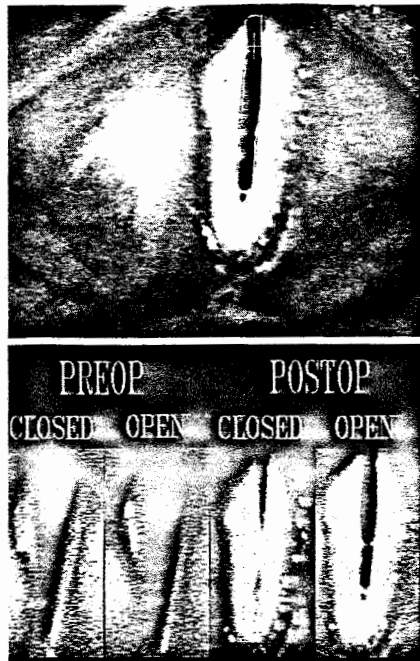


Fig 4.—Videostroboscopic image obtained from patient 1 with outlining of the glottis. Software provides area in pixels. Bottom, Composite image of patient 1 showing the most open and most closed images before (preop) and after (postop) laryngoplasty. The patient had minimal glottic closure preoperatively.

indicating that the surgery reduced the glottal gap during the most closed portion of the cycle to 7.6% of its preoperative value. However, postoperatively there was still a glottal gap during the most closed portion of the vocal cycle: the patient's most closed area was still 20% of the most open area.

Another potentially important measure is the symmetry of vocal cord excursion, which can also be obtained with the digitized most open and most closed images. The images in Fig 5 depict this procedure for patient 3 who had right-sided vocal cord paresis. The thickness of the vocal fold for the most closed and most open images is measured on each side in pixels. The maximal medial-to-lateral excursion is length *a* minus *c* (25.0 pixels) on the patient's left side and length *b* minus *d* (8.0 pixels) on the right side, as shown in Fig 5. The symmetry ratio in patient 3 is 25.0/8.0 pixels, or 3.1, indicating that the normal left cord's excursion is 3.1 times greater than the weak right cord's excursion. Such excursion asym-

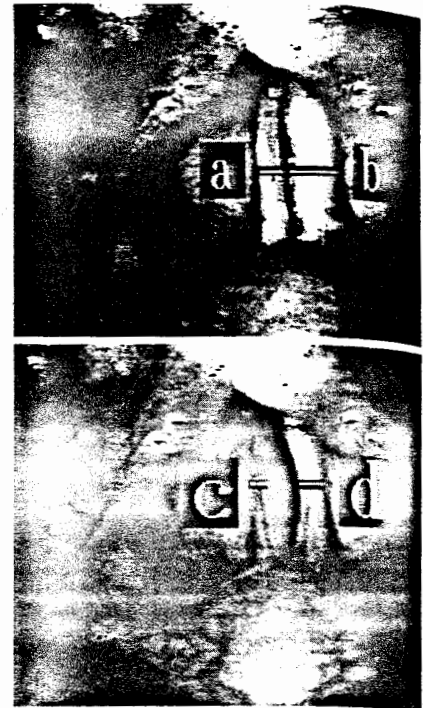


Fig 5.—Digitized videostroboscopic image of the most open phase of the cycle in patient 3 with right-sided vocal cord paresis showing measurement of the symmetry ratio: *a* equals 53.0 pixels and *b* equals 25.0 pixels. Measurement of the symmetry ratio in patient 3 during the most closed phase of the cycle: *c* equals 28 pixels and *d* equals 17.0 pixels.

metry is an important factor of the patient's voice disorder.

To compare ratios between patients with either left- or right-sided paralysis, we have arbitrarily determined that the vocal cord with greater excursion is always placed in the numerator. A second variable is then used in cases of laryngeal paralysis to indicate whether the normal "N" or paralyzed "P" cord demonstrates the greatest lateral excursion. For example, a ratio of 1.5N indicates that the normal cord demonstrates 1.5 times the excursion of the paralyzed cord.

COMMENT

The system described herein has two major clinical applications. First, the system allows quantitative analysis of laryngeal variables, including relative glottic area, interarytenoid distance, vocal cord area, and symmetry of traveling wave motion. We have demonstrated how these measures ap-

ply to patients with unilateral vocal cord paralysis before and after laryngoplasty. Second, the system permanently archives digitized video images, thereby permanently storing the images on the disk memory of a relatively inexpensive computer. This feature is important for conditions of the larynx that require the comparison of serial video images and eliminates time-consuming searches through videotape. Approximately six frames can be archived on one high-density diskette, and up to nine images can be printed by the video printer.

Compared with previous methods of objective analysis of stroboscopic images,¹⁴ the methods described herein have the advantages of ease of operation, availability, and accuracy. Most laboratories lack the capability to quantify laryngeal variables through high-speed photography, but centers with stroboscopic units can readily ob-

tain the computer software and hardware necessary for the system depicted. Additional equipment costs for laboratories with videostroboscopic capability are listed in Table 2.

Our group currently has only preliminary clinical experience with the symmetry ratio and the ratio of the most closed to most open portions of the cycle presented herein. However, a number of studies comparing quantitative videostroboscopic measures to acoustic, aerodynamic, and glottographic measures are under way.

The technique is limited by the fundamental nature of both stroboscopy and the computer software. Stroboscopy evaluates only a "snapshot" of any single glottic cycle and is therefore not suited to the assessment of patients with severe laryngeal aperiodicity, because extensive cycle-to-cycle variations in vocal periodicity cannot be evaluated with stroboscopy.⁸ In ad-

dition, the software system has two main limitations. First, it is somewhat labor-intensive, requiring a trained individual to carefully outline the laryngeal variable to be measured. Although software for automatic analysis is not yet available, such automatic analysis methods would save considerable time and expense. Second, the resolution of the videostroboscopic images following digitization is reduced but still adequate, as shown in Figs 2 and 3.

This system provides the opportunity to quantitatively analyze individual videostroboscopic images. Further advances in videostroboscopic technology and computer software promise to improve the ease and accuracy of the system described herein.

This investigation was supported by grant DC 00855-01 from the National Institutes of Health, Bethesda, Md. US Dept of Health and Human Services.

References

1. Ward PH, Hanson DG, Gerratt BR, Berke GS. Current and future horizons in laryngeal and voice research. *Ann Otol Rhinol Laryngol.* 1989;98:145-152.
2. Berke GS, Moore DS, Hantke DR, Hanson DG, Gerratt BR, Burstein F. Laryngeal modeling: theoretical, in vitro, in vivo. *Laryngoscope.* 1987;97:871-881.
3. Alberti PW. The diagnostic role of laryngeal stroboscopy. *Otolaryngol Clin North Am.* 1978; 11:347-354.
4. Padovan IF, Christman NT, Hamilton LH, Darling RJ. Indirect microlaryngostroboscopy. *Laryngoscope.* 1973;83:2035-2041.
5. Gould WJ, Kojima H, Lambiase A. A technique for stroboscopic examination of the vocal folds using fiberoptics. *Arch Otolaryngol.* 1979; 105:285.
6. Hirano M, Kakita Y. Cover-body theory of vocal fold vibration. In: Daniloff RG, ed. *Speech Science.* San Diego, Calif: College-Hill Press; 1985:1-46.
7. vonLeden H. The electronic synchron-stroboscope: its value for the practicing laryngologist. *Ann Otol Rhinol Laryngol.* 1961;70:881-893.
8. Fex S, Elmqvist D. Endemic recurrent laryngeal nerve paresis: correlation between EMG and stroboscopic findings. *Acta Otolaryngol (Stockh).* 1973;75:368-369.
9. Fex S. Judging the movements of the vocal cords in larynx paralysis. *Acta Otolaryngol (Stockh).* 1970;263:82-83.
10. Kitzing P. Stroboscopy: a pertinent laryngological examination. *J Otolaryngol.* 1985;14:151-157.
11. Bless DM, Hirano M, Feder RJ. Videostroboscopic evaluation of the larynx. *Ear Nose Throat J.* 1987;66:48-58.
12. Hayden EH, Koike Y. A data processing scheme for frame by frame film analysis. *Folia Phoniatr (Basel).* 1972;24:169-181.
13. Tanabe M, Kitajima K, Gould WJ, Lambiase A. Analysis of high-speed motion pictures of the vocal folds. *Folia Phoniatr (Basel).* 1975;27:77-87.
14. Koike Y, Imaizumi S. Objective evaluation of laryngostroboscopic findings. In: Fujimura O, ed. *Vocal Physiology, Voice Production, Mechanisms and Functions.* New York, NY: Raven Press; 1988.
15. Berke GS, Hanson DG, Trapp TK, Moore DM, Gerratt BR, Natividad M. Office-based system for voice analysis. *Arch Otolaryngol Head Neck Surg.* 1989;115:74-77.