Videostroboscopic images associated with glottographic waveforms in an in vivo canine model of phonation

Gerald S. Berke, Terrence K. Trapp, Bruce R. Gerratt, David G. Hanson, David P. Arnstein, and Manuel Natividad

Citation: The Journal of the Acoustical Society of America **85**, 1789 (1989); doi: 10.1121/1.397923
View online: http://dx.doi.org/10.1121/1.397923
View Table of Contents: http://asa.scitation.org/toc/jas/85/4
Published by the Acoustical Society of America

TECHNICAL NOTES AND RESEARCH BRIEFS

Rudolph H. Nichols
Physics Department, Code 61/No, Naval Postgraduate School, Monterey, California 93943

Editor's Note: Original contributions to the Technical Notes and Research Briefs section are always welcome. Manuscripts should be double-spaced, and ordinarily not longer than about 1500 words. There are no publication charges, and consequently, no free reprints; however, reprints may be purchased at the usual prices.

Videostroboscopic images associated with glottographic waveforms in an *in vivo* canine model of phonation[a] [43.80.Ev, 43.80Jz, 43.80.Nd]

Gerald S. Berke, Terrence K. Trapp, Bruce R. Gerratt, David G. Hanson, David P. Arnstein, and Manuel Natividad

VA Medical Center, West Los Angeles, California 90024 and UCLA School of Medicine, Division of Head and Neck Surgery, Los Angeles, California 90024, 62-132 CHS

A technique correlating videostroboscopic images with glottographic vocal fold vibratory events in an in vivo canine model of phonation is reported. The system involved digitizing glottographic signals simultaneously with the video camera vertical retrace signal and a 5-ms synchronization pulse that was also recorded on the audio channel of a video recorder. The details and validation of the technique are discussed.

INTRODUCTION

Although noninvasive measures such as photoglottography (PGG) and electroglottography (EGG) provide useful information on vocal fold movement, inferences are often required regarding the corresponding glottal morphology to a particular segment of the glottographic signal.[1] Knowledge of vocal fold vibratory morphology in normal and pathologic phonation is fundamental to the understanding of glottal airflow and acoustics. Thus confirmation of the relationship between objective measures of vocal function and vocal fold configuration is desirable and important in studying laryngeal physiology.

Gerratt *et al.* reported correlation of 35-mm photographic images from stroboscopic flashes with their positions on glottographic waveforms.[2,3] This technique was limited, however, by the ability to document only one image of the glottis during a single, sustained phonation. Anastaplo and Karnell correlated electroglottographic data to simultaneously obtained videostroboscopy. The technique used an oscilloscopic tracing of an EGG signal triggered by the stroboscopic light flash.[4] Video images of both the strobed larynx and the EGG tracings from the oscilloscope were then merged and recorded on videotape to correlate the EGG with vocal fold configuration. However, a few problems in this method may reduce its utility. For example, the oscilloscopic waveforms were recorded on videotape; therefore, subsequent waveform analysis and processing were not possible without making tracings of the waveforms from the monitor screen. Additionally, the EGG signal was reconstructed from a number of glottal cycles, providing a composite EGG waveform rather than the raw signal. The present study describes a technique in which digitized, glottographic signals were correlated with their corresponding videostroboscopic images, allowing subsequent waveform analysis.

I. METHODS

Figure 1 depicts the experimental setup. The technique was developed using a previously reported *in vivo* canine model of phonation[5]; therefore, the details of the model will not be discussed.

Phonation was produced by independent electrical stimulation of the recurrent and superior laryngeal nerves while insufflating warm humidified air through the glottis. The animal respired through a tracheostomy under general anesthesia. Vocal fold vibration was monitored by photoglottography (PGG) and electroglottography (EGG). A photo sensor (Centronics OSD-50) was placed on the trachea, and a supraglottic xenon light source was used for PGG. EGG was obtained with a Synchrovoice unit. The reference electrode was applied to the strap muscles and the recording electrodes placed on either side of the thyroid cartilage. The analog signals from the EGG and PGG were low-pass filtered at 3 kHz and digitized with a 12-bit A/D board at 10 kHz for 5 s (Fig. 1).

The elements used for videostroboscopic synchronization included (1) a record of the video camera vertical retrace signal, (2) a pulse to synchronize the video image to the digitized glottographic data, (3) a video time or frame coder, and (4) a record of the strobe flashes transduced by a light sensor (PGG).

A record of the video signal from a charged coupled device (CCD) video camera (SONY industrial model 101) was obtained by passing it through the coupled TV vertical trigger input of an oscilloscope (Hitachi, V-1050F) for retrieval of the vertical retrace signal from the gated output for concurrent digitization with the glottographic signals (Fig. 1). The video camera output was also simultaneously recorded on a SONY VO5800 video recorder along with the signal from a time coder (Panasonic, WJ-810). A 5-ms, 50 mV square wave pulse (SWP) (at approximately 1 pulse per second) was digitized and simultaneously recorded on the audio channel of the video recorder to synchronize the tape recorded and digitized video fields.

Although a low-level constant xenon PGG light was present during the stroboscopic imaging, no blurring or distortion of the stroboscopic images was observed. The position of the strobe image on the digitized glottographic waveforms was identified by the left border of the high-amplitude strobe light spike superimposed on the PGG waveform.[2,3]

II. AUDIO SYNCHRONIZING PULSE AND STROBE FLASH POSITION

This study addressed two concerns regarding system operation. The first was the use of an audio SWP to synchronize the tape recorded video signal and digitized video fields. A video SWP could not be routinely used in this system due to distortion of subsequent CCD signal line transfer timing by the pulse. In order to validate using an audio SWP for synchronization of video signals to digitized data, the SWP was simultaneously recorded on the audio and video channels of the tape recorder. Playback of the recorded pulse caused an audible click in the audio channel that coincided with a 5-ms white band generated on the monitor screen. The audible click and white band invariably occurred within the same video field, indicating that an audio SWP could reliably be used to synchronize the video and digitized fields.

Another primary concern was the timing relationship of the strobe flash to when the image produced by the flash occurred in the tape-recorded video signal. To examine this, the Bruel and Kjaer (B&K) 4914 laryngostroboscope 35-mm photographic flash option was utilized. This option produced a relatively brighter 24-klux and longer 80-μs flash when compared to the 5-μs strobe light that occurs during normal strobe operation. Therefore, the 80-μs flash could be identified in the tape-recorded video

[a] This paper was presented at the 115th meeting of the Acoustical Society of America, Seattle, Washington, May 1988 [J. Acoust. Soc. Am. Suppl. 1 **83**, S111 (1988)].

 0001-4966/89/041789-05$00.80

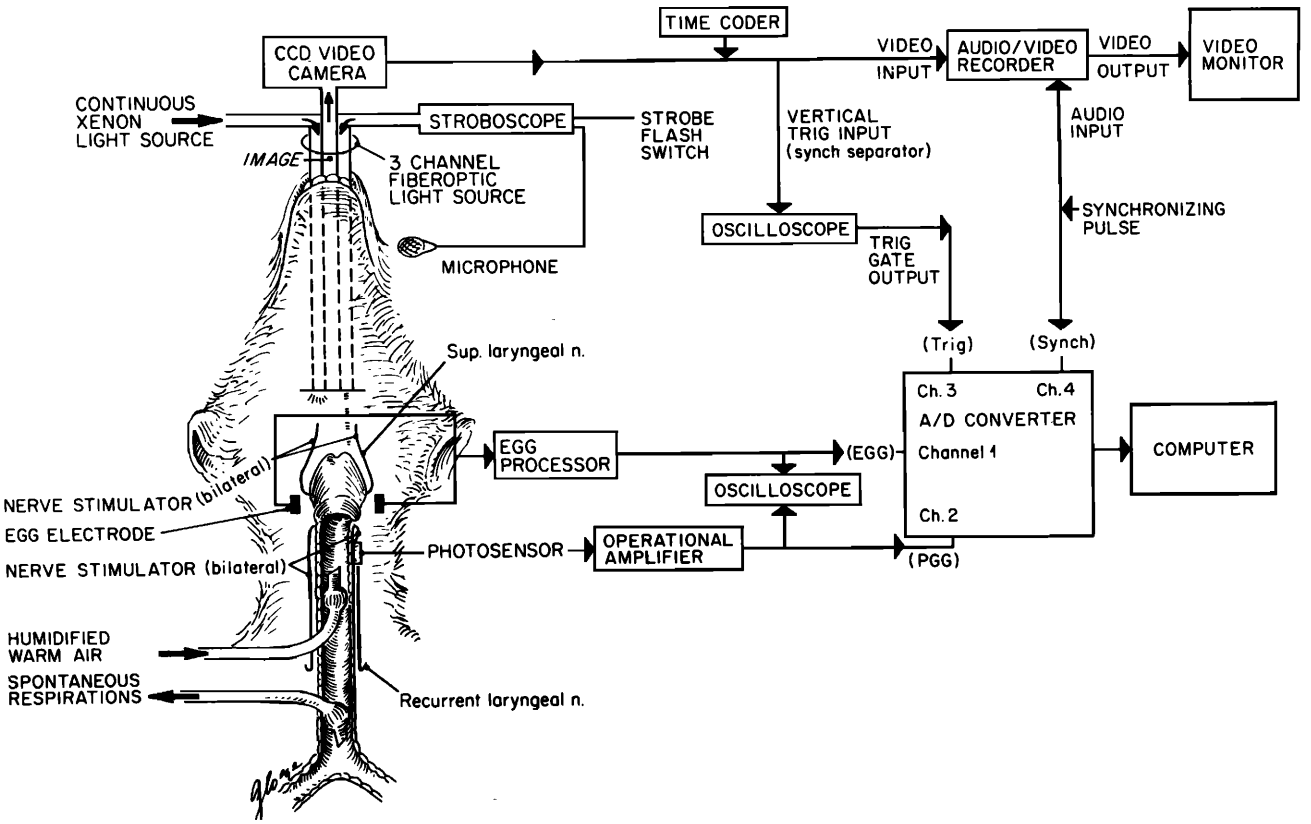


FIG. 1. Diagram of the experimental setup.

signal because the flash was brighter and was also distinguished in the digitized PGG waveform by its longer duration than the 5-μs flashes. Using the experimental setup of Fig. 1, the 80-μs flash was activated during videostroboscopy of a stationary test pattern. After synchronization of the digitized and tape recorded video fields, the faint images generated by the 5-μs light were sequentially tracked to determine when the video field containing the bright 80-μs strobe flash appeared relative to when the flash actually occurred as demonstrated in the digitized signals.

Figure 2 depicts data from such a procedure. It was observed that the bright image or field from the camera flash option was identified in the video field following the one in which the flash occurred, as identified in the digitized PGG signal. After this brightly illuminated field, a second field with a bright image was also observed. If the 5-μs strobe light did not flash during the time these two fields were recorded (due to the delay in the resumption of normal strobe operation), the third field was dark. The fourth and subsequent fields again had dim images as the 5-μs strobe light resumed flashing during recording of the third field.

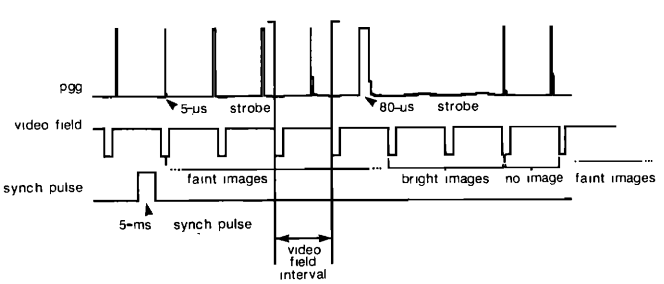


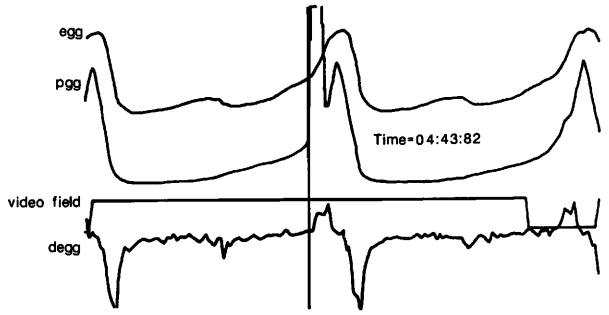
FIG. 2. Validation of timing of recorded images relative to strobe occurrence. Bright images were produced by the 80-μs flash. Digitized waveforms: PGG recording of strobe, video fields, and synchronization pulse. Cursors depict the video field interval.

These results can be interpreted within the context of the operational characteristics of the SONY 101 CCD camera signal line transfer. The latch of the even or odd line array occurs at 12 horizontal lines after the initiation of the vertical retrace.[6,7] At the latch, all odd or even array lines are sent to the output register and the CCD odd or even array is reset. However, because there is a two-to-one interlace (odd-even-odd-even..., etc.), each odd or even array integrates light for two fields or one frame (1/30th of a second). Thus the two sequential bright images resulting from a single 80-μs flash were caused by overlap of light integration from an even and odd array. Consequently, if strobing occurs at greater than 30 flashes per second, multiple composite images may occur in one video screen. When they do occur, it is usually perceived during the glottal opening phase, when a darker open glottis is superimposed on the lighter closed vocal folds. In our experience, strobing at 60 flashes per second, the rate for most commercially available laryngostroboscopes, has only occasionally produced noticeable composite images. However, to assure separate images, the strobe should be set at no more than 30 flashes per second. A few laryngostroboscopes such as the Storz (model 8000) permit this adjustment. It is important to understand that 30 flashes per second does not imply a fixed rate.

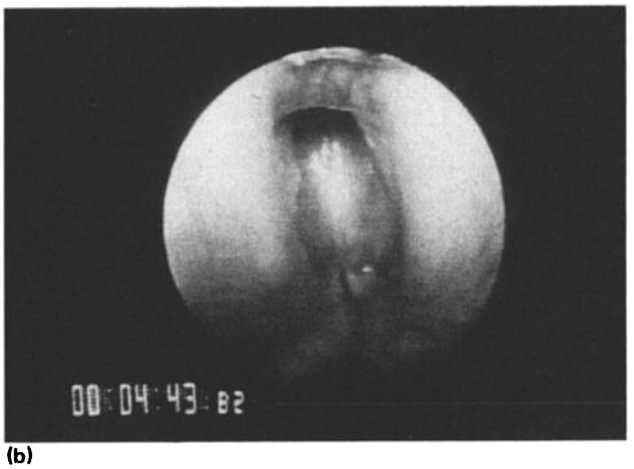
The Storz unit alters the phase angle advance between successive strobe flashes such that a slow motion rate of 0.5, 1, 2, or 4 Hz can be observed (1.4 Hz for the B&K unit). Therefore, if the fundamental frequency is 120 Hz and the stroboscope is set at 30 flashes per second and a slow motion rate of 1 Hz, then one flash occurs every fourth cycle with a phase angle advance of 12 deg per strobe flash.

III. CORRELATION PROCEDURE

The procedure used for correlation can be most easily explained by an example. If we want to find the glottal configuration that corresponded to the peak of the PGG, first we would scan the entire digitized segment to determine where the left border of a strobe flash occurred at the peak of a

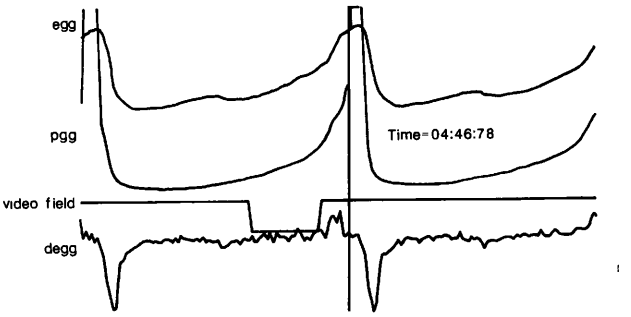


(a)

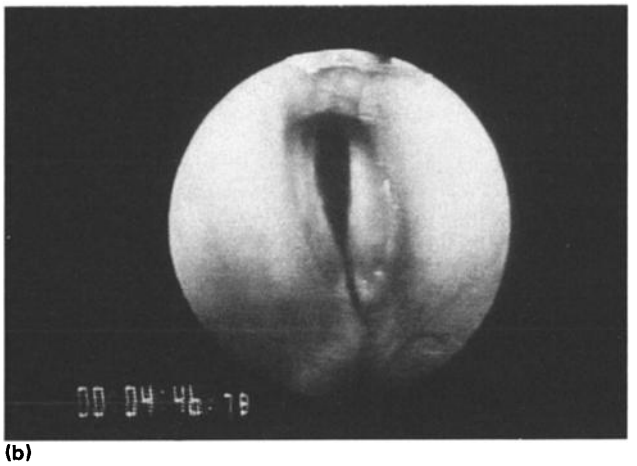


(b)

FIG. 3. (a) Glottographic waveforms and (b) the corresponding video image at the beginning of horizontal separation: egg is the electroglottograph; pgg is the photoglottograph, video field camera timing information; degg is the differentiated electroglottograph. The cursor depicts the position of image on the waveforms.

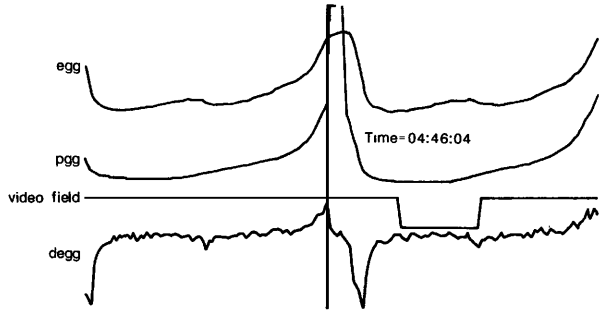


(a)

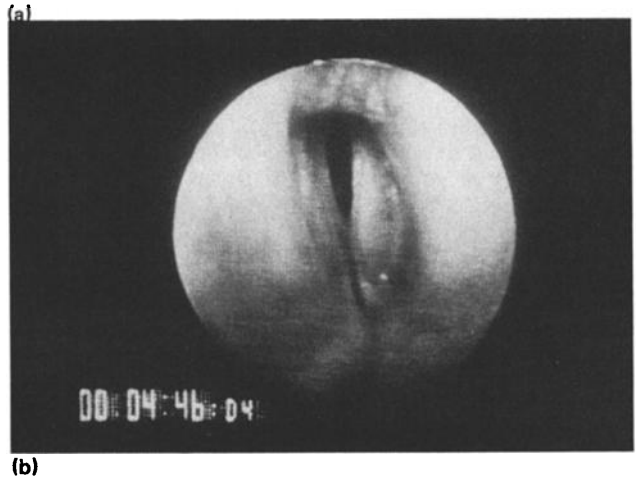


(b)

FIG. 5. (a) Glottographic waveforms and (b) the corresponding video image at the peak of the PGG waveform.

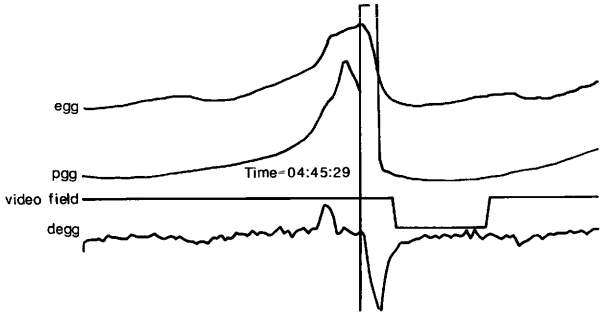


(a)

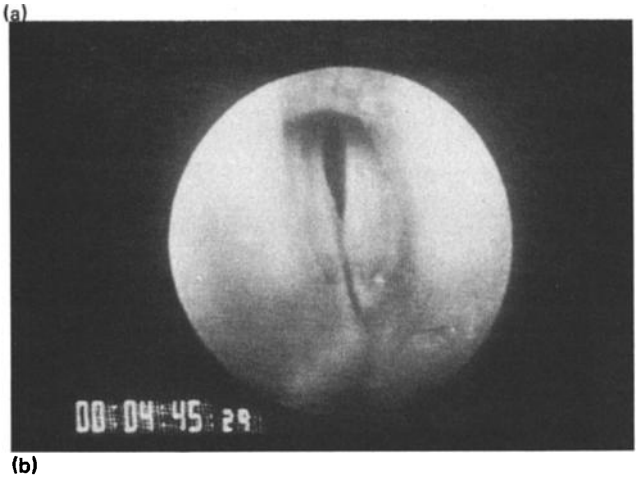


(b)

FIG. 4. (a) Glottographic waveforms and (b) the corresponding video image at the peak of the differentiated egg.



(a)



(b)

FIG. 6. (a) Glottographic waveforms and (b) the corresponding video image at the end of the plateau phase of the EGG waveform.

PGG waveform. Then we would find the nearest field containing an SWP to the one of interest (in this case, the field where a PGG peak is crossed by a strobe flash). Having found the nearest SWP field, it would then be necessary to determine the time interval from the *field containing* the SWP to the *field following* the one of interest. Then the digitized field containing the SWP would be synchronized to the video field containing the same SWP. The previously determined time interval could then be used to determine the number of video fields between the one containing the SWP to the one containing the image representing the glottal configuration of interest (in this case, the one at the peak of the PGG). It should be emphasized that, in determining the time interval from the digitized SWP field to the field of interest, one needs to use the field following the one of interest. This is because the recorded image does not appear on the video screen until one field after the one in which the flash actually occurred (see Fig. 2). In a 5-s digitized segment, three to five SWPs were usually actuated. The position of a specific SWP (i.e., 1st SWP, 2nd SWP,..., etc.) was determined by reviewing the entire digitized segment. Similarly, the audio SWP's location was determined by reviewing the audio channel of the videotape during playback in the search mode. This scanning procedure usually takes about 100 s. The system could be used in either direction to correlate images with their position on waveforms or points on waveforms with their corresponding vocal fold images. This system also permitted correlation of all recorded images with their corresponding positions on glottographic signals.

IV. RESULTS

The images and their position on glottographic waveforms displayed in Figs. 3–8 were determined using the above synchronization technique. In (a) in Figs. 3–8, the cursor has been placed on the left border of the strobe to indicate the glottographic position of the images. Figure 3(a) and (b) shows the beginning of horizontal separation. Figure 4(a) and (b) demonstrates the vocal fold configuration at the peak of the differentiated EGG waveform. Figure 5(a) and (b) displays the lateral excursion of the vocal folds at the peak of the PGG waveform. Figure 6(a) and (b) displays the closing of the lower margin and its position at the end of the plateau phase of the differentiated EGG waveform. Figure 7(a) and (b) depicts the lower margin further closing as the nadir of the differentiated EGG is approached. Figure 8(a) and (b) shows the nadir of the differentiated EGG and the dissolution of the traveling mucosal wave.

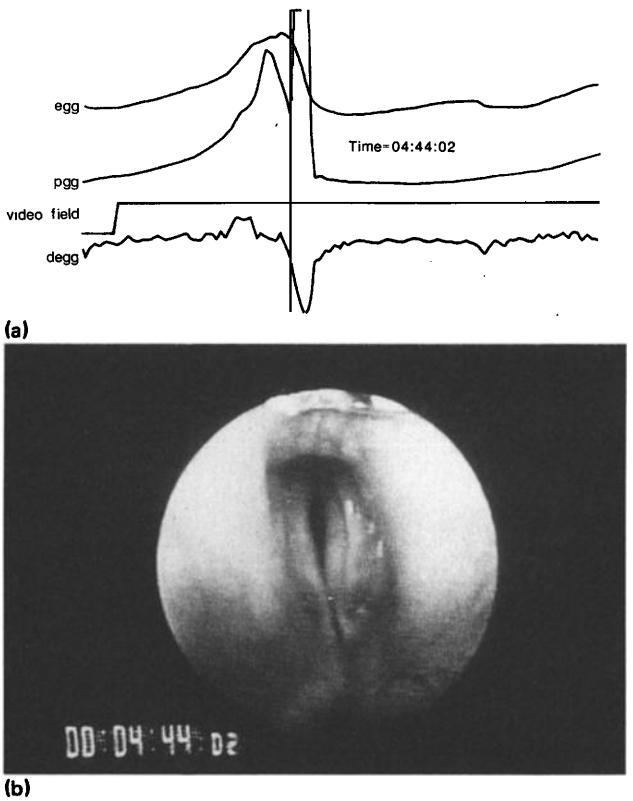


FIG. 7. (a) Glottographic waveforms and (b) the corresponding video image prior to the nadir of the differentiated EGG.

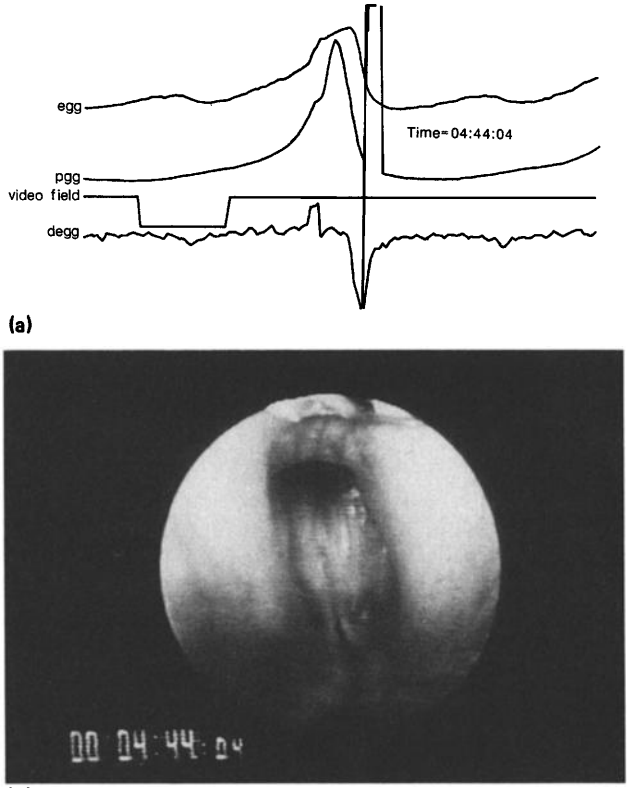


FIG. 8. (a) Glottographic waveforms and (b) the corresponding video image prior to the nadir of the differentiated EGG waveform.

V. DISCUSSION

This synchronization technique was developed using an *in vivo* canine model of phonation. However, application to human phonation should be possible using the same correlation methodology. Although *in vivo* canine phonation was produced by quite invasive means, the actual synchronization technique utilized methods common to the study of human laryngeal physiology (videostroboscopy, PGG, EGG) in a number of research laboratories and clinics.

This study demonstrated that stroboscopic operation at greater than 30 flashes per second may produce a recorded image on the video screen that is a montage composed of multiple strobe light flashes (i.e., multiple vocal fold configurations). However, even when the montage effect occurs, it may be difficult to perceive if the stroboscope is operated at greater than 60 flashes per second. This is due to the small phase shift present between each successive strobe flash at high rates of strobing. Although Anastaplo and Karnell[4] demonstrated the lack of a delay between the strobe light and trigger output, they did not control the number of strobe flashes occurring per video field and, therefore, their recorded images may have been a montage. Their technique is thus more akin to the average vocal fold configuration associated with a place on the EGG rather than the specific configuration as we have reported.

There are a number of advantages to the system presented in this report. In contrast to single-lens reflex photography, many images can be sampled during a sustained phonation. The method can be adapted to most objective measures of vocal fold movement, and waveforms obtained can be subsequently analyzed. In addition, the system is relatively inexpensive in comparison to high-speed filming, and is clinically applicable. The disadvantages are those related to the use of a stroboscope: Usually only one image can be obtained per cycle and very irregular vibration may be difficult to reconstruct.

In summary, we have presented an imaging system correlated to digitized glottographic data that utilizes the video camera vertical retrace signal and a synchronizing pulse that is digitized and simultaneously recorded on the audio channel of a videotape recorder.

ACKNOWLEDGMENTS

This work was supported by a VA Technical Merit Review Grant and a PHS Grant R01 NS-20707 from the National Institute of Neurological and Communicative Disorders and Stroke.

[1]B. R. Gerratt, D. G. Hanson, and G. S. Berke, "Glottographic measures of laryngeal function in individuals with abnormal motor control," in *Proceedings of the 4th International Conference on Vocal Fold Physiology*, edited by T. Baer (College Hill, 1986), pp. 521–531.

[2]G. S. Berke, D. Hantke, and B. R. Gerratt, "An experimental model for the testing of vocal fold tension and length effects upon phonation," in *Proceedings of the IALP Conference on Voice*, edited by M. Hirano (Kurume, Japan, 1987), pp. ii, 1–8.

[3]B. R. Gerratt, D. G. Hanson, and G. S. Berke, "Laryngeal configuration associated with glottography," Am. J. Otolaryngol. **9**, 173–179 (1986).

[4]S. Anastaplo and M. P. Karnell, "Synchronized videotroboscopic and electroglottographic examination of glottal opening," J. Acoust. Soc. Am. **83**, 1883–1890 (1988).

[5]D. M. Moore and G. S. Berke, "The effect of laryngeal nerve stimulation on phonation: A glottographic study using an *in vivo* canine model," J. Acoust. Soc. Am. **83**, 705–715 (1988).

[6]R. H. Stafford, *Digital Television: Bandwidth Reduction and Communication Aspects* (Wiley, New York, 1980).

[7]Consultation with the video imaging section, SONY Corporation of America.

Effects of surface riblets on the reduction of wall pressure fluctuations in turbulent boundary layers [43.28.Ra, 43.50.Gf]

Timothy W. Lancey

California State University at Fullerton, Fullerton, California 92634

Laurel W. Reidy

Naval Ocean Systems Center, San Diego, California 92152-5000

For this study, pressures at the surface of a flat plate below a turbulent boundary layer were measured to allow comparison of the wall pressure for a smooth wall to that of a surface covered with riblets of triangular cross section running streamwise. Also measured were mean and fluctuating air velocities across the boundary layers. It was found that the wall noise was reduced approximately 2.8 dB by the riblets, while the fluctuating velocity was also reduced, but to a lesser extent.

INTRODUCTION

Recent studies have shown that fine riblets, with sharp peaks running streamwise along a flat plate, reduce drag by decreasing the effective wall shear stress.[1–5] It was also reported that the turbulence fluctuation levels near the wall were reduced by the riblets.[1] Wall pressure fluctuations under normal zero pressure gradient turbulent boundary layers have been measured, with highest intensity fluctuations occurring near a wall unit frequency of $\omega^+ = 0.52$, where ω^+ is obtained from the radian frequency (ω) by $\omega^+ = \omega(\nu/U_\tau^2)$; here, ν is the kinematic viscosity and U_τ the wall friction velocity.[6–8] The rms wall pressure (p) was found to be related to the free-stream dynamic pressure (q) and the skin friction coefficient (C_f) by approximately $(\overline{p^2})^{1/2}/q = 3.4\ C_f$, where $q = \frac{1}{2}\rho U_0^2$. Then, for a smooth wall, using a microphone having a pinhole aperture flush with the wall, the rms pressure is near 0.01 q (see Refs. 9 and 10). This is valid for a wall unit aperture diameter (d^+) up to approximately 100, where $d^+ = dU_\tau/\nu$ relates actual length to wall unit length, in this case diameter.[10] For the present study fluctuating wall pressures and velocities (near $y = 0$) were measured with a smooth vinyl surface and with a surface of parallel, triangular riblets, running streamwise along the test section of a wind tunnel. Narrow and 1/3-octave-band analyses of the pressures and narrow-band analyses of the fluctuating velocities were performed.

I. EXPERIMENTAL APPARATUS AND METHODS

A 0.0762-mm-thick sheet of smooth vinyl was glued to a flat plate, 12.7 mm thick, next to a 0.229-mm-thick vinyl sheet, manufactured by the 3M Company, having riblets of triangular cross section of 0.152-mm height (h) and lateral distance peak to peak (s) yielding $s^+ = h^+ = 12$ (for 30.5 m/s free-stream speed). A type 4138 Bruel & Kjaer microphone, 3.18 mm ($\frac{1}{8}$ in.) in diameter, was flush mounted to the upper surface of the plate, from below. The microphone terminated at a 0.762-mm-diam ($d^+ = 60$) pinhole aperture, necessary for high-frequency discrimination of the pressure. The microphone was designed to slip fit into each side of the plate. This is presented in Fig. 1. A TSI 1218-20 single sensor boundary layer hot film probe was used for velocity measurements.

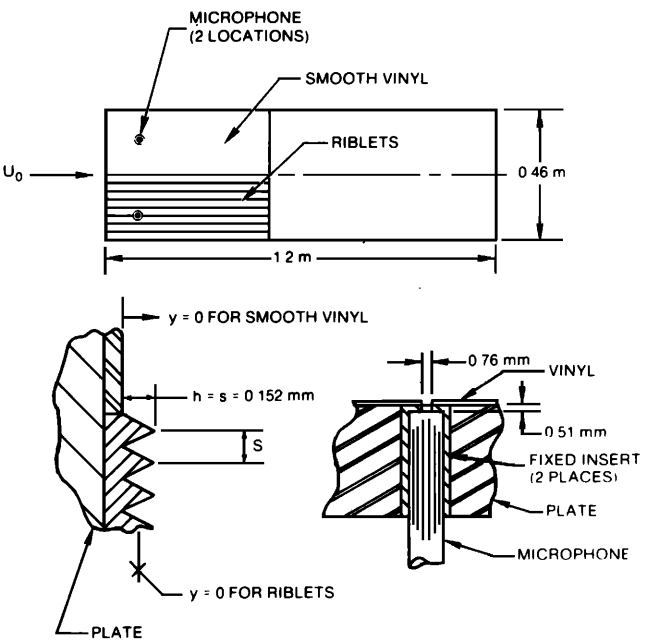


FIG. 1. Plate configuration.

The plate was mounted in a wind tunnel that has a 0.762-m-wide×0.51-m-high×1.22-m-long test section. Air free-stream velocity (U_0) was set at 30.5 m/s, and the fluctuating velocity near $y = 0$ and the wall pressure were measured at the riblets side of the plate; this was repeated at the smooth side of the plate by moving the microphone from the riblets side to the smooth side, and by laterally traversing the hot film probe, with the flow conditions held fixed. The $y = 0$ datum is defined for each surface

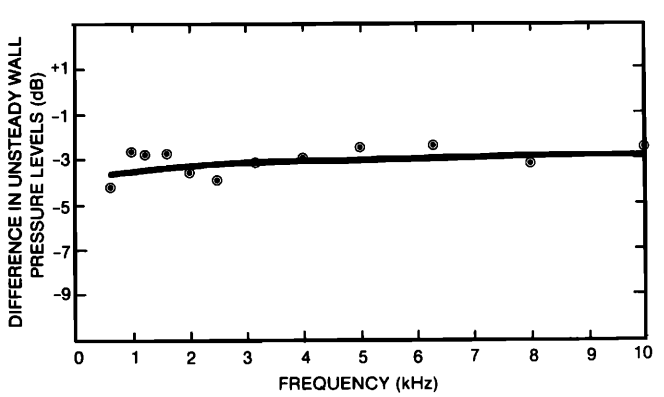


FIG. 2. Wall noise reduction with riblets, relative to a smooth surface, using 1/3-octave-band analysis. Solid line is the least-squares best-fit curve for data.