A COMPUTERIZED METHOD OF FRAME-BY-FRAME FILM ANALYSIS FOR FIBERSCOPIC MEASUREMENT OF THE GLOTTIS

S. Miyazaki, S. Sekimoto, H. Ishida and M. Sawashima

Frame-by-frame analysis of motion pictures is an important experimental procedure in the physiological study of speech. A technical difficulty in this procedure is that it requires a large amount of time-consuming manual work. Although fully automated picture processing system is still not practical, a considerable improvement in the procedure can be achieved by employing a computer-aided, i. e., man-machine interactive, system.[1)]

This paper reports on a new technique combining a 16-mm film projector and a computer input device consisting of a "control ball"[2, 3)] and a scope display. This system is being applied to measurements of the glottal opening in fiberscopic films.

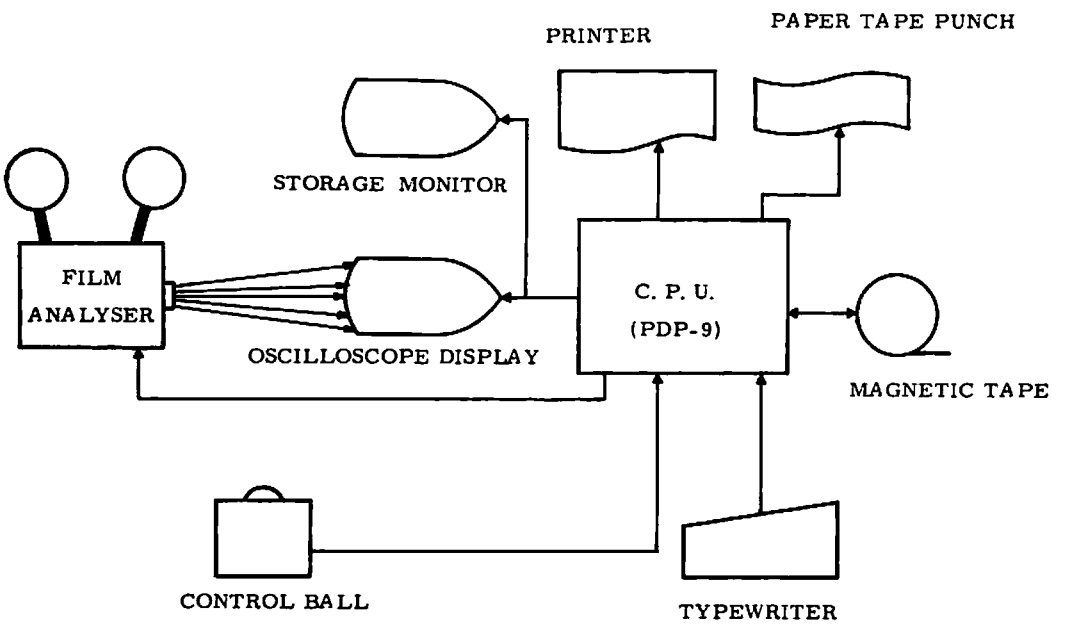

Fig. 1. A schematic diagram of the film data registration and editing system.

Structure and Operation of the System

A schematic diagram of the system is shown in Fig. 1. The system is composed of a frame-by-frame film projector and a PDP-9 computer with peripheral equipment, including a flat face oscilloscope display (Tektronix RM 503), a storage oscilloscope for monitoring, and the control ball.

The film is directly projected on the oscilloscope surface where a small cross as a two-dimensional cursor is displayed by the computer. Since both the film image and the cross mark appear on the same fluorescent surface, there are no parallax problems.

By manipulation of the ball surface of the control ball, the cross mark can be moved in any direction on the oscilloscope surface for determining the points of measurement in the frame image. The vectorial rotation of the ball is resolved into X and Y components, the increment angles being measured digitally by X and Y counters, respectively. The values of the two counters are fed into the accumulator of the central processor to determine the coordinates of the cross mark. Each of the S and Y counters holds 9 bits and distinguishes 512 values, giving corresponding coordinate steps on the oscilloscope display. Since the effective area for the measurement on the display was designed to be 90 x 90 mm^2, a unit step in the displacement of the cross mark on the X or Y axis is 0.18 mm.

When the center of the cross mark is visually judged to be positioned at the desired point in the picture, the experimenter pushes a control button on the control ball, and a fixed dot replaces the cross mark on the scope, making the cross available for selecting a second point in the picture. When the button is pushed for the second time, the cross mark is deleted, leaving two light spots at the selected points. At this point of the procedure, the distance between the two points is automatically calculated and plotted against the frame number on the storage-type monitor oscilloscope. The current frame number of the film is also displayed in a corner of the monitor scope.

When the button is pushed for the third time, the two light spots are deleted and the cross mark is again displayed, the film being advanced by one frame for the next measurement. The plotted points on the storage monitor remain in display. Special marks such as voice onset, articulatory closure or release can be given for the pertinent film frames through the

keyboard. There is another push-button attached to the control ball. This is used for erasing the fixed light spot for correction.

After a series of measurements, the data are stored on magnetic tape with a file name. The data can be displayed either as a graph on the oscilloscope or printed out either numerically or in a graphic form.

Measurement of Fiberscopic Laryngeal Film

The system is currently in use for measuring glottal width during articulations of voiceless consonants on fiberscopic films. The distance between the film projector and the oscilloscope display is set at 100 cm, and the film image is magnified by a factor of 17. Under this condition, the diameter of the picture element transmitted by a single glassfiber of our fiberscope is approximately 0.75 mm on the projected plane.

As a measure of the glottal width, the glottal distance at the tips of the vocal processes of the aryteroid cartilages was estimated. Fig. 2 shows an example of the graphic print-out of glottal width values for consecutive frames. The utterance represents a Japanese word /seQkeH/ (snow valley) embedded in a frame "sore o _____ to yuu" (We call that _____). The interval between consecutive plots along the abscissa is 20 msec (corresponding to 50 frames/sec), and an arbitrary linear scale is used for the ordinate.

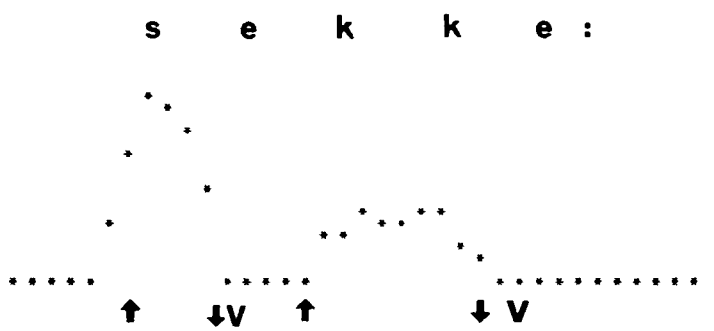

Fig. 2. A print-out of glottal width against time for the utterance of /seQkeH/ in a frame "sore o _____ to yuu".

The upwards arrows at the bottom indicate the beginning moments of the oral closure or constriction, the downwards arrows the oral releases of the consonants, and the V's voice onset for the following vowels, as determined by spectrograms. In the graph, it is observed that the glottis abducts to a considerable extent for a time span of 140 msec for [s] while a smaller glottal aperture with a larger duration of 200 msec is shown for the geminate [kk]. No separation of the arytenoids is revealed for the vowel segments.

Some data are discussed in relation to the laryngeal gestures for the voiceless consonants in a separate article in this issue.[4)]

References

1) E. H. Hayden and Y. Koike, "A Data Processing Scheme for Frame by Frame Film Analysis," Folia Phoniatrica, Vol. 24, 169-181 (1972).

2) H. Ishida, S. Tairadata, and K. Umezawa, "On a Display Setup for Controlling Experiments," Proc. of Graphic Display Symposium, 13-22, Information Processing Society (1970) (in Japanese).

3) E. Goto et al., "A Control Ball and Graphic Input through a High-precision CRT," Proc. of 11th Programming Symposium, A35-A49 (1970).

4) M. Sawashima and S. Miyazaki, "Glottal Opening for Japanese Voiceless Consonants," Annual Bulletin (Research Institute of Logopedics and Phoniatrics, University of Tokyo) this issue.