X-RAY MONITORING OF THE POSITION OF THE FIBERSCOPE IN THE PHARYNX

Shigeru Kiritani

In the previous issue,[1] we reported that our pilot system of the computer controlled radiography was being used effectively in monitoring the position of the fiberscope during pronunciation. In this report, some modifications of the computer program are described and preliminary results of the movement of the fiberscope observed for some test utterances are presented.

Algorithm for Estimation of the Fiberscope Position

Fig. 1 shows an x-ray image of a typical situation in this monitoring experiment, with the tip of the fiberscope hanging down in the pharynx. For tracking the movement of the fiberscope, its position is determined by locating the anterior edge (a vertical line segment on the left side of the fiberscope) and the lower edge (a horizontal line segment) of the fiberscope. The x- and y-coordinates of the image field in the program to be described are chosen in such a way that the y-axis is approximately parallel to the outline of the back pharyngeal wall. In typical situations of the observation of the larynx, the fiberscope is located in the pharynx approximately along the back pharyngeal wall. Therefore, its longitudinal axis is also approximately parallel to the y-axis. Movements of the fiberscope are measured as the x- and

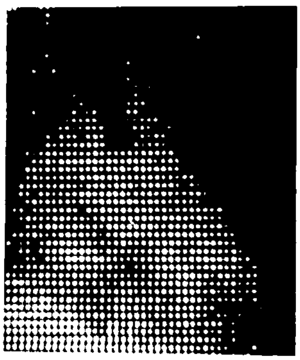

Fig. 1. Example of an x-ray image of the fiberscope in the pharynx.

y-displacements of a point fixed on the tip of the fiberscope. The spatial sample points for locating the fiberscope are selected 1 mm apart both in x- and y-directions.

Sample points to be exposed for the position measurement of one time frame are selected as shown in Fig. 2. For each time frame, first a horizontally extended rectangular mesh field of 4 x 14 sample points is exposed ("exposure field" henceforth) and by using the radiopacity data at these points, the position of the anterior edge is determined by the following method. A regular mesh of 4 x 6 sample points is selected as a sub-field within the exposure field and the difference between the average x-ray intensities in the right- and left-halves of the sub-field is examined. By trying seven different (horizontally shifted) positions for this sub-field, the position for which this difference is maximum is determined and the middle line of the sub-field for this position is taken as the vertical edge of the fiberscope. The horizontal line segment for the bottom edge is determined in a similar way by maximizing the difference between the average x-ray intensities for two vertically consecutive halves of a 4 x 4 mesh area sub-field by trying the seven different vertical positions within an 11 x 4 exposure field. The positions of the exposure fields are updated for each time frame according to the last detected position of the fiberscope. Thus, at each time frame, one hundred sample points are exposed and any movement with a speed of not larger than ± 3 mm per frame in x- and/or y-directions can be tracked continuously.

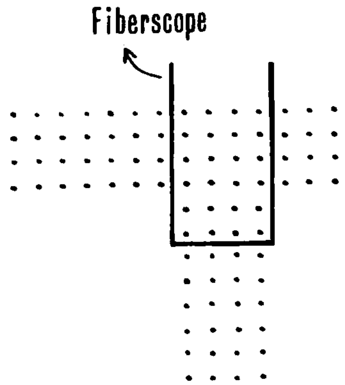

Fig. 2. Sample points to be exposed for 'one frame' determination of the position of the fiberscope during tracking of its movement.

In the present program, the exposure time per sample point is set at 230 microsec and additional computing time of 50 microsec per sample point is required on the average. Thus, the measurement time for one time frame is 28 msec in total and the resulting frame rate is 36 frames per second. The accuracy of the position measurement was calibrated by a model experiment and was estimated at approximately ±6 mm both in x- and y-directions.

In our previous program, the horizontal and vertical positions of the fiberscope were determined by a single horizontal and vertical line scan, respectively. In that case, occasional measurement errors in the horizontal position significantly affected the positioning of the vertical scan line and often made it difficult to locate the lower edge of the fiberscope correctly. The present modification of the program is intended to make up for this defect. By this modification, the total amount of exposure per frame is made nearly twice as large as before; the exposure time per sample point is one fourth of that before, taking the advantage of higher redundancy in the image information.

Movement of the Fiberscope Observed for Some Test Utterances

As a preliminary but practical application of the present method, movements of the fiberscope due to velar movements in some utterances have been actually observed in real time. Pronunciations of a continuous sound sequence [i: ĩ: i: ĩ: . . .] and utterances of the Japanese words [i˩ʃi], [i˩dʒi], [i˥ʃi], and [i˥dʒi] in isolation have been tested.* The glottal images through the fiberscope were simultaneously recorded by a cinecamera at a rate of 24 frames per second. Examples of the observed movements are shown in Fig. 3 and Fig. 4. In what follows, the results of these tests are summarized.

(1) The fiberscope was located close to the back pharyngeal wall when the velum was at a lowered position during the [ĩ] pronunciation and moved in the forward-upward direction as the result of the elevation of the velum towards the [i] gesture. The amount of displacement between the nasalized and non-nasalized vowel articulations was 4 mm in the x-direction and 6 mm

* The mark "˩" and "˥" indicate the rise and drop of pitch, respectively, as manifestations of different word accent patterns.

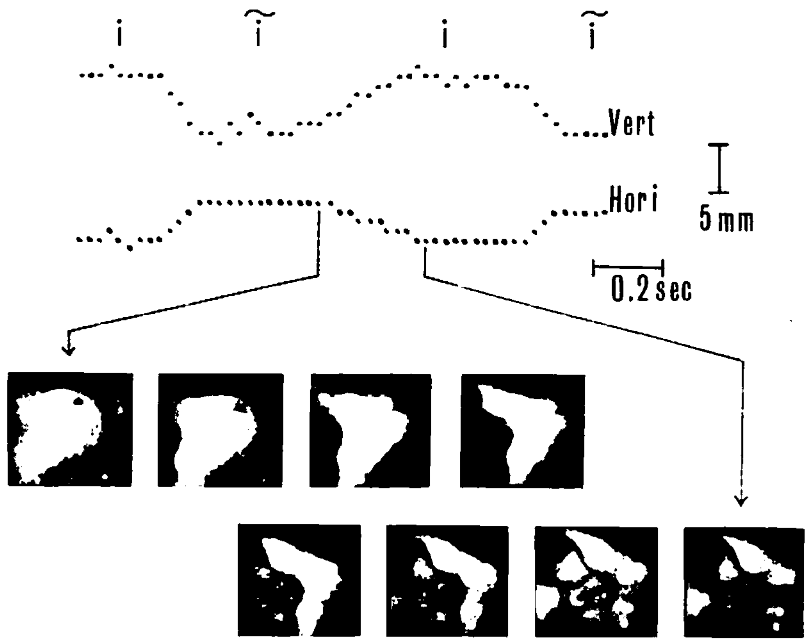

Fig. 3. Horizontal and vertical displacements of the fiberscope associated with the velar movement during the pronunciation of the test sounds sequence [i: ĩ: i: ĩ: . . .].

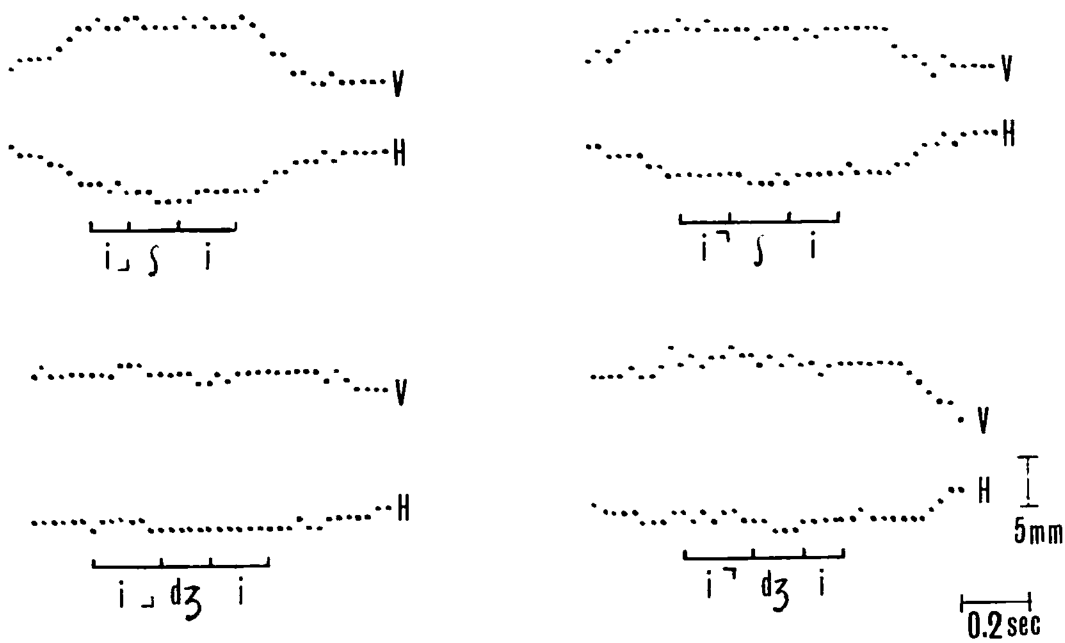

Fig. 4. Horizontal and vertical displacements of the fiberscope observed for the utterances of the isolated words.

in the y-direction. It may be assumed that there were no appreciable changes in the laryngeal gestures during this pronunciation,* and that the apparent movement of the image was mainly due to the movement of the fiberscope. The magnification factor of the recorded image varies with the distance between the object plane and the object lens of the fiberscope. This change, due to the movement of the fiberscope, was estimated by this result of x-ray observation at approximately 20% for the conditions of this test utterance.**

(2) An apparent movement of the image in the picture field was observed during the [ĩ] to [i] transition. It was estimated that the position of the optical axis of the fiberscope on the object plane moved approximately 12 mm during the 4 mm x-displacement of the tip of the fiberscope. The image movement was thus largely due to a change in the direction of the fiber axis rather than its displacement. The cable of the fiberscope can be considered as an elastic bar which is pressed against the back pharyngeal wall by the rear edge of the velum. As the velum is elevated and the nasopharyngeal path is closed, the fiberscope not only moves upward but also the curvature of the bending near the nasal port becomes somewhat more marked. This would cause a clockwise rotation of the fiberaxis at the tip in a situation like Fig. 1.

(3) In the case of [iʃi] shown in Fig. 5, there was observed just after the end of the utterance a movement of the fiberscope that was apparently associated with the velar gesture towards its respiratory position. The magnitude of this displacement was about the same as that observed during the pronunciation of [i: ĩ: . . .]. A similar velar movement is seen in the beginning of this utterance. For [idʒi], no movement was observed before the utterance and the movement after the utterance was smaller compared to that of [iʃi]. Generally, there were considerable inter-utterance variations in

---

* No discernible pitch change accompanied the pronunciation.

** It is generally the case that the distance between the fiberscope and the vocal cords is 3-4 cm and the size of the image field is 2 x 2 cm^2 at a distance of 3.5 cm.[2)] The changes in the magnification factor and the displacements of the image due to movements of the fiberscope were estimated by assuming this lens-object distance. Details of these correction factors do not affect the present qualitative discussions.

the magnitude and timing of these mode-transitional movements.

(4) During the pronunciation of the isolated words, slight forward movements of the fiberscope (x-direction) were often observed around the middle of the utterance, but no consistent movement in the y-direction was observed, except for some occasional (apparently random) displacements. Y-displacements greater than 1 mm were observed in two (of eight) utterances. The change in the magnification ratio of the image that would be caused by this 1-mm movement in the y-direction is approximately 3 percent.

Within each utterance, the apparent size of the vocal cords varied considerably. In order to obtain a tentative estimation of the apparent length of the vocal cords, the two extreme points of the visible portion of the vocal cords illustrated in Fig. 5 (a) were examined and the distance was measured frame by frame on the phtographic images. The position of the point A, which represents the anterior commissure of the vocal cords, was measured for those frame images in which the view of this point was not observed by the epiglottic tubercle. Measurement of the point B was made only when the vocal cords were in the adducted position.

(5) The apparent length of the vocal cords was always short at the beginning of the utterance. For the utterances of [i˩dʒi] and [i˩ʃi], it became maximum during the initial period of the second vowel. For [i˥dʒi] and [i˥ʃi], the shortening of the vocal cords appears to start during the consonantal period. Near the end of the consonant towards the second vowel, the epiglottis moves backward (towards the center of the image field) to cover the frontal portion of the vocal cords. Generally, a 20-40% change of the apparent vocal cords length was observed during the course of individual utterance. The possible error due to the change of the image magnification as stated above is of course too small to account for this change, and the apparent lengthening of the vocal cords is considered to reflect the laryngeal gestures associated with the pitch control, viz., the change in the height of the larynx and/or the actual lengthening of the vocal cords (see the pitch curves in Fig. 5 (c)).

(6) It is difficult, with the measurement accuracy presently available, to compare the movements of the point A and B for different utterances, since the effect of the 1-2 mm x-displacement of the fiberscope may not be negligible. If we assume that the change of the direction of the fiberscope axis could be expressed as a function of the x-displacement, apparent movement

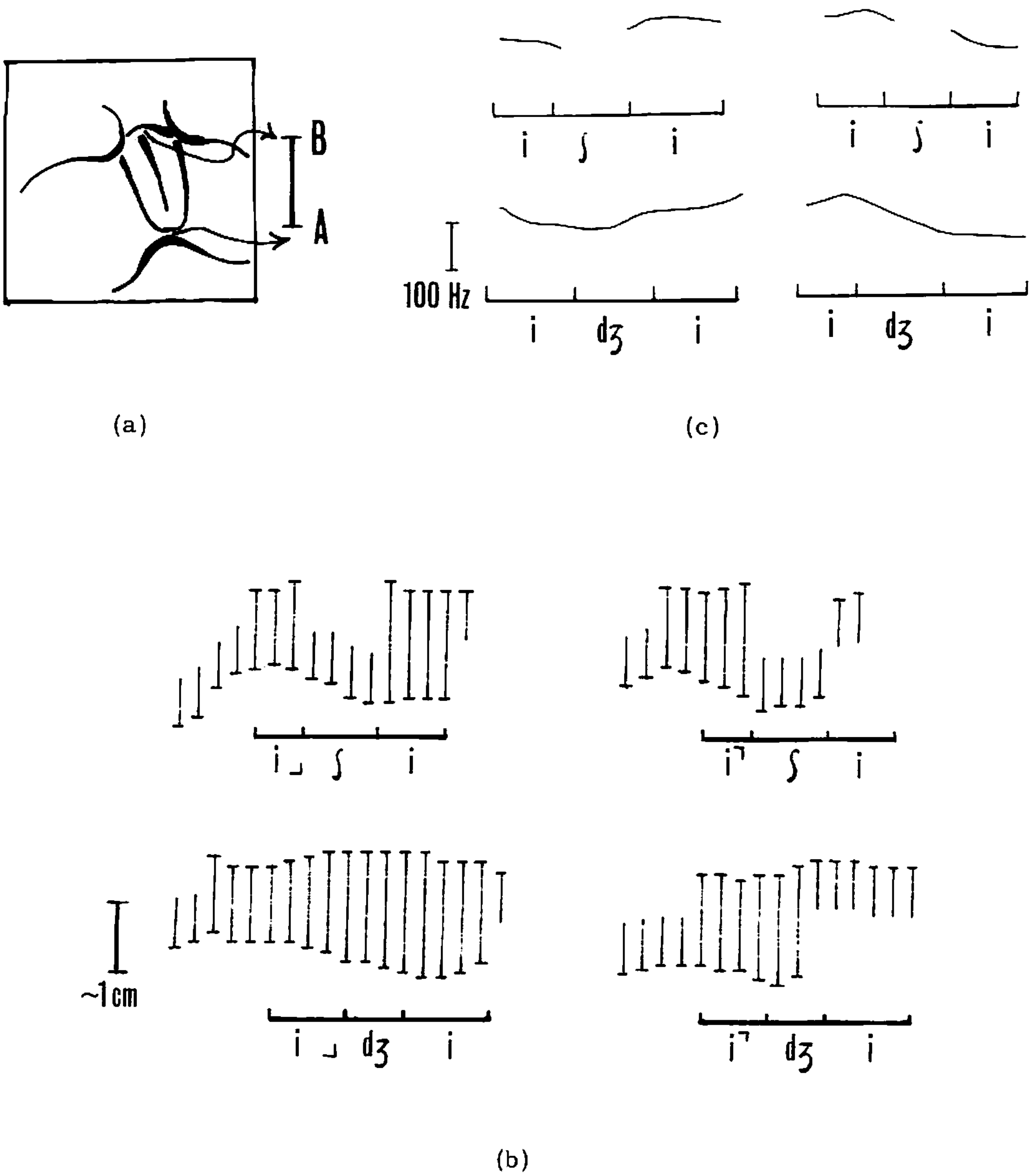

Fig. 5. Tentative measurements of the time variations of the apparent vocal cord length.

(a) Two extreme points of the visible portion of the vocal cords selected for measuring the apparent vocal cords length.

(b) The vertical distance between the point A and B and their relative positions in the image field are measured on each movie frame and plotted against time (24 frames per second).

(c) Pitch curves traced on the narrow band sound spectrograms.

of the image per 1 mm x-displacement was estimated at approximately 3 mm for the pronunciation [i: ĩ: i: ĩ:] (see supra). It can be said, however, that the up-down movement of the larynx can be effectively observed by measuring the magnification factor of the image and estimating the changes in the distance between the vocal cords and the fiberscope, if we can find some undeformable shape on the object plane for scaling.

References

1) S. Kiritani, "X-Ray Monitoring of the Position of the Fiberscope by Means of Computer Controlled Radiography," Annual Bulletin (Research Institute of Logopedics and Phoniatrics, University of Tokyo) No. 5, 35-39 (1971).

2) M. Sawashima and T. Ushijima, "Use of the Fiberscope in Speech Research," Annual Bulletin (Research Institute of Logopedics and Phoniatrics, University of Tokyo) No. 5, 25-34 (1971).