

Problem: Current research in the field of vocal physiology uses high speed video imaging to analyze the motion of the vocal folds during phonation. High speed digital cameras capture images of the vocal folds at rates up to 4000 frames per second. In order to gain knowledge of the movement of the vocal folds over time, the edges of the folds must be extracted from series' of several thousand images automatically. From this the area of the glottis, the opening between the vocal folds, can be calculated with respect to time. This data is then used to find irregularities in the vocal fold motion, and determine how irregular the motion is.

Currently, several methods can be used to detect the glottal edges, including histogram based adaptive thresholding methods, region growing methods, and active contour methods. However, each of these methods contains flaws, and improvement is necessary. I will discuss the advantages and disadvantages of each of these methods, and then investigate a new method for the extraction of vocal fold edges.

Approach: In our new method we use differentiation along each pixel column of the vocal fold image. In our video data, the glottal axis is always oriented horizontally in the image frame. Thus we can assume that there are at most two edges of the glottis lying in each pixel column of the grayscale image. Thus we take the derivative of each column, and conclude that the edges are located at the points where the derivative is maximum or minimum. This is where the column changes from light (vocal fold) to dark (glottis). Using these edges, a binary image is created where the glottis is white (0) and the vocal folds are black (1). However, due to blurring along the vocal fold edges, the edge varies from column to column, making the composite edge very jagged. To address this, a Canny edge detection operator is applied to the binary image to smooth the edge. From this smoothed edge, the glottal width is calculated from each column. The glottal width curve is filtered to address differentiation errors that occur when there is a bright spot in the image. After filtering, the glottal width curve is integrated to determine the glottal area. This process is applied to thousands of frames so that data can be analyzed. Results of each step of the process will be shown, and results from several vocal fold videos will be shown.

References:

Marendic, B.; Galatsanos, N. and Bless, D. "A new active contour algorithm for tracking vibrating vocal folds." *Proceedings 2001 International Conference on Image Processing.*, pt. 1, vol.1, pp. 397-400. 2001.

Yuling, Y., Chen, X., and Bless, D. "Automatic tracing of vocal fold motion from high speed digital images." *IEEE Trans. Biomed. Engr.* Vol. 53-7 pp. 1394-1400. Jul 2006.

Wittenberg, T., Moser, M., Tigges, M., and Eyesholdt, U. "Recording, processing, and analysis of digital high -speed sequences in glottography." Vol. 8 pp. 399-404. 1995.

Xin Chen, Bless, D., Yuling Yan. "A segmentation scheme based on Rayleigh distribution model for extracting glottal waveform from high-speed laryngeal images." *2005 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4, 2006.