Integration of Ultra High-Speed Color Videoendoscopy with Time-Synchronized Measures of Vocal Function

Daryush D. Mehta, SM[1,2]; Dimitar D. Deliyski, PhD[3]; Steven M. Zeitels, MD[1,4]; Matías Zañartu, MS[5]; and Robert E. Hillman, PhD[1,2,4]
[1]Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital; [2]Speech and Hearing Bioscience and Technology Program, Harvard-MIT Division of Health Sciences and Technology; [3]Department of Communication Sciences and Disorders, University of South Carolina; [4]Harvard Medical School; [5]School of Electrical and Computer Engineering, Purdue University

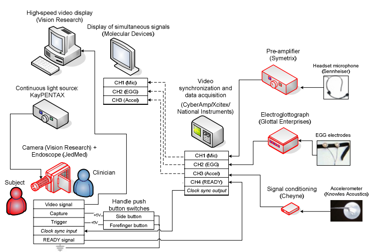
Objective

To develop an **innovative voice assessment system** that can provide new insights into mechanisms of normal and disordered voice production by integrating advances in **ultra high-speed color imaging** with objective **measures of vocal function**.

Study Design

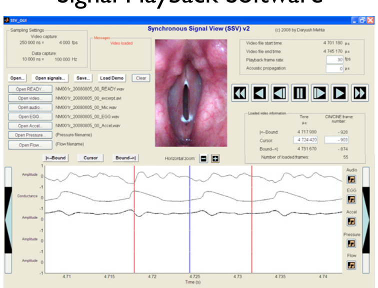
Iterative development and testing of an integrated system that simultaneously acquires ultra high-speed color endoscopic images of vocal fold vibration with time-synchronized recordings of the acoustic, electroglottographic, and neck accelerometer signals.

Synchronous Video and Signal Acquisition



Approaches for time-synchronization of simultaneously recorded signals were established, verified, and implemented using a high-precision hardware/software solution, displayed in the schematic above.

Synchronous Video and Signal Playback Software



A graphical user interface has been developed to visualize the high-speed video and synchronous signal data in an integrated playback format.

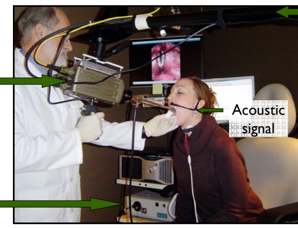
This allows the user to **cross-correlate observations** between video images and signals as a basis for in-depth investigations of how vocal fold vibratory motion relates to the production of sound. Of special interest is better delineation of relationships between **asymmetries** in vocal fold vibration and the acoustic generation of **noise/perturbation** in disordered voices.

High-Speed Videoendoscopy Setup

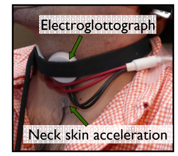
High-speed video camera

- Up to 10,000 frames per second
- 320 pixels x 480 pixels
- Approximately 0.05 mm^2 per pixel
- 42-bit RGB color

Xenon light (300 watt)



Laryngeal recordings



The system was built around a new high-speed digital camera that has sufficient light sensitivity to record high-resolution color images through a transoral endoscope at rates up to 10,000 frames per second. Specially-designed devices for suspending and controlling the 7 lb. camera were fabricated to facilitate its uses for endoscopy.

Examples of Quantitative Voice Assessment

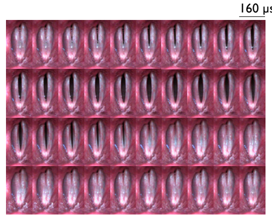
Normal phonation

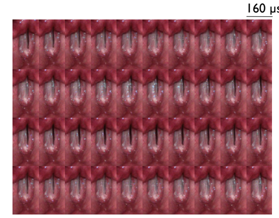
Adult male, no history of voice disorders

Disordered phonation

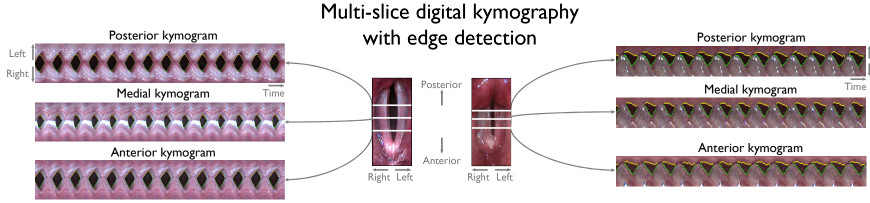
Adult female, post-operative recording following surgical removal of hemorrhagic polyp on left vocal fold

High-speed videoendoscopy image sequence





Multi-slice digital kymography with edge detection



Kymographic measures of vocal function from multiple slices

Posterior	Medial	Anterior		Posterior	Medial	Anterior
14.4%	24.0%	36.6%	Left-right amplitude asymmetry	77.1%	75.9%	80.9%
1.7%	4.1%	2.5%	Left-right phase asymmetry	27.5%	22.5%	29.8%
7.2%	8.0%	11.1%	Axis shift	5.9%	8.2%	6.2%
0.04	0.05	0.08	Integrative asymmetry	0.43	0.43	0.53
54.0%	53.5%	38.4%	Open quotient	79.7%	65.8%	77.5%
151 Hz	151 Hz	151 Hz	Fundamental frequency	221 Hz	221 Hz	221 Hz
1.1%	1.1%	1.4%	Period irregularity (jitter)	3.3%	1.9%	3.6%

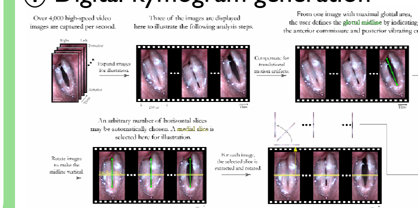
Acoustic measures of vocal function

30.6 dB	Harmonics-to-noise ratio	26.6 dB
0.16%	Jitter	0.32%
1.5%	Shimmer	3.5%
151 Hz	Fundamental frequency	221 Hz

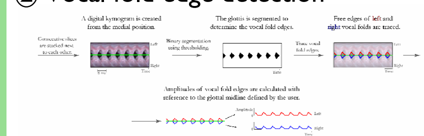
Above are examples of simultaneously-obtained measures from multi-slice kymographic processing of the high-speed endoscopic images and standard acoustic analyses of the microphone signal for one normal and one disordered voice.

Digital Processing of Endoscopic Images

① Digital kymogram generation

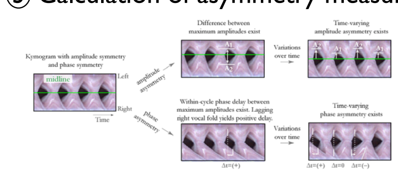


② Vocal fold edge detection



Vocal fold amplitude waveforms (red and blue) extracted from the high-speed video images enable the quantification of cycle-to-cycle vocal fold vibratory (spatial) asymmetry and (temporal) irregularity measures.

③ Calculation of asymmetry measures



- **Amplitude asymmetry**: Difference in the amplitudes of the left and right vocal folds. Amplitude differences potentially vary from period to period, and this time-varying phenomenon can be quantified.
- **Phase asymmetry**: Phase difference in each period can be quantified by measuring time delays between maximum abduction of the left and right vocal folds.
- **Axis shift**: Difference in the location of the glottal midline from opening to closure.
- **Integrative asymmetry**: Asymmetry measure reflecting the combined effects of amplitude asymmetry, phase asymmetry, axis shifts, and vocal fold midline crossings. Derived from the arithmetic difference between the vocal fold amplitude waveforms.

Conclusions

This new system makes possible the application of more powerful digital image and signal processing approaches to better elucidate and quantify vocal fold sound source mechanisms, particularly in cases where vocal fold vibration is irregular and/or asymmetric.

Acknowledgements

This project is supported by grants from the **NIH National Institute on Deafness and Other Communication Disorders** (T32 DC00038 and R01 DC007640) and by the **Institute for Laryngology and Voice Restoration**. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. We would like to thank **Thomas F. Quatieri** for valuable discussions regarding signal and image processing algorithms.