Title: Spatiotemporal Fourier analysis of high-speed videoendoscopy
Authors: Yue Gao[1]; Yu Zhang[2], Ph.D.; Alicia J. Spreche[2]r; Jack J. Jiang[2], M.D., Ph.D.
Running Title: Spatiotemporal Fourier analysis of HSV
Article/Category: Original Research
Section/Category: Voice

[1]Department of Computer Science, University of Wisconsin, Madison, WI, 53706
[2]Department of Surgery, Division of Otolaryngology - Head and Neck Surgery, University of Wisconsin School of Medicine and Public Health, Madison, WI, 53792

Corresponding Author: Jack J. Jiang, M.D., Ph.D.
Mailing Address: 1300 University Avenue
5745 Medical Sciences Center
Madison, WI 53706
Email Address: jiang@surgery.wisc.edu
Phone Number: 608-265-9854
FAX: 608-265-2139

ABSTRACT

Objective/Hypothesis: In voice analysis, fundamental frequency (f_0) is of great importance. Many applications such as Electroglottography and videokymography analysis have applied temporal f_0 as an indicator of the voice condition. Since vocal fold vibration exhibits spatial as well as temporal dynamics, f_0 may reflect spatial characteristics of the vocal fold in addition to its temporal vibratory pattern. Therefore, understanding variations in both dimensions is essential for studying voice production in laryngeal pathology. This study applies a Spatiotemporal Fourier Analysis (SFA) method to high-speed videoendoscopy (HSV) data to delineate spatial and temporal dynamics simultaneously. *Study Design*: This was a methodological study designed to provide a unique approach to the interpretation of HSV data. *Methods*: SFA was applied to examine the vibration of ten normal and ten pathological voices. The variable spatial bandwidth was defined to categorize the spatial distribution of the f_0. *Results*: Patients with vocal fold nodules show more complex spatiotemporal Fourier spectrum properties than normal subjects. The spatial bandwidth derived from SFA was significantly different between normal and pathological groups ($p<0.001$). *Conclusions*: SFA applied to HSV data represents an improvement over traditional Fourier analysis as SFA simultaneously delineates both the spatial and temporal characteristics of the vocal fold vibration and the spatial distribution of the f_0. The spatial bandwidth derived from SFA can successfully differentiate normal and pathological voices. These findings indicate that SFA could be helpful in the clinical assessment of laryngeal pathology.
Keywords: High-Speed Videoendoscopy; Spatiotemporal; Fourier Transformation, decision support parameter

I. INTRODUCTION

Electroglottography (EGG) and acoustic analysis represent commonly used noninvasive voice recording methods which compliment the direct observation of the vocal folds.[1-4] Objective parameters such as fundamental frequency (f_0), jitter, and shimmer can be derived from these two measurements. However, EGG and acoustic analysis are based solely on the change in time series and contain no spatial information, information which is important for vocal fold analysis. As a consequence, laryngeal visualization has received considerable interest in recent years. The most common clinical technique for laryngeal visualization is videoendoscopy with stroboscopy[5]. Although stroboscopy represents an improvement in laryngeal assessment, aperiodic vibrations make it inapplicable, as the system uses the measured f_0 to time the strobe light.[6, 7] Another recording tool is videokymography (VKG). It has been used to examine the effects of elongation and subglottal pressure on the f_0 and mucosal wave with less computing time.[5, 8, 9] However, as the traditional VKG scans a single linear portion of the glottis, it provides only partial information on the spatial characteristics of the vocal fold. All these procedures lack the ability to reveal both the spatial distribution and the f_0 of the laryngeal dynamic motion.

High-speed videoendoscopy (HSV) has shown the potential to resolve this problem.[5] HSV provides a high-speed, frame-by-frame based image recording, which presents more parameters for classification and diagnosis. It captures the temporal as well as spatial vibratory properties of the vocal folds irregardless of periodicity, overcoming the limitations of videostroboscopy and VKG.[5, 7, 10, 11, 14] Consequently, HSV has emerged as a popular tool for the assessment of voice disorders. Larsson et al developed an analysis system based on HSV and acoustic data that studied the relationship between vocal fold vibration and associated airflow.[13, 14] Braunschweig et al. described different types of oscillation and their onset mathematically to classify HSV data from subjects with functional voice disorders.[12] Meanwhile, using HSV data, Bless et al. implemented active contour algorithms and automatic tracing methods to extract the change in glottal area over time for normal and pathological subjects.[10, 16] Jiang et al. introduced an interpolation method of extracting mucosal wave parameters from HSV data based on a simulation of VKG information.[9] Few studies have applied f_0 to HSV to combine extensive spatial information with time series analysis.[1-3, 18-20] Therefore, developing a method which reveals both spatial and temporal characteristics and the spatial distribution of the fundamental frequency is necessary for obtaining global instead of local representations. This study proposes the use of Spatiotemporal Fourier analysis (SFA) to generate a spatiotemporal representation of HSV data and a spatial range of the f_0. SFA was applied to HSV data from normal subjects and subjects with vocal fold nodules. We define the spatial bandwidth to represent the spatial distribution of the f_0. By comparing SFA with temporal Fourier analysis (TFA), particularly when analyzing nodule subjects, we hypothesize that vocal fold vibratory behavior in spatial domain can be more clearly revealed.

II. MATERIALS AND METHODS

A. Material Collection Procedures and Image Segmentations.

The protocol used in this study was approved by the IRB of Fudan University, EENT Hospital Shanghai, China. Twenty subjects participated; ten had normal voices and ten

had vocal nodules as diagnosed by attending physicians. The gender, age and diagnosis of the subjects are shown in Table 1. All subjects were asked to sustain the vowel /a/ at a conformable pitch for as long as possible. A high-speed digital camera (Photron fastcam ultima SE, San Deigo, CA) was used to record vocal fold vibrations at a rate of 4000 frames per second with the resolution of 120×256 pixels. Matlab 7.3 was used for pre-processing the HSV data input and obtaining the Spatiotemporal Fourier analysis pattern.

Table 1: Detail information of the twenty human subjects including ten normal and ten pathological subjects with vocal fold nodules. The table includes their age, gender, and diagnosis result.

Diagnosis	Gender Ratio M:F	Mean Age	Total number
Normal	2:3	39.3 +/- 15.6	10
Nodules	2:3	38.5+/- 10.6	10

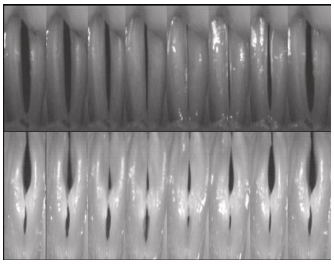


Figure 1: HSV images of one cycle of vocal fold oscillation. The upper panel of the image consists of eight frames collected from a normal subject. The lower panel contains eight frames collected from a subject with vocal fold nodules

Figure 1, which contains a normal subject at the upper panel and a nodule subject at the lower panel, shows eight consecutive frames of HSV. An automatic threshold based edge detection method, in which pixel intensities above the threshold are considered to be background and pixel intensities below the threshold are considered to be the open glottal phase, was applied on a frame-by-frame basis.[21] From the anterior side (thyroid cartilage) to the posterior side (arytenoid cartilage), the image of the glottis was divided into M horizontal segments ($M >>1$). Each segment shows a temporal oscillatory behavior. Thus, the spatiotemporal data from HSV can be described as,

$$u(j,t_n)\ ,\quad j=1,2,\cdots,M \ \text{ and } \ t_n=n\Delta \ \ (n=1,2,\cdots,N), \qquad (1)$$

where j denotes a particular spatial region, t_n denotes the sampling time with the interval $\Delta=1/f_s$, and N is the total number of frames. $u(j,t_n)$ is a single line through the glottis at time t_n. From the spatiotemporal data, $u(j,t_n)$, the glottal area can be described as,

$$s(t_n)=\sum_{j=1}^{M}u(j,t_n), \qquad (2)$$

which follows the spatial changes over time during vocal fold vibrations.

B. Temporal Fourier Analysis

The traditional approach of laryngeal image analysis uses vocal-fold edge segmentation and glottal area calculation. For the glottal area time series *s(t),* the temporal Fourier transformation can be applied to describe the frequency [23]:

$$T(u) = \frac{1}{N}\sum_{t=1}^{N} s(t)e^{-2\pi i\frac{u}{N}} \quad where\ u = 1,2,.....\mathrm{N} \qquad (3)$$

where $i = \sqrt{-1}$. *s(t)* is the glottal area at the given time *t*. $T(u)$ denotes the calculated amplitude for each frequency *u*, after the transformation. Every frequency level *u* is evaluated based on the entire input domain.

C. Spatiotemporal Fourier Analysis

SFA is derived from a Two-Dimensional Discrete Fourier Transform, which is an extension of the TFA analysis. The output of the transformation is an image containing both the temporal frequency domain and the spatial frequency domain. It is generated from spatiotemporal information $u(j,t)$ instead of the glottal area $s(t)$. The function below is responsible for the generation of the two dimensional SFA [23].

$$F(u,v) = \frac{1}{MN}\sum_{j=1}^{M}\sum_{t=1}^{N} u(j,t)e^{-2\pi i(\frac{uj}{M}+\frac{vt}{N})} \qquad (4)$$

$$for\ \mathrm{j} = 1,2,.....M\ \text{and}\ \mathrm{t} = 1,2,.....N$$

The spatiotemporal graph is essentially a matrix with *t* as the time axis and *j* as the spatial axis. The function $u(j,t)$ returns the open glottal length for a specified time *t* and spatial column value *j*. N is the total number of frames and M is the total number of spatial regions. The temporal frequency and spatial frequency are represented by *u* and *v*, respectively. The SFA performs a one-dimensional Fourier Transformation with each value of the spatial domain and another one in the temporal domain, on the resulting values.

F_0 represents an important parameter for the assessment of laryngeal pathology.[18-20, 22] Physical abnormalities of the vocal fold may have an impact on the f_0 and its spatial distribution. In order to quantify the energy distribution of the glottal motion after applying the SFA, we defined the spatial bandwidth δ of the f_0 as

$$\delta = \left| S^{-1}(0.5\times\max(F(f_0,x)\big|_{x>0}) - S^{-1}(0.5\times\max(F(f_0,x)_{x\le 0})\right| / L \qquad (5)$$

where *L* is the glottal length, $S(x)$ is the function of the frequency in the spatial domain which returns the corresponding amplitude of the frequency *x*. $S^{-1}(\varepsilon)$ maps the amplitude ε to a set of corresponding frequencies on the spatial domain, and δ represents the spatial bandwidth at the f_0. Figure 2 is a demonstration of equation 5 at the f_0.

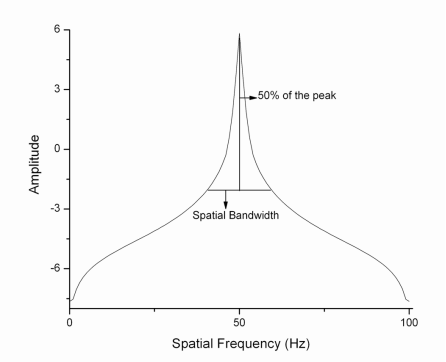


Figure 2: Graphical demonstration of spatial bandwidth variable. This measure focuses on the width of the f_0 peak instead of the amplitude (height). Spatial bandwidth characterized the spatial distribution of the f_0.

In order to show that a SFA transformation can be effectively applied to HSV, we consider the following synthesize data

$$f(x,t) = \max(\sin(2\pi f_0 \cdot t)\cdot \sin(\frac{\pi x}{L}),0), \quad t \in (0,0.1), x \in (0,10)$$

where *f(x,t)* is a function of amplitude. In the temporal domain, N=1000 and M=100 in the spatial domain. The glottal length, L, is 10 pixels. Shown in the upper panel of figure 3 is the glottal area time series *s(t)* of the synthesized model. We recorded the temporal time series change and applied TFA to it. The arrow in the lower panel of figure 3 at 100 Hz indicates the f_0 of TFA. Furthermore, we included the spatiotemporal analysis of the synthesized data *f(x,t)* in figure 4 which contains four images. Figure 4a is a three dimensional model representing the glottal motion respect to a time-space domain and the repeating pattern is due to the natural periodical vibration. Figure 4b is a contour plot of the spatiotemporal model, the intensity of each pixel is the amplitude of the movement. After the SFA procedure, the result of the synthesized model is shown in Figure 4c. The x-axis is the temporal frequency in Hz, the y-axis is the spatial frequency and the z-axis is the amplitude. Similarly, figure 4d is the contour plot of the SFA analysis and the spatial bandwidth of the f_0 is 0.178. After studying the synthesized spatiotemporal model, we confirmed that SFA is capable of generating a measure the spatial range of the f_0 in a global view and that spatial bandwidth can reveal the spatial distribution of f_0. We, therefore, applied TFA and SFA methods to the previously described twenty human HSV samples.

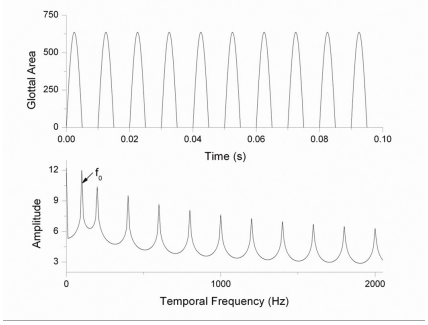
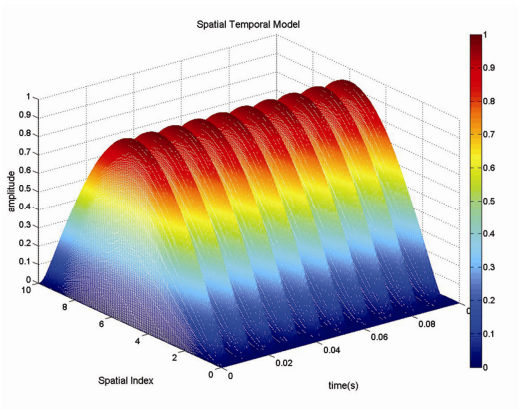
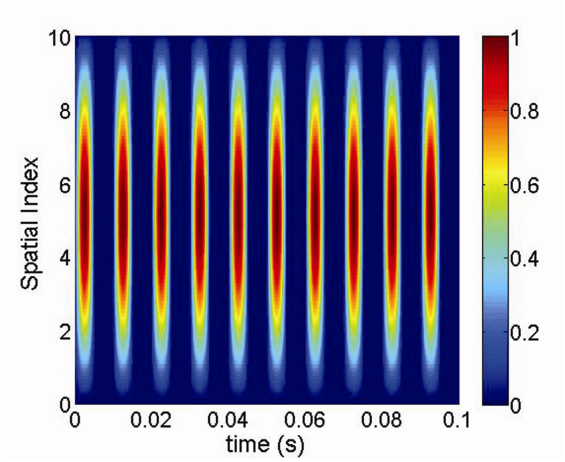


Figure 3: TFA transformation based on the synthesized model. The upper panel shows the glottal area over ten cycles before the transformation. The lower panel is the result of the TFA applied to the glottal area and the arrow indicates f_0.





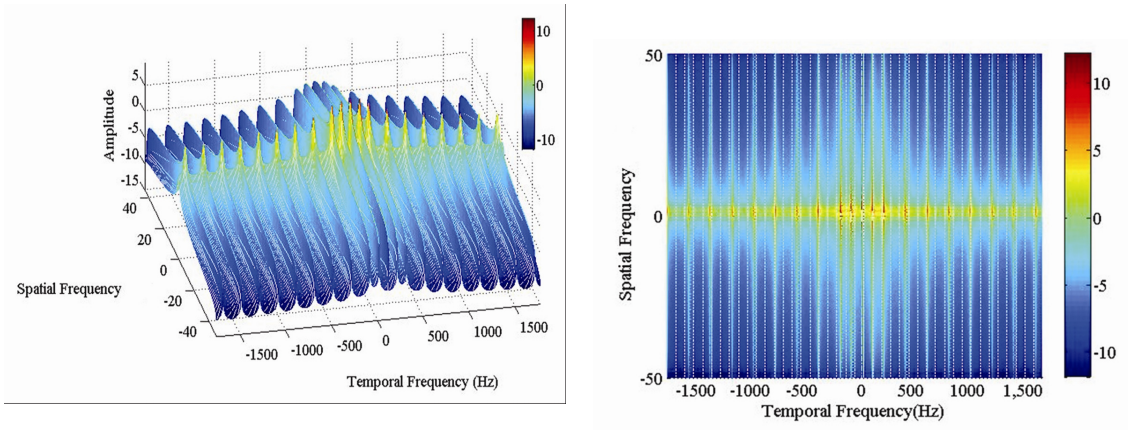


Figure 4: Figure 4a is a three dimensional model representing the glottal motion with respect to a time-space domain. Figure 4b is a contour plot based on the data seen in figure 4a, the intensity of each pixel is the amplitude of the movement. Figure 4c is the SFA transformation result in a three dimension domain. The x-axis is the temporal frequency in Hz, the y-axis is the spatial frequency and the z-axis is the amplitude. Figure 4d is the resulting contour plot from figure 4c.

D. Statistical Analysis

A Mann-Whitney Rank sum test was used to compare our two HSV data groups, as this test is robust to non-normally distributed samples [24]. The f_0 derived from TFA was evaluated based on its ability to differentiate normal and nodule subjects. In addition, spatial bandwidth derived from SFA was assessed by comparing the same twenty subjects.

III. RESULTS

A. One-dimension time series analysis

The results from the TFA from a normal subject are shown in figure 5. The upper panel is the glottal area which was extracted from approximately 300 consecutive frames through the application of regular threshold segmentation. The x-axis represents time in seconds, while the y-axis represents the overall glottal area as the total number of pixels in the open glottis. The periodic dynamic motion over time is clearly shown. The TFA result is presented in the lower panel of figure 5, where the x-axis is the temporal frequency and the y-axis is the corresponding amplitude. The first peak is called the f_0 which is estimated at 267Hz. The peaks are caused by the periodic vibration of the glottal area. Figure 6 displays the glottal area and TFA results of subject with vocal nodules generated using the same approach described above. For the nodule subject, the f_0 is at 304Hz marked by the arrow in the lower panel of graph 6.

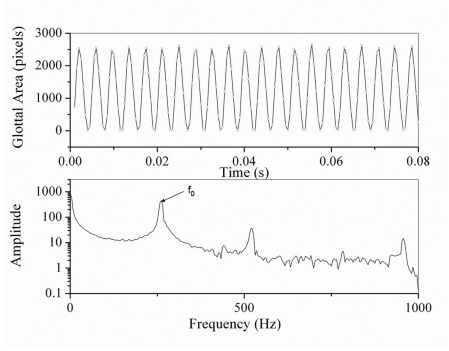


Figure 5: TFA results of a normal subject based on glottal area for consecutive 300 frames. The upper panel is the glottal area and the lower panel is the TFA transformation. The arrow indicates the f_0 which is at 267Hz.

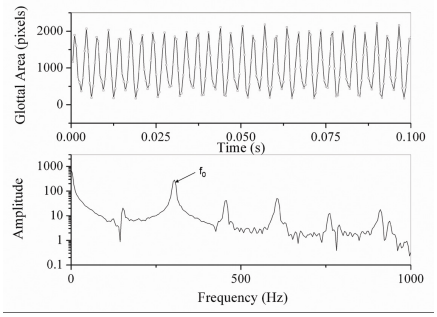


Figure 6: TFA results of a nodule subject based on glottal area for consecutive 300 frames. The upper panel is the glottal area and the lower panel is the TFA transformation. The arrow indicates the f_0 which is at 304Hz.

B. Two-dimension spatiotemporal series analysis

The spatiotemporal graph and SFA results for a normal subject are shown in figures 7 and 8, respectively. In figure 7, the x-axis is the temporal domain consisting of approximately 200 frames and the y-axis is the spatial domain including all the columns of the image. In figure 8, the x-axis is the temporal frequency and the y-axis is the spatial frequency. The intensity on each pixel represents the amplitude according to the color bar assignment. The spatial bandwidth is 0.082 indicated by the δ symbol in figure 8. Figures 9 and 10 are the spatiotemporal graph and SFA analysis results of a nodule subject with spatial bandwidth of 0.16. The calculations and display followed the same procedures as in figures 7 and 8.

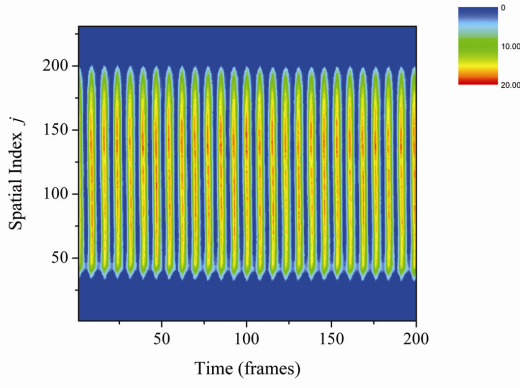


Figure 7: Spatiotemporal Analysis of a normal subject based on HSV images for approximately 200 frames. The x-axis is the time represented in frames and the y-axis is the spatial index of the glottis. The amplitude is reflected by the color of the image.

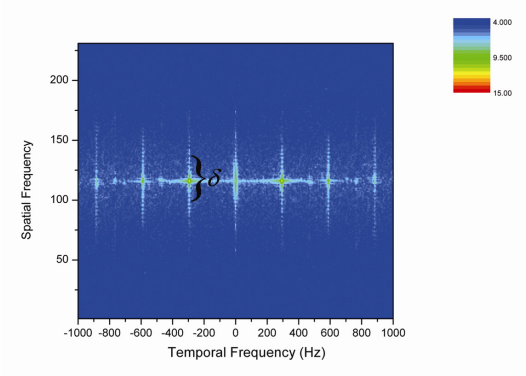


Figure 8: Spatiotemporal Fourier analysis results based on the data displayed in figure 7. The x-axis is the temporal frequency (Hz) and the y-axis is the spatial frequency. The color of the image represents different amplitude values. The bracket marks the spatial bandwidth of the f_0 which is equal to 0.082.

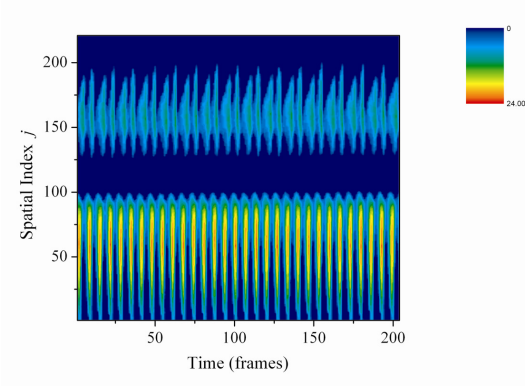


Figure 9: Spatiotemporal Analysis for a nodule subject based on HSV images for approximately 200 frames. The x-axis is the time represented in frames and the y-axis is the spatial index of the glottis. The colors of the image indicate the amplitude.

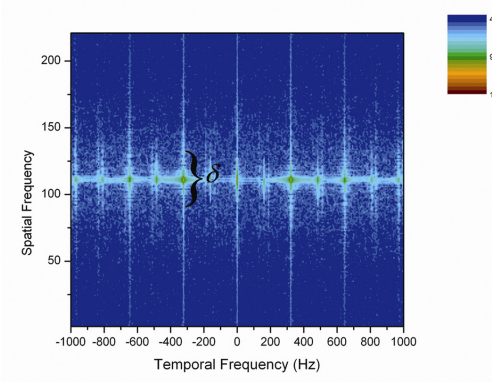


Figure 10: Spatiotemporal Fourier analysis result based on the data displayed in figure 9. The x-axis is the temporal frequency (Hz) and the y-axis is the spatial frequency. The color of the image represents different amplitude values. The bracket marks the spatial bandwidth of the f_0 which is equal to 0.16.

C. Statistical Evaluation of Spatial Bandwidth

Traditional analysis of f_0 is based on its amplitude which is indicated by the arrows in lower panels of figures 5 and 6 [16, 20]. We obtained the f_0 values from twenty subjects by applying TFA based on their glottal area; however, a Mann-Whitney Rank sum test revealed no statistical significance between the normal and the abnormal subjects ($p=0.65$). The comparison is shown in figure 11. A bar plot of the spatial bandwidth, δ, for the same twenty subjects as above is shown in figure 12. The mean δ for normal and nodule subjects is 0.0833 and 0.154, respectively. The Mann-Whitney Rank sum test result shows a significant difference between the normal and nodule patients with $p<0.001$.

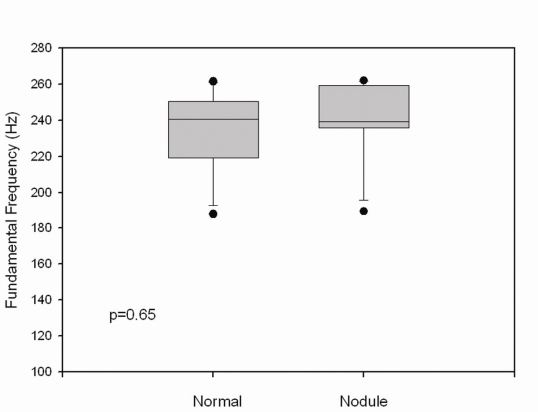


Figure 11: Bar plot based the TFA results using the glottal area of ten normal and ten nodule subjects. The boxes indicate the amplitude of f_0 calculated using TFA. The line within the box indicated the median while the upper and lower edges of the box denote the first and third quadrants of the data. The whiskers indicate the range and outliers are represented as dots.

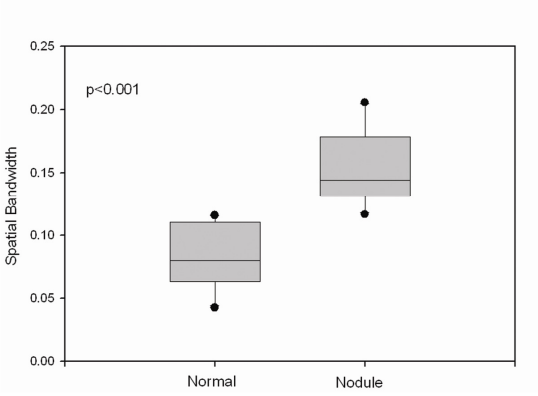


Figure 12: Bar plot based on the spatial bandwidth derived from SFA using of the same normal and nodule subjects analyzed in Figure 11. The mean spatial bandwidth for normal and nodule subjects is 0.0833 and 0.154, respectively. The line within the box indicates the median while the upper and lower edges denote the first and third quadrants. The range of the data is indicated by lines above and below the box.

IV. DISCUSSION

In this study, both Temporal Fourier Analysis (TFA) and Spatiotemporal Fourier Analysis (SFA) were applied to HSV data acquired from twenty subjects. A temporal frequency domain is defined through TFA based on glottal area extraction. Meanwhile, the SFA method maintains the spatial characteristics of the HSV data by generating both spatial and temporal frequency domains. These methods allow us to examine the relationship between vocal fold abnormalities and temporal and spatial behavior.

The importance of f_0 is emphasized in this study. During acoustic analysis, one of the basic tasks for computer laryngeal diagnostics is to determine the f_0 of sustained vowels. The calculation of the f_0 in the time domain is highly relevant in understanding the functionality of voice production.[18-20, 22] Furthermore, in image analysis, f_0 is an important indicator of voice condition. Many image processing tools have been developed to analyze HSV data clinically [23]. Some of these programs simulate kymography from HSV data by taking a single line scan of the image over time which provides a measure of f_0; however, this method does not incorporate the spatial information available in a HSV sample. Another procedure applies a Fast Fourier transformation on a single image, which loses the temporal aspect of the HSV data.[25] Although these analysis features are commonly incorporated as part of a decision support system for clinical diagnosis, neither feature combines spatial information with f_0. However, with our approach, we can provide a global overview of the dynamic

movement as well as the spatial distribution of the f_0. Using the spatial bandwidth variable, we offer a decision threshold to discriminate normal versus nodule groups.

Abnormal voice subjects often have an aperiodic vibration in the temporal domain; however, in our nodule subjects, the vocal fold vibration moved in a periodic manner over the time series. In the latter cases, TFA would not provide enough information to differentiate between normal and abnormal voices. Comparing the results in the lower panels of figures 5 and 6, the pattern of the f_0 on temporal domain reveals a periodical vibration; the abnormal spatial movement caused by the lesion on the glottal edge is not reflected in the transformation. As a consequence, traditional f_0 comparison fails to provide significance in differentiating normal and nodule subjects. A conclusion supported by the lack of a significant difference between normal and nodule subjects as shown in figure 11 ($p=0.65$).

In contrast, the SFA result is capable of detecting abnormalities in either the temporal or spatial domain. For normal voices, after SFA transformation, the frequency energy is more concentrated on the temporal domain and very few energy peaks are found in the high spatial domain. However, in pathological cases, the subject might appear to have abnormal vibrations either in the spatial or temporal dimensions, which usually leads to a high amplitude on the specified domains. Many nodule subjects displayed near-cyclic motion in the temporal dimension, while non-cyclic motion was detected in the spatial domain, as shown in the lower panel of figure 1. After the segmentation phase, all ten nodule subjects display an empty region observed in figure 9. The gap indicates that the tissue has separated the glottal area into two parts. The discontinuity in the spatiotemporal pattern leads to high amplitude in the spatial frequency domain. As the spatial bandwidth quantitatively delineates the spatial distribution of the f_0, it can differentiate samples based on these spatial abnormalities in addition to frequency aperiodicities. Our results support this assertion as there is a significant difference in spatial bandwidth between the normal and nodule groups indicated in figure 12 ($p<0.001$).

Although SFA is very promising, it cannot yet differentiate among various pathologies. The development of a method capable of assisting in the diagnosis of vocal fold pathology based on SFA could be a topic for future research. Meanwhile, the accuracy of SFA analysis is still limited by the quality of HSV video and image segmentation procedures. Overcoming such difficulties will depend on technological innovations.

V. CONCLUSION

In this study, we compared Spatiotemporal Fourier analysis with the tradition Temporal Fourier analysis approach of analysis HSV data. We discussed the advantages of SFA as it provides a spatiotemporal representation of HSV data as well as the spatial distribution of the f_0. In addition, we define the spatial bandwidth which shows significance when comparing to normal and nodule groups, demonstrating its potential as an indicator of voice condition. In contrast, we found that f_0 derived from glottal area cannot discriminate between pathologic and normal voices. Consequently, SFA could be helpful in clinical assessment because is reveals the global behavior of HSV recordings. The change of amplitude and overall trend in the high spatial domain are good indicators of vocal fold pathology. The proposed algorithm could be helpful in providing another evaluation tool, permitting clinicians to easily analyze HSV data.

ACKNOWLEDGEMENTS

The research was supported by NIH grant number 1-R01 DC05522 and DC008850 from the National Institute on Deafness and Other Communication Disorders.

REFERENCES

[1] Schoentgen J, Guchteneere,R. Time series analysis of jitter. *J phonetics* 1995:23: 89-102.

[2] Jiang JJ, Zhang Y, MacCallum JK, Zhou L. Objective Acoustic Analysis of Pathological Voices From Patients with Vocal Nodules and Polyps, *Folia Phoniatr Logo,* 2007; [In Press].

[3] Manfredi C, D'Aniello M, Bruscaglioni P, Ismaelli A. A comparative analysis of f0 estimation methods with application to pathological *Med Eng Phys*: 2000: 22: 135-47.

[4] Mitev P, Hadjitodorov S. F0 estimation of voice of patients with laryngeal disorders *Informatics and Computer Science*: 2003:156:3-19.

[5] Deliyski DD, Petrushev PP Bonilha HS. Clinical Implementation of Laryngeal High-Speed Videoendoscopy: Challenges and Evolution. *Folia Phoniatr Logo*. 2008:60:33-44.

[6] Schutte HK, Svec JG, Sram F. First results of clinical application of videokymography. *Laryngoscope.* 1998:108(8):1206-10

[7] Granqvist, S. Lindestad, P. A method of applying Fourier analysis to high-speed laryngoscopy. *J Acoust Soc Am* 2001:110(6):3193-7.

[8] Qiu Q, Schutte HK. Real-time kymographic imaging for visualizing human vocal-fold vibratory function. *Rev Sci Instrum*. 2007 ;78(2) :024302.

[9] J.J. Jiang, Chang CI, Raviv JR, Gupta S, Banzali FM Jr, Hanson DG. Quantitative study of mucosal wave via videokymography in canine larynges. *Laryngoscope*. 2000:110(9): 1567-73.

[10] Yan Y, Ahmad K, Kunduk M, Bless D. Analysis of Vocal-fold Vibrations from high-Speed laryngeal Images using a Hilbert Transform-Based methodology. *J Voice*. 2005:19(2):161-75.

[11] Lohscheller J, Eysholdt U. Phonovibrogram visualization of entire vocal fold dynamics. *Laryngoscop.* 2008:118(4):753-8

[12] Braunschweig T, Schelhorn-Neise P, Dollinger M. Diagnosis of Functional Voice Disorders by Using the High Speed Recording Technics. *Laryngo Rhino Otole.* 2008:87(5):323-30.

[13] Hertegard S, Larsson H, Wittenberg T. High-speed imaging: applications and development. *Logoped Phoniatr Vocol*: 2003:28(3):133-9.

[14] Larsson H, Hertegard S, Lindestad PA, Hamarberg B. Vocal fold vibrations: high-speed imaging, kymography, and acoustic analysis: a preliminary report. *Laryngoscope.* 2000:110(12):2117-22.

[15] Braunschweig T, Schelhorn-Neise P, Dollinger M. Diagnosis of functional voice disorders by using the high speed recording techniques. *Laryngo Rhino Otole.* 2008:87(5):323-30.

[16] Marendic B, Galatsanos N, Bless D. A new active contour algorithm for tracking vibrating vocal folds. *Proceedings of the IEEE International Conference on Image Processing.* 2001:1:397-400.

[17] Jiang JJ, Zhang Y, Ford C.N. Nonlinear dynamic of phonations in excised larynx experiments, *J Acoust Soc Am*. 2003:114: 2198–2205.

[18] Titze IR, Liang H. Comparison of Fo extraction methods for high precision voice perturbation measurements, *J Speech Hear Res* 1993:36(6):1120-33

[19] Bielamowicz S, Kreiman J, Gerratt BR, Dauer MS, Berke GS. Comparision of voice analysis systems for perturbation measurement. *J Speech Hear Res* 1996:39(1):126-34.

[20] Titze IR. Workshop on Acoustic Voice Analysis. *Summary statement, National Center for Voice and Speech, Denver*, CO 1995: 36.

[21] Zhang Y, Bieging E, Tsui H, Jiang JJ. Efficient and Effective Extraction of Vocal Fold Vibratory Patterns from High-Speed Digital Imaging. *J Voice*. 2008; [In Press]

[22] Manfredi C, D'Aniello M, Bruscaglioni P, Ismaelli A. A comparative analysis of f0 estimation methods with application to pathological voices. *Med Eng Phys*.2000: 22(2):135-47.

[23] Steven W. S. The Scientist & Engineer's Guide to Digital Signal Processing. *California Technical Publishing*. 1997.

[24] Bruning J, Kintz B. Computational Handbook of Statistics (4th Edition). *Allyn & Bacon publishing*. 1997.

[25] Kay Elemetrics Corp. - Description of Kay's Image Processing Software Description. Model 9181. http://www.kayelemetrics.com/Product%20Info/9700/KIPS%20overview.pdf