Chapter 8

Working with Wolf Ltd. HRES Endocam 5562 analytic system for high-speed recordings

Mette Pedersen, Martin Eeg, Anders Jønsson & Sanila Mamood

Abstract

We report on our experience of using Wolf Ltd. HRES Endocam 5562 analytic system for high speed recordings of the vocal folds (VF) in our clinical setting. The system uses high-speed videophotography and is able to capture 4000 or more images per second. It is therefore superior to clinical laryngo-stroboscopy in many areas of voice diagnostics. Such fast resolution of digital high-speed videophotography makes it possible to observe variability in individual VF kinetics and to observe the workings of the VF in pathological phonation. Post-recording analyses are performed using built-in software systems combined with kymography and electroglottography. Future development includes calculations of the lateral features and cycle-based features as well as phonovibrograms of the right and left vocal cord, from cycle to cycle, reproduced online for further software calculations.

Keywords: *High Speed Digital Imaging (HSDI), High speed films, vocal folds, quantitative voice measures, EGG, kymograms, phonovibrograms*

Introduction

High-speed (HS) recordings generate thousands of frames per second, hence this technology provides a powerful means for evaluating rapid movements, i.e. such as those happening during voice production. Hence, HS is a great tool for investigating the interrelations between the mechanical aspects of vocal fold (VF) vibration [1]. Results of HS recordings are subject to frame rate, pixel resolution, and to the dynamic ranges of the HS camera. In our clinic, we have used the Wolf Ltd. HRES Endocam 5562 analytic system (Richard Wolf GmbH, Pforzheimer Strasse 32, 75438 Knittlingen, Germany). Here we describe our experience in using this technology for the last eight years (Figure 1). The system can be delivered on the dedicated cart.

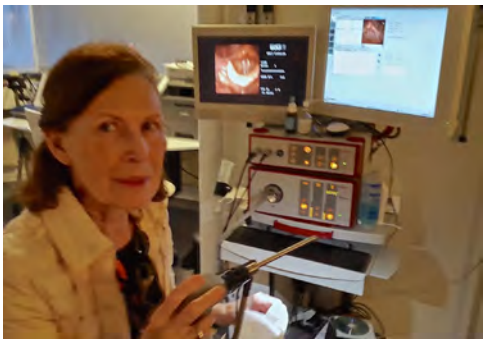

Figure 1. Mette Pedersen in front of her Wolf 5562 high--speed recording tower in her Copenhagen (DK) voice clinic.

Method

The typical recording speed we use is between 2000-4000 frames per second (f/s) for a 2 second recording duration with full color option, played back in a slow motion. Recording in full color, at this speed shows a dynamic resolution of 20 images per VF oscillation (at phonation of 200 Hz), allowing visualization of voice fluctuations and of aperiodic VF vibrations.

Using this system we can evaluate transient and rapid VF vibratory behaviors such as phonatory breaks, laryngeal spasms, and the onset and offset of phonation. The system can also be used to study vocal attack, coughing, throat clearing, laughing, and other activities involving rapid laryngeal maneuvers. It can be used to study different singing techniques and styles.

Post-exam software application permits immediate analysis of the recorded signals. This analysis can incorporate segmentation, kymography, acoustical curves, electroglottography (EGG), and can quantize open quotients. To create a representative kymograph picture, the Wolf software allows the user to place a line across any point of the glottis. Furthermore, movement analysis of the left versus the right VF is possible and a representative glottogram can be created. The user is in control of which regions of interest (ROI) are to be defined and plotted as a curve.

When EGG and acoustic signals are recorded simultaneously with the visual signals, delay between some of these signals is present. Figure 2 demonstrates the delays between acoustic and EGG tracing of a voiced signal.

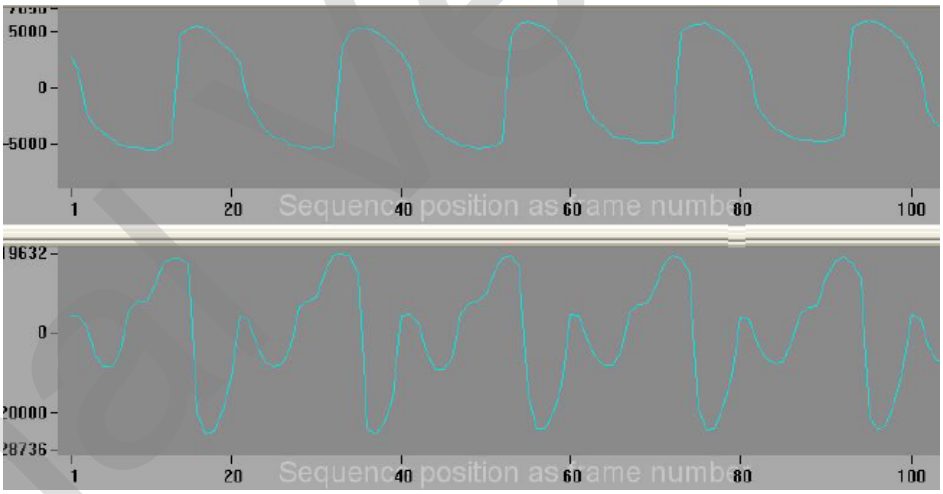

Figure 2. Acoustical delay compared to Electroglottography (EGG) for the normal function of VF. Electroglottography (upper tracing) shows the tone generator and the acoustical curve (lower tracing) shows the resonance.

Kymography

The Wolf system provides an excellent presentation of kymography. This is illustrated in Figure 3 in relation to the EGG signal. There is a direct connection between the closed phase on the EGG and the kymography signal. The closed phase of the EGG corresponds to the kymography, where the VF are seen from above.

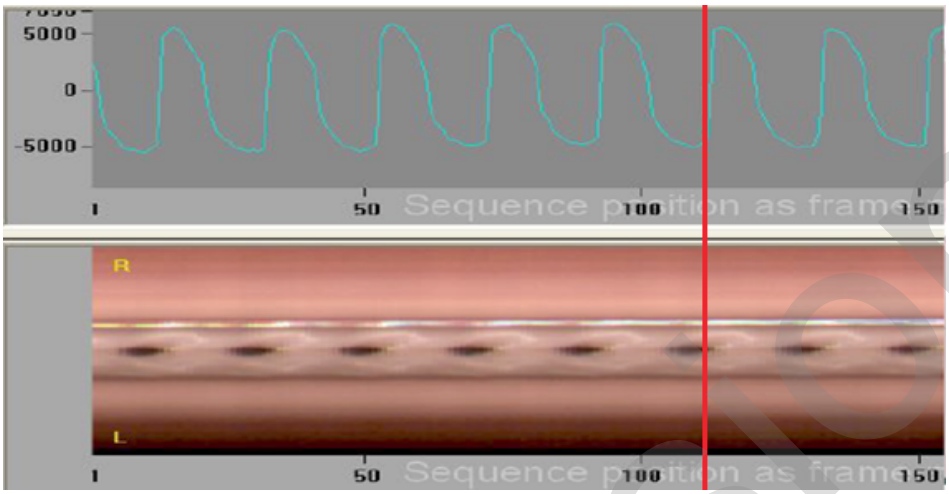

Figure 3. Good correlation between Kymography & EGG signals for VF activity.

Quantitative open quotients

The Wolf system also provides for a simple way to calculate the glottic quotients. Table 1 shows open quotients derived for both the male and the female subjects during different vocal tasks.

Table 1. The quantitative open quotients for a normal voice, based on high-speed frames up to 8000 frames in 2 seconds.

Range					
	Average	**Lowest**	**Highest**	**SD**	Normative values in our clinic, measured on 18 females and 12 males (aged 20-40 years) of high-speed films, sustained tone (/a/) for two seconds. Open quotient between the vocal cords in: - Front - Middle - Rear - Area Calculations
Open Quotient Front					
Male	0.45	0.14	0.92	0.32	
Female	0.48	0.37	1.0	0.49	
Open Quotient Center					
Male	0.51	0.09	1.0	0.27	
Female	0.58	0.12	1.0	0.29	
Open Quotient Rear					
Male	0.59	0.07	0.99	0.32	
Female	0.48	0.0	1.0	0.31	
Area Between Vocal Cords					
Male	0.6	0.04	1.0	0.43	
Female	0.68	0.13	1.0	0.3	

Clinical examples

We now present some of the results obtained with this system.
Normal Function – Male, trained amateur: analysis is shown in Figures 4a and 4b.

Figure 4a. High-Speed picture of a normal functioning voice of a male subject. The VF have smooth margins, they close nicely in the rear part, the arytenoid region is without edema, and passage to esophagus is not swollen.

Recording date:	2012-Jan-09	
Number of frames:	2243	
Image resolution:	256 x 256	Pixel
Recording speed:	4000	Bilder/s
Max. sound pressure:	77	dB
Min.sound pressure:	74	dB
Fundamental	142	Hz

Figure 4b. Shows recording data of the sound of the male glottis shown in Figure 4a.

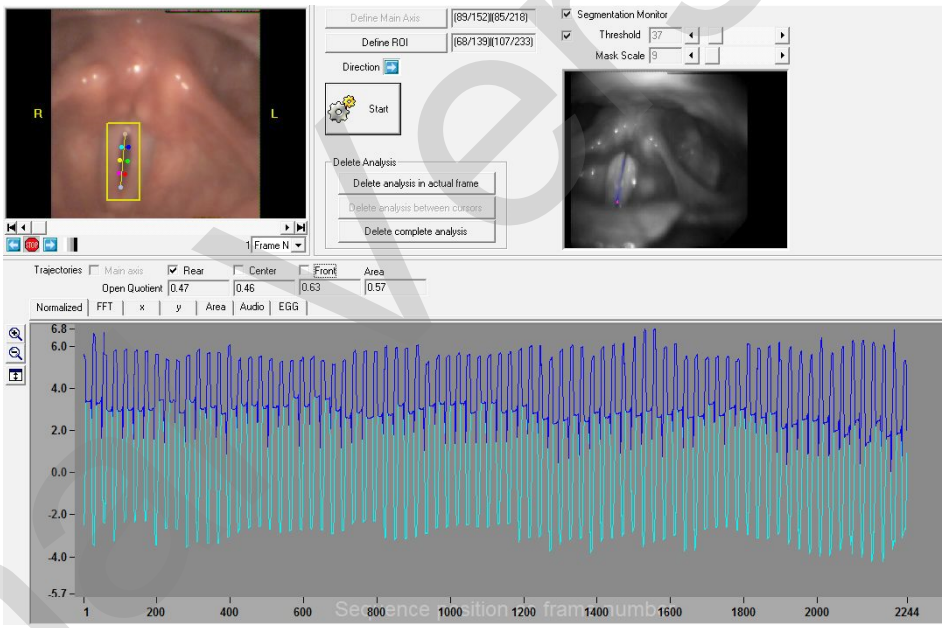

Figure 5. The system includes a routine possibility for marking of the edges and slow motion reproduction of the original video, which also can be shown in slow motion. The segmentation shown here is of a normal functioning voice with marking of the edges of the vocal folds—male. Based on the segmentation the opening quotients and the area are calculated (of the stored recording).

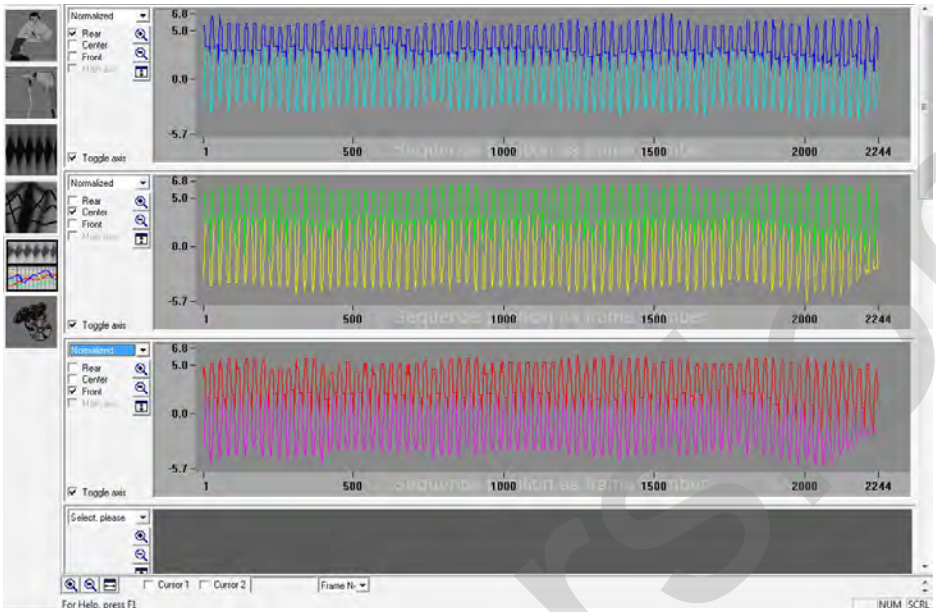

Figure 6. Segmentation of a normal functioning voice—male: right and left vocal fold presented in front, center, and rear parts (of the stored recording). There is a possibility in the software to zoom in and out for greater detail in both axes.

Electroglottography and other measures

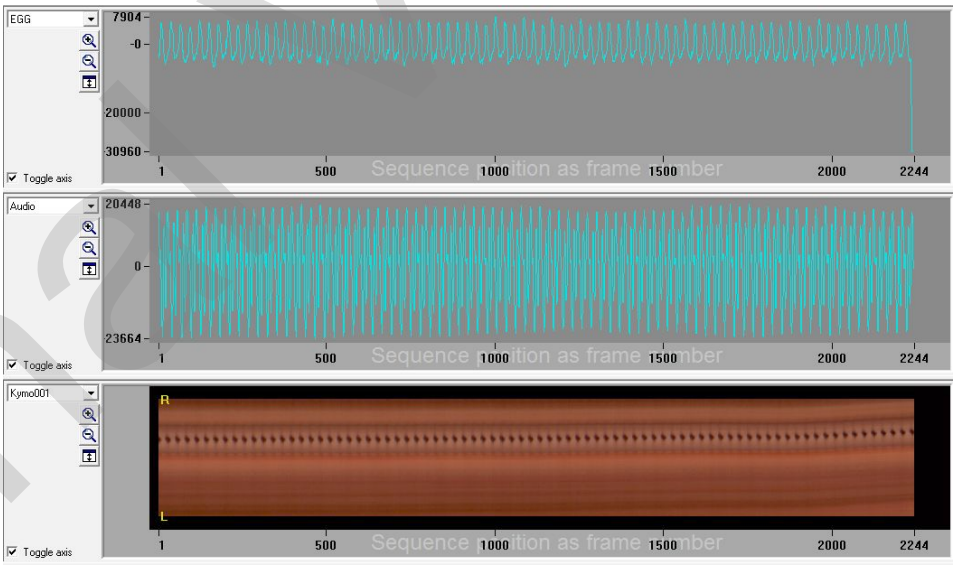

Figure 7. The EGG (upper tracing) shows the single curves of the VF in time at the same time as the acoustical wave (middle tracing) and the kymography picture (lower tracing) presenting the VF as seen from above—male.

Normal function – Female non professional

Figure 8a. High-Speed picture of normal functioning voice—female. The vocal folds have smooth margins, they close nicely in the rear part, the arytenoid region is without edema, and passage to esophagus is not swollen.

Recording date:	2009-Jul-21	
Number of frames:	2095	
Image resolution:	256 x 256	Pixel
Recording speed:	4000	Bilder/s
Max. sound pressure:	93	dB
Min.sound pressure:	93	dB
Fundamental	317	Hz

Figure 8b. Shows recording data of the sound of the female glottis shown in Fig 8a.

Analysis of the left vs. the right VF movements

The system allows for quick and efficient analysis of the differences in the vibratory movements of the left (L) and the right (R) VF movements. This is illustrated in Figure 9 and in Figure 10. Figure 9 shows how the location is chosen, while Figure 10 shows segmentation results at the three different locations, each for the L vs. R movements.

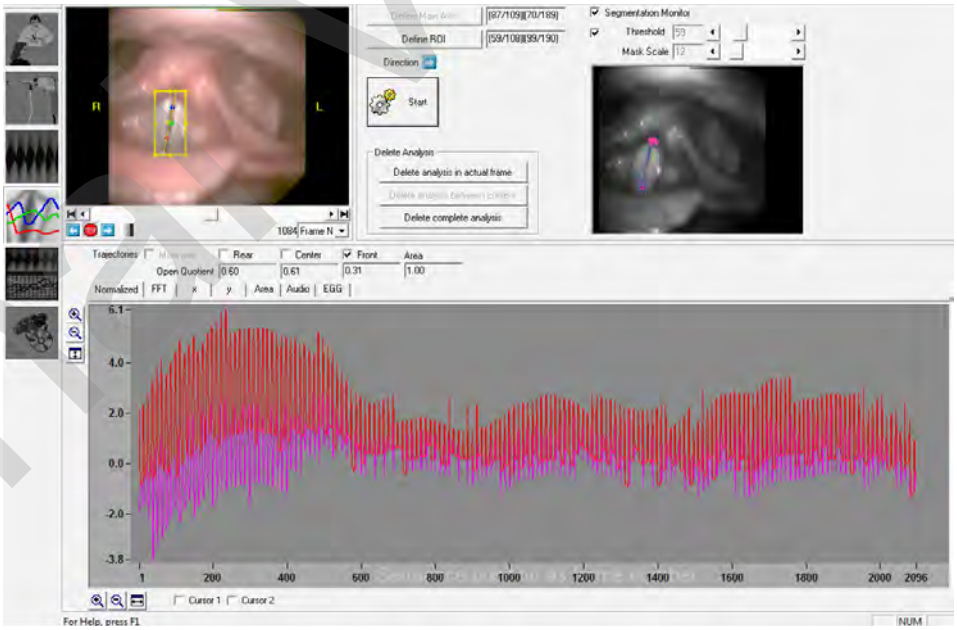

Figure 9. The segmentation is presented of a normal functioning voice with marking of the edges of the VF – female. The system includes a routine possibility for marking of the edges and slow motion reproduction of the original video, which also can be shown in slow motion. Based on the segmentation the opening quotients and the area are calculated (from the stored recordings).

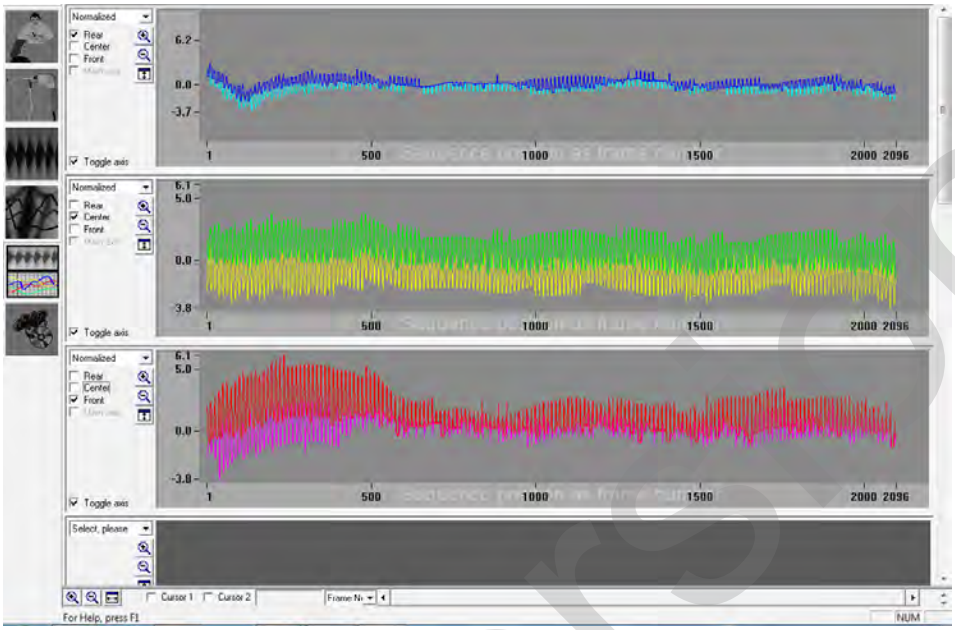

Figure 10. Segmentation of a normal functioning voice – female: right and left VF presented in front, center, and rear parts (of the stored recording). There is a possibility in the software to zoom in and out for greater detail in both axes.

Electroglottography (EGG) and other measures

How EGG corresponds to other measures (here to audio and to kymogram) is shown in Figure 11.

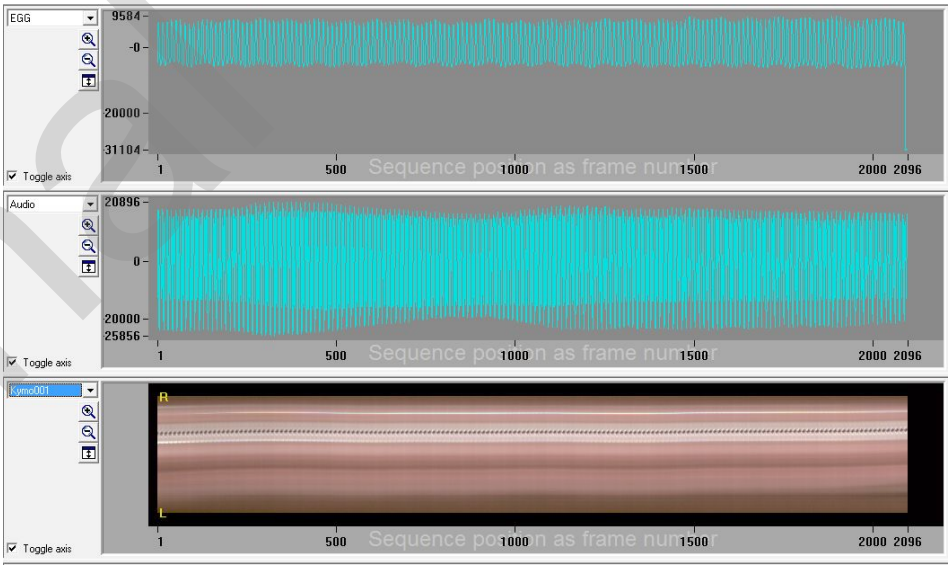

Figure 11. The EGG (upper tracing) shows the single curves of the VF in time at the same time as the acoustical wave (middle tracing) and the kymography picture (lower tracing) presenting the VF as seen from above—female.

Laryngovideotroboscopy (LVS) versus high-speed digital phonoscopy (HSDP)

Because of the relatively slow frame rate (+/- 30 f/s), LVS misses the details of mucosal movement, while HSDP, because of the fast recording speeds, provides the true view of how the glottis vibrates. However, LVS shows the anatomy of the glottis in more details than HSDP. Therefore, HSDP will show the correct movement of right and left VF, including VF irregularity and closure patterns. The possibility to zoom and to observe sequences longer than 2 seconds at a time help a lot, especially when examining the voices of singers [2].

Online curves on HS films

The great value of HS films is the combination of vocal segmentation with kymography, EGG, and with acoustical curves. The more irregular the curves are, the worse the voice is. The EGG curve shows time relations on the VF level. The acoustical curve includes the resonance.

Quantitative measures on HS films

Based on segmentation, with marking of the edge of the VF during intonation, quantitative VF measures can be calculated of the open quotient in the front, center, and rear parts. This is of great diagnostic value. For example, in untreated reflux patients there will be a gap in the posterior part due to the acid and gastric enzyme (pepsin) irritation of the arytenoids. For trained singers with reflux the opening quotients are often reduced in the rear parts of the vocal area, because of pressure exerted on the arytenoids, relating to their unhealthy singing compensating techniques.

The Wolf system was further developed by Döllinger et al. [3] and Lohscheller et al. [4] based on the segmentation analysis of the VF. Presentation of the lateral features of the VF can be calculated. Cycle-based features can also be made (i.e., time periodicity, maximum flow declination rate, speed quotient, speed index, rate quotient, asymmetry, and amplitude quotient). The software development includes measures of stiffness of the VF (Figure 12a and 12b).

Conclusion

This chapter is based on our instruction courses we presented at the International Federation of Oto-rhino-laryngological Societies (IFOS), Sao Paulo, Brazil (2009). We believe that this system improves our examinination and treatment of dysphonias beyond the traditional LVS.

Acknowledgements

The authors thank Dr. M. Döllinger for help in preparation of this chapter. He is developing further measurements for clinical use and we are grateful to be part of his experimental settings. We are also thankful to Dr. Izdebski for organizing this chapter and for all corrections and suggestions.

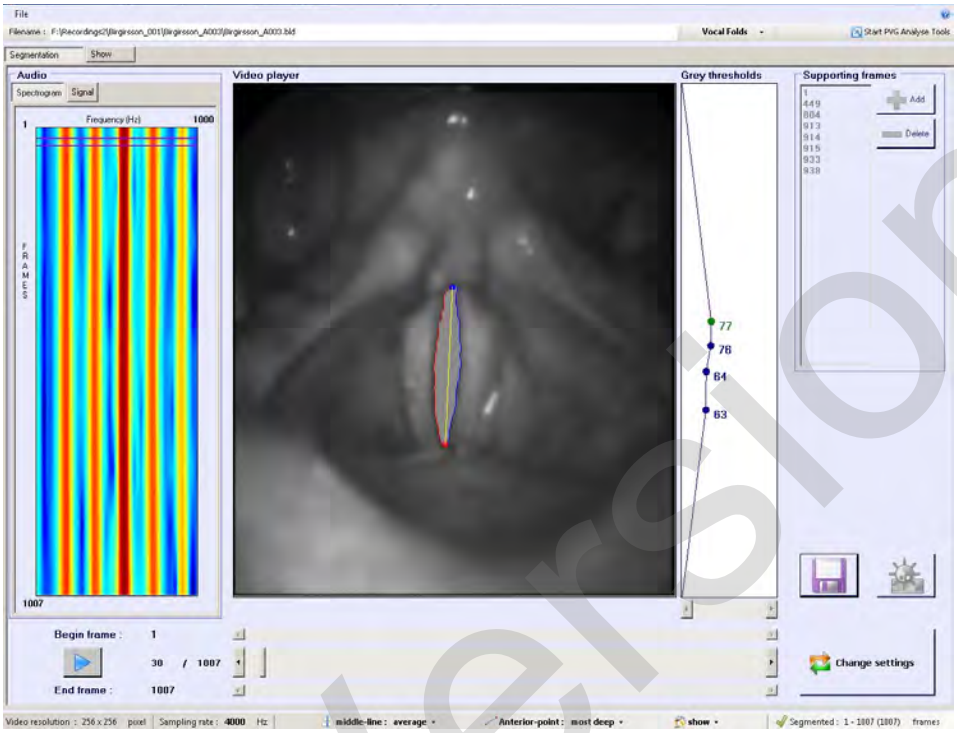

Figure 12a. Shows the imported recording and the segmentation in Glottis Analysis Tools (M. Döllinger). The possibilities are many to ensure an accurate segmentation: varying black/white balance and segmentation area during the length of the film are just some of them.

Shimm(%)	5,048				
HNR(dB)	11,098				
Jitt (%)	0,542				
		Mean	Std	Min	Max
ClosingQuotient(CQ)		0,4149	0,0602	0,2727	0,5455
AsymmetrieQuotient		0,4872	0,0721	0,3333	0,5833
Stiffness	Left	0,4919	0,1902	0,254	0,9897
Stiffness	Right	0,4769	0,1664	0,2625	0,8411

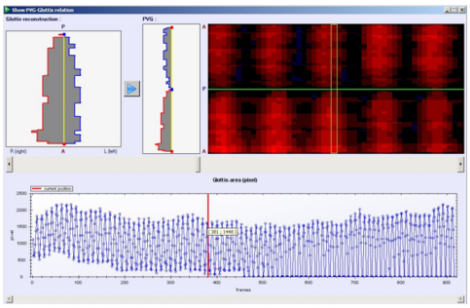

Figure 12b. This figure represents a Glottal Analysis Tool. An increased control with contrast ensures accurate segmentation and the area is calculated for each VF cycle. This also ensures accurate jitter, shimmer, and opening quotients. Here this process is shown for a 13-year-old male soloist.

References

1. Deliyski, D., Hillman, R., 2010. State of the art laryngeal imaging: Research and clinical implications. Curr. Opin. Otolaryngol. Head Neck Surg. 18, 147-152.
2. Pedersen, M., Eeg, M., Jønsson, A., 2013. Video stroboscopy and high-speed films of pathological voices. Presented at 4th Symposium of Advanced Voice Assessment, Copenhagen.

3. Döllinger, M., Hoppe, U., Hettlich, F., Lohscheller, J., Schubert, S., Eysholdt, U., 2002. Vibration parameter extraction from endoscopic image series of the vocal folds. IEEE Trans. Biomed. Eng. 49, 773-781.
4. Lohscheller, J., Eysholdt, U., Toy, H., Döllinger, M., 2008. Phonovibrography: mapping high-speed movies of vocal fold vibrations into 2-d diagrams for visualizing and analyzing the underlying laryngeal dynamics. IEEE Trans. Med. Imaging 27, 300-309.