

초고속 후두내시경 영상을 이용한 평면 스캔 비디오카이모그래피 영상 생성

부산대학교병원 이비인후과, 의생명연구원,¹ 부산대학교 의학전문대학원 이비인후과학교실,²
U-MEDICAL,³ 아주대학교 의학전문대학원 이비인후과학교실⁴
차원재^{1,2} · 왕수건^{2,3} · 장전엽⁴ · 김근효¹ · 이연우¹

Post-Processing of High-Speed Video-Laryngoscopic Images to Two-Dimensional Scanning Digital Kymographic Images

¹Department of Otorhinolaryngology-Head and Neck Surgery and Biomedical Research Institute,
Pusan National University Hospital, Busan; and ²Department of Otorhinolaryngology-Head and Neck Surgery,
Pusan National University School of Medicine, Yangsan; and ³U-Medical Co., Ltd, Busan; and
⁴Department of Otolaryngology, Ajou University School of Medicine, Suwon, Korea
Wonjae Cha^{1,2}, Soo-Geun Wang^{2,3}, Jeon Yeob Jang⁴, Geun-Hyo Kim¹ and Yeon-Woo Lee¹

Background and Objectives : High-speed videolaryngoscopy (HSV) is the only technique that captures the true intra-cycle vibratory behavior of the vocal folds by capturing full images of the vocal folds. However, it has problems of no immediate feedback during examination, considerable waiting time for digital kymography (DKG), recording duration limited to a few seconds, and extreme demands for storage space. Herein, we demonstrate a new post-processing method that converts HSV images to two-dimensional digital kymography (2D-DKG) images, which adopts the algorithm of 2D videokymography (2D VKG). **Materials and Methods** : HSV system was used to capture images of vocal folds. HSV images were post-processed in Kay image-process software (KIPS), and conventional DKG images were retrieved. Custom-made post-processing system was used to convert HSV images to 2D-DKG images. The quantitative parameters of the post-processed 2D-DKG images was validated by comparing these parameters with those of the DKG images. **Results** : Serial HSV images for all phases of vocal fold vibratory movement are included. The images were converted by the scanning method using U-medical image-process software. Similar to conventional DKG, post-processed 2D DKG image from the HSV image can provide quantitative information on vocal fold mucosa vibration, including the various vibratory phases. Differences in amplitude symmetry index, phase symmetry index, open quotient, and close quotient between 2D-DKG and DKG were analyzed. There were no statistical differences between the quantitative parameters of vocal fold vibratory movement in 2D-DKG and DKG. **Conclusion** : The post-processing method of converting HSV images to 2D DKG images could provide clinical information and storage economy.

KEY WORDS : Vocal fold vibration · High-speed image · Two-dimensional scanning digital kymography · Image post-processing · Voice.

Introduction

Examination of the vibratory movement of the vocal fold mucosa is essential to understand the mechanism of voice production and to diagnose various vocal fold disorders. Since the 1960s, videostroboscopy has been the primary

method used to evaluate vocal fold vibration and is the clinical standard for laryngeal imaging. Videostroboscopy is widely used to study the vibration of the vocal folds in clinical practice, and is used most frequently because it provides full-color images with high spatial resolution at a relatively low cost.¹⁾ However, this technique usually shows somewhat illusory slow-motion images of the vibrating vocal folds and

Received: October 14, 2017 / Revised: November 7, 2017 / Accepted: November 23, 2017

Address for correspondence: Soo-Geun Wang, MD, PhD, Department of Otorhinolaryngology-Head and Neck Surgery and Biomedical Research Institute, Pusan National University Hospital, 179 Gudeok-ro, Seo-gu, Busan 49241, Korea
Tel: (051) 240-7336, Fax: (051) 246-8668, E-mail: entwangsg@gmail.com

provides a clear image only when vocal fold vibrations are periodic and have a stable phonation frequency.²⁾ To address such technologic shortcomings, high-speed imaging (HSI) was applied to laryngoscopy to more clearly visualize mucosal wave mechanics.³⁾ By capturing at least 2,000 images per second, HSI can capture at least 10–20 frames per vibratory cycle depending on the fundamental frequency.³⁾ As it captures multiple frames within each cycle, HSI is not dependent on periodic vibratory motion and can describe vibratory behavior beyond the limitations of videostroboscopy.⁴⁾

Ultimately, high-speed videolaryngoscopy (HSV) is the only technique that captures the true intra-cycle vibratory behavior of the vocal folds by capturing a full image of the vocal folds.⁵⁾ HSV overcomes the limitations of videostroboscopy to provide more accurate objective quantification of vocal fold vibratory behavior. It has a broader range of applications and a higher successful interpretation rate compared with videostroboscopy.¹⁾ Another advantage of HSV is that a more advanced analysis is possible. Data from videostroboscopy are usually evaluated subjectively, whereas HSV allows the use of a variety of methods, such as laryngotopography (LTG)⁶⁾ and phonovibrography.^{7,8)} In addition, HSV allows more reliable quantitative analysis, and most of the vibratory parameters routinely evaluated subjectively by videostroboscopy can be quantified by HSV thus allowing objective documentation of the severity of vocal disturbance.¹⁾ Since the HSV system was introduced commercially in the 1990s, HSV has been considered to be the most accurate tool for visualization of vocal fold vibratory movement. However, commercial HSV systems are not in common clinical use owing to unsolved technical, methodological, and practical limitations, and an associated lack of information regarding the validity and clinical relevance of HSV.⁹⁾ Recently, researchers demonstrated that the methods used to analyze HSV data can be clinically useful in documenting the characteristics of vocal fold vibrations in patients with various vocal fold pathologies and in exploring vibratory disturbances to estimate the severity of dysphonia.^{1,9-11)}

There are still problems with the use of HSV in clinical practice, including a lack of immediate feedback during examination, considerable waiting time before kymographic visualization, a recording duration limited to seconds, and extreme demands for storage space.¹²⁾ The recorded HSV data can be analyzed frame by frame to evaluate the entire vocal fold vibration.¹³⁾ However, this requires a high level of concentration from the examiner because of the extremely large storage capacity. It is also difficult to compare images taken at different times on playback in a busy outpatient clinic. Also, the ky-

mographic conversion has the fundamental limitation that only one or several linear portions of the vocal fold mucosa can be visualized. In addition, DKG is extracted from images obtained using laryngeal high-speed imaging playback and shows the real vibratory image of vocal fold mucosa.^{14,15)} Also, methods such as LTG⁶⁾ and glottal area waveform analysis¹⁴⁾ may require complicated computerized manipulation and are difficult to interpret.

Two-dimensional scanning videokymography (2D VKG) was developed to visualize the entire vocal fold vibratory movement.¹⁶⁾ This system can extract dynamic images of the entire vocal fold in real time and analyze the whole vibratory mucosal movement simultaneously. A single still image generated by 2D VKG provides information on the dynamic vibratory motion of the vocal folds.¹⁷⁾ These images can also measure various objective parameters that have previously been reported in videokymography.¹⁷⁾ This imaging technique uses a low frame rate (30 frames/s), and the video clip is small enough to store in a commercial picture archiving and communication system (PACS) and be uploaded without technical problems.

A new post-processing method of HSV images, which can provide compressive and clinical information to laryngologists at a glance, is necessary to overcome the limitation of HSV as a clinical application. The 2D DKG image type, which adopted the algorithm of 2D VKG, can provide clinically relevant information with low storage needs and might be suggested as the imaging protocol for the new post-processing method. In this study, we demonstrate a new post-processing method that can convert HSV images to 2D DKG images.

Materials and Methods

1. The participant

A healthy 34-year-old man with no history of laryngeal disorders or surgery participated in this study. He was a speech-language pathologist and could produce various registers. He phonated sustained /e/ vowels in a modal voice, which were recorded using an HSV imaging system.

2. HSV imaging system

A high-speed color videolaryngoscopy system (model 9710, KayPENTAX, Montvale, NJ) was used to capture images of the subject's vocal folds together with a rigid endoscope (5.8 mm, 70 degrees, 8700CKA, Storz, Germany) and a 300 W xenon light source (NOVA 300, Storz, Germany). This system was used to visualize the vocal folds in their entirety with the resolution of 352×512 pixels and 3000 frames per second. The

HSV images were post-processed in Kay image process software (KIPS), and the conventional DKG images were retrieved.

3. Post-processing system for conversion of HSV images to 2D DKG images

The post-processing system [a custom-made program for post-processing imaging, U-medical image process software (UIPS), version 1.0], and its algorithm and software were established by Dr. Wang. This software runs on the Windows Operating System (Microsoft Corp., Redmond, WA, USA), was developed in Visual C# (Microsoft Corp.) and operates within a single window. The left side of the field contains the function buttons for HSV image uploading and conversion to 2D DKG, the settings bar for glottal area selection, and options for image modulation. The converted video is displayed in the main field, and some optional buttons are listed at the bottom of the field (Fig. 1). To convert an HSV image into a 2D DKG image, the extracted video files are uploaded to the software. The converted images can be exported as a video file (avi type). The conversion algorithm from the scanning method of 2D VKG was adopted.^{16,17)} In the selected vibratory area of vocal folds, a horizontal line or lines of HSV images are serially reconstructed into the new 2D DKG images. The custom-made post-processing system was used to con-

vert HSV images to 2D DKG images.

4. Validation of the quantitative parameters of post-processed 2D DKG images compared to DKG images

Post-processing should not skew the original information of vocal fold mucosa vibration. To validate the quantitative parameters of the post-processed 2D DKG images, we compared these parameters with those of the DKG images. Each ten cycles of vocal fold vibration were analyzed in both 2D DKG and DKG images.

5. Statistical analysis

A Student's t-test was used to evaluate the differences between the two groups ; p values of <0.05 were considered significant. All statistical analyses were performed using R, version 3.4.0 (The R Foundation for Statistical Computing, Vienna, Austria) and RStudio 1.0.143 (RStudio Inc., Boston, MA, USA).

Results

1. Voice characteristics of the participant

After acoustic analysis, the fundamental frequency, jitter, shimmer, and noise-to-harmony ratio of the participant's

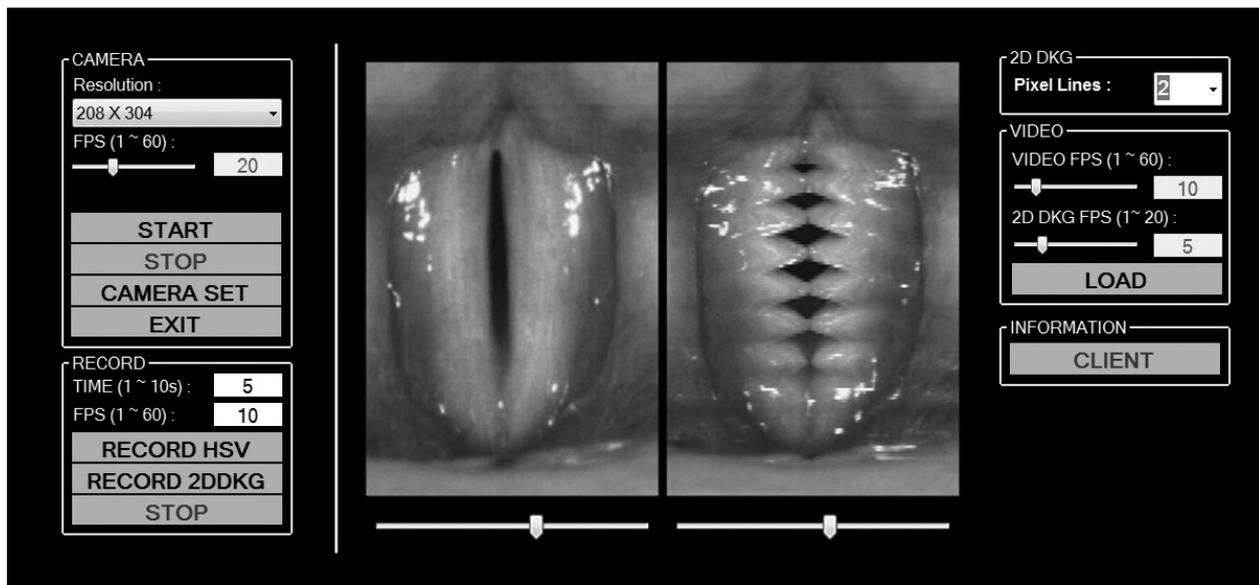


Figure 1. Custom-made post-processing program [U-medical image process software (UIPS), version 1.0]. The software runs on the Windows Operating System (Microsoft Corp., Redmond, WA, USA), was developed in Visual C# (Microsoft Corp.) and operates within a single window. The left side of the field contains the function buttons for HSV image uploading and conversion to 2D DKG, the settings bar for glottal area selection, and options for image modulation. The converted video is displayed in the main field, and some optional buttons are listed at the bottom of the field. To convert an HSV image to a 2D DKG image, the extracted video files are uploaded, the glottal area is manually selected, and the conversion is executed. The converted images can be exported as a video file (avi type). The conversion algorithm was adopted from the line scanning method of the 2D DKG system. In the selected glottal area, a horizontal line or lines of HSV images are serially reconstructed into the new 2D DKG images. The custom-made post-processing system was used to convert HSV images to 2D DKG images.

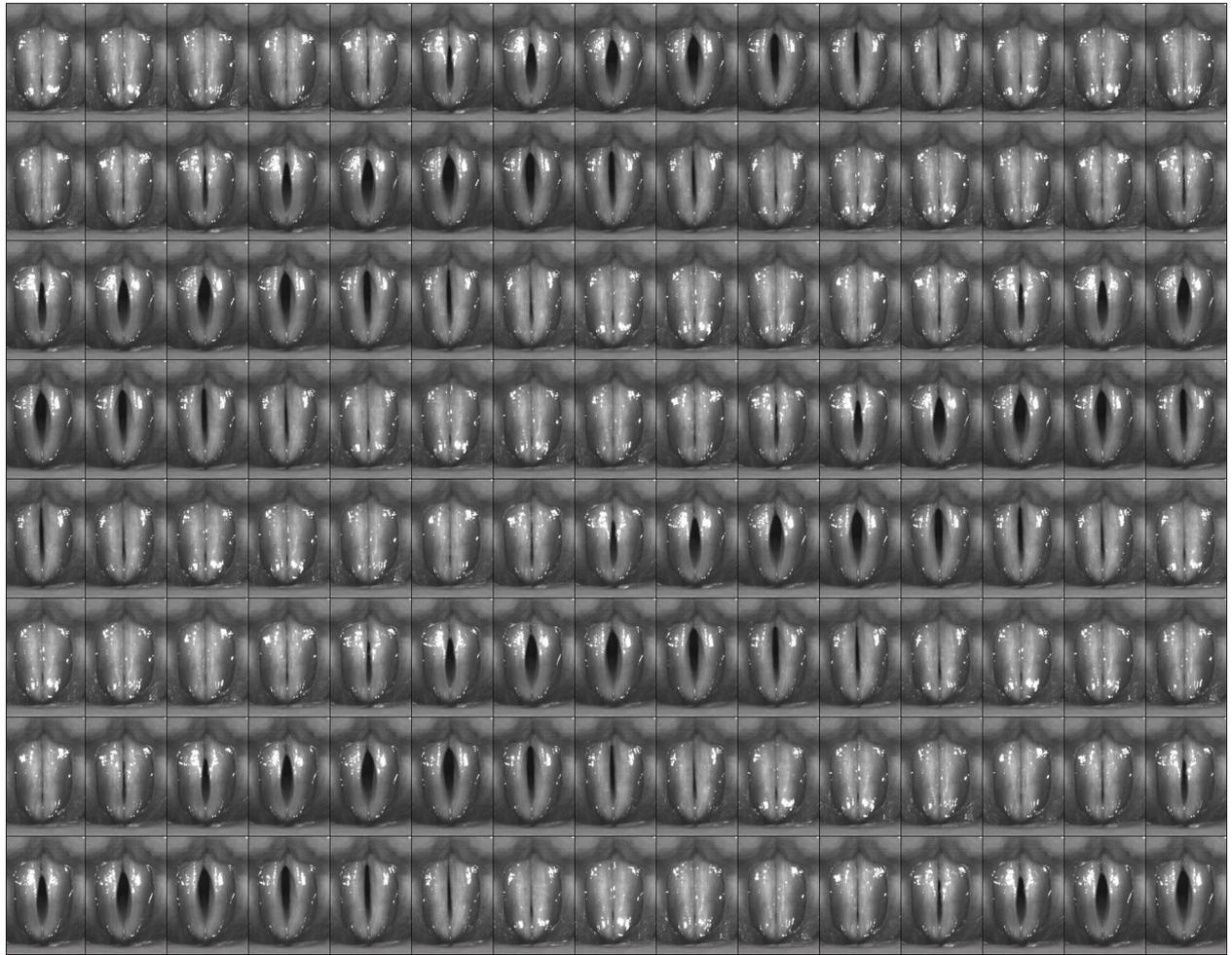


Figure 2. The serial HSV images for all phases of vocal fold vibratory movement.

modal voice were determined to be 123.5 Hz, 1.201, 2.428, and 0.04, respectively.

2. HSV and DKG images

The sequential HSV images for all phases of vocal fold vibratory movement are included in Fig. 2 ; the total size of the HSV video file (4 sec) was 6.3 GB (6291 MB). The converted multi-line DKG images are shown in Fig. 3.

3. Conversion of captured HSV images to 2D DKG images

The HSV video file was uploaded using UIPS and converted to 2D DKG images by the scanning method using 1 to 4-pixel lines. The total size of a post-processed 2D DKG video with four types of pixel line scanning (1 to 4 lines) was approximately 6.2 Megabits. The post-processed 2D DKG images with various line scanning methods are shown in Fig. 4. Similar to 2D VKG, the post-processed 2D DKG image from the HSV image can provide quantitative information on vocal fold mucosal vibration. This includes the vibratory phases of the open phase, opening phase, closing phase, and closed

phase, the kymographic markers of the lateral and medial peaks, and the upper and lower lips of the vocal fold mucosa.

4. Validation of the quantitative parameters of post-processed 2D DKG images compared with conventional DKG images

Differences in the amplitude symmetry index (ASI), phase symmetry index (PSI), open quotient (OQ), and close quotient (CQ) between 2D DKG and DKG images were analyzed in each of the ten cycles. There were no statistical differences between the quantitative parameters of vocal fold vibratory movement of 2D DKG and DKG images (Table 1).

Discussion

The clinical utility of videostroboscopy was established as an evaluation method for vocal fold mucosal vibration.¹⁸⁾ Currently, videostroboscopy is the recommended examination modality for specialized evaluation of dysphonic patients.³⁾

Videostroboscopy is widely used to show “illusory” slow-motion images of the vibrating vocal folds. However, a clear image can usually only be obtained when vocal fold vibrations are periodic and have a stable phonation frequency. Gall et al. first devised strip kymography as an alternative tool in 1984 which enables images of vibrating vocal folds to be acquired while the slit shutter is fixed to one point of the vocal fold and the film is moved rapidly.¹⁹⁾ Since Gall’s prominent work, there have been notable advances in imaging technology. Currently, there are three kymographic techniques to examine vocal fold vibration; VKG, DKG, and strobovideokymography.¹²⁾

HSV has been considered the most accurate tool to visualize vocal fold vibratory movement since its invention. The HSV system can record just a few seconds of movement, and it requires an enormous storage capacity. In this study, an HSV file approximately 4 seconds in length was obtained, and its file size was over 6 Gigabits. In research, the accuracy and quality of the data or image are the most important aspects. However, in clinical practice, physicians and speech language pathologists must consider factors related to time, resources, manpower, and physical data storage of their study tools. Due to its high storage and analysis time requirements, it is difficult to use the HSV system in daily clinical practice.

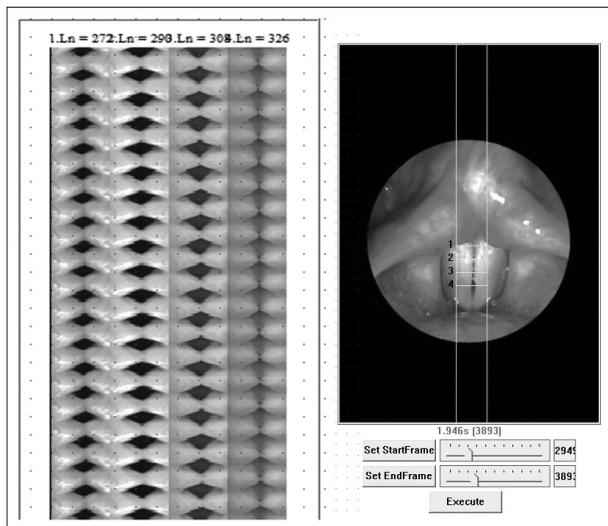


Figure 3. The converted multi-line DKG images.

Thus, most institutes have used DKG images to interpret the HSV images for clinical or research purposes. However, because DKG is fundamentally a one-dimensional analysis of single or multiple lines, the entire vocal fold configuration cannot be visualized, and continuous strip images are not easy to interpret in busy clinics.

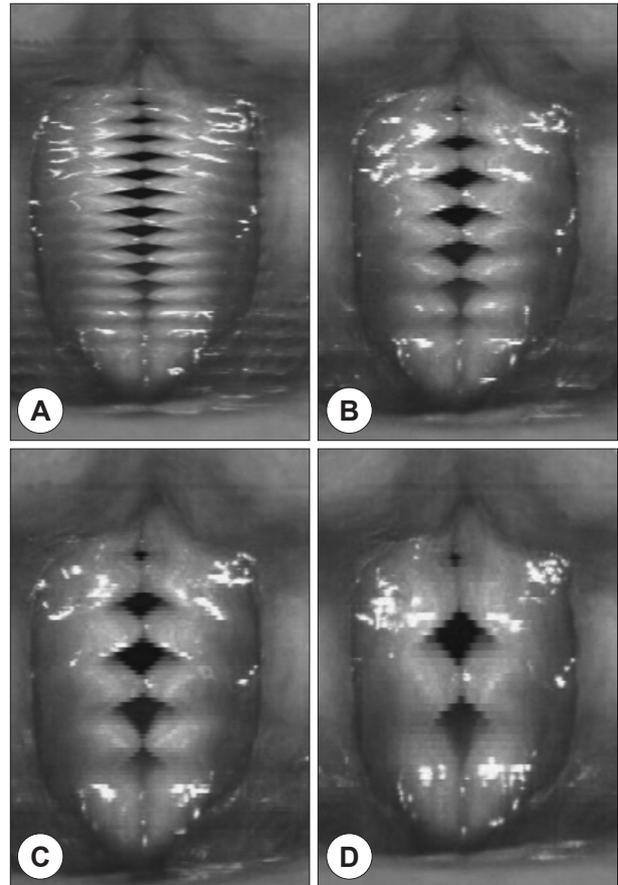


Figure 4. Conversion of captured HSV images to 2D DKG images. The HSV video file was uploaded to the software and converted to 2D DKG images. The images were converted by the scanning method using 1 to 4-pixel lines. Similar to 2D DKG, the post-processed 2D DKG image from the HSV image can provide quantitative information about vocal fold mucosa vibration, including the vibratory phases of the open phase, opening phase, closing phase, and closed phase, the kymographic markers of the lateral and medial peaks, and the upper and lower lips of the vocal fold mucosa. A : 1-pixel line method. B : 2-pixel line method. C : 3-pixel line method. D : 4-pixel line method.

Table 1. Quantitative parameters of post-processed 2D DKG images compared with DKG images of ten cycles

Parameters	Post-processed 2D DKG (n=10)	Conventional DKG (n=10)	p
ASI	0.04±0.02 (0.02–0.09)	0.04±0.01 (0.03–0.05)	0.9371
PSI	0.07±0.01 (0.06–0.08)	0.07±0.00 (0.06–0.07)	0.1442
OQ	0.58±0.04 (0.53–0.64)	0.56±0.05 (0.53–0.64)	0.1788
CQ	0.42±0.04 (0.36–0.47)	0.44±0.05 (0.36–0.50)	0.1778

2D DKG : two-dimensional scanning digital kymography, DKG : digital kymography, ASI : amplitude symmetry index, PSI : phase symmetry index, OQ : open quotient, CQ : close quotient

Recently, 2D DKG has been developed to visualize the vibratory movement of the vocal fold mucosa.^{16,17)} Photokymography allows the entire area of the vocal fold to be recorded as the slit shutter moves in the inferior-to-superior direction in front of a fixed film.²⁰⁾ In 1984, strip kymography was devised, that enables images of vibrating vocal folds to be acquired by fixing the slit shutter on one point on the vocal folds and moving the film rapidly.^{16,19)} The principle of Gall's photokymography was applied to 2D DKG.^{16,20)} The results of 2D DKG are multiphasic functional images of the vibrating mucosa, and several key images can be enough to analyze the status of vocal fold mucosa. Previous studies have suggested that 2D DKG images can provide vocal fold vibratory patterns, and various parameters such as the fundamental frequency, OQ, SQ, PSI, ASI, and glottal area index.¹⁷⁾ The 2D DKG image is not a real image but can be considered a functionally compressed image type.

After the development of the prototype 2D VKG system, we initially applied the system to an ex vivo canine larynx model.¹⁶⁾ In the ex vivo study, we compared 2D VKG with videostroboscopy. The new system has several advantages. It can extract dynamic images of the entire vocal fold in real time and analyze the entire vibratory movement simultaneously. It can be used in real time without time-consuming image processing, which should be useful in busy clinical settings. Also, as 2D VKG records the images at a speed of 30 frames/s, much less data storage is required.

In this study, we attempted to convert the HSV image file to a post-processed 2D DKG image. We found that converted 2D DKG images have several advantages compared to DKG images. As mentioned, 2D DKG shows the two-dimensional information of the entire vocal fold, which can easily provide clinical information of laryngeal diseases to physicians and speech language pathologists. Its storage capacity requirements are small, and a single 2D DKG image can provide sufficient quantitative information. Images may also be uploaded to medical PACS without additional upgrades to either storage or the computer system. To validate whether post-processed 2D DKG could provide optimal information for vibratory movement from HSV images, its quantitative parameters were compared with those of conventional DKG. As there was no difference between the two image types, post-processing to 2D DKG was considered to be a suitable conversion method. Also, this system can convert any type of HSV image because our protocol does not implement whole image processing and instead uses the line scanning method.

However, post-processing 2D DKG also has certain disadvantages.^{16,17)} The motion of the vocal fold at a specific, pre-

cise location cannot be evaluated in contrast with VKG or digital kymography, and anterior-posterior phase differences cannot be evaluated in contrast with multi-line digital kymography or phonovibrography. A combination of spatial and temporal information could make it difficult to determine whether the irregularities shown in the 2D DKG images are due to a spatial abnormality on the vocal folds or to the irregular nature of the vibrations. The static images in routine videolaryngoscopy can provide good information about the mucosal lesion or status.

Although post-processed 2D DKG is not perfect compared to HSV in evaluating the vibratory movement of vocal fold mucosa, it has advantages in data storage efficiency and in providing clinically relevant information. Due to its clinical advantages, we expect that physicians and speech language pathologists will use 2D DKG images in clinical practice.

Conclusion

The post-processing method of converting HSV images to 2D DKG images could provide useful clinical information and storage economy.

This work was supported by a clinical research grant in 2016 from Pusan National University Hospital.

REFERENCES

- 1) Yamauchi A, Yokonishi H, Imagawa H, Sakakibara K, Nito T, Tayaama N, et al. *Quantification of Vocal Fold Vibration in Various Laryngeal Disorders Using High-Speed Digital Imaging*. *J Voice* 2016;30:205-14.
- 2) Kendall KA, Browning MM, Skovlund SM. *Introduction to high-speed imaging of the larynx*. *Curr Opin Otolaryngol Head Neck Surg* 2005;13:135-7.
- 3) Mendelsohn AH, Remacle M, Courey MS, Gerhard F, Postma GN. *The diagnostic role of high-speed vocal fold vibratory imaging*. *J Voice* 2013;27:627-31.
- 4) Jiang JJ, Yumoto E, Lin SJ, Kadota Y, Kurokawa H, Hanson DG. *Quantitative measurement of mucosal wave by high-speed photography in excised larynges*. *Ann Otol Rhinol Laryngol* 1998;107:98-103.
- 5) Deliyski DD, Petrushev PP, Bonilha HS, Gerlach TT, Martin-Harris B, Hillman RE. *Clinical implementation of laryngeal high-speed videoendoscopy: challenges and evolution*. *Folia Phoniatr Logop* 2008;60:33-44.
- 6) Yamauchi A, Imagawa H, Sakakibara K, Yokonishi H, Nito T, Yamasoba T, et al. *Phase difference of vocally healthy subjects in high-speed digital imaging analyzed with laryngotopography*. *J Voice* 2013;27:39-45.
- 7) Lohscheller J, Eysholdt U, Toy H, Dollinger M. *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* 2008;27:300-9.
- 8) Kunduk M, Doellinger M, McWhorter AJ, Lohscheller J. *Assessment of the variability of vocal fold dynamics within and between recordings with high-speed imaging and by phonovibrograph*. *La-*

- ryngoscope* 2010;120:981-7.
- 9) Yamauchi A, Yokonishi H, Imagawa H, Sakakibara KI, Nito T, Tayaama N, et al. *Characterization of Vocal Fold Vibration in Sulcus Vocalis Using High-Speed Digital Imaging. J Speech Lang Hear Res* 2017;60:24-37.
 - 10) Yamauchi A, Yokonishi H, Imagawa H, Sakakibara K, Nito T, Tayaama N, et al. *Visualization and Estimation of Vibratory Disturbance in Vocal Fold Scar Using High-Speed Digital Imaging. J Voice* 2016;30:493-500.
 - 11) Yamauchi A, Yokonishi H, Imagawa H, Sakakibara KI, Nito T, Tayaama N. *Quantitative Analysis of Vocal Fold Vibration in Vocal Fold Paralysis With the Use of High-speed Digital Imaging. J Voice* 2016;30:766 e13-e22.
 - 12) Svec JG, Schutte HK. *Kymographic imaging of laryngeal vibrations. Curr Opin Otolaryngol Head Neck Surg* 2012;20:458-65.
 - 13) Yamauchi A, Imagawa H, Yokonishi H, Nito T, Yamasoba T, Goto T, et al. *Evaluation of vocal fold vibration with an assessment form for high-speed digital imaging: comparative study between healthy young and elderly subjects. J Voice* 2012;26:742-50.
 - 14) Wittenberg T, Tigges M, Mergell P, Eysholdt U. *Functional imaging of vocal fold vibration: digital multislice high-speed kymography. J Voice* 2000;14:422-42.
 - 15) Larsson H, Hertegard S, Lindestad PA, Hammarberg B. *Vocal fold vibrations: high-speed imaging, kymography, and acoustic analysis: a preliminary report. Laryngoscope* 2000;110:2117-22.
 - 16) Wang SG, Park HJ, Cho JK, Jang JY, Lee WY, Lee BJ, et al. *The First Application of the Two-Dimensional Scanning Videokymography in Excised Canine Larynx Model. J Voice* 2016;30:1-4.
 - 17) Park HJ, Cha W, Kim GH, Jeon GR, Lee BJ, Shin BJ, et al. *Imaging and Analysis of Human Vocal Fold Vibration Using Two-Dimensional (2D) Scanning Videokymography. J Voice* 2016;30:345-53.
 - 18) Sataloff RT, Spiegel JR, Hawkshaw MJ. *Stroboscopy: results and clinical value. Ann Otol Rhinol Laryngol* 1991;100:725-7.
 - 19) Gall V. *Strip kymography of the glottis. Arch Otorhinolaryngol* 1984;240:287-93.
 - 20) Gall V, Gall D, Hanson J. *[Laryngeal photokymography]. Arch Klin Exp Ohren Nasen Kehlkopfheilkd* 1971;200:34-41.