Chapter 48
3D video reconstructions of normal and abnormal human vocal folds

Jarosław Sova, Joanna Cieszyńska, Jarosław Kijewski & Paweł Jesikiewicz

Abstract

The aim of this study is to present a 3D digital reconstruction of normal and pathological vocal folds (VF) at rest and during phonation as well as discuss research and clinical applications of this emerging technology. Video recordings of the human glottis were recorded using a newly developed endoscope (ENT Viewer®, Sinutronic, Gdansk, Poland). This system is capable of obtaining images with stroboscopic (LVS) or without stroboscopic (LV) illumination. The LVS and LV images were processed digitally and were processed into a 3D video reconstruction. The 3D video reconstructions of VF kinematics are presented. The images obtained during LV and LVS are compared in a 1:1 scale to the images obtained from such imaging techniques as MRI, CT scans, and 3D prints; dimensions and distance scaling are estimated digitally.

Keywords: 3D video reconstruction, VF 3D imaging, LVS, LV, VF kinematics, distance scaling

Introduction

Viewing an object in the third dimension (3D) is desirable as it enhances image perception, especially its depth. Therefore, 3D reconstructions are now being frequently implemented in medicine from CT and MRI scans [1-2] or even from ultrasound imaging [3]. 3D imaging can be accomplished by post-production reconstruction technology.

When applying this 3D reconstructive technology to study the VF in action, new observations about VF kinematics are emerging. These observations are not visible in stroboscopic and or even in High Speed Digital Phonoscopy (HSDP) illumination and capture.

Method

The 3D reconstruction technology captures the shape and appearance of real objects either by active or passive methods. In our case, we used an active spatio-temporal reconstruction method based on color differences to construct 3D images of the VF. Spatio-temporal, or so called non-rigid reconstruction, is used when a change of image shape is seen in time.

To get the video material subjected to 3D processing treatments, a newly developed endoscope ENT Viewer® (Sinutronic Sp z oo, Gdansk, Poland) was used. This ENT Viewer® is an electronic and optical device (Figure 1). The ENT Viewer® comprises a spatula with a video camera and a base used for signal reception. A compact camera (1/8 inch matrix) is embedded at the distal end of the spatula and registers the images with a resolution of 800 x 600 pixels. The video image from this unit is transmitted to the base wirelessly. The spatula also contains a broadband analogue transmitter.
The examined object is lit by two white LED diodes with the optimally matched color temperature and light intensity. Additionally, the light may be modulated by sampling the patient’s voice fundamental frequency (F0), which results in a strobing effect, allowing for a LVS examination. The response frequency is from 90-1000 Hz. The software EndoMaster, which comes with the ENT Viewer®, enables the user to preview, record, edit, and archive the examination on the screen.

Open source program, which makes use of photometric method, is used in creating the 3D reconstruction. The photometric method consists of using the digital assessment of surface area with diverse illumination intensity in controlled light conditions (Figure 2). As a result, the content of a particular brightness area is subjected to calculations and this allows us to obtain the information about the heights of the points on the reconstructed area. By using brightness differences the system can create a 3D simulation of a block of data we are interested in.

**Figure 1.** Overall view of the Sinutronic ENT Viewer® device.

**Figure 2.** The reconstruction of 3D image from planar image. The computer program creates a 3D image according to the rule: the further from the source of light the object is, the darker it is; and the closer it is, the lighter it is.
Challenges in the 3D reconstruction method used

Glares

Glares present in the image (mostly from secretions) cause significant 3D distortions. When the reflected reflexes are bright, this makes a computer program to interpret these images as a big size area function. To further improve the process of mapping the images with a great number of glares, an original program was created. This program reduces the glares that appear in the examination. This negative effect is partly eliminated by implementing filters that calculate median brightness. The method used in the program calculates median brightness for each source image frame and the size of the frame is defined. If the brightness of a pixel in a source image is way bigger than in a median image, it is replaced with a pixel from a median image. This process is demonstrated in Figure 3. However, this process degrades the quality of the obtained video material. To prevent such distortions, we use two independent diodes to spotlight the image obtained by means of ENT Viewer®, hence we eliminate many unwanted reflections by active filters and improve the quality of the final 3D reconstruction.

Figure 3. A) Image with glares; B) image after software glare filtration; C) 3D reconstruction with glares; and D) 3D reconstruction after software filtration eliminated glares.

Distance scaling

Distance scaling or size measurements are big challenges in derivation from optical images. We believe that VF 3D visualization will add significantly to distance scaling possibilities through the creation of custom programs. This process still requires much researching, so in the study described here we only report on our preliminary methodology and results.
Results

Figure 4 shows a 3D reconstruction of normal VF in a lateral and an anterior view. A separate challenge is to evaluate the distance and the dimensions of obtained images. We are trying to obtain these measures by comparing the images obtained in 3D reconstruction to the images obtained from graphic examinations (e.g., CT, MRI). As the larynx moves during sound production, it is essential to find the constant-size elements, regardless of the position of larynx. Only in such a situation can the comparison of the distance be made securely. Examples of such constant places in the larynx include the distance between the corniculate cartilages, over the muscular process of arytenoid cartilage, or laryngeal lesions visible in radiologic or in LV examinations as demonstrated in Figure 5.

![Figure 4. 3D reconstruction of normal VF in lateral (A) and anterior (B) views.](image)

![Figure 5. Right VF carcinoma. Example of 3D reconstruction (A) compared to the CT scans (B,C) and printing reconstruction (D). Color arrows show the same distances in constant-size elements.](image)
We believe that the possibility to measure the distance in the examination of larynx can be very useful. It is significant during the evaluation of tumorous or other lesions. Aside from scientific research, it can be meaningful in the assessment of the size of VF, for example when patients are being treated with hormones, or with ones that have hormone levels changed [3]. Thereby, it would be possible, for example, to monitor the progress in rehabilitation of functional dysphonia as well as the effects of hyaluonic acid treatment [4]. It might also be an element of an assessment in disorders affecting the professional voice as well as in difficult jurisdiction cases.

Conclusions

The use of a novel approach to create 3D reconstruction of the VF is described here. By using the 3D reconstruction program to process moving images obtained during glottis imaging (LV or LVS), we found a completely new way of interpreting VF kinematics. The observation of the VF at rest and during vibration gives us valid and novel information. This gives a new opportunity to look at the glottis from a different perspective.

Due to the fact that this method requires considerable research as well as further programming works, we present this material as an initial report. Designing a program that will create a possibility of making precise measurements within glottis will result in creating a tool useful in many clinical and jurisdiction cases.

References

2. Grodd, W., Dannenmaier, B., Petersen, D., Gehrke, G., 1987. 3-dimensional (3-D) image reconstruction of the facial skull and skull base in computerized tomography. Radiol. 27, 502-510.