Fitting synthetic to clinical kymographic images for deriving kinematic vocal fold parameters: Application to left-right vibratory phase differences

Sridhar Bulusu[1], S. Pravin Kumar[2], Jan G. Svec[3] , Philipp Aichinger[1]

[1]Division of Phoniatrics-Logopedics, Department of Otorhinolaryngology, Medical University of Vienna, Austria
[2]Department of Biomedical Engineering, Sri Sivasubramniya Nadar (SSN) College of Engineering, Chennai, India
[3]Voice Research Laboratory, Department of Biophysics, Faculty of Science, Palacky University Olomouc, Czechia

1. Background and Objectives

- Vocal fold vibrations are recorded through endoscopy using a videokymographic camera (Fig. 1)
- Videokymographic camera produces high-speed kymograms (Fig. 2, right) from a single line of the vocal fold image (Fig. 2, left) [1,2]
- Kinematic mucosal wave model simulates vocal fold vibration and produces synthetic kymograms (Fig. 3) [3]
- Objectives :
  - Fitting a clinical corpus of kymograms
  - Generation and fitting of synthetic corpus of kymograms
  - Visual assessment of phase differences by three observers
  - Comparison of visual assessment with automated calculation of Phase Difference
  - Employment of ROC-analysis (Receiver Operating Characteristic)
  - Calculation of optimal thresholds for prediction of phase differences

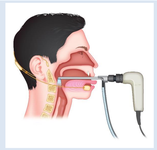

Fig. 1: Endoscopy

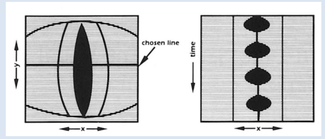

Fig. 2: Laryngoscopic view of the vocal folds (left) with a selected line for obtaining the kymogram (right) [1]

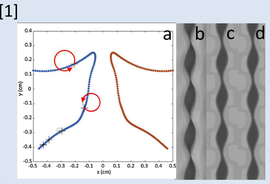

Fig. 3: (a) Geometry of model for synthesizing kymograms [3]. (b) Clinical kymogram exhibiting moderate phase difference. (c,d) Synthetic kymograms obtained by fitting the model with CUC (c) and DSSIM (d) error measure.

2. Model, Parameter Extraction and Corpora

2.1 Kinematic Model

Simulation of wavelike and circular motion of points of vocal fold contour (Fig. 3) [4]. Equations governing the motion of the vocal fold contours in the model are:

$$x_n^j = x_{on}^j + A_n^j \cdot \sin(2\pi f_0 X + \phi_j)\; y_n^j = y_{on}^j + A_n^j \cdot \cos(2\pi f_0 X + \phi_j)$$

2.2 Error Measures

Employment of two error measures to quantify agreement between a clincial or synthetic Kymogram and its fit: Structural Dissimilarity Index Measure (DSSIM) based on Structural Similarity Index Measure (SSIM), and Cross Uncorrelation (CUC) based on Cross Correlation (CC). Errors vary between 0 and 1.

$$\mathrm{SSIM(Y,Z)} = \frac{(2\mu_Y\mu_Z+C_1)(2\sigma_{YZ}+C_2)}{(\mu_Y^2+\mu_Z^2+C_1)(\sigma_Y^2+\sigma_Z^2+C_2)},\ \mathrm{DSSIM(Y,Z)} = \frac{1-\mathrm{SSIM(Y,Z)}}{2},\ C_1 = (0.01\cdot \mathrm{L})^2,\ C_2 = (0.03\cdot \mathrm{L})^2$$

$$\mathrm{CC(Y,Z)} = \frac{\sum_{ab}(Y_{ab}-\mu_Y)(Z_{ab}-\mu_Z)}{\sqrt{\sum_{ab}(Y_{ab}-\mu_Y)^2\sum_{ab}(Z_{ab}-\mu_Z)^2}},\ \mathrm{CUC(Y,Z)} = \frac{1-\mathrm{CC(Y,Z)}}{2}$$

2.3 Corpora

Clinical corpus of 55 kymograms fitted with respect to error measures in order to yield distribution means and covariance. These values are used to generate a synthetic corpus using a multivariate truncated normal distribution:

$$N(\vec{x};\vec{\mu},\Sigma) = \frac{1}{2\pi^{D/2}}\frac{1}{|\Sigma|^{1/2}}exp\left(-\frac{1}{2}(\vec{x}-\vec{\mu})^T\Sigma^{-1}(\vec{x}-\vec{\mu})\right)$$

3. Rating, Regression and ROC

2.1 Rating of Kymograms for Phase Differences

Four rating labels: '0' (negligible phase difference), '1' (low phase difference), '2' (moderate phase difference) and '3' (strong phase difference)

2.1 Regression

Modified Weber Fechner's law is used to carry out regression of relationship between perceived, y, and objective stimuli, x.

$$x(y) = \theta_3 + \theta_1\left(\exp\left(\frac{y}{\theta_2}\right) - 1\right).$$

2.2 Receiver Operating Characteristic

Three thresholds are used to distinguish between labels. Number of True Positives (TP), False Positives (FP), True Negatives (TN) or False Negatives (FN) are calculated based on candidate threshold. "Positive" and "Negative" refer to whether the kymogram is above or below the threshold of the corresponding rating, respectively. "True" and "False" reflect whether the kymogram's threshold-based rating agrees with the subjective rating.

4. Regression and ROC Curves

Coefficients for the regression for all corpora (Fig. 4) are θ_1 = 0.77, θ_2 = 3.36 and θ_3 = 0.19. The curvature is small indicating that the relationship is almost linear. Differential sensitivity slightly decreases with stimulus size in accordance with Weber Fechner's Law.

The ROC curve for the clinical corpus when using the CUC as well as DSSIM as error measures (Fig. 5 and Fig. 6) indicates that the performance parameters Sensitivity, Specificity, Accuracy and AUC are better for larger phase differences.

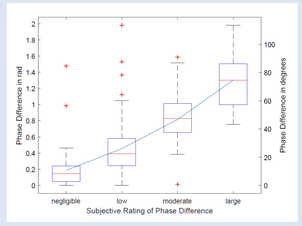

Fig. 4: Boxplots of the phase differences (in fractions of π) obtained from fitting for all error measures vs. the subjective phase difference rating for all the corpora.

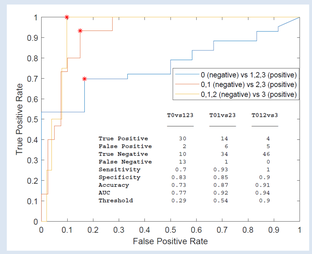

Fig. 5: ROC curves and rating performance obtained with the clinical measurement using the CUC

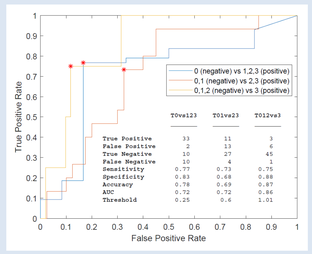

Fig. 6: ROC curves and rating performance obtained with the clinical measurement using the DSSIM

5. Thresholds

Fig. 7 and Fig. 8 show boxplots of objectively estimated phase differences with regard to the subjective ratings of clinical corpus for CUC and DSSIM respectively.
Thresholds vary nonlinearly: For CUC the difference of first to second threshold is 0.25 and 0.36 for second to third threshold. For DSSIM the differences are 0.35 and 0.41

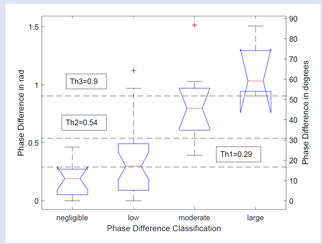

Fig. 7: Boxplots of the phase difference for the CUC error and clinical corpus

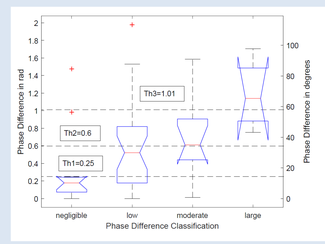

Fig. 8: Boxplots of the phase difference for the DSSIM error and clinical corpus

6. Conclusion

The regression shows that the relation between the perceived phase difference and actual phase difference can be modelled by a modified version of Weber-Fechner's law, although the curvature is very small. The performance parameters of the ROC analysis are better for larger phase differences. The thresholds vary non-linearly, which can be traced back to Weber-Fechner's law.

References and Acknowledgments

[1] Svec JG and Schutte HK. Videokymography: high-speed line scanning of vocal fold vibration. J Voice 10: 201-205, 1996.
[2] Qiu Q and Schutte HK. Real-time kymographic imaging for visualizing human vocal-fold vibratory function. Rev Sci Instrum 78: Art. No. 024302, 2007.
[3] Kumar SP and Svec JG , "Kinematic model for simulating mucosal wave phenomena on vocal folds," Biomed. Signal Proces., vol. 49, pp. 328–337, 2019.
[4] Scherer R, Shinwari D, de Witt K, Zhang C, Kucinschi B, and Afjeh A, "Intraglottal Pressure Profiles for a Symmetric and Oblique Glottis with a Divergence Angle of 10 Degrees," The Journal of the Acoustical Society of America, vol. 109, no. 4, pp. 1616–1630, 2001, [Online]. Available: http://search.proquest.com/docview/85574322/.
[5] Bulusu S, Kumar SP, Švec JG, Aichinger P. Fitting synthetic to clinical kymographic images for deriving kinematic vocal fold parameters: Application to left-right vibratory phase differences. Biomedical Signal Processing and Control 63: 02253, 2021.

Thanks to Dr. Jitka Vydrova from the Voice and Hearing Centre Prague, Medical Healthcom, Ltd., for providing clinical kymograms

This work was supported by the Austrian Science Fund (FWF): KLI 722-B30 and the Czech Science Foundation (GA CR) project no. 19-04477S.

FWF Der Wissenschaftsfonds.