AUTOMATIC CLASSIFICATION OF VOLUMES OF WATER USING SWALLOW SOUNDS FROM CERVICAL AUSCULTATION

Siddharth Subramani[*], Achuth Rao M V[*], Divya Giridhar[*], Prasanna Suresh Hegde[†], Prasanta Kumar Ghosh[*]

[*]SPIRE LAB, Electrical Engineering, Indian Institute of Science (IISc), Bangalore, India
[†]Dept. of Head and Neck Surgery, Health Care Global Enterprises Ltd., Bangalore, India

Poster Session: WE3.PG.12: Biomedical Signal Processing
Wednesday, 6 May, 2020, 16:30-18:30

Flow of presentation

- Introduction
- Data
- Proposed Study
- Results
- Conclusion

Plan

Swallow physiology and Dysphagia

- Swallow - Movement of food bolus from mouth through pharynx to the esophagus
- Process in sync with respiration
- Swallowing disorders (Dysphagia) can occur due to
  - Neurological disorders - Parkinson's disease, stroke
  - Irregularities in esophageal and pharyngeal muscles
  - Amyotrophic lateral sclerosis
  - Head and Neck cancer, etc.
- Common clinical assessments - Videofluoroscopy, Fiber-optic endoscopy - **invasive and expensive methods**
- Cervical Auscultation[1] (CA) characterizes swallow in terms of signal - **non-invasive**

---

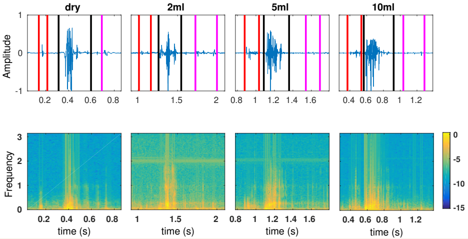
[1]Geovana de Paula Bolzan, Mara Keli Christmann, Luana Cristina Berwig, Cintia Conceição Costa, and Renata Man-copes Rocha, "Contribution of the cervical auscultation in clinical assessment of the oropharyngeal dysphagia," *Revista CEFAC*, vol. 15, no. 2, pp. 455–465, 2013.

Applications of Swallow Sound analysis

- Wearable MIB (Monitoring of Ingestive Behaviour) devices for food intake characterization
- Identification of dysphagia through Spontaneous Swallowing frequency analysis
- Volume specific models can help in studying the severity of dysphagia

Swallow signal characterization

- Swallow signals are characterized by their **S**wallow **C**omponents[2]



Red, black and magenta lines represent SC1, SC2 and SC3 regions, respectively

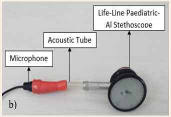
- **Why volume-specific features?**: Signatures of **SC**s are found to vary with bolus volume
- **Objective**: Learn features for bolus volume characterization through classification of swallows of water against dry swallows

[2] Divya Giridhar, Achuth Rao, Prasanna Hegde, and Prasanta Ghosh, "Analysis of swallow sounds of healthy controls for different volumes of water," in Int. Conf. on Engg. in Med. and Life Sci., PSG College of Technology, Coimbatore, India, 2019.

Plan

Swallow Sound Recording

- 56 subjects: 34 male, 23 female (20 - 30 years of age)
- Subjects made to swallow water of volumes 2ml, 5ml and 10ml, and also perform dry swallow - each 6 times per subject
- CA setup
  - Life-Line Paediatric-Al Stethoscope
  - Acoustic tube
  - Sorella'z portable 3.5mm microphone (frequency range of 30Hz - 15000Hz)
- Device placement site[3]: Posterior inferior to the cricoid cartilage encircling the trachea



[3]Q Pan, Naoto Maeda, Yousuke Manda, Naoki Kodama, and Shougo Minagi, "Validation of the optimal site in the neck region for detecting swallowing sounds, "Journal of oral rehabilitation, vol. 43, no. 11, pp. 840–846, 2016.

Plan

Feature sets

- **Baseline Feature set (BLF)**: 12 features - duration of SSW1, SSW2 ,SC1, SC2, SC3, intervals between swallow components (I1, I2), duration to peak intensity (DPI), peak intensities (PI_SSW1, PI_SSW2, PI) and total duration (TD) of the swallow signal.
- **ComPARE 2016 Feature set (OSF)**: 6373 features - Voice source based (group A) and glottal excitation based (group B) features

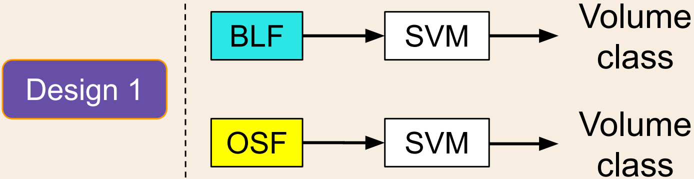
Classes	Feature name / Abbreviation / Dimension	Total
Group A (59 features + 59 delta features)	Loudness (Ldns)(1), RASTA (26), MFCC(14), RMS energy (RMSe)(1), Modulated Loudness (MLdns)(1), ZCR(1), Band energy (BE)(2), Spectral: ROP (SR)(4), Flux (SF)(1), Centroid (SC)(1), Slope (SSl)(1), Entropy (SE)(1), Variance (SV)(3), Harmonicity (SH)(1), Sharpness (Shs)(1)	59 X 54 functionals + 59 delta X 46 functionals = 5900 features
Group B (6 features + 6 delta features	F0(1), Prob. Voicing (PV)(1), Jitter (J)(2), Shimmer (Shr)(1), log(HNR)(1)	6 X 39 functionals + 6 delta X 39 functionals = 468 features

Experimental setup

- Baseline features computed from swallow signals annotated by a clinical expert
- ComPARE 2016 features computed using OpenSMILE Audio Extraction and Analysis Tool - **require no manual annotation**
- 10-fold cross-validation step (no common subjects across folds)
- Three classification tasks: Dry-vs-2ml, Dry-vs-5ml, Dry-vs-10ml
- Classifier: Linear SVM
- Grid search for optimal C-parameter selection
- Evaluation metric: F-score
- Significance test: Wilcoxon signed rank test for zero median

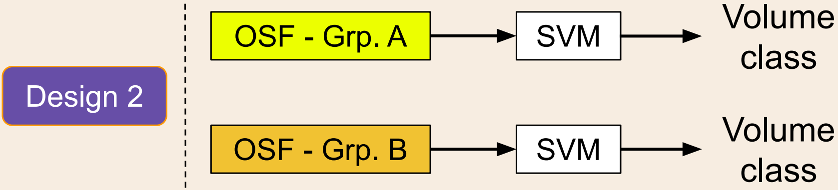
Study Design 1 - BLF vs OSF

- Assessment of overall performance of BLF and OSF
  Features requiring manual annotation vs features not requiring manual annotations



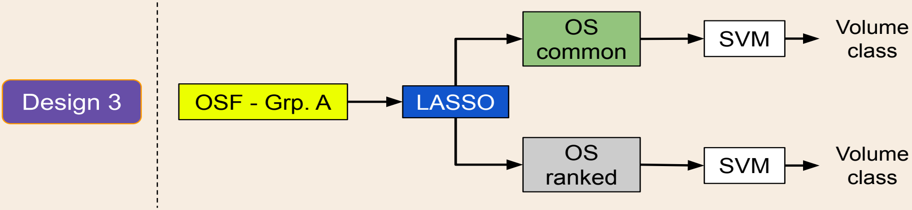
Study Design 2

- Assessment of performance of subgroups of OSF
  vocal tract related features vs glottal excitation related features



Study Design 3

- Feature ranking for each fold and each task can be different
- Joint features: Forward Feature Selection is computationally expensive!
  - LASSO[4] feature selection algorithm used to rank order OSF features
  - Features selected in two ways
    - OS-ranked: 12 features for each of the three classification tasks
    - OS-common: 12 features common across all three classification tasks



[4]Valeria Fonti and Eduard Belitser, "Feature selection usinglasso," *VU Amsterdam Research Paper in Business Analytics*, pp. 1–25, 2017

Plan

Results of Study 1: BLF **vs** OSF

Table: Mean **F-scores** (%) of the three classification tasks

Classification task	BLF	OSF	p-value
Dry vs 2ml	37.26 (15.05)	**70.14** (8.56)	0.03
Dry vs 5ml	57.77 (15.05)	**73.44** (3.03)	0.0098
Dry vs 10ml	69.10 (9.67)	**77.45** (6.65)	0.002

- OSF outperformed BLF by an average of around 18.9%
- Reduced standard deviation in OSF indicates the robustness of OSF, across subjects and volumes

Results of Study 2: OSF - Group A vs Group B

- Features unrelated to voice source outperformed features related to voice source by **21.02%**, BSF by **24.65%** and also OSF by **5.7%**
- Top performing features - MFCC, RASTA filtered audio spectrum and RMS energy

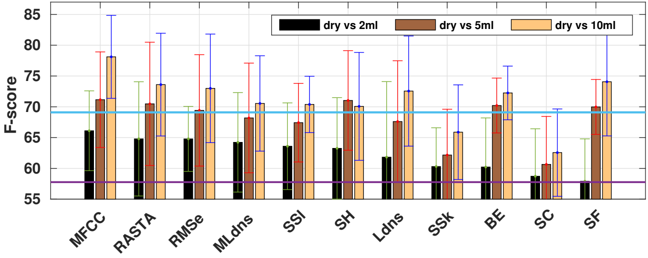


Figure: Mean F-scores(%) of different Group A features (ranked by F-scores of dry-vs-2ml classification); violet and blue horizontal lines are the baseline F-scores of dry-vs-5ml and dry-vs-10ml respectively

Results of Study 3

Mean F-scores (%) of OS-common and OS-ranked features with standard deviations in (.)

Volume / Feature set	OS-common	OS-ranked	p-value
Dry vs 2ml	73.55 (6.78)	74.84 (4.85)	0.6523
Dry vs 5ml	75.88 (9.05)	77.67 (8.17)	0.4316
Dry vs 10ml	80.68 (4.88)	80.36 (7.61)	1.0

- Performance of OS-common & OS-ranked features similar to each other

Top 12 OS-common features with Pearson correlation coefficient (PCC) between feature value and the volume of water in [.] P_k indicates k percentile

Loudness	**Derivative of Loudness**
Inter-quartile range [0.14] ($P_{75} - P_{50}$)	P_1 [0.39], Standard deviation[0.41], Kurtosis [0.03], Mean segment length [0.25], Minimum segment length [0.18], Up-level time 25,50,75,90 [0.03] (%) Risetime [0.023], Left curvature time [0.06]

- LASSO selected features seem to indirectly indicate swallow components' boundaries

Plan

Key Takeaways

- Eliminated need for expert manual annotation
  - Improved F-score of OSF over BLF
  - LASSO selected functionals provided cues & indirectly indicated the boundaries of SCs
- 12 functionals of Loudness outperformed (significantly different from) the 12 baseline features
- Lower standard deviations - proposed features more robust to variations in signal characteristics with volume

Future Works

- Exploring swallow sounds specific features unlike the generic OS features designed for speech analysis purposes
- Expanding the swallow sounds dataset with the inclusion of patient data of dysphagic swallows
- Classification of healthy and dysphagic swallows

References

- Tsuyoshi Honda, Takuro Baba, Keiko Fujimoto, TakaharuGoto, Kan Nagao, Masafumi Harada, Eiichi Honda, and TetsuoIchikawa, “Characterization of swallowing sound: Preliminaryinvestigation of normal subjects,”PloS one, vol. 11, no. 12, pp.e0168187, 2016
- Sylvain Morinière, Patrice Beutter, and Michèle Boiron, ”Sound component duration of healthy human pharyngoesophageal swallowing: a gender comparison study,” *Dysphagia*, vol. 21, no. 3, pp. 175–182, 2006.
- Florian Eyben, Martin Wöllmer, and Björn Schuller, ”openSMILE - the Munich Versatile and Fast Open-Source Audio Feature Extractor,” 01 2010, pp. 1459-1462

ACKNOWLEDGEMENTS

The authors thank **all the subjects** for participating in the swallow sound data collection.

The authors also thank **Pratiksha Trust** and **DST, Govt. of India** for their support.

THANK YOU

Have questions/suggestions?

Write to us at `spirelab.ee@iisc.ac.in`