WAVELET FEATURES FOR CLASSIFICATION OF VOTE SNORE SOUNDS

Kun Qian[1], Christoph Janott[2], Zixing Zhang[5], Clemens Heiser[3], Björn Schuller[4,5]
[1]MISP group, MMK, Technische Universität München, Germany [2]Institute for Medical Engineering, Technische Universität München, Germany
[3]Department of Otorhinolaryngology/Head and Neck Surgery, Klinikum rechts der Isar, Technische Universität München, Germany
[4]Machine Learning Group, Imperial College London, UK [5]Chair of Complex & Intelligent Systems, University of Passau, Germany

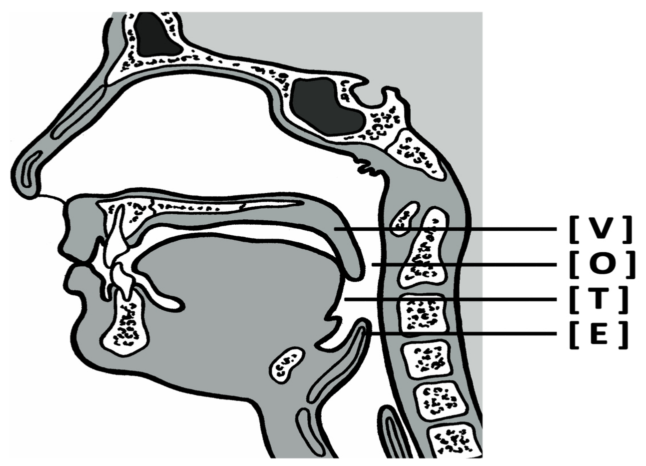
Introduction

- **VOTE** snore sounds (SnS) are respectively corresponding to
  - **velum**, **oropharyngeal lateral walls**, **tongue base**, and **epiglottis** SnS labelled during drug-induced sleep endoscopy.
- **Issues:**
  - Obstructive sleep apnea (OSA) diagnosis and evaluation
  - Location and form of the upper airway obstruction
  - Targeted therapy of OSA
  - Finding suitable acoustic features
- **Our Work:**
  - Present *wavelet packet transform (WPT)* features in classification of VOTE SnS
    - presentation of new features: *WPT features*
    - comparison with other 'old' features: *formants, MFCC, power ratio, crest factor, F0*

Database

- **Provider**: Klinikum rechts der Isar, Technische Universität München
- **Subjects**: 24 (all are males), 2 groups, 117 episodes segmented into basic instances (200 ms), subject-independent 2-fold cross validation
- **Sleep Type**: drug-induced sleep

VOTE Positions:

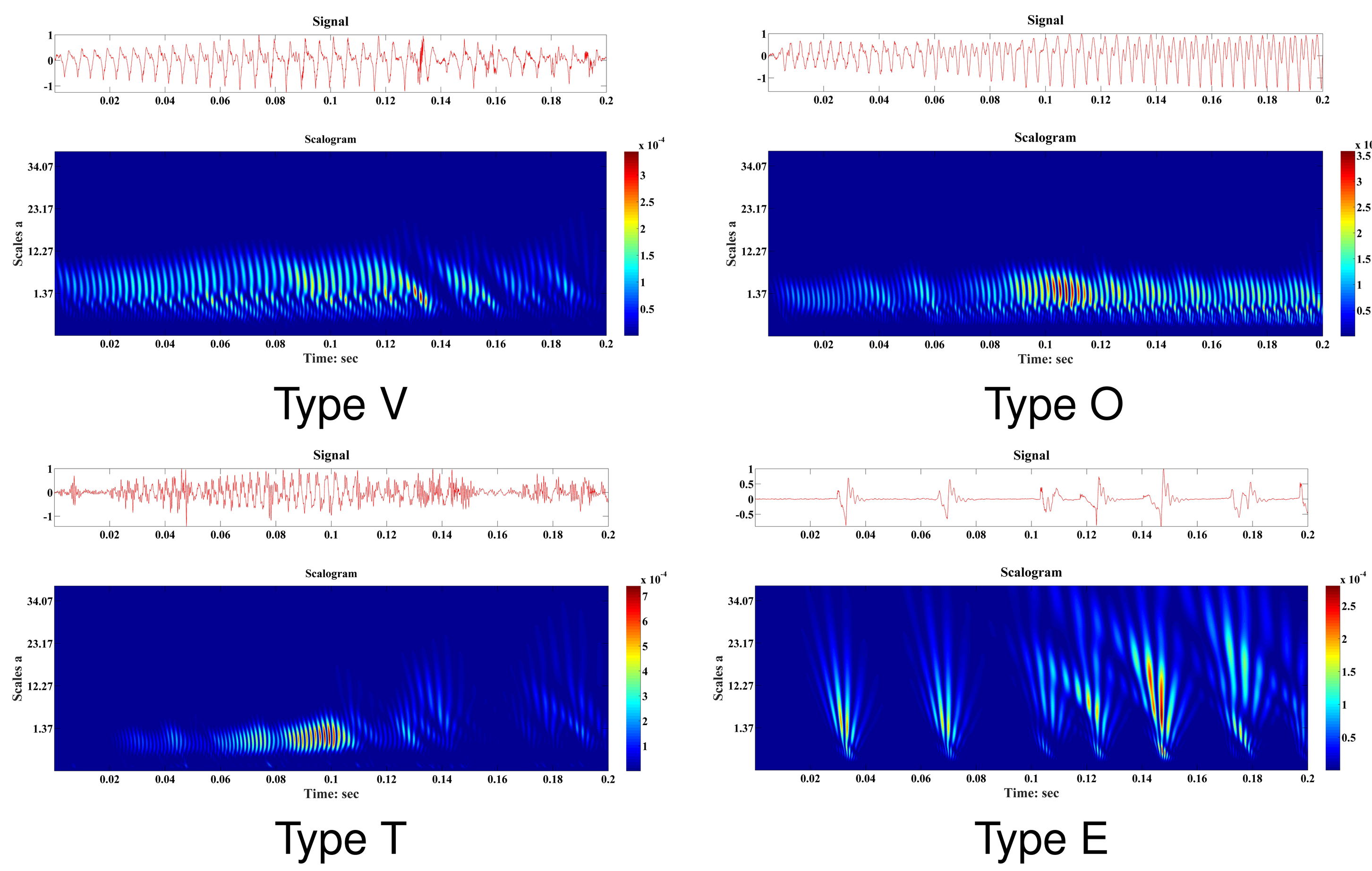


Data Collection:

	mean	std	range
Age $[years]$	46.2	13.2	26 – 72
BMI $[kg/m^2]$	26.8	2.9	18.9 – 31.5
AHI $[events/h]$	20.4	10.7	6.2 – 45.6

	V	O	T	E	All
Group 1	376	132	18	125	651
Group 2	434	111	46	141	732
Total	810	234	64	266	1 383

Scalograms of the VOTE SnS



The scalogram indicates the energy percentage for each wavelet coefficient (the wavelet function here is 'db10', and the decomposition level is 6.

Wavelet Packet Transform (WPT) Features

- Low-Level Descriptors (LLDs): Proposed by Khushaba et al., 2011.[a]

$$E_{V_{j,k}} = \sqrt{\frac{\sum_n (\mathbf{w}_{j,k,n})^2}{N_k}}, \quad (1)$$

where $\mathbf{w}_{j,k}$ represents the wavelet-packet transform coefficients evaluated from the signal at the subspace $V_{j,k}$, and N_k is the number of wavelet coefficients in the k-th subband; therefore, $E_{V_{j,k}}$ denotes the normalised bank filter energy in k-th subband with the j-th decomposition level. Furthermore, the subband energy percentage is defined as:

$$E_{V_j} = \frac{\sum_k \sum_n (\mathbf{w}_{j,k,n})^2}{\sum_{j=1}^{J_{max}} \sum_k \sum_n (\mathbf{w}_{j,k,n})^2}. \quad (2)$$

- Statistical Functions:

LLDs (155)	**Statistical functionals** (9)
$E_{V_{j,k}}$ (127)	max, min, mean,
E_{V_j} (7)	range, standard deviation,
Variance of E_{V_j} (7)	slope, bias (linear
Waveform length of E_{V_j} (7)	regression approximation)
Entropy of E_{V_j} (7)	skewness, kurtosis

[a]R. N. Khushaba, S. Kodagoda, S. Lal, and G. Dissanayake, "Driver drowsiness classification using fuzzy wavelet-packetbased feature-extraction algorithm", *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 1, pp. 121–131, 2011.

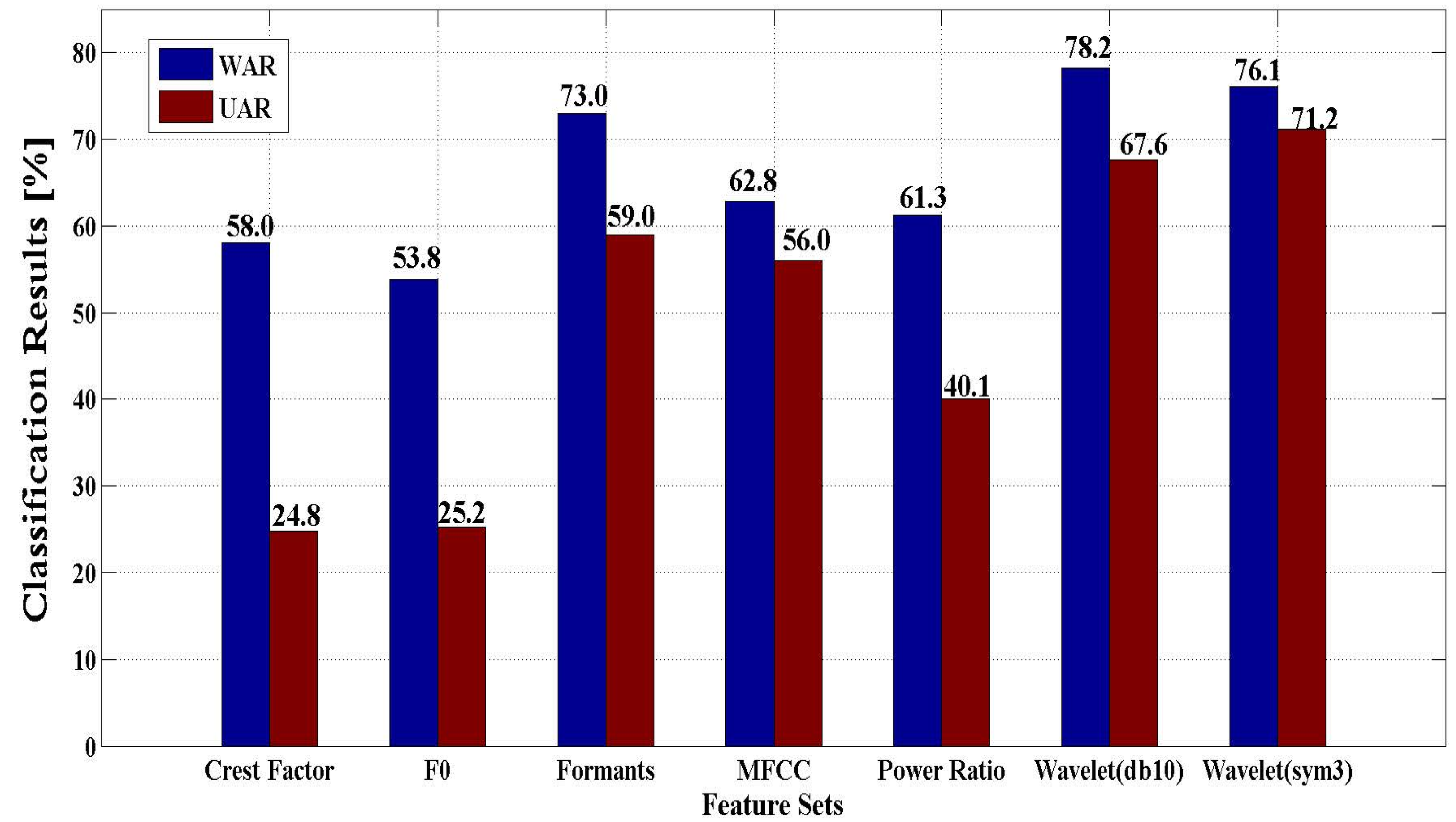
Experiments and Results

- Classifier: Support Vector Machines
- Evaluation: WAR (weighted average recall), UAR (unweighted average recall)
- Subject-independent two-fold cross validation

Maximun WAR and UAR of Different Wavelet Functions

Family	Optimum	WAR_{max} [%]	UAR_{max} [%]
BiorSplines	bior3.5	76.8	70.1
Coiflets	coif1, coif5	77.0 (coif5)	68.9 (coif1)
Daubechies	db3, db10	**78.2** (db10)	71.1 (db3)
Dmeyer	dmey	74.9	68.2
Haar	haar	66.6	61.0
ReverseBior	rbio1.5	76.4	69.7
Symlets	sym3	76.1	**71.2**

Comparison With Other Features



(We extract the formants , MFCC, power ratio, crest factor, and F0 from SnS and calculate the statistical values with the same approach as WPT features.)

Conclusion

- WPT features achieve excellent performance: 78.2 % WAR and 71.2 %UAR
- WPT features outperform other features: formants, MFCC, power ratio, crest factor, and F0
- Future work: collection of much more data, imbalanced data problem, deep learning model