MULTI-VIEW REPRESENTATION LEARNING VIA GCCA FOR MULTIMODAL ANALYSIS OF PARKINSON'S DISEASE

J. C. Vásquez-Correa, J. R. Orozco-Arroyave, R. Arora, E. Nöth, N. Dehak, H. Christensen, F. Rudzicz T. Bocklet, M. Cernak, H. Chinaei, J. Hannink, Phani Sankar Nidadavolu, M. Yancheva, A. Vann, N. Vogler

University of Antioquia UdeA, Medellín, Colombia.
Friedrich Alexander University, Erlangen, Germany.
Center for language and speech processing, Johns Hopkins University, USA

PROBLEM

Information from different bio–signals such as speech, handwriting, and gait have been used to monitor the state of Parkinson's disease (PD) patients, however, all the multimodal bio–signals may not always be available

CONTRIBUTIONS

- A multimodal monitoring of PD patients using features from speech, handwriting, and gait is proposed.
- We apply a method for multi-view learning based on the generalized canonical correlation analysis (GCCA), which transforms features from different modalities into a different feature space, where only one modality is available.

Speech Handwriting Gait

- Three scenarios are considered
  1. Classification of PD vs. HC speakers.
  2. Prediction of the neurological state of the patients (UPDRS-III score).
  3. Prediction of scale to assess only the speech deficits of patients (m-FDA score).

MULTI-VIEW LEARNING

The multi-view learning is performed using GCCA, with the aim of obtaining a feature embedding that represents the maximally correlated projection from the multimodal information and the speech, respectively. This projected feature space can be used also when the multimodal information is not available.

DATA

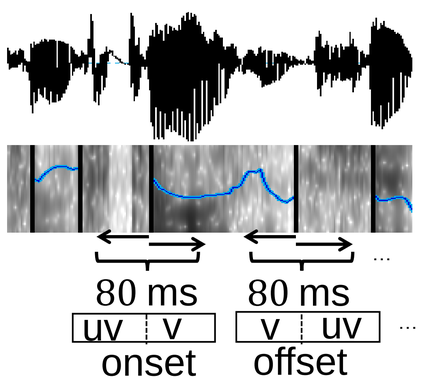
Train:
Recordings from speech in Spanish, handwriting and gait from 30 PD patients.

Test:
Speech in Spanish, German and Czech from PD patients and HC speakers.

FEATURE EXTRACTION

Speech

- **Articulation:**



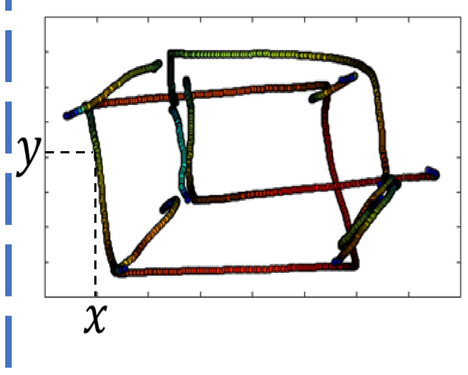
12 MFCC and 22 Bark band energies.

- **Prosody:**

Energy, duration and fundamental frequency based features

Handwriting

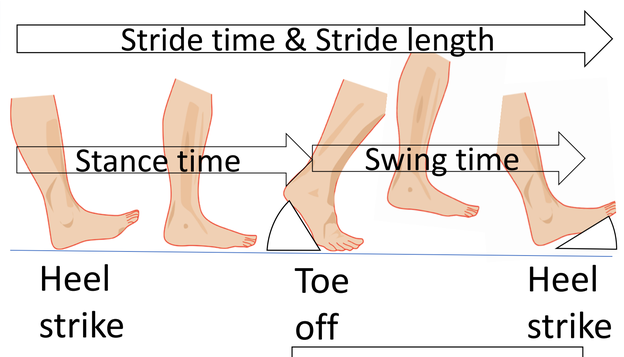
Kinematics and pressure.



- Speed of the stroke
- Velocity
- Acceleration
- Jerk
- In-air movement
- Pressure of the pen

Gait

Kinematics and bio-mechanical



- 3d accelerometer
- 3d gyroscope
- 2 feet

RESULTS

- **Baseline Multimodal Spanish**

Features & Task	UPDRS score	M-FDA score
Speech modality		
Art. /pa-ta-ka/	-0.33	**0.40**
Art. Monol.	-0.39	0.19
Art. Read text	0.19	0.13
Pros. Monol.	-0.23	0.22
Gait modality		
4x10 left	0.68	0.49
4x10 right	0.66	0.32
4x10 both	**0.72**	0.39
Handwriting modality		
Cube	0.48	-0.18
ID	0.47	0.25
Spiral	0.12	-0.22

- **Test results: Before and after GCCA**

	Before GCCA			After GCCA		
Features & Task	Class	UPDRS	M-FDA	Class	UPDRS	M-FDA
Spanish						
Art. /pa-ta-ka/	77%	0.34	0.67	78%	0.40	0.72
Art. Monol.	70%	0.32	0.39	73%	0.30	0.40
Art. Read text	78%	0.28	0.56	78%	0.39	0.59
Pros. Monol.	69%	-0.43	0.41	70%	0.14	0.40
German						
Art. /pa-ta-ka/	70%	0.11	-	71%	0.14	-
Art. Monol.	73%	0.01	-	74%	-0.03	-
Art. Read text	79%	0.03	-	76%	-0.69	-
Pros. Monol.	76%	-0.69	-	76%	0.40	-
Czech						
Art. /pa-ta-ka/	82%	0.29	-	82%	0.46	-
Art. Monol.	77%	-0.51	-	77%	0.12	-
Art. Read text	80%	-0.59	-	80%	-0.59	-
Pros. Monol.	69%	0.00	-	69%	0.51	-

CONCLUSION

- A method based on GCCA is applied to map features from three modalities into a different dataset that contains only features from one modality (speech).
- An improvement in the performance of the three tasks is observed.
- The proposed method provides additional information for the PD analysis, even when the language of the test data is different.

FUTURE DIRECTION

- New features from gait and handwriting with the aim of improving the results.
- Collecting multimodal data from healthy controls.
- Other approaches for multi-view learning

42nd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2017.

Contact.
jcamilo.vasquez@udea.edu.co