Multimodal i-vectors to Detect and Evaluate Parkinson's Disease

Nicanor García[1], **Juan Camilo Vásquez-Correa**[1,2], Juan Rafael Orozco-Arroyave[1,2], and Elmar Nöth[1,2]

[1] Faculty of Engineering, University of Antioquia, Medellin, Colombia

[2] Pattern Recognition Lab, Friedrich-Alexander University of Erlangen-Nürnberg

September 4, 2018

Introduction: Parkinson’s Disease (PD)

- Second most prevalent neurological disorder worldwide.
- Patients develop several motor and non-motor impairments.
- Patients are affected by gait, handwriting, and speech disorders, e.g., freezing of gait, micrographia, dysarthria.

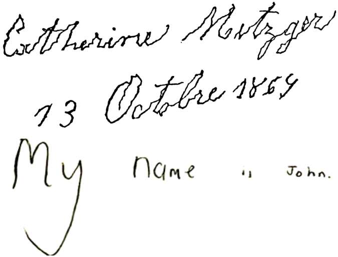
Introduction: Parkinson's Disease (PD)

- Second most prevalent neurological disorder worldwide.
- Patients develop several motor and non-motor impairments.
- Patients are affected by gait, handwriting, and speech disorders, e.g., freezing of gait, micrographia, dysarthria.
- The diagnosis and assessment of the progression of the disease are subject to clinical criteria.
- The neurological condition of the patients can be assessed using the MDS-UPDRS scale.

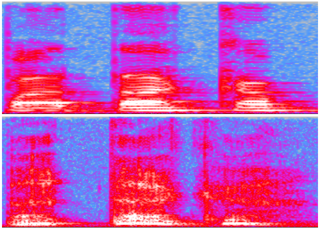
Introduction: motor disorders

Gait: Freezing of gait

Handwriting: Tremor and micrographia



Speech: Hypokinetic dysarthria



Introduction: Motivation and Hypothesis

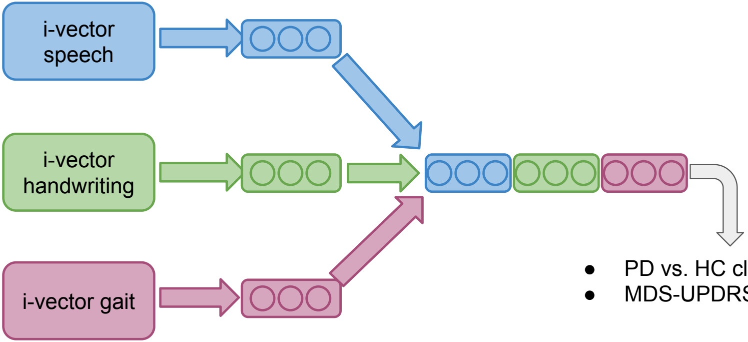
- i-vectors are considered the state-of-art in speaker verification, and also have proofed to be accurate to detect other traits from speech, including the presence of PD[1].
- The i-vector approach has been adapted for other biometric verification tasks considering handwriting and gait.
- Related studies suggest that i-vectors are able to capture the traits of a person in different bio-signals.

We believe that i-vectors can also capture the effect of PD in handwriting and gait, and such information is complementary to that one provided by speech signals to detect the presence of the disease and to evaluate the neurological state of the patients.

[1] N. Garcia et al. (2017). "Evaluation of the neurological state of people with Parkinson's disease using i-vectors". In: *Proc. of the 18th INTERSPEECH*.

Introduction: Aims

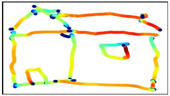
- Multimodal assessment of PD.
- Classification of PD patients and healthy control (HC) subjects.
- Evaluation of the neurological state of the patients.
- i-vectors are extracted from different bio-signals.
- Two fusion strategies are proposed to combine multimodal information.

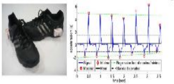


- PD vs. HC classification
- MDS-UPDRS prediction

Materials and Methods

Bio-signals → Feature extraction → i-vector extraction → i-vector post-processing → Fusion of modalities





- PD vs. HC classification
- MDS-UPDRS prediction

Materials and Methods: Multimodal data

- Speech, Handwriting and Gait from:
  - 49 patients (average age 60 ± 10.0 years). Most of them in early to mid-stages of the disease.
  - 41 healthy subjects (average age 65.1 ± 10.8) years.
- Gait signals captured with inertial sensors attached to the lateral heel of the shoe (100 Hz, 12-bit resolution).
- Handwriting signals captured with a digitizing tablet with a sampling frequency of 180 Hz and 12-bit resolution.
- Several exercises are performed by the participants in each modality.
  - Speech: ten sentences.
  - Handwriting: name, signature, sentence, and different drawings.
  - Gait: 40 meters walk in straight line with stops every 10 meters.

Materials and Methods: Feature Extraction

- **Gait:**
  Eight modified MFCCs extracted for frames with 320 ms length from the triaxial accelerometers and gyroscopes from both foot[2].

  Non-linear spectral representation with more resolution in the lower frequency bands.

- **Handwriting:**
  x, y, and z-positions; azimuth and altitude angles; pressure of the pen.

  In addition with their first two derivatives[3].

- **Speech:**

  20 MFCCs (including MFCC_0) with their first two derivatives extracted for frames with 25 ms length with a time-shift of 10 ms.

---

[2] R. San-Segundo et al. (2016). "Feature extraction from smartphone inertial signals for human activity segmentation". In: *Signal Processing* 120, pp. 359–372.

[3] P. Drotár et al. (2016). "Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease". In: *Artificial Intelligence in Medicine* 67.C, pp. 39–46. ISSN: 0933-3657.

Materials and Methods: i-vector extraction

- Universal background models were trained for the features extracted from each bio-signal.
- i-vectors were extracted for each subject and for each task.
- The dimension of the i-vector is given by[4]:

$$\dim_w = N \cdot \log_2(M)$$

N: number of features.
M: number of Gaussian components.

[4] N. Garcia et al. (2017). "Evaluation of the neurological state of people with Parkinson's disease using i-vectors". In: *Proc. of the 18th INTERSPEECH*.

Materials and Methods: i-vector post-processing

- i-vectors of the different tasks of a given subject are averaged to obtain one i-vector per subject.
- Principal Component Analysis (PCA) is applied to the subject i-vectors to perform a whitening transformation[5].

[5] D. Garcia-Romero and C. Espy-Wilson (2011). "Analysis of i-vector Length Normalization in Speaker Recognition Systems.". In: *Proc. of the 12th INTERSPEECH*.

Materials and Methods: Fusion of modalities

1. Super i-vector $\mathbf{w}_f$: concatenating the i-vectors from each bio-signal.

$$\mathbf{w}_f = \begin{bmatrix} \mathbf{w}_h \\ \mathbf{w}_g \\ \mathbf{w}_s \end{bmatrix}_{(H+G+S)\times 1}$$

   H, G, and S are the dimension of each modality i-vector.

2. Score fusion: the scores of the predictions obtained from each bio-signal are averaged.

Methodology: Classification and neurological state assessment

- **Classification**: A soft margin Support Vector Machine (SVM) with Gaussian kernel is used.
- **Neurological state assessment**: comparison between the subject's i-vector and a set of N reference i-vectors using the cosine distance:

$$\mathrm{d}(\mathbf{w}_{\text{test},j}) = \frac{1}{N}\sum_{i=1}^{N}\left(1 - \frac{\mathbf{w}_{\text{test},j}\cdot\mathbf{w}_{\text{ref},i}}{||\mathbf{w}_{\text{test},j}||\,||\mathbf{w}_{\text{ref},i}||}\right)$$

Methodology: Classification and neurological state assessment

- **Classification**: A soft margin Support Vector Machine (SVM) with Gaussian kernel is used.
- **Neurological state assessment**: comparison between the subject's i-vector and a set of *N* reference i-vectors using the cosine distance:

$$\mathrm{d}(\mathbf{w}_{\text{test},j}) = \frac{1}{N}\sum_{i=1}^{N}\left(1 - \frac{\mathbf{w}_{\text{test},j}\cdot\mathbf{w}_{\text{ref},i}}{||\mathbf{w}_{\text{test},j}||\,||\mathbf{w}_{\text{ref},i}||}\right)$$

Validation

- A five-fold cross-validation scheme is implemented for the classification experiment.
- To minimize possible bias due to the different microphones and shoes used to capture the signals, the patients of each fold were balanced according to these condition.

Results

Table: Classification of Parkinson's patients and healthy subjects

Signal	Acc. (%)	Sens. (%)	Spec. (%)	AUC
Gait	76.9 ± 9.1	77.1 ± 11.5	76.8 ± 12.5	0.83
Handwriting	75.1 ± 3.7	79.3 ± 7.4	70.0 ± 17.0	0.82
Speech	79.4 ± 7.8	83.1 ± 15.2	75.0 ± 17.7	0.87
Super i-vector	**85.0 ± 9.6**	**81.3 ± 12.4**	**89.6 ± 9.5**	**0.92**

- Fusion of modalities provides the highest accuracy.
- Among the three bio-signals, speech is the modality that provides the best accurate results.

Results

Table: Spearman's correlation between the cosine distance and the MDS-UPDRS-III

Signal	ρ young healthy subjects ref.	ρ elderly healthy subjects ref.	ρ Patients ref.
Gait	−0.14	−0.11	−0.25
Handwriting	0.20	−0.07	−0.18
Speech	−0.14	0.30	−0.33
Super-i-vector	0.03	−0.08	−0.26
Score fusion	**0.31**	**0.20**	**−0.41**

- Positive correlation with respect to healthy subjects reference i-vectors.
- Negative correlation with respect to patient's reference i-vectors.
- Score fusion is the most correlated with the neurological state of the patients.

Conclusion

- A multimodal analysis of Parkinson's disease is proposed considering i-vectors extracted from different bio-signals: speech, handwriting and gait.
- Two fusion strategies were evaluated to combine information from different bio-signals.
- The super i-vector fusion method improved the accuracy of classification between PD and HC; however, it is not suitable to assess the neurological state of the patients.
- The score fusion slightly improved the correlation with the neurological state of the patients.

Conclusion

- Additional features need to be explored to model the gait and handwriting signals.
- The i-vector approach might need to be adapted in its core to model other bio-signals.
- Other fusion strategies could be addressed to improve the results.

Thanks for attending.

Any questions?

`juan.vasquez@fau.de`

`www5.cs.fau.de/en/our-team/vasquez-camilo`

TAPAS

Training Network on Automatic Processing of PAthological Speech (TAPAS)
Horizon 2020 Marie Sklodowska-Curie Actions Initial Training Network European Training Network (MSCA-ITN-ETN) project.

References I

San-Segundo, R. et al. (2016). "Feature extraction from smartphone inertial signals for human activity segmentation". In: *Signal Processing* 120, pp. 359–372.

Garcia-Romero, D. and C. Espy-Wilson (2011). "Analysis of i-vector Length Normalization in Speaker Recognition Systems.". In: *Proc. of the 12th INTERSPEECH*.

Garcia, N. et al. (2017). "Evaluation of the neurological state of people with Parkinson's disease using i-vectors". In: *Proc. of the 18th INTERSPEECH*.

Drotár, P. et al. (2016). "Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease". In: *Artificial Intelligence in Medicine* 67.C, pp. 39–46. ISSN: 0933-3657.

Multimodal i-vectors to Detect and Evaluate Parkinson's Disease

Nicanor García[1], **Juan Camilo Vásquez-Correa**[1,2], Juan Rafael Orozco-Arroyave[1,2], and Elmar Nöth[1,2]
[1] Faculty of Engineering, University of Antioquia, Medellin, Colombia
[2] Pattern Recognition Lab, Friedrich-Alexander University of Erlangen-Nürnberg
September 4, 2018