PhD subject:

# DAMoS: Deep Analysis of Motor Symptoms for Dementia with Lewy Bodies

## Hosting institute

ICube Laboratory (Le laboratoire des sciences de l'ingénieur, de l'informatique et de l'imagerie: The Engineering science, computer science and imaging laboratory) at the University of Strasbourg is a leading research center in Computer Science, with more than 300 permanent researchers, with the recently opened AI graduate school supported by the French government.

## Work place

The PhD will take place in MLMS (Machine Learning, Modeling & Simulation, 2021–) research team located at the hospital site, 10 min walking distance to the city center of Strasbourg, a UNESCO world heritage site.

## Funding source

l'Agence Nationale de la Recherche (ANR: French National Research Foundation) through AI graduate school & University of Strasbourg.

## Staring date

October 2021.

## Supervisors

Supervisor: Hyewon Seo (ICube, Univ. Strasbourg), co-supervisor: Frédéric Blanc (ICube, Univ. Strasbourg and Univ. Hospital of Strasbourg)

## Context, objectives, and positioning

The diagnosis of dementia with Lewy bodies (DLB), a disease associated with abnormal deposits of synuclein (specific protein) in the brain, can be challenging. Its early symptoms are often confused with similar symptoms found in other brain diseases like Alzheimer's disease or in psychiatric disorders like schizophrenia. The most common DLB signs and symptoms are changes in cognition, movement, and behavior. In this thesis, we will develop a learning-based approach to model face and eye movement from video input, with a specific focus on modeling the motor symptoms of DLB [1][2]. Our primary objective will be on modeling and detecting facial motor symptoms, such as reduced facial expression, and facial expressions showing behavioral symptoms (depression, apathy, or agitation). Other motor symptoms related to cognitive symptoms will also be considered, which requires analyzing the eye movement: Troubles or unpredictable changes in visual attention, executive, visual and spatial abilities (judging distance and depth or misidentifying objects), and movements showing cognitive fluctuations or visual hallucinations.

Robust three-dimensional reconstruction, analysis, and characterization of shape and motion of individuals or groups of people from one or more video images have been open problems for decades, with many exciting application areas such as early abnormality detection in predictive clinical analysis. A common way to acquire necessary 3D data and model is to use calibrated multi-view passive cameras [3-6] to merge a sparse or dense set

reconstructed depth images into a single mesh, but the size and cost of such multi-view systems prevent their use in consumer applications. In more unconstrained and ambiguous settings, such as in the monocular image or video, priors in the form of template or parametric model derived from a large dataset are often used, which help to constrain the problem significantly [7][8][9]. While *generative* methods reconstruct the moving geometry by optimizing the alignment between the projected model and the image data [10][11], *regressive* methods [9][12][13][14] train deep neural networks to infer shape parameters of a parametric body model from a single image. Despite remarkable progress, the reconstruction and analysis of facial models from video has not been fully addressed yet, with most existing algorithms operating on 'normal' faces, and in a frame-by-frame manner. In this study, we will (1) address the relatively unspoken problem and data, i.e. abnormal face and eye movements, and (2) include a temporal aspect of the facial movement into a learned model, a work that has not been done before.

## Work plan

We will deploy recent deep learning techniques to approach the challenging problem of detection and analysis of facial motor symptoms of DLB from video. Devoted learning-based models will be developed to model the face and eye movements, which will then be integrated into a DLB diagnoser. In both cases, we aim to build our models in 3D, meaning that (1) A 2D-to-3D reconstruction will be preceded for the facial modeler, and (2) 2-dimensional visual stimuli for the eye tracking tests will be generated from 3D models, and the 2D fixation map back-projected to 3D. The work is articulated in three parts:

1. *Face movement modeler.* A model-based DNN (deep neural network) will be developed, which will learn to jointly regress the 3D facial shape and movement (head pose, and pose-dependent shape change) from the monocular video input. Following our recent success on the DNN-based facial animation modeling [15], a recurrent neural network will be adopted, which has shown to achieve promising results in modeling sequential, time-series data.

2. Eye *movement modeler* will be trained to model and analyze the sequence data of saccades and fixations on observed visual stimuli, which we will acquire by using an eye-tracker.

3. *DLB diagnoser.* Both aforementioned modelers will be integrated into a DLM diagnoser, with a capability of detecting some of the known motor-, cognitive-, and behavioral symptoms.

The observation data of patients and normal aged populations will be collected in collaboration with the University hospital, and other publicly available resources.

## Candidate profile

— Master's degree in Computer Science, Electrical Engineering or Applied Mathematics
— Solid programming skills in Python/Matlab
— Solid knowledge in deep learning with programming experience in Tensorflow or Pytorch
— Working knowledge in geometry modeling and statistics
— Good communication skills

## Application

Please send your CV, academic transcriptions (undergraduate and master courses with rankings) and one or more recommendation letters to seo@unistra.fr and f.blanc@unistra.fr.

## References

[1] Botzung, Anne & Philippi, Nathalie & Noblet, Vincent & de Sousa, Paulo & Blanc, Frederic. (2019). Pay attention to the basal ganglia: a volumetric study in early dementia with Lewy bodies. Alzheimer's Research & Therapy. 11. 10.1186/s13195-019-0568-y. Diagnosis and management of dementia with Lewy bodies.

[2]  McKeith IG, Boeve BF, Dickson DW, Halliday G, Taylor JP, Weintraub D, Aarsland D, Galvin J, Attems J, Ballard CG, Bayston A, Beach TG, Blanc F, Bohnen N, Bonanni L, Bras J, Brundin P, Burn D, Chen-Plotkin A, Duda JE, El-Agnaf O, Feldman H, Ferman TJ, Ffytche D, Fujishiro H, Galasko D, Goldman JG, Gomperts SN, Graff-Radford NR, Honig LS, Iranzo A, Kantarci K, Kaufer D, Kukull W, Lee VMY, Leverenz JB, Lewis S, Lippa C, Lunde A, Masellis M, Masliah E, McLean P, Mollenhauer B, Montine TJ, Moreno E, Mori E, Murray M, O'Brien JT, Orimo S, Postuma RB, Ramaswamy S, Ross OA, Salmon DP, Singleton A, Taylor A, Thomas A, Tiraboschi P, Toledo JB, Trojanowski JQ, Tsuang D, Walker Z, Yamada M, Kosaka K (2017) Diagnosis and management of dementia with Lewy bodies: Fourth consensus report of the DLB Consortium. Neurology 89, 88-100

[3]  J. Carranza, C. Theobalt, M. A. Magnor, and H.-P. Seidel. Free-viewpoint video of human actors. In ACM Transactions on Graphics, volume 22, pages 569–577. ACM, 2003.

[4]  J. Starck and A. Hilton. Surface capture for performance-based animation. IEEE Computer Graphics and Applications, 27(3), 2007.

[5]  E. De Aguiar, C. Stoll, C. Theobalt, N. Ahmed, H.-P. Seidel, and S. Thrun. Performance capture from sparse multiview video. In ACM Transactions on Graphics, volume 27, page 98. ACM, 2008.

[6]  B. Allain, J.-S. Franco, and E. Boyer. An Efficient Volumetric Framework for Shape Tracking. In IEEE Conf. on Computer Vision and Pattern Recognition, pages 268–276, Boston, United States, 2015. IEEE.

[7]  V. Blanz and T. Vetter. A morphable model for the synthesis of 3d faces. In Proc. SIGGRAPH, pages 187–194. ACM Press/Addison-Wesley Publishing Co., 1999.

[8]  Tran, Luan & Liu, Xiaoming. (2018). Nonlinear 3D Face Morphable Model. 7346-7355. 10.1109/CVPR.2018.00767.

[9]  Deng, Yu & Yang, Jiaolong & Xu, Sicheng & Chen, Dong & Jia, Yunde & Tong, Xin. (2019). Accurate 3D Face Reconstruction with Weakly-Supervised Learning: From Single Image to Image Set. WORKSHOPS (CVPRW) (2019): 285-295.

[10] F. Bogo, A. Kanazawa, C. Lassner, P. Gehler, J. Romero, and M. J. Black. Keep it SMPL: Automatic estimation of 3D human pose and shape from a single image. In European Conf. on Computer Vision. Springer International Publishing, 2016

[11] Thiemo Alldieck, Marcus Magnor, Weipeng Xu, Christian Theobalt, and Gerard Pons-Moll. (2018). Video Based Reconstruction of 3D People Models. CVPR 2018 Spotlight, IEEE Conference on Computer Vision and Pattern Recognition 2018 (CVPR)

[12] Dibra, Endri & Jain, Himanshu & Oztireli, Cengiz & Ziegler, Remo & Gross, Markus. (2017). Human Shape from Silhouettes Using Generative HKS Descriptors and Cross-Modal Neural Networks. 5504-5514. 10.1109/CVPR.2017.584.

[13] Richardson, Elad & Sela, Matan & Or-El, Roy & Kimmel, Ron. (2017). Learning Detailed Face Reconstruction from a Single Image. 5553-5562. 10.1109/CVPR.2017.589.

[14] Tewari, Ayush & Zollhöfer, Michael & Kim, Hyeongwoo & Garrido, Pablo & Bernard, Florian & Pérez, Patrick & Theobalt, Christian. (2017). MoFA: Model-based Deep Convolutional Face Autoencoder for Unsupervised Monocular Reconstruction.

[15] Seo H. and Luo G., "Generating 3D Facial Shapes and Expressions with Neural Networks", to appear in the book "Intelligent Scene Modelling and Human Computer Interaction", in press, Springer, 2021.