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 and salary

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

Salary: 1,500€/month approximately for a duration of 36 months (2021 - 2024).

Funding source

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

Staring date

October 2021.

Supervisors

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

Context

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. Initially, efforts were focused on facial reconstruction [1][2][3] and analysis, and later evolved into the body [4-11]. A common way to acquire necessary 3D data and model is to use calibrated multi-view passive cameras [4-8] 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. While *generative* methods reconstruct the moving geometry by optimizing the alignment between the projected model and the image data [9][10], *regressive* methods [3][11][12][13] 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 human models from video has not been fully addressed yet, with most existing algorithms operating in a frame-by-frame manner.

In this thesis, we will focus on the analysis-by-synthesis of human face and body from video input, with a specific focus on the motor symptoms of dementia with Lewy bodies [14]. The diagnosis of Lewy body dementia (LBD), a disease associated with abnormal deposits of a protein in the brain, can be challenging, as its early symptoms are often confused with similar symptoms found in other brain diseases like Alzheimer's or in psychiatric disorders like schizophrenia. The most common LBD signs and symptoms are changes in cognition, movement, and behavior. In particular, we are concerned with movement symptoms: a change in handwriting, gait disturbance, muscle rigidity or stiffness, reduced facial expression, slow movement, frozen stance, tremor or shaking, balance problems and repeated falls. Other motor symptoms related to cognitive symptoms will also be considered: trouble with attention, visual and spatial abilities (judging distance and depth or misidentifying objects), and movements showing cognitive fluctuations (unpredictable changes in concentration, attention) or visual hallucinations. Last but least, we will model facial movements showing behavioral symptoms: depression, apathy, anxiety, anxiety, or agitation.

PhD objectives

We will approach the challenging problem of detection and analysis of motor symptoms of DLB from video by leveraging recent deep learning techniques. A devoted learning-based model will be developed for the face and body, and in both cases, we will perform our study in 3D, meaning that a 2D-to-3D reconstruction will be preceded. The observation data of patients and normal aged populations will be collected in collaboration with the University hospital, and other publicly available resources.

Given the context described above, the work will be articulated in three parts:

(1) 3D facial expression modeling, reconstruction, and classification from videos. The main goal is to model-based recurrent neural network that will learn to jointly regress the head pose, identity shape, and pose-dependent shape change from the input video.

(2) Eye movement analyzer model. A sequence of saccades and fixations on observed visual stimuli will be analyzed, to detect some of the cognitive symptoms such as troubles or unpredictable changes in visual attention, and the presence of visual hallucinations. An eye-tracker will be deployed, as one main way of studying the visual attention.

(3) 3D body pose and movements estimation from videos, and detection of motor symptoms: We will train a deep neural network with a few daily motions such as walking or standing, with a capability of detecting some of the known motor symptoms: slow movement, frozen stance, tremor or shaking, balance problems and falls.

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.

References

[1] 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.

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

[3] 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.

[4] 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.

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

[6] 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.

[7] N. Hasler, H. Ackermann, B. Rosenhahn, T. Thormahlen, and H.-P. Seidel. Multilinear pose and body shape estismation of dressed subjects from image sets. In IEEE Conf. on Computer Vision and Pattern Recognition, pages 1823–1830. IEEE, 2010.

[8] 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.

[9] 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

[10]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)

[11]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.

[12] 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.

[13] 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.

[14] 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.