Huỳnh Văn Thống

vthuynh@ieee.org | pr.ai.vn/vthuynh | 0000-0002-6240-2959 | Google scholar

Summary

Over the past few years, my research has focused on facial behavior analysis in video, with a particular emphasis on investigating the complex interplay between human emotions and digital media. Through my work, I have developed advanced algorithms and models to accurately detect and interpret emotions from visual content. Looking ahead, I am eager to apply my expertise and expand my horizons in diverse areas such as human-computer interaction, finance, healthcare, etc. With a track record of successful research projects and a commitment to innovation, I am well positioned to make significant contributions to drive advancements in various domains with AI.

Education

Chonnam National University *Gwangju, South Korea, Mar 2020 - Feb 2023*

PhD in Artificial Intelligence Convergence *Overall GPA: 4.18/4.5*

Thesis: *Emotional Analysis of Video Clips by Utilizing Temporal Dependency Modeling*. Advisor: Prof. Kim Soo-Hyung

Chonnam National University *Gwangju, South Korea, Mar 2018 - Feb 2020*

MEng in Computer Science *Overall GPA: 4.38/4.5*

Thesis: *Gaze Tracking for Facial Emotion Analysis*. Advisor: Prof. Kim Soo-Hyung

Ho Chi Minh City University of Technology (HCMUT) *Ho Chi Minh City, Vietnam, Sep 2013 - Feb 2018*

BEng in Computer Science *Overall GPA: 3.20/4.0*

Thesis: *Extracting the table region in document images*. Advisor: Dr. Nguyen An Khuong and Dr. Tran Tuan Anh

Experience

Dept. of ITS, FPT University HCMC *HoChiMinh, Vietnam, Oct 2023 - Present*

Lecturer and Researcher - `https://pr.ai.vn/`

- Affective computing in healthcare with video and physiological signals.
- Multimodal representation learning – integrating information from multiple heterogeneous sources of data.

Dept. of AI Convergence, Chonnam National University *Gwangju, South Korea, March 2023 - Sept 2023*

Postdoctoral researcher, Pattern Recognition Lab - `https://prlabjnu.github.io/`

- Long-term video understanding.
- Multimodal representation learning – integrating information from multiple heterogeneous sources of data.

Dept. of AI Convergence, Chonnam National University *Gwangju, South Korea, March 2018 - Feb 2023*

Research assistant, Pattern Recognition Lab - `https://prlabjnu.github.io/`

- Estimating student engagement intensity in online classroom, and Eye gaze estimation on mobile (Android). Technology transfer: MayfarmSoft - September 2020.
- Pain intensity estimation from facial expressions.
- Developed a system for emotion recognition from videos of Korean based on face, audio and text emotion.
- Analyzing facial behaviors to detect discrete/continus emotion and action units.
- Eye semantic segmentation in 2D (images) and 3D (point cloud) for AR/VR environment.
- Estimating evoked expression from untrimmed videos.

DEK Technologies *HoChiMinh city, Vietnam, Jun 2016 - Aug 2016*

Software Engineer Internship

Selected Publications

Patents

1. Kim, S. H., Lee, G. S., Yang, H. J., **Huynh, V. T.**, & Oh, A. R. (2021). Apparatus and Method for Identifying Emotion by Gaze Movement Analysis. Korea Patent Registration Number: *10-2204743*. Technology transfer: MayfarmSoft - September 2020.

Journal

1. **Huynh, V. T.**, Yang, H. J., Lee, G. S., & Kim, S. H. Generic Event Boundary Detection in Video with Pyramid Features, 2023, **Under review**.
2. **Huynh, V. T.**, Yang, H. J., Lee, G. S., & Kim, S. H. Prediction of Evoked Expression from Videos with Temporal Position Fusion. *Pattern Recognition Letters*, 172, 245-251, doi:10.1016/j.patrec.2023.07.002. (**ISI/Scopus: Q1**)
3. **Huynh, V. T.**, Yang, H. J., Lee, G. S., & Kim, S. H. End-to-end learning for multimodal emotion recognition in video with adaptive loss. *IEEE MultiMedia*, 28 (2), 59–66. doi:10.1109/MMUL.2021.3080305.
4. **Huynh, V. T.**, Yang, H. J., Lee, G. S., & Kim, S. H. Semantic segmentation of the eye with a lightweight deep network and shape correction. *IEEE Access*, 8, 131967–131974. doi:10.1109/ACCESS.2020.3010011.
5. Jacomini, J.P., Liu, S., Shen, Y., **Huynh, V.T.**, Rickli, J.L., Yang, H.J., Kim, S.H. and Kim, K.Y., Empirical study for human engagement in collaborative robot programming. *Journal of Integrated Design and Process Science*, 26 (2), 159-181. doi:10.3233/JID-221012.

Conferences

1. Vu, N. T., **Huynh, V. T.**, Yang, H. J., & Kim, S. H. Multiscale Transformer-based for Multimodal Affective States Estimation from Physiological Signals. The 7th Asian Conference on Pattern Recognition, Lecture Notes in Computer Science, vol 14408. Springer, Cham.
2. Vu, N. T., **Huynh, V. T.**, Yang, H. J., Zaheer, M. Z., Nawaz, S., Nandakumar, K., & Kim, S. H. DCTM: Dilated Convolutional Transformer Model for Multimodal Engagement Estimation in Conversation. *The 31st ACM International Conference on Multimedia*, 2023, to be available. doi:10.1145/3581783.3612857
3. Nguyen, H. H, **Huynh, V. T.**, & Kim, S. H. An Ensemble Approach for Facial Behavior Analysis in-the-wild Video. *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2022, 2511-2516. doi:10.1109/CVPRW56347.2022.00281
4. **Huynh, V. T.**, Yang, H. J., Lee, G. S., & Kim, S. H. Multimodality pain and related behaviors recognition based on attention learning. *IEEE International Conference on Automatic Face and Gesture Recognition*, 2020, 814-818. doi:10.1109/FG47880.2020.00034
5. **Huynh, V. T.**, Yang, H. J., Lee, G. S., & Kim, S. H. Engagement intensity prediction with facial behavior features. *ACM International Conference on Multimodal Interaction*, 2019, 567-571. doi:10.1145/3340555.3355714
6. **Huynh, V. T.**, Le, B. K. T., Tran, T. A., Nguyen, A. K.,Yang, H. J., & Kim, S. H. Learning to detect tables in document images using line and text information. *International Conference on Machine Learning and Soft Computing*, 2018, 151-155. doi:10.1145/3184066.3184091

Skills

Technical skills

- Proficiency in programming languages as Python, MATLAB.
- Proficiency in OpenCV or CV libraries & deep learning frameworks such as TensorFlow, PyTorch.
- Familiar with classification algorithms, CNNs, and related machine learning/deep learning techniques.
- Experience with developing and implementing algorithms for video analysis (e.g., segmentation, object detection, behavior analysis).
- Experience with Linux commands.

Soft Skills Self-study, Time Management, Teamwork, Problem-solving, Documentation.

Honors and Awards

Jun 2023	**2nd place**, Generic Event Boundary Detection Challenge in Long-form Video Understanding (CVPRW)	*Canada*
Jun 2022	**3rd place**, Action Unit detection Challenge in Affective Behavior Analysis in-the-wild (CVPRW)	*USA*
Feb 2022	**Scholarship**, BK21 Senior Fellowship - Chonnam Nat'l Univ.	*South Korea*
Jun 2021	**1st place**, Evoked Expressions from Videos (EEV) Challenge in Affective Understanding in Video (CVPRW)	*USA*
Nov 2021	**2nd place**, 3D Point Cloud Segmentation Challenge - Facebook Reality Labs@OpenEDS 2021 (ICCVW)	*Canada*
Nov 2020	**1st place**, Pain Intensity Estimation from Facial Expressions in EmoPain challenge (FG)	*Argentina*
Feb 2020	**Scholarship**, Strategic Researcher Scholarship - Chonnam Nat'l Univ.	*South Korea*
Oct 2019	**1st place**, Engagement prediction in the Wild in 7th Emotion Recognition in the Wild Challenge (ICMI)	*China*
Feb 2018	**Scholarship**, Global Scholarship - Chonnam Nat'l Univ.	*South Korea*

Languages

English Professional proficiency

Korean Elementary proficiency

Vietnamese Native proficiency