Rahul Sharma

3740 McClintock Ave, EEB 400, Los Angeles, CA, 90007
Email: rahul.sharma@usc.edu
Mobile No.: +12132456486

ACADEMIC DETAILS

Degree	Institute	Year	GPA
PhD in Electrical and Computer Engineering	Univ. Southern California	2017-Present	3.8/4
Integrated BTech-MTech in Electrical Engineering	IIT Kanpur	2012-17	8.8/10

FIELDS OF INTEREST

Primarily interested in ***multimodal signal processing***, more inclined towards the ***visual signal***, to understand human actions and behavior in multimedia content. Furthermore, I am keenly interested in ***semi-supervised systems*** and the notion of ***weaker-than-full supervision***.

SCHOLASTIC ACHIEVEMENTS

- **Viterbi Graduate Student Fellowship (2017)**: Viterbi School of Engineering, USC.
- **Cadence Gold Medal (2017)**: Awarded by IIT Kanpur, for best master's thesis work across all departments.
- **GATE Fellowship (2016-2017):** Awarded by HRD India, towards a stipend during the master's program.
- **Merit Cum Means Scholarship (2013 - 2016)**: Awarded by IIT Kanpur to support the tuition at the institute.

EXPERIENCE

- **Internship at Trustworthy Alexa, Amazon (May'21 – Aug'21):** Worked on understanding the effects of noisy labels **i**n federated learning setting. Proposed a novel strategy to train a system in self-supervised federated setting, involving the user generated noisy dataset, and evaluated the system for the task of text classification.
  R. Sharma et. al, "*Federated Learning with Noisy User Feedback*": NACCL 2022

RELAVENT PUBLICATIONS

- **R. Sharma, S. Narayanan,** *"Audio-Visual Activity Guided Cross-Modal Identity Association for Active Speaker Detection*": Under review at Open Journal on Signal Processing
- **R. Sharma, S. Narayanan**, "*Unsupervised active speaker detection in media content using cross-modal information* ": Under review at IEEE Transactions on Image Processing
- **R. Sharma, S. Narayanan**, "*Cross modal video representations for weakly supervised active speaker localization*": IEEE Transactions on Multimedia, 2022.
- **A. Hebbar, R. Sharma,** "*Vocal tract articulatory contour detection in real-time magnetic resonance images using spatio-temporal context*": Proceedings of ICASSP, May 2020
- **R. Sharma, S. Narayanan**, "*Towards Visual Voice Activity Detection for Unconstrained Videos*": Proceedings of International Conference on Image Processing (ICIP), September 2019

CURRENT RESEARCH

Computational Media Intelligence: Developing tools to study representation and character portrayals in entertainment media, especially Hollywood movies and TV shows. My current research focuses on devising multimodal strategies to detect which character is speaking in the visual frames at any time. It includes utilizing the visual activity in the mouth region of the visible character faces along with character identity information available in faces and concurrent speech. Furthermore, my work also involves collecting all instances of each character through the video, known as the character diarization problem.

MASTER'S THESIS

Towards Multimodal Assessment of Speaker Performance in Public Speaking (2017).
We propose a computational framework for quantifying speaker performance in the context of public speaking. For this purpose, we created a database consisting of more than two thousand Technology, Entertainment, Design (TED) conference videos along with associated metadata (number of likes/dislikes, views, comments) from YouTube. We do not consider the content of the talk, analyze the speech and the visual content to capture the verbal and non-verbal behavior of the speaker. We have established baselines which can predict the performance rating with correlation coefficient 0.68.
R. Sharma, T. Guha, G. Sharma, *"Multichannel Attention Network for Analyzing Visual Behavior in Public Speaking"*: Proceedings of Winter Conference on Applications of Computer Vision, February 2018