

# Rahul Sharma

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

## ACADEMIC DETAILS

Degree	Institute	Year	GPA
PhD in Electrical and Computer Engineering	USC	2017-Present	3.9/4
Integrated BTech – MTech in Electrical Engineering	IIT Kanpur	2012-17	8.8/10

## FIELDS OF INTEREST

Primarily interested in applying **machine learning** algorithms to **Audio-Visual** signals to understand human actions, emotions and behavior. Interests further extends to video understanding and media analytics.

## 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.
- **Joint Engineering Exam (2012)**: Secured All India Rank **438 (99.9 percentile)**.
- **CBSE (2009)**: Awarded 0.1% merit certificate by **CBSE** for standing in top **0.1%** students across the nation.

## PUBLICATIONS

- **R. Sharma, S. Narayanan**, "[\*Crossmodal video representations for weakly supervised active speaker localization\*](#)": Under review at IEEE Transaction on Multimedia
- **R. Sharma, S. Narayanan**, "[\*Audio visual character profiles for detecting background characters in entertainment media\*](#)": Submitted to ICIIP 2022
- **R. Sharma, S. Narayanan**, "[\*Using Active Speaker Faces for Diarization in TV shows\*](#)": Submitted Interspeech
- **R. Sharma, S. Narayanan**, "[\*Towards Visual Voice Activity Detection for Unconstrained Videos\*](#)": Proceedings of International Conference on Image Processing, September 2019
- **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
- **R. Sharma, T. Guha**, "[\*A Trajectory Clustering Approach to Crowd Flow Segmentation in Videos\*](#)", Proceedings of International Conference on Image Processing, September 2016

## COURSE WORK @USC

- **EE503: Probability for Engineering (A)**
- **EE510: Linear Algebra for Engineering (A)**
- **EE483: Digital Signal Processing (A)**
- **EE562: Random Processes (B+)**
- **EE546: Mathematics of High Dimensions (A)**
- **CSCI677: Advanced Computer Vision (B+)**
- **CSCI534: Affective Computing (A-)**
- **CSCI544: Natural Language Processing (A)**

## CURRENT RESEARCH

- **Active Speaker Detection in Unconstrained Videos**: Working towards detecting an active speaker in each frame of the videos from Hollywood movies. We formulate it with a primary objective of visual voice activity detection (VAD) and a lateral objective of active speaker detection. We hypothesize that, to decide for VAD, the network should consider the active speakers as most salient. We exploit the interpretability of the employed 3-D convolutional network to localize the active speakers in the frames.

## 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.

*Published in WACV' 2018*