Enabling Multimodal Pervasive Computing Systems for Agriculture and Transportation Applications

Mohan M Trivedi (mtrivedi@ucsd.edu)        Rajesh M Hegde (rhegde@iitk.ac.in)
University of California at San Diego             Indian Institute of Technology Kanpur

Vision and Research Agenda:

The primary focus of this proposal is on using multiple modes (vision, audio, gestures, and text) of information and integrate (fuse) them in a sustained fashion for agriculture and environment (transportation and emission modelling) in the Indo-US context. Important questions underlying realization of practical and deployable pervasive computing systems, which this collaborative research will address include,

a) How can pervasive computing technologies be used to reduce collisions, traffic delays, fuel consumption, and emissions?
b) How can context aware driving and emission control systems be designed using an array of active sensors?
c) How can multi modal signal processing and recognition techniques be utilized to help the common folk in India to make modern technology accessible in the real sense? For e.g. utilization of internet kiosks and mobile phones although easily affordable can be very cumbersome for an Indian farmer to query the remedy to the disease his crop is effected with given the current state of the input technology used for the same. Can multi modal inputs (combination of speech, keyboard, and images) be effectively used to make this information easily accessible?
d) How can one address the larger question of effectively fusing information in different sensors (sensor fusion) and modalities (modal fusion) like vision, keyboard, location from gps devices, etc. to come up with a multi modal solution to the problem of agriculture, transportation, and emission?

Our research groups at UCSD and IIT, Kanpur have collaborated closely over the past five year period. The collaboration was initiated as a part of a NSF sponsored ITR (Information Technology Research) project focused on technologies which can be used by various first responder teams as they respond to emergency situations. After successful completion of that project, Professor Hegde joined IIT, Kanpur faculty and has established a new research laboratory to pursue research in multimodal signal processing. Professor Trivedi has recently expanded his research in the multimodal systems especially for monitoring and mitigating emissions from vehicular traffic in the urban areas. The new collaboration among our teams will allow for exchange of students, researchers on a regular basis and also for conducing two workshops (one in the US and another in India), where representatives from the application domains will participate and contribute.

Overview:

a) Transportation and Environment
Emissions studies, congestion studies and safety/security related studies have identified vehicle detection and tracking, traffic flow characterization, vehicle activity detection, vehicle type classification and criticality assessment as highly relevant information for higher level understanding of transportation related issues. In this research we propose the development and experimental evaluation of efficient computer vision, machine learning, and multi sensory fusion
algorithms to extract such needed information using a large network of video cameras and other sensory networks. The capabilities of existing sensors and traditional techniques, e.g. inductive loops, will be augmented by information obtained from video sensors and will be integrated in a hierarchical fusion framework with contextual information, such as GPS and GIS, to provide real-time fleet composition, vehicle activity information, and site analysis while overcoming individual modality shortcomings. By incorporating varied sensing and information sources, the overall robustness of transportation analyses will be directly impacted to improve highway efficiency through evaluation of congestion events such as bottlenecks and collisions, provide more complete environmental emissions studies, knowledge of infrastructure load, design of operational control strategies, and improve general road safety with early incident detection for better response. The proposed research will result in the development of algorithms and systems including those for real-time video analysis and hierarchical information assimilation which should be general enough for duplication at other sites and can be expanded to provide additional functionalities in the future.

b) Agriculture
Agricultural information access via resource constrained devices is socially and economically viable in a country like India. Providing agro commodity prices to the Indian farmer on his cell phone using keyboard and speech inputs is common place. However using multiple modalities in the form of images and voice is an interesting aspect of pervasive computing that can provide more reliable and accurate information to make technology accessible to people with limited literacy. Several pervasive computing issues will come into play in this context. One example would be with respect to how speech inputs can be handled at the device. The pervasive aspect of this research will address the question of, whether speech recognition can be done at the client (when mobile devices are used) or the server (when kiosks or desktops or used). It will also address network issues that assume importance in such scenarios. The inclusion of multiple modes of information like the keypad, stylus, speech, ambient audio and vision will bring in many pervasive aspects of research in the context of information retrieval or dissemination. Issues of robustness like the use of multiple sensors on pervasive devices and source localization prior to processing multi modal signals play an important role in this aspect. The effective fusion of multi modal information at different levels of abstraction namely, the early and late fusion will be addressed to provide more reliable retrieval of information.

**Background and Experience:**

**Mohan Trivedi** is a Professor of Electrical and Computer Engineering at the University of California in San Diego. He established the Computer Vision and Robotics Research Laboratory and LISA: Laboratory for Intelligent and Safe Automobiles promoting multidisciplinary research at UCSD. His team is pursuing research in machine and human perception, machine learning, distributed video systems, multimodal affect and gesture analysis, human-centered interfaces and intelligent driver assistance systems active safety systems for automobiles. CVRR Lab has played key role in several major research collaborative initiatives. These include an autonomous robotic team for Shinkansen track maintenance, a human-centered collision avoidance system, panoramic vision system for incident detection and also for driver assistance, vision based occupant protection system for “smart airbags. Trivedi’s team also designed and deployed the “Eagle Eyes” system on the US-Mexico border in 2006. He served on a panel dealing with the legal and technology issues of video surveillance organized by the Constitution Project in Washington DC as well as at the Computers, Freedom and Privacy Conference. Trivedi is served as an “Expert Panellist” for the Strategic Highway Research Program of the Transportation Research Board of the National Academy of Sciences. He has given over 65 Keynote/Plenary talks at major conferences. Trivedi has received the Distinguished Alumnus Award from the Utah State University, Pioneer (Technical Activities) and Meritorious Service Awards from the IEEE Computer Society. He is a co-author of a
number of papers winning “Best Papers” awards. Two of his students were awarded Best Dissertation Awards by the IEEE ITS Society (Dr. Shinko Cheng 2008 and Dr. Brendan Morris 2010). Trivedi is a Fellow of the IEEE (“for contributions to Intelligent Transportation Systems field”), IAPR (“for contributions to vision systems for situational awareness and human-centered vehicle safety”), and the SPIE (“for contributions to the optical engineering field”).

Selected Relevant Publications:
(for complete list please see http://cvrr.ucsd.edu/publications/index.html)
Brendan Morris and Mohan M. Trivedi, "Contextual Activity Visualization from Long-Term Video Observations" IEEE Intelligent Systems, 2010 (pdf).
Brendan Morris and Mohan M. Trivedi, "Learning, Modelling, and Classification of Vehicle Track Patterns from Live Video", IEEE Transactions on Intelligent Transportation Systems, Sep 2008. (pdf)

Rajesh M Hegde is an Asst. Professor and P K Kelkar Research Fellow with the Dept of Electrical Engineering at IIT Kanpur, where he leads the multi modal information processing systems lab (http://mipslab.tk). His current areas of research interest include multi media signal processing, multi microphone speech processing, speech recognition, speaker identification, pervasive multi media computing, applications of signal processing in wireless networks with specific focus on emergency response. He has earlier worked on NSF funded projects on speech processing and ICT applications for emergency response at the University of California San Diego, USA, where he was a researcher and lecturer in the Dept of Electrical and Computer Engineering between 2005 -2008. He actively publishes in IEEE signal processing society conferences and journals. He is on the review board for various international journals including IEEE Transactions, Elsevier, and Springer. He is also a member of the National working group of ITU-T (NWG-16) on multi media standards.

Selected Relevant Publications: