

With multi-perspective, multi-modal cameras

# Tracking human activity

Computer vision researchers at the University of California, San Diego, have developed new techniques to improve recognition of human activity by using cameras that operate at different wavelengths than those used in human vision. The algorithms could be of use in applications ranging from surveillance, automotive safety, smart spaces, and human-computer interfaces.



"The new systems we are developing are multi-perspective and multi-modal," said Mohan Trivedi, professor of electrical and computer engineering in UCSD's Jacobs School of Engineering. "They allow observation of a space and occupants from various viewpoints and emitted energies. The results of testing the technology

have been very encouraging."

The multi-perspective approach involves two or more cameras observing the same person from different angles. "Multi-modal" means more than one type of camera—e.g., thermal infrared, and color.

Recent results of Trivedi's research are recounted in two new papers co-authored with researchers in his Computer Vision and Robotics Research (CVRR) laboratory, an affiliate of Calit2 on the UCSD campus. The papers are published in the latest edition of a special issue of *Computer Vision and Image Understanding* devoted to "Vision Beyond Visible Spectrum."

Tracking and analyzing a person's movements are complex tasks, especially in unconstrained environments (such as inside and outside a vehicle) where the lighting is unreliable and where there is "noise" in the environment.

UCSD researchers in the smart-car lab developed a system consisting of four multi-perspective, multi-modal (i.e., both thermal infrared and color) cameras. The equipment was installed on the LISA-Q, an Infiniti Q45 carrying a number of cameras, sensors, and processors. The vehicle had earlier been used in a number of automotive computer-vision experiments. The video-based system was then tested on the road to see how well it did tracking of the driver in real time—particularly his or her head, arms, torso, and legs.

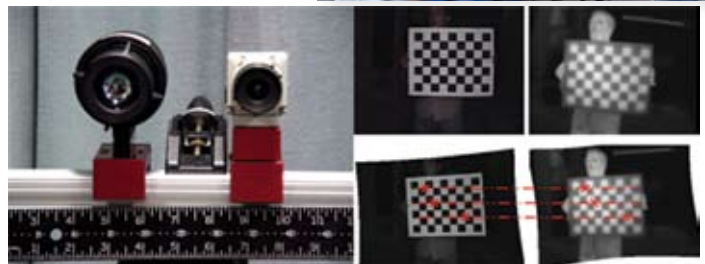
"The multi-perspective capabilities of the system help identify redundant trajectories of various parts of the human body, while its multi-modality provides robustness and reliability of feature detection and tracking," reported the scientists. The bottom line—the proposed system is effective in real-world

street driving; it tracks the driver's head and hands regardless of the level of illumination, and the system is fairly accurate in noisy outdoor driving situations.

The UCSD scientists also found a new approach to finding accurate correspondence between objects that are simultaneously seen by a stereo head with one eye sensitive to thermal infrared and a second to color wavelengths. Existing algorithms do not operate well on data representing multiple objects at multiple depths that are significant relative to the objects' distance from the camera. The new algorithm developed at UCSD offers substantial benefits, especially in instances of close-range surveillance and pedestrian detection.

"It provides robust and accurate pedestrian detection, as well as tracking and analysis for active safety systems in a vehicle, and also for operating surveillance systems on a 24/7 basis," explained Ph.D. student Steve Krotosky [right & below].

The research was mainly supported by the CVRR lab's grant from the Technical Support Working Group (TSWG), a federal, inter-agency



institution responsible for overseeing technology development to help in the fight against terrorism. TSWG also supported the Eagle Eyes system developed by the UCSD researchers, which is used by the Police Department in Eagle Pass, TX. (See Related Links below for more on the Eagle Eyes project.)

#### Research papers

- » S.Y. Cheng, S. Park, M.M. Trivedi, Multi-spectral and Multi-perspective Video Arrays for Driver Body Tracking and Activity Analysis, *Computer Vision and Image Understanding* (2006).
- » S.J. Krotosky, M.M. Trivedi, Mutual Information Based Registration of Multimodal Stereo Videos for Person Tracking, *Computer Vision and Image Understanding* (2006).