A wireless network for non-obtrusive continuous assessment of astronaut fatigue

George Zouridakis, Edgar Gabriel, Ioannis Kakadiaris, Shishir Shah, Xiaojing Yuan, and Rong Zheng

ABSTRACT—UH researchers have developed a non-obtrusive tracking system based on a wireless network of smart cameras and processing components that can achieve continuous, multi-object identification for purposes of assessing astronaut fatigue. The framework includes direct camera-to-camera communication and software components that respect the memory and processing limitations of smart cameras. The system utilizes 3-D-aided 2-D face recognition that has proved reliable for different poses and under varying light conditions. UH researchers also have developed a means of fast facial feature extraction via new methods of detecting regions of interest in the foreground of an image, before any features are extracted.

INTRODUCTION

Fatigue, whether physical or mental, is known to affect the cognitive functions and overall performance of humans working on Earth as well as in space. The long-term objective of this project is to develop a non-obtrusive, real-time, continuous monitoring system for assessing human fatigue and stress in an effort to monitor the cognitive function of humans working in space. The system can continuously assess fatigue via real-time analysis of brain activity and of the facial expressions of subjects working in a confined space, such as the International Space Station. The system relies on two main modules:

1. A non-obtrusive, wireless tracking system that can continuously track all subjects present in a confined space, monitor brain activity via wireless sensors placed on each subject's head, and analyze facial expressions continuously via a network of wireless cameras; and
2. A decision-making system that analyzes a subject's fatigue indicators and makes recommendations to fatigued subjects who operate critical equipment or make critical decisions. Activities performed by NASA astronauts can be mentally and physically demanding and may last for hours. Therefore, continuous assessment of the crew's cognitive abilities and the effects of fatigue are of paramount importance.

METHODOLOGY

The main effort of this project has focused on the development of a visual tracking framework, a communications framework, subject identification, and feature extraction Algorithms.

Visual tracking framework

To provide distributed processing on individual cameras and integration with the underlying communications framework, we have developed a tracking framework that leverages a modular, component-based architecture, as shown in Figure 1. The architecture realizes two groups of components: camera components that manage and communicate knowledge about objects and the scene, and processing components that provide image and video processing capabilities to achieve multi-object tracking and identification.

Processing components encapsulate several processing steps necessary to achieve robust tracking and identification of a target object. The first processing step is that of background modeling. The background model is learned online on a continual basis to account for environmental and scene variability. This is accomplished by probabilistic modeling of the scene using a combination of Gaussian models. The unrecognized foreground is segmented and refined to detect the target object. Since it is not feasible to calibrate cameras for each deployment, we also incorporate processing components to estimate the ground plane and localize the camera coordinates. Ground plane is recovered based on estimation of the vanishing geometry through each camera view.

The object identified from foreground segmentation is further analyzed under constraints of the ground plane and a code book of object models. Tracking is achieved based on a prediction filter that uses several features obtained from the segmented object. These include multiple measures of object appearance and provide robustness in tracking across occlusions and scene variability. At each step, relevant object and background information is communicated with camera components. A database stores all information relevant to each object and broadcasts it when an object exits the field of view of a camera.

All local trajectory and appearance data are collected at a central processing unit that integrates information across multiple cameras to generate global motion tracks for each object based on interpolation of local trajectories across local camera coordinates. In our preliminary work, we have implemented

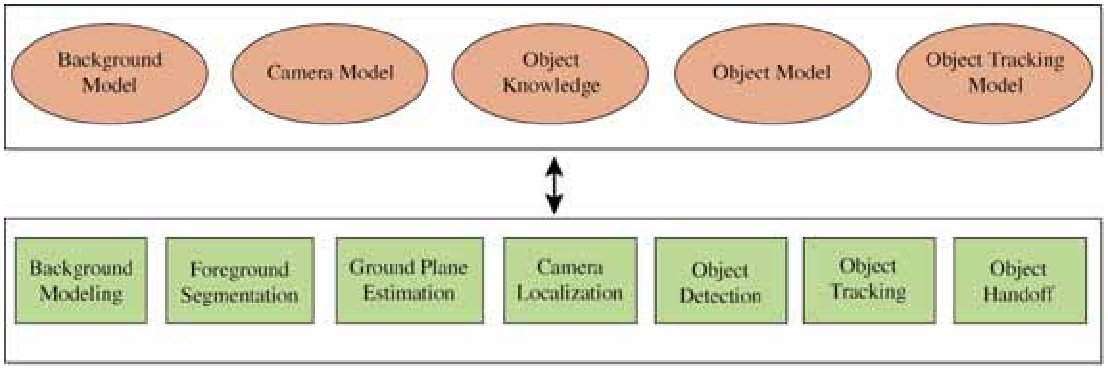

Figure 1. Component-based framework for object tracking

and tested two processing components: background modeling and single-camera object tracking. Typical challenges faced in background extraction are changes in illumination and the merging of new foreground objects into the background over time.

We employed a probabilistic model capable of predicting the association of each individual pixel in the image. The background model comprises *k* Gaussian distributions for each pixel of the image $I(x,y)$. The background modeling process employs a two-stage approach for updating background pixels based on statistical properties and temporal persistence and recurrence frequency of each pixel. Meanshift has been successfully applied to track an image region. We identify image regions based on background subtraction using the background model generated from the background modeling component. Meanshift is a procedure used for updating the probability density of an image region by computing the gradient of the density and following the gradient to find the local maximum within a few steps. When applied to the problem of target tracking, the target detected in the first frame of an image sequence is used to initialize the probability density. The maximum of the density is identified as the vehicle position, and the position is searched for in the following frames using simplified meanshift equations.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i \cdot P(x_i, y_i)}{\sum_{i=1}^{n} P(x_i, y_i)}, \bar{y} = \frac{\sum_{i=1}^{n} y_i \cdot P(x_i, y_i)}{\sum_{i=1}^{n} P(x_i, y_i)}, \quad (1)$$

where $(\bar{x},\bar{y})$ is the new vehicle position and (x_i,y_i) is the image coordinate of each pixel within the foreground object region. We implemented this simple approach to evaluate the tracking performance and understand its limitations.

Communications framework

Our work focused on the overall design of the communications and process management framework, which is required in order to build a network of smart cameras. The main requirements for such a framework include direct, camera-to-camera communication; lightweight software components that respect the limitations in main memory and processing power of smart cameras; and definition and development of user level API functions, which hide much of the complexity of the distributed smart camera network from the application.

In addition to the overall design, significant progress has been made in the development of individual software components, including (1) a communication module enabling synchronous and asynchronous data exchange between processes, and (2) a launcher application for starting and monitoring a given number of processes/applications on remote cameras. Furthermore, several experimental components have been developed in order to evaluate the implications and the usability of certain techniques. The list of experimental components includes (1) a framework for dynamically detecting, loading and unloading shared libraries, and (2) a functionality detection component, which is capable of parsing the source code of a given application, report back to the end-user a list of library functions used for a previously registered library (e.g., MPI), and automatically generate a header with appropriate pre-processor directives, which can be used to generate a personalized version of the according library. This component is the major building block for automatic optimization to further reduce the memory footprint of the mobile component architecture itself.

Bandwidth provisioning in multi-hop wireless networks

Robust channel assignment for extreme environments. We investigated the problem of link-level resource provisioning in multi-radio, multi-channel wireless networks. A robust channel assignment algorithm was proposed that explicitly takes into consideration link-level demands. Our solution utilizes the dual decomposition and Gibbs sampler techniques, which effectively incorporate the channel allocation (combinatorial constraints), network resource, and link-level demand (continuous constraints) in a single optimization framework. A key advantage of the proposed scheme is the robustness to channel variability and external interference sources.

Starvation modeling and identification of multi-hop 802.11 wireless networks. With the growing number of spontaneously deployed WiFi hotspots and home networks, end-users often experience significant performance degradation or even starvation. However, we observed that tuning individual sys-

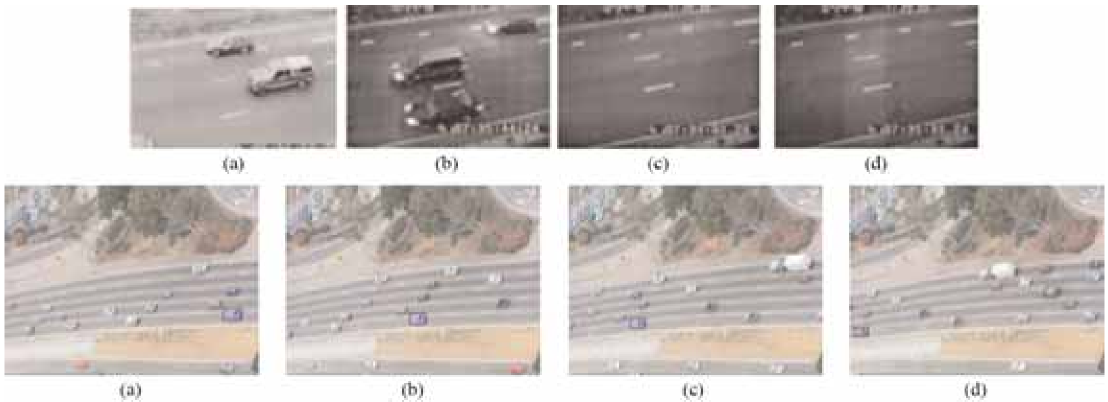

Figure 2. (Top) Background modeling in the presence of illumination changes: (a) initial background model, (b) frame in which the ambient lighting has changed, (c) the recovered background model within 4 seconds, and (d) the background model after 8 seconds. (Bottom) Tracking of a single vehicle across an 8-second video represented in four frames at 2-second intervals.

tem parameters, such as channel, Tx power, carrier sense threshold, and transmit rate, is insufficient and in some cases may lead to starvation. We have developed a comprehensive analytical model to characterize throughput of individual flows in dense IEEE 802.11 wireless community networks. The proposed model subsumes existing models for the IEEE 802.11 MAC in multi-hop wireless networks by accounting for heterogeneous transmission power levels and carrier sense thresholds, as well as various sources of packet collisions. Based on the insight from the theoretical analysis and simulation results, we propose a simple identification mechanism that determines the sources of starvation using local measurements. Both the theoretical model and the identification algorithm are validated using ns-2 simulations.

2-D face recognition

It is well established that even the leading 2-D face recognition methods (including commercial products) suffer a significant performance penalty when employed in nonfrontal images with lighting variations. We have developed a novel method for 3-D-aided 2-D face recognition under large pose and illumination changes. Specifically, our method employs 3-D data for enrollment and 2-D data for verification/identification. During the enrollment process, the subject's raw 3-D data are fit using a subdivision-based deformable model. During authentication, only a single 2-D image is necessary. The subject-specific (in the gallery), fitted, annotated face model is used to lift the texture of the face from the 2-D probe image, and a bidirectional relighting algorithm is employed to change the illumination of the gallery texture to match that of the probe. Then, the relit texture is compared to the gallery texture using the complex wavelet structural similarity index metric. We showed that this approach yields significantly better recognition rates, with an error rate that is less than half of the best-performing commercially available 2-D face recognition software.

Feature extraction

To assess the cognitive state of a subject and the effects of fatigue, a variety of features can be extracted from multichannel EEG signals and from facial and video images. However, the features tend to be washed out if they are extracted from the whole image, which usually includes a great deal of background information. To obviate this problem, we developed new methods for region-of-interest detection and identification before any features are extracted. In particular, we developed a generic framework whereby we transformed the segmentation problem into a numerical optimization problem and used evolution strategies with an ellipsoid search structure. Additionally, new feature extraction algorithms are being developed to solve the problem of non-existing training data sets for overlapping features.

RESULTS

Using several videos, we test-tracked vehicular traffic on U.S. Highway 101 in the Universal City area of Los Angeles, California. Each component was implemented in Matlab and video-captured at 30 frames per second. The image size for all the videos was 640 × 480 pixels. The background modeling component resulted in reliable estimation of the background model. Figure 2 (top) shows results for one video sequence. The dataset represents vehicle trajectory data on a 2,100-foot, six-lane segment of southbound U.S. 101. We restricted our tracking initially to only one vehicle blob and tested the scalability of the approach to up to six vehicles. The first two seconds of the video were used to recover the background model.

Figure 2 (bottom) shows the frames with single vehicle tracking at 2, 4, 6, and 8 seconds. Overall, we were able to achieve single-vehicle blob tracking at 15 frames per second as long as the vehicle blobs did not show any occlusion. Occlusions resulted in loss and failure to recover the object because of similar features in multiple objects. The appearance model used for tracking was the image color distribution within the foreground region, and in all videos analyzed, the typical size ranged from 30-50 KB. This would be the information recorded in the database that would then be broadcast to neighboring cameras when the target exits one camera's field of view.

Additionally, the proposed channel allocation algorithm was implemented in a distributed manner with limited local information exchange among neighboring nodes. We have evaluated our scheme using traces collected from a wireless mesh test bed. Our experiments show that the proposed channel assignment algorithm is superior to existing schemes in providing larger noise margin and reducing outage probability.

Finally, since various stages of the project underwent simultaneous development by different team investigators, we decided to apply our feature selection methods to a set of skin lesion images used in another of the team's projects. We found that the evolution strategy-based algorithm was more robust with regard to artifacts and noise and achieved better performance than other existing methods. The computational savings were also significant, considering that typical images have 60-90 percent of background.

CONCLUSIONS

A tracking system based on a wireless network of smart cameras has been developed that can achieve continuous multi-object identification and tracking. The underlying communications framework addresses issues of limited memory and processing power of smart cameras and the problem of link-level resource provisioning in multi-radio, multi-channel wireless networks. Subject identification and facial feature extraction rely on a novel, robust method for 2-D face recognition and on algorithms for detecting appropriate regions of interest in the foreground before any features are extracted. This tracking system can be used to detect astronaut fatigue, which can play a crucial role in long-term missions in space.

PUBLICATIONS

Hua, C. and Zheng, R. Starvation modeling and identification in dense WLAN networks. *Proceedings* of the 27th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM). Phoenix (April 15–17, 2008).

Hua, C., Zheng, R., and Wei, S. Robust channel assignment for link-level resource provisioning in multi-radio multi-channel wireless networks. UH-CS Technical Report. Houston: University of Houston (2007).

Kakadiaris, I., Toderici, G., Passalis, G., and Theoharis, T. 3D-aided face recognition. *Computer Vision and Pattern Recognition* (2008) (*submitted*).

Omiwade, S., Zheng, R., and Hua, C. Butterflies in the mesh: Lightweight localized wireless network coding. *Proceedings* of the Fourth Workshop on Network Coding, Theory, and Applications (NetCod). Hong Kong, 1-6 (Jan. 3–4, 2008).

Situ, N., Yuan, X., and Zouridakis, G. Segmentation of skin lesion images using a narrow band graph partitioning method. *Proceedings* of the Houston Society of Engineering in Medicine and Biology, Houston, (Feb. 7–8, 2008).

Situ, N., Yuan, X., Zouridakis, G., and Mullani, N. Automatic segmentation of skin lesion images using evolutionary strategy, *Proceedings* of IEEE International Conference on Imaging Processing. San Antonio, **6**, VI - 277-VI - 280 (Sept. 16–19, 2007).

Wu, Y., Chen, J., Yuan, X., and Zouridakis, G. A programmable DSP development platform for automated detection of melanoma. IEEE *Transactions on Consumer Electronics* (2007) (*submitted*).

Wu, Y., Stotzer, E., Chen, J., Yuan, X., Mullani, N., and Zouridakis, G. A programmable DSP development platform for automated detection of melanoma. *Proceedings* of the IEEE Dallas Engineering in Medicine and Biology Workshop, Dallas, 106-109 (Nov. 11–12, 2007).

Yuan, X., Situ, N., and Zouridakis, G. A narrow band graph partitioning method for skin lesion segmentation. *International J. Pattern Recognition* (2007) (*submitted*).

Yuan, X., Situ, N., and Zouridakis, G. Automatic segmentation of skin lesion images using evolution strategies. *Biomedical Signal Processing and Control* (*in press*).

PRESENTATIONS

Zheng, R. Modeling and resource management in wireless networks. Texas State University-San Marcos, San Marcos, TX, Jan. 25 (2008).

PROPOSALS

Gabriel, E. and Co-PI: Shah, S. SRS: A modular adaptive software infrastructure for runtime optimization of real world software systems, National Science Foundation, $460,431 (*pending*).

Shah, S., Co-PIs: Gabriel, E., Garbey, M., and Zheng, R. DURIP: Heterogeneous smart camera networks for collaborative missions, U.S. Army Research Office, $278,656 (*pending*).

Shah, S., Co-PIs: Gabriel, E., and Zheng, R. CyberSystems: A wireless network of smart cameras for real-time event analysis, National Science Foundation, $547,722 (*pending*).

Zouridakis, G., Co-PIs: Chen, J., and Yuan, X. Integration of digital dermoscopy and hardware design for the development of a handheld skin cancer screening device, National Science Foundation, $448,974 (*pending*).