

on the framing structure design: a lower density with a sensor spacing of 20m (still reasonable for seismic acquisitions) can make WiMedia feasible for WGN applications with minimal modifications. On the other hand, higher geophone densities, with a sensor spacing of less than 11m, require further MAC modifications. More details are given at the WisyGeo Web site.

**Links:**

Land seismic exploration:  
<http://www.oilfieldreview.com/>

Wireless Geophone Network:  
<http://www.wisygeo.com>  
<http://scitation.aip.org/tle>

WiMedia:  
<http://www.wimedia.org>

**Please contact:**

Stefano Savazzi, Umberto Spagnolini,  
 DEI, Politecnico di Milano, Italy  
 E-mail: [savazzi@elet.polimi.it](mailto:savazzi@elet.polimi.it),  
[uspagnolini@elet.polimi.it](mailto:uspagnolini@elet.polimi.it)

Vittorio Rampa  
 IEIIT-CNR, Italy  
 E-mail: [rampa@elet.polimi.it](mailto:rampa@elet.polimi.it)

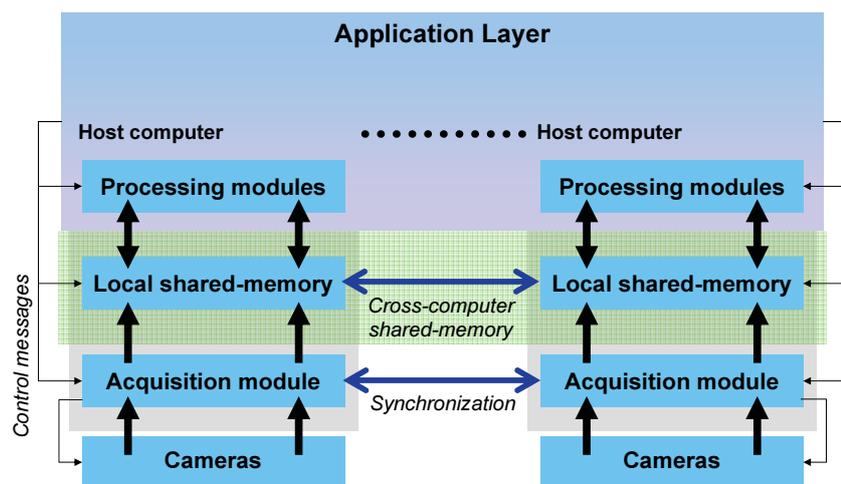
## A Software Platform for the Acquisition and Online Processing of Images in a Camera Network

by Thomas Sarmis, Xenophon Zabulis and Antonis A. Argyros

*Applications related to vision-based monitoring of spaces and to the visual understanding of human behaviour, require the synchronous imaging of a scene from multiple views. We present the design and implementation of a software platform that enables synchronous acquisition of images from a camera network and supports their distribution across computers. Seamless and online delivery of acquired data to multiple distributed processes facilitates the development of parallel applications. As a case study, we describe the use of the platform in a vision system targeted at unobtrusive human-computer interaction.*

Camera networks are increasingly employed in a wide range of Computer Vision applications, from modelling and interpretation of individual human behaviour to the surveillance of wide areas. In most cases, the evidence gathered by individual cameras is fused together, making the synchronization of acquired images a crucial task. Cameras are typically hosted on multiple computers in order to accommodate the large number of acquired images and provide the computational resources required for their processing. In the application layer, vision processing is thus supported by multiple processing nodes (CPUs, GPUs or DSPs). The proposed platform is able to handle the considerable technical complexity involved in the synchronous acquisition of images and the allocation of processes to nodes. Figure 1 illustrates an overview of the proposed and implemented architecture.

The platform integrates the hardware and device-dependent components employed in synchronous multi-camera image and video acquisition. Pertinent functionalities become available to the applications programmer through conventional library calls. These include online control of sensor-configuration parameters, online delivery of synchronized data to multiple distributed processing nodes, and support for the integration and scheduling of third-party vision algorithms.



*Figure 1: Platform architecture for camera network supported by multiple computers. Central sensor control is provided by a message-passing communication infrastructure. Acquisition modules place the images on a synchronized shared-memory space, making them available to multiple processing nodes across computers. Processing modules read these data and perform parallel execution of computer vision algorithms. The same modules also have access to this space, allowing them to synchronously view intermediate computation results.*

System modules can communicate in two modes. Communication through message-passing addresses control messages to targeted or multicast recipients. The diversity of communicated information types is accommodated by data-structure serialization. Communication through shared-memory spaces provides visual data or intermediate computation results to the nodes of the host or of multiple computers. The large

bandwidth requirements imposed by image transmission are accommodated by a Direct Memory Access channel to a local shared-memory space. For cross-computer availability of images, memory spaces are unified over a network link. The latency introduced by this link is compensated for by notification of nodes, regarding the partial or total availability of a synchronized image set. In this way, per-frame syn-

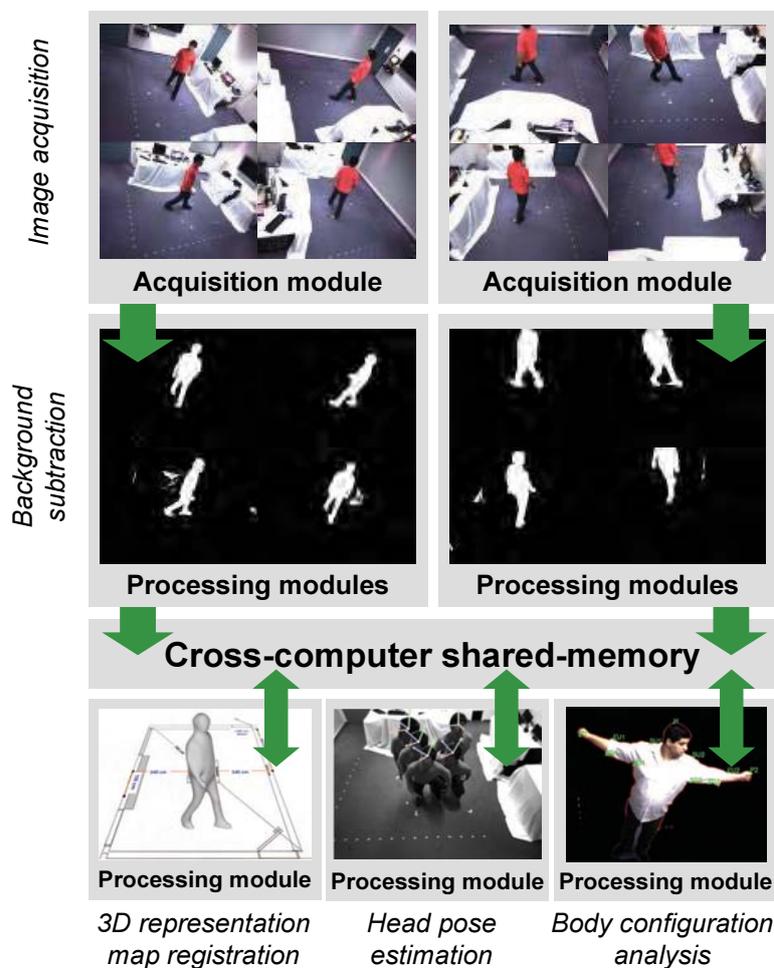


Figure 2: Person silhouettes in synchronous images are segmented in parallel. Individual processes fuse them into a registered 3D representation of the person, recognize coarse gestures as expressed by body configuration, and estimate the spatial direction in which the person's head is facing.

chronization of modules is achieved, but at the same time, processing of partially available input is also supported. Shared-memory spaces across processing nodes are essential, as large data capacity and frequent input rate demand the parallelization and pipelining of operations.

Acquisition modules encapsulate the complexity of sensor-specific, synchronization, and shared-memory configurations. Online sensor configuration and command is implemented through message-passing, while image transmission utilizes shared-memory communication. A range of off-the-shelf sensor types is supported through an extensible repository of device-specific wrappers. To facilitate testing of applications, input may be prerecorded.

Processing modules run vision algorithms that are transparent to the computer and provide access to images and intermediate computation results. During the applications development stage, an Application Programming Interface enables synchronization and message coordination. Articulated application

development is facilitated by support for 'chaining' of processes.

Being in the format of a binary library, this platform can be invoked, independent of the programming language used. As an additional utility, the developed platform provides a GUI for the control of generic camera networks and the recording of image sequences. Forthcoming extensions involve additional capabilities for cooperation with middleware infrastructures in systems where vision is integrated with other sensory modalities (aural, tactile etc).

The platform is currently employed in the development of a vision system (illustrated in Figure 2), targeting unobtrusive and natural user interaction. The development of this system is part of a broader project funded internally at FORTH-ICS on Ambient Intelligence (AmI) environments. The system employs multiple cameras that jointly image a wide room. Two computers host eight cameras and a dedicated bus for their cross-computer synchronization, and utilize a LAN connection for communication. Upon image acquisition, a sequence of image processing

and analysis operations is performed in parallel on each image, to detect the presence of humans through background subtraction in the acquired images. Using the shared memory across computers, segmentation results are fused into a 3D volumetric representation of the person and registered to a map of the room. Two other processes run in parallel and access the same data to recognize the configuration of the person's body and estimate the pose of the person's head. The utilization of the proposed platform facilitates the modular development of such applications, improves the reusability of algorithms and components and reduces substantially the required development time.

This work has been partially supported by the FORTH-ICS RTD programme 'AmI: Ambient Intelligence Environments'.

**Link:**

<http://www.ics.forth.gr/cvrl/miap/doku.php?id=intro>

**Please contact:**

Xenophon Zabulis

FORTH-ICS, Greece

E-mail: [zabulis@ics.forth.gr](mailto:zabulis@ics.forth.gr)