Distributed Architecture for Efficient Indoor Localization and Orientation

Félix J. Villanueva, Julio Dondo Gazzano, David Villa, David Vallejo, Cesar Mora,
Carlos Gonzalez Morcillo, Juan Carlos López
School of Computer Science
University of Castilla-La Mancha, Ciudad Real, Spain
Email: felix.villanueva@uclm.es

***Abstract*—In last decade, indoor GPS-like navigation has devoted attention from research community since is a key point for advanced services. Assistance, direct marketing, localization, etc. are examples of services that an appropiate navigation infrastructure could enable. With the evolution of cameras integrated in consumer electronics devices (mobile phones, notebooks, etc.) a new method, based on video streamming analisys gotted from mobile device videocamera, can provide us with the two necessary features for a sucessfull navigation experience, localization and orientation. In this paper, we present an architecture for videocamera streamming analisys joining existing algorithms for providing, in a robust way, to the consumer electronic device with its orientation and localization.**

I. Introduction

Indoor localization of users enable a great variety of services in order to help people to find the way to a meeting, to find a specific shop, to find the office for an administrative task, to know position of other users, etc. There are a great variety of methods and technologies which have been used for indoor localization, RFID tags, Wifi, Bluetooth, etc. The wireless technologies use or RSSI methods, with low accuracy and without orientation information or Time of Arrival and/or Angle of Arrival methods with special and costly hardware.

We think that a promising approach is the analysis of video streaming captured by the consumer device owned by the user. This analysis try to recognize interesting points in the video streaming and infers the position of the device who is sending the video. Obviously the environment has to be previously parametrized in order to lookup for patterns (some works use visual tags).

The solution we propose in this paper is based on the grid computing paradigm, with the purpose of integrating heterogeneous processing devices (partially reconfigurable or not) under the umbrella of the distributed object paradigm to make efficient implementation of Indoor localization and orientation services.

We model a multi-user architecture for include in user devices a generic application which send video streaming to the GPU/FPGA Grid getting its position and orientation. The application running in the user's mobile device is generic and can be used for position and navigation in any building.

We have a set of GPU/FPGA (dimensionated according to the environment and number of posible users) in which we previously store a set of interesting points according to the tracking algorithms. The user takes, for example, its notebook starting to send the streaming of the camera of its device to the GPU/FPGA grid. According to the processing of the streaming video, the GPU/FPGA grid compares with previously store information. Finally, the GPU/FPGA grid sends the calculated position and orientation to the device.

This works is embedded in a global architecture [1] which provides support to handicapped people. The current architecture is an autonomus location provider of that global information system.

II. High performing processing architecture

Image processing is normally decomposed in a processing pipeline where a set of different operations are applied over a stream of data with the possibility of having multiple inputs and outputs. FPGAs are specially well suited for this kind of process acceleration, in special when the type and size of data is not the standard in general purpose computing. On the other side parallelism can be exploited to the last consequence, not only for fine grain repetitive operations applied to a pixel neighborhood, but also for the global concurrent processing of several images at a time.

In our approach heterogeneous resources composed of GPU and FPGAs are used to accelerate part of the tracking algorithms of the indoor localization process. The main target of the proposed architecture is to provide a fast, energy efficient and escalable product involving different technologies. FPGAs offer the possibility of implementing tracking algorithms in two different fashions. One enterely software, using embedded processor such as Power PC, microblaze or Cortex 9 ARM; or a mixed approach accelerating some critical part or time demmanding part of the tracking algorithm. In our case we use the second approach. In this approach biggest time consuming parts of the different algorithms used for tracking methods are accelerated through the design of specific hardware modules, and implemented using the partial reconfiguration capability of the FPGAs. Each hardware module is implemented in one reconfigurable area (RA) (see Fig 1).

The design is formed by two main parts: an static one that have the control of the reconfiguration process, is composed of a reconfiguration controller plus a the reconfiguration engine,

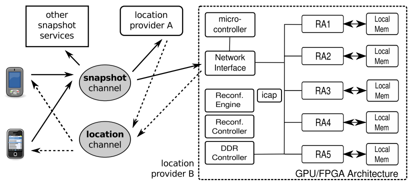

Fig. 1. Architecture overview

both implemented completely in hardware. In this way we obtain very short partial reconfiguration time. Besides, a network interface is provided to have access to the outside of the FPGA. This architecture allows to instantiate or replicate modules according to application demand when computational needs require, reducing power consumption. ARToolkit and PTAMM tracking methods were accelerated in our design. In the former the Binarization, Labeling and Pattern extraction subalgorithms were accellerated in hardware modules. In the latter the Features from Accelerated Segment Test (FAST) corner detection algorithm [2] was implemented using pattern compression. The hardware module designed to accelerate binarization process is formed by a component that perform the sum of the red, green, and blue values of each pixel, a component that perform the comparison with the threshold value, and a memory controller that takes data from DDR3 memory. This memory is used as a buffer for images to binarize. The image result is also stored in this memory. To accelerate Labeling sub algorithm a hardware module was developed using several blockrams in order to save data to be processed. The information in these blockram includes a row of binarized pixels to be processed taken from binarized images from DDR3 memory, the history buffer, the auxiliary values such as area, position, and the list of connected components. The processor sends information to the hardware labeling module in several format depending on the capacity of the blockram. All these modules can be replicated in different dinamically reconfigurable areas depending on the amount of information to be processed.

III. Communication infrastructure

The communication model is based on message events. The video sources (on smartphones) provide captured images tagged with timestamp and unique node identifiers. These *snapthot messages* are sent to a well-know specific event channel called *location-snapshot*. Any entity in the system is able to subscribe to it and receive these events.

Usually, the subscribers of this channel are *location providers* in turn. Location providers sent location and orientation information of especific physical entities in the environment. It may be different kind of location providers depending on their input information, but all of them send exactly messages with same format.

As snapshot messages, location events are sent to event channels too. This way the information can reach different consumers for different purposes. In summary, mobile nodes are snapshot publishers and location subscribers, but not limited to their own *queries*. That event-based model has many advantages respect to a RMI synchronous invocation model as, for example, a simpler programming models for providers, non blocking invocations, stateless location providers, decoupling publishers and subscribers, etc.

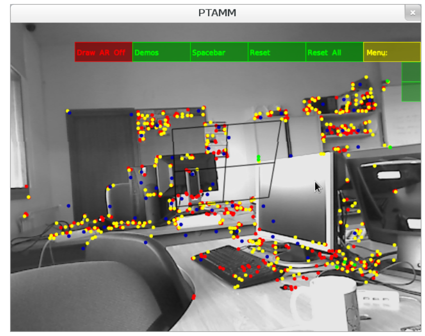

Fig. 2. PTAM interesting points detected

Snapshot messages are composed by three parameters: an image of 640x480 pixels and 24 bits color depth, a globally unique identifier for the device capturing the image and an time-stamp. Note that it is not required an accurate global clock reference due to this field is mainly relevant for the issuer device.

Location algorithms inputs are single images in all cases. Mobile nodes must produce images in the appropriate rate to ensure correct location feedback. Around 3-5 snapshots/location events by second must be enough for pedestrians. Note that same snapshot may be processed by several algorithms at different computing nodes. Depending on algorithm complexity, node performance and network load the location result may reach to mobile nodes in an arbitrary order, however this point should not be a problem at all.

Location messages are based on the MLP standard. It uses MLP::Position that includes an identifier for the provider, an identifier for the information source (the smartphone in this case), the timestamp and a shape due to just a point is not realistic.

IV. Conclusion

Our GPU/FPGA based architecture enable an efficient multi-user location and orientation position provider. In figure 2 we can see points detected using PTAM algorithm with our architecture. Next steps are stress test in order to model number of users vs necessary resources.

Acknowledgment: This research was supported by the Spanish Ministry of Economy under the project DREAMS (TEC2011-28666-C04-03) and by the Catedra Indra UCLM.

References

[1] Villanueva, F.J.; Martinez, M.A.; Villa, D.; Gonzalez, C.; Lopez, J.C.; , *Elcano: Multimodal indoor navigation infrastructure for disabled people* Consumer Electronics (ICCE), IEEE International Conf. on, pp.611-612, 2011

[2] Rosten, Edward and Drummond, Tom, *Machine learning for high-speed corner detection*, Proceedings of the 9th European conference on Computer Vision, ECCV'06, pp 430-443, 2006.