

Limits of a Distributed Intelligent Networked Device in the Intelligence Space

Gyula Max, Peter Szemes

Budapest University of Technology and Economics, H-1521, Budapest, Po. Box. 91. HUNGARY, Tel: +36 1 463 2870, fax: +36 1 463 2871, e-mail: max@aut.bme.hu, Szemes@hlab.iis.u-tokyo.ac.jp

Abstract: The paper analyses an existing DIND (Distributed Intelligent Networked Device) in an Intelligent Space (iSpace), which has ubiquitous sensory intelligence including sensors, cameras, microphones, haptic devices (for physical contact) and actuators with ubiquitous computing background. Devices use high speed network communication to flow information among them. The various devices are made for welfare support. They communicate to each other autonomously using ubiquitous computing intelligence. Intelligent Space can guide and protect humans in a crowded environment by the aid of the devices supported by the DIND. The paper tries to find the boundaries of the the recent Intelligence System to map the possibilities.

1 Brief History of the Intelligent Space

Hashimoto Lab. in University of Tokyo has proposed 'Intelligent Space' since 1996 [1]. At the beginning it consisted of two sets of vision cameras and computers with a home made 3D tracking software, this was written in C and tcl/tk under Linux. Later, a large-sized video projector (100 inches) was added to the Intelligent Space as an actuator. Mobile robots were located in the Intelligent Space for supporting people as well as for being supported. Vision cameras and computers sets were arranged around an entire room and it changed into the Intelligent Space. Conventionally, there is a trend to increase the intelligence of a robot operating in a limited area. The Intelligent Space concept is the opposite of this trend. The surrounding space has sensors and intelligence instead of the robot. A robot without any sensor or own intelligence can operate in an Intelligent Space. In the conventional solution the robot measures, calculates and decides. The heart of the iSpace concept is that the robots must not measure, calculate or make decision. They just carry out, execute commands getting information from the distributed devices called Ubiquitous Sensory Intelligence which is realised by Distributed Intelligent Networked Devices (DIND).

The Intelligent Space consists of humans not only sensors cameras or robots. In the Intelligent Space DINDs monitor the space, achieve data and share them through the network. Since robots in the iSpace are equipped with wireless network devices, DINDs and robots together organize a network.

The basic concept of Intelligent Space has extended with its development. The iSpace is a system for supporting people in it. Events, which happen in it, are understood. However, to support people physically, the intelligent space needs robots to handle real objects. Mobile robots become physical agents of the Intelligent Space and they execute tasks in the physical domain to support people in the space. Task includes movement of objects, providing help to aged or disabled persons etc. Thus, the Intelligent Space is an environmental system, which supports people in it electrically and physically. Another interesting application here is that the room can serve as a high level, context sensitive interface to robots. The Intelligent Space is a platform to which desultory technologies are installed.

2 Basic Elements of Ubiquitous Sensory Intelligence

In the Fig.2-1. three interesting elements of the current Intelligent Space with Ubiquitous Sensory Intelligence are selected and briefly described:

- Distributed Intelligent Network Device (DIND)
- Virtual Room
- Ubiquitous Human Machine Interface (UHMI)

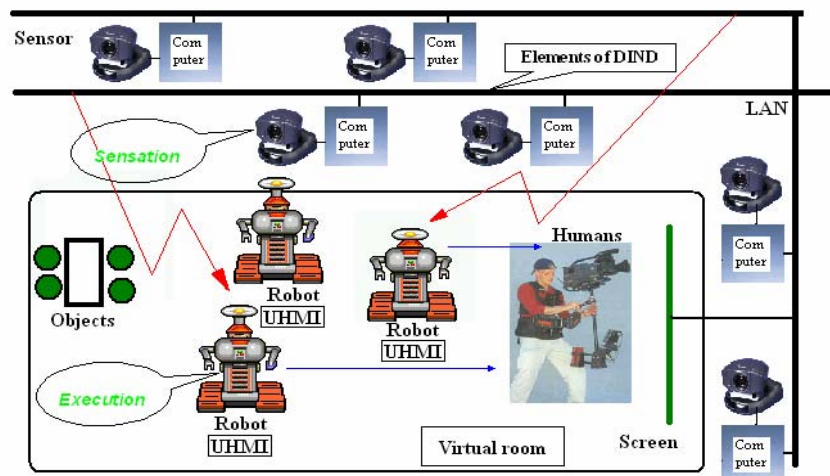


Fig.2-1. Basic Elements of Ubiquitous Sensory Intelligence

2.1 Distributed Intelligent Network Device

We can use as a definition: A space becomes intelligent, when Distributed Intelligent Network Devices (DINDs) are installed in it [2]. DIND is very fundamental element of the Intelligent Space. It consists of three basic elements:

- sensors: camera with microphone
- processor: computer
- communication device: LAN

DIND can communicate with other DINDs or robots through the network. DIND uses these elements to achieve four functions:

- the sensor monitors the dynamic environment, which contains people and robots
- the processor deals with sensed data and makes decisions
- the LAN organize communication among the elements
- DIND can communicate with other DINDs or robots through the network.

In actual system where number of the sensors are above 20, six Sony EVI D30 pan-tilt CCD camera and general bt848 based image capture board is adopted as a sensor [3]. For the processor, industrial standard Pentium III 500MHz PC is used and general 100baseT LAN card is used as a network device. Robots are able to use resources of DINDs as their own parts. However, robots with their own sensors may be considered mobile DINDs.

2.2 Virtual Room

The aim of Virtual Room (VR) research project is to recreate an environment of a physical experimental space Fig.2-2. for studying different motion control and vision algorithms for a given robot before real world implementation.

The room currently contains the following objects:

- Passive objects: Desks, chairs
- Active objects: Three robot agents
- Sensors: CCD cameras
- Actuators: Large Screen

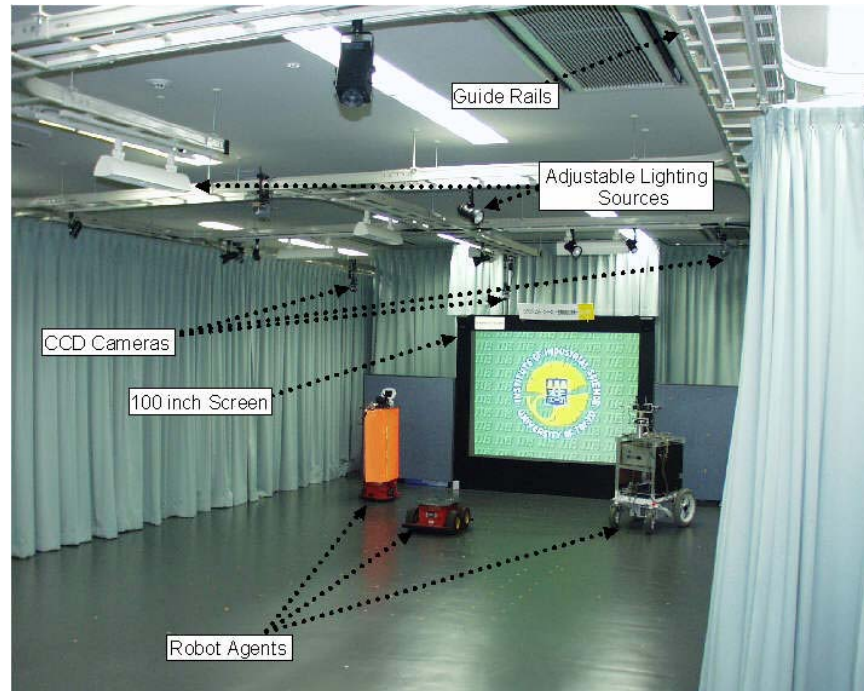


Fig. 2-2. Physical realization of the Virtual Room

2.3 Ubiquitous Human Machine Interface

There are three mobile robots in the current Intelligent Space. The most interesting one is a special mobile human-machine interface [4]. There are three basic communication channels, which the people use in daily conversation: audio, visual, and haptic. All three communication-channels are implemented on the UHMI. The human user, who is in the Intelligent Space has an aim in his mind which he wants to realize using different type of commands. Some commands are associated with certain parts of the human body. UHMI has special devices to make connections with certain part of the human body. A video Camera and a TV Screen are mounted on the UHMI for visual communication. Speaker and microphone realize the audio communication. Haptic Interface is mounted on the robot to realize a physical connection. The UHMI can be seen on Fig.2-3. This UHMI is able to move to the user or it can guide him. A very special application could be to guidance a blind or a deaf people.

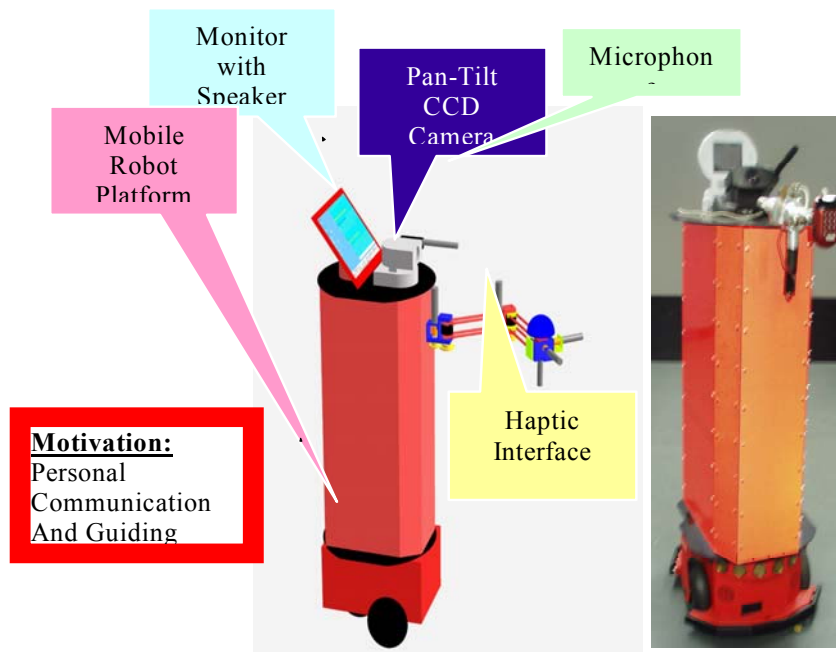


Fig. 2-3. An Ubiquitous Human Machine Interface

3 What Can Be Done In Intelligent Space?

3.1 3D Positioning of Human

Our aim is to support humans in our Intelligent Space. In order to support them first iSpace must recognize them. Recognition of a human is done in two steps [5]:

- the area or shape of a human is separated from the background (Fig.3-1.)
- features of the human as head, hands, feet, eyes etc. are located (Fig.3-1.)

Taking the images of three pairs of cameras (see Fig.2-1 and Fig.2-2.), the 3D position of the human can be calculated. The scanned areas of each parallel camera pairs are overlapped. To calculate 3D from several camera views point correspondences are needed. To establish these correspondences directly from the shape of the human is difficult. Instead of it, the head and hands of the human beings are found first and their centres are used for matching.

A second motivation to further analyse the shape is that adaptive background separation in complex scenes detects recently displaced objects. The algorithms of the recognition are implemented in three different software modules (Camera Server, 3D Reconstruction Module, Calibration Client) of the Intelligent Space.

The error of the estimated position of an object changes with the distance from and pose of the camera. The error is influenced by several factors; the performance of each camera, the method of image processing, etc. Kalman filter is applied to smooth the measured data.



Fig. 3-1. Separation of Human beings from the background

4 Technical environment

As you can see in the Fig.2-1. sensors are located above the space. In actual system two times two Sony EVI D30 pan-tilt CCD cameras hang parallels on the ceiling and the third pair are turned by 90 degrees. They are connected to the general bt848 based image capture board that are adopted as sensors in an industrial standard Pentium III 500MHz PCs as well as general 100baseT LAN cards are used as network devices.

According to the concept of the Intelligent Space the sensors must collect information and control robots. Their limits must be known. In case of the simplest systems we want to know how fast the system is or how many cameras are need to observe the events exactly. What do we call as events? Do we have enough cameras? How do we locate cameras to get the most information using less cameras? If numbers of the cameras are increasing do we get more information or will the processing time less? Do we have enough time to evaluate images in a real-time process? Do the cameras process images in the spot or send them to a “central” computer? And so on ...

What measurements do we make by them and what do we have to transfer to the other computer?

The basement of all questions is the bottleneck? Where are the bottlenecks of our system?

4.1 Cameras

In our system cameras are located parallel which imitate human eyes. This arrangement is efficient to create images, to measure near distance or to evaluate colours and inefficient to make different between human beings and objects. As it was said in the 3.1. the method of the exact positioning was difficult. It needed to much calculation because of the overlapping or views point correspondences of several cameras.

Do we have to evaluate whole 3D image? Do all cameras must work all time? If PCs with 500 MHz processors are used and the average process time is 8 clock signal / operation then 62.5 million operation are done in 1 sec. A normal camera makes 25 images / second. It means 2.5 million operations can be done during this time. According to the User's Manual of Sony EVI D30 camera (Fig.4-1.) it makes images of 786 x 492 pixels. Overlooked the technical parameters of the general Bt848 based image capture board [6], this card is fast enough to send and receive data from the camera. If all pixels are wanted to evaluate 6.5 operations are available for each pixel. Is this time enough for the evaluation using not the RS-232 cable on the camera?

Take an example: Our aim is to recognize persons collapsed in a metro station. Information come to the monitors of the guards' room. In case of emergency a signal must be given and the stuff will decide next steps. Since not this system makes the decision and execution, signal could be faulty. (Human decision making is not avoided.)

In order to get good result whole picture coming from a camera is not needed to evaluate. Our searching method could be the following: searching human face above the ground upto 40 cm. The other parts of image is not interesting. To get relevant information one image is not enough. Since two parallel cameras are focused to the events two images can be made. If these two pictures show give almost the same result the process can be ended. Or not? Imagine a bag on the ground leaned a colour newspaper to its side. The cameras will recognise a human laying on the ground. The number of the dimensions must be increased. It means one camera is not enough. Neither is two parallel ones. Imagine a collapsed man lying on the ground. Generally he has good contrast from two directions. The third direction is covered by hair. For the sake of the good result at least two cameras are needed for the recognition and cameras are turned 90 degrees to each other as a 2D system of coordinates. As it can be seen above according to the example the bottom part of the images are interesting . First man is to be located. To decide whether a human is lying on the floor or not is enough to find a head. All human head have a special color spectrum. This spectrum must be find. Since a head is big enough from this distance (appr. 2-3 m) not whole image of the bottom part of

the image must be tested. If a human head is found the environment of this part is belonged to the head, too. The bottom part of the images must be splitted into 2-4 cm strips. If none of the lines contain human spectrum then no lying men is on the floor. Using this searching method in the first step only 786 x 20 (15,720) pixels are examined during one sample time period instead of 786 x 492 (386,712) pixels. It means appr. 160 operations / pixels.

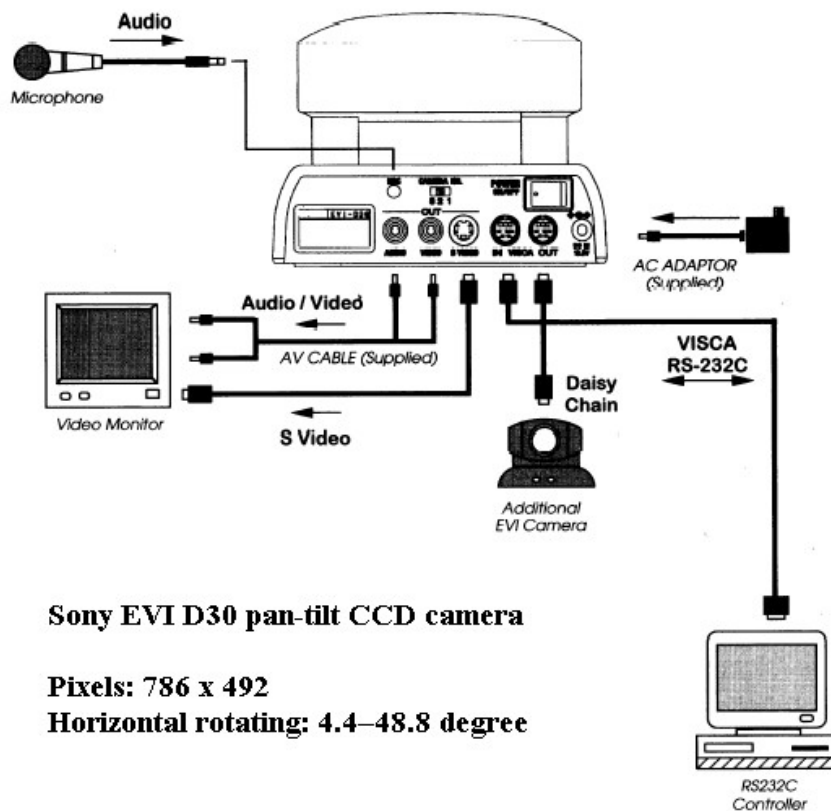


Fig.4-1. Sony EVI D30 camera and its environment

4.2 LAN

Sensors and its additional are connected to the LAN by general 100baseT LAN cards. These cards supply us 100 Mbit/s transfer speed. This speed is the absolute speed of the LAN. If frame making or delay of the transfer protocol is considered then this speed is less than 12,5 MB/s. As Fig.4-1 shows one image is over 386,000 pixels. Supposed 256 colors/pixels it means 25 x 386 kB data/sec/camera. It is almost the full capacity of the transfer channel. The transfer channel could be

the bottleneck of the system. Because of the capacity of the transfer channel central processing is unimaginable. Images must be evaluated nearby the camera. It could be done, but then cameras must inform each other on the events. No news is good news, but it means not null information flows. If events is happened then at least (See. Chapter 4.1) two cameras must cooperate to each other. Coordinantes transformations, overlapping and pattern matching must be done in a very sort period.

Conclusions

Our aim was to find the limits of a Distributed Intelligent Networked Device in the Intelligence Space. In order to find the limits bottleneck of system must be found. In the iSpace tasks are distributed. Sensors connected to DIND are responsible for finding information sources. After having found them they have to be collected and evaluated. During these procedures local comuters and LAN are used. According to the technical environment making samples and image recognition are available in the local computers. If the result of the evaluation is positive than the communication must be started among the DIND devices. This communication means the bottleneck of this system. Remember the more devices use the transfer cannel the less effective transfer rates. Fortunately there are many helping methods to reduce transfer time trough the channel.

- image compressions
 - o advantage:
 - useage time of the transfer cannel reduces significantly
 - o disadvantage:
 - compression and decompression take time
- double transfer channels
 - o advantage:
 - transfer rate is doubled
 - o disadvantage:
 - increasing costs of development
- image partitioning
 - o advantage:
 - parts of the images must be transferred
 - o disadvantage:
 - takes time

This paper dealt with finding limits and bottlenecks in an existing DIND in the iSpace. We focused on the starting phase of the process i.e. when the information must be completed. The automatization of these procedures are complicate because of the many participants, the huge amount of information as well as the length of the calculation processes. Since the recognition methods are either not fast enough or not safe the human intervention is unavoidable yet. While this fact can not be changed process automation is too difficult. If the human decision making is not wanted to avoid than the task of the iSpace is “just” to call attention to the events. To make the process faster tasks have to be splitted among the sensors while an event happens. One of the sensors are responsible for the recognition of the heads, another is for the hands or the fingers.

Remember that these procedures are the starting faces of a complet task. After the recognition phase the answers must be found and exected them. To carry them out resources and time are needed again.

References

- [1] <http://dfs.iis.u-tokyo.ac.jp/~leejoocho/inspace/>
- [2] J.-H. Lee and H. Hashimoto, “Intelligent Space - Its concept and contents –“, *Advanced Robotics Journal*, Vol. 16, No. 4, 2002.
- [3] T. Akiyama, J.-H. Lee, and H. Hashimoto, “Evaluation of CCD Camera Arrangement for Positioning System in Intelligent Space”, *International Symposium on Artificial Life and Robotics*, 2001.
- [4] Peter Korondi, Hideki Hashimoto, ”INTELLIGENT SPACE, AS AN INTEGRATED INTELLIGENT SYSTEM”, **Keynote paper** of International Conference on Electrical Drives and Power Electronics, Proceedings pp. 24-31. 2003.
- [5] Kazuyuki Morioka, Joo-Ho Lee, Hideki Hashimoto, “Human Centered Robotics in Intelligent Space” 2002 IEEE International Conference on Robotics & Automation (ICRA'02), pp.2010-2015, May, 2002
- [6] Booktree Division, Rockwell Semiconductor System, Inc.: Bt848/848A849A Single-Chip Video Capture for PCI, February 1997.
apps@booktree.com