

논문 98-7-2-07

## 다중 TMS320C31 DSP를 사용한 3-D 비전센서 Implementation.

V.옥센헨들러\*, A.벤스하이르\*, P.미셰\*, 이 상국(李相國)\*\*

## A 3-D Vision Sensor Implementation on Multiple DSPs TMS320C31

V. OKSENHENDLER\*, Abdelaziz BENSRAHAI\*, Pierre MICHE\*, and Sang-Goog LEE\*\*

## 요 약

독립적인 로봇이나 자동차 제어 응용을 위하여 고속 3-D 비전시스템들은 매우 중요하다. 이 논문은 다음과 같은 세가지 과정으로 구성되는 stereo vision process 개발에 대하여 논술한다 : 왼쪽과 오른쪽 이미지의 edges 추출, matching corresponding edges와 3-D map의 계산. 이 process는 VME 150/40 Imaging Technology vision system에서 이루어졌다. 이것은 display, acquisition, 4Mbytes image frame memory와 세 개의 연산 카드로 구성되는 modular system이다. 40 MHz로 작동하는 프로그래머블 연산 모듈은 64×32 bit instruction cache와 두 개의 1024×32 bit RAM을 가진 TMS320C31 DSP에 기초를 두고 있다. 그것들은 각각 512 Kbyte static RAM, 4 Mbyte image memory, 1 Mbyte flash EEPROM과 하나의 직렬 포트로 구성되어있다. 모듈간의 데이터 전송과 교환은 8 bit globalvideo bus와 세 개의 local configurable pipeline 8 bit video bus에 의하여 이루어졌고, system management를 위하여 VME bus가 쓰였다. 두 개의 DSP는 왼쪽 및 오른쪽 이미지 edges 검출을 위하여 쓰였고 마지막 processor는 matching process와 3-D 연산에 사용되었다. 512×512픽셀 이미지에서 이 센서는 scene complexity에 따라 1Hz정도의 조밀한 3-D map을 생성했다. 특수목적의 multiprocessor card들을 사용하면 결과를 향상시킬수 있을 것이다.

## Abstract

High-speed 3D vision systems are essential for autonomous robot or vehicle control applications. In our study, a stereo vision process has been developed. It consists of three steps : extraction of edges in right and left images, matching corresponding edges and calculation of the 3D map. This process is implemented in a VME 150/40 Imaging Technology vision system. It is a modular system composed by a display, an acquisition, a four Mbytes image frame memory, and three computational cards. Programmable accelerator computational modules are running at 40 MHz and are based on TMS320C31 DSP with a 64x32 bit instruction cache and two 1024x32 bit internal RAMs. Each is equipped with 512 Kbytes static RAM, 4 Mbytes image memory, 1 Mbytes flash EEPROM and a serial port. Data transfers and communications between modules are provided by three 8 bit global video bus, and three local configurable pipeline 8 bit video bus. The VME bus is dedicated to system management. Tasks between DSPs are distributed as follows: two DSPs are used to edges detection, one for the right image and the other for the left one. The last processor computes the matching process and the 3D calculation. With 512x512 pixels images, this sensor generates dense 3D maps at a rate of about 1 Hz depending of the scene complexity. Results can surely be improved by using a special suited multiprocessors cards.

\* L.C.I.A., INSA de Rouen, BP 08, 76 131 Mont-Saint-Aignan cedex, FRANCE

\*\* 경북대학교 센서기술연구소 (Sensor Technology Reserch Center, Kyungpook National University)  
<접수일자 : 1997년 12월 30일>

## 1. Introduction

Stereo vision allows to find distances of objects from two images acquired from different positions by studying the difference of locations of stereo

corresponding points. Several techniques have been developed to infer 3-D information from a set of brightness images. Among them, passive stereovision is a very attractive approach for ranging applications especially because of its ability of working in various illumination conditions and in a large depth range<sup>[1]</sup>.

In our approach, a fast and automatic stereovision process has been developed<sup>[2]</sup>. The applications of this system are mobile robot guidance and autonomous vehicle navigation. This needs high speed response. So, to increase the speed of our system, we propose an implementation of our stereovision algorithms on specialised architecture which comprises 3 DSPs TMS320C31. A 150/40 vision system from Imaging Technology has been chosen for this task because it is a modular system which allows parallelism of many image treatments.

Some experimental results are presented at the end of this paper.

## 2. Stereovision process

In general case, stereovision process consists of three steps: feature extraction in right and left images, feature matching and calculation of the 3-D map.

To reduce the computation time and to simplify the stereovision algorithms, we chose a special configuration. The two cameras have parallel optical axes, and the  $i$ th scanline of the right and left CCDs are in the exact extension of each other. Given this configuration, the epipolar lines are confused with the scanlines of the CCDs. Consequently, the stereovision algorithms may be processed line by line.

The first step of our stereovision system is based on a new concept that we call *declivity*<sup>[3]</sup>. A *declivity* (see Fig. 1) is defined in an image line as a cluster of contiguous pixels, limited by two end-points ( $X_i$  and  $X_{i+1}$ ) which correspond to

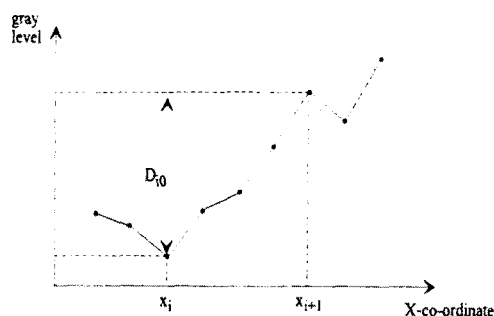


Fig. 1. An example of declivity.

two consecutive local extrema of gray level intensity, i.e. one maximum and one minimum.

Then 3-D reconstruction is based on detection and matching of characteristic declivities which are detected by a thresholding of their amplitude  $D_i$ . Furthermore, to be matched this primitive is characterised by its position and the photometric characteristic of its neighbourhood. In a pair of epipolar line, a gain is associated at each possible matching between characteristic elements of right and left images. Gains are representative of photometric similarity of primitives and the set of the best associations is selected by using dynamic programming method<sup>[4]</sup>.

Once found the matching declivities, the 3-D information are deduced using disparity values which are the differences between the positions of the right and left matched primitives. Then, due to the cameras configuration, the depth is given by :

$$depth = \frac{f \times d}{disparity}$$

where  $d$  is the distance between cameras and  $f$  is their focal length.

In the final step, the depth map is completed by interpolation.

## 3. System

The system used is a 150/40 VME vision system from Imaging Technology. This is a

modular system allowing general purpose image treatment and specific processes depending on the chosen configuration. It is composed of mother boards at the VME format on which can be plugged specific modules.

Our configuration (see figure 3) is composed of two mother boards: an Advanced Image Manager (IMA) and a Computational Module Controller (CMC) one which are plugged modules. Modules are: a Color Acquisition Module (AM\_CLR) and a Pseudo Color Display Module (DM\_PC) plugged on the IMA, and three Programmable Accelerator Computational Modules (CM\_PA) based on a Texas Instruments floating point TMS 320C31 DSP, one of them plugged on the IMA and the others on the CMC.

- Advanced Image Manager:

The IMA includes four Megabytes of reconfigurable image memory, a cross-port switch for data routing and three supports for plug-in modules.

Memory is organised as four frames of 1K x 1K x 8 bits. Each frame is provided with two asynchronous 8 bits video ports: one for input and the other for output. In our case, we only use three frames: one for display and two for acquisition from cameras.

The cross-port switch is a specific circuit which allows to configure connections between the six 8 bits inputs and the six 8 bits outputs of IMA, and inputs and outputs of modules of IMA.

- Computational Module Controller:

As for the IMA, it provides supports for three plug-in computational modules and a cross-port-switch to configure data communication for modules.

The cross-port-switch allows to configure video connections with external boards: six 8 bits inputs and three 8 bits outputs as well as video connections between modules of the CMC.

The CMC also provides local connections between serial ports of computational modules.

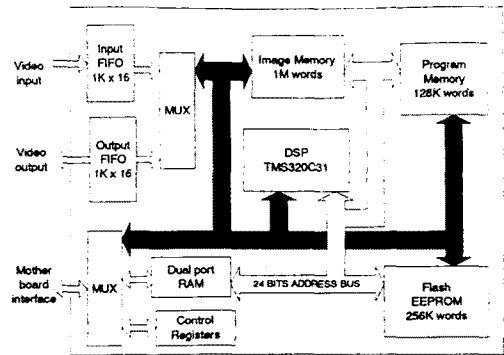


Fig. 2. CM-PA block diagram.

- Programmable Accelerator Computational Module:

The CM\_PA (Fig.2) is based on a Texas Instrument floating point TMS320C31 DSP. The TMS320C31 has an addressing space of 24 bits, a 32 bits data width, two 2K x 32 bits single cycle RAM, a 64 x 32 bits instruction cache, a single cycle floating point multiplier, a single cycle ALU, two address generators and a DMA controller.

The DSP addresses four different types of external memory:

- Program memory: it is a 128K x 32 bits of zero wait state static RAM for program store, extended stack, coefficient storage and general purpose memory.

- Dual port RAM: it is a 4K x 16 bits memory shared by the DSP and the CPU host. It allows communications without halting the DSP. In our process, it is used only during initialisation when the host loads program code and configures the CM-PA.

- Image memory: it is 1M x 32 bits and can be addressed by DSP, by VME host and by video data bus. Accessing image memory simultaneously by different process is not possible.

- Flash EEPROM boot memory: it is a 256K x 32 bits that allows DSP to boot itself. This memory is not used in our 3-D vision sensor application.

- External communication devices:
  - Dual port RAM is previously described.
  - One 16 bits video input port allowing connection with video bus. This port is not separable in two 8 bits ports but the use of only 8 bits is possible. The same characteristics are available for video output.
  - One serial link.

The general scheme of the system used for 3-D vision is provided below.

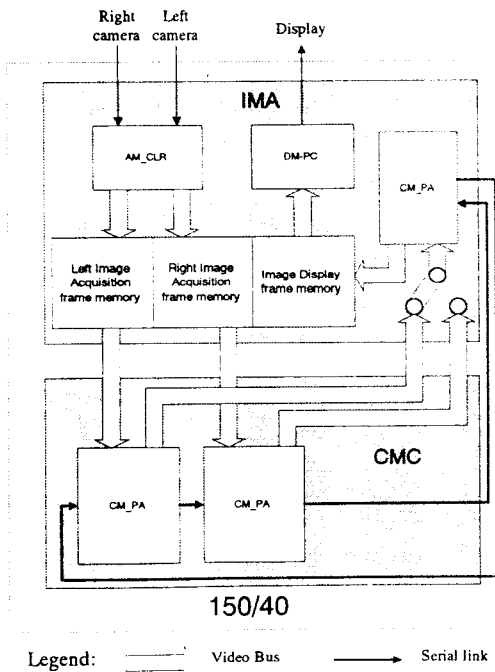


Fig. 3. General configuration.

All the system is configured by the host at the initialisation. In our case, the CM\_PA of the IMA becomes the server unit. During the computation process, it controls the cross-port switch and the modules of the IMA.

Imaging technology vision system is provided with software utilities and libraries allowing to configure boards and modules and to manage computation of CM\_PAs. Program language used for host and DSPs is C. Debugging DSP program

is possible only for the CM\_PA of IMA when it is not server. To develop the 3D sensor programs, serial port of CM\_PAs is used to control data flows and computations.

#### 4. Tasks allocation

As mentioned above, the stereovision process consists in three steps. In order to take advantage of the three DSPs, these steps have been separated in tasks which are distributed as follows. The two DSPs on the CMC are allocated respectively for right and left image segmentation. The DSP on the IMA, which is server, matches characteristic elements issued from DSPs of CMC. It also manages acquisition and data transfer.

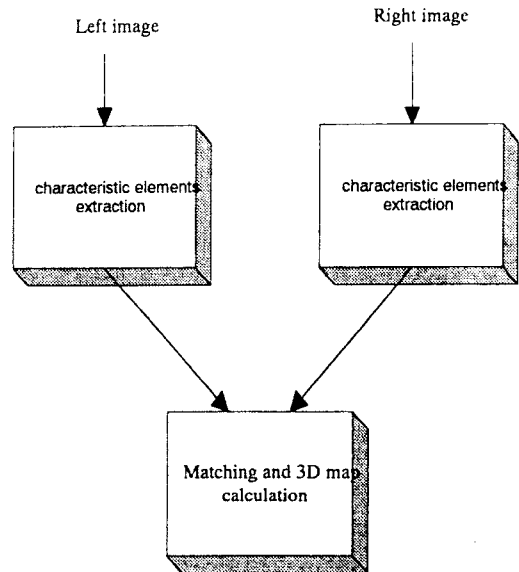


Fig.4 Tasks allocation and data transfer.

Data communications are made via the video buses. All data transferred by video buses must have the same image format, here imposed by the acquisition which is 512 x 512 x 8 bits. The data transfer step is time critical due to the amount of data communications.

During a loop (excluding the first one since it

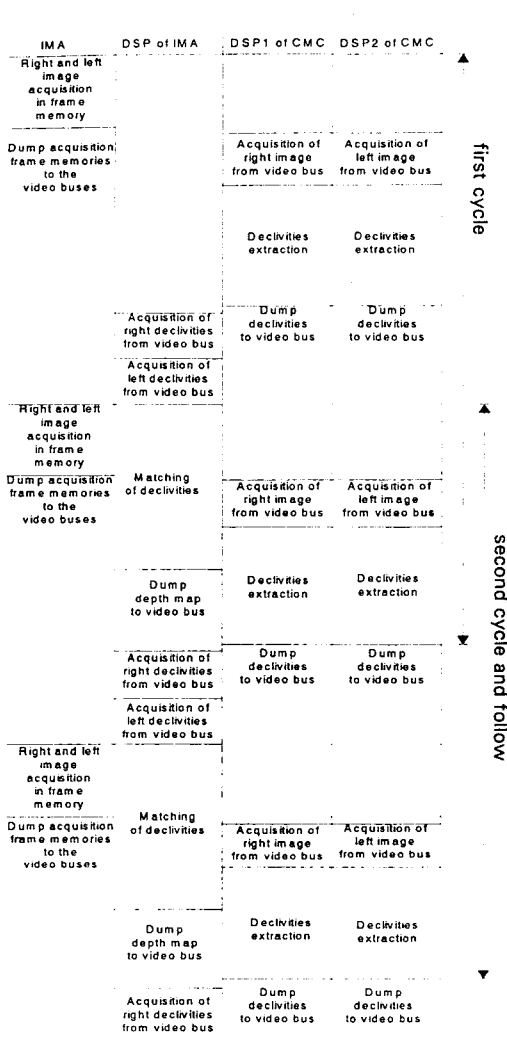


Fig. 5. Chronogram.

is particular) (see figure 5), right and left images are acquired synchronously in the IMA frame memory during the matching of characteristic elements issued of the previous stereo image pair. This implies no time penalty because acquisition in IMA memory frame is managed by the IMA board without halting the DSP. Next, images are transferred from the IMA memory frame to the image memories of CM-PAs of the CMC. The DSPs must wait for the end of image loading before extracting declicivities. Characteristic elements are then converted in an image format,

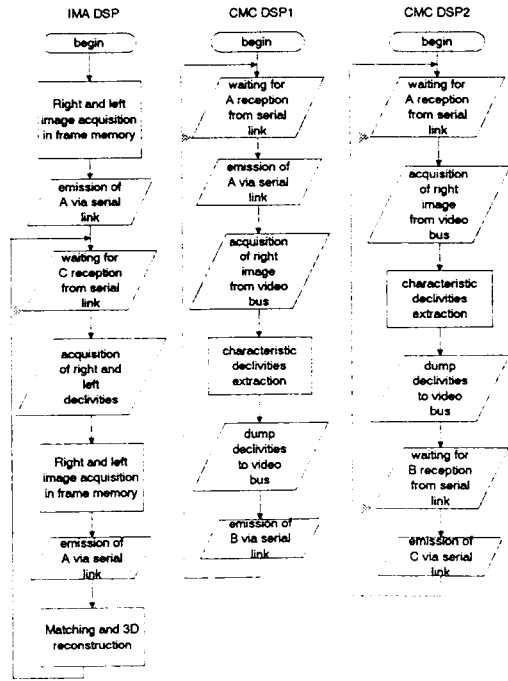


Fig. 6. Synchronisation by serial link

loaded in image memory and dumped to the video bus connected to the IMA board. To get back characteristic elements, the DSP of the IMA connects its entry port to the video bus linked with one of the DSPs of the CMC and acquires the characteristic elements image. Then it connects its entry port to the video bus linked with the other DSP of the CMC and acquires the other characteristic elements. Characteristic element images are converted and loaded in DSP RAM (which is one wait state memory) and are matched in order to calculate 3-D map. Results are then converted in an image format, written in DSP image memory and dumped to the display.

This process is looped in each DSPs and is illustrated by the figures 5 and 6.

The serial link is dedicated to synchronisation task. Before the beginning of the stereo process, DSPs must wait the end of the load of the program code in each DSP. For this, the IMA DSP emits a zero on the serial link to the first CMC DSP. Next, it emits the received number

incremented by one until a three is emitted and received. In case of erroneous transmission (reception of a number different of 0, 1, 2, or 3) the process comes back to the beginning by emitting a zero.

For synchronisation of the different tasks, acquisition, segmentation and matching, the same way is used. Synchronisation words corresponding at the beginning or at the end of each task are transmitted by the serial link.

## 5. Results

The sensor was tested with indoor and outdoor scenes. Rate of the sensor is calculated from 200 successive processes on images which have a format of  $512 \times 512 \times 8$  bits. The computation times are comprised between 1 and 1.5 second depending on scene complexity. The following result (fig. 7, 8 and 9) is corresponding to an indoor scene which involves a chair in front of a white board.

In the resulting depth map, dark pixels are meaning far points and white pixels are meaning near points.

The computation time for this scene is about 1.1 second. This time allows to consider specific applications like robot guidance.

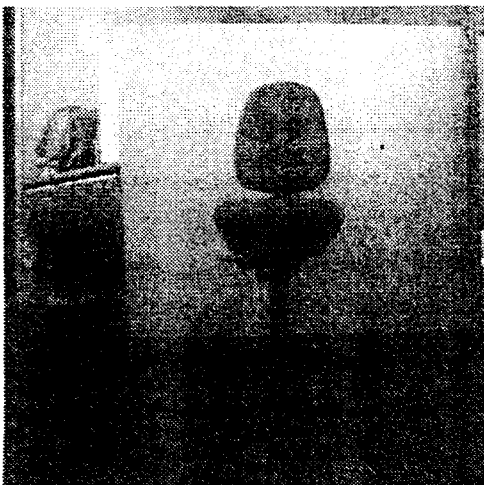


Fig.7 Left image



Fig.8 Right image.

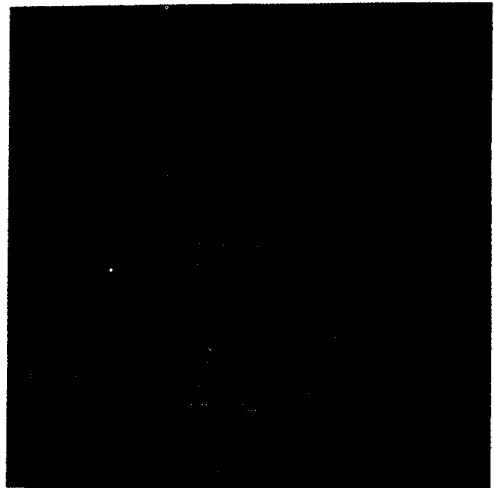


Fig.9 Depth map.

## 6. Conclusion

Results obtained with the 150/40 vision system from Imaging Technology show that DSPs are well suited for our 3D vision sensor algorithms. We can hope that the use of new DSPs generation like Texas Instrument C80 will reduce significantly the processing time. System with fast shared memory will exchange data (characteristic elements) without needing any image format conversion. Moreover, the treatments can be made line by line allowing a high parallelism level.

References

- [1] Horn B. K. P., (1991). Robot vision. MIT Press, 299-326,1991
- [2] P. MICHE, A. BENSRAIR, S.G. LEE : " High Speed and Self-Adaptive stereo Vision Algorithms for Implementation in a 3-D vision sensor" Journal of The Korean Sensor Society, vol.6, No.2, March 1997.
- [3] Miché P. and Debrie R., (1994). Fast and self-adaptive image segmentation using extended declivity. Annals of Telecommunications, 50, No.3-4, 401-410, 1995
- [4] Bensrhair A., Miché P. and Debrie R.,(1996). Fast and automatic stereovision matching algorithm based on dynamic programming method. Pattern Recognition Letters, 17, 457-466, 1996.

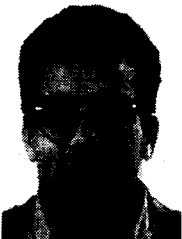
著 者 紹 介



**Vincent OKSENHENDLER**  
 1966년 프랑스 생. 1997년 프랑스 국립루앙대학교 졸업(공학박사), 현재 프랑스 국립 응용 과학원 (Institut National des Sciences Appliquees de Rouen) 센서 및 전자 기기 연구소(L.C.I.A. sensor instrumentation & analysis lab.) 연구원.



**Pierre MICHE**  
 1947년 프랑스 생. 1975년 프랑스 국립 루앙 대학교 졸업(공학박사). 현재 프랑스 국립 응용 과학원(Institut National des Sciences Appliquees de Rouen) 센서 및 전자 기기 연구소 (L.C.I.A. sensor instrumentation & analysis lab.) 소장, 프랑스 국립 루앙대학교 공과대학 학장



**Abdelaziz BENSRAIR**  
 1960년 2월 15일 모로코 생. 1992년 루앙대학교 전자·전기공학부 computer vision 전공 졸업(공학박사). 현재 프랑스 국립 응용 과학원(Institut National des Sciences Appliquees de Rouen) 센서 및 전자 기기 연구소(L.C.I.A. sensor instrumentation & analysis lab.) 연구원, 프랑스국립 루앙 대학교 전자·전기공학부 조교수.

**Sang-Goog LEE**  
 『센서학회지 제6권 제2호』 논문 97-6-2-06, p.130 참조. 현재 프랑스 국립 응용 과학원(Institut National des Sciences Appliquees de Rouen) 센서 및 전자기기 연구소(L.C.I.A. sensor instrumentation & analysis lab.) 연구원, 경북대학교 센서기술연구소 프랑스현지연구센터(IJRL) 운영실장, 프랑스국립 루앙대학교 전자·전기공학부 조교수.