TOWARDS ACTIVE VISION SENSORS

Harvey A. Cohen and David Suter
Computer Science Department
La Trobe University
Bundoora Victoria
AUSTRALIA 3083

SUMMARY The concept of an active vision sensor is explained and the advantages of integrating in the sensor both image capture and significant processing discussed. Various active sensor designs and design features are described and limitations noted.

1 INTRODUCTION

The operations of a robotic vision system involve

(a) sampling of the image by an array of pixel sensors
(b) scanning of the sensor array
(c) transmission and capture of scan data
(d) image processing - edge detection
(e) object recognition and scene analysis

Contemporary sensors integrate the functions of image sampling and pixel scanning and output pixel data in serial form via a single data stream as video (or XY pixel data). Such sensors are essentially passive as no significant image processing is performed during the pixel scan. An active sensor performs some image processing function in addition to sensing the visual field.

The obvious limitation to the development of active vision sensors is the scarcity of chip real estate: existing single chip vision sensors operate at close to today's design limits for in-chip complexity. However recent work aimed at producing an effective doubling of resolution via a piezoelectric shift-and-wobble operation shows that using current fabrication techniques existing resolution can be combined with in-chip processing. Nevertheless real estate on vision sensor chips is in short supply so the question arises as to what image processing functions can be effectively calculated and how. In this paper, after reviewing sensors with respect to capability of in-chip processing, we recount two different approaches to the design of Active vision processors of which one was recently patented by Nishizawa et al (1984).

2 THE SCOPE FOR ACTIVE VISION SENSORS

2.1 Active Image Processing by Optical Modulation

The term active image processing has been recently used by Ikuta (1985) who applied it to a conventional camera used in conjunction with a frame buffer for the impinging image is modulated by passage through a vibrating optical

Accordingly Ikuta utilises mechanical means to *perform convolutions* and in particular his apparatus effectively performs a Laplacian type operation equivalent to the 3*3 mask L4:

$$L4 = \begin{matrix} 0 & -1 & 0 \\ -1 & +4 & -1 \\ 0 & -1 & 0 \end{matrix}$$

This operation replaces each pixel value with the difference between four times that value and the sum of four neighbouring cells. L4 is a useful edge enhancer which works best with vertical and horizontal lines. Ikuta's apparatus can be programmed to perform more elaborate convolutions. However we believe that this scheme of optical modulation via a vibrating mirror is inherently slow, due to mechanical limitations, and does not offer the possibilities of speed and practical utility that one-chip sensor chips promise.

2.2 Solid-state sensors

Solid-state vision sensors detect light by the production of a cloud of charge proportional to the (pixel) intensity. Traditional CCD sensors offer no on-chip storage of pixel values and inherently hold pixel data in serial form. MOS sensors involve individual photo cells holding clouds of charge which can be sensed non-destructively so that a MOS sensor is intrinsically XY-addressable. Consequently in a MOS sensor an in-chip processor can in principle perform convolutions.

2.3 Analogue arithmetic

Where the visual sensor on a chip produces a cloud of charge it is far more economical of chip real estate to provide on-chip analogue processing of such discrete units of sampled data. By using CTD analogue arithmetic the need to perform much large range A/D is obviated, and only further down the calculation path is A/D required for chip output.

As long ago as 1976 Whitehouse (1976) and notably Butler et al (1976) pointed out that charge transfer devices provide on-chip analogue memory that can be manipulated more easily than by digital signal processing. This early work was applied to certain 1 dimensional convolutions.

The same concept of processing data analogically by physical manipulation of discrete bundles of charge carriers has recently being successfully applied to the design of an MNOS sensing chip. Yamasaki and Ando (1985) designed an MNOS single chip sensor with in-chip capability for 'optical multiplication'. *This* chip cannot perform any useful convolution, but represents a significant step in chip evolution.

3 SINGLE CHIP ACTIVE VISION SENSORS

3.1 The Nishizawa-Konishi Sensor

In their recent patent (dated 1984 - filed 1981) Nishizawa and Konishi describe a MNOS image sensor that includes an array of photocells, with a non-destructive read-out. However in contrast to the simple scanning (akin to key-board scanning) in an ordinary MOS sensor this sensor features a scanner that accesses in sequence the cells involved in the convolution, but inserts appropriate delays so that groups of cells can be simultaneously added. (The addition is purely analogue). The patent data shows the Laplacian convolution L4 and a low pass (smoothing) filter as being readily implementable. The patent is assigned to Fuji Photo Film Co. There is no doubt that this design represents a milestone in the development of active sensors.

3.2 XY Shift chip

Yosida et al (1985) and Tanuma et al (1985) all of the Toshiba corporation have described a CCD chip which is flexibly mounted on a conventional chip base with motion perpendicular to the DIP axis controlled by piezoelectric actuators. By synchronising the chip shift - the X shift - with scanning an effective doubling of the horizontal pixel number was achieved. A tendency for image flicker was alleviated by superimposing a small 'wobble' on top of the shift.

It is far from obvious how to reliably mount a chip with both X and Y shift control. A MOS chip with such programmable XY motion would be an exceedingly simple convolving chip. Consider firstly the performance of a 3*3 convolution involving only positive (or zero) coefficients (such as low pass filtering): each cell would move in a square about the positions of each of its neighbours staying at each position for a time proportional to the corresponding mask coefficient. At the end of the tour the cell would be read. A mask involving negative coefficients requires two (selectable) XY shift cycles, one cycle for positive coefficients, one for the negative mask coefficients. Convoluted pixel values would be determined in an (external) frame buffer by subtraction.

4 CONCLUSIONS

It is notable that both Nishizawa-Konishi and our proposed Shift XY active chip designs are probably effectively limited to 3*3 convolutions. For edge detection in some circumstances a 3*3 Laplacian is adequate, but a more appropriate filter for edge detection combines the Laplacian with a smoothing operator, such as the DOGS (difference of Gaussian) favoured by Marr (1982) and is not realisable with a 3*3 mask.

The approach here presented may be contrasted with various proposals to develop specialised convolution chips designed to process an entire image frame by transformation methods. In such chips a transformation such as FFT, Z-transform,

number theoretic, transform, is used to map from pixel space onto another space in which the convolution has lower complexity. This is a VLSI version of an old idea in signal processing (see, e.g., Stockham 1969).

Middlehoek and Hoogerwerf have used the term "smart sensor" to describe sensors with a signal conditioning and compensating capability. Accordingly active light sensors are smart.

The Nishizawa chip is of especial interest in that the scheme of piping and routing data is reminiscent of the processing scheme in systolic arrays for convolution as in Kung (1982).

5 REFERENCES

BUTLER, W.J. et al, "Charge-Transfer Analog Memories for Radar and ECM Systems", IEEE Trans Electron Devices Vol ED-23 No 2 Feb 1976 pp 161-164. IKUTA, (1985) "Active Image Processing", Applied Optics, Vol 24 No 18 pp 2807-2913.

KUNG, H.T. and SONG, S.W. (1982). "A Systolic 2-D Convolution Chip" in Preston, K. and Uhr, L., "Multicomputers and Image Processing", Academic Press, New York.

MARR, D. (1982). "VISION: A Computational Investigation into the Human Representation and Processing of Visual Information", W.H. Freeman, San Francisco.

MIDDLEHOEK, *S.* and HOOGERWERF, A.C. (1985). "Smart Sensors: Where and When?" Sensors and Actuators, Vol. 8 pp 39-48.

NISHIZAWA, et al (1984). "Two-dimensional Solid-State Image Sensor", U.S. Patent Number 4,472,638 Sept 18, 1954.

STOCKHAM, T.G. (1969). "High-speed Convolution and Correlation with Applications to Digital Filtering" in Rader and Gold (eds),"Digital Signal Processing': McGraw Hill 1969 pp 203-252.

TANUMA, C. et al. (1985). "SYNCHRO VISION CCD - High-Resolution Image Reproduction for a CCD Imager Using Swinging Operation", Toshiba Review, No 152 pp 34-37.

WHITEHOUSE, H.J. (1976). "Signal Processing with Charge-Transfer Devices" IEEE Trans Electron Devices, ED-23 No 2 pp 132-3.

YAMASAKI, H. and ANDO, T. (1985). ."Optical Multiplication in Solid-State Imaging Devices with an Inherent MNOS Memory Gate", IEEE Electron devices Letters EDL Vol 6 No 2 Feb 1985 pp 88-90.

YASHIDA, O. et al (1985). "Image Flicker Control by CCD-Chip-Shift and Wobble Operation at Double-Enhanced Nyquist Regions" IEEE Trans Elect Dev ED32 No 8 Aug 1985 pp 1505-1510.