IAPR Workshop on CV - Special Hardware and Industrial Applications OCT.12-14, 1988, Tokyo

Real-time stereo matching using transputer arrays for close-range applications

J-P Muller[1], A Anthony[1], A T Brown[2], A T Deacon[2]
S A Kennedy[2], P M Montgomery[2], G W Robertson[2], D M Watson[2]

[1] Department of Photogrammetry & Surveying, University College London,
Gower Street, London WC1E 6BT, U.K. (ARPANET: jpmuller@cs.ucl.ac.uk)

[2] THORN EMI Central Research Laboratories, Dawley Road,
Hayes, Mddx UB3 1HH, U.K. (ARPANET: pmm1%uk.alvey@nss.cs.ucl.ac.uk)

ABSTRACT

The Alvey MMI-137 project is concerned with the development of real-time 2.5D vision systems using reconfigurable transputer arrays. Over the last two years, a Close-Range Vision Cell (CRVC) based on active stereo using CCIR frame-transfer CCD cameras and a texture light projection system has been designed and installed at THORN EMI Central Research Laboratories. This CRVC has been developed for research into direct entry of surface shape into CAD systems for free form objects at a standoff range of 500-850mm with a rms accuracy of 0.1mm in plan and range. CRVC images have been used to evaluate a variety of stereo matchers for their accuracy, reliability and sampling density. The CRVC is currently integrated into a special-purpose real-time image processing module (the THORN EMI Modular Image Processing System, TEMIPS) which allows image data to be transferred to a Sun-3™ workstation in 0.5s and in the future to a PARSYS™ SN1000 at CCIR video frame-rates. This PARSYS™ transputer array is predicted to be able to estimate depth values using an adaptive least-squares patch-fitting algorithm at rates of 50 patches/second/ transputer. Hence for a surface model to be produced every 5th pixel using 32 T800 transputers should take 6.5 seconds.

1. INTRODUCTION

Co-ordinate Measuring Machines (CMM) have been in use for many years for the determination of surface shape of a wide variety of industrial objects from machined components to automobile clay models and airframes (Higashimoto, 1983). Typical tolerances required for these applications are on the order of 0.1mm. There are a number of significant drawbacks for such CMMs for the precise determination of the shape of surfaces of free form objects. These include their high cost, both in capital and in manpower, their potential to impact and damage the surface of an object, their very slow speed and their consequent inability to measure dynamically- changing structures.

A number of alternative non-contact techniques have been investigated. These have tended to concentrate on systems employing active vision (see recent comprehensive review by Besl, 1988). However, most of these types of systems suffer from problems associated with trading-off accuracy against speed, specularity, non-extensibility to other problem domains and eye safety. An alternative approach is to use active stereo or multiple-image matching to extract a surface model.

To provide sufficient features for image matching, unstructured light illumination may be used. Systems incorporating this approach include the PRISM system developed at MIT (Nishihara, 1984) which was applied to robotic grasping of parts from a pile (Ikeuchi et al., 1986) and the Zeiss INDUSURF™ system installed at Volkswagen for measuring styling models, originally developed at Stuttgart University (Schewe & Foerstner, 1986). Claus (1988) recently reported on a comparison of CMM technology and the INDUSURF™ system for surface shape determination of a Volkswagen styling model. He showed that in the former case, 30-40 days was required using CMM technology whilst the active stereo matching system incorporating digitisation of stereo photographs reduced this time to 0.5 days. In the latter case, rms accuracies of 0.1 pixels or 1:5000 in plan have been achieved (Claus, 1987).

Besl (1988) recently proposed a sensor performance parameter, M, for active vision systems biased towards accuracy. In this, M is proportional to $1/\sqrt{T}$ where T is a linear combination of the pixel dwell and processing time. The pixel dwell time in the case of frame-rate acquisition is some 67.5ns for CCIR video. The processing time for adaptive least squares patch-matching (Otto & Chau, 1988) is about 200ms on a Sun-3™ workstation. These figures emphasize that in order for stereo vision to have comparable or better performance than other active vision systems, processing times must be increased by several orders of magnitude. It is one of the objectives of our Alvey project to reduce this stereo-processing bottleneck by several orders of magnitude using arrays of floating-point transputers (Muller et al., 1988, Collins & Roberts, 1988).

2. CLOSE RANGE VISION CELL(CRVC)

Anthony et al. (1988) describe the system configuration and camera calibration used in the Close Range Vision Cell constructed at THORN EMI Central Research Laboratories . Figure 1 is a picture of the cell, captured on one of our CCD cameras. In the foreground the rigid aluminium framework, one of the COHU 4710 cameras and the mirror used to project textured light onto the object can be seen whilst in the background the two lower monitors are displaying the left and right images captured by the two COHU cameras.

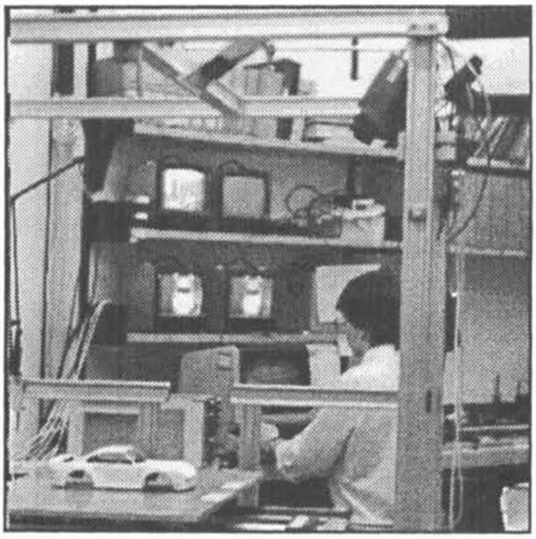

Figure 1. Close Range Vision Cell showing rigid framework, CCD cameras, mirror for projection of textured illumination and TEMIPS display of digitised frames.

The CRVC has been designed to capture surface shape with a precision of 0.1mm at a standoff distance of 500-850mm with an object-space pixel size varying from 0.3-0.5mm. Single pixel disparities, depending on the Base-to-Height ratio vary from 0.15-1.8mm corresponding to the aforementioned depth-of-field. The limited sensor size (hence field of view) of CCD cameras (e.g. COHU 4710 is 6.4mm) compared with metric photographic cameras (e.g.230mm^2) restricts the measuring volume in the CRVC to some 350mm. Extending the CRVC to larger volumes would require either multiple cameras or larger CCD pixel arrays. The system is currently designed for 12-bit range resolution taken from 8-bit greyscale images. This requires that specular highlights be removed using matte painting of the surface. We hope to remove this restriction in the future by a combination of modulating the intensity of the light source(s) (see example in Doemens et al., 1986) and/or increasing the dynamic range from some other CCD sensor.

A number of authors have demonstrated that the highest accuracy image matching system for continuous surfaces is one based on least squares matching of grey-levels where rms accuracies of 0.05-0.1 pixels have been achieved (Gruen & Baltsavias, 1986; Ackermann et al., 1986; Day & Muller, 1988; Wrobel, 1988). The most significant drawbacks of this type of algorithm is its limited pull-in range (<2 pixels for <3 iterations) and its inability in its existing configuration to handle disparity jumps. A number of solutions exist to address these problems which are not discussed here. Suffice it to say, that for certain ranges of disparity, the sheet-growing variant of the least-squares matcher (see Otto & Chau, 1988) does produce dense rangemaps, comparable or greater in density than laser radar (see Lougheed & Sampson, 1988). Figure 2 shows an example from the CRVC of the output disparity image from the textured illuminated image of a model Porsche motor-car. Edge information in this disparity image is missing. Schewe and Foerstner (1986) discuss using taped lines to measure edges for large objects. Current plans call for the use of a pointable laser system to provide a scanning light-stripe. This can be used to acquire further active stereo information on occluding contours and edges.

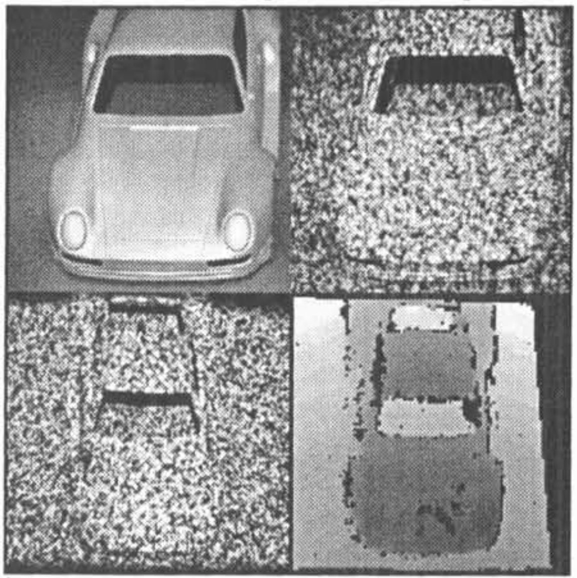

Figure 2. Examples of automated stereo matching using least-squares techniques. Upper panel shows effect of textured illumination. Lower panel shows dense disparities.

3. TEMIPS

A special-purpose real-time image processing system has been designed and constructed at THORN EMI Central Research Laboratories to provide the following functions:

- synchronisation of camera pair
- frame-rate video data acquisition and display
- hardware pre-processing
- high-speed data transfer to/from host (Sun™ and PARSYS™ SN1000 reconfigurable transputer array)

This image processing system has been assembled using a selection of appropriate modules from the range provided by the THORN EMI Modular Image Processing System (TEMIPS).

TEMIPS is an extendable range of hardware and software system components configurable to specific applications at a low cost and avoiding the limitations of general purpose systems. A modular approach has been adopted to provide maximum flexibility in implementation whilst limiting the incorporation of potentially expensive hardware.

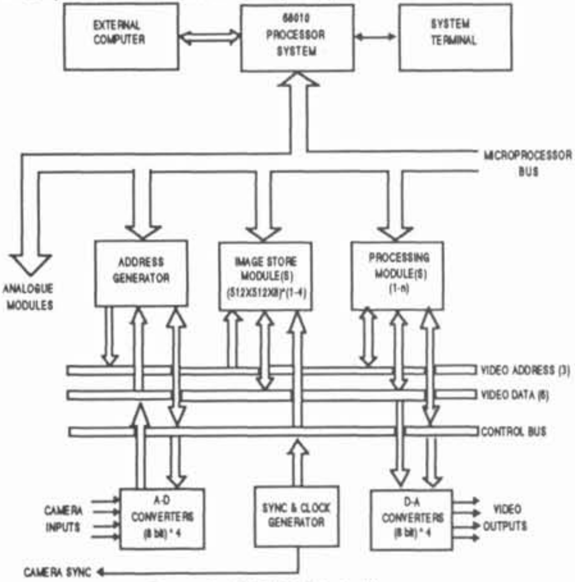

Figure 3. TEMIPS block diagram

Figure 3 shows a schematic diagram of the four main subsystems: Microprocessor, Analogue, Image and Processor. All control functions and user and host interfacing are supervised by the 68010-based microprocessor susbsystem which comprises various controller, memory and I/O modules. The Analogue subsystem provides sync generators, ADC and DAC modules (currently up to 8-bits), whilst the Image subsystem provides address generator and image store modules. The Processor subsystem comprises modules which implement various image processing kernel functions (e.g. frame-rate resampling) in hardware.

The bus structure used to link the modules of an image processing system together ultimately places limits on its processing capabilities both in terms of speed and flexibility. TEMIPS uses a custom bus, split into three sections:

- microprocessor subsystem
- analogue subsystem
- Image & Processor bus

The last one of these provide three independent 18-bit video address busses and six independent 8-bit busses, thereby providing a very flexible, high bandwidth interconnection scheme for these modules.

The TEMIPS hardware provided for the stereo matching system incorporates frame-rate digitisation and display of CCIR video for up to four camera inputs operating independently with 512x512x8-bit sampling. For the sub-pixel acuities required to achieve our target specification of 0.1mm range resolution, timing errors in the acquired images must be kept to a minimum. For this reason, the camera line and frame syncs are locked to the TEMIPS system clock. COHU 4710 cameras were chosen since they had the required external synchronisation facility.

Figure 4 shows the final system configuration of the CRVC Exemplar. Interfaces are provided for the Sun-3™ workstation, the transputer array (Link interface) and the two CCD cameras. The Link interface, which is desgned around a TEMIPS image memory module, will support bi-directional image data transfer using four INMOS serial links running at 20Mbits/sec. This gives a transfer rate of approximately 6.5MB/sec from TEMIPS to the processor array. A pair of these modules may be used to transfer the stereo pair simultaneously.

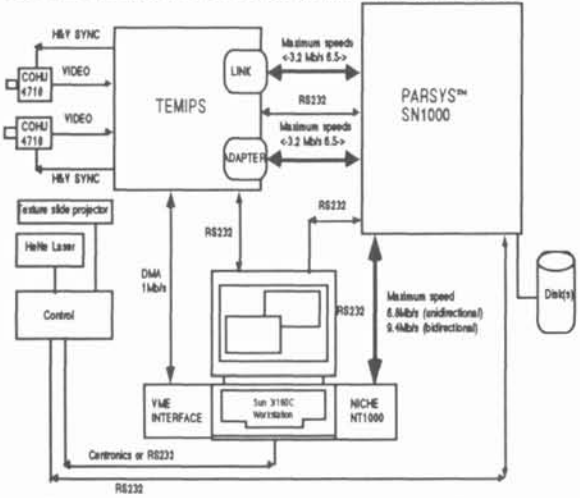

Figure 4. Close Range Vision Cell Transputer Exemplar

4. DIRECT INPUT FOR CAD

To construct a CAD representation of an object for eventual "3D xerography" (Bieman, 1988) on either a numerically controlled machine (NCM) or a solid modelling graphics system for later input to a model-based vision system (Sharma & Scrivener, 1986), information on both surface reflectance and geometry must be obtained. Surface reflectance information can be obtained using a least-squares approach with multiple images (see Wrobel, 1988) and will not be discussd further here. Complete surface geometry must be built using multiple range images.

The construction of accurate mathematical models from range data have been investigated by several researchers. Sharma & Scrivener (1986) investigated the construction of voxel based representations. Svetkoff et al., (1984) segmented planar facets detected in range images for interfacing to CAD. Connolly (1984) obtained octree models from range data. Bogaert & Cornez

(1985) used segmentation of 3D data to interact closely with the acquisition system to provide extra scanning of regions of low confidence. This approach is similar to the active interaction planned in our synergistic system.

Of the representation schemes available, boundary representations appear to be best suited to processing rangemaps (Requicha & Voelcker, 1982). The processing of rangemaps takes three principle forms:

i) Building of a free form B-rep model of a sculptured surface using either manual interaction to provide form lines (see discussion in Schewe & Foerstner, 1986) or occluding contour information derived from other active stereo sources.

ii) Construction of a complete model where the data points have been segmented into primitives such as planes, spheres, cylinders and quadrics (Pratt,1987)

iii) Comparison of the perceived depth information with that predicted by analysis of an existing CAD model.

If the scope of a CAD representation scheme included the precise physical form of the surface being measured, the object could, in theory, be modelled exactly. In general, however, the uneven micro profile of the surface and noise attributable to the measurement system conspire to produce many differences between a CAD representation and the "true" physical form. Representation of these differences in 3D shape can be tackled in two ways: a) Goodness of fit metric - how well does the primitive fit the data. (This is required in the model building process to choose between the various possibilities) and b) incoporation in a CAD system that is able to include explicit tolerance information.

Certain dimensions may have very tight tolerances and representation will need the application of geometrical constraints so that the concept of tolerance zones can be explicitly conatined within the CAD system (see discussion in Requicha, 1983; Requicha & Chan, 1986). Conversion between CSG and B-rep systems will also need to be studied for model-based vision studies (see Xie & Calvert, 1988). Finally, the investigation of conversion between B-rep and finite element representation systems will need to be investigated if thermal and mechanical simulations are to be done using as input the geometrical description obtained from our CRVC.

5. CONCLUSIONS

A Close Range Vision Cell has been described which uses active stereo vision incorporating frame-transfer CCD cameras and a textured light projection system. Target specification parameters of this system are rms accuracies of 0.1mm with a depth-of-field of 500-850mm equivalent to rms disparity errors of 0.06-0.7 pixels. In order to achieve real-time stereo matching, special-purpose hardware has been designed and constructed for grabbing CCIR frames and inserting them into an array of transputers using the custom-built TEMIPS machine and subsequent automated stereo matching on this transputer machine. The difficulties entailed in object representation of range information in existing CAD systems has been discussed and possible solutions outlined.

Future studies will include the complete integration of the PARSYS™ SN1000 transputer array into a real-time loop involving a laser system to provide occluding contour and formline information coupled with a manipulation system for the creation of multiple views of an object's surface.

6. ACKNOWLEDGEMENTS

We would like to thank the UK Science and Engineering Research Council and the Department of Trade and Industry for support of Alvey MMI-137 and SERC for a research studentship (AA). MMI-137 is a university-led collaboration between the Departments of Photogrammetry & Surveying and Computer Science, University College London; THORN EMI Central Research Laboratories; Laser-Scan Laboratories Limited and the Royal Signals & Radar Establishment, Malvern.

7. REFERENCES

Ackermann, F., Schneider, W., Vosselman, G. (1986) *Empirical investigation into the precision of digital image correlation* Int. Arch. Photogrammetry & Remote Sensing, **26** (3/1):115-130

Anthony, A., Deacon, A.T., Muller, J-P (1988) *A Close Range Vision Cell for Direct Input to CAD Systems* IAPR Workshop on Computer Vision - Special Hardware and Industrial Applications, October 12-14, 1988, Tokyo.

Besl,P.J. (1988) *Active, Optical Range Imaging Sensors*, Machine Vision and Applications, **1**:127-152

Bieman, L.H. (1988) *Three-dimensional Machine Vision (Designer's Handbook)*, Photonics Spectra, May 1988: 81-92

Bogaert, M., Cornez, P. (1985) *Calibration and segmentation algorithms for an accurate and fast range-finder* SPIE Computer Vision for Robots **595**:214-224

Claus, M. (1987) *No-contact automatic 3D measurement of car bodies*, Zeiss INDUSURF™ background paper

Claus, M. (1988) *Experience with INDUSURF in 3D measurement of industrial surfaces* Int. Arch. Photogrammetry & Remote Sensing, **27** (5B)

Collins, K.A., Roberts, J.B.G. (1988) *Stereo matching of satellite images with transputers* Proc. IEEE Int. Conf. on Systolic Arrays:175-182, San Diego, May 25-28, 1988

Connolly, C.I. (1984) *Cumulative generation of Octree Models from range data* Atlanta GA USA 13-15 March 1984:25-32

Day, T. Muller, J-P (1988) *Digital elevation model production by stereo-matching SPOT image-pairs: A comparison of algorithms*, Proc. 4th Alvey Vision Conference:117-122. To appear in Image & Vision Computing (1989)

Doemens, G., Buerger, R., Goebel, W.W., Haas, G., Schneider, R. (1986) *A fast 3-D sensor with high dynamic range for industrial application*, Proc. ROVISEC-6 Conf.:197-207, Paris, France, August 1986

Gruen, A.W., Baltsavias, E.P. (1986) *High precision image matching for Digital Terrain Model generation*, Int. Arch. Photogrammetry & Remote Sensing, **25** (3):254-276

Higashimoto, A. (1984) *On-site 3-D Coordinate measuring function which assists CAD/CAM system*. Proc. 2nd Biennial Int. Machine Tool Technical Conference, Chicago. Illinois, 5-13 September, 1984.

Ikeuchi, K., Nishihara, H.K., Horn, B.K.P., Sobalvarro, P., Nagata, S. (1986) *Determining Grasp configurations using photometric stereo and the PRISM binocular stereo system*, Int.J. Robotics Research, **5** (1):46-65

Lougheed, R.M., Sampson, R.E. (1988) *3-D imaging systems and high-speed processing for robot control*, Machine Vision & Applications, **1**(1):41-58

Muller, J-P., Collins, K.A., Otto, G.P., Roberts, J.B.G. (1988) *Stereo matching using transputer arrays*, Int. Arch. Photogrammetry & Remote Sensing, **27** (B3):559-586. To be submitted to Machine Vision & Applications.

Nishihara, H.K. (1984) *Practical real-time imaging stereo matcher*, Optical Engineering, **23**(5):536-545

Otto, G.P., Chau, T.K.W. (1988) *A "region-growing" algorithm for matching of terrain images*, Proc. 4th Alvey Vision Conference:123-128. To appear in Image & Vision Computing (1989)

Pratt, V. (1987) *Direct Least-Squares Fitting of Algebraic Surfaces* ACM SIGGRAPH **21**(4):145-152

Requicha, A.A.G. (1983) *Toward a Theory of Geometric Tolerancing* Int. Jour. Robotics Research, **2**(4):45-60

Requicha, A.A.G., Voelcker, H.B. (1982) *Solid Modeling: A Historical Summary and Comtempory Assessement* IEEE Computer Graphics & Applications, **3**:7-29

Requicha, A.A.G., Chan, S.C. (1986) *Representation of Geometric Features, Tolerances, and Attributes in Solid Modelers Based on Constructive Geometry.* IEEE Jour. Robotics and Automation, **RA-2**(3):156-166

Schewe, H., Foerstner, W. (1986) *The program PALM for automatic line and surface measurement using image matching techniques*, Int. Arch. Photogrammetry & Remote Sensing, **26** (3/2):608-622

Sharma, A., Scrivener, S.A.R. (1986) *Constructing 3-D object models using multiple simulated 2.5D sketches*, Int. J. Man-Machine Studies, **24**:633-644

Svetkoff, D.J., Leonard, P.F., Sampson, R.E. (1984) *Techniques for Real-Time, 3D, Feature extraction Using range Information* SPIE Intelligent Robots and Computer Vision **521**

Xie, S-E, Calvert, T.W. (1988) *CSG-EESI: A new solid representation scheme and a Conversion Expert System*, IEEE Trans. Patt. Anal. Mach. Intell., **10**(2):221-234

Wrobel, B.P. (1988) *Least-squares methods for surface reconstruction from images*, Int. Arch. Photogrammetry & Remote Sensing, **27** (B3):806-821