The Active Stereo Probe : Dynamic Video Feedback

JPSiebert , CW Urquhart , DF Wilson, JP McDonald, PH Mowforth & RJ Fryer
The Turing Institute, George House, Glasgow, G1 2AD & University of Strathclyde, Department of Computer Science, Glasgow, G1 1XH

Abstract

This paper sets out to overview the progress of a binocular robot vision project titled the "Active Stereo Probe" (ASP) and present the novel aspects of an active controllable illumination system to be integrated with a robot vision head now under development.

1 Introduction

The purpose of the ASP development is to investigate the broad issues of stereo camera configuration and control for both active and passive robot vision. In particular, we wish to investigate algorithms to implement opto-motor reflexes such as the triple response, real-time algorithm issues including the use of foveated image tessellation structures, real-time vision executive issues, attentive search, intelligent high-level control and the integration of a complete system built from these components. We also intend to investigate the use of generalised structured illumination and section 3 of this paper sets out the basis of this illumination technique. Two differing approaches to robot head design have been adopted: active/passive head design based on the use of precision actuators, and passive anthropomorphic head design based on much faster but consequently less accurate actuators.

The principal goal of this research effort is to develop robot vision heads capable of actively interacting with, or at least responding to, the environment in which they are operating. Hence these robot heads should be able to orientate and verge, i.e. *self configure*, as required by some visual task and also undertake some degree of *self calibration*. The ASP robot head described below is one instance of one of the several demonstrators being developed in the effort to achieve the above goal.

2 Current Status

The ASP project has now been running for approximately 1 year and during this period we have developed a generic simulation software suite by which to model the stereo camera head geometry and the Active Illumination Projection System (AIPS) geometry. This modular simulation suite is capable of modelling arbitrary actuator configurations and is notable in providing a representation for the types of system errors exhibited by real actuators as specified by their manufacturers. The purpose of this simulation effort has been to explore the space of head design possibilities in order to converge on a realisable design for a given performance specification, to predict the implications of the chosen configuration with regards to algorithm design for tasks such as stereo matching and predict the impact of system errors on the final head performance. We expect to achieve millimetre depth error accuracy when the stereo cameras have been correctly verged to a single point at a nominal range of 5 meters and an extended working volume of $1.5 \times 1.5 \times 1.5$ meters. A complete description of the modelling suite is given in [1].

Development of a Real-time Sensory Virtual Prototype (RSVP) software executive has proceeded in parallel with the head design. The RSVP executive provides a software environment that supports fast prototyping, scheduling, chaining and analysis/data logging of real-time vision algorithms. A generic video/robotics device interface is also

BMVC 1991 doi:10.5244/C.5.56

supported to assist in maintaining the portability of the RSVP. System V UNIX provides the underlying process control and inter-process communication mechanisms of this execution environment. A fuller account of the RSVP is given in [2].

Detailed head/projection system design is now in progress and we hope to complete the fabrication of the precision head within the very near future.

3 The ASP and Dynamic Video Feedback (DVF)

3.1 DVF Background

An original goal of the background research that led to the ASP project was to develop a range image acquisition device using stereo imaging for human body metrology applications, e.g. human head modelling. However, the lack of resolvable texture on the human body made stereo matching difficult. This was overcome by projecting a gaussian texture field onto the subject as reported in [3] to yield dense range maps of the body surface by scale-space stereo signal matching techniques (MSSM, [4]) as reported in [5]. The projected texture spectral characteristics were subsequently tuned for a particular projector/camera configuration [6] to optimise the matching process. The original texture projection experiments utilised a conventional transparency projector to illuminate the scene. The logical extension was to replace the fixed pattern transparency with a LC Spatial Light Modulator (SLM) and thereby achieve computer control of the illumination pattern. Hence, the desired spectral response of illumination texture could be optimised dynamically in a closed loop fashion by examining the 2D spectral response of an acquired image using textured light, computing a "whitening" factor and then reprojecting this spectraly shaped pattern.

3.2 ASP Head Configuration for DVF.

The ideal configurations for mono (cyclopean) DVF and the extension to stereo DVF (requiring the use of 3 cameras) are shown in figures 1 & 2 respectively. The ideal cyclopean located beam spliced camera & projector combination share congruent fields of view & projection, hence the camera exactly views the image presented by the projector. For the purposes of stereo imaging, the stereo camera pair would in effect operate independently from the camera/projector processing loop. However, in our initial configuration we intend to dispense with the cyclopean camera and directly utilise one or both of the stereo cameras to complete the video feedback path. (This clearly complicates the processing as some degree of perspective shift/distortion must be corrected to maintain correspondence of the camera/projector world planes.) Accordingly, the ASP robot head configuration comprises a stereo pair of cameras mounted on azimuth and elevation precision actuators. An active projection system modulated by a Liquid Crystal (LC) array provides arbitrary spatial illumination patterns under computer control. In order to maintain this active illumination within the field of view of the cameras, the projection beam is steered by means of a mirror mounted on azimuth and elevation precision actuators and set between the cameras in the cyclopean position.

3.3 Single Cycle DVF

To generalise upon the concept of DVF, single cycle DVF can be thought of as the single application of a control cycle. In most basic form Dynamic Scene Interaction comprises the following sequence: illuminate the scene and simultaneously capture a *reference* image of the scene (strobe illumination with synchronised gated cameras is being employed); apply some image processing function to the captured reference image of the scene using a digital computer (this function is dependent on the information to be extracted, e.g.

image negation or texture whitening); re-illuminate or *overlay* the scene with the *modified image of the scene itself* and simultaneously capture a new image of this scene.

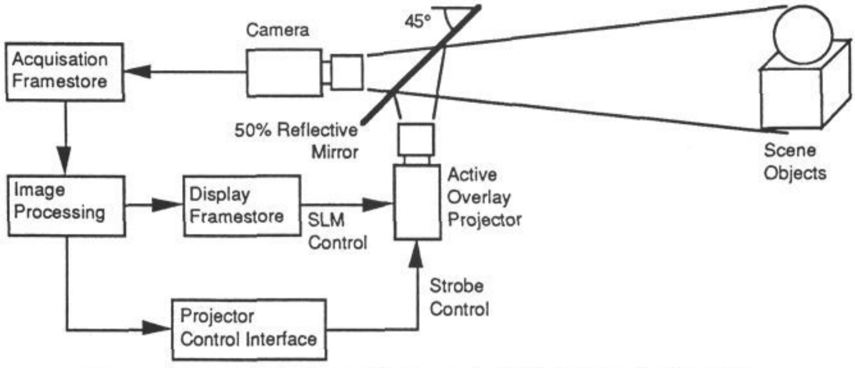

Figure 1: Ideal Mono (Cyclopean) DVF System Configuration

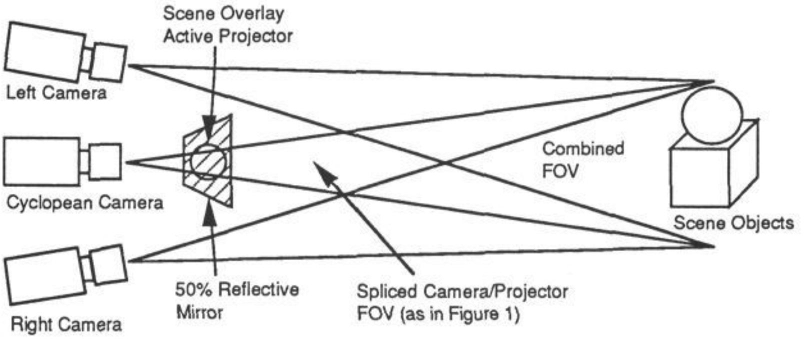

Figure 2: Ideal Stereo Head Configuration for DVF

This sequence essentially describes the basic cycle by which we propose to generate scene dependent texture overlay illumination. However, all active vision systems are susceptible to the effects of high image contrasts caused either by extremes of surface reflectivities or simply high angles of surface normal to the illumination source boresight. Hence all active vision approaches, e.g. grids or stripe based systems, could benefit from dynamic scene interactive cycles that capture a reference image under normal uniform illumination and use this to reduce or cancel the above image contrast effects by modulating the subsequently projected structured light pattern accordingly. The benefit is that such systems will be able to compute depth reliably over a greater dynamic range of reflectances in the scene.

3.4 Multiple Cycle Dynamic Scene Interaction

In order to cater for situations that involve large scale object motions within the scene or temporally changing scene components (e.g. surface reflectance changes) the cycle of activity must be modified to execute repetitively and thereby allow for the periodic reacquisition, modification and projection of a *reference* image.

Clearly, the Single Cycle sequence could be repeated to accommodate changing scene contents. If however, the overlay image of the previous modify-project-illuminate cycle is now utilised as the new reference, a rather more interesting situation arises. The acquired image of the scene illuminated by the modify-project cycle is then modified again and the cycle repeated. Hence the modified image will not result from a single cycle of processing but the entire sequence. Clearly, this *infinite* effect form of cycle can be used to establish *closed-loop dynamic video feedback.*

3.5 Closed-loop Video Feedback

Consider an application where an image of the scene has been sampled, negated and reprojected in order to highlight any object movement by imperfect cancellation. In the case of the finite effect cycles, this processing is simply repeated every cycle to display a continuously updated set of difference regions onto the surface of motive scene objects (and their immediate backgrounds).

Where infinite effect cycles are adopted, then rather than simply negate the input image, the *difference* between each pixel value of the input image and some desired (neutral grey) value can be computed and the new output image will be a function of this difference or *error* image. Hence control theory techniques can be employed to reduce the error value for each pixel on successive cycles. The form of the error control function could be chosen as desired, e.g. simple proportionality, three term PID, auto regressive etc. The benefit of this approach is that closed loop feedback iteratively drives the measured error to as close to zero as can be achieved within the constraints of the system.

Each camera/projector pixel thus serves as the input and outputs respectively of an array of closed loop feedback paths, each path having associated state variables stored in the computer that undertakes image modification.

4 Conclusions

We believe that the next generation of vision systems will comprise binocular robot vision heads capable of dynamic interaction with their environments by means of motive sensors. Accordingly our work centres on researching issues basic to robot vision head design and reflex control. The novel technique of DVF provides an optical feedback probe between the imaged surface properties and the vision head sensors through dynamic illumination spatial light modulation. Where use of active illumination is appropriate, we expect that DVF will be an important development for machine vision with implications for tasks such as depth ranging, surface reflectance estimation (monochromatic and colour), change detection and system calibration. Accordingly, many other applications for DVF are expected to follow as the technique matures.

References

[1] Urquhart C.W., McDonald J.P., Siebert J.P. and Fryer R.J. *Geometrical Modelling of the ASP Active Vision Head.* in preparation.

[2] Wilson D.F. Solbakken S. and Undbekken K. *Software Workbench for the Control of a Robot Vision Head.* in preparation.

[3] Siebert J.P. and Urquhart C.W. *Active Stereo: Texture Enhanced Reconstruction.* Electronics Letters, 26(7):427-430, March 1990.

[4] Jin Z.P. and Mowforth P.H. *A Discrete Approach to Signal Matching.* Research Memo TIRM-89-036, The Turing Institute, Glasgow, UK, January 1989.

[5] Mowforth P.H., Siebert J.P., Jin Z.P. and Urquhart C.W. *A Head Called Richard* Proceedings of the British Machine Vision Conference, pp361-366, University of Oxford, 24-27 September, 1990.

[6] Urquhart C.W. *An Investigation into Active and Passive Methods for Improving the Performance of Scale-Space Stereo.* MEng Thesis, BBN Systems & Technologies and Heriot-Watt University, Edinburgh, UK, March 1990.