Underwater Technology 2000 conference
23-26 May 2000, Tokyo, Japan

A STEREOSCOPIC RANGING SYSTEM USING STANDARD PC TECHNOLOGY

Nick Nelissen*, Rolf Koch* and Alec Duncan**

*Physics and Energy Studies Program,
Murdoch University,
South Street, Murdoch WA 6153 AUSTRALIA
E-mail: nelissen@central.murdoch.edu.au
**Australian Maritime Engineering CRC,
Curtin University of Technology,
Kent Street, Bentley, Western Australia, 6102.
E-mail: a.duncan@cmst.curtin.edu.au

Abstract - A stereoscopic ranging system is currently being developed as a key source of positional information for an underwater ROV Station Keeping System. Advancements in PC technology make it possible to use a relatively simple image capture card and a PC as a platform for the fast capture and processing of video images. We make use of the extensive capabilities of fast databuses and the high processing power of fast PCs with Pentium II or III processors. Using this approach we are developing an image processing system that is largely manufacturer independent and promises a good path for both hardware and software upgrading in the future.

I. INTRODUCTION

Underwater Remotely Operated Vehicles (ROVs) play a significant role in offshore petroleum production facilities. A ROV station keeping system is being developed by Curtin University Centre for Marine Science and Technology and Murdoch University. The aim of this system is to hold vehicle position in fluctuating underwater currents. A combined system using ranging information from stereoscopic TV images and an on-board inertial system provides positional information which is used as input to thruster control algorithms.

This paper describes the visual ranging system using computer processing of stereo video image pairs. The processing method basically consists of feature matching in consecutive stereo image pairs and extracting a 3D displacement vector which is subsequently processed by a station keeping system for thruster control. A first prototype has been developed in a previous project, using dedicated hardware. This system had a processing rate of 3 positions per second in good seeing conditions. The further development of the system is the goal of this project and aims at an improved stability of positional information as well as an increased position update rate.

Our first prototype system was based on a Matrox Image 640 system. The problems faced by a developer using such a proprietary system are twofold: (i) systems become quickly obsolete and new products use different architectures and hardware. (ii) image software packages cannot be moved to operate on equipment from other manufacturers. Off-the-shelf image processing software packages cannot be used for this project because these packages operate on image files and are slow, whereas we aim at “real-time” processing at a rate of several frames/sec. Most image processing, analysis and machine vision tools are also limited to the capabilities provided by the hardware of the same manufacturer. In order to provide an easier path to hardware upgrade and further software development we have investigated the capabilities of standard PCs. Central processors and data buses of modern PCs (using fast Pentium II or III and the PCI bus) have reached a stage where they have become serious competitors to specialised image processing hardware. We will show how a standard PC together with a relatively inexpensive standard image capture card can be used to process images at speeds that are adequate for many applications, such as the "real-time" tracking of features in stereo image pairs.

II. BASIC SYSTEM CONSIDERATIONS

Digital image processing is a two-step process, capturing the image and then processing it. The capture of an analog video frame consists of digitisation using an Analog-to-Digital Converter (ADC) followed by storing this digital data file. Data transfer directly to PC memory was slow in previous PC generations. Frame grabbers were used which consisted of a flash ADC connected by a fast bus to local memory. Images could then be transferred to the PC memory or hard disk at low rates, limited by the speed of the PC bus. Once the images were accessible to the PC they could be processed. The digitisation of even a single video frame results in a fairly large data file. Processing of this amount of data was slow. If computationally expensive calculations or the processing of large numbers of images was required, fast image processing systems were used with on-board fast processing architectures and fast data buses. Recent developments in PC architecture and processors have vastly improved computing power and data transfer rates. It is now possible to do serious image processing using a simple capture card interfaced to a PC.

The advantages of the generic PC approach can best be shown by a comparison with the dedicated hardware approach. Although dedicated systems often have the advantage of overall processing speed, it can be difficult to make best use of this. There are two possible cases where this arises.

In the first case maximum performance must be achieved from given hardware. Our application is an instance of this case. The processing must be done so as to provide real-time positional information. Software is not available off the shelf for this. Increased speed of processing has the effect of improving the reliability of our tracking algorithm and reduces synchronisation problems with the station keeping system. In situations such as this, compared to offline processing, delays degrade the system as opposed to just being annoying to the operator who waits in an offline situation.

To achieve the maximum processing speed, it is required that native code be written at the lowest level. Dedicated hardware, by definition, lacks portability. It is unlikely that code written for one manufacturer's hardware will run natively on another system. Furthermore these hardware systems often lack a clearly defined upgrade path. There is a reasonably probability that the manufacturer may discontinue the entire product line at some stage. This could result, at best, in heavy reworking of the software or at worst having to start from scratch. By using generic PC hardware, this danger can be avoided. The PC has an almost 20 year track record of physical backward compatibility. Code from the original 8086 can still be run successfully on the latest Pentium series, although of course much faster.

In the second case is where the ultimate speed of processing is not the most important requirement. This is often the case where photogrammetric work can be done offline from pre-recorded images. These may be single frames or movie sequences. The key point here is that the work is not done in real-time, although the task should not take so long as to be uneconomic or unrealistic. Delays here are annoying but not fatal and generally do not degrade algorithms, as they do in our real-time example. To facilitate this most manufacturers have created software libraries that allow a certain degree of portability between their systems. This level of abstraction allows manufacturers to change their hardware, while retaining backward compatibility in the consumer's code. A further application layer of code can be provide that eliminates the need for program development. This application layer may or may not be part of the manufacturer’s complete package. For example, Adobe Photoshop is independent of any particular hardware. This application layer software is not always available for real-time applications.

There are three major disadvantages with this approach. Firstly, there is a processing overhead involved in abstract software libraries, thus the optimum speed is usually not achieved. Secondly, the libraries only allow operations that the designers of the library have included. If operations are required that are not in the library, and this is often the case with innovative developments, then the non-portable, low level approach, with all its drawbacks, must be used. Thirdly, the libraries are only portable between the different hardware of only one manufacturer, thus none of the advantages of cross-platform compatibility are attained.

The generic PC approach has distinct advantages in both real-time and offline image processing.

With this approach almost any manufacturer's hardware can be used with minimal or no modification to the software. Usually the only difference will be in the selection of the video capture card and the appropriate device drivers. The capabilities and operation of these cards are very similar across the range as they basically all just deposit image data in the selected area of memory.

PCs have a clear upgrade path. According to Moore's law, "the processing power of a given chip should double every 18 months". It is highly unlikely that PC manufacturers will break with backward compatibility, thus current software should run faster and without modification on subsequent PC platforms.

A further advantage is that much of the support software has already been written. For offline processing there are packages available, such as Photoshop and GIMP (GNU image manipulation Program), that provide the user with interface, image viewing, storage and basic manipulation facilities. These basic facilities can be extended by the use of plugins which provide additional operations. Custom plugins can easily be written to provide innovative processing.

Real-time processing does not make as much use of these packages during actual operations. However, they are still useful for software development and testing. In our system we made extensive use of GIMP to approximate the results of various algorithms before any code was written.

III. THE SYSTEM

Our system is a Dell Poweredge 4200 server with dual PII333MHz processor, 128MB of RAM and dual PCI bus. It should be noted that typical desktop speeds are approaching 500MHz and 128MB RAM is now considered standard. There are two capture cards, one on each PCI bus. One capture card will be used to pipe stereo video directly to the VGA display for the operator and is not involved in the image processing. This aspect of the system is isolated from the rest by being on the second PCI bus. The advantage is that the user interface can be overlayed on the live image without requiring external video mixing as was the case previously. The other advantage is that the refresh rate is higher for the VGA monitor and thus disturbing flicker is reduced.

The actual image processing system is on the primary PCI bus. This consists of the capture card, the PCI-to-memory bridge, the processors and their on board caches. The only item that is not normal for a standard PC is the capture card. The stereo video input is interlaced in field sequential mode. This is separated by the capture card and dumped into linear memory areas. When a processor first accesses this memory, it is loaded into its onboard cache. Thus the processor can work on the data at high speed even while the capture card is dumping more data into memory. Typical cache sizes start at 256kB, this is sufficient to store the entire region of interest.

In our case we extended the generic approach to the software. This was written for the Linux operating system. This operating system started as a clone of Unix but has branched out into the real-time and embedded field as well. The open source nature of Linux allowed us to write a customised device driver to fully exploit the hardware.

IV. RESULTS

We use a convolution approach to feature tracking. This involves pre-processing an image to remove noise and extract stable edges. Then an image of the feature is overlayed and compared at each position in the region of interest and the best match is taken as the location of the feature within the region of interest. The region of interest is of course larger than the feature image. A typical region of interest is 128 X 128 pixels square with a feature size of 64 x 64 pixels. The size of these two images is critical, if it is too large then the algorithm runs too slow to follow the object, if it is too small then the object may exit the search region between iterations of the algorithm. The search region is moved to the centre of the last matched feature position, thus the search region follows the feature.

As a benchmark, this system tracks a 64 x 64 pixels squared feature in a 128 x 128 pixels search region at 7 frames per second and 48 x 48 pixels in a 96 x 96 pixels search region at 14 frames per second. This compares favourably with the original dedicated hardware system which tracked 64 x 64 pixels in a 128 x 128 pixels search region at only 2 frames per second. Currently only a single processor is used on the dual processor machine.

At 14 frames per second the system will track objects that move at up to 336 pixels per second. Thus the object can cross the camera's field of view, in approximately 2 seconds and still be tracked.

V. DISCUSSION AND FUTURE DEVELOPMENT

Some of the PC's advantages that we have not yet exploited are the inbuilt parallelism within the MMX series of processors and the use of multiprocessor motherboards. MMX offers the opportunity to increase speeds eight fold while still using only one processor. Pentium class processors are designed for multiprocessor motherboards and the major operating systems such as Linux and Windows NT support this.

PC-type components are readily available in full size industrial enclosures and embedded formats, thus the relative bulk and fragility of typical desktop systems is avoided.

ACKNOWLEDGEMENTS
The authors acknowledge project funding and the granting of an MPhil scholarship for one of the authors (N.N.) by the Australian Maritime Engineering CRC.