A HIGH RESOLUTION FULL-FIELD RANGE IMAGING SYSTEM FOR ROBOTIC DEVICES

D. A. Carnegie[1], M. J. Cree[2], A. A. Dorrington[2], and A. D. Payne[2]

[1]School of Chemical and Physical Sciences
Victoria University, Wellington,
New Zealand
Dale.Carnegie@vuw.ac.nz

[2]Department of Physics and Electronic Engineering
University of Waikato, Hamilton,
New Zealand
cree@phys.waikato.ac.nz
adrian@waikato.ac.nz

ABSTRACT
There has been considerable effort by many researchers to develop a high resolution full-field range imaging system. Traditionally these systems rely on a homodyne technique that modulates the illumination source and shutter speed at some high frequency. These systems tend to suffer from the need to be calibrated to account for changing ambient light conditions and generally cannot provide better than single centimeter range resolution, and even then over a range of only a few meters. We present a system, tested to proof-of-concept stage that is being developed for use on a range of mobile robots. The system has the potential for real-time, sub millimeter range resolution, with minimal power and space requirements.

KEY WORDS
Robot Sensing, Computer Vision, Full-Field Ranging

1. Introduction

The authors have developed a full-field image ranger that has application for mobile robotics and automated process machinery[1]. The authors are interested in augmenting the sensing capabilities of a number of mobile semi-autonomous robots designed for a variety of terrains developed by the Mechatronics Group of the University of Waikato[2]. Each of these robots contains their own power supply, processing system and intelligence. Currently they rely on a combination of laser scanning (1 – 8 m), infrared psd sensors (0.3 – 1.5 m) and whiskers (0 – 0.3 m) in order to localize, navigate and avoid obstacles. These robots are illustrated in Figure 1. Figure (1a) is MARVIN, an indoor security robot, differential wheelchair drive configuration, speed approximately 8 ms^{-1}; (1b) is the TANK, an outdoor multi-terrain robot utilizing self-laying tracks, maximum speed 0.5 ms^{-1}; (1c) are two identical robots to investigate cooperative robotic behavior (affectionately called "itchy" and "scratchy") – that employ a tricycle arrangement with rear-mounted steering wheel, typical speed 2 ms^{-1}.

Figure 1: Autonomous mobile robots developed by the Mechatronics Group (a) MARVIN, (b) TANK, (c) Itchy and Scratchy

Traditionally ranging systems for robotics or automated machine processes are classified as being either laser scanning or full-field (simultaneous) image acquisition. The former has advantages of high precision ranging and

x-y scanning, but as it is generally attained by physically moving a laser dot over a scene, it tends to take a considerable amount of time to scan the field of interest. It is commonly necessary to employ software interpolation and edge detection algorithms in order to obtain a clear image of the features in the environment.

Full field techniques do not suffer these time constraints and can be configured to operate in near real-time and do not suffer the inconvenience of requiring moving parts. However, such systems do incur a penalty in terms of exhibiting a significant loss of resolution and x-y positioning compared to the laser scheme. Full-field image ranging systems can also be classified by the ranging method used, the most common being direct time-of-flight measurement, AM/FM modulation, structured lighting, and time of flight measurements of picosecond pulses. Of these, the AM/FM modulation gives the best ranging resolution[3], primarily due to the limitations of the speed of operation of the electronic circuitry required for accurate time-of-flight measurement. A survey of full-field imaging systems [4, 5, 6, 7, 8] yields a best range resolution of 10 mm over a range of 2 to 3 meters. Beyond this, the resolution of these systems drops off inversely proportional to distance.

The authors present a novel full-field ranging system, suitable for mobile robots that incorporates standard off-the-shelf components, is portable, inexpensive, CPU non-intensive and capable of millimeter range resolution. Further, techniques for obtaining the range of every object in a standard camera lens' field of view (irrespective of whether it is telephoto or wide-angle) in real-time (sub 100 ms) are presented. The system has been proved to proof-of-concept stage, limitations identified, and modifications implemented for it to be installed as the primary long-range sensor on our fleet of mobile robots. This paper details the principles of the ranging system, presents results to date, and discusses the modifications underway for robotic implementation. These results will be updated during the paper presentation.

2. Review of Full-Field Rangers

As mentioned, the primary advantage of a full-field imaging ranger (compared to a laser scanner) is that it can acquire a full-field image in one measurement process in a comparatively short length of time and without mechanical scanning. Most full-field systems use pulsing or modulation methods to encode range into some other signal parameter that is easier to measure than a direct time-of-flight process. A generalized form of the hardware is illustrated in Figure 2.

The objects to be imaged are illuminated by a light source that is continuously modulated at some (high) frequency f_1. The reflected light incident on the ranger is delayed by a phase angle ϕ given by:

$$\phi = 2\pi f_1 \tau = \frac{4\pi f_1 d}{c} \qquad \text{Equation 1}$$

where $\tau = 2d/c$ is the time-delay of the signal, d is the range to the object and c is the speed of light. Most described full-field rangers (for example [7, 9]) use a homodyne configuration, in which the shutter is modulated with the same signal as, and in phase with, the light source. This causes the phase shift due to distance to be encoded into both a very high-frequency component (which is filtered out) and into the dc component. However, this dc component is contaminated by the reflectance properties of the object, the intensity of the transmitted light as received at the particular point of the scene, and by any ambient background light. Calibration images with the modulation switched off are needed to cancel out these extraneous intensity changes in the dc component. A more comprehensive review is provided in [3].

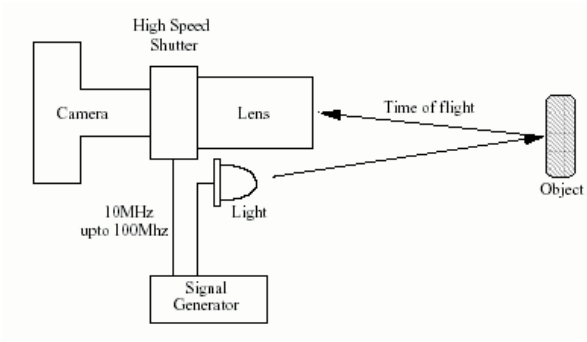

Figure 2: A Generalized block diagram of a modulation-based full-field range imaging system

In our system, the modulated light is incident on a high speed shutter that is modulated at some slightly different frequency f_2. The two signals mix to produce a substantially lower frequency signal ($f_1 - f_2$) that has ϕ encoded in it in a manner similar to the homodyne configuration, but without the disadvantages detailed above. This low frequency signal can be as low as single cycles per second, so (for example) a standard CCD camera with a frame rate of 30 frames per second (fps) can be used as the light sensing element, even though the typical shutter and illumination modulation rates maybe in excess of 100 MHz.

3. Our Range Imaging System

Our hardware implementation closely follows that described by Christie et al.[7] and Scott[9], with a bank of Agilent HLMP EL series LEDs modulated at radio frequencies illuminating the scene. A Photek MCP125 25 mm diameter, 1 MCP, image intensifier is used as the

high speed shutter. The image is acquired by a manual focus Nikon 80 – 200 mm focal length lens, and projected onto the photo-cathode of the image intensifier. An 8-bit Pulnix TM9701 CCD camera was originally used to view the phosphor screen of the image intensifier via a 25 mm focal length lens, although this was later changed to a Dalsa Pantera 1M60 12-bit camera for reasons described in section 4. Modulation rates between 10 MHz and 25 MHz have been investigated with this system, though modifications (refer section 3.6) are currently being implemented that will shift this bandwidth up to 50 MHz, and potentially as high as 100 MHz.

3.1 Signal Processing

A large number of algorithms exist to detect the phase of a sinusoidal signal[10]. Many, such as the phase-stepping algorithms used in profile measurement with fringe patters, are only suitable for pure uncontaminated signals. As the components in our system are inherently non-linear as well as noisy, an algorithm is required that can detect the phase of a single frequency in the presence of harmonics and noise. The Fourier method is particularly good at detecting the phase of a signal when significant noise is present. It provides an optimal estimate of the phase provided the noise is white and Gaussian and no harmonics are present[11]. If harmonics are present, as with the hardware described here, the Fourier method does not provide an optimal estimate, however it will still yield a reasonably accurate result[10] and consequently was selected for implementation in our system.

We apply a discrete Fourier transform (DFT) in time separately to each pixel and calculate the angle of the complex quantity in the bin of the Fourier spectrum corresponding to the beat signal. This gives the phase of the beat signal from which the range can be determined using equation (1). Our ranger produces video sequences sampled at 29.97 fps with a resolution of 768 × 484 pixels, although the image intensifier phosphor screen fills only about 55% of the camera's field of view.

As the camera's frame rate is not an integer multiple of the modulation signal frequency difference, it is likely that the heterodyne beat frequency will not precisely land on a single bin in the Fourier domain. We therefore have also analyzed the data by zero-padding the sequences to eight times their original length before Fourier transforming and then locating the peak in the magnitude of the Fourier spectrum. Quadratic interpolation was used on the bin corresponding to the peak and the two neighboring bins to better locate the peak to sub-bin accuracy. This frequency estimate was repeated over the 4000 or so pixels with the strongest signal in the sequence and an average taken to give a precise estimate of beat frequency. The phase was calculated for each pixel by linear interpolation of the phase between the two neighboring bins to the frequency estimate.

Currently our system suffers from an arbitrary offset error due to the 2π cycling of the returned signal. For example, using a 10 MHz gating signal, then from:

$$\lambda = \frac{c}{f} \qquad \text{Equation 2}$$

we find a 2π cycling of the range measurement over a return path length of 30 meters. The consequence of this is that with this single frequency, we would be unable to uniquely determine if an object was 1 meter away, or 16 meters way, there being an unknown distance bias of 0, 15, 30, …., n × 15 meters. Relative distances within this 15 meter window can however, be precisely determined. The implementation of a second modulation frequency (section 5) will eliminate this problem, however, in the interim, a reference block is situated at a known distance. A region of interest (ROI) of area between 1600 and 2500 pixels has been used as this reference block for the initial experiments and the relative range between this block and the objects to be imaged calculated.

3.2 Hardware Configuration

As explained previously, a Photek MCP125 image intensifier is used as the high speed shutter. This image intensifier is a three terminal device comprising a low voltage photocathode, a single microchannel plate (MCP) and phosphor screen. To use the image intensifier as a fast electronic shutter, a voltage pulse is applied to the photocathode. The on/off gain/attenuation ratio is typically greater than 10^{10}, ensuring that an efficient optical block is created. Specifications for the device indicate that a modulation frequency of 100 MHz could be attained, though to date, we have only tested our system up to 25 MHz. A better image resolution with less noise and higher contrast can be obtained by gating the microchannel plate (MCP) rather than the photocathode, however this is problematic at our frequencies of interest as this typically requires a voltage swing of several hundred volts.

The modulation frequency for the image intensifier is of a similar frequency to the illumination source, but vitally, the frequency difference (a few Hz) must not drift. For initial development, two frequency synchronized Agilent 8648B signal generators were employed. However, for implementation on the mobile robotic devices a custom signal generator board was designed.

3.3 Signal Generator Board

A first version of this board utilized two AD9852 Analog Devices digital synthesizer chips interfaced to an Atmel AT89LS8252 microcontroller[3]. These digital synthesizer chips use DDS (direct digital synthesis) technology and a high-speed, high-performance D/A converter to form a digitally programmable agile

synthesizer function. When referenced to an accurate clock source, the AD9852 generates a highly stable, frequency-phase-amplitude-programmable cosine output. Up to 48-bit frequency resolution is provided, so (for example) 1 microHertz tuning is possible with a 300 MHz SYSCLK. Possible output frequencies range from dc to 150 MHz (1/2 SYSCLK).

A significant disadvantage of the '9852 chips is their considerable power dissipation (922 mA max per chip). A subsequent version of the signal generator board was constructed that uses three Analog Devices AD9952 digital frequency synthesizer chips. These ICs offer increased bandwidth over the AD9852, operate at a lower voltage, and importantly, have only 10% of the power dissipation. This is achieved at the expense of functionality, many of the '9852's output options are not provided, and the frequency resolution is reduced to 32 bits. However, the AD9952 includes synchronization inputs and outputs, allowing three of these devices to be synchronized together.

Two of these chips output f_1 and f_2 as before, a third chip outputs a synchronized pulse at an integer multiple of the beat frequency δ that initiates the frame grabbing. This will ensure that the beat frequency is a known bin in the FFT. The revised board uses only a few tenths of an ampere, has dimensions 165 mm × 105 mm, and hence can easily be accommodated in most robotic devices.

3.4 Power Supply

As well as the modulated photocathode, the image intensifier requires approximately 750 V applied to the microchannel plate, and 5.5 kV to the phosphor screen. For initial testing, a dedicated bench power supply has been constructed. This is being miniaturized for portability utilizing EMCO C01 (0 – 100 V, 10 mA), C10 (0 – 1000 V, 1 mA), and C60 (0 – 6000 V, 0.166 mA) modules, with indicative dimensions of 50 mm × 30 mm. The robots of section 1, have two 12 V FLA batteries wired in series to provide 24 V dc and are connected to an ACE-828C industrial ATX power supply. Currently the motors draw 10 – 12 A continuous and 17 – 20 A peak current, and the on-board computer system and other electronics draw approximately 5 A continuous and 10 A peak. To power the image intensifier at the voltages of interest, the EMCO modules will consume a total maximum power of 1.7 W, an insignificant percentage of the robots' existing power consumption.

The camera system also draws minimal current. Although the existing camera runs from mains (230 V ac), either a dc powered camera could be employed, or the ACE power supply swapped for a standard UPS/230 V ac supply (that incidentally was initially employed to power the robots' PC board before the ACE was implemented). The modulated LEDs draw approximately 30 mA from a 24 V source and consequently power supply issues are not a limiting factor in the implementation of this system.

3.5 Processing Issues

Each of the robots described in section 1 currently host an AMD 64 bit 3000+ processor, 512 MB RAM and a 160 GB HDD. Windows XP operating system is employed, and Visual C, Matlab and LabVIEW are installed. We reserve a maximum of 256 MB for the image acquisition, which permits the acquisition of 334 frames (at 768 × 484 resolution) i.e. slightly over 11 seconds at 30 fps. These 11 seconds are well in excess of the minimum sample time that could be tolerated on the mobile robots. However, the longer the acquisition time, the higher the range resolution. Ideally a full-field acquisition would not take more than 300 ms in order for the information to be of use for obstacle avoidance for a robot traveling at a slow walking speed (1 ms^{-1}). This is discussed further in section 4.

3.6 Other acquisition issues

The camera and the lens systems are off-the-shelf components, are readily available, and in the case of the lens system, easily interchangeable if a wider-angle or telephoto image is required. A fundamental issue is the illumination source. With the first implementation of the modulated illumination source taking the form of LEDs, the range of the system is limited to approximately 10 meters (depending upon the reflectivity of the imaged objects). Also the system is constrained to operating in the absence of ambient background light due to the potential of affecting burn-in on the phosphor screen of the image intensifier.

We are investigating filtering for the LED illumination and expect that we will be able to operate the device in the presence of fluorescent lighting. A second version of the illumination source is currently under construction utilizing laser diodes. These laser devices have the advantage of increased range (potentially 30 + meters), and being of narrow bandwidth, a selective filter can be placed in front of the image intensifier allowing operation in naturally illuminated environments.

4. Results

The excellent linearity (0.99996 with a norm of residuals = 0.0566) of the system for objects between 1 and 5 meters distant is illustrated in Figure 3 for a 10 MHz modulation frequency and a beat frequency $\delta = (f_1 - f_2)$ of 1 Hz.

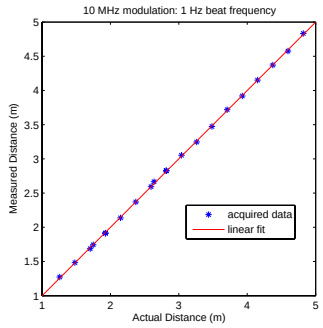

Figure 3: Measured distance vs actual distance for the full-field image ranging system

Using the Pulnix camera, the standard deviation of the error between measured range and actual range is between 1.2 and 2.6 cm depending upon the value of δ, for measurements performed over a 10 second time frame. Padding the data by a factor of eight prior to Fourier analysis did not significantly alter the determined range.

From equation (2) it can easily be determined that an increase in the modulation frequency should result in a proportional increase in the range resolution of the imaging system. Until the laser diode illumination source is completed, we are limited to a bandwidth of approximately 25 MHz as the LEDs can not be effectively operated at higher frequencies. Figure 4 illustrates the improvement in range resolution when the modulation frequency is increased from 10 MHz to 20 MHz.

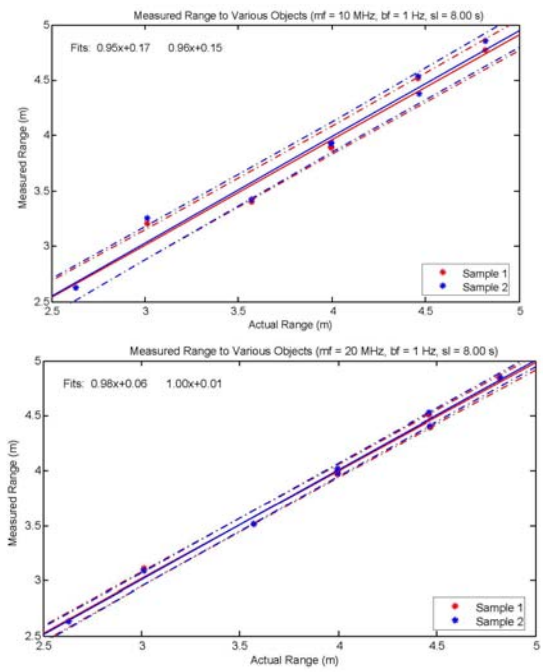

Figure 4: Range resolution vs modulation frequency (a) 10 MHz (b) 20 MHz

It should be noted, that although an 8 bit digital camera is employed, the maximum intensity over the region of interest is generally 4 bits or even as low as 2 bits, depending upon the color and range of the imaged object. Traditional homodyne techniques would be completely unable to resolve the range for such objects.

This loss of bit resolution does however, have a dramatic effect on the range resolution. An analysis by Zhao and Surrel[12] demonstrates that quantizing the intensity data to 6 bits will result in an approximately five times increase in the standard deviation of the phase error compared to an 8 bit quantization. This is comparable to the error margins obtained in our results. An increase in the number of bits to 10 reduces this error by more than an order of magnitude (compared to a 6 bit system), and our calculations indicate that at a 12 bit dynamic range our system should have a sub-millimeter precision (for a 100 pixel averaged region). Practically however, at this resolution issues such as the geometrical arrangement of the illumination and camera system, the non-linearity of the irising effects, and radial and decentering distortions in the lens system would need to be taken into account.

The implementation of the 12-bit Dalsa camera at modulation frequencies of 10 MHz and 25 MHz yielded the results summarized in Table 1. These were obtained by imaging 9 objects in the camera's field of view between 1.1 and 5.2 meters distant from the lens.

δ (Hz)	Standard Deviation (cm) 10 MHz	Standard Deviation (cm) 25 MHz
1	1.4	0.5
7	1.5	0.4
11	1.4	0.4

Table 1: Standard deviation of difference between measured and actual range for 10 second sample of 10 MHz and 25 MHz modulation frequency.

The acquisition time is a function of several parameters. Sufficient frames must be acquired in order to adequately resolve the phase shift. With a standard camera operating at 29.97 fps and δ = 1 Hz, at least one second of data must be acquired. If the beat frequency is increased (to say 5 Hz), then this time could in principle be reduced to 200 ms. However, reducing the number of frames does detrimentally affect the resolution, although this can be somewhat compensated for by averaging over adjacent pixels.

Trials were conducted using the Dalsa camera operating at 97 fps, incorporating 2 × 2 binning with acquisition times of 1, 2 and 5 seconds. These results are summarized in Table 2 that presents the results of 4 acquisition times for δ = 7 Hz, at 10 MHz and 25 MHz modulation frequencies. As Table 2 illustrates, sub-centimeter precision was achieved for a 2 second acquisition time for the 25 MHz modulation frequency. It

is expected that with better light coupling between the image intensifier and the camera system, that this level of precision could be achieved for 300 ms samples.

Acquisition Time (seconds)	Standard Deviation (cm) 10 MHz	Standard Deviation (cm) 25 MHz
1	4.1	1.2
2	3.4	0.9
5	2.0	0.5
10	1.5	0.4

Table 2: Standard deviation of difference between measured and actual range for various acquisition times and modulation frequencies.

5. System Improvements

A fundamental limitation to the data acquisition time (as well as the range resolution) is the amount of light being captured by the camera. Increasing the illumination strength (whilst still ensuring eye safety) or improving the coupling between the image intensifier and the camera (for example by using direct fiber coupling to the CCD element of the camera) would significantly enhance the system's performance. Indeed, we are exploring commercial options of collaboration with a chip designer to directly incorporate a high speed shutter into an imaging element with the potential of attaining the millimeter resolution in considerably less than 100 ms.

Finally the range ambiguity discussed in section 3.1 can be simply resolved by using two modulation frequencies. With the integrated signal generator board introduced in section 3.3, all that is necessary is for the PC to communicate with the Atmel embedded controller that a change of frequency is required. The controller then sends out a revised 32 bit frequency to the DDS chips, and the modulation frequency of the illumination source and high speed shutter will be altered several microseconds later. Whilst we have not tested this on the fixed bench signal generators, we do not anticipate any problems implementing this code for the mobile robots.

6. Conclusions

In this paper we have presented a novel full-field image ranging system employing a heterodyning technique that has substantial application to those fields of automation and robotics that rely on some form of image ranging. Without accounting for the geometry of the lighting system, irising effects or lens calibration, sub-centimeter precision has been demonstrated over a range of 1.1 – 5.2 meters. With the improved illumination source and filtering proposed, it is expected that this system will be able to operate in normal light to a range of 30+ meters.

References:

[1] A. A. Dorrington. Range sensing system, 2003. New Zealand Patent 525241.

[2] D. A. Carnegie. Towards a Fleet of Autonomous Mobile Mechatrons. In Proceedings of *the Second International Federation of Automatic Control Conference on Mechatronics Systems*, Berkeley, Ca, 2002, 673 – 678.

[3] D. A. Carnegie, M. J. Cree, and A. A. Dorrington. A high resolution full-field range imaging system. R*eview of Scientific Instruments*, *76*, (2005).

[4] http://www.3dvsystems.com/products/zcam.html

[5] G. Yahav and G. J. Iddan, USA Patent No. 6,100,517 (2000).

[6] T. C. Monson, J. W. Grantham, S. W. Childress, J. T. Sackos, R. O. Nelums, and S. M. Lebien. Characterisation of scannerless ladar. *Proceedings of the SPIE* 3707, 409 (1999).

[7] S. Christie, S. L. Hill, B. Bury, J. O. Gray, and K. M. Booth. Design and development of a multi-detecting two dimensional ranging sensor. *Measurement Science and Technology*, *6*, 1301 (1995).

[8] P. Gulden, M. Vossiek, P. Heidi, and R. Schwarte. Novel opportunities for optical level gauging and 3-d imaging with the photoelectronic mixing device. *IEEE Transactions on Instrumentation and Measurement*, *51*, 679 (2002).

[9] M. W. Scott. Range imaging laser radar. USA Patent No. 4,935,616 (1990).

[10] P. O'Shea. Phase measurement. *Electrical Measurement, Signal Processing and Displays*, chapter 5 (CRC Press, Boca Raton, FL., 2003).

[11] D. Rife and R. Boorstyn. Single-tone parameter estimation from discrete-time observations. *IEEE Transactions on Information Theory*, *20*, 591 (1974).

[12] B. Zhao and Y. Surrel. Effect of quantisation error on the computed phase of phase-shifting measurements. *Applied Optics*, *36*, 2070 (1997).