Toward 1-mm depth precision with a solid state full-field range imaging system

Adrian A. Dorrington, Dale A. Carnegie, Michael J. Cree

Dept. of Physics and Electronic Eng., Univ. of Waikato, Private Bag 3105, Hamilton, New Zealand
School of Chemical and Physical Sciences, Victoria Univ. of Wellington, PO Box 600, Wellington, New Zealand

ABSTRACT

Previously, we demonstrated a novel heterodyne based solid-state full-field range-finding imaging system. This system is comprised of modulated LED illumination, a modulated image intensifier, and a digital video camera. A 10 MHz drive is provided with 1 Hz difference between the LEDs and image intensifier. A sequence of images of the resulting beating intensifier output are captured and processed to determine phase and hence distance to the object for each pixel. In a previous publication, we detailed results showing a one-sigma precision of 15 mm to 30 mm (depending on signal strength). Furthermore, we identified the limitations of the system and potential improvements that were expected to result in a range precision in the order of 1 mm. These primarily include increasing the operating frequency and improving optical coupling and sensitivity. In this paper, we report on the implementation of these improvements and the new system characteristics. We also comment on the factors that are important for high precision image ranging and present configuration strategies for best performance. Ranging with sub-millimeter precision is demonstrated by imaging a planar surface and calculating the deviations from a planar fit. The results are also illustrated graphically by imaging a garden gnome.

Key words: solid-state, heterodyne, full-field, imaging, three-dimensional, range finding, range imaging, surface profiling.

1. INTRODUCTION

1.1 Three dimensional imaging

The acquisition of three-dimensional images has been the subject of research and development for some time with a great variety of proposed methods and a wide variety of applications. Some general application examples include machine vision, surface profiling, surveying, metrology, real-time multi-media integration, and three-dimensional recordings of objects and artifacts, each having unique requirements for measurement precision, speed, and coverage. Methods for three-dimensional image capture, or range imaging, can be divided into the four broad categories of laser scanning, stereo vision, structured light, and imaging lidar systems, all of which will now be briefly summarized.

Laser scanning systems utilize a single point laser range finder that is mechanically scanned across a scene and distance for every measured point is recorded in sequence. Although very good measurement precision and a large number of measurement points are achievable with these systems, the acquisition time (in the high precision modes) can be very long. Such systems are numerous both in the literature, and in models commercially available.

Stereo vision, on the other hand, relies on two or more images of a the scene to be acquired with standard (usually digital) photographic or video cameras from different angles and processed with triangulation techniques to determine three dimensional locations. Implementations of stereo vision systems vary considerably from two camera low precision natural scene processing systems mimicking human vision to multi-camera high precision photogrammetric metrology systems for scientific analysis.

Structured light systems also use standard cameras and triangulation principles, but can determine three-dimensional shape with only one camera and a known illumination pattern. There is also a wide variety of implantations of this...
approach ranging from simple single line illumination patterns, to full field complex illumination patterns, and high precision Fourier based processing for scientific applications\(^1\).

All of the first three categories have been well researched in the past and are commercially available in numerous configurations. However, the final category of imaging lidar is comparably in its infancy with a relatively small number of systems demonstrated with varying degrees of success\(^5\text{-}12\). Although there are a variety of implantation methods, all of these types of systems have common operating principles and hardware configurations. This operating principle is the expansion of time-of-flight single point laser range finders to operate simultaneously over a full field of view. The result is an image similar to a digital photograph, but instead of brightness and color information, each pixel contains distance (or range) information to the objects in the field of view. Some systems also provide functionality to augment the traditional intensity and color information as well. The Image Ranging system presented later in this paper falls into the imaging lidar category, so the principles of imaging lidar will now be described in more detail.

### 1.2 Imaging lidar systems

Conceptually, the basic components, configuration, and operation of all imaging lidar systems are the same, as illustrated in figure 1. An illumination source emits light that is scattered by objects in the scene and collected by a camera system. This illumination is intensity modulated in one of a variety of ways, and the light collected by the camera is also intensity modulated by a high-speed (non-mechanical) shuttering mechanism. The major differences between imaging lidar systems arise in the modulation control signals used for the light source and the shuttering mechanism, and the implementation of the shuttering system itself. The most common shutter implementation is an image intensifier with high-speed photo-cathode modulation capability\(^5,10\). Image intensifiers have a number of disadvantages, so there is a move to integrate high-speed shuttering capabilities in custom image sensors. A number of these types of sensors have been demonstrated, but they are currently limited in spatial resolution\(^11,12\). The types of modulation philosophies can be grouped into the three main categories of Pulsed, Homodyne, and Heterodyne, which will now be explained in more detail.

![Figure 1 – Imaging lidar principle](image)

In the pulsed systems, both the illumination source and the high-speed shutter are controlled with a single pulse in the nano-seconds region\(^5\). Consider the situation where both the illumination and shutter are pulsed simultaneously. The scattered illumination light pulse arriving at the camera is delayed in time due to propagation delay. If scattered from a nearby object the time delay will be minimal and almost all of the received light will enter the camera while the shutter is open, resulting in a bright image. Light scattered from a distant object, however, will arrive significantly later and will be “cut-off” by the shutter resulting in a darker image. Therefore, the brightness of a pixel is correlated with range. Depending on the situation, the shutter pulse can be delayed from the illumination pulse (and its width adjusted), to customize the sensitivity to specific depth regions.
Homodyne systems are similar in principle to pulsed systems, but the illumination and shutter modulation signals are a continuous square, sinusoidal, or triangle wave generally in the 10 MHz to 100 MHz region\textsuperscript{7,10}. In this situation, the propagation delay to the objects in the scene is manifested as a phase change to the illumination modulation envelope. The modulation signal of the light entering the shutter can be considered as being mixed with the shutter modulation signal. Because the illumination and shutter signals are of the same frequency, the mixing results in a DC brightness level for each pixel corresponding to range. Some decoding is required to derive actual range values from brightness, and often multiple measurements are performed at different relative phases for the modulation signals to perform quadrature or phase-shift keying (PSK) type decoding. Because of the respective nature of phase change, range ambiguities occur at multiples of half of the wavelength of the modulation signal. These ambiguities can be resolved by performing two measurements with differing modulation frequencies.

Heterodyne systems are different from homodyne in that the modulation frequencies applied to the illumination and shutter differs slightly in frequency\textsuperscript{10,13}. This means the mixing process at the shutter produces a beat frequency equal to the difference between the modulation frequencies. The phase of the collected light’s modulation envelope (which is proportional to propagation delay and hence range) is also encoded in the beat signal. The image captured by the camera therefore appears to be “flashing”, and objects at different distances will flash at different times. Range information for each pixel can be determined by acquiring a video sequence of the scene and calculating the beat signal phase or frequency (over time) for each pixel. The heterodyne approach (with phase detection) is used by the Range Imager described latter in this paper, so it will now be explained in more detail.

1.2 Heterodyne imaging lidar

Heterodyne systems generally achieve a higher range precision than both the pulsed and homodyne approaches. This is basically because pixel brightness level encodes range for both the pulsed and homodyne systems. However, other effects such as object color and background lighting also affect pixel brightness and can contaminate the range results. This is usually compensated for by imaging the scene with no illumination to find background offsets, and with un-modulated illumination to find object reflectivity values independently from range measurement. These corrections are straightforward and effective, but act to reduce the dynamic range available in the image sensor, thereby limiting the range measurement resolution and precision.

Because the heterodyne approach determines range from analysis of time varying intensities independently from absolute or average intensity, the range determination is not (directly) affected by background lighting and object colors. However, signal to noise ratio does impact on range precision, so excessive levels of background lighting and high contrast objects can have an indirect influence on precision. Furthermore, the heterodyne method is much less sensitive to dynamic range limits than the homodyne or pulsed methods.

The simplest heterodyne implementation uses fixed modulation frequencies with a known difference, and relies on measuring the phase change of the beat signal to determine range. Another more complicated example is Frequency Modulation Continuous Wave (FMCW), where the modulation signals are swept in frequency (or chirped) causing a varying frequency beat signal depending on the propagation delay of the illumination\textsuperscript{8}. In this case the frequency of the beat signal is analyzed to determine range. Because the Range Imager described later in the paper uses the fixed frequency method, this method will be the focus from here on.

A further advantage of the fixed frequency heterodyne approach is the capability to perform multiple measurements simultaneously. From a given beat signal, phase is used to determine range as explained above, but also the amplitude of the beat signal can be used to find object brightness, and a time average of the beat signal will reveal background lighting conditions. Furthermore, by superimposing additional modulation frequencies on the illumination (and possibly on the shuddering signal) with different beat frequencies, multiple beat signals can be created. These multiple signals can be processed individually to perform ambiguity resolution and if multiple red, green, and blue illumination sources are used, generate color range images.

Digital video cameras are used to capture the flashing beat signal image in heterodyne systems, but they are generally limited to between 10 Hz and 100 Hz maximum frame rate. Therefore, the beat signal (and modulation frequency difference) must be limited to the tens or singles of Hertz (or can even be sub-Hertz) to stay within the Nyquist sampling
criteria. Stability is therefore important when generating two stable modulation frequencies with a difference of 7 or 8 orders of magnitude lower than the base frequency. Such stability is generally achieved by employing frequency locked signal generators. It is also useful to generate a camera synchronization signal from the same set of frequency locked signal generators, and thereby ensure that the sampling video frame clock is also synchronized, and the beat signal lies exactly on a Fourier transform bin allowing simplification of the phase determination process and high-speed processing.

1.3 The University of Waikato Range Imager

In a previous publication we detailed a novel heterodyne imaging lidar system dubbed the “Range Imager”. This system comprises of a bank of modulated LEDs for illumination, a photo-cathode modulated image intensifier for image gain modulation, and a monochrome digital video camera for image acquisition. A frequency-locked signal generator is used to provide a 10 MHz drive for the LED modulation and a 10.000001 MHz drive for the photo-cathode modulation. The result at the image intensifier output is an image containing a 1 Hz beat signal, which has a phase component proportional to the time-of-flight of the light from the illumination light source, to the object, and back to the image intensifier. After image sequence acquisition, the phase of the beat signal for each image pixel is determined with Fourier analysis, from which the distance to the object for each pixel can be calculated from knowledge of the modulation signal’s wavelength (and speed of light).

With this system a one-sigma range determination precision of 15 mm to 30 mm (depending on signal strength) was demonstrated. As the goal of 1 mm precision was not achieved, the performance limitations of the system were identified and improvements to overcome these limitations were recommended. It was expected that these improvements would result in a range precision approaching our goal of 1 mm. The two primary limitations are low dynamic range, and low modulation frequencies.

Low dynamic range in the image can cause quantisation related phase determination errors. It was proposed that by upgrading the camera and intensifier to camera coupling the dynamic range of the beat signal could be improved from 5 to 11 or 12 bits. A more sensitive 12-bit camera system (rather than 8-bit as previously used), and better intensifier output to camera coupling should therefore result in almost an order of magnitude improvement in precision.

Furthermore, low modulation frequencies limit the range precision because range is derived from wavelength. Increasing the modulation frequency was expected to lead to a proportional improvement in precision (for a fixed phase discrimination precision) as the wavelength decreases. With new equipment, we expect to be able to operate the system between 50 and 100 MHz, leading to a 5 to 10 times improvement in ranging precision. This involves replacing the LED illumination source with laser diodes that can be modulated at significantly higher frequencies, and replacing the image intensifier and driver electronics to allow higher frequency photo-cathode modulation, and hence image gain modulation.

In the remainder of this paper, we will report on the implementation of these improvements and show the results of systems characterisations experiments of the upgraded systems. Using ranging results from a flat panel we demonstrate sub-millimetre ranging precision, achieving and exceeding our goal by approximately a factor of two. We also show examples of range images obtained with the system using our standard test article, “Stumpy” the garden gnome.

2. IMPROVED HARDWARE SET-UP

Although the principle of operation and general system configuration remains unchanged, compared to the previously published hardware set-up, the system has been effectively re-built “from the ground up”. This included replacing the digital video camera and PC frame grabber hardware, image intensifier, intensifier driver circuit, illumination source, and upgraded synchronized signal generators. A photograph of the improved hardware set-up is shown in figure 2.

The video camera used previously was deficient in several respects, but most importantly dynamic range and frame rate. There are a number of commercial options for digital video cameras with a high dynamic range of 12 bits of more, and also a number of options for cameras with a high frame rate around 60 Hz. However, there are very few options
available with both of these features present. We decided on the Pantera TF 1M60 (DALSA, Waterloo ON, Canada) camera as the best suited at the time.

Unfortunately the image intensifier used as the high-speed shutter in the previously published system was both limited in operation speed and had suffered some burn-in damage. As identified in the previous publication, a higher modulation frequency results in proportionally better precision. Therefore a replacement image intensifier was required. The new intensifier was purchased (by special order) from Photek Ltd (East Sussex, United Kingdom), who also manufactured the previous intensifier. This new intensifier is almost identical to the previous one except that it has a response almost an order of magnitude faster, and is capable of being photo-cathode gated down to 1 ns. Although a smaller size would have been easier to electronically drive, we continued with the 25 mm diameter form-factor because of ease of lens coupling the video camera. Lens coupling provides flexibility in the set-up, but better results could be achieved with a smaller size intensifier fiber-optically coupled to the camera.

In addition to upgrading the intensifier, the intensifier driver circuit was also upgraded. Previously, the photo-cathode drive signal was in the order of 5 V peak-to-peak. Although this level provided near linear modulation capability, the low bias level of the photo-cathode relative to the micro-channel-plate (MCP) did not allow for maximum image quality. Normal operation bias levels are in the order of -200 V and provide sufficient acceleration of the photo-electrons from the photo-cathode to the MCP, however, lower levels of bias can cause the photo-electronic to “wander off course” resulting in a blurred image. The new drive circuit (designed in-house) is capable of delivering a 50 V peak-to-peak modulation drive at up to 90 MHz.

Increasing the operational speed of the image intensifier by itself does not provide a significant advantage because the previous illumination source was also limited in maximum frequency, and therefore the illumination source also needed to be upgraded. The previous LED lighting could only be modulated efficiently to 10 MHz, and up to 20 MHz with reduced efficiency. A new laser-diode based illumination module has been fabricated that is capable of digital intensity modulation up to frequencies in excess of 100 MHz (beyond the upper limit of the intensifier driver), and provides capability for up to four laser-diodes to be operated in parallel. This module is currently fitted with two ML120G21 laser-diodes (purchased from Thorlabs, Newton NJ, USA) and can provide a peak output of 320 mW in total.

Figure 2 – Photograph of the range imager set-up showing the intensifier camera, LED illumination board, power supplies, and synchronized signal generator.
Because the new video camera is capable of external asynchronous triggering, the signal generator was modified to produce a third frequency-locked output to drive the camera acquisition frame rate. This allowed software configuration of the modulation and camera frame rate frequencies, but more importantly, provided that capability to ensure that both an integer number of beat signals and an integer number of video frames were present in the acquisition period. Consequently, it can be ensured that the beat signal will fall exactly on a Fourier bin simplifying the phase determination algorithm and speeding up the processing.

3. SYSTEM PERFORMANCE

A number of system characterization experiments have been undertaken to determine the improvement in performance over the previously published system. The results from these performance improvement experiments fall into the same two categories of performance limitations identified in the previous publication summarized in section 1.3 above. For consistency, all measurements presented in this paper consist of a 10 second acquisition at 1 Hz beat frequency and a 29 Hz frame rate.

3.1 Improved dynamic range

Upgraded dynamic range through camera and intensifier to camera coupling was expected to significantly improve system performance. The previous camera used an 8-bit digitizer, however, limitation due to intensifier to camera coupling in that system limited usable dynamic range to 5 bits. The new camera has a 12-bit digitizer and a sensor with higher efficiency. Also, the new camera is fitted with a lens system with a lower f-number and a closer operational distance, and as a result, the camera sensor now saturates before the intensifier output, and the full dynamic range of the camera is usable. From calculations performed in the previous publication it was expected that approximately an order of magnitude improvement in measurement precision would be achieved with a 12-bit camera.

An experiment was performed to find the ranging precision when measuring a white flat panel object. The precision was determined by fitting a plane to a 20-by-20 pixel region of the range image (around the area of the image with the strongest beat signal), and determining the standard deviation of the measurement errors between the fitted plane and the actual measurements. These results, graphed in figure 3, show the new system has a precision of approximately 1.5 mm, thereby achieving the order of magnitude improvement over the previous system that demonstrated at best 15 mm precision (for approximately equivalent single frame integration times of 30 ms).

However, all of this improvement can not be solely attributed to a larger dynamic range. Figure 3 also graphs the precision obtained when processing the data after it has been truncated to a lower dynamic range, and shows that the precision does not deteriorate significantly until the dynamic range drops to 5 bits or below. Even at such a low dynamic range, the precision is still several times better than that obtained with the previous systems. Truncating the dynamic range does not affect the signal to noise ratio, and therefore, the improvement in precision obtained over and above the improvement in dynamic range must be due to better signal to noise ratio. This improvement in signal to noise ratio is due to several efficiency improvements of the new set-up over the old one. These improvements include:

- better image intensifier to camera coupling (providing higher intensities and better photon statistics on the CCD),
- better camera sensitivity (due to larger pixel and sensor size),
- better noise performance of the camera itself,
- high photo-cathode voltage drive, providing approximately twice the quantum efficiency (and better photo-electron statistics), and
- higher illumination drive levels, providing approximately twice the illumination and even more improved photon statistics on the photocathode).

Although figure 3 shows that not much is gained by way of precision for a 12 bit system over an 8 or 6 bit system, it does show that a 12 bit system has a much larger head room for signal level variation. The graph shows that the signal level can vary over more than an order of magnitude without change in precision, assuming a constant signal to noise ratio. However, the assumption of a constant signal to noise ratio is not likely to be valid, because, with the exception of photon noise, the noise level is generally independent of signal level. If the intensity from one object in a scene is half that of another, it is likely that the signal to noise ratio for the first object will also be close to half. To investigate
the effect of dropping signal level an experiment was performed to determine the ranging precision (in the same was as described above) as the intensity of the received signal was reduced (with a pulse skipping function in the intensifier driver). The results, graphed in figure 4, closely follow an inverse square-root relationship, and show that the received light intensity can drop by approximately 70% (from full scale) while only affecting the precision by a factor of 2.

![Figure 3](image1.png)

**Figure 3** – Precision obtained with the image ranging system vs. dynamic range for three different integration times.

![Figure 4](image2.png)

**Figure 4** – Precision obtained with the image ranging system vs. intensity of the received signal, and an inverse square-root trend line.

### 3.2 Improved operating frequency

The previously published system operated at a frequency of 10 MHz\(^1\), and the results presented so far here are also with 10 MHz modulation frequencies. Because range precision is proportional to phase determination precision and modulation wavelength, it was predicted that reducing the wavelength (by increasing modulation frequency) should result proportionally improved range precision. The LED illumination source used for the result reported so far has a maximum operating frequency of 10 MHz before the modulation depth starts to deteriorate (resulting in a lower signal to noise ratio). Also, in the previously published system, the photo-cathode response of the image intensifier was also limited to approximately 10 MHz before its response began to degrade.
In the new system, the image intensifier has been replaced with a model that is capable of photo-cathode response times of approximately an order of magnitude faster. Also, by replacing the LED illumination with a laser-diode light source, much higher illumination modulation frequencies are possible. The new system can now respond up to approximately 90 MHz, but power requirement and dissipation issues in the image intensifier and driver circuit limit the practical maximum usable modulation frequency to 65 MHz at this time. The pulse skipping function can be used to extend the usable frequency range to 80 MHz, but uneven rise and fall times of the devices cause loss of collected light and associated reduction in signal to noise ratio negate the improvements in precision obtained from operating at higher frequencies. For the range of 10 MHz to 65 MHz, the precision of the system was determined with the flat panel measurement approach as used in the previous section. The results obtained are graphed in figure 5.

![Figure 5](image.png)

Figure 5 – Plot of precision (in both milli-radians and millimeters) for operating frequencies from 10 to 65 MHz, along with a linear fit trend line for the milli-radian data points, and an inverse relationship fit for the millimeter data points.

The plot in figure 5 shows that the phase determination precision remains consistent at approximately 1.6 milli-radians over the range of operating frequencies. When converted to distance precision, the expected inversely proportional relationship with modulation frequency is shown. The precision obtained at 10 MHz does not correlate with the results shown in figure 3 when LED illumination was used because the laser-diodes emit a lower power density than the LEDs, leading to a lower signal to noise ratio and consequently worse precision. With the ranging precision approaching 0.5 mm, the total precision improvement over the previously published system is at least 30 times.

3.3 Measurement Example

To graphically demonstrate the precision improvements obtained with the system enhancements detailed above, measurements were acquired of our universal test structure – a garden gnome – in both high precision 65 MHz modulation mode, and lower precision 10 MHz mode. Three-dimensional reconstructions of these measurements are shown in figure 6 along with a photograph of the object. In the 65 MHz measurement, the structures of the gnome’s facial features are clearly visible with little apparent noise. At a modulation frequency of 10 MHz however, only the general shape of the gnome’s face is apparent. Measurements with the previously published system would contain too much noise for any viable shape, so are not shown here. At this time, the repetitive horizontal structure visible in the 10 MHz measurement is not well understood; we suspect it is caused by interaction of the camera’s internal 40 MHz clock with harmonics of the modulation signals. Similar image distortions have been noted at other modulation frequencies that are sub-harmonics of the camera’s clock.
ACKNOWLEDGEMENTS

The authors would like to thank Adrian Jongelen, Richard Conroy, John Godbaz, and WaiktoLink Limited for their assistance with funding, equipment upgrades, and operation. Also, this research would not have been possible without post-doctoral funding from the Foundation for Research, Science, and Technology (New Zealand).

REFERENCES