Computing depth under ambient illumination using multi-shuttered light

Hector Gonzalez-Banos James Davis

Honda Research Institute
Mountain View, CA 94041 USA

Abstract

Range imaging has become a critical component of many computer vision applications. The quality of the depth data is of critical importance, but so is the need for speed. Shuttered light-pulse (SLP) imaging uses active illumination hardware to provide high quality depth maps at video frame rates. Unfortunately, current analytical models for deriving depth from SLP imagers are specific to the number of shutters and have a number of deficiencies. As a result, depth estimation often suffers from bias due to object reflectivity, incorrect shutter settings, or strong ambient illumination such as that encountered outdoors. These limitations make SLP imaging unsuitable for many applications requiring stable depth readings. This paper introduces a method that is general to any number of shutters. Using three shutters, the new method produces invariant estimates under changes in ambient illumination, producing high quality depth maps in a wider range of situations.

1. Introduction

Range sensors are important components of many computer vision applications, including surveillance, face recognition, object modelling, target tracking, and human modelling. A number of technologies exist for recovering range images, including stereo triangulation, laser scanners, structured light, and shuttered light-pulse (SLP) imaging. Although each technology has its own advantages, SLP imaging is particularly interesting because it can theoretically provide high resolution range images at video frame rates. In addition, several vendors have indicated that they will soon produce commodity CMOS versions of their sensors. It is foreseeable that in the near future many cameras will have embedded SLP chips and return a depth channel in addition to the usual RGB color signal.

SLP imagers recover depth by emitting and shuttering a short pulse of light. Objects in the scene reflects this pulse back into the imager. Since the speed of light is finite, light reflected from closer objects returns to the imager before light reflected by more distant objects (see Figure 1). A fast shutter can be placed in front of the imager, and closed before all of the incoming light arrives. Thus, the imager will capture more light from near-by objects while distant ones will appear dim. The depth of objects is then proportional to the amount of integrated light I_h.

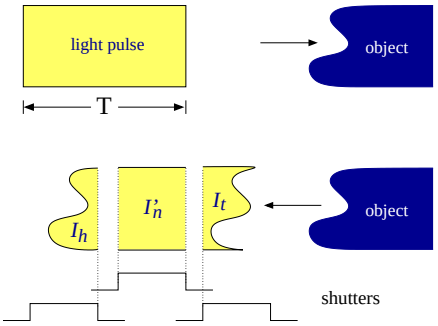

Figure 1: A light pulse of duration T radiates an object and is reflected back to the sensor. The signal is shuttered at the head, center and tail of the signal. The measured intensities I_h and I_t are functions of the distance travelled by the pulse, while the intensity I'_n is a constant fraction of the unshuttered value I_n.

In practice, object reflectivity (albedo) is unknown. Thus, a second signal is required in order to estimate the two unknowns: depth and reflectivity. Depth can be recovered by using an un-shuttered pulse of light I_n to normalize the first signal I_h. Depth becomes proportional to the ratio I_h/I_n, and depth recovery occurs under the *single-shutter condition*. On the other hand, the *double-shutter condition* involves a signal I_t produced by a second shutter that is opened *after* the incoming light from near-by objects arrives. Distant objects will now appear brighter in the imager, and depth becomes proportional to $I_h/(I_h + I_t)$.

One common depth estimation error is a bias due to object reflectivity. Although both the single- and double-shutter models theoretically eliminate this bias, it is nevertheless a common artifact. Bias is often due to accidental misconfiguration of the hardware (e.g., calibrating using a single-shutter model when the sensor operates under double-shutter condition), and less-frequently due to light-scattering effects when the shutter is closed [12]. Unfortunately, current implementations of SLP imaging make use of relatively crude analytical models which are difficult to calibrate as they require precise understanding and configu-

ration of the hardware.

A more fundamental difficulty is that existing models do not account for ambient illumination. Specifically, they assume that the emitted pulse is the only source of illumination (except for a possible term that is constant to every pixel). Current devices use a narrow bandpass color filter in the range of the device's active IR illumination. While this is sufficient for laboratory conditions, it is often inadequate for real world tasks. Strong ambient IR emitters such as sunlight and halogen lamps greatly reduce the accuracy of estimated depth. In general, the effect of ambient illumination is pixel-dependant, which brings the number of unknowns factors to three (depth, reflectivity, and illumination). A third shutter is then required.

This work contributes a mathematical model for determining depth. The main advantages of the model are twofold: it handles a wide range of shutter settings, and it is scalable to any number of shutters. Bias due to the incorrect selection of a recovery model is thus eliminated. Furthermore, bias due to light-scattering effects or ambient illumination is corrected by using three shutters with no additional calibration effort.

2. Previous work

Most previous research in SLP imaging has focused on building the hardware device itself. For example, 3DV Systems, Ltd. and Canesta, Inc. have both developed and patented commercial implementations [11, 1]. Hardware enhancements are also possible. For instance, Kawakita et. al have shown a MCP image intensifier that increases light sensitivity and thus depth sensitivity of their imager [5].

Basic SLP implementations make use of the single-shutter model of depth estimation [4, 3, 1]. This model directly factors out object reflectivity by integrating unshuttered light in order to normalize shuttered measurements with respect to object reflectance. Although intuitively straightforward this model suffers from relatively poor precision.

The double-shuttered model of depth estimation shutters the reflected light on both the head and tail ends of the returning pulse, and has been shown to improve depth precision [6, 9, 2].

Both of the previous models suffer from sensitivity to ambient light. Although most implementations have made no effort to address this issue Schroeder et. al. directly measured the ambient light level per pixel so that it could be pre-subtracted from the shuttered measurement [10].

This work generalizes the single- and double-shutter methods into a single model which can be used to calibrate devices with shutters in either of the above configurations. Further, it allows calibration, even when the precise placement of shutters is unknown, possibly matching neither condition above.

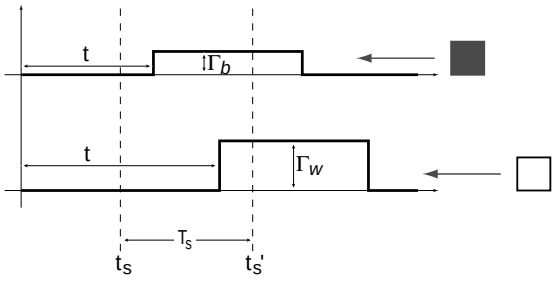

Figure 2: The effect of reflectivity and depth on measured intensity. Brighter objects will have a higher value of Γ, and further objects will have a higher value of t.

The above discussion relates to shutter placement, there has also been a wide variety of experimentation with regard to the shape of the lamp pulse. Most researchers have assumed a square wave [11, 1, 9, 2, 10]. Others have used linearly increasing and decreasing ramps [6] or sinusoidal modulation [7]. This variation does not substantially change the essential characteristics of depth recovery, and has not been the focus of our work. Since our hardware device approximates a square pulse, we use that model for the analysis given in this paper.

3. Basic principles

Shuttered light pulse imagers rely on the finite speed of light to measure depth. A pulse of light is emitted into the scene. The reflected light pulse returns to the sensor as a wavefront that encodes object depth, as shown in Figure 1. The returning light is integrated by a standard CCD imager. By activating a fast shutter after a portion of the light has returned, but before the entire pulse arrives, object depth can be measured. Because light reflected by near-by objects returns sooner, the imager will collect more light from closer objects. That is, object depth is related to measured intensity, with closer objects appearing brighter.

If the integrated light intensity was only a function of depth it would be possible to measure depth directly. However, the measured intensity is also related to object reflectivity. The relationship between object depth, object reflectivity, and measured intensity can be seen in Figure 2. A pulse is reflected from each of two objects, a bright block and a dark block at different depths. The front of the light pulse begins to arrive at time t, which is related directly to object depth. Since the dark object is closer, the value of t is smaller. The pulse has intensity Γ_b and Γ_w for the black and white object respectively ($\Gamma_b < \Gamma_w$). The CCD integrates the returning light pulse during the interval while the shutter is open. Since no light returns ahead of the returning pulse the integrated value is $I_h = (t'_s - t) \cdot \Gamma_{b,w}$, related to both depth and reflectivity. A normalization measurement can be taken with the shutter open, integrating the whole

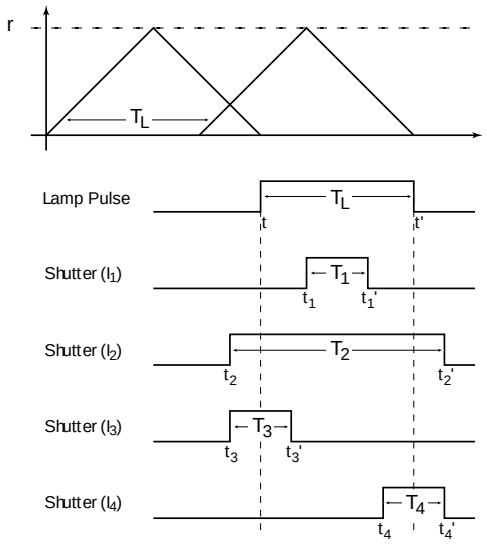

Figure 3: Schematic of the shutter timings with respect to an arriving light pulse. The front of the pulse is emitted at $t = 0$, and its reflection arrives at a time $t = 2r/c$, where r is depth and c is the speed of light.

pulse, resulting in $I_n = T \cdot \Gamma_{b,w}$. Depth is computed as the ratio $t = t'_s - T \cdot I_h / I_n$, which is independent of reflectivity.

3.1. Multiple shutters

In general it is possible to place a shutter anywhere during the returning pulse. Figure 3 shows four possible categories of shutter locations. This diagram plots depth versus time, and represent the shutter timings with respect to a travelling light pulse of duration T. The pulse is emitted at $t = 0$, and its reflection returns to the sensor at a time $t = 2r/c$, where c is the speed of light. Shutters are activated at times t_1, t_2, t_3 and t_4 with a duration T_1, T_2, T_3 and T_4, respectively.

The discussion so far has presumed a head shutter, labelled here as I_3, and a surrounding shutter, labelled as I_2. But a shutter in the middle (shutter I_1) or on the tail of the pulse (shutter I_4) is equally possible. Although the traditional model has been to use shutters I_2 and I_3, depth can be estimated using other combinations. The equations that relate depth to different combinations of shutters will be described in the next section.

In most cases, two unknown values, object reflectivity and object depth, need to be estimated. Therefore measurements from at least two independent shutter locations will be required to solve for a unique solution. A single measurement can estimate only one value. For example, a standard camera can be thought of as a single measurement at I_2, which estimates the single value Γ. Alternately, if object reflectivity is known a priori, a single measurement at I_3 would be sufficient to estimate depth.

Using measurements I_3 and I_4 involves shuttering twice, at the head and tail of the pulse. In contrast, measurements

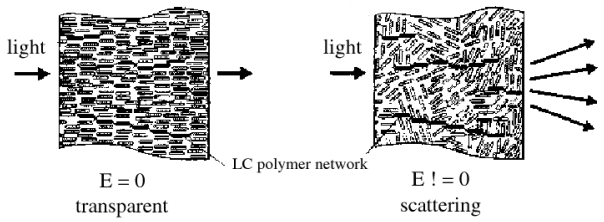

Figure 4: An example of a liquid crystal polymer shutter: The electrical field E controls the light transmission through the crystal; when $E \neq 0$ few photons go through, but those that do are scattered (from Anteryon LCP 250 data sheet).

I_2 and I_3 technically require only I_3 to be shuttered. For this reason these methods have been labelled in previous work as double- and single-shuttering, respectively. However, we feel that this naming somewhat obscures the actual situation since both methods use two measurements to estimate two unknowns. In order to maintain consistency with previous work we label measurements I_2 and I_3 as "single" shuttering, and measurements I_3 and I_4 as "double" shuttering, keeping in mind that the labels refer to particular configurations rather than to the number of measurements.

3.2. Solid state shutters

For a SLP device measuring depth within a range of a few meters the required shutters and light pulses have durations on the order of nanoseconds. It seems surprising that such shuttering speeds are possible, but in fact such shutters exist [8]. These shutters are solid state devices whose exact nature is often a patented technology or a trade secret. One embodiment are liquid crystal polymer (LCP) shutters. Although not capable of nanosecond rise times, LCP shutters are simple to understand and give us a good insight on the difficulties of using shuttered light to measure depth.

Consider the diagram shown in Figure 4. The molecules in the polymer crystal are well-aligned in the absence of an electric field ($E = 0$). However, the material becomes turbid when an electric field is applied. This effect is due to the fact that not all molecules in the crystal respond to the electric field in the same way (but they all have the same *relaxed* state). The result is a controllable light scatterer.

When the shuttered is open light leaves the crystal along the same direction it enters. Thus, an imager placed on one side of the shutter is capable of imaging a scene relayed by a lens system on the opposite side. When the shuttered is closed very little light goes through, but a small fraction does. It is important to realize that this residual light is *scattered* –i.e., light does not necessarily leave the crystal along the incoming direction. This scattered light may or may not be a problem. If the imager is placed very near the crystal then a "closed" shutter will effectively scatter each

ray over the entire image. Image contrast is then destroyed and a DC signal remains that can be corrected by electronic means. However, when the imager is separated from the crystal some image contrast remains. The "off" value for a pixel becomes spatial-dependant, and is not merely a fraction of the "on" value.

The above discussion refers to LCPs, other types are based on different physical principles. The reader should keep in mind that light scattered by the shutter in its "off" state may be a problem in some embodiments. It is desirable that this effect is accounted for by the recovery model.

4. Depth recovery models

Consider the shutter settings shown in Figure 3. The amount of light received by each camera is given by the following equations:

$$\begin{aligned} I_1 &= T_1\,\Gamma + (T_{\rm L} - T_1)\lambda\,\Gamma, && (1)\\ I_2 &= T_{\rm L}\,\Gamma, && (2)\\ I_3 &= (t_3 + T_3 - t)(1-\lambda)\Gamma + T_{\rm L}\,\lambda\,\Gamma, && (3)\\ I_4 &= (t - t_4)(1-\lambda)\Gamma + T_{\rm L}\,\Gamma. && (4) \end{aligned}$$

Here, Γ represents the pulse intensity reflected back into the sensor, which is different for every pixel. The term $\lambda > 0$ represents the fraction of reflected intensity that reaches the imager when the shutter is "off". Ideally, λ is a constant for all pixels, but if light-scattering effects are large, then λ is pixel dependant.

Since $t = 2r/c$, where r is depth and c is the speed of light, recovering t amounts to recovering depth. Consider equations (1) and (3). This is equivalent to the traditional single shuttering configuration. After some algebra, the value of t can be computed as follows:

$$t = t_3 + T_3 - T_1\,\frac{I_3}{I_1} + T_{\rm L}\,\frac{\lambda}{\lambda+1}\frac{I_1 - I_3}{I_1}. \quad (5)$$

Notice the dependancy with λ, which has not been accounted for in the past. Thus, if light-scattering effects are large, the estimated depth will have a bias because λ will not be a constant.

In fact, we can use other shutter pairs to produce different estimates as follows:

Equations (1) & (4) $\Rightarrow$

$$t = t_4 - T_{\rm L} + T_1\,\frac{I_4}{I_1} - T_{\rm L}\,\frac{\lambda}{\lambda+1}\frac{I_1 - I_4}{I_1}, \quad (6)$$

(2) & (3) $\Rightarrow$

$$t = t_3 + T_3 - T_{\rm L}\,\frac{I_3}{I_2} + T_{\rm L}\,\frac{\lambda}{\lambda+1}\frac{I_2 - I_3}{I_2}, \quad (7)$$

(2) & (4) $\Rightarrow$

$$t = t_4 - T_{\rm L} + T_{\rm L}\,\frac{I_4}{I_2} - T_{\rm L}\,\frac{\lambda}{\lambda+1}\frac{I_2 - I_4}{I_2}, \quad (8)$$

and (3) & (4) $\Rightarrow$

$$t = (t_4 - T_{\rm L})\frac{I_3}{I_3 + I_4} + (t_3 + T_3)\frac{I_4}{I_3 + I_4} + T_{\rm L}\,\frac{\lambda}{\lambda+1}\frac{I_4 - I_3}{I_3 + I_4}. \quad (9)$$

This last equation describes the double-shutter method.

4.1 Eliminating light-scatter

At this point it should be obvious that we can factor out the term λ by combining three measurements. For instance, consider equations (1), (3) and (4). After several algebraic steps we obtain the following:

$$t = (t_4 + T_1 - T_{\rm L})\frac{I_3 - I_1}{I_3 + I_4 - 2I_1} + (t_3 + T_3 - T_1)\frac{I_4 - I_1}{I_3 + I_4 - 2I_1}, \quad (10)$$

where the factor λ has been eliminated. This last equation has the alternative form:

$$t = a'\frac{I'_3}{I'_3 + I'_4} + b'\frac{I'_4}{I'_3 + I'_4}, \quad (11)$$

where $I'_3 = I_3 - I_1$ and $I'_4 = I_4 - I_1$. Note that this is similar to the double-shutter equation (9) when $\lambda = 0$.

To understand why equation (11) is invariant to light scatter consider the term I'_3. From equations (1) and (3) we get $I'_3 = (t_3 - t + T_3 - T_1)\Gamma'$, where $\Gamma' = (1-\lambda)\Gamma$. Similarily, $I'_4 = (t - t_4 + T_{\rm L} - T_1)\Gamma'$. That is, the effect of light scatter has been lumped together with the intensity Γ into the instrumental variable Γ', and we are back to the ideal double-shutter condition.

4.2. Eliminating ambient illumination

Suppose the light scatter effect is negligible and that λ is constant for all pixels. The recovered depth map using two shutters will have a constant offset which is of no consequence. We can then use the third shutter to eliminate the effect of ambient illumination.

Let $\lambda = 0$ in order to simplify the discussion, and assume that an ambient term β is present, as shown in figure 5. The ambient energy collected by camera k will be $\beta \cdot T_{\rm k}$, where $T_{\rm k}$ is the duration of shutter k. This energy is independent of object distance but dependant on each pixel because ambient illumination in general does not affect all scene points equally.

The energy collected by each camera under the shutter settings shown in Figure 3 is as follows:

$$\begin{aligned} I^b_1 &= T_1\,\Gamma + T_1\,\beta, && (12)\\ I^b_2 &= T_{\rm L}\,\Gamma + T_2\,\beta, && (13)\\ I^b_3 &= (t_3 + T_3 - t)\Gamma + T_3\,\beta, && (14)\\ I^b_4 &= (t - t_4 + T_{\rm L})\Gamma + T_4\,\beta. && (15) \end{aligned}$$

As before, let us use shutters 1, 3 and 4 to calculate depth. Let $I_3'' = T_1\, I_3^b - T_3\, I_1^b$ and $I_4'' = T_1\, I_4^b - T_4\, I_1^b$. By computing the ration I_3''/I_4'' we arrive at the following equation:

$$t = (\, t_4 \, + T_4 \, - T_{\rm L}\,) \frac{I_3''}{I_3'' + I_4''} + \, t_3 \, \frac{I_4''}{I_3'' + I_4''}, \qquad (16)$$

which is again similar to the double-shutter equation.

Equations (16) and (11) become identical when all three shutters have the same duration ($T_1 = T_3 = T_4$). That is, the same model that eliminates light-scatter also eliminates ambient illumination. Both effects are *indistinguishible* from each other when all three shutters have the same duration.

4.3. General model

Depth can be recovered from a variety of situations using the equations presented so far. We basically chose a model that matches the shutter settings, calibrate the necessary parameters, and then use it to recover depth. The drawback of this approach is that the shutter settings must exactly match our assumptions, otherwise the model will perform poorly. For instance, equation (5) (single-shutter) is a bad choice if the sensor operates under double-shutter conditions. And what about two distinct tail shutters, or two distinct head shutters? Neither of these scenarios is described by any of the above equations.

Instead of deriving every possible shutter condition, and then guessing the correct shutter settings, we postulate a general equation $r = r(I_1, I_2, \ldots, I_n)$ encapsulating the recovery process for any shutter setting. Note that all recovery models described so far consist on a linear combination of n distinct shutters divided by a different linear combination of the same n shutters. In other words,

$$\begin{aligned} r &= \frac{a_0 + a_1 I_1 + a_2 I_2 + \ldots a_n I_n}{b_0 + b_1 I_1 + b_2 I_2 + \ldots b_n I_n}, \qquad (17) \\ \text{or} \qquad r &= \frac{\mathbf{a} \cdot [1,\, \mathbf{Y}]}{\mathbf{b} \cdot [1,\, \mathbf{Y}]}, \end{aligned}$$

where n is the number of shutters employed, and $\mathbf{Y}$ is a row-vector of intensities obtained by imaging the scene under each shutter setting. We must, of course, calibrate the parameter vector $[\mathbf{a}, \mathbf{b}]$. A calibration procedure is described in the next section.

So far we have neglected the effect of camera gain and electronic (DC) bias. Because these parameters are intrinsic to the imager they can be lumped into the parameters of equation (17). In fact, a_0 and b_0 are introduced especifically to account for DC bias. In the idealized case this bias is zero and parameters a_0 and b_0 can be disregarded.

5. Model calibration

In equation (17), any multiple of $[\mathbf{a}, \mathbf{b}]$ will produce the same result. We thus set $b_0 = 1$ and obtain: ,

$$r \;=\; \mathbf{a} \cdot [1,\, \mathbf{Y}] - r\, [b_1, b_2, \ldots b_n] \cdot \mathbf{Y}. \qquad (18)$$

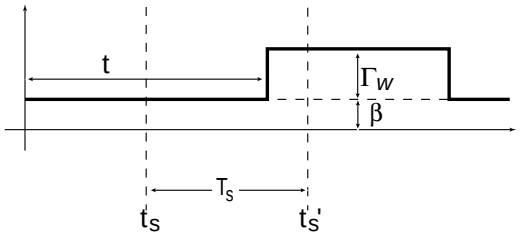

Figure 5: The effect of ambient illumination on the measured light level. In addition to the reflected light level Γ, there will be an additional background lighting level, β which will be integrated for the entire shutter period regardless of object depth.

Suppose that m measurements of the vector $\mathbf{Y}$ are available, along with a vector of *known* depths $\mathbf{r} = [r_1, r_2, \ldots, r_m]^T$, where one depth is associated with each measurement. m is typically very large because it is the number of pixels used during calibration, whereas n is very small (it is the number of shutters). The data set must satisfy the following matrix constraint:

$$\mathbf{r} \;=\; \begin{bmatrix} 1 & \mathbf{Y}_1 & -r_1\mathbf{Y}_1 \\ 1 & \mathbf{Y}_2 & -r_2\mathbf{Y}_2 \\ 1 & \mathbf{Y}_3 & -r_3\mathbf{Y}_3 \\ \vdots & \vdots & \vdots \\ 1 & \mathbf{Y}_m & -r_m\mathbf{Y}_m \end{bmatrix} \begin{bmatrix} a_0 \\ \vdots \\ a_n \\ b_1 \\ \vdots \\ b_n \end{bmatrix}, \qquad (19)$$

$$\Rightarrow \quad 0 \;=\; \mathbf{r} - \mathbf{M}\,\theta. \qquad (20)$$

Clearly, the above equation will not hold exactly in the presence of uncertainties. We compute our model parameters by solving this matrix equation in the least-squares sense. The solution is given by $\hat{\theta} = \mathbf{M}^{+}\mathbf{r}$, where $\mathbf{M}^{+}$ is the pseudo-inverse of $\mathbf{M}$. If $(\mathbf{M}^T\mathbf{M})$ is full rank the solution takes the familiar form $\hat{\theta} = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{r}$.

Alternatively, we could enforce the condition that $\|(\mathbf{a}, \mathbf{b})\| = 1$ and solve an homogenous system of equations instead of setting $b_0 = 1$. This can be accomplished using singular value decomposition. In our preliminary results this method is more fragile than the previous one, although a more thorough analysis is necessary for future work.

5.1. Algorithm

The least-squares method is an adequate procedure provided we exercise care in the computation of $\mathbf{M}$. The columns of $\mathbf{M}$ can easily differ by orders of magnitude, which results in ill-conditioned computations of $\mathbf{M}^{+}$. To avoid this, we ensure that the intensity data is normalized so that on average the value for each shutter is 0, while its absolute value is on average 1. We apply a similar conditioning to the data $\mathbf{r}$.

The overall calibration algorithm is as follows:

Algorithm *Multi-shutter calibration*
Input: 1.- Given n independent shutters,
2.- a vector of known depths $\mathbf{r}$ of length m,
3.- and correponding readings $\mathbf{Y}_1, \ldots, \mathbf{Y}_m$
Output: A parameter vector $[\mathbf{a}, \mathbf{b}]$ solving equation (20) in the least-squares sense.

1. $\mathbf{Y}_k \leftarrow (\mathbf{Y}_k - c_k)/s_k \ \ \forall k = 1, 2, \ldots, n$, where $c_k = \text{avg}(\mathbf{Y}_k)$ and $s_k = \text{avg}(|\mathbf{Y}_k - c_k|)$.
2. $\mathbf{r} \leftarrow (\mathbf{r} - c_r)/s_r$, where $c_r = \text{avg}(\mathbf{r})$ and $s_r = \text{avg}(|\mathbf{r} - c_r|)$.
3. Assemble $\mathbf{M}$ as per equation (19).
4. Compute $\mathbf{M}^+$ and evaluate $\hat{\theta} = \mathbf{M}^+\mathbf{r}$.
5. Evaluate $\hat{\mathbf{r}} = \mathbf{M}\hat{\theta}$. Discard any row in $\mathbf{M}$ yielding an error larger than ϵ (e.g., $\epsilon = 2$).
6. Recompute $\mathbf{M}^+$ and evaluate $\hat{\theta} = \mathbf{M}^+\mathbf{r}$.
7. $(a_0, \ldots, a_n, b_1, \ldots, b_n) \leftarrow \hat{\theta}$.
8. $a_k \leftarrow a_k/s_k$, and $b_k \leftarrow b_k/s_k \ \ \forall k = 1, \ldots, n$.
9. $a_0 \leftarrow (a_0 - \sum_1^n a_k c_k)$, and $b_0 \leftarrow (1 - \sum_1^n b_k c_k)$.
10. $a_k \leftarrow (s_r a_k + c_r b_k) \ \ \forall k = 0, \ldots, n$.
11. Return $[\mathbf{a}, \mathbf{b}]$.

5.2. Calibration with ambient light

Although correct, the calibration procedure presented so far requires a large amount of data to be collected when used with three shutters and ambient illumination. The process of calibration allows us to find the model parameters $[\mathbf{a}, \mathbf{b}]$ which best explain a set of training data. If our model is appropriate, we hope that these parameters will allow us to recover depth from new observations at a later point in time. Unfortunately, we can only expect that the model will perform well if the initial calibration data spans the space of expected observations. When we expect ambient lighting, this implies that calibration data must be captured under a wide range of lighting conditions. Although this is possible, it would be tedious.

We can eliminate the need to capture a large data set that covers all possible ambient lighting conditions by using the mathematical model itself as a constraint to cancel any ambient lighting effects. Given the relatively modest assumptions that all imagers have equal gain, and that all shutter intervals, T_1, T_3, T_4, are of equal duration, ambient illumination can be removed by making the substitutions given for equation (16). Using the substituted measurements, I_3'' and I_4'', together with equation (17) allows us to calibrate without ensuring that variations in ambient lighting are present in the training data. As a result much fewer calibration images need to be captured.

6. Experimental evaluation

Each of the previously presented models of depth estimation can be calibrated with similar data. A planar target printed with a black and white checkerboard pattern is moved to each of several known depths, $[r_1, r_2, ..., r_m]$.

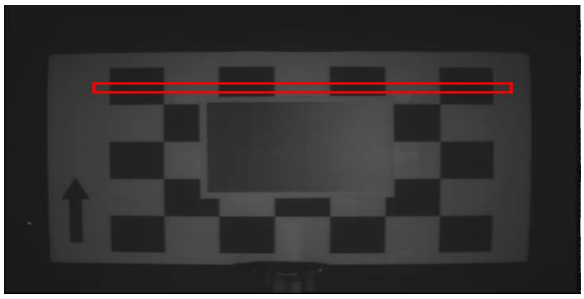

Figure 6: Test target. The region inside the red rectangle was used to evaluate depth. The mean calculated depth of each column under several different conditions is plotted in Figures 7 and 8.

Although we have obtained useable calibration under some circumstances with as few as two depth planes, in this work we use 11 depth planes spaced 10cm apart, covering a 1m working volume.

Each pixel which lies on the target is observed under several shuttering conditions generating an observation vector, $\mathbf{Y}$, which is used to generate one row of the constraint matrix in equation (19). The target typically covers thousands of pixels in each of 11 imaging conditions, while only two or three shutters are used. Thus, the matrix $\mathbf{M}$ is tall and narrow, serving to well constrain the solution space of $[\mathbf{a}, \mathbf{b}]$.

We evaluated a total of three shuttering conditions. Single shuttering makes use of shutters I_1 and I_3, double shuttering uses shutters I_3 and I_4, and triple shuttering uses I_1, I_3, and I_4. The single and double shuttering models of calibration were taken from equations (5) and (9). In order to match the models used in most prior publications, the scattering term, λ, is assumed to be zero. The triple shuttering model make use of equation (17). The measurement substitution discussed in section 5.2 is utilized. All calibration data was captured without the presence of significant ambient illumination.

We use the 3DV systems Mini-Z as our hardware platform. This device has only two physical shutters. Thus for all experiments reported here that require three shutters we use a single imager in the device and adjust the shutter timing to each of the three shuttering conditions to take measurements. For all experiments we keep the shutter interval, T_s, constant.

6.1. Reliable operation

The single and double shuttering conditions are both theoretically correct with regard to reflectivity bias, i.e. an object's color should not effect its estimated depth. Nevertheless these models often perform poorly, with noticeable artifacts due to reflectivity. The primary reason for this deficiency is that it can be difficult to know precisely when

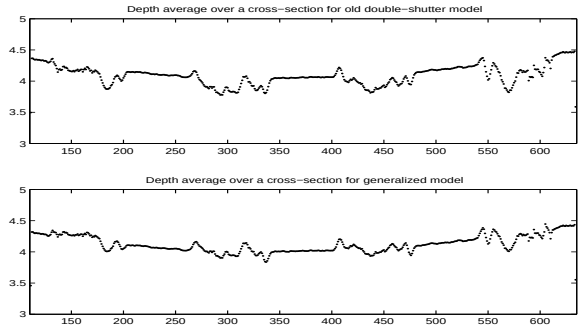

Figure 7: Calculated depth when the shutters are configured in a known good configuration. Both the traditional and generalized double shuttering models are able to recover depth.

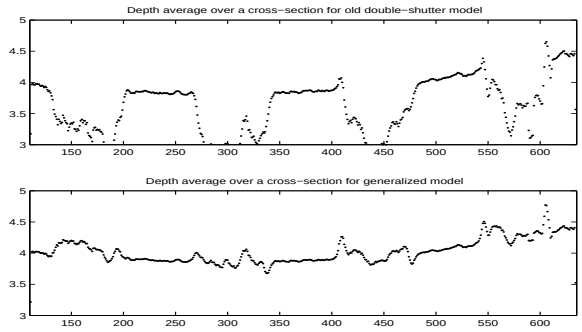

Figure 8: Calculated depth when the shutters are configured in a known bad configuration. The traditional double shuttering model does a poor job estimating the depth of dark objects, while the generalized model is still capable of estimating depth.

shutters are being triggered relative to the returning pulse.

Given a particular arrangement of shutters, it may be true that near objects are in a single shutter configuration while distant objects are in a double shutter configuration. Ideally the same conditions would apply to all objects in the working volume. This requires specifying a shuttering arrangement to match the desired working volume.

In order to accommodate variable working volumes, our experimental device has a user configurable slider for setting shutter range. While this is a useful option, it is far from intuitive. For example, the manufacturer default settings place the shutter values t_1 and t_3 at the relatively cryptic values of -87 and 100 respectively, in unknown units. In practice, we have found it extremely difficult to reliably choose appropriate settings for a given working volume, with accidental misconfiguration relatively common.

Both the single and double shuttering models of depth estimation require that the device is correctly configured to match the model. Since in practice this is often not true, a calibration model that does not require a priori knowledge of the shutter settings is desirable. The generalized double shuttering model presented as equation (17) was developed to satisfy this need.

In order to show the improvement from generalized double shuttering, we first placed the sensor in a configuration known to work with the traditional double shuttering method. Figure 6 shows an image of our test target. As shown in figure 7 both the tradition and generalized models of double shuttering are able to consistently estimate depth. We next changed the shuttering model to a configuration known to cause difficulties. As shown in figure 8, the traditional model does a poor job estimating depth in dark regions, while the generalized model continues to consistently estimate depth. The plots appear slightly curved rather than flat because we estimate the distance from a point sensor, which is more similar to $radius$ in a spherical coordinate system than to Z in a Euclidean space.

6.2. Ambient light

The intensity of ambient light can vary on a per pixel basis, according to the local presence of highlights or shadows. Since ambient lighting may also be time varying, it is necessary to estimate its intensity at each pixel, along with depth and object reflectivity. With the addition of ambient light, there are three values to be estimated, and thus three measurements are required. The triple shuttering model introduced by this work allows depth to be recovered even in the presence of strong ambient lighting conditions.

In order to evaluate the ability of triple shuttering to correctly estimate depth when ambient lighting is present, we compared it against the single and double shutter models. As shown in Figure 9, a test scene was constructed in which a halogen lamp illuminated objects from the side. The left image shows the scene illuminated only from the SLP illumination, with the halogen lamp turned off. The right image shows the scene illuminated only by the halogen lamp, with the SLP turned off. Note the shadows cast by the lamp.

The depth of this test scene was measured first without, and then with, ambient illumination, as shown in Figure 10. Both the single and double shutter models were adversely affected by the ambient light. The triple shutter model was able to correctly calculate depth under both conditions. Notice in particular that the single and double shuttered models calculated the correct depth in the ”ambient shadow”. However the rest of the scene has been interpreted as being closer (darker) than reality.

7. Conclusion

This work contributes a generalized model of depth estimation for SLP imagers. The new model allows for more reliable depth estimation because it does not require a priori knowledge of the precise configuration of shuttering conditions. In addition the new model easily generalizes to triple shuttering, allowing accurate estimation of depth, even in the presence of ambient light.

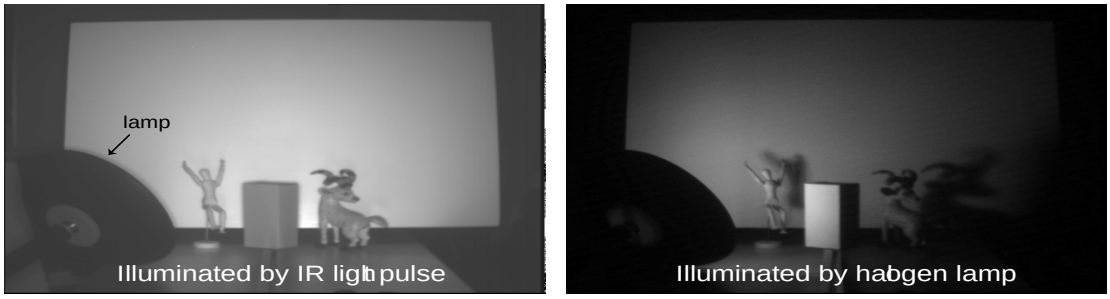

Figure 9: This test scene is shown first under illumination from the SLP device, and then under ambient illumination from a halogen lamp. The ambient illumination is weeker than the SLP illumination. Nevertheless it is sufficient to interfere with depth estimation.

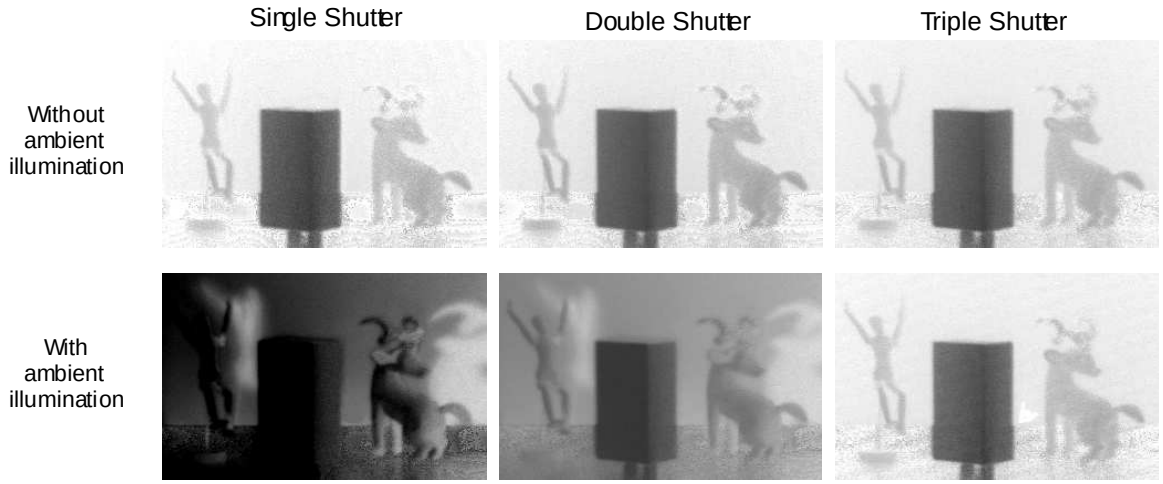

Figure 10: The test scene is evaluated first without ambient illumination, and then with ambient illumination. All models of calibration are capable of determining depth when no ambient illumination is present. However when there is ambient illumination, the single and double shutter models of depth recovery produce incorrect results, while triple shuttering is unaffected.

In the future we would like to better investigate and characterize the shutter configurations which are sufficient or optimal with regard to the quality of depth estimation.

References

[1] C. Bamji, "CMOS-Compatible 3-Dim. Image Sensor IC," *United States Patent*; no. US 6,323,942 B1; Nov. 27, 2001.

[2] J. Davis and H. Gonzalez-Banos, "Enhanced Shape Recovery with Shuttered Pulses of Light," *IEEE Workshop on Projector-Camera Systems*, 2003.

[3] R. Gvili and A. Kaplan and E. Ofek and G. Yahav, "Depth Keying," *SPIE Electronic Imaging 2003 Conference*, 2003.

[4] G. Iddan and G. Yahav, "3D Imaging in the Studio," *Proc. of SPIE Videometrics and Optical Methods for 3D Shape Measurement*, SPIE vol. 4298, pp. 48-55, 2001.

[5] M. Kawakita and T. Kurita and H. Kikuchi and S.Inoue, "High-definition three-dimension camera: HDTV version of an Axi-vision camera," *NHK Laboratories Note* , no.479, p.cover1-12, Aug. 2002.

[6] M. Kawakita and K. Iizuka and T. Aida and H. Kikuchi and H. Fujikake and J. Yonai and K. Takizawa, "Axi-Vision Camera (real-time distance-mapping camera)," *Applied Optics* , Opt. Soc. America, vol.39, no.22, p.3931-9, Aug 2000.

[7] R. Lange and P. Seitz, "Solid-State Time-of-Flight Range Camera," *IEEE Journal of Quantum Electronics* , vol. 37, no. 3, March 2001.

[8] A. Manassen and G. Yahav and G. Iddan, "Large Aperture Optical Image Shutter," *United States Patent*; no. US 6,331,911 B1; Dec. 18, 2001.

[9] R. Miyagawa and T. Kanade, "CCD-Based Range-Finding Sensor," *IEEE Transactions on Electron Devices*, vol. 44, no. 10, October 1997.

[10] W. Schroeder and E. Forgber and S. Estable, "Scannerless laser range camera," *Sensor Review*, MCB University Press, vol. 19, no. 4, pp. 285-291, 1999.

[11] G. Yahav and G. Iddan, "Optimal Ranging Camera," *United States Patent*; no. US 6,057,909; May 2, 2000.

[12] "LCP 250 Fast Scattering Liquid Crystal Polymer Shutter," *Data Sheet*; Anteryon B.V.; Netherlands.