

Enhanced Shape Recovery with Shuttered Pulses of Light

James Davis

Hector Gonzalez-Banos

Honda Research Institute
Mountain View, CA 94041 USA

Abstract

Computer vision researchers have long sought video rate sensors which return a channel of depth, in addition to color. One promising technology to make this possible is based on a projector-camera pair that generate shuttered pulses of light. Commercial implementations of the hardware technology are available today. Unfortunately, the software models that allow recovery of depth measurements suffer from relatively high noise and bias. This paper describes a mathematical recovery model for this class of shuttered sensors. The model is useful for understanding the behavior of these sensors, and is validated against empirical data. Based on our model, we introduce two specific methods for improving the quality of recovered depth. Multi-intensity estimation makes use of several observations with varying lamp intensities, and double shuttering suggests that improved performance can be obtained using two shutters instead of one.

1. Introduction

Range images are widely used in many computer vision applications, including surveillance, robotic navigation, motion capture, and gesture recognition. Since these images store shape as well as color information they enable tasks that are ill-posed when working with color alone.

Many technologies for obtaining range images exist. One particularly promising class of technologies are shuttered light-pulse (SLP) sensors. These sensors project a pulse of light of known dimensions, and observe the reflected illumination with a shuttered camera, as shown in Figure 1. The advancing wavefront of the light pulse is reflected from objects in the scene back towards the camera. The reflected wavefront encodes the shape of objects. Since the speed of light is finite, the portion of the returning pulse that is reflected from closer objects will arrive back at the sensor at an earlier time than those portions of the pulse that are reflected from more distant objects. Using a fast opto-electronic shutter, such as the one described in [3], the CCD can be blocked before the returning wavefront arrives entirely. Since light from near-by objects returns before the shutter closes, these will appear brighter to the CCD. Con-

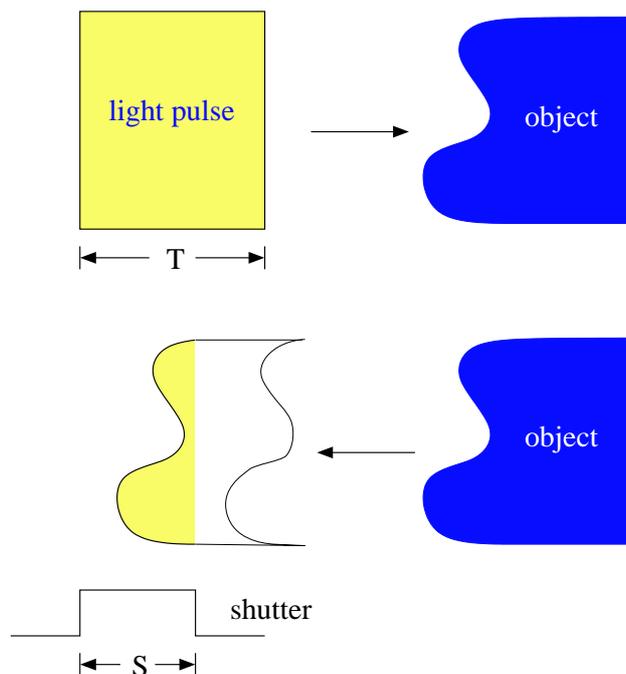


Figure 1: A light pulse of duration T radiates an object. The reflected pulse is shuttered at the sensor upon arrival. The measured intensity is a function of the distance traveled by the pulse.

versely, only small amounts of light returning from sufficiently distant objects will be observed, since the shutter closes before its arrival. Under these conditions, the intensity recorded by the CCD is correlated with object depth.

Observed intensity on the CCD is correlated with both object depth due to the shutter, and with the reflectivity of observed objects. That is, observed intensity is a function of both distance and object color. This ambiguity limits the accurate reconstruction of depth. Current implementations attempt to factor the effects of object reflectivity by using an unshuttered camera [1, 4] to record object reflectivity alone. A normalized estimate of depth is then calculated as the ratio between shuttered and unshuttered measurements.

The published methods of depth reconstruction attempts to obtain accuracy by computing the strict ratio of two measurements. However, this naive model is insufficient to ac-

count for the actual responses of available hardware. This paper describes an alternative model that more accurately predicts the response of real SLP sensors. In addition, our model handles the case when a second pulse shuttered at the trailing edge of the light pulse is used as a normalization measurement.

Based on our enhanced model we propose two specific methods for improving the quality of reconstructed depth:

Multi-intensity estimation The precision of depth computation is correlated with object reflectivity. The estimated depth of dark objects is inflicted with a much greater amount of noise than is the estimated depth of light objects. Multi-intensity estimation uses several observations, taken with variable illumination intensity, to produce a much more precise estimate of depth in dark regions.

Double-shuttering Existing SLP sensors theoretically employ single shuttering, with one shuttered CCD and one unshuttered CCD. Although this arrangement allows normalization, it does not allow optimal separability between objects at different depths. Double shuttering, in which both cameras are shuttered, one on the head of the pulse and one on the tail, improves depth estimation.

2. Previous work

Under ideal conditions SLP sensors are capable of relatively high-quality measurements as seen in Figure 2. However, ideal conditions rarely exist and their depth measurements are often corrupted by both noise and bias as a function of object intensity.

SLP sensors are available commercially [1, 2]. 3DV Systems, Ltd. and Canesta, Inc. have both developed SLP technologies. In these implementations the required components, a projector, shutters, and two cameras, have been packaged together as a single sensor. Canesta, Inc. has employed their technology [1] in the “Canesta Keyboard Perception Chipset,” which creates a virtual keyboard to be used in hand-held devices. This chipset is sold to OEMs. In contrast, 3DV Systems, Ltd. uses their shutter technology [2, 3] in an actual 3-D camera that is available to non-OEMs. In both cases, the reconstruction method generates depth using a shuttered camera and an unshuttered normalization camera. This simple reconstruction model results in excess noise and bias in the recovered depth.

Existing implementations of SLP sensors focus primarily on the underlying hardware technology, rather than algorithms for best reconstructing depth. All published implementations that we are aware of use a naive reconstruction model [5, 4, 1] that performs adequately for TV video keying applications, but poorly for object-shape recovery.

Although 3DV Systems uses a simple reconstruction model, one of their products (the Z-Mini) contains two shutters, each in front of one imager. These shutters can be



Figure 2: An example of a depth image produced by an SLP sensor. Far objects are shaded with increasingly darker tones of grey.

controlled independently, a fact which we exploit in our reconstruction algorithms.

3. Recovery models

Suppose a light pulse of constant intensity is emitted at $t = 0$ for a duration T . To simplify things, assume that the projector and cameras are collocated, so that this pulse leaves the camera along the imager’s optical axis. The pulse radiates the entire scene and is reflected back into the imager. Each pixel in the imager represents a ray originating at the focal point and intersecting the imager at the pixel location. Consider one ray, and let r be the distance to the closest scene object along this ray. The first photons of the pulse along this ray will arrive at the imager at a time $t = 2r/c$, where c is the speed of light. Let $t' = t + T$ be the time when the final photons arrive; see Figure 3.

Consider a shutter in front of the imager that opens at a time t_p and closes at a time $t'_p = t_p + S_p$, where S_p is the time the shutter remains open. If $t_p < t < t'_p < t'$, then the imager will receive a total radiation of $I_p = I_r(t'_p - t)$, where I_r is the reflected intensity of the pulse. I_r is a function of the pulse intensity and the reflectivity of the

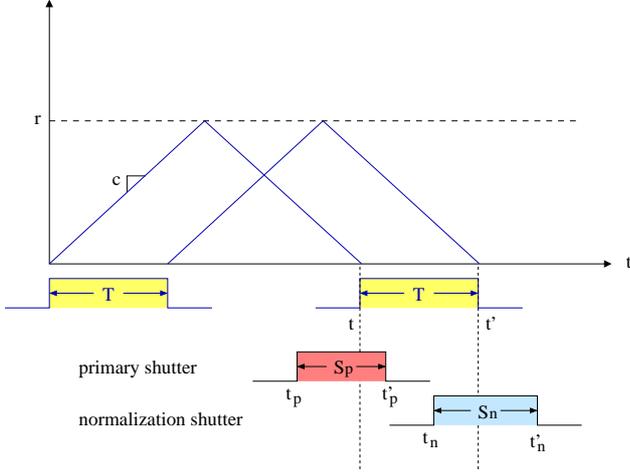


Figure 3: A representation of a the shutter timings with respect to an arriving light pulse. The pulse is emitted at $t = 0$, and its reflection arrives at a time $t = 2r/c$, where c is the speed of light.

object. Let I_p be the *primary measurement*.

In order to recover r we need to recover t . But I_r is unknown. If we repeat the above experiment, but this time we leave the shutter open, the imager will receive a total radiation of $I_n = I_r T$. Therefore, the recovery equation is:

$$\frac{2r}{c} = t'_p - T \frac{I_p}{I_n}. \quad (1)$$

The above represents the simplest recovery model that estimates depth as the ratio between shuttered and unshuttered measurements. We call I_n the *normalization measurement*. The measurement of I_n can be done in sequence after I_p . However, if the measurement device is equipped with two sensors, both I_n and I_p can occur simultaneously. In the rest of this paper we assume that this is the case.

Suppose the normalization measurement is also shuttered. Let t_n be the time when the second shutter opens, and $t'_n = t_n + S_n$ when it closes. Now the total radiation received by the imager will depend on the timings (t_n, t'_n) . We consider two cases: $t < t_n < t'_n < t'$ and $t < t_n < t' < t'_n$. The latter case is shown in Figure 3.¹

3.1. Single-shuttered recovery

If the second shutter settings satisfy $t < t_n < t'_n < t'$, then the total radiation received by the imager will be $I_n = I_r(t'_n - t_n) = I_r S_n$. That is, $2r/c = t'_p - S_n(I_p/I_n)$, which is similar to equation (1). This case is equivalent to an unshuttered normalization measurement. The recovery equation takes the following form:

$$r = a_1 + a_2 m, \quad (2)$$

¹A third possibility, $t_n < t < t'_n < t'$, is identical to that for I_p described above.

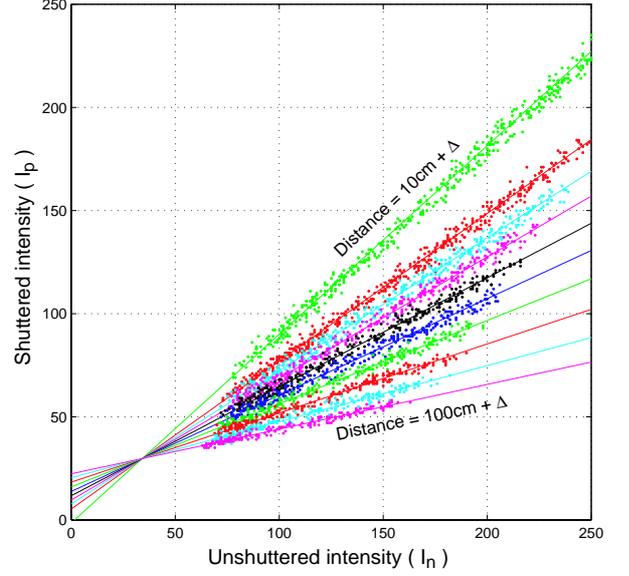


Figure 4: Plot of I_p vs. I_n as the flat target shown in Figure 9 is placed at different locations from the sensor (un-shuttered case).

where $m = I_p/I_n$. Depth is related to the ratio I_p/I_n .

3.2. Double-shuttered recovery

Now consider the case $t < t_n < t' < t'_n$. The measured radiation will be $I_n = I_r(t' - t_n) = I_r(t + T - t_n)$. Taking the ratio $m = I_p/I_n$, and after some algebraic manipulations, we obtain that $2r/c = (t'_p + T - t_n)/(1 + m) - (T - t_n)$. The recovery equation becomes:

$$r = b_1 + b_2 \frac{1}{1 + m}. \quad (3)$$

Depth is now related to $I_n/(I_p + I_n)$. Note that the model becomes non-linear in m , the intensity ratio often used to compute depth. Double shuttering has several advantages over the single-shutter case which will be explained in Section 5.

3.3. Offset compensation

The naive method of depth reconstruction suffers from bias as a function of object intensity. That is, black and white objects at the same depth will be reconstructed as if they were at different depths. This bias can be corrected using *offset compensation*, resulting in improved reconstruction.

The theoretical model for depth reconstruction predicts that the ratio I_p/I_n remains constant for objects at equal depth. In order to validate this assertion we observed a flat target at ten known depths ranging from $10\text{cm} + \Delta$ to $100\text{cm} + \Delta$, where Δ is some unknown distance to the sensor center. The target, shown in the upper portion of Figure 9,

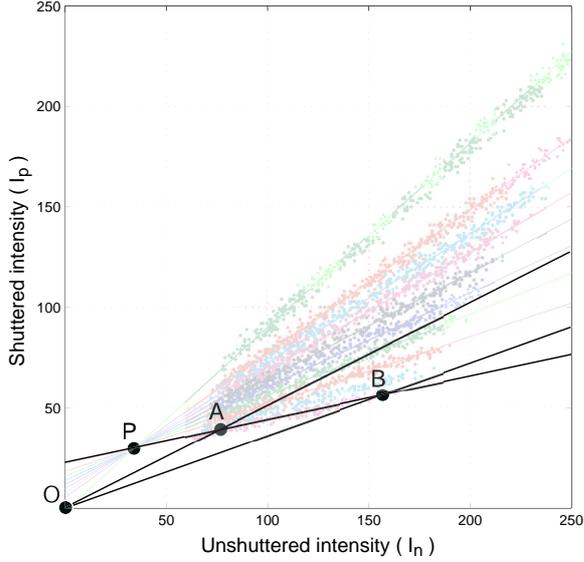


Figure 5: Two pixels with identical depth but different recorded intensities. Notice that darker pixels are the most affected by intensity bias.

was constructed with variable reflectivity. Figure 4 shows a scatter plot relating I_p to I_n for each observed pixel. We expect that all pixels observed at the same target location will generate points along a line of slope I_p/I_n .

As can be seen, points of equal depth do indeed lie on a line, however this line does not pass through the origin, as predicted by the model. Instead, the lines cross at some other point, P , offset from the origin. This offset point essentially encodes a constant unmodelled bias in our measurements. In order to include this in our model, we redefine m as $m = (I_p - P_p)/(I_n - P_n)$.

Figure 5 illustrates the computation performed with and without correction for the offset point. Consider two pixels, A and B , of identical depth relative to the sensor, one of which is darker than the other. In the naive model depth is calculated by drawing a line between each measured point and the origin, O . The slope of each line dictates the object's depth at that pixel. Clearly the slope of \overline{OA} and \overline{OB} are not the same if a measurement bias exists, and thus objects of different intensity will be reconstructed at different depths. The model introduced in this paper accounts for a constant offset in camera measurements, considering lines which intersect the offset point, P . If the location of P has been correctly determined then \overline{PA} and \overline{PB} are identical, and the computed depth of A and B will be equal.

3.4. Experimental calibration

If we know the operating range, we can set the shutter timings to be in either the single-shuttered or double-shuttered

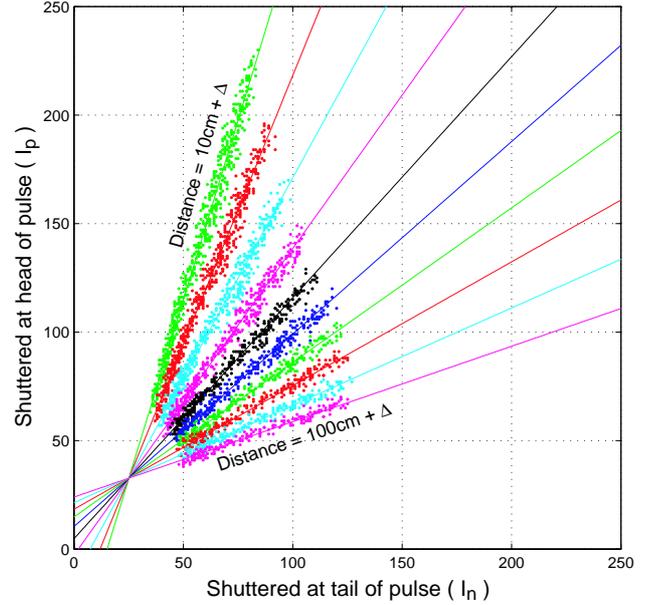


Figure 6: Plot of I_p vs. I_n for the flat target shown in Figure 9 (double-shuttered case).

case. Therefore, model calibration is simply a matter of choosing a set of conditions, selecting the appropriate model and estimating a set of coefficients, (a_1, a_2, P_p, P_n) or (b_1, b_2, P_p, P_n) .

Figure 6 shows the scatter plot relating I_p to I_n for measurements taken with double shuttering, but otherwise under conditions similar to those for Figure 5. Lines were fitted to each group of pixels, and the slopes were used to estimate the coefficients b_1 and b_2 of equation (3) using a linear regression. The result is plotted in Figure 7. The regression error is very small.

In general, despite careful configuration, it is likely that some points in the scene are single-shuttered while others are double-shuttered. Consider a point that is double-shuttered. As the distance r increases, the trailing edge of the pulse t' becomes larger, until possibly $t' > t'_n$. The conditions become single-shuttered. Likewise, if r decreases, the leading edge of the pulse t becomes smaller, until possibly $t < t_p$. The primary measurement becomes unshuttered, and we again have a single-shuttered scenario, except that I_p and I_n are reversed.

Figure 8 shows the data points for the single-shutter experiment. This time, equation (2) was used as the regression model, and a line was fit to the farthest seven data points. Notice that the values corresponding to the closest depths do not fit the computed line. These points in fact are in double-shutter condition.

It is not easy to detect when a double-shuttered condition becomes single-shuttered. Also, shutters have non-zero

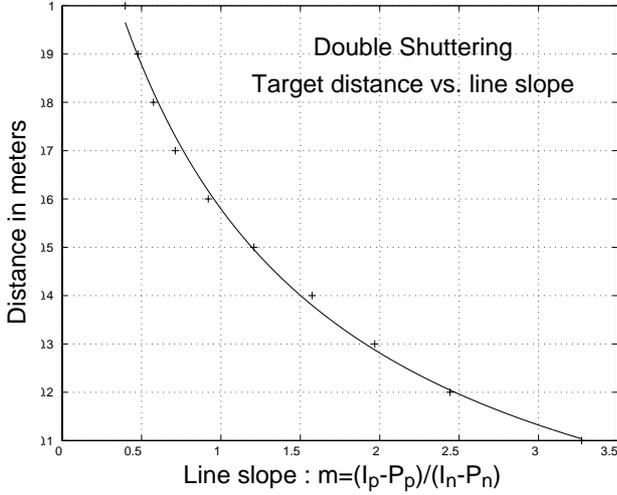


Figure 7: Plot of true depths against the ratio m for the double-shutter experiment. The continuous curve is the depth predicted by equation (3).

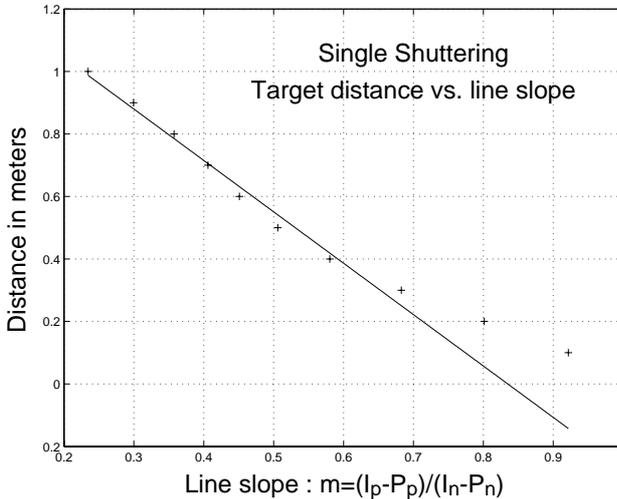


Figure 8: Plot of true depths against the ratio m for the single-shutter experiment. The values corresponding to the smallest depths are in double-shutter condition and do not obey the single-shutter equation.

fall and rise times. The rising time may well be within 1 nsec [5], but light travels 30 cm during this period. This tail effect was not considered in our model, which becomes most noticeable when the edge of an arriving pulse falls in the vicinity of a falling or rising shutter.

In practice, we increase the order of our model by adding quadratic and cubic terms to account for un-modeled tail effects. The recovery models become:

$$r = Q_3(m), \quad \text{single-shutter case,} \quad (4)$$

$$r = Q_3\left(\frac{1}{1+m}\right), \quad \text{double-shutter case.} \quad (5)$$

Here $Q_3(\cdot)$ is polynomial of degree 3. Alternatively, we could also fit a model of the type:

$$r = a + bm + c\frac{1}{1+m}, \quad (6)$$

which is simply the linear combination of the single- and double-shutter cases.

4. Multi-intensity estimation

Depth recovery precision is a function of the observed object intensity. In a scene with both dark and light objects, it is expected that the darker objects will exhibit more noise in their reconstructed depth, because imager noise has a larger effect on m when the divisor I_n is small.

Intuitively the dependance on object intensity is clear from Figure 4. Dark objects will result in measurements near the offset point. In this region of the graph many depth lines come together and noise will have a larger effect on the depth calculation.

Multi-intensity estimation improves the precision of depth in darker areas by aggregating the data from several images captured at various lamp intensities. This expanded data set yields more reliable depth estimates.

The dependance of precision on object intensity is shown in Figure 9. A target textured with a continuous ramp from white to black is placed at a constant depth in front of the sensor. The upper portion of this figure shows the target as seen by the sensor, while the lower portion shows the computed depth as a function of object intensity. It is clear that precision degrades as the object becomes darker.

Increasing the camera gain or lamp intensity will brighten the dark areas, thus increasing precision. Unfortunately, brightening may saturate the sensor in light areas, preventing the determination of any meaningful depth.

The precision of depth estimates can be analyzed in terms of their standard deviation. Figure 10 shows a plot of the standard deviation of the estimated depth as a function of object reflectivity as the lamp brightness varies. Note that as lamp brightness increases, so does precision in darker regions. However, if the lamp brightness is increased too greatly, the CCD saturates in light regions and no values can be calculated.

The traditional strategy is to set the lamp to the brightest value such that no pixels saturate, labelled "Medium lamp brightness" in Figure 10. Higher lamp brightness is of course possible, and would result in lower curves on the dark end of the plot. However, higher brightness would also result in saturation on the light pixels, with no subsequent depth calculation possible. By using the "medium

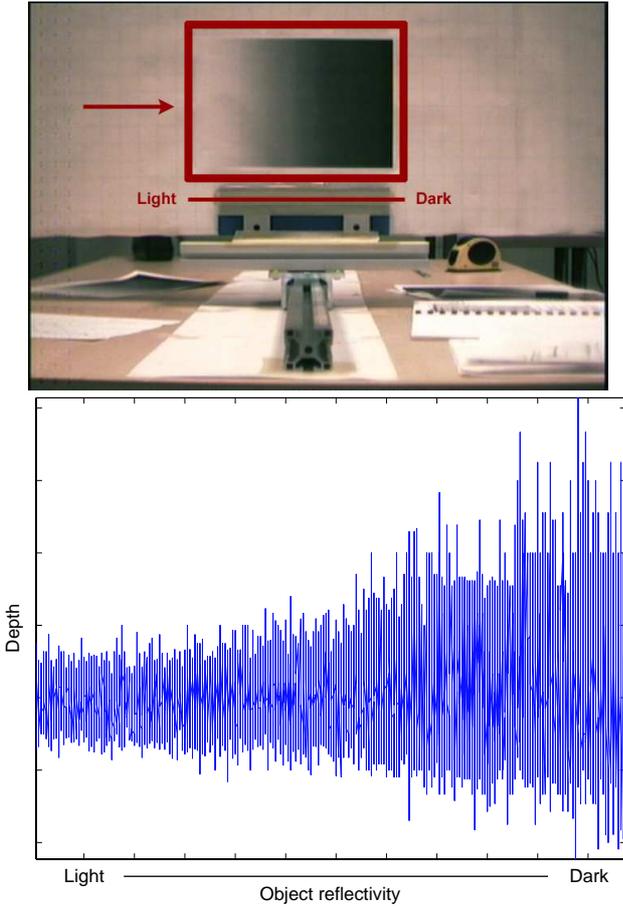


Figure 9: A textured target with a continuous ramp from white to black placed at a constant depth, and the associated precision of the estimated depth.

lamp brightness” we obtain the best single curve which extends over all object intensities. Nevertheless, by using multiple images captured under variable lighting, better results are possible. By treating pixels independently, the depth of each pixel can be calculated from the image with brightest non-saturating lamp intensity, thus higher precision can be obtained.

Figure 11 shows a plot of the observed intensity values for three different pixels as lamp brightness is increased. Note that the observations for a given pixel fall along a line. The method proposed above, of using the brightest non-saturating measurement to determine depth is equivalent to computing depth based on the slope of \overline{PA} . This slope can be estimated more reliably than (for example) the slope of \overline{PB} . It is also possible to aggregate many observations by fitting a line to all data points associated with a given pixel. A topic of future research is to analyze which method gives better results.

The location of P may be corrupted with noise. For in-

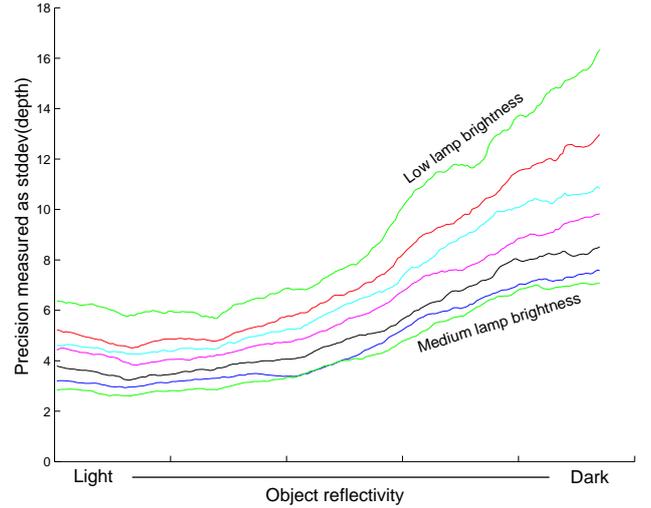


Figure 10: Standard deviation of estimated depth as a function of object reflectivity for different lamp brightnesses.

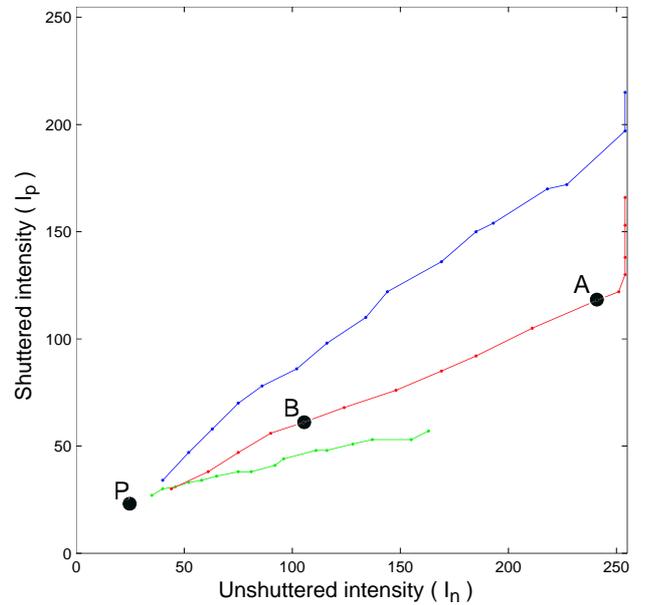


Figure 11: Plot of I_p vs. I_n for three different pixels as lamp brightness increases. Pixels with different depths move along a lines of different slopes.

stance we have found that our sensor has a cyclical low amplitude shift in the position of P . We have not yet found a way to reliably calibrate this shift, so we treat the effect as noise. Under these conditions *multi-intensity estimation* is important. The location of a point near to the location of P results in noisy estimates of slope, and thus depth. However points distant from from P result in more reliable estimates of the slope, and thus better computed depths.

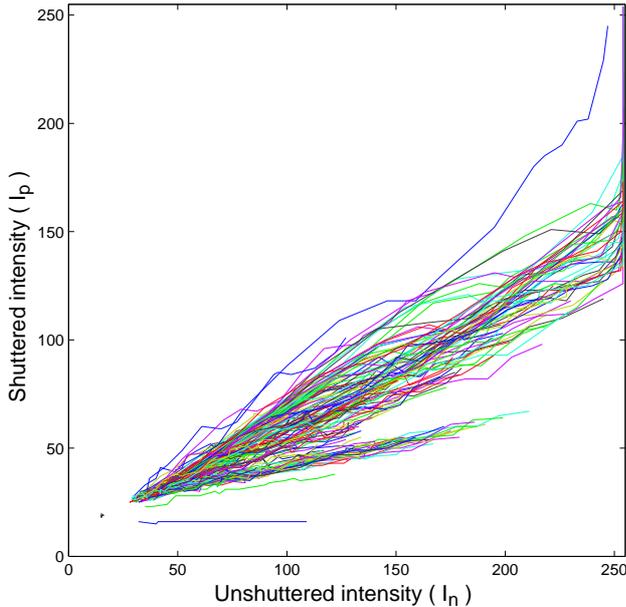


Figure 12: Plot of I_p vs. I_n for many pixels as brightness is increased. The pixels move along curves that intersect at the offset point P .

Theoretically, the location of the offset point can be calculated as a by-product of multi-intensity estimation, avoiding the need for careful physical positioning of a calibration target. Figure 12 shows a plot of the change in observed intensity for many pixels as lamp brightness is increased. Assuming that the observed scene contains objects at a variety of depths, the offset point can be estimated as the intersection of all lines. We have not yet carefully evaluated the quality of calibration obtained in this manner.

5. Double Shuttering

Depth computation relies on the ability to reliably estimate the ratio between the intensities of two images. Unfortunately noise from various sources corrupts the measurements. The effects of this noise can be minimized by making use of *double shuttering*, a method by which both cameras are shuttered, rather than only one.

The task of estimating the ratio of image intensities can be equivalently stated as the task of classifying to which line in Figure 4 a point belongs. Note that in this figure all lines are oriented between 0-45 degrees. This is because only a single shutter is employed. The unshuttered camera gathers all returned light while the shuttered camera observes only a fraction of the light, and thus can not possibly observe a greater value. However if we wish to obtain maximum discriminating ability, this narrow range is not desirable. We should arrange for lines of equal depth to expand to fill the entire quadrant of 0-90 degrees.

The desired increase in range can be obtained by shuttering both cameras. It is possible to shutter on either the head or the tail of the returning light pulse. Shuttering on the head of the pulse results in greater intensity when the object is closer. Shuttering on the tail results in the opposite condition in which intensity is greater when the object is further from the camera. By shuttering one camera on the front of the pulse and the other camera on the tail of the pulse, we obtain the desired expansion of intensity ratio range.

Figure 6 shows a plot of measurements taken with double shuttering, but otherwise under similar conditions as Figure 4. A target was moved to each of ten different depths, and the ratio of observed intensity was plotted. Note that depth is still related to the ratio between image intensities, but that the measured ratios have expanded to fill more of the available range.

In order to validate that double shuttering does in fact improve the precision of depth estimation, we evaluated both the single and double shuttered scenarios. The planar target shown in figure 9 was placed at several known depths, and each shuttering model was used to calculate the depth of each pixel on the target. The target pixels were subdivided into uniform regions so that light and dark regions of the target could be evaluated separately. Precision was evaluated as the standard deviation of the calculated depth for all pixels in a given region. Figure 13 shows a plot of true object depth versus calculated depth precision. Double shuttering performs better than single shuttering for objects at all depths. Figure 14 shows a plot of object intensity versus calculated depth precision. In this case the target was placed at $70\text{cm} + \Delta$. As previously discussed, precision is better for light colored objects and degrades for darker objects. However, double shuttering results in more precise estimates of depth in all cases.

6. Future work

Although the methods introduced in this paper are widely applicable, there are many opportunities for further enhancement. Offset compensation requires calibration of the offset point prior to use, and this paper introduces two possible methods for calibrating this point. However, since a-priori calibration is not always possible we are interested in methods for estimating this parameter directly from measured data.

During the course of this work we have empirically verified that the methods presented improve the quality of depth measurements. However careful quantitative analysis is an ongoing effort.

7. Conclusion

This paper has contributed an improved model of depth recovery using shuttered light-pulse sensors, as well as two

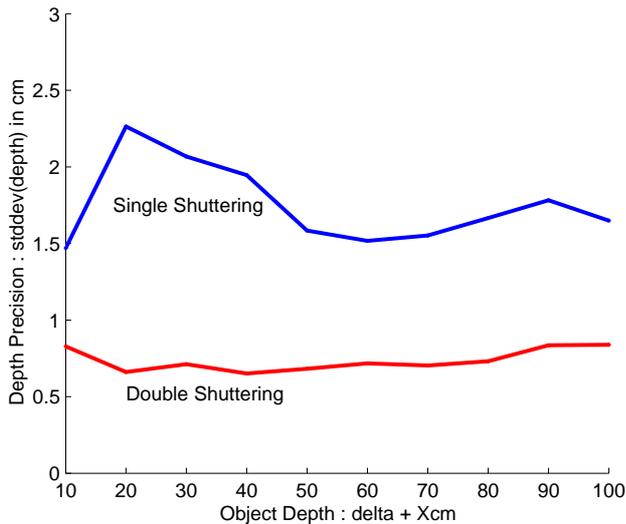


Figure 13: Object depth versus precision for both single and double shuttering. Precision is measured as $\text{stddev}(\text{computed-depth})$ for clusters of pixels on the same depth plane. Notice that double shuttering always computes depth with higher precision.

specific methods for improving the quality of recovered depth. Our model adds terms for offset compensation as well as correctly predicting sensor behavior in both the single and double shuttering scenarios.

Double shuttering is an entirely new technique and in addition to developing an analytical model we show its benefits empirically.

Multi-intensity estimation improves the precision with which depth can be estimated by using the optimal measurement from a set taken under varying lamp intensity.

Together these contributions allow enhanced shape recovery using shuttered light pulse sensors.

References

- [1] C. Bamji, "CMOS-Compatible 3-Dim. Image Sensor IC," *United States Patent*; no. US 6,323,942 B1; Nov. 27, 2001.
- [2] G. Yahav and G. Iddan, "Optimal Ranging Camera," *United States Patent*; no. US 6,057,909; May 2, 2000.
- [3] 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.
- [4] R. Gvili and A. Kaplan and E. Ofek and G. Yahav, "Depth Keying," *SPIE Electronic Imaging 2003 Conference*, 2003.
- [5] 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.

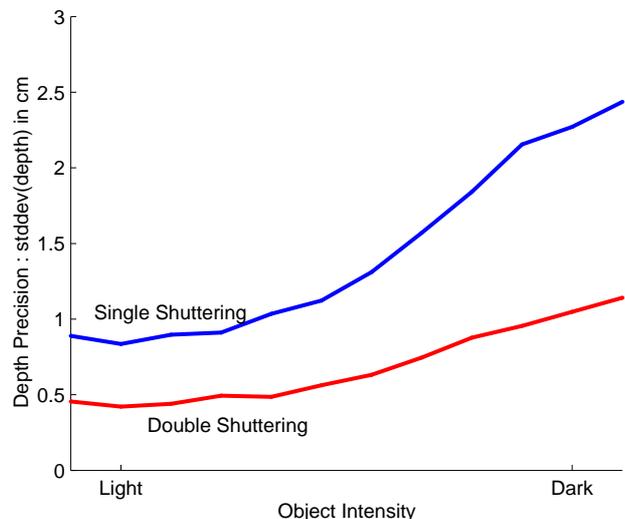


Figure 14: Object intensity versus precision for both single and double shuttering. The depth of darker objects is calculated with lower precision than the depth of light colored objects. However, notice that double shuttering always computes depth with higher precision.