Ambient Removal Theory for TetraVue Camera

Lorenzo Venneri, TetraVue

September 2017

Contents

1 Abstract

We present the problem of ambient light for the TetraVue camera, and show the reduction and subtraction strategies developed at TetraVue so far.

2 Effect of ambient light

2.1 Ambient Light Magnitudes

We can calculate the expected ambient for our system. The solar signal depends only on the solar intensity and reflectivity, not the distance to the target due to the difference between intensity and flux. If this point is not congealing, consider taking a photograph of an infinitely large white wall illuminated by the sun. If you were to image it with the same lens, it would have the same signal regardless of how far away from the wall you stood.

We can use the basic approximation for sensor illumination

$$I_{sensor} = \frac{I_{scene}}{4F^2}$$

where F is the F number (used F/1.5).

The solar signal measured at each pixel is

$$s_{solar} = \frac{I_\lambda w A_{sensor} QE}{4F^2 E_\lambda n}$$

where I_λ is the solar flux ($1\frac{W}{m^2 nm}$ at 800 nm), w is the bandpass (10 nm), A_{sensor} is the area of the sensor ($.67cm^2$), QE is quantum efficiency (30 to 40 percent), E_λ is the photon energy ($h\nu = 2.510^{-19}J$), and n is the number of pixels in the sensor (2.2 MP). This yields about 40 photons per micro second per pixel in our system.

Tests were conducted to measure how much ambient light the camera has to deal with, by measuring the reflected solar light off of a white target, which is distance invariant as the incident intensity is constant over distance. The distance invariance was verified by moving the target through range without an observed change in intensity. I used the StingRay 16mm (40 dFOV) lens at F/1.5 on a CMV2000 sensor which has KADC of 2.16 e-/DN and 40 percent QE. For 50 us exposures, the sensor measured ambient at 600 DN, or 3300 impinging photons. For a single pixel then, our system receives 65 photons per us, of which it measures about 30.

The predicted and measured ambient photon fluxes are close enough to make the prediction function useful. Deviation is likely due to transmission properties of the optics, exposure time misreporting, and gross approximation. A more useful formulation in terms of the pixel pitch area is shown below.

$$s_{solar} = \frac{I_\lambda w A_{pixel} QE}{4F^2 E_\lambda}$$

2.2 Effect of ambient light

Ambient light is a linear offset to the intensity of the images we capture. Because the ambient is spread evenly over the entire exposure, and is not pulsed, it does not contain timing or range data when modulated by the PC. If we consider the ratio, where we normalize the modulated signal, and then add ambient to the component images, we can see that as the ambient becomes large relative to the pulse signal, the ratio approaches the ratio of the ambient signal. You can think of the ambient ratio as sink, which all ratios approach as you add ambient light.

$$ratio = \frac{C1 + C1_a}{C1 + C1_a + C2 + C2_a}$$

$$\lim_{A \to\ large} ratio = \frac{C1_a}{C1_a + C2_a}$$

The ratio of the ambient signal is not necessarily .5 because when a high contrast Pockels Cell transmits light through one polarization, it does not transmit to the other. The ambient light is a CW source, and will pick up the modulation of the PC over the sensor exposure. This means the placement of the modulation ramp within the exposure affects how much ambient light goes to each polarization image. This can be used to gate one image more than another. In general, the ramp timing relative to the exposure should be constant during all operation. There are also reasons for placing the ramp at the end of the exposure, so that any pulsed sources (non modulated) inside the sensor exposure are measured in the same ratio as the ambient and can be subtracted. We are referring to the multiple camera confusion phenomenon. Another reason to place the ramp at one end of the exposure is to force ambient ratio sinks to be outside of the useful ratio region defined by amin and amax.

3 Analysis

3.1 General framework

For convenience, normalization for the input beam is assumed. For a real system however, a normalizing measurement must also be made so that variations in the input beam do not affect temporal reconstruction.

The time frame is the exposure time, going from 0 to t_0, the length of the exposure.

An input beam of light, $f(t)$, is defined on the same interval as the exposure.

A series of N modulators, $p_n(t)$, is defined on the same interval as the exposure, with modulation possibly varying over time. There are N matching integrating measurement devices, the pixels, whose measurement is defined by m_n.

The beam passes through the modulator medium, thus picking up a modulation. Each pixel then integrates the entirety of the modulated light to measure a single number representing the total measured flux of the beam after undergoing a particular modulation.

$$m_n = \int_0^{t_0} f(t_e - t) p_n(t) dt$$

$f(t)$ is reversed as it is being convolved with the modulation function, so it is evaluated on the modulation frame in reverse.

3.2 Case 1: A known input and a single known modulator

This is the normal TetraVue case in which the input beam and modulators are known functions due to our calibration procedure, and they are constants for every single frame.

For clarity, the known variables are:

- $f(t)$, defined only by the laser diode pulse shape and measured during calibration
- $p_n(t)$, defined by the Pockels Cell modulation ramp and measured during calibration
- m_n, defined the pixel's integrated measurement of f(t) modulated by $p_n(t)$

And the camera seeks to find:

- α in $f(t+\alpha)$, the timing delay that characterizes the pulse position relative to the modulation function and by extension, the time of flight and distance

The measurement can then be restated as

$$m_n = \int_0^{t_0} f(t_e - (t + \alpha))p_n(t)dt$$

Assuming a normalized input (which can be made normalized by a measurement without modulation, or indeed any independent modulation), α is the only unknown, and we require only a single modulation function. We can solve for the case that $f(t)$ is a Gaussian and $p_1(t)$ is a half wave sinusoidal (have not done yet).

This should reveal that the single modulation must be monotonic in order to uniquely find α.

3.3 Case 2: A known input with an unknown constant offset and a single known modulator

When the TetraVue camera is taken outdoors and the ambient light overpowers the bandpass filter and exposure limiters, Case 2 is relevant because the input beam is the sum of the known laser diode pulse as well as a linear offset from the ambient illumination. But now, the ambient light complicates $f(t)$ such that

$$f(t) = f_p(t) + f_a(t)$$

where $f_a(t)$ is the portion of the input beam from the ambient light which we do not know.

For clarity, the known variables are:

- $f_p(t)$, the laser diode pulse shape portion of $f(t)$

- $p_n(t)$, defined by Pockels Cell modulation ramp
- m_n, defined the pixel's integrated measurement of f(t) modulated by $p_n(t)$

And the camera seeks to find:

- α in $f_p(t{+}\alpha)$, the timing delay that characterizes the pulse position relative to the modulation function
- $f_a(t)$, the ambient portion of the input function

A simplification can be made by assuming $f_a(t)$ is a time independent offset because we do not usually expect the solar flux to change appreciably during a short exposure ($<$ 50 us), though it easily varies frame to frame or pixel to pixel. We can then rewrite $f(t)$ as

$$f(t) = f_p(t) + f_a$$

$$m_n = \int_0^{t_0} (f_p(t_e - (t + \alpha)) + f_a)p_n(t)dt$$

$$m_n = \int_0^{t_0} (f_p(t_e - (t + \alpha))p_n(t)dt + f_a \int_0^{t_0} p_n(t)dt$$

which indicates that the measurement is the sum of the integrated modulated pulse and integrated modulated ambient, but that the ambient has no time dependence, and so the ambient component of the measurement simplifies to the ambient times a known factor. It is a known factor in the TV camera case because we can measure the modulation function of a time independent input signal which is simply the integral of the modulation function. For example, we can measure this integral by pointing the camera towards a bright blank wall and averaging frames that are captured in the same way as normal frames but without pulsed illumination. The exposure length and timing settings must remain unchanged to include the true modulation any time dependent modulation artifacts like ringing.

There are then two unknowns, and two independent equations are needed to solve the system. We assumed a normalized input; therefore, in practice, a third equation and measurement for the no modulation case (or an independent modulated pixel) is needed.

Calling

$$m_1 = \int_0^{t_0} (f_p(t_e - (t + \alpha))p_1(t)dt + f_a \int_0^{t_0} p_1(t)dt$$

$$m_2 = \int_0^{t_0} (f_p(t_e - (t + \alpha))p_2(t)dt + f_a \int_0^{t_0} p_2(t)dt$$

$$f_a = \frac{m_2 - \int_0^{t_0} (f_p(t_e - (t + \alpha))p_2(t)dt}{\int_0^{t_0} p_2(t)dt}$$

$$-m_1+\int_0^{t_0}(f_p(t_e-(t+\alpha))p_1(t)dt+\left(m_2-\int_0^{t_0}(f_p(t_e-(t+\alpha))p_2(t)dt\right)\left(\frac{\int_0^{t_0}p_1(t)dt}{\int_0^{t_0}p_2(t)dt}\right)=0$$

Which is an equation with a single unknown, α.

3.4 Case 3: An unknown input and N known modulators

Unrelated to the TetraVue ranging application, but worth ,mentioning for its possible applications in the observation of time variable phenomena within a single exposure using 2d integrating arrays, is the extension to N modulators and an unknown input function.

For clarity, the known variables are:

- $p_n(t)$, defined by a particular modulation scheme
- m_n, defined the pixel's integrated measurement of f(t) modulated by $p_n(t)$

And the camera seeks to find:

- $f(t)$, the unkown input

Motivated by Fourier Series, we consider how to construct the elements of the Fourier Sine Series (FSS) using the integrated measurements to finally obtain a reconstruction of the original input, $f(t)$.

We design the modulators to produce sinusoidal shapes varying between 0 and 1.

$$p_n(t)=\frac{sin(\frac{\pi}{2}nt)+1}{2}$$

We can then adjust the measurement which is now

$$m_n=\int_0^{t_0}f(t_e-(t+\alpha))\frac{sin(\frac{\pi}{2}nt)+1}{2}dt$$

to match the FSS coefficients.

$$m_n'=4(m_n-m_0/2)$$

where m_0 is an unmodulated integration of $f(t)$.

The FSS representation of $f(t)$ is then approximated by

$$FSS(f(t))=\sum_{n=1}^{N}m_n'p_n(t)$$

and we can discern the dominant temporal features of $f(t)$; specifically, the frequency components up to $\frac{N}{t_0*2}$, by the Nyquist limit. This requires being able to modulate the signal on a per pixel basis, perhaps using LCD technology or pixels with different response curves. Furthermore, the super-pixel containing the N separately modulated pixels must receive the same input beam, requiring a rather large point spread function.

4 Ambient Reduction Strategies

The ambient reduction strategies seek to lower the contribution of f_a relative to f_p such that the offset in range is within error bars or tolerable by the user. Below are briefly described approaches.

1. Bandpass Filter
   The tighter the band, the better the reduction in ambient. Beyond a 15 nm bandpass, reductions in band have marginal improvements in ambient reduction while potentially interfering with diode wavelength drift (¡.5 nm per degree) and spread (2-10 nm) and so lowering the measured signal.

2. Implicit Exposure Control by a Single High Contrast Pockels Cell
   The Pockels Cell ramp is on the order of 100 ns, which is much less than the total exposure time of 1 to 20 us. The ramp is usually placed in the middle of the exposure, so that for a given aligned analyzer polarization, the PC is off for the first half and on for the second, thereby halving the effective ambient exposure. To be precise, it is the product of the PC over the entire exposure period, which includes any ringing and contrast of the PC.

3. Exposure Control by Sensor Exposure

4. Exposure Control by Second Pockels Cell or Other Fast Gate

5 Ambient Subtraction Strategies

Here, we seek to subtract the ambient photons to recover the lone pulse signal. All the methods somehow measure f_a which can be subtracted from each image in the standard ratio.

5.1 3 State Polarizer

The theory behind the ambient subtraction via 3 state polarizer was shown above where a system of two independent equations is required to solve for α, the time delay, and the f_a, ambient. We present here a practical quantity that fits into the current pipeline that is intensity independent but retains time dependence without an ambient error.

$$ratio_{AA} = \frac{m_1 - a_{m_1/m_3} \cdot m_3}{m_1 - a_{m_1/m_3} \cdot m_3 + m_2 - a_{m_2/m_3} \cdot m_3}$$

where

$$a_{m_1/m_3} = \frac{\int_0^{t_0} p_1(t)dt}{\int_0^{t_0} p_3(t)dt}, a_{m_2/m_3} = \frac{\int_0^{t_0} p_2(t)dt}{\int_0^{t_0} p_3(t)dt}$$

where a_α and a_β are the fractions of ambient exposure relative to the θ state, for the p_1 and p_2 modulations relative to p_3. Note that each term in this ratio contains an intensity factor (variability of the scene), so that the factor cancels and the ratio is, in effect, computed with normalized quantities.

The ambient fraction that each state measures relative to the other states allows us to subtract the ambient from the components of the ratio. That third state must simply be a distinct state from the others, otherwise we would always subtract to zero or not get any new information.

One knows that the ambient scales linearly for all three states, and thus the relative fractions of ambient light are constant between the three states. That is, if one state receives a third as much ambient as another, it will always do so. The ambient component of one state can then be eliminated by subtracting the value of another state multiplied by the relative ambient fraction. The ambient component is removed, and we are left with a composite modulated signal which has the same kind of timing information as the original single modulated signal except that it is ambient free by definition and has higher noise. Note, that we cannot add image operations without also adding noise. The point of dividing is to remove the scale factor from the modulation signal which depends on the reflectivity and distance of the object. As long as the scale factor is divided out, and the signal changes based on timing (i.e. contains timing information), we have a calibration ready measurement. That it no longer varies between 0 and 1 is not obviously important.

The sensor must behave linearly for both pulsed and CW sources in order for these linear operations to work. Perhaps this is best accomplished by a multi pulse method whereby the single highly energetic pulse is spread thin.

5.2 High Frequency Pattern Illumination

This method obtains a number for the ambient offset f_a by finding the global component of the illumination. It doubles as a correction to multipath.

We aim to correct multipath by separating the total measured light into direct and global components, those being the single path and multipath signals, respectively. The most well known lighting based techniques were first expounded by [Nay06] and usually involve multiple images. The single image techniques involve patterned illumination, such as high contrast dark and bright stripes or dark and bright speckle patterns with about twice the angular spacing as the pixels, to reconstruct a lower resolution version of the direct and global components. The illumination pattern must have the same FOV as the imaging assembly in order for the stripes to have a constant angular width with distance. The dark and bright points in the scene serve as control and test points respectively. The bright test points measure the entire signal, both the direct and the global illumination. The dark control spots measure only the global component. It is assumed that nearby areas have the same behaviour, and so the global component in the dark can be subtracted from the bright spot to find the direct component, which necessitates some loss in resolution. In practice, we take into consideration the contrast and fraction of the scene illuminated to create a useful functional form of the problem.

Ambient light falls into the global component, and so high frequency pattern illumination offers another method for ambient subtraction.

The first version of patterned illumination for direct lighting estimation has a resolution loss. [Sub17] outlines a method that maintains resolution and has consistently lower RMSE. The concept is the same, expect the analysis occurs in frequency space. When illuminating a scene with a high frequency pattern, the direct component of the light preserves the high frequencies of the illumination. The global and so ambient light component is predominantly of low frequency makeup, regardless of the illuminating frequency. This difference in the frequency domain allows the separation of the two components using basis representation.

5.3 Separate dual exposures

As explored in the overlapping dual exposure case in a separate paper, ambient frames should be recorded at longer exposure times to increase their S/N and minimize the noise added by additional image operations. Care must be taken to avoid saturation, and perhaps a variable exposure time can be used depending on scene. Linearity for different exposures and pulsed source problems are also present here. The α coefficient should be calculated on a per pixel basis as an additional step during calibration, and a LUT used to divide the long exposure ambient signal before it is subtracted from the pulse or data exposure. The background or dark level frame is different for different exposure lengths.

6 Aggregate Solution

There are strategies for ambient reduction: the bandpass filter, short exposures, gating; and there are strategies for ambient subtraction: ambient frame, 3 state polarizer, and high frequency illumination. The former must be used to conserve the sensor's dynamic range so that most of the signal measured comes from the useful diode light rather than the ambient. The alpha prototype achieved ambient levels of 15 percent of dynamic range for 100 percent reflective objects in outdoor environments, which renders at least 90 percent of the DR useless for range measurement. If no ambient subtraction solution is provided, a reduction of 300 times compared to the alpha prototype is required to make 95 percent of the DR useful for range measurement. These numbers come from simulating different ambient effects on range and where useful data deviate from ground truth by less than 10 percent. The 300 times ambient reduction might come from a 20x decrease in the sensor exposure time, 2-4x reduction due to PC gating, and a 2-3x bandpass improvement. This is difficult to achieve. Ambient subtraction is needed. More probably we can expect up to a 100x reduction, or .15 percent of the dynamic range.

Subtraction must be used to eliminate whatever dilution the ambient signal adds, which will have increasingly dramatic effects for lower pulse signals. Given sufficient DR, the subtraction tools should prove effective for ambient levels 5-10 times the pulse levels (i.e. if the pulse is already half the dynamic range, there is at most the same amount of ambient light as pulse light without saturating). A white target at its full ambient-free range, has signal just above the noise floor which is about 0.2 percent of the dynamic range. To subtract ambient without adding excessive noise, the ambient level can be no more than 5-10x the minimum pulse signal, and so it follows that ambient reduction strategies must lower the worst case ambient level to 2 percent or less of the dynamic range. This two percent is a function of the entire optical assembly: QE, lens, optics, gating, etc. The dynamic range, or the signal for which a range measurement can be made, is reduced by 4 percent relative to an ambient free measurement.

References

[Nay06] Shree K. Nayar. "Fast Separation of Direct and Global Components of a Scene using High Frequency Illumination". In: (2006).

[Sub17] Art Subpa-asa. "Direct and Global Component Separation from a Single Image Using Basis Representation". In: (2017).