UCRL-ID-154786

Motion Detection using Mean Normalized Temporal Variance

C. Chan

U.S. Department of Energy

Lawrence
Livermore
National
Laboratory

August 4, 2003

Approved for public release; further dissemination unlimited

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

Approved for public release; further dissemination unlimited

Motion Detection using Mean Normalized Temporal Variance

Carri Chan 8/4/03

Introduction

Scene-Based Wave Front Sensing uses the correlation between successive wavelets to determine the phase aberrations which cause the blurring of digital images.[1] Adaptive Optics technology uses that information to control deformable mirrors to correct for the phase aberrations making the image clearer.

The correlation between temporal subimages gives tip-tilt information. If these images do not have identical image content, tip-tilt estimations may be incorrect. Motion detection is necessary to help avoid errors initiated by dynamic subimage content.

With a finely limited number of pixels per subaperature, most conventional motion detection algorithms fall apart on our subimages. Despite this fact, motion detection based on the normalized variance of individul pixels proved to be effective.

Method

Motion detection is achieved by thresholding the normalized variance of each pixel. Previous work has been done using image statistics to isolate moving objects. While these methods are effective for their applications, they either require some knowledge of the image motion which we do not have,[2] require images over 10 times larger than our resolutions,[3] or require more computation power and memory than we can afford.[4] By looking at the statistics of each pixel independently and by performing pixel-by-pixel operations on a finite number of stored frames, motion detection can be done in $O(n^2)$ time and $O(n^2)$ space for an nxn image.

A pixel's value is modeled as the sum the realization of a Poisson process with mean and variance equal to its noise-free intensity value and Poisson read-noise. The mean-normalized variance provides more useful information for motion detection because pixels of higher intensity have higher variance, but equal normalized variance, while pixels with varying intensity have higher normalized variance.

Types of Images

The information we will take from the normalized variance is in two parts. First we must differentiate what type of image is being viewed: a stationary image in which camera and scene are static, a panning image in which the camera is moving over the image, or a moving image in which the camera is static but objects, such as cars, are mobile within it.

The next step is to determine how to estimate phase aberrations. If the scene is stationary, SBWFS will continue to work. If it is a panning image, the tip-tilt sensors will be shut off because either the mirror positioning will max out or the system will go unstable. Finally, if it is a moving image, one may be able to adapt the SBWFS algorithm to account for movement; otherwise, the scene will be treated similarly to a panning image.

Pixel Models

Image motion can cause edges to initiate false alarms. In order to differentiate truly moving pixels, we categorize pixels as stationary, edge, or moving. The ratio of moving to non-moving pixels allows us to determine what type of image is being viewed.

The motion detection threshold is determined based on the probability model and expected normalized variance of each pixel type. An accurate approximation for the expectation of the normalized variance can be calculated based on the expected value of the sample mean and the

sample variance. Where n is the number of samples, x_i is the ith sample and is a random variable with mean μ_i and σ_i^2, the expected value of the sample mean. M_n, is:[5]

$$E[M_n] = 1/n * \sum_{i=1}^{n} \mu_I$$

and the expected value of the sample variance, S_n^2, is [5]

$$E[S_n^2] = 1/n * E[\sum x_I^2] - 1/n^2 * \sum_{i=1}^{n}\sum_{i=1}^{n} E[x_i x_j]$$

and the approximation of the expected value of the sample variance, N_n^2, is:

$$E[N_n^2] = E[S_n^2] / E[M_n]$$

Stationary Pixel

The noise model of a stationary pixel is purely Poisson. The pixel value is characterized as a Poisson process with its mean and variance equal to the actual temporally static pixel value of the noise-free image. Poisson read-noise is added on top of that. One must note that the noise in this system is independent.

Where n is the number of samples, the expectation of the mean-normalized variance of a stationary pixel is:

$$E[N_n^2 | \text{ pixel is stationary}] = (n - 1) / n \approx 1$$

Edge Pixel

Random subpixel image shifts can cause edge pixels to be incorrectly flagged as moving. The noise model of edge pixels is the same as that of the stationary pixels. An edge is modeled as being one pixel wide. The intensity of an edge pixel is a random variable that shifts from its mean which is equal to the average of the intensity levels, μ and μ, on each side of the edge. Each shift is a zero-mean Gaussian random variable with variance σ_s^2, where σ_s^2 has a maximum of .01 pixels. Where n is the number of samples, and σ_n is the Poisson read-noise parameter, the expectation of the mean-normalized variance of an edge pixel is:

$$E[N_n^2 | \text{ pixel is edge}] = (n - 1) / n * [\mu/4 + \mu/4 + \sigma_n + (\mu - \mu)^2 * \sigma_s^2] / [\mu/2 + \mu/2 + \sigma_n]$$

For an edge that goes from 0 counts to 250 counts, the expectation of the normalized variance is over 5 times larger than the expectation for a stationary pixel.

Moving Pixel

A moving pixel begins at some intensity level and increases with a constant slope, k, for n samples. The noise model of a moving pixel is identical to that of edge and stationary pixels. Where n is the number of samples, h is the sum of starting intensity value and σ_n which is the Poisson read-noise parameter, and k is the slope of the ramp, the expectation of the mean-normalized variance of a moving pixel is:

$E[N_n^2 | \text{ pixel is moving}] =$

$$[(k+2kh)*(n+1)*(1-1/n)/2 + k^2*(2n+1)*(n+1)/6 - k^2*(n+1)2/4 - hk(n^2-1)/n + h - h/n] / [k*(n+1)/2 + h]$$

For a moving pixel that moves from 0 to 250 counts with a slope of 50, the expectation of the normalized variance is nearly 8 times larger than the expectation for an edge pixel with the same range. This large discrepancy suggests that we should be very successful using the normalized variance for motion detection.

Results

Over a thousand realizations of each random process of a pixel value was run for each pixel type. These simulations, which varied in range of peak and base values and number of samples, verified our approximations for the expected values of the normalized variances.

Choosing an n

With standard Shack-Hartmann subimages being 16x16 or 32x32 pixels, most moving objects, such as cars, are likely to be on the order of 3 to 5 pixels. As such, a reasonable size for n is 5 because a car with a size of 5 pixels would take 5 frames to fully traverse a single pixel.

While more frames would make the sample variance and sample mean a more accurate measurement, requiring more frames would not be beneficial. In order to use SBWFS to find slopes, the frames must be correlated. However, frames separated by many time steps are not likely to be correlated. Also, if the AO system must wait for more frames, it might not be able to correct the phase aberrations quickly enough.

Choosing a Threshold

The expected normalized variance for the edge and moving pixels change with varying ranges and background levels (Chart 1). The expected normalized variance for stationary pixels is not shown because it is always lower than the expected values for edge and moving pixels. A good threshold depends on the illumination and background levels of the subimage. Too low a threshold will result in many false alarms while too high of one will result in many misses. For a given background level, the expected normalized variance of moving pixels is over 5 times larger than that of an edge pixel. In order to minimize false alarms and misses, the motion detection threshold is set to 1.25 times the maximum expected normalized variance of each pixel.

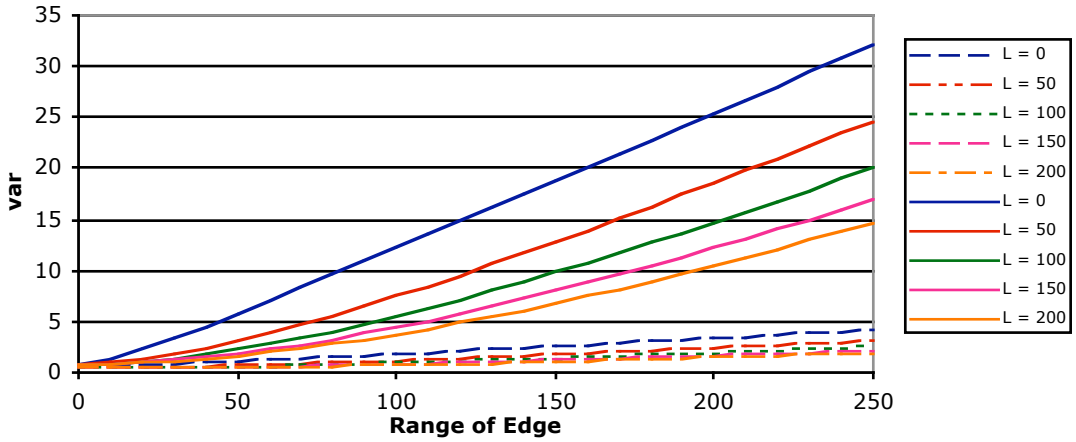

Chart 1: Expected normalized variance of edge and moving pixels. Edge pixels are dotted lines, moving pixels are solid lines. L is the background level of the pixel.

How does this work on actual images?

By thresholding the normalized variance of individual pixels, motion detection is successfully achieved. For noise-free images, the motion detection is extremely accurate for 32x32 and 16x16 images (Fig. 1a, b). These images are of a car driving down a street. The white with black border isolate the pixels of the moving car which the motion detection algorithm detected.

Unfortunately, noise and image shifts challenge the accuracy of this technique. The threshold is based on the expected normalized variance of edge pixels. It is very likely that the normalized variance of some edges will go above the threshold causing false alarms. Also, stationary pixels may be detected as moving simply because they have high variances due to the randomness of

noise. These false alarms are quite regular for noisy and shifted images, especially as the background level of the image increases and the SNR decreases. Figure 1c show 8 false alarms that are dispersed randomly throughout the image. Because we know that for our resolution, most cars are 3 to 5 pixels large, it is unlikely that a single pixel with high normalized variance, stranded by itself is part of a moving object. In order to find clumps of pixels with high normalized variance, a lowpass filter was used on an image mask where moving pixels are equal to 1 and non-moving pixels are equal to 0. The lowpass filter filters out single pixels with high normalized variance and leaves behind the pixels which are most likely to be part of a moving car. The results of this lowpass filter technique can be seen in figure 1d. Despite the noise and image shifts, the moving car is detected just as accurately in 1d as in the original image of 1b.

A. 16x16 image **B.** 32x32 image **C.** 32x32 noisy/shifted image with no lowpass filter **D.** 32x32 noisy/shifted image with lowpass filter

Figure 1: Motion Detection of a car driving on a street. The white with black border highlight the pixels detected as moving.

Illumination and background levels significantly affect SNR and, hence, the performance of motion detection. The percentage of false alarms is defined as the number of pixels falsely marked as moving divided by the number of non-moving pixels in the noise-free version of the same image. The percentage of misses is defined as the pixels not marked as moving in the noisy image divided by the total number of pixels marked as moving in the noise-free version of that frame. Motion detection is very consistent with exposure level, only degrading for fewer than 50 counts per pixel (Chart 2a). When the background level increases beyond the illumination level, false alarms increase rapidly (Chart 2b). Fortunately, beyond this point, motion detection is not necessary since images may be too degraded to use for slope detection using SBWFS.

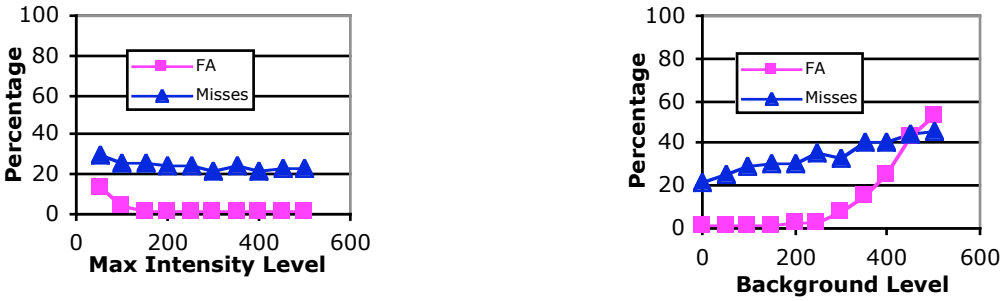

Chart 2: Percent of False Alarms/Misses due to A. Changing Illumination B. Changing Background

Thresholding mean-normalized variance to achieve motion detection is generally more successful when a small fraction of the image is moving and when there are few sharp contrast edges. The number of false alarms decreases significantly for these types of images. This disparity is most likely due to image shifts. Image shifts cause the variance of a pixel to increase. For pixels that are stationary and are surrounded by mostly stationary pixels, these increases are not large enough to initiate a false alarm. However, edge pixels may shift enough to cause false alarms and moving pixels may shift to cause pixels previously traversed to be marked. Fortunately, this phenomenon does not hurt our usage of this motion detection algorithm. If images have enough motion to cause so many false alarms, SBWFS will be unsuccessful for slope estimation.

Conclusion

Using a threshold on the mean-normalized temporal variance of individual pixels proved to be an effective method for motion detection. While many standard image-processing techniques failed on our low-resolution images, this method successfully detected moving objects on our 16x16 and 32x32 images. The next step will be to take this information and determine how to use it to minimize errors in tip-tilt information when there are moving objects within surveyed scenes.

References

[1] Poyneer, L., "Scene-based Shack-Hartmann Wave-Front Sensing: Simulation and Analysis," submitted to Applied Optics

[2] Y. Ren, C. Chua, Y. Ho, "Motion Detection with Non-stationary Background," IEEE Image Analysis and Processing Proc., pp. 78-83, 2001

[3] Haritaoglu, D. Harwood, L. Davis. "A Fast Background Scene Modeling and Maintenance for Outdoor Surveillance," IEEE Pattern Recognition Proc., 4, pp 179-183, 2000

[4] S. Blostein, T. Huang. "Detecting Small, Moving Objects in Image Sequences Using Sequential Hypothesis Testing," IEEE Trans Sig. Proc. 39, pp 1611-1629, 1991

[5] Drake, A., *Fundamentals of Applied Probability Theory*, (McGraw-Hill, Inc., NY, 1967)