

Illumination Compensation Based Change Detection Using Order Consistency

Vasu Parameswaran, Maneesh Singh, Visvanathan Ramesh
Real Time Vision and Industrial Imaging
Siemens Corporate Research, Siemens Corporation, Princeton, NJ 08540

Abstract

We present a change detection method resistant to global and local illumination variations for use in visual surveillance scenarios. Approaches designed thus far for robustness to illumination change are generally based either on color normalization, texture (e.g. edges, rank order statistics, etc.), or illumination compensation. Normalization based methods sacrifice discriminability while texture based methods cannot operate on texture-less regions. Both types of method can produce large missing regions in the distance image which in turn pose problems for higher-level processing tasks that may be shape or region-based and require accurate foreground masks (e.g. person detection and tracking, crowd segmentation, etc.). Texture based methods have an additional problem in that they produce false alarms due to textures induced by local illumination effects (e.g. cast shadows). In this paper we propose a compensation based approach for change detection. Prior work on compensation has largely taken an empirical approach, and has not dealt with the important problem of rejecting outliers when they dominate the scene. In contrast, our generative approach and systematic handling of outliers enables us to achieve robustness to illumination change while eliminating the problems mentioned above. Furthermore, the computational complexity of our method is low enough for real-time performance. Results comparing images taken under strongly different illumination conditions, demonstrate the power and generality of the proposed method.

1. Introduction

Change detection is a fundamental problem for visual surveillance with a static camera and forms an important pre-processing step for various applications in the domain. Although the goal is simple, i.e. quantify significant deviations from a given background image, perturbations arising from nuisance variables from the scene and camera pose obstacles. A key perturber is illumination change, arising from variations in scene lighting as well as camera gain

changes designed to improve the dynamic range of the image. Not surprisingly, there is a large body of research work on change detection algorithms resistant to illumination change, that has achieved relative success. Among these, *order consistency* based methods, which are based on the relative ordering of pixels in an image patch, form a class of robust non-parametric methods for removing the effect of illumination change on the patch. While such approaches are sound and fast, they suffer from two fundamental effects that limit their power and which have not been tackled so far in the literature. Firstly, pixel order is meaningful and robust only when a block of pixels is considered. Secondly, pixel order is meaningful only when the block of pixels has texture (i.e. where at least a few pixels are of statistically different intensity). Together, these two phenomena render the resulting foreground mask to be coarse and with missing areas (e.g. the bottom left image in figure 1). Additionally, spurious rank order changes can be induced by cast shadows for blocks that are on the shadow boundary, causing false alarms. In this paper we circumvent these phenomena by using order consistent blocks to calculate an illumination correction which can be applied to an input image before change detection. The method can be used for correcting for global illumination changes (e.g. ambient light change, camera gain change) as well as local illumination changes (e.g. those resulting from a change in the number of illuminants and/or their directions).

Besides change detection, a key benefit of the method is in background maintenance. In outdoor applications, backgrounds are typically updated constantly to handle fast illumination changes. Foreground objects corrupt the background when updates are done naively. Sophisticated algorithms control updates based on higher level detection. However, when foreground pixels cover a large part of the scene, updates are delayed, leading to change detection errors. The proposed method can obviate or reduce the need to frequently update the background.

The rest of the paper is organized as follows: We discuss prior work in section 1.1 and review rank order consistency in section 2. Section 3 develops a generative model for illumination change between a pair of images and shows how

we can use the model and order consistency to compensate images for global illumination change. Section 4 discusses approximations to the model, and section 5 covers the local illumination case. We present results in section 6 and conclude in section 7.

1.1. Prior Work

Radke et al [18], in their survey on change detection, cover a number of papers on illumination invariance. Rather than repeat them here, we highlight the main approaches, which are based on normalization, texture, and illumination compensation. An early attempt at illumination invariance, intensity normalization, can at best handle global illumination and does not work well when foreground dominates the image. Another early attempt, normalized color, is designed for a single illuminant. Comprehensive color normalization [10] which combines normalized color and ‘grey-world’ normalization, can only work with a single illuminant and cannot remove shadows. Finlayson et al [8] remove shadows by first calculating an illumination invariant image (free of shadow), the gradient image, and finally integrate the gradient after shadow reasoning, to get a shadow free image. While their integration step is computationally expensive for real-time application, the invariant calculation step is cheaper. We compare the proposed method to one based on the invariant image and demonstrate better performance. Texture based methods exploit the illumination invariance of local features to detect change ([3], [24], [22], [17], [15], [20]). As pointed out before, these methods cannot work on texture-less regions and produce blocky output.

Methods based on ‘scattergrams’ (also referred to as ‘comparagrams’ or ‘joint intensity histograms’) are closer in spirit to this work. A scattergram is a square matrix showing the number of times a pixel of intensity i in one image was mapped to intensity j in another one. Bromiley and Thacker [4] and Kita [14], use the scattergram to classify pixel pairs of low co-occurrence probability as change. They do not model the generation of the scattergram or deal with the severe outlier case (i.e. when foreground dominates and biases the scattergram). Other work (e.g. [7], [5], [23]) use a linear model, only handle global changes, and do not handle the severe-outlier case.

2. Preliminaries

We use the Lambertian model for reflectance where the intensity reported by the camera at pixel location (x, y) is given by the equation (see [11] for a block diagram of the major steps in the imaging process):

$$I(x, y) = f(R(x, y)) \quad (1)$$

$$R(x, y) = \int E(x, y, \lambda) S(x, y, \lambda) Q(\lambda) d\lambda \quad (2)$$

where $f(\cdot)$ is the camera response function, an intrinsic property of the camera that maps a given irradiance $R(x, y)$ into an intensity. $E(x, y, \lambda)$ is the irradiance into the scene point corresponding to the pixel, $S(x, y, \lambda)$ is the reflectance function at the scene point, $Q(\lambda)$ is the camera sensor sensitivity, and λ is the wavelength. For narrow-band (or spectrally sharpened [9]) cameras, and a specific color channel k , one can simplify the above as:

$$I(x, y) = f(E_k(x, y) S_k(x, y) Q_k) \quad (3)$$

As was shown in [22], for a small neighborhood of pixels, smoothly varying illumination, and a monotonic camera response function, the relative ordering of pixel intensities mirrors the relative ordering of their reflectances, which is illumination independent. Devising a practical illumination invariant based on this property has been the subject of several papers (e.g. [3], [24], [22], [17]) including recent ones [15], [20] that have used generative probabilistic models to characterize and use this property effectively. Regardless of the effectiveness and efficiency of these methods, the two fundamental limitations described earlier make them useful only for a limited class of application. Figure 1 shows an example from the publicly available iLIDS dataset [2]. The top left image shows the background and the top right image shows an input image after having been transformed by a monotonic transformation in intensity (simulating a strong change in global illumination). The bottom left image is the change mask obtained using order consistency (we used the method in [20]). Note the presence of holes in the foreground for texture-less patches and the overall ‘blocky’ character of the mask. Such a segmentation, while robust to illumination change and perhaps sufficient for some applications, will pose problems for many higher level tasks such as people detection, localization, tracking, action/activity recognition etc. (e.g. [6], [25], [19]). The method we describe below produces a much better change mask (shown on bottom right) while remaining robust to illumination change.

3. Global Illumination Compensation

Consider pixel intensities at two different illuminations at a given location (subscripts denote location, superscripts denote illumination) and under the absence of foreground objects. We can express the intensities as follows:

$$\begin{aligned} I_i^{(b)} &= f\left(\left(E^{(b)} + E_i^{(b)}\right) S_i\right) \\ I_i^{(c)} &= f\left(\left(E^{(c)} + E_i^{(c)}\right) S_i\right) \\ &= f\left(\frac{E^{(c)} + E_i^{(c)}}{E^{(b)} + E_i^{(b)}} f^{-1}\left(I_i^{(b)}\right)\right) \\ &= f\left(k_i f^{-1}\left(I_i^{(b)}\right)\right) \end{aligned} \quad (4)$$

The superscripts b and c denote a reference background image and a current input image. We have used the Phong reflection model with $E^{(b)}$ and $E^{(c)}$ denoting the ambient illumination levels, and $E_i^{(b)}$ and $E_i^{(c)}$ denoting the effective irradiances into the scene point (the dot product between the illumination direction and surface normal being absorbed into these). Note also that $f^{-1}(\cdot)$ exists because f is monotonic. Let us consider the case of global illumination change. Here, k_i becomes constant ($= k$, the illumination ratio between the levels) and equation 4 can be simplified as:

$$\begin{aligned} I_i^{(c)} &= f\left(kf^{-1}\left(I_i^{(b)}\right)\right) \\ &= \tau\left(I_i^{(b)}\right) \end{aligned} \quad (5)$$

where $\tau(x) = f(kf^{-1}(x))$ is a well defined monotonic function we call *illumination transfer function* (ITF for short). Note that a form of this intensity-to-intensity mapping also arises in camera response function estimation [12], the quantities of interest there being the exposure-ratio k and $f(\cdot)$, given that the two images have identical scene-content. In contrast, our problem is one of detecting changes in the two images having possibly different scene-content in a manner invariant to illumination and $f(\cdot)$.

Note that the intensity order of a group of textured pixels will be the same across the two illumination levels, effectively removing dependence on τ . Now allowing foreground objects, blocks with texture in either of the two images, and that are occluded by foreground in the current image, will undergo order violation and be detected as changed. Every block can be classified into one of four types: (type 1: texture-less and unchanged, type 2: texture-less and changed, type 3: textured and unchanged, type 4: textured and changed). Among these sets of blocks, only blocks of type 3 (call this set B_3) can be checked for and be consistent in pixel order. The other sets consist of blocks that have either undergone order violation (B_4), and so are clearly part of the foreground, or blocks whose status is unknown from the perspective of order violation (B_1 and B_2). We note that B_3 contains important information about the illumination change that has occurred. Accordingly, we pool all blocks in set B_3 and estimate $\tau(\cdot)$ from equation 5 given that $I^{(b)}$ and $I^{(c)}$ are both known. Note that this is a critical *outlier rejection step*. τ , which is effectively an illumination compensation between the two images, can then be applied to pixels in *all* blocks in $I^{(c)}$ to equalize its illumination to that of $I^{(b)}$. The bottom right image in figure 1 shows the resulting foreground mask, i.e. a thresholded distance measure between the background image (top left) and a given input image (top right) although the illumination levels in both of them are very different.

4. Approximations for ITFs

For practical purposes, τ can be represented as a 256 dimensional vector (e.g. 8 bit color channel). B_3 may not contain pixels of all possible intensities (especially when foreground objects occupy a significant proportion of the scene) leaving gaps in the estimate of τ . However, a parametric form $\tau(\cdot; \theta)$ can be assumed where the parameter vector θ is of significantly smaller dimension than the number intensity levels. This makes it possible to determine θ robustly and avoid any gaps in the mapping even with a relatively small number of pixels (e.g. 5 – 10, see below).

Inspired by Grossberg and Nayar’s work [11] where a database of real camera responses (DoRF) was compiled and Principal Component Analysis (PCA) was used to create a low dimensional empirical model of a camera response function, we proceed in a similar fashion to construct a low dimensional model for all possible ITFs. Accordingly, we simulated different possible illumination ratios k (varying it from 10^{-2} to 100 - this range can be selected by an application dependent prior that tells us the range of illumination variation that can be expected in the scene) for all the camera response curves $f(\cdot)$ in the DoRF. Next, we empirically constructed sample ITFs: $\tau(\cdot) = f(kf^{-1}(\cdot))$. We used PCA on the sample space of ITFs and found that the first 5 eigenvectors carry 99.8% of the total energy thus forming a reasonably accurate basis. An ITF can now be approximated as:

$$I^{(c)} = \tau(I^{(b)}) = \tau_0(I^{(b)}) + \sum_{q=1}^T c_q \tau_q(I^{(b)}) \quad (6)$$

where T is the number of basis eigenvectors, $\theta = [c_1 \dots c_T]$ is the parameter vector, and $\tau_0(\cdot)$ is the mean ITF (this is typically just the identity function $\tau_0(x) = x$, if the k values in the simulations were sampled symmetrically around 1, the case where no illumination change is observed). Given N samples of the pair $(I^{(b)}, I^{(c)})$ from blocks in set B_3 , and using equation 6, one obtains an overdetermined set of linear equations in the coefficients $[c_1 \dots c_T]$ which can be solved in a least-squares sense to obtain an estimate of the ITF. $T = 5$ gave us a pretty good fit for any ITF. Figure 2 shows the scattergram between $I^{(b)}$ and $I^{(c)}$ along with an overlay of the estimated ITF for the example shown in figure 1. It can be seen that the ITF estimate approximates well the relationship between $I^{(b)}$ and $I^{(c)}$. It should be noted that the solution for the ITF as obtained need not be monotonic. We enforced the monotonicity constraint, setting up a linearly constrained quadratic programming problem on the coefficients. As this is straightforward, we omit details of this process.

In the special case of an exponential camera response function (as is typically the case for many cameras), it can be shown that the ITF becomes just a straight line passing



Figure 1. Global illumination change example (see text for details)

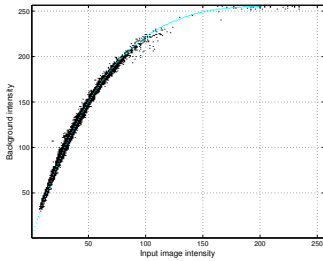


Figure 2. τ estimate for example in figure 1

through the origin (assuming dark current offset has been removed from both images). Hence the number of parameters for an ITF reduces to 1 and the problem becomes greatly simplified. The ITF is given by $I^{(c)} = \tau(I^{(b)}) = k' I^{(b)}$. Whether or not this particular form of the ITF is appropriate for a given camera (without knowing the camera response function upfront) can be determined by propagating the variance from the image domain to the slope of the estimated line, and comparing it to that predicted by the image noise distribution (see equation 8).

5. Local Illumination Change

Consider now the case of scene lighting changes which in effect, cause local illumination variation across the image. While we consider light sources that change in number, direction, and intensity, we assume that they are all distant and are point sources of light (allowing us to ignore penumbras in shadow regions). We also assume that the scene can be approximated by a finite number of planes. Under these assumptions, the factor $E_i^{(c)}/E_i^{(b)}$ assumes a finite set of values (each plane normal direction will typically assume exactly one value but when there are cast shadows there will be one value for each combination of light sources that have changed their impingement on the plane). Call this set of illumination factors $K = \{k_1, k_2, \dots, k_m\}$. The counterpart

to equation 5 in this case becomes:

$$I^{(c,j)} = \tau_j \left(I^{(b)} \right) \quad (7)$$

where $\tau_j(\cdot) = f(k_j f^{-1}(\cdot))$. Note that the above equation remains independent of pixel and block locations as in the global illumination case. The price for this independence is the need to consider multiple possible ITFs ($\tau_j, j = 1..m$). One can, in principle, focus analysis on a small neighborhood (e.g. a single image block), possibly allowing for a single ITF with local support. However, due to the possibility of a block straddling a shadow boundary or a depth discontinuity, a single ITF is inadequate for modeling illumination change within a block. Also, two points could be far from each other in the image but could be on parallel planes, making their local ITFs the same. We should exploit this property as well. Due to these considerations, we continue our analysis with the location independent formulation of equation 7. We could calculate the PCA based ITF for each block in set B_3 as we did for the global case, and cluster the coefficients to obtain a set of m ITFs. We found this approach computationally prohibitive for a real-time application. To simplify matters for this more challenging local illumination case we could either (a) assume that the camera response function is exponential (but unknown), which allows us to use the linear form of the ITF (the advantage in this case is that, as mentioned earlier, the assumption can be verified) or (b) estimate the camera response function offline using any of an array of methods (recent ones being [16], [13], [21]) first, and then work in the radiance domain. Analysis similar to the one presented below (assuming approach (a)) also applies in the irradiance domain.

Each pair of pixels from set B_3 will provide one estimate of the slope k of the ITF¹. We assume that the sensor noise distributions for $I^{(b)}$ and $I^{(c)}$ are isotropic, independent, and zero-mean Gaussians with the same variance σ_0^2 . It is more convenient to work with slope-angle θ . One can show that the variance in θ is:

$$\sigma_\theta^2 \approx \left(\frac{1}{1+k^2} \right) \left(\frac{\sigma_0}{I^{(b)}} \right)^2 \quad (8)$$

The observed distribution over slope angles can be approximated by a mixture of Gaussians whose components can be extracted (e.g. using Expectation Maximization). However, we followed a simpler recursive approach that worked well in practice. Specifically, we extracted the components of mixture recursively, by extracting the global mode, removing all blocks in B_3 that are compensated by the mode, and continuing this process until the total number of elements in the distribution fell below a threshold. For every pixel in

¹For brevity, we re-use the variable k to denote the ratio of intensities although we have used it before for the ratio of irradiances

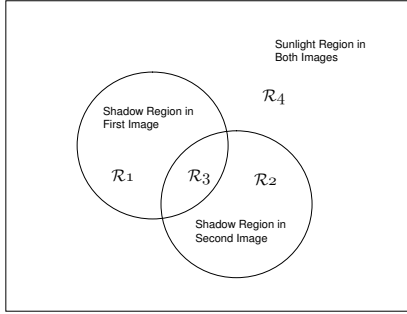


Figure 3. Possible Illumination Changes in an Outdoor Scenario

the current image, we estimate its closest distance to the set of illumination-compensated images that correspond to the extracted modes. This becomes our estimated distance image after illumination compensation. Each pixel can also be labeled based on its closest mode, which allows us to group them together based on illumination change.

5.1. Outdoor Scenario

In an outdoors scenario under daylight conditions, there is only one primary light source, the sun. The amount of light impinging on a plane is the sum of skylight and possible direct sunlight². Consider two images of a plane taken at different times of day. Figure 3 shows the possible classes of illumination conditions under which every pixel falls: Pixels in class \mathcal{R}_1 are in shadow in the first image and in sunlight in the second image. Pixels in \mathcal{R}_2 are in sunlight in the first image and in shadow in the second. Pixels in class \mathcal{R}_3 are in shadow in both images. Finally pixels in class \mathcal{R}_4 are in sunlight in both images. Let τ_i be the illumination transfer function that transforms pixel intensities in class \mathcal{R}_i from the first to the second image. Thus there are up to four ITFs possible for a plane. For the *ground plane* at low to moderate latitudes, one could approximate skylight as a constant factor of sunlight (skylight irradiance being highest when the sun is directly overhead and lowest when the sun is at the horizon). The irradiance at a point on the ground plane in sunlight at time t can be approximated as $L(t) = L^{sky}(t) + L^{sun}(t) = L^{sun}(t)(c + 1)$ where c is a constant. Similarly the irradiance at a point in shadow can be approximated as $L(t) = L^{sky}(t) = cL^{sun}(t)$. It is easy to show in this case that the all-sunlight region \mathcal{R}_4 and the all-shadow region \mathcal{R}_3 have the same ITF $\tau_4 (= \tau_3)$. Thus the total number of ITFs for the ground plane is three. This makes intuitive sense because shadow-to-shadow and sunlight-to-sunlight transformations can be expected to be the same. We also found this to be true in practice (see fig-

²Note that we use the term *skylight* and *sunlight* for the total amount of skylight/sunlight energy impinging at the point

ure 4). It can also be shown that:

$$\tau_3\tau_1^{-1} = \tau_2\tau_3^{-1} \quad (9)$$

These two results put together show that there are only up to two independent illumination transfer functions for the ground plane in an outdoor environment. Note that when the camera response is an exponential function, we know that the illumination transfer functions are straight lines passing through the origin. In this case, $\tau_i(x) = k_i x$, and it is easy to see that for this case, equation 9 reduces to $k_3^2 = k_1 k_2$. Figure 4 shows an example where the two images being compared have different shadow positions (images in the upper row). Note that the ITF slope histogram (bottom left) shows three peaks as predicted. Also, notice that except for the penumbra and compression induced artifacts in the far distance, the system is able to detect the person at the bottom right inspite of severe illumination changes due to shadowing.

6. Results

We present validation and exemplar results. We present validation of the proposed global and local illumination compensation approach on synthetically modified real images. We follow this up with real image examples showing change detection under strongly different illumination conditions.

For validation of the global case, we took 50 images from the iLIDS abandoned object detection dataset [2], applied random global illumination to each of them and compared the proposed method (i.e. based on PCA with 5 dimensions - see equation 6) against three other methods, namely the order consistency (OC) method of [20], the illumination invariant (II) image calculation method of [8], and the comprehensive color normalization (CN) approach of [10]. The original images from the dataset are from an indoors scenario with virtually no illumination change. Manually generating ground truth for this moderately crowded scenario is a laborious task. We chose instead to generate approximate ground truth, by running the L_2 distance based change detector on the original images. We made the reasonable assumption of normally distributed image noise with $\sigma = 4$ gray levels of standard deviation. This makes the normalized L_2 distance $(I^{(b)} - I^{(c)})/\sigma^2$ central χ^2 distributed with one degree of freedom. Choosing a false alarm rate of 0.01 we created binary images flagging outlier pixels. While these binary images are only approximate ground-truth, they are very close to it and sufficient for relative comparison of methods. The first two images in figure 5 shows a sample. The left image in figure 6 shows the ROC curves for the three methods. As we can see, the proposed method compensates for the global change almost exactly. OC is not able to capture the complete foreground due to

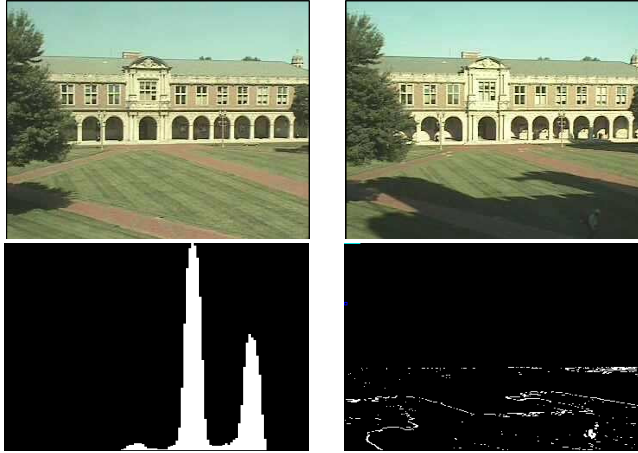


Figure 4. Outdoor scenario (see text for further details): Top row: Background and Current image. Bottom Row: Histogram of ITF slopes and change mask. Note that the region of interest is the ground plane.



Figure 5. Validation Examples: From left (Global background, Sample input, Local background, Sample input)

missing internal homogenous regions. As the scene has little color, the II and CN methods are not able to capture the entire foreground. For the local illumination case, we chose a pair of images from the AMOS dataset [1] taken at different times of the day, one with an empty scene and one with small foreground. We applied synthetic local illumination at random regions in the image and created a set of 50 images. The local illumination was simulated by multiplying each channel by a different scale factor to simulate light sources of different intensity and wavelengths. The images in figure 5 shows a sample. We set the region of interest to the ground plane and manually labeled the foreground. The right image in figure 6 shows the ROC curves for the four methods. We observe that OC works reasonably well for this scenario because the object is small (leading to smaller miss detections on texture-less regions). CN on the other hand does not produce good results. We specifically excluded II from the comparison because the simulated changes are non-Planckian.

Finally, we present exemplar results on four challenging scenarios from AMOS dataset. The input images are of poor quality with significant compression induced artifacts. Figure 7 shows several examples from this dataset. The first column shows an outdoor scene where the shadows have changed significantly between the background and input image. It can be seen that except for penumbras, as

expected, the proposed method (bottom-most image in column) is able to detect the two persons very well and more completely than the other methods. The remaining columns show different scenes, where again, except for penumbra effects, the change masks from the proposed method delineate the silhouettes better than the other methods. In practice, the penumbra problem can be handled by reasoning over the foreground regions (e.g. filtering out long connected strips of foreground pixels that all mapped to the same ITF). One could apply such filters over the low level output and produce even better results. However, please note that our goal here is to present our approach for illumination compensation rather than a complete "change detection" system.

7. Conclusions

We made the following contributions: First, we modeled illumination change between a given pair of registered images and showed the effect of relative global and local illumination changes on image intensities. The illumination change is represented as an illumination transfer function (ITF). Second, we showed how we could use the database of camera responses [11] to create a low dimensional approximation of the ITF in the general case. We observed that for specific cameras the ITF reduces to a linear function. We showed that the number of independent ITFs for planes in outdoor scenes under daylight conditions is constrained. Finally we showed how we could exploit rank order consis-

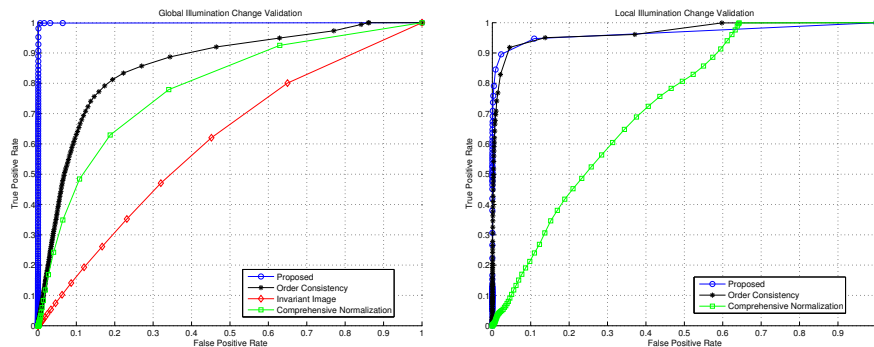


Figure 6. Validation Results: Global change (left), Local change (right)

tency to remove outliers, estimate the ITF and compensate for local and global illumination change. We validated the method and demonstrated superior results on exemplar images. For future work, on the theoretical side, we are working towards extending the analysis to nearer light sources, curved surfaces, and penumbras by exploiting local structures in the scattergram. Handling camouflage (where foreground and background pixels intensities are close) is another important area for future work on the practical side. Camouflage causes missed detections for any change detection algorithm. In the case of a large number of local illumination changes, this effect could be severe if all extracted ITFs are applied *globally*, the way our approach is implemented currently. We are working towards reducing this effect by a spatial grouping of ITFs and performing reasoning in local neighborhoods.

References

- [1] Archive of many outdoor scenes. http://www.cse.wustl.edu/~jacobsn/projects/webcam_dataset/.
- [2] Imagery library for intelligent detection systems. <http://scienceandresearch.homeoffice.gov.uk/hosdb>.
- [3] D. N. Bhat and S. K. Nayar. Ordinal measures for image correspondence. *IEEE PAMI*, 20(4):415–423, 1998.
- [4] P. A. Bromiley, N. A. Thacker, and P. Courtney. Non-parametric image subtraction using grey level scattergrams. *Image and Vision Computing*, 20(9-10):609 – 617, 2002.
- [5] J.-H. Cho and S.-D. Kim. Object detection using multi-resolution mosaic in image sequences. *Signal Processing: Image Communication*, 20(3):233 – 253, 2005.
- [6] L. Dong, V. Parameswaran, V. Ramesh, and I. Zoghlami. Fast crowd segmentation using shape matching. *Proc. ICCV*, 2007.
- [7] M. E. Farmer. "a chaos theoretic analysis of motion and illumination in video sequences". *Journal of Multimedia*, 2(2), 2007.
- [8] G. Finlayson, S. Hordley, C. Lu, and M. Drew. On the removal of shadows from images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(1):59–68, Jan. 2006.
- [9] G. D. Finlayson, M. S. Drew, and B. V. Funt. Spectral sharpening: sensor transformations for improved color constancy. *Journal of the Optical Society of America*, 11(5), 1994.
- [10] G. D. Finlayson, B. Schiele, and J. L. Crowley. Comprehensive color image normalization. *Proc. European Conference on Computer Vision*, 1998.
- [11] M. Grossberg and S. Nayar. Modeling the space of camera response functions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(10):1272–1282, Oct. 2004.
- [12] M. D. Grossberg and S. K. Nayar. What can be known about the radiometric response function from images ? In *Proc. European Conference on Computer Vision*, 2002.
- [13] S. J. Kim and M. Pollefeys. Radiometric alignment of image sequences. *Proc. IEEE CVPR*, 2004.
- [14] Y. Kita. A study of change detection from satellite images using joint intensity histogram. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4, Dec. 2008.
- [15] A. Lanza and L. D. Stefano. Detecting changes in grey level sequences by ml isotonic regression. *IEEE Conf. on Advanced Video and Signal Based Surveillance*, 0:4, 2006.
- [16] S. Lin and L. Zhang. Determining the radiometric response function from a single grayscale image. *Proc. IEEE CVPR*, 2005.
- [17] A. Mittal and V. Ramesh. An intensity-augmented ordinal measure for visual correspondence. In *Proc. IEEE CVPR*, 2006.
- [18] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam. Image change detection algorithms: A systematic survey. *IEEE Trans. on Image Processing*, 14(3):294–307, 2005.
- [19] J. Rittscher, P. Tu, and N. Krahnstoever. Simultaneous estimation of segmentation and shape. *Proc. IEEE CVPR*, 2:486–493, 2005.
- [20] M. Singh, V. Parameswaran, and V. Ramesh. Order consistent change detection via fast statistical significance testing. *Proc. IEEE CVPR*, 2008.
- [21] J. Takamatsu, Y. Matsushita, and K. Ikeuchi. Estimating camera response functions using probabilistic intensity similarity. *Proc. IEEE CVPR*, 2008.
- [22] B. Xie, V. Ramesh, and T. E. Boult. Sudden illumination change detection using order consistency. *Image and Vision Computing*, 22(2):117–125, 2004.

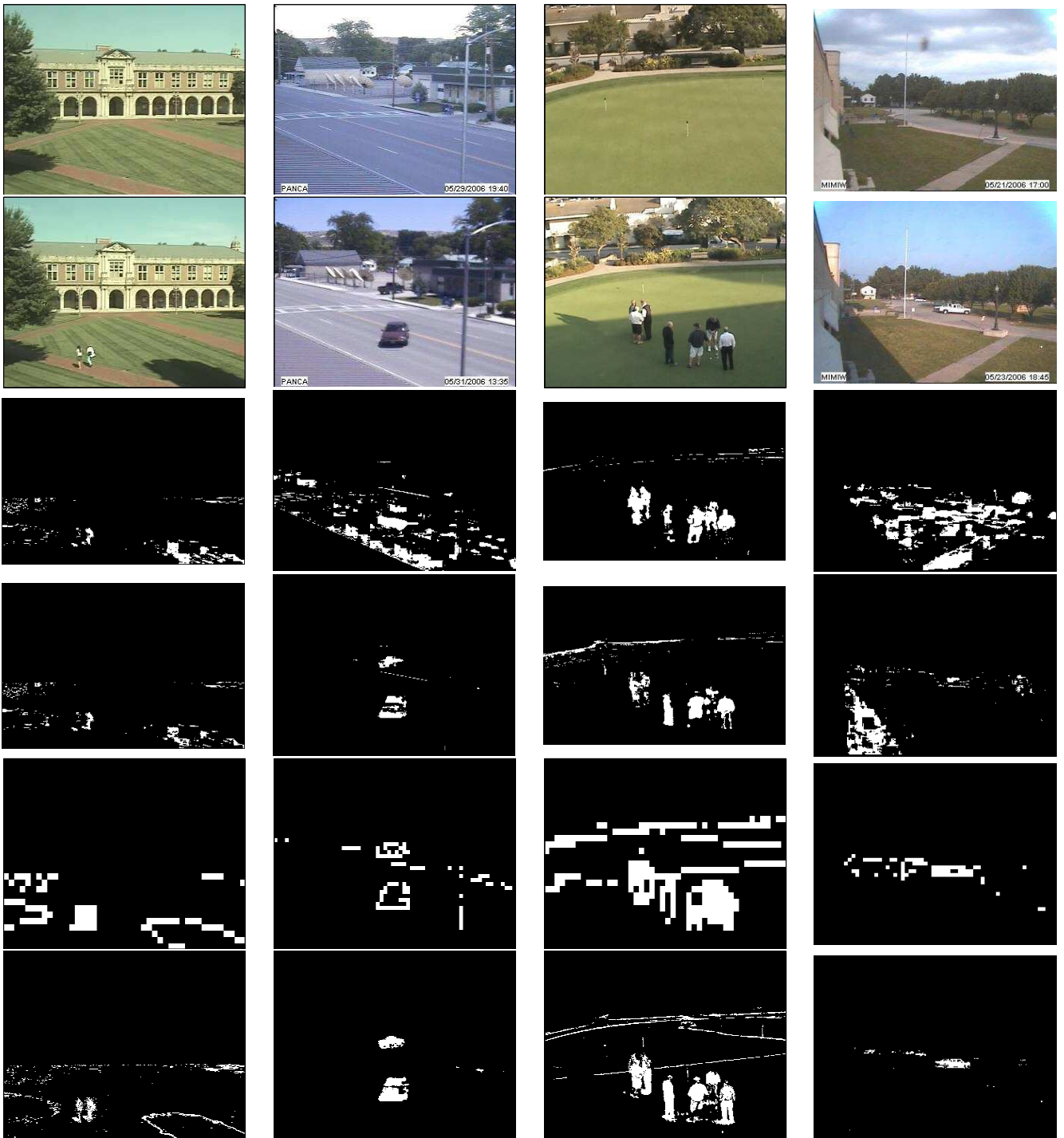


Figure 7. Top to bottom(1) Background, (2) input, (3) II mask, (4) CN mask, (5) OC mask, (6) Proposed method.

[23] H. Yalcin, R. Collins, and M. Hebert. Background estimation under rapid gain change in thermal imagery. *Computer Vision and Image Understanding*, 106(2-3):148 – 161, 2007. Special issue on Advances in Vision Algorithms and Systems beyond the Visible Spectrum.

[24] R. Zabih and J. Woodfill. A non-parametric approach to

visual correspondence. In *Proc. European Conference on Computer Vision*, pages 82–89, 1994.

[25] T. Zhao and R. Nevatia. Bayesian human segmentation in crowded situations. *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2:18–20, 2003.