PIRM: FAST BACKGROUND SUBTRACTION UNDER SUDDEN, LOCAL ILLUMINATION CHANGES VIA PROBABILISTIC ILLUMINATION RANGE MODELLING

Parthipan Siva†, Mohammad Javad Shafiee‡, Francis Li‡, and Alexander Wong‡

† Aimetis Corp., Waterloo, Ontario
‡ University of Waterloo, Waterloo, Canada

ABSTRACT

We present an illumination-compensation method to enable fast and reliable background subtraction under sudden, local illumination changes in wide area surveillance videos. We use Probabilistic Illumination Range Modeling (PIRM) to model the conditional probability distribution of current frame intensity given background intensity. With this model, we can identify a continuous range of current frame intensities that map to the same background intensity, and scale all pixels within that range in the current frame appropriately to enable illumination-compensated background subtraction. Experimental results using a standard academic dataset as well as very challenging industry videos show that PIRM can achieve improvements in compensating for sudden, local illumination changes.

Index Terms— background subtraction, illumination-compensation, probabilistic modelling

1. INTRODUCTION

Background subtraction is a crucial component in a wide range of video surveillance applications, e.g., traffic monitoring, detecting interactions between people, and identifying safety hazards in industry plants and school yards. A major challenge to accurate background subtraction is sudden illumination changes in the scene (e.g., indoor light flickering, shadows, overhanging clouds passing by, and strong sunlight). In these situations, the background is no longer stable and could be mistakenly classified as the foreground, giving false positives. Earlier statistical background subtraction methods [1234], while robust to noise and dynamic backgrounds, depend on a learning rate to update the background model to account for gradual illumination changes, and thus are prone to large errors (false alarms) when subject to sudden illumination changes.

A number of methods have been proposed to improve background subtraction under sudden illumination changes. In RaBS [5] and ViBe [6], sudden illumination changes are handled by re-initializing the background frame if the illumination difference is too large. A challenge with this strategy is the need to select a threshold at which to abandon the old background model, which can be extremely difficult. Furthermore, during background re-initialization objects in the scene cannot be tracked.

Several methods [7891011] use the ratio between current frame intensities and background frame intensities to map illumination changes from background model to current frame. The work by [8] assumed a uniform illumination change in the scene, and used the median ratio over all pixels as the illumination scaling factor for the background model. Such a linear assumption does not always hold true, especially when there are local illumination changes (e.g., part of the frame is under direct sunlight). To address this issue, other methods have employed nonlinear intensity mapping [910111213]. Paruchuri et al. [9] segments the current frame into regions and finds a nonlinear mapping for each segment. Bales et al. [11] proposed the estimation of ratios on large and permanent background objects identified in the scene. Vosters et al. [1213] extended upon the work in [7] by incorporating an eigenbackground [14] model to better handle local illumination changes.

The two main limitations to existing methods are that many are limited to a one-to-one mapping between current frame intensities and background intensities (which may not hold true for a large number of situations) and that many are computationally complex, making them unwieldy for industry video surveillance applications with real-time requirements.

Motivated to tackle these two limitations, we propose a new Probabilistic Illumination Range Modeling (PIRM) approach to fast and reliable background subtraction under sudden, local illumination changes in wide area surveillance. By modelling the conditional probability distribution of current frame intensity given background intensity, PIRM allows us to identify a continuous range of current frame intensities that map to the same background intensity. This enables us to scale all pixels within that range in the current frame appropriately to enable illumination-compensated background subtraction.

The authors would like to thank Aimetis Corp., NSERC, Ontario Ministry of Research and Development, and Canada Research Chairs program.
Fig. 1: Scatter plots relating current frame intensities to background frame intensities.

Fig. 2: Example of modelling illumination change using median [8], mean (Mmap), and PIRM ($\bar{\mu} \pm T \hat{\sigma}$). Best viewed in colour.

2. METHODOLOGY

The PIRM approach is described in detail below. While PIRM can be combined with any background modelling method, we will combine PIRM with the Gaussian Mixture Model (GMM) presented in [15] for background modelling because GMMs are computationally efficient and have been successfully used in many industry applications.

2.1. Gaussian Mixture Model

At time $t$, in the GMM used here for background modelling, we have a set of pixels $\mathcal{X}_T = \{\bar{x}_1, \ldots, \bar{x}_{T-1}\}$, where $T$ is an adaptation period and $\bar{x}$ is the colour representation of a pixel. The estimated density using a GMM with $M$ components can be expressed by

$$p(\bar{x}_t | \mathcal{X}_T) = \sum_{m=1}^{M} \tilde{p}_m N(\bar{x}_t; \tilde{\mu}_m, \tilde{\sigma}^2_m)$$  \hspace{1cm} (1)

where $\tilde{\mu}_1, \ldots, \tilde{\mu}_M$ are the estimates of the means and $\tilde{\sigma}^2_1, \ldots, \tilde{\sigma}^2_M$ are the estimates of the variances, $\tilde{\pi}_m$ is the weight of each component, and $I$ is the identity matrix. If the $M$ components are sorted in descending order of component weight $\tilde{\pi}_m$, then the background is represented by the first $n$ components.

Given $\bar{x}$, if $\bar{x}$ is a close match to one of the $n$ background modes, then it belongs to the background, otherwise it belongs to the foreground. The pixel is close to a component if:

$$D^2_m(\bar{x}) = \frac{(\bar{x} - \tilde{\mu}_m)'(\bar{x} - \tilde{\mu}_m)}{\tilde{\sigma}^2_m} < \tau$$  \hspace{1cm} (2)

From the GMM, an estimate of the current background image $B$ can be obtained based on the mean of the last active background GMM component. Our goal is to use $B$ and the current frame $I$ to estimate the lighting change that has occurred in the scene.

2.2. Probabilistic Illumination Range Modelling

As with [7, 8, 9, 10], we assume the current pixel colour $\bar{x}$ has been scaled from the modelled background value ($\bar{\mu}$) by some factor $1/f(\cdot)$ due to a lighting change. In order to test if $\bar{x}$ belongs to the background using (2), the scaling to the current pixel colour must first be compensated for (i.e., $f(\cdot)\bar{x}$) before testing if it belongs to the background model:

$$D^2_m(\bar{x}) = \frac{(f(\cdot)\bar{x} - \tilde{\mu}_m)'(f(\cdot)\bar{x} - \tilde{\mu}_m)}{\tilde{\sigma}^2_m} < \tau$$  \hspace{1cm} (3)

In PIRM, we take a probabilistic modelling approach to estimating $f(\cdot)$. Consider Fig. 1 where scatter plots relating background pixel intensities to current frame pixel intensities for different situations are shown.

No illumination change: (Fig. 1) Without any foreground objects, the relationship between background intensities and the current frame intensities is linear. In the presence of foreground objects, the relationship is still fairly linear with foreground object pixels being outliers.

Sudden local illumination change: (Fig. 1) The relation-ship is no longer linear. Furthermore, there is no one-to-one mapping $b$ for different situations are shown.

In PIRM, we assume the current pixel colour $\bar{x}$ is a probabilistic function of background intensity $b$ and current frame intensity values map to the same background intensity value. As a result, it is impossible to find a one-to-one mapping function $b = g(i)$, where $f(b) = b/g(i)$, which maps the current frame pixel intensity to the background pixel intensity.

Unlike existing approaches [7, 8, 9, 10] that attempt to find a one-to-one mapping function $b = g(i)$, we model the range of current frame intensity values $i \in T$ that maps to the same, unique background intensity $b$ using conditional probability distribution $P(i|b)$. Based on $P(i|b)$, one can define the mapping $f(\cdot)$ as a probabilistic function of background intensity $b$ and current frame intensity $i$:

$$f(b, i) = \begin{cases} b/i & \text{with prob. } P(i|b) \\ 1 & \text{with prob. } 1 - P(i|b) \end{cases}$$  \hspace{1cm} (4)

In practice we model $P(i|b)$ parametrically as a Gaussian distribution $P(i|b) \sim N(\mu(b), \sigma^2(i|b))$ (Fig. 2)

$$\mu(b) = \frac{1}{|T|} \sum_{i \in T} i, \quad \sigma^2(b) = \frac{1}{|T|} \sum_{i \in T} (i - \mu(b))^2$$  \hspace{1cm} (5)
where $\mathcal{I}$ is the set of all pixel intensities in the current frame $I$ that maps to the background pixel intensity $b$, $|\mathcal{I}|$ is the number of pixels in set $\mathcal{I}$, and $b = 0 \ldots 255$ is all the possible background intensities. A Gaussian distribution for $P(i|b)$ is computationally efficient and allows for real-time applications. To reduce the influence of the foreground objects in modelling the relationship between current frame intensities and background intensities, we smooth $\mu(b)$ and $\sigma^2(b)$ across the background intensity range $[b - N, b + N]$ (Fig. 2).

$$\mu(b) = \frac{1}{2N} \sum_{k=-N}^{N} \mu(b + k) \quad \sigma^2(b) = \frac{1}{2N} \sum_{k=-N}^{N} \sigma^2(b + k) \quad (6)$$

Our model assumes that the video is a wide area surveillance video where the objects are relatively small in pixel size compared to the frame size. If the objects are relatively large compared to the frame size, then the means $\mu(b)$ will be affected by the foreground pixel intensities.

Given the constructed model $P(i|b)$ as characterized by $\bar{\mu}(b)$ and $\bar{\sigma}^2(b)$ [6], we simplify the mapping $f(\cdot)$ (Eq. 4) into a deterministic form as:

$$f(b, i) = \begin{cases} b/i & \text{if } \bar{\mu}(b) - T_1 \bar{\sigma}(b) \leq i \leq \bar{\mu}(b) + T_2 \bar{\sigma}(b) \\ 1 & \text{else} \end{cases} \quad (7)$$

where $T_1 = 2$ and $T_2 = 3$ if $\bar{\mu}(b) > b$ and $T_1 = 3$ and $T_2 = 2$ if $\bar{\mu}(b) \leq b$, as it was found to provide strong performance.

### 3. RESULTS

Two datasets are used to test PIRM. The first, Alley, is a custom data obtained from Aimetis Corp. [16] with groundtruth. It consists of a 4-min video sequence with very strong sudden illumination changes due to direct sunlight. The second, PETS, is a PETS 2001 Dataset [17] with groundtruth obtained from Kyushu University [18].

PIRM is compared to: [8] – the linear median-based mapping method proposed in [8], Mmap – a linear mean-based mapping method ($f = mean(b/i)$), and NLmap – a nonlinear mapping method where $f = b/\bar{\mu}(b)$.

Fig. 3 shows some of the illumination-compensation results on Alley and Table 1 reports the recall, precision, and F1-Measure as implemented by ChangeDetection.net [19]. Based on F1-Measure, we can see that PIRM provides significant improvements over linear methods like [8] and Mmap. Furthermore, PIRM performs better than using just a non-linear mapping like NLmap. These results illustrate that modelling illumination change as a probabilistic distribution, as in PIRM, may be more effective than modelling it as a one-to-one mapping, as in existing works [2][8][9][10]. PIRM has significantly higher precision than the other methods, which is very important as it means that there are fewer false objects being detected.

The F1-measure is high for PIRM, but the recall is lower than the other methods. While this is a concern, what is of more practical importance is whether all foreground objects in every frame are detected. Even a partial detection on each frame can allow us to track the object. For each foreground object in each frame we also look at the percentage of groundtruth foreground pixels detected as foreground after illumination-compensation. The mean percentage of correctly detected foreground pixels per object and the number of objects with at least 10% of pixels detected as foreground are shown in Table 1. All objects are still detected (by at least 10%) by PIRM on the Alley dataset. Furthermore, on the PETS dataset only 2 objects are missed by PIRM.

Timing evaluation on the 720x576 resolution Alley video, showed that the runtime of incorporating PIRM (33ms) to GMM background subtraction is comparable to other methods such as [8] (32ms). This shows that PIRM can be used to facilitate fast and reliable background subtraction under sudden, local illumination changes.

### Table 1: Performance results: Per-pixel recall (Rec.), precision (Prec.), F1-Measure (F1), mean and standard deviation ($\mu \pm \sigma$) of the percentage of foreground object area (in pixels) correctly detected as foreground, and the number of foreground objects ($\geq 10\%$) with at least 10% of the area correctly detected as foreground.

<table>
<thead>
<tr>
<th></th>
<th>Alley</th>
<th>PETS</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Comp</td>
<td>0.78 ± 0.06</td>
<td>0.11 ± 0.11</td>
</tr>
<tr>
<td>[8]</td>
<td>0.75 ± 0.11</td>
<td>0.20 ± 0.19</td>
</tr>
<tr>
<td>Mmap</td>
<td>0.66 ± 0.09</td>
<td>0.17 ± 0.19</td>
</tr>
<tr>
<td>NLmap</td>
<td>0.71 ± 0.14</td>
<td>0.23 ± 0.38</td>
</tr>
<tr>
<td>PIRM</td>
<td>0.56 ± 0.55</td>
<td>0.38 ± 0.39</td>
</tr>
</tbody>
</table>

**Fig. 3:** Results of the various illumination-compensation methods on the datasets. Alley is a 5 FPS video and PETS is a 25 FPS video.
4. REFERENCES


