

Image-based fusion for video enhancement of night-time surveillance

Yunbo Rao,^a Weiyao Lin,^b and Leiting Chen^a

^aUniversity of Electronic Science and Technology of China, School of computer science and engineering, Chengdu, China, 611731

^bShanghai Jiao Tong University, Institute of Image Communications and Information Processing, Department of Electronic Engineering, Shanghai, China, 200240
E-mail: cloudrao@gmail.com

Abstract. In this paper, a novel image-based fusion video enhancement algorithm is proposed for night-time video surveillance applications by a combination of illumination fusion and based on moving objects fusion. The proposed algorithm fuses video frames from high quality day-time and night-time background with low quality night-time videos. For improving the perception quality of the moving objects, based on moving objects of region fusion method is proposed. Experimental results show the effectiveness of the proposed algorithm. © 2010 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.3520553]

Subject terms: video enhancement; illumination fusion; image-based fusion.

Paper 100440LR received Oct. 6, 2010; revised manuscript received Oct. 25, 2010; accepted for publication Mar. 11, 2010; published online Dec. 29, 2010.

1 Introduction

Video surveillance is often more important in the dark environment since many activities of interest often occur in the dark environment.^{1,2} Video enhancement plays a key part in night-time video surveillance so that the objects or activities of interest can be clearly monitored. The problem of video enhancement for low quality video has become increasing by acute.³ The goal of the enhancement⁴ is to improve the visual appearance of the video, or to provide a “better” representation for further processing, such as analysis, detection, segmentation, and recognition. However, it is still a challenging problem for night-time video applications. To traditional algorithms, when the background is enhanced, the contrast often remains low, or the noise is often greatly enhanced. Since day-time videos are also available in many surveillance applications, there are many attempts to take this advantage to combine the day-time and the night-time scenes to enhance the night-time videos.⁵⁻⁷ However, since there are different moving objects in the day-time video and night-time video, and also the condition in the background may be different, it is not easy to produce good results with fusion.

To further improve the enhancement quality, it is desirable to give larger weight to the moving objects. However, accurate moving foreground extraction is difficult, especially with low contrast and noisy videos. Reference 8 used a multi-color background model per pixel to enhance the foreground mov-

ing objects. However, the method suffers from slow learning at the beginning, especially under busy background. It also cannot distinguish moving shadows from moving objects. In addition, the model does not update with time and therefore often fails under outdoor environments where the scene lighting changes frequently with time. Reference 9 presented a method which improves this adaptive background mixture model.

In this paper, a new method is proposed to fuse video frames from high quality day-time and night-time backgrounds with low quality night-time video frames. With the proposed algorithm, day-time images and night-time images are combined together to provide a much more enhanced background. In order to enhance the moving objects as well, we also propose a moving objects of region fusion method for improving the sharpness of the moving objects.

2 The Proposed Algorithms

The detailed procedures of the proposed method can be described as in Fig. 1. It should be noted that in Fig. 1, the ‘day-time image’ as well as the ‘night-time image’ are the high-quality images for providing a better enhanced background and the night-time video is the actually low-quality input that we need to enhance. In the following, we will describe algorithm.

2.1 Illumination Segmentation

We first decouple an input color image $f(x, y)$ into intensity $I(x, y)$, and color layer $C(x, y)$ where our algorithm is mainly processed on the intensity layer $I(x, y)$. Then an image is separated into the illumination layer $L(x, y)$ and the reflectance layer $R(x, y)$ of the day-time and night-time background image can be obtained by Retinex theory.¹⁰ It is assumed that the available luminance data in the image is the product between illumination and reflectance. The input image $I(x, y)$ is represented by the product of the illumination and the reflectance as follows:

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

The illumination $L(x, y)$ is assumed to be contained in the low frequency components of the image while the reflectance $R(x, y)$ mainly represents the high frequency components of the image. The Gaussian low-pass filtered result of the intensity image is used as the estimation of the illumination. The filtering process is actually a 2D discrete convolution with Gaussian kernel, which can be mathematically expressed as¹¹:

$$L(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n) G(m+x, n+y) \quad (2)$$

where G is the 2D Gaussian function with size $M \times N$. Gaussian kernel G is defined as:

$$G(x, y) = q \cdot \exp\left(\frac{-(x^2 + y^2)}{c^2}\right) \quad (3)$$

q is a normalization factor:

$$\sum_x \sum_y q \cdot \exp\left(\frac{-(x^2 + y^2)}{c^2}\right) = 1 \quad (4)$$

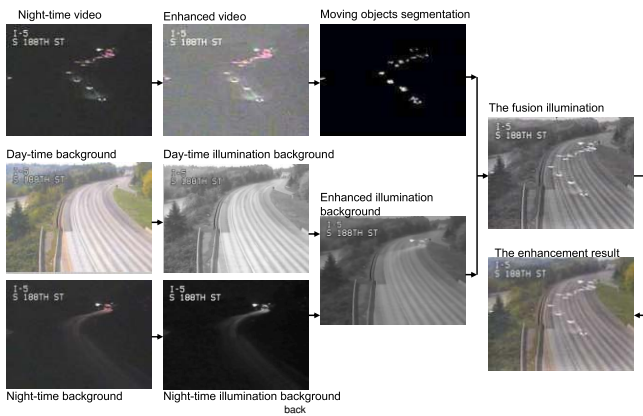


Fig. 1 A block diagram of the proposed algorithm

and c is a scale constant ($c = 2 \sim 5$ is commonly used). Here, c is set to 3.

2.2 Enhanced Background

We adopt the weighted-average image-fusion algorithm to enhance nighttime background using illumination images $L(x, y)$. The proposed fusion equation is as follows.

$$B_L(x, y) = \alpha * N_L(x, y) + (1 - \alpha) * D_L(x, y) \quad (5)$$

where $B_L(x, y)$ is the final background image, $N_L(x, y)$ is the nighttime illumination image, $D_L(x, y)$ is the daytime illumination image. The weight α is in the range $[0, 1]$. In our algorithm, α is empirically determined by the global mean $N_L(x, y)$ and $D_L(x, y)$ of the illumination image based on the image enhancement experiments.

2.3 Enhanced (Night-time) Video

Due to the low contrast, we can not clearly extract moving objects from the dark background. We propose an enhanced-video step to facilitate the extraction moving objects. The tone-mapping approach is used to enhance the video frames and to separate an image into details and large scale features. More specifically, the nonlinear tone mapping function is used to attenuate image details and to adjust the contrast of large scale features,¹² as in Eqn. (6).

$$m(x, \psi) = \frac{\log\left(\frac{x}{xMax}(\psi - 1) + 1\right)}{\log(\psi)} \quad (6)$$

The white level of the input illumination is set by $xMax$ and ψ controls the attenuation profile. This mapping function exhibits a similar characteristic as the traditional Gamma correction.

2.4 Motion Segmentation

After enhanced the night-time videos, motion detection is performed to extract the foreground moving objects,⁹ as in Eqn. (7). Each pixel in the scene is modeled by a mixture of K Gaussian distributions. The probability that a certain pixel has a value of X_N at time N can be written as.

$$p(X_N) = \sum_{j=1}^K w_j \eta(X_N; \theta_j) \quad (7)$$

where w_k is the weight parameter of the K^{th} Gaussian component. $\eta(X_N; \theta_j)$ is the normal distribution of K^{th} component represented by

$$\begin{aligned} \eta(X; \theta_k) &= \eta\left(X; u_k, \sum_K\right) \\ &= \frac{1}{(2\pi)^{D/2} |\sum_K|^{1/2}} e^{-\frac{1}{2}(X-\mu_k)^T \sum_K^{-1} X-\mu_k} \end{aligned} \quad (8)$$

where u_k is the mean and $\sum_K = \sigma_k^2 I$ is the covariance of the K^{th} component.

The K distributions are ordered based on the fitness value w_k/σ_k and the first B distributions are used as a model of the background of the scene where B is estimated as

$$B = \arg \min_b \left(\sum_{j=1}^b w_j > T \right) \quad (9)$$

The threshold T is the minimum fraction of the background model. Under this method, a pixel will be detected as a foreground pixel if it is more than 2.5 standard deviations away from any of the B distributions.

2.5 Final Fusion and Enhancement

After getting the weighting based fusion background image and the resulting motion detection video frames, we will perform the final video enhancement by a combination of illumination and based on moving objects region fusion. The proposed combination of illumination and region fusion mathematic equation is as follows:

$$\begin{aligned} F_L(x, y) &= \beta M(x, y) + \gamma N_L(x, y) \\ &+ (1 - \gamma) B_L(x, y) \end{aligned} \quad (10)$$

where $F_L(x, y)$ is the final illumination image, $M(x, y)$ is motion-detection video frame, $N_L(x, y)$ is the night-time illumination image and $B_L(x, y)$ is enhanced background illumination as in Eqn. (5). β and γ are the weights for these input images, respectively. In our algorithm, β and γ are determined in the following way.

$$\begin{cases} \beta = 1 \text{ and } \gamma = 1 \text{ if } M(x, y) + N_L(x, y) + B_L(x, y) \geq 1 \\ \beta = 0 \text{ and } \gamma = K \text{ if } M(x, y) + N_L(x, y) + B_L(x, y) < 1 \end{cases} \quad (11)$$

When $M(x, y) + N_L(x, y) + B_L(x, y) < 1$, it assumes that there are no moving objects in the current pixel. Then the coefficient $\beta = 0$ and γ can be decided by the illumination



Fig. 2 Enhancing a traffic night video. (a) A low quality night-time video frame. (b) A frame from the final result of the proposed algorithm.

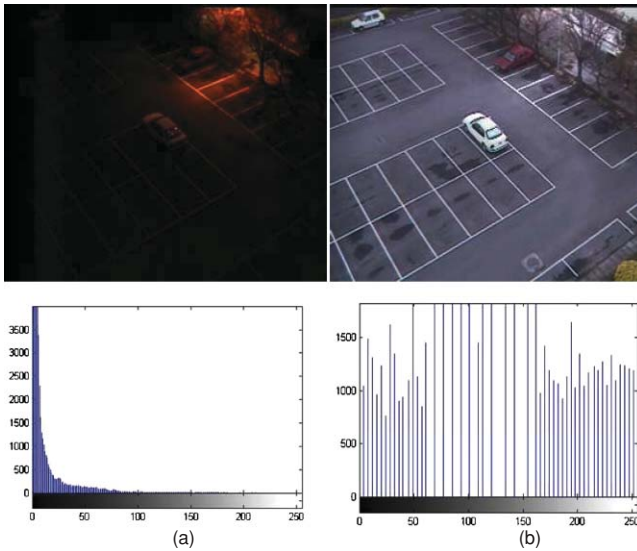


Fig. 3 (a) Original low quality night-time video frame and histogram, (b) the enhanced result of the proposed algorithm and histogram.

images $N_L(x, y)$ and $B_L(x, y)$. In our experiments, we set the values $\gamma = K = 0.4$. On the other hand, if $M(x, y) + N_L(x, y) + B_L(x, y) \geq 1$, it is assumed that there are moving objects in the current pixel. And in this case, both β and γ are set to be 1. Furthermore, in order to avoid $F_L(x, y)$ exceeds the illumination range of $[0, 255]$, a pixel will be set to be 255 if its value exceeds 255.

3 Final Experimental Results and Conclusions

In this paper, we propose a novel algorithm to enhance night-time video. We focus on addressing the following two key issues for video enhancement. (1) illumination-based fusion for enhancement background image. (2) Moving objects of region fusion for improving sharpness of the moving objects. Figures 2 and 3 show the experimental results. The result demonstrates that the proposed algorithm can use the color

resources (i.e. color levels) more efficiently and it is robust and effective.

Acknowledgments

This work is partly supported by National High-Tech Program 863 of China (Grant No. 2007AA010407 and 2009GZ0017), National Research Program of China (Grant No.9140A06060208DZ0207), National Science Foundation of China (61001146), and China Scholarships Council.

References

1. W. Lin, M.-T. Sun, R. Poovendran, and Z. Zhang, "Human activity recognition for video surveillance," ISCAS, pp. 2737–2740, 2008.
2. W. Lin, M.-T. Sun, R. Poovendran, and Z. Zhang, "Group Event Detection with a Varying Number of Group Members for Video Surveillance," *IEEE Transactions on Circuits and Systems for Video Technology*, **20**(8), 1057–1067(2010).
3. X. Dong, Y. Pang, and J. Wen, "Fast efficient algorithm for enhancement of low lighting video," SIGGRAPH 2010, Los Angeles, California, July 25–29, 2010.
4. S.S. Agaian, B. Silver, and K.A. Panetta, "Transform coefficient histogram-based image enhancement algorithms using contrast entropy," *IEEE transactions on image processing*, **16**(3), 741–758 (2007).
5. M. H. Asmare, V.S. Asirvadam, Iznita, A. Fadzil, and M. Hani, "Image enhancement by fusion in contourlet transform," *International journal on electrical engineering and informatics*, **2**(1), 2010.
6. A. Ilie, R. Raskar, and J. Yu, "Gradient domain context enhancement for fixed cameras," *Pattern recognition and artificial intelligence*, **19**(4), 533–549 (2005).
7. A. Yamasaki, H. Takauji, Shun'ichi Kaneko, T. Kanade, and H. Ohki I, "Denighting: enhancement of nighttime image for a surveillance camera," *Proc. ICPR, San Diego, CA, USA* (2008).
8. C. Stauffer and G. Wel, "Learning patterns of activity using real-time tracking," *Pattern analysis and machine intelligence*, **22**(8), 747–757 (2000).
9. P. KaewTraKulPong and R. Bowden, "An improved adaptive background mixture model for real-time tracking with shadow detection," In *Processing 2nd European workshop on advanced video based surveillance systems, AVBS01*, 2001.
10. E.H. Land and J.J. McCann, "Lightness and retinex theory," *Journal of the optical society of america*, **61**, 1–11 (1971).
11. C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," In *Proceedings of ICCV*, pp. 836–846, 1998.
12. F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic range images," *ACM transactions on graphics*, **21**(3), 257–266 (2002).