

# ADAPTIVE DEBLURRING OF SURVEILLANCE VIDEO SEQUENCES THAT DETERIORATE OVER TIME

*Konstantinos Vougioukas, Bastiaan J. Boom and Robert B. Fisher*

University of Edinburgh  
School of Informatics,  
10 Crichton St, Edinburgh, EH8 9AB, U.K

## ABSTRACT

We present a method for restoring the recordings obtained from surveillance cameras whose quality deteriorates due to dirt or water that gathers on the camera's lens. The method is designed to operate in the surveillance setting and makes use of good quality frames from the beginning of the recorded sequence to remove the blur at later stages caused by the dirty lens. A background subtraction method allows us to obtain a stable background of the scene. Based on this background, a multiframe blind deconvolution algorithm is used to estimate the Point Spread Function (PSF) of the blur. Once the PSF is obtained it can be used to deblur the entire scene. This restoration method was tested on both synthetic and real data with improvements of 15 dB in PSNR being achieved by using clean frames from the beginning of the recorded sequence.

*Index Terms*— Blind Image Deconvolution, Video Deblurring, Point Spread Function

## 1. INTRODUCTION

The quality of recordings obtained by surveillance cameras usually deteriorates over time due to the dirt or water that gathers on the camera lens. This problem is more intense in the case of underwater surveillance because of the large number of floating particles and algae that also accumulate on the lens. This build up of dirt causes the video sequences to appear blurred, so surveillance cameras have to be cleaned regularly. Even so the quality of the recordings will exhibit a significant deterioration until the cleaning. Certain environments make cleaning the lens harder due to the fact that the camera is not easily accessible. The quality of the recordings could be improved with the use of video processing techniques called *video restoration* techniques. These techniques can be used both on-line or to post-process the footage so that the archived videos are of a higher quality.

Video restoration methods usually require a model to explain the blurring effect that is present in the frames. To model the blur at time instant  $t$ , the blurry frame  $g_t$  at that time can be described as the result of the convolution of the original

frame  $f_t$  with an unknown PSF  $h_t$  and added noise  $\eta$ :

$$g_t(x, y) = h_t(x, y) \star f_t(x, y) + \eta_t(x, y), \quad (1)$$

where we assume  $f_t, g_t$  are 2D images, and  $\mathbf{p} = \{x, y\}$  is the pixel position in them.

If the PSF is known then deconvolution techniques such as *Wiener Filtering* can be used to restore the frames back to their original quality. However, in most practical cases the PSF of the blur  $h$  as well as the original frames  $f_t$  of the scene are unknown and have to be estimated. Methods that estimate both the PSF of the blur and the original images are called *blind image deconvolution* methods.

In the surveillance problem, the monitored scene does not change drastically in time. In most cases, frames recorded in the past with a clean lens will still have important information about the scene. This paper proposes to use past frames recorded with a clean lens so as to improve the estimate of the PSFs. This is achieved by modifying the Multichannel Blind Deconvolution (MBD) method proposed by Sroubek and Flusser in [1, 2] to achieve deblurring in video surveillance recordings. The method uses multiple consecutive frames of a video sequence to estimate the underlying clean frames as well as the PSFs that cause the blur. We have extended this method allowing it to also use frames taken from the beginning of the sequence which do not yet contain the blurs due to the dirt. This makes full use of all the information available in the recording and improves the methods performance. An effort was made during the design of this method to make use of the clean frames only indirectly so as to gain information about the PSFs and not directly in our frame estimation. This ensures that no temporal information in the scene, for instance a moving object, is corrupted.

## 2. PRELIMINARIES

The blind deconvolution problem is a very ill-posed problem because it tries to solve (1) with respect to both  $h$  and  $f$ . As mentioned in [3] there are two main approaches to performing blind image deconvolution. One way is to extract the PSF based on external information and then proceed to perform

non-blind deconvolution. In this case a parametric blur model may be used to identify the most likely PSF from observation or based on some identification methods [4]. The other approach tries to simultaneously estimate the PSF and original image and is much more widespread. Most algorithms that do this use an alternating approach to iteratively identify the PSF and the image [5, 6]. Representative blind deconvolution methods can be found in [3, 7, 8].

Multichannel methods like the one proposed in [9] deals with the deconvolution problem by using multiple blurred images. It is assumed that the blurred images  $g_i$  correspond to the same original image  $f$  but have been blurred with different PSFs  $h_i$ . This allows the problem to be described by the following system of equations:

$$\begin{aligned} g_1 &= h_1 \star f + \eta_1 \\ g_2 &= h_2 \star f + \eta_2 \\ &\vdots \\ g_m &= h_m \star f + \eta_m \end{aligned} \quad (2)$$

This has significantly narrowed-down the search for the unknowns due to the fact that the system of equations has  $m$  equations for  $m + 1$  unknowns and is now less ill-posed.

A method that uses this approach is the one proposed in [1, 2, 10] which uses multiple blurred images of the same scene shot consecutively. This method solves the regularized minimization problem:

$$\begin{aligned} E(f, h_1, \dots, h_m) &= \frac{1}{2} \sum_{i=1}^m \|h_i \star f - g_i\|^2 \\ &+ \lambda Q(f) + \gamma R(h_1, \dots, h_m), \end{aligned} \quad (3)$$

where the distance is calculated between the original image  $f$  convoluted with the PSFs  $h_i$  and multiple consecutive frames  $g_i$  together with regularization terms  $Q(f)$  and  $R(h_1, \dots, h_m)$  and their respective weights  $\lambda$  and  $\gamma$ . Examples of the regularisation terms  $Q$ ,  $R$  are given in [1]. In order to minimize (3) alternate minimizations of  $E$  are performed with respect to  $f$  and  $h_1, \dots, h_m$ . The two steps that are repeated by the Alternating Minimization (AM) algorithm are:

$$h_p^n = \arg \min_{h_p} E(f^{n-1}, h_1, \dots, h_m) \quad (4)$$

$$f^n = \arg \min_f E(f^{n-1}, h_1^n, \dots, h_m^n), \quad (5)$$

where  $f^n$  and  $h_i^n$  are the estimates of the  $n$ -th iteration of the AM algorithm. This minimization algorithm is a steepest descent algorithm applied alternately to the subspaces of the images and the PSFs.

### 3. METHODOLOGY

The MBD method described in the previous section was originally intended for image deblurring. The modification

proposed in this work allows it to be used on consecutive frames of video sequences and in combination with clean frames recorded in the past. The proposed method is designed to work in surveillance situations and assumes that the background of the scene doesn't greatly vary. We maintain a dictionary of the background at each pixel. A background subtraction method is used to detect pixels that belong to the foreground.

By replacing the values of the foreground pixels with median of those in the background dictionary, the method obtains frames that are quite similar to each other, even in cases of increased activity in the foreground. The clean past frames are processed in a similar manner to produce clean frames that are similar to the recent blurry ones. Based on these consecutive blurry frames and the past clean images, the corresponding PSFs can be computed. The original frames can then be deblurred using these PSFs. This leads to the method which comprises of the following 3 steps:

1. Background subtraction to ignore moving objects in scene (Section 3.1)
2. Dictionary to reconstruct missing background pixels (Section 3.2)
3. Computation of PSF based on past frames for deburring (Section 3.3)

#### 3.1. Background Subtraction

In order to use the MBD method on video the consecutive frames have to be very similar. In certain cases such as underwater surveillance videos there is a heavy traffic in the scene from fish that swim by. Consecutive frames therefore might not be very similar in these cases so a background subtraction algorithm can be used to detect the set of foreground pixels  $F_g$  and background pixels  $B_g = \forall \mathbf{p} \notin F_g$  in a image  $g$ . Several methods of background subtraction can be used to detect the foreground  $F_g$  but they must be adaptive and able to deal with jitter in the scene. A suitable background subtraction algorithm and the one used in our experiments is the Background Subtraction proposed in [11], which uses a Gaussian Mixture to model the background.

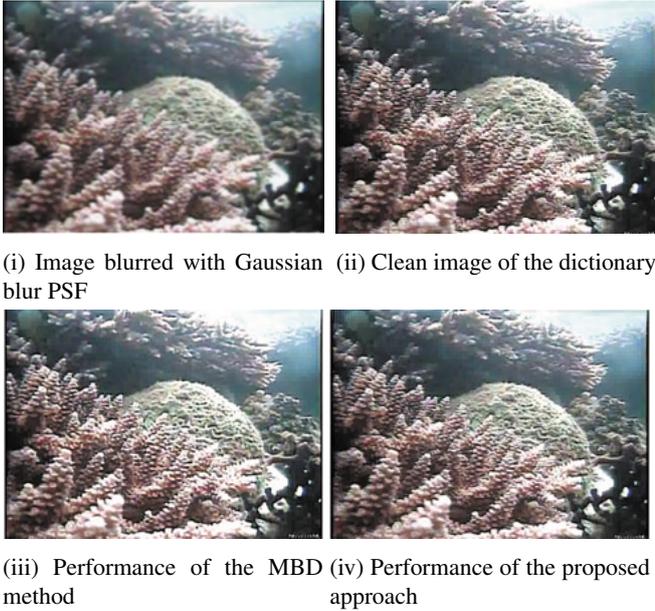
#### 3.2. Using a Dictionary to Reconstruct Missing Background Pixels

The detected foreground pixels  $F_g$  can be replaced by corresponding pixels in the dictionary. In order to estimate the corresponding background pixels of an image  $g_t$ , we propose maintaining a dictionary of the  $n$  most recent frames  $\{t-1, \dots, t-n\}$  and using the median pixels of the dictionary to replace those of the foreground. The estimated background image can be constructed by

$$g_t^B(\mathbf{p} \in B_{g_t}) = g_t(\mathbf{p} \in B_{g_t})$$

$$g_t^B(\mathbf{p} \in F_{g_t}) = \text{median}(g_{t-1}(\mathbf{p} \in F_{g_t}), \dots, g_{t-n}(\mathbf{p} \in F_{g_t}))$$

The fact that the dictionary is constantly updated by inserting the most recent frame and discarding the oldest helps to maintain any temporal information such as the illumination of the scene. Once the foreground objects have been removed the MBD method can be used on a dictionary of consecutive frames  $g^B$  to produce an estimate of the PSF. For off-line applications, this dictionary is centred around the frame we wish to deblur. The size of the dictionary can vary depending on the video (especially the frame rate), but in general needs to contain only a few frames so that the algorithm is more adaptive. Usually a dictionary of 3-5 frames is adequate for producing a good estimate of the blur PSF.



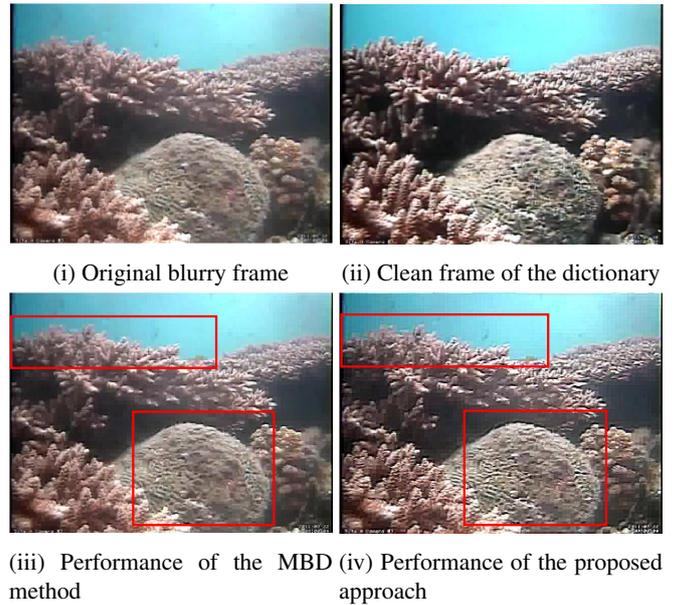
**Fig. 1:** Performance of our proposed method and the classical MBD approach [2] for a frame corrupted with synthetic gaussian blur

### 3.3. Using Clean Past Frames

We propose modifying the MBD algorithm to include past clean frames  $g_{clean}^B$  of the background in its dictionary. This will help the algorithm to better estimate the PSF. Clean frames can already be included in Equation 3, in this case however a PSF will be computed for the clean frame as well. In order to further limit the search of the minimization problem, we propose an additional constraint (second term in Equation 6) that limits the search even more based on the clean frames  $g_{k,clean}^B$ . By taking into account this constraint the minimization problem of Equation 3 therefore becomes

$$E(f, h_1, \dots, h_m) = \frac{1}{2} \sum_{i=1}^m \|h_i \star f^B - g_i^B\|^2 + \sum_{k=1}^l \|g_{k,clean}^B - f^B\|^2 + \lambda Q(f^B) + \gamma R(h_1, \dots, h_m) \quad (6)$$

In Equation 6,  $l$  is the number of clean frames that are used and  $m$  is the number of blurry ones. Typically few blurry frames and a single clean frame are used. By minimizing this equation in a similar fashion as Equation 3, we obtain the PSFs  $h_i$  as well as an estimate  $f^B$  of the original background scene. Once the PSFs  $h_i$  has been calculated then the original blurred frames  $g_i$  of the sequence which contain foreground can be deblurred using non-blind deconvolution methods like Wiener filtering. This produces a final estimate of the original image  $f$ , containing both background and foreground. Typically a single clean frame, which is representative of the scene should be used in the dictionary. Equation 6 shows that adding additional clean frames is equivalent to using the mean as the representative clean frame.



**Fig. 2:** Performance of our proposed method and the classical MBD approach [2] for a frame of the underwater recordings

## 4. EXPERIMENTS & RESULTS

In this paper, two different kinds of experiments are performed: In the first experiment, we added synthetic Gaussian blur and white noise to clean recordings shot in the underwater scene. This allows us to verify and measure the performance of our method on synthetic data. In the second experiment, we look at some real scenarios: In the case of underwater cameras, a set of recordings has taken over a period of several days, where the quality degrades rapidly. For the indoor surveillance video, it was difficult to obtain any recordings over time. However, a dirty translucent plastic film was placed in front of the camera, mimicking the effect of a dirty lens. In these real scenarios, measures of the integrity against the original signal  $f$  are not possible, so improvement

will have to be based on the reader’s visual inspection.

### 4.1. Synthetic Blur

The first experiment was done using a synthetic Gaussian blur. A clean recording was corrupted using a Gaussian blur PSF. In order to better simulate the fact that the blur might not be identical in two consecutive frames the coefficients of the PSF were also corrupted by additive zero mean gaussian noise with a standard deviation of 0.02. The resulting blurred image had a PSNR of 21.55 dB. For our proposed method a similar clean frame was taken from the clean recording at a different moment. The results obtained by using the MBD method and our approach are shown in Figure 1. The PSNR for the classical MBD approach was 24.75 dB whereas our method yielded 35.1 dB.

### 4.2. Underwater Recordings

The algorithm was initially tested on underwater surveillance videos. The quality of the recordings deteriorates over time due to the fact that particles and algae are gathered on the protective glass covering the lens. As the amount of dirt increases the frames become blurrier. Experiments were held using a sliding dictionary containing 3 recent blurry frames and one past frame shot with a clean lens (Figure 2ii.). The clean frame was chosen so that the illumination conditions were similar to those of the blurry sequence. The effects of the deblurring can be seen in the images shown in Figure 2, where more details are visible in the regions highlighted by the red rectangles by using the proposed approach.



(i) Original blurry frame (ii) Clean image of the dictionary



(iii) Performance of the MBD method (iv) Performance of the proposed approach

**Fig. 3:** Performance of our proposed method and the classical MBD approach [2] for a frame of the indoor recordings



(i) Details of MBD method's reconstruction (ii) Details of the proposed approach's reconstruction

**Fig. 4:** Zoomed in details of reconstructions of the MBD method and our proposed approach

### 4.3. Indoor Recordings

Another experiment was performed in a simulated similar situation. A camera is used to record videos in a room in order to simulate the surveillance camera set-up. Similar recordings were afterwards obtained with a dirty film in front of the camera. The dictionary used contained 3 frames shot with the dirty film and one shot without the film. The results are shown in Figure 3 and can be compared to the approach that only uses blurry frames in the dictionary. The blurred frames however contain a person in the foreground that is not in the clean frame. The results from the two methods are quite similar, however close inspection reveals for instance high frequency artifact under the door handle in Figure 3iii. Figure 4 shows that the face (which was not part of the clean image) is deblurred correctly, however Figure 4i looks less realistic than Figure 4ii because of the sharpness of edges and the higher contrast in the face.

## 5. CONCLUSIONS

Based on the results shown in the previous section it is clear that when dealing with surveillance videos better results can be achieved when using an algorithm that takes into account past frames that are cleaner. Our proposed method, which is specially designed for the case of the surveillance setting is able to produce better results than those of the standard MBD approach by using a few frames from archived good quality recordings. The proposed algorithm can be used in many cases where a high standard of footage is desired without requiring constant maintenance for the camera lens. In many cases, the computationally expensive blur PSF calculation step does not need to be calculated for every frame since we can assume that the blur introduced by the dirty lens will not drastically change for a large number of frames. Therefore by reducing the frequency of this calculation the computational complexity introduced by this set will drop to a constant making the method suitable for real-time application.

## 6. REFERENCES

- [1] Flusser Jan Sroubek Filip, “Multichannel blind iterative image restoration,” *Image Processing, IEEE Transactions on*, vol. 12, no. 9, pp. 1094 – 1106, 2003.
- [2] Flusser Jan Sroubek Filip, “Multichannel blind deconvolution of spatially misaligned images,” *Image Processing, IEEE Transactions on*, vol. 14, no. 7, pp. 874–883, 2005.
- [3] D Kundur and D. Hatzinakos, “Blind image deconvolution revisited,” *Signal Processing Magazine, IEEE*, vol. 13, no. 6, pp. 61 – 63, 1996.
- [4] B Chalmond, “PSF estimation for image deblurring,” *CVGIP: Graphical Models and Image Processing*, vol. 53, no. 4, pp. 364–372, 1991.
- [5] RL Lagendijk, “Iterative identification and restoration of images,” in *Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on*, 1988, pp. 992 – 995 vol.2.
- [6] AK Katsaggelos and KT Lay, “Maximum likelihood blur identification and image restoration using the EM algorithm,” *Signal Processing, IEEE Transactions on*, vol. 39, no. 3, pp. 729 – 733, 1991.
- [7] D Kundur and D. Hatzinakos, “Blind image deconvolution,” *Signal Processing Magazine, IEEE*, vol. 13, no. 3, pp. 43 – 64, 1996.
- [8] Campisi P. and Egiazarian K., *Blind image deconvolution: theory and applications*, 2007.
- [9] U Javed, MM Riaz, and TA Cheema, “Multichannel blind image deconvolution,” in *Applied Sciences and Technology (IBCAST), 2012 9th International Bhurban Conference on*, 2012, pp. 95 – 99.
- [10] F Sroubek and P Milanfar, “Robust Multichannel Blind Deconvolution via Fast Alternating Minimization,” *Signal Processing, IEEE Transactions on*, vol. 21, no. 4, pp. 1687–1700, 2012.
- [11] Z Zivkovic, “Improved adaptive Gaussian mixture model for background subtraction,” in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, 2004, pp. 28 – 31 Vol.2.