

Stabilizing Atmospheric Turbulence in a Video Surveillance using Dual Tree Complex Wavelet Fusion

Abhijit R. Mange
Research Scholar

Government Engineering College Aurangabad

Chitra M. Gaikwad
Asst. Professor

Government Engineering College Aurangabad

ABSTRACT

Many valuable data in a video can be lost because of atmospheric distortion such as mirage, haze, fog, etc. This is a challenge in video surveillance system. We propose complex wavelet fusion for atmospheric turbulence algorithm for this purpose. This algorithm consists of series of steps to extract the information from the video. Frame selection method is used for selection of informative ROI only from good quality frame. These ROI were then fused to get the required objects. The fusion is performed in the Dual Tree Complex Wavelet Transform (DT-CWT) which employs two different real discrete wavelet transforms (DWT) to provide the real and imaginary parts of the CWT. Finally Haze removal is done by using dark channel prior mechanism. The proposed method is shown to significantly enhance the visual quality in a range of video surveillance.

General Terms

Dual tree complex wavelet transform (DT-CWT).

Keywords

Turbulence, Restoration, Mitigation, Segmentation

1. INTRODUCTION

Atmospheric turbulence is a naturally occurring phenomenon that can severely degrade the visual quality of video signals during acquisition. There are various types of atmospheric distortions such as fog or haze which reduce contrast and turbulence due to temperature variations or aerosols. Video footage in public areas is affected by such atmospheric distortions which result in blurring, wavering and warping of image of objects in the scene. In strong turbulence, blurring effects which are present in the video imagery, scintillation producing small-scale intensity fluctuations in the scene as well as shearing effect is observed. Due to shearing effect, different parts of objects are perceived to be moving in different direction. These effects are mainly found at locations such as hot roads and deserts. This is particularly a problem close to the ground in hot environments and also in long range surveillance applications where images can be acquired over distances up to 20 km. It is found difficult to interpret information behind the distorted layer due to turbulence effects. This leads to faster and greater microscale changes in the air's refractive index. In situations when the ground is hotter than the air above it, the air is heated and forms horizontal layers. Due to this, increment in the temperature difference of ground and air is observed so thickness of layer shrinks and air layer move upwards leading to change in the air's refractive index. Hence, there has been significant research activity attempting to faithfully reconstruct this useful information using various methods. The perfect solution however seems impossible since this problem is irreversible, although it can be simply written as Eq. 1.

$$I_{obv} = DI_{idl} + \varepsilon \quad (1)$$

Here I_{obv} and DI_{idl} are vectors that contain the observed and ideal images, respectively. Matrix stands D for geometric distortion and blur, while ε stands for noise [1]. Different approaches have attempted to solve this problem by modeling it as a point spread function (PSF), in which D is considered as a convolution matrix, and then employing deconvolution with an iterative process to estimate I_{idl} . The PSF is generally unknown for the atmospheric distortion case so blind deconvolution is employed [2]. Method for removal of spatio temporal distortions utilizes a set of images to construct one enhanced image as it not possible with a single image. Fig. 1 represents multi frame method where most approaches employ all functions or subset of them. There are 2 different ways for image restoration in the given fig 1. The first one which is shown by red line employs an image registration technique with deformation estimation attempting to align objects by solving small movements of camera and temporal variation. The image fusion block may optionally be employed (yellow line) combining several aligned images. Several aligned images can be combined using image fusion block. Finally deblurring process is applied to combined image.

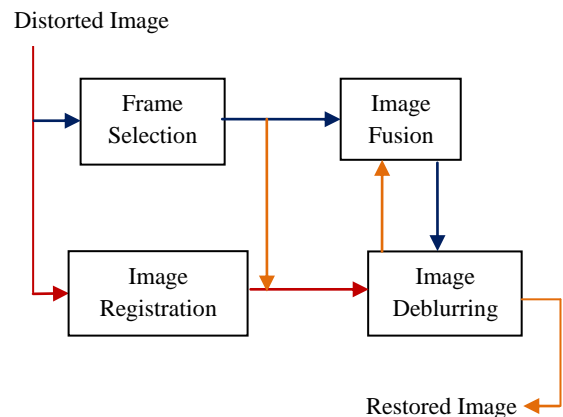


Fig 1: Block diagram of image restoration for atmospheric turbulence

The other route for image restoration is from frame selection and image fusion block (blue line). This route is often called as “lucky region” Using image quality metric best quality image frames are selected. Selection of images is normally applied in the spatial frequency domain, to extract the best quality frames (less distorted and least blurred). These are then combined in an intelligent manner. This process can be improved with ROI alignment technique as region of interest containing important information gets aligned. Finally deblurring is applied [1].

Yet, atmospheric turbulence mitigation is an important challenging problem. Blind deconvolution methods works well in real atmospheric images but cannot remove high atmospheric turbulence. It also suffers from spatial and temporal variation due to PSF. Fusion is employed for mitigating atmospheric distortion caused by air turbulence. Image fusion is done at the feature level by a region-based scheme. To produce a set of regions, an image is initially segmented in some way. For determination of various features from which images are included in the fused image various properties of these regions are calculated so that more intelligent semantic fusion rules based on actual features in the image, rather than on single or arbitrary groups of pixels can be considered [6]. The fusion is performed in the Dual Tree Complex Wavelet Transform (DT-CWT) domain as it provides near shift invariance and directional selectivity [3]. Additionally, the phase of a CWT coefficient is robust to noise and temporal intensity variations which provide an efficient tool to remove the distorting ripple. Before applying fusion, a set of selected images or ROIs must be aligned. An ROI alignment approach is used for distorted images. As randomly distorted images do not provide identical features, conventional methods cannot be used to find matching features. So, a morphological image processing technique is employed. Subsequently the ROI (or whole image) from only the informative frames measured by a novel quality matrix, based on sharpness, ROI size and intensity similarity is selected. Sigmoid function is used so that one factor should not dominate the other one. Then a non-rigid image registration is applied. We propose a new method for haze removal. After the fusion, haze removal is done by using dark channel prior mechanism.

2. PROPOSED MITIGATION SCHEME

The proposed process is shown in Fig. 2. Details of each step are described below.

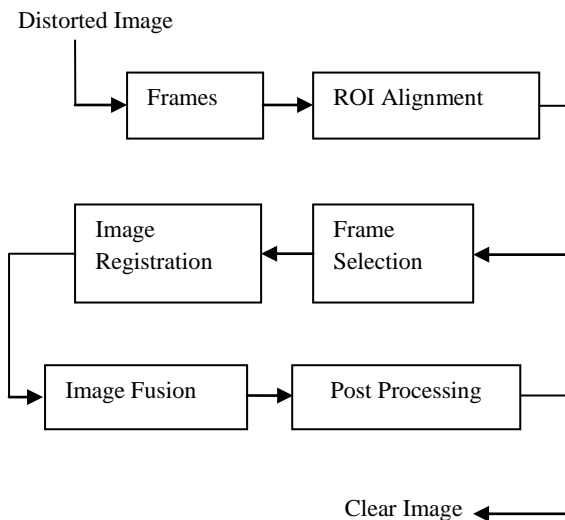


Fig 1: Block diagram of the proposed method

2.1 ROI Alignment

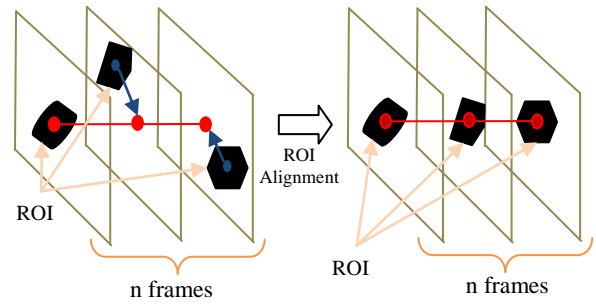


Fig 3: ROI alignment technique

The ROI in each frame become misaligned in long surveillance while capturing video in the presence of atmospheric turbulence. Due to camera movements, high variation in the image is observed. The inter frame distance between the shaking objects may be so large that image registration cannot be applied directly as it significantly impacts on image quality more than the turbulence. Fig 3 explains the ROI alignment technique. A simple ROI alignment method might choose a subset of the original frames where the inter frame distance is less. One cannot use matching algorithm which use feature detection as it is not suitable for air turbulence problem because the strong gradient within each frame distorts randomly. So a simple approach using morphological image processing is proposed. The first containing ROI is marked. The histogram, generated from the selected ROI and the surrounding area, is utilized to find an Otsu threshold which is used to convert the image to a binary map [4]. Areas which are connected to the edge of image are removed by applying erosion process. This process is repeated until the area near ROI is isolated. Otsu threshold and the number of iterations are same for other frames also. The centre position of each mask has to be computed. The area closest in size and location to the ROI in the first frame is used in case of more than one isolated area. Finally ROI is aligned by using center of mask across set of frames. Frame selection process removes all the incorrectly detected ROIs. Frames which are selected in this process are significantly different from other frames.

2.2 Frame Selection

In this method, only selected frames are used for the restoration of the undistorted image; the bad images (e.g. the very blurred ones) are not taken into consideration as they would possibly deteriorate the fused result. Three factors; sharpness G_n , intensity similarity S_n and detected ROI size A_n are used for selection of set of images.

2.2.1 Sharpness

To determine the amount of details which image can convey, sharpness G_n is used and it is consider being the most important factor to improve the quality of image. Here, the sharpness parameter G_n is computed from the summation of intensity gradient or the high pass coefficient magnitudes.

2.2.2 Intensity Similarity

Intensity similarity S_n is used to remove outliers. There is a high probability that the frames with significantly different contents than the others are greatly distorted. To compute S_n , the average frame of the whole sequence is used as a reference for calculating the mean square error (MSE) for frame n. Then

MSE^{-1} represents the similarity of each frame. Illumination changes cannot be adopted by this approach.

2.2.3 Detected ROI Size

Larger ROIs are likely to contain more useful information so it is important to calculate the size of ROI containing more details about image. Detected ROI size A_n gives the total number of pixels in the ROI.

The cost function C_n [1] for frame n is computed using Eq. 2.

$$C_n = \frac{w_G G_n}{\lambda_G + |G_n|} + \frac{w_S S_n}{\lambda_S + |S_n|} + \frac{w_A A_n}{\lambda_A + |A_n|} \quad (2)$$

Where w_k and λ_k are the weight and slope control of the factor $k \in \{G, S, A\}$, respectively [1]. The sigmoid function is used here to prevent one factor dominating the others, e.g. a blocking artifact may cause significantly high values of sharpness, yet this frame should not be included in the selected data set. The λ_k is set to equal the mean of factor k so that at the mean value, the cost value of such factor is 0.5. The cost C_n is ranked from high to low. To find how many frames should be included in the selected set, the Otsu method can be applied [4].

2.3 Image Registration

The phase shift properties of the DT-CWT are used for registration of non-rigid bodies [5]. Phase based multidimensional volume registration is robust to noise and temporal intensity variations. Coarser level complex coefficients and finer level coefficients are done iteratively for motion estimation. Coarser level complex coefficients are used to determine large motion components whereas finer level coefficients are used to refine the motion field.

2.4 Image Fusion

The region based image fusion technique using complex wavelets is being used to remove atmospheric turbulence. This method first reconstructs each image into the DT-CWT domain. The DT-CWT is used in image fusion as it has shift invariance, orientation selectivity and multiscale properties. In image fusion useful information from a number of source images are selected and combined into a new image [1]-[9]. To divide the image into R regions, version of the combined morphological spectral unsupervised image segmentation and a multiscale watershed segmentation [7] is employed. The average of the lowpass values of all images is used for construction of the lowpass DT-CWT coefficients of the fused image. Activity measurement indicating the importance of that region is categorized according to highpass coefficient. Operating on each band separately is necessary to produce proper sharper result. The priority P of region $r_n^\theta \in R$ in image n is computed with the detail coefficients $d_n^{\theta,l}(x,y)$ of level l and sub-band θ as shown in Eq. 3, where $|r_n^\theta|$ is the size of such region used for normalization [1]. Priority map is employed for consequent construction of fused image.

$$P(r_n^\theta) = \frac{1}{|r_n^\theta|} \sum_{\forall l, (x,y) \in r_n^\theta} |d_n^{\theta,l}(x,y)| \quad (3)$$

Due to turbulence distortion segmentation boundaries which separate inhomogeneous region vary significantly frame to frame in this air turbulence scenario. This scenario is very much different from other image fusion problems. The ultimate goal is to get the sharpest and most temporally

consistent boundaries. Instead of using $P(r_n^\theta)$, the maximum of DT-CWT coefficient magnitudes from all frames are used at the boundary of each region. The phase of the complex wavelet coefficients corresponds to the type of dominant directional features in its support regions. Therefore, it is very helpful for reducing the distortion due to rippling. Hence, the DT-CWT coefficients $d_n^{\theta,l}(x,y)$, of the fused image are adjusted with a unit vector representing the average phase from all frames, N , used in the fusion process (Eq. 4). Pixels deviate from their actual positions in turbulent motion so the average phase can be used with approximately zero mean and with a quasi-periodic motion [1].

$$d^{\theta,l}(x,y) = \frac{\sum_n d_n^{\theta,l}(x,y)}{\left| \sum_n d_n^{\theta,l}(x,y) \right|} |d_n^{\theta,l}(x,y)| \quad (4)$$

Shrinkage function is applied to reduce noise. This function derived as Maximum A Posteriori (MAP) estimators. To boost high pass coefficient magnitudes, a gain $A_g > 1$ can be applied only when sharpening is required. The drawback of this technique is that it might cause the noise to become clearly visible again. Due to this reason it is recommended to only enhance the high pass coefficient when they form regions of large magnitude. Now, the binary map $M^{\theta,l}$ for each subband is created. The $M^{\theta,l} = 1$ if $\tau > \tau_0$, where τ_0 is a predefined threshold. The isolated pixels are subsequently removed from $M^{\theta,l}$. Eq. 5 depicts the modified high pass coefficients [1].

$$d^{\theta,l} = A_g^{\theta,l} M^{\theta,l} A_s^{\theta,l} d^{\theta,l} \quad (5)$$

2.5 Haze Removal

A problem that often accompanies turbulence is haze or fog. We are using dark channel prior mechanism for haze removal. Dark channel prior is based on the statistics of outdoor haze free images, we find most of the local region of the haze free image that contain some type of pixels that are called as dark pixels whose intensity becomes very low in atleast any one of the color channel. Using this prior, we can directly estimate the thickness of the haze and recover a high quality haze-free image [8]. In haze free images the intensity of these dark pixels in that channel is mainly contributed by the atmospheric light, therefore these dark pixels can directly provide an accurate estimation of haze transmission.

3. CONCLUSION

This paper has introduced a method for mitigating atmospheric distortion in long range surveillance imaging. Using region-based fusion in the DT CWT domain, visibility of ROI in image sequence is improved. We have used a simple ROI alignment method and a cost function for frame selection to pre-process the distorted sequence. We propose dark channel prior technique for haze removal. This helps us to significantly enhance the visual quality in a range of video surveillance.

4. ACKNOWLEDGMENTS

I feel a deep sense of gratitude to Prof V. P. Kshirsagar HOD of Computer Science and Engineering Department for his motivation and support during this work. I am also thankful to the Principal, Government College of Engineering, Aurangabad for being a constant source of inspiration.

5. REFERENCES

- [1] Nantheera Anantrasirichai, Member, IEEE, Alin Achim, Senior Member, IEEE, Nick G. Kingsbury, Fellow, IEEE, and David R. Bull, Fellow, IEEE, "Atmospheric Turbulence Mitigation Using Complex Wavelet-Based Fusion" *IEEE Transactions On Image Processing*, Vol. 22, No. 6, June 2013.
- [2] S. Harmeling, M. Hirsch, S. Sra, and B. Scholkopf, "Online blind image deconvolution for astronomy," in *Proc. IEEE Conf. Comp. Photogr.*, Apr. 2009, pp. 1–7.
- [3] I.W. Selesnick, R.G. Baraniuk, and N.G. Kingsbury, "The dual-tree complex wavelet transform," *Signal Processing Magazine, IEEE*, vol. 22, no. 6, pp. 123 – 151, nov. 20.
- [4] N. Otsu, "A threshold selection method from gray-level histograms," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 9, no. 1, pp. 62 –66, jan. 1979.
- [5] H. Chen and N. Kingsbury, "Efficient registration of nonrigid 3-d bodies," *Image Processing, IEEE Transactions on*, vol. 21, no. 1, pp. 262–272, jan. 2012.
- [6] J. J. Lewis, R. J. O'Callaghan, S. G. Nikolov, D. R. Bull, C. N. Canagarajah, and Essa Basaeed, "Region-based image fusion using complex wavelets," in *Proc. 7th Inter. Conf. Info. Fusion*, 2004, pp. 555–562.
- [7] R.J. O'Callaghan and D.R. Bull, "Combined morphological-spectral unsupervised image segmentation," *Image Processing, IEEE Transactions on*, vol. 14, no. 1, pp. 49 –62, jan. 2005.
- [8] Kaiming He ; Dept. of Inf. Eng., Chinese Univ. Of Hong Kong, Hong Kong, China ; Jian Sun ; Xiaoou Tang, "Single image haze removal using dark channel prior", *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on 20-25 June 2009*.
- [9] T. Wan, N. Canagarajah, and A. Achim, "Segmentation driven image fusion based on alpha-stable modeling of wavelet coefficients," *IEEE Trans. Multimedia*, vol. 11, no. 4, pp. 624–633, Jun. 2009.