

Dual-Tree DWT based turbulence mitigation in visual surveillance

M.NAWAZ SHAREEF¹, S.SWARNA LATHA²

M.Tech Student, Dept of Electronics & Communication Engineering, SVUCE, Tirupati, India¹

Associate Professor, Dept of Electronics & Communication Engineering, SVUCE, Tirupati, India²

Abstract: A long-distance imaging system can be strongly affected by atmospheric turbulence. In this paper, we propose a novel method for mitigating the effects of atmospheric distortion on observed images, particularly airborne turbulence which can severely degrade a region of interest (ROI). In order to extract accurate detail about objects behind the distorting layer, a simple and efficient frame selection method is proposed to select informative ROIs only from good quality frames. The ROIs in each frame are then registered to further reduce offsets and distortions. We solve the space-varying distortion problem using region-level fusion based on the dual tree complex wavelet transform. Finally, contrast enhancement is applied. We further propose a learning-based metric specifically for image quality assessment in the presence of atmospheric distortion. This is capable of estimating quality in both full and no-reference scenarios. The proposed method is shown to significantly outperform existing methods, providing enhanced situational awareness in a range of surveillance scenarios.

Keywords: Dual tree complex wavelet transform (DT-CWT), image restoration, quality metrics, region-level fusion.

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. The perfect solution however seems impossible since this problem is irreversible, although it can be simply written as Eq. 1.

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

Here equation of two vectors that contain the observed and ideal images, respectively. Matrix stands for geometric distortion and blur, while stands for noise [1]. Different approaches have attempted to solve this problem by modeling it as a point spread function (PSF), in which is considered as a convolution matrix, and then employing deconvolution with an iterative process to estimate. 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..

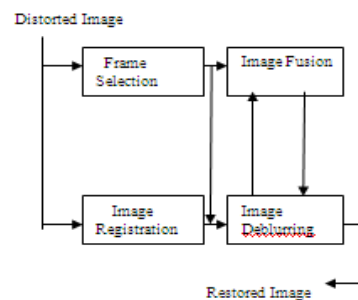


Fig 1: Block diagram of image restoration for atmospheric turbulence

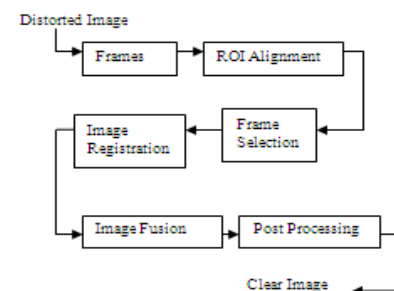


Fig 2: Block diagram of the proposed method

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). 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. 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.

II. PROPOSED MITIGATION SCHEME

We propose a new fusion method for reducing the effects of atmospheric turbulence as depicted in Fig. 2. First, before applying fusion, a subset of selected images or ROIs must be aligned. Here we introduce a new alignment approach for distorted images. As randomly distorted images do not provide identical features, we cannot use conventional methods to find matching features. Instead, we apply a morphological image processing technique, namely erosion, to the ROI (or whole image) based only on the most informative frames. These are selected using a quality metric based on sharpness, intensity similarity and ROI size. This has advantages over pixel-based processing

since more intelligent semantic fusion rules can be considered based on actual features in the image, rather than on single or arbitrary groups of pixels. 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. Two fully decimated trees are produced, one for the odd samples and one for the even samples generated at the first level this increases directional selectivity over the DWT and is able to distinguish between positive and negative orientations giving six distinct sub-bands at each level, corresponding to $15^\circ, 45^\circ, 75^\circ$. Additionally, the phase of a DT-CWT coefficient is robust to noise and temporal intensity variations thereby providing an efficient tool for removing distorting ripples. Finally, the DT-CWT is near-shift invariant—an important property for this application. After fusion, the effect of haze is reduced using locally-adaptive histogram equalization. For convenience, we refer to this algorithm as CLEAR (Complex waveLET fusion for Atmospheric tuRBulence). Details of each step in our algorithm are described below.

A. ROI Alignment

Capturing video in the presence of atmospheric turbulence, especially when using high magnification lenses, may cause the ROI in each frame to become misaligned. The displacement between the distorted objects in the successive frames may be too large for conventional image registration, using non-rigid deformation, to cope with. Equally, matching using feature detection is not suitable since strong gradients within each frame are randomly distorted spatially. Hence, an approach using morphological image processing is proposed.

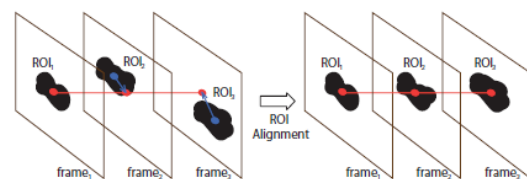


Fig.3. ROI alignment technique

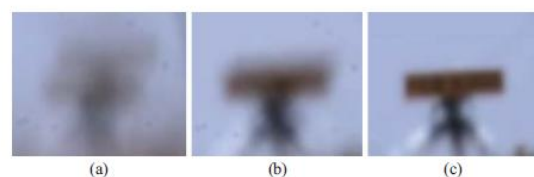


Fig. 4. Average frame of *Number Plate* sequence by using (a) 200 original frames, (b) 50 frames most similar (lowest MSE) to (a), and (c) 200 frames after applying ROI alignment approach.

The ROI (or ROIs) is manually marked in the first frame. Then the histogram, generated from the selected ROI and the surrounding area, is employed to find an Otsu threshold, which is used to convert the image to a binary map. An erosion process is then applied and the areas connected to the edge of the sub-image are removed. This step is performed iteratively until the area near the ROI is isolated. The same Otsu threshold with the same number of iterations is employed in other frames. The center position of each mask is then computed. If there is more than one isolated area, the area closest in size and location to the ROI in the first frame is used. Finally, the center of the mask in each frame is utilized to shift the ROI and align it across the set of frames (Fig. 3). Note that the frames with incorrectly detected ROIs will be removed in the frame selection process (Section II-B). These frames are generally significantly different from others. Fig. 4 demonstrates the improvement due to the proposed ROI alignment approach. The left image represents the average frame of the whole *Number Plate* sequence (see B4 in Table II and Fig. 9) and it reveals high variation due to camera movement which significantly impacts on image quality more than the turbulence. A simple alignment method might choose a subset of the original frames where the inter-frame distance is not too large. Fig. 4(b) shows an example of this where, based on mean square error (MSE), 50 original frames with the smallest difference from the average image are selected. The alignment of these frames is consequently improved as shown by their average image in Fig. 4(b). However, this truncation may lead to loss of some useful information. In contrast, the proposed ROI alignment approach keeps all frames so that an intelligent approach can be employed later to select the best subset. The average frame of the whole sequence of our method is shown in Fig. 4(c), which is clearer and sharper than the other two images [Fig. 4(a) and (b)]. This means that the aligned frames are more stable thereby producing better image registration results.

B. Frame Selection

In CLEAR, not all frames in the sequence are used to restore the image since the low quality frames would possibly degrade the fused result. A subset of images is carefully selected using three factors: sharpness, intensity similarity and detected ROI size.

1) Sharpness: G_n is one of the most important image quality factors since it determines the amount of detail an image can convey. Here, the sharpness parameter G_n is computed from the summation of the high pass coefficient magnitudes. Intensity gradients can also be used as the result is insignificantly different from high pass coefficients.

2) Intensity Similarity: S_n is employed to remove outliers. This operates under the assumption that most frames in the sequence contain fairly similar areas. Frames with significantly different content to others are likely to be 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. It should be noted that this approach is not robust to illumination changes.

3) Detected ROI Size: A_n is the total number of pixels contained in the ROI. This is used because, from observation, larger ROIs are likely to contain more useful information. The cost function C_n for frame n is computed using

$$C_n = \frac{\omega_G G_n}{\lambda_G + |G_n|} + \frac{\omega_S S_n}{\lambda_S + |S_n|} + \frac{\omega_A A_n}{\lambda_A + |A_n|}$$

Where wk and λk are the weight and slope control of the factor $k \{G, S, A\}$, respectively. 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 probably not be included in the selected data set. The λk is set to equal the mean of factor k so that at the mean value, its cost value is 0.5. The cost C_n is ranked from high to low. The Otsu method can then be applied to find how many frames should be included in the selected set.

C. Image Registration

Registration of non-rigid bodies using the phase-shift properties of the DT-CWT, as proposed, is employed. This algorithm is based on phase-based multidimensional volume registration, which is robust to noise and temporal intensity variations. Motion estimation is performed iteratively, firstly by using coarser level complex coefficients to determine large motion components and then by employing finer level coefficients to refine the motion field.

D. Image Fusion

Due to its shift invariance, orientation selectivity and multi-scale properties, the DT-CWT is widely used in image fusion where useful information from a number of source images are selected and combined into a new image.

E. Post-Processing

1) Contrast Enhancement: In many cases, atmospherically degraded images also suffer from poor contrast due to severe haze or fog. In such cases, pre- or post-processing is needed to improve image quality. Numerous techniques have been proposed for haze reduction using single images. Here we employ a simple and fast method using contrast limited adaptive histogram equalization (CLAHE). The method enhances intensity

locally, so it is suitable for applications which consider the ROI and its information content.

2) Other Possible Enhancements:

Generally the embedded parameter A_g in our approach produces sharp results; however, in cases which are out-of-focus or which lack a “lucky region”, post-processing may be required to further sharpen the images.

III. QUALITY ASSESSMENT

Image quality assessment is used to measure perceived image degradation, typically compared to an ideal or perfect image. This is important when assessing the performance of individual systems or for comparing different solutions. Image quality metrics can be classified according to the availability of a reference (distortion-free) image, with which the distorted image is to be compared. Most existing approaches are classed as full-reference (FR), meaning that a complete reference image is available.

A. No-Reference Image Quality Assessment

The JPEG quality score was one of the first NR quality assessment methods introduced. It attempts to align image quality with HVS perception by characterizing blockiness and blurring. Subsequently, the JPEG2000 (JP2K) score was proposed for blind assessment of images compressed by wavelet based coding. The Anisotropic Quality Index (AQI) is another NR metric based on measuring the variance of the expected entropy of a given image in a set of predefined directions. Recently, the combination of five distortion types, namely JPEG, JPEG2000, white noise, Gaussian blur and fast fading, were used in the Blind Image Quality Index (BIQI).

B. Proposed NR Method for Atmospheric Distortion

The methods described in Section III-A do not work well with atmospheric distortion, since they are usually based on prior knowledge of the distortion characteristics and none are derived from spatially varying distortions. In this paper, we therefore introduce a new blind image quality assessment metric specifically for this scenario. We employ support vector regression (SVR) to model and predict image quality scores using the features listed in Table I. There are three groups of features.

1) **Individual Scale:** The magnitude of highpass coefficients relate to details and sharpness of the image, while the phase can be linked to edge information. We therefore employ the mean and variance of both values to compute the feature vectors at each scale level. We decompose the image into 3 levels using the DT-CWT.

2) **Inter-Scale:** Weighted mean and variance at level l are computed using the magnitudes of the next coarse level to calculate a weight. Moreover, moderate atmospheric turbulence consistently manifests itself as clearly visible ripples along object edges.

3) **Existing NR Methods:** Existing NR scores are also included in the regression to model image quality, since these have been developed for, and generally work well for, general natural images. The SVR parameters are selected by cross validations the number of features is not large, principal component analysis is not required for dimensionality reduction. Instead, to find the best features for regression, different combinations of features were tested. We found that the combination including all features performs best. We refer to this new metric as QSVR.

IV. RESULTS AND DISCUSSION

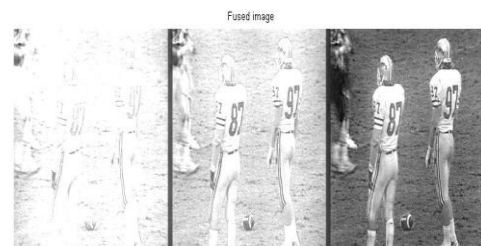


Fig.5.Average frames of Football

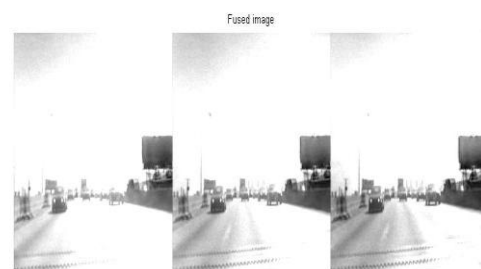


Fig.6.Average frames of visual surveillance

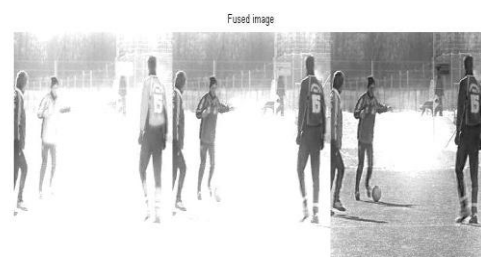


Fig.7.Average frames of soccer

Table.1: Objective Results Using FR Methods

SEQUENCE	METHOD	PSNR	NC
FOOTBALL	SVOLA	31.93	0.212
	ZHU	28.90	0.201
	CLEAR	32.02	0.221
	DT-DWT	37.34	0.231
VISUAL SURVEILLANCE	SVOLA	24.88	0.313
	ZHU	23.83	0.332
	CLEAR	25.18	0.442
	DT-DWT	35.63	0.552
SOCCER	SVOLA	19.39	0.222
	ZHU	20.34	0.332
	CLEAR	25.36	0.382
	DT-DWT	36.06	0.552

From the table.1 it can be observed that the Proposed method gives fair results compared to the other methods in terms of PSNR, NC. The results have been observed for different types of videos.

V.CONCLUSION

This paper has introduced a new method for mitigating atmospheric distortion in long-range surveillance imaging. Significant improvements in image quality are achieved using region-based fusion in the DT-CWT domain. This is combined with a new alignment method and cost function for frame selection to pre-process the distorted contrast enhancement to reduce haze interference. CLEAR offers class-leading performance for off-line extraction of enhanced static imagery and has the potential to achieve high performance for on-line mitigation for full motion video—this is topic of ongoing research. Experiments with real data show superior performance compared with existing methods. Using simulated data, full reference metrics clearly show the superiority of this method. We have also introduced a new metric, QSVR, based on support vector regression for blindly assessing image quality.

This learning-based method shows higher correlation with the FR methods than existing NR methods.

REFERNCES

- [1] J. Gilles, T. Dagober, and C. Franchis, "Atmospheric turbulence restoration by diffeomorphic image registration and blind deconvolution," in *Proc. 10th Int. Conf. Adv. Concepts Intell. Vis. Syst.*, 2008, pp. 400–409.
- [2] C. S. Huebner, "Compensating image degradation due to atmospheric turbulence in anisoplanatic conditions," *Proc. SPIE*, vol. 7351, p. 735106, Apr. 2009.
- [3] X. Zhu and P. Milanfar, "Removing atmospheric turbulence via space invariant Deconvolution," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 157–170, Jan. 2013.
- [4] I. Selesnick, R. Baraniuk, and N. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Process. Mag.*, vol. 22, no. 6, pp. 123–151, Nov. 2005.
- [5] NPTEL lectures by ministry of HRD.
- [6] Complex wavelet transforms (Wikipedia)
- [7] WAVELETS a primer by Christian Blatter
- [8] <http://research.ijcaonline.org/volume88/number5/pxc3893693.pdf>
- [9] <http://www.rohan.sdsu.edu/~jegilles/doc/acivs2008.pdf>
- [10] http://www.kyb.tue.mpg.de/fileadmin/user_upload/files/publications/attachments/vobd_paper_mpitechreport_6328%5b0%5d.pdf.

BIOGRAPHIES



M. Nawaz Shareef has received graduation degree from RAJEEV GANDHI MEMORIAL COLLEGE OF ENGINEERING in the year 2008. He received his post-graduation in the specialization of COMMUNICATION SYSTEMS from S.V.U.C.E, TIRUPATI in the year 2014. The major areas of interest are image fusion and signal processing.



S.Swarna Latha received her graduation degree from JNTUCEA in the year 2000. She received her post-graduation degree from JNTU in the year 2004. She is pursuing her Ph.D. from S.V.U.C.E. TIRUPATI in the image processing domain. She worked as lecturer at J.N.T.U.C.E.A, Anantapur for 4 years and she worked as assistant and Associate professor in the department of E.C.E., at M.I.T.S, Madanapalle, She worked as Associate professor at C.M.R.I.T, Hyderabad and presently she is the Associate Professor at S.V.U.C.E, Tirupathi.