

# MITIGATING THE EFFECTS OF ATMOSPHERIC DISTORTION USING DT-CWT FUSION

*N. Anantrasirichai, Alin Achim, David Bull*

Visual Information Laboratory  
University of Bristol, UK

*Nick Kingsbury*

Signal Processing Laboratory  
University of Cambridge, UK

## ABSTRACT

This paper describes a new method for mitigating the effects of atmospheric distortion on observed images, particularly airborne turbulence which degrades a region of interest (ROI). In order to provide accurate detail from objects behind the distorting layer, a simple and efficient frame selection method is proposed to pick informative ROIs from only good-quality frames. We solve the space-variant distortion problem using region-based fusion based on the Dual Tree Complex Wavelet Transform (DT-CWT). We also propose an object alignment method for pre-processing the ROI since this can exhibit significant offsets and distortions between frames. Simple haze removal is used as the final step. The proposed method performs very well with atmospherically distorted videos and outperforms other existing methods.

*Index Terms*— Image restoration, fusion, DT-CWT

## 1. INTRODUCTION

Various types of atmospheric distortion can influence the visual quality of a captured video signal. Distortion types include fog or haze which reduces contrast, and atmospheric turbulence due to temperature variations or airborne contaminants. Such obscurations in the lower atmosphere reduce the ability to see objects at a distance. A variation in temperature causes different interference patterns in the light refraction, leading to unclear, unsharp, waving images of the objects. Examples of this effect are found in areas such as hot roads and deserts, as well in the proximity of aircraft jet exhausts. Examples are shown in columns 1 and 2 of Fig. 3.

Since turbulence in the captured images makes it difficult to interpret information behind the distorting layer, there has been significant research activity trying 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.

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

where  $I_{obv}$  and  $I_{idl}$  are the observed and ideal images respectively.  $D$  represents geometric distortion and blur, while  $\varepsilon$  represents noise. Various approaches have attempted to solve

this problem by using blind deconvolution (BD) [1, 2]. The results, however, still exhibit artefacts since the point spread function (PSF) is assumed to be space-invariant.

It is obvious that using a single image is not sufficient to remove the visible ripples and waves, while utilising a set of images to construct one enhanced image makes more useful information available. There are two types of restoration process that use multiple images. The first employs image registration with deformation estimation to align objects temporally and to solve for small movements due to atmospheric refraction [3, 4]. Then a deblurring process is applied to the combined image (which is a challenging task as this blur is space-variant). The other group employs image selection and fusion, known as ‘*lucky region*’ techniques [5]. The regions of the input frames having the best quality in the temporal direction are selected and then are combined in an intelligent way. Recently this method has been improved by applying image alignment to those lucky regions [6].

In this paper, we propose a new fusion method to reduce image distortion caused by air turbulence. We employ a region-based scheme to perform fusion at the feature level. This has advantages over pixel-based processing as more intelligent semantic fusion rules can be considered based on actual features in the image. The fusion is performed in the Dual Tree Complex Wavelet Transform (DT-CWT) domain as it provides near shift-invariance and directional selectivity [7]. Additionally, the phase of a CWT coefficient is robust to noise and temporal intensity variations thereby providing an efficient tool for removing the distorting ripples.

Before applying fusion, a set of selected images or ROIs must be aligned. We introduce an object alignment approach for distorted images. As randomly distorted images do not provide identical features, we cannot use conventional methods to find matching features. We, instead, employ a morphological image processing technique. Subsequently we select the ROI (or whole image) from only the informative frames measured by a novel quality metric, based on sharpness, intensity similarity and ROI size. Then a non-rigid image registration is applied. After the fusion, haze and fog are removed using a locally adaptive histogram equalisation.

The remaining part of this paper is organised as follows. The proposed scheme for mitigating the atmospheric distortion is described in details in Section 2. The performance of

the method is evaluated on a set of images and is compared with other techniques in Section 3. Finally, Section 4 presents the conclusions of the study.

## 2. PROPOSED MITIGATION SCHEME

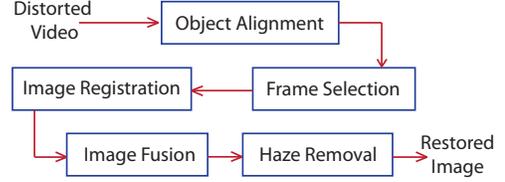
The proposed process is depicted in Fig. 1. Details of each step are described below.

**Object Alignment** Shooting video in the far distance may cause the ROI in each frame to become misaligned. The inter-frame distance between the shaking objects may be too large to apply image registration. Unfortunately matching algorithms which use feature detection are not suitable for the air turbulence problem as strong gradients within each frame are distorted randomly. Here a simple approach using morphological image processing is proposed. The ROI is marked in the first frame. The histogram, generated from the selected ROI and the surrounding area, is utilised to find an Otsu threshold [8] 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 image are removed. This step is done iteratively until the area near the ROI is isolated. The same Otsu threshold and the number of iterations are employed in other frames. The centre position of each mask is 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 centre of the mask in each frame is utilised to shift the ROI to align across the set of frames. Note that the frames with the incorrectly detected ROIs will be removed in the frame selection process. These frames are generally significantly different from others.

**Frame Selection** In our proposed method, not every frame in the sequence is used to restore the undistorted image since the bad images (e.g. the very blurred ones) would possibly deteriorate the fused result. A set of images are carefully selected using three factors; sharpness  $G_n$ , intensity similarity  $S_n$  and detected ROI size  $A_n$ .  $G_n$  can be computed from a summation of intensity gradients or the magnitude of high-pass coefficients. For  $S_n$ , the average frame of the whole sequence is used as a reference for calculating the mean square error (MSE). Then  $MSE^{-1}$  represents the similarity of each frame.  $A_n$  is the total number of pixels contained in the ROI. This is used because larger ROIs contain more details. The cost function  $C_n$  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  $\lambda_k$  are the weight and slope control of the factor  $k \in \{G, S, A\}$ , respectively. The sigmoid function is used here to prevent one factor dominating the others, e.g. a blocking artefact may cause significantly high values of sharpness, yet this frame should not be included in the selected data set.



**Fig. 1.** Block diagram of the proposed method

The  $\lambda_k$  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. The Otsu method can be applied to find how many frames should be included in the selected set.

**Image Registration** Registration of non-rigid bodies using the phase-shift properties of the DT-CWT proposed in [9] is employed in this paper. The algorithm is developed from the ideas of 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.

**Image Fusion** We have adapted the region-based image fusion technique using complex wavelets proposed in [10] to address the air-turbulence problem. The method first transforms each image into the DT-CWT domain. Then it employs an adapted version of the combined morphological spectral unsupervised image segmentation and a multiscale watershed segmentation [11] to divide the image into  $R$  regions. The lowpass DT-CWT coefficients of the fused image are simply constructed from the average of the lowpass values of all images, while the highpass coefficients are selected according to an activity measurement indicating the importance of that region. In this paper, to produce sharper results, we operate on each sub-band separately. 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  $\theta$  as shown in Eq. 3. The fused image is consequently constructed based on this priority map.

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

The air-turbulence scenario is different from other image-fusion problems as the segmentation boundaries which separate inhomogeneous regions vary significantly from frame to frame (due to turbulence distortion). The aim is to provide the sharpest and most temporally consistent boundaries. Therefore, at the boundary of each region, we use the maximum of DT-CWT coefficient magnitudes from all frames instead of using  $P(r_n^\theta)$ . To reduce the distortion due to rippling, the phase of the complex wavelet coefficients play an important role since it corresponds to the type of dominant directional features in its support regions. Hence, the DT-CWT



**Fig. 2.** Original image (left) and its contrast enhancement with global HE (middle) and CLAHE (right)

coefficients,  $d^{\theta,l}(x, y)$ , 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). The average phase can be used because, in turbulent motion, pixels deviate from their actual positions with approximately zero mean and with a quasi-periodic motion [4].

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

To reduce noise, we apply a shrinkage function,  $A_s$ , derived as Maximum A Posteriori (MAP) estimators as in [12]. Then, if sharpening is required, a gain  $A_g > 1$  can be applied to boost highpass-coefficient magnitudes. However this simple technique might cause remaining noise to become clearly visible again. Therefore we recommend to enhance only the high-pass coefficients when they form regions of large magnitude. We create the binary map  $M^{\theta,l}$  for each subband. The  $M^{\theta,l} = 1$  if  $|d^{\theta,l}| > \tau$ , where  $\tau$  is a predefined threshold. The isolated pixels are subsequently removed from  $M^{\theta,l}$ . The modified highpass coefficients are finally rewritten as Eq. 5.

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

**Haze Removal** A problem that often accompanies turbulence is haze or fog. Simple and fast methods such as histogram equalisation (HE) can be used to reduce this effect. However, in this paper, we consider the ROI and the meaning it carries. In this case, a contrast limited adaptive histogram equalisation (CLAHE) is more suitable than global adjustment. The method calculates local histograms derived from a neighbourhood region. A threshold is predefined to limit the cumulative values thereby preventing overamplifying small amounts of noise in largely homogeneous regions. An example of haze removal with HE and CLAHE is shown in Fig. 2. The number at the top of the car is much clearer in the CLAHE result.

### 3. RESULTS AND DISCUSSIONS

**Datasets** Four sequences captured from long distance exhibiting significant turbulence distortions are used in our tests. Two datasets, *Hill House* and *Hot Road*, are greyscale only,

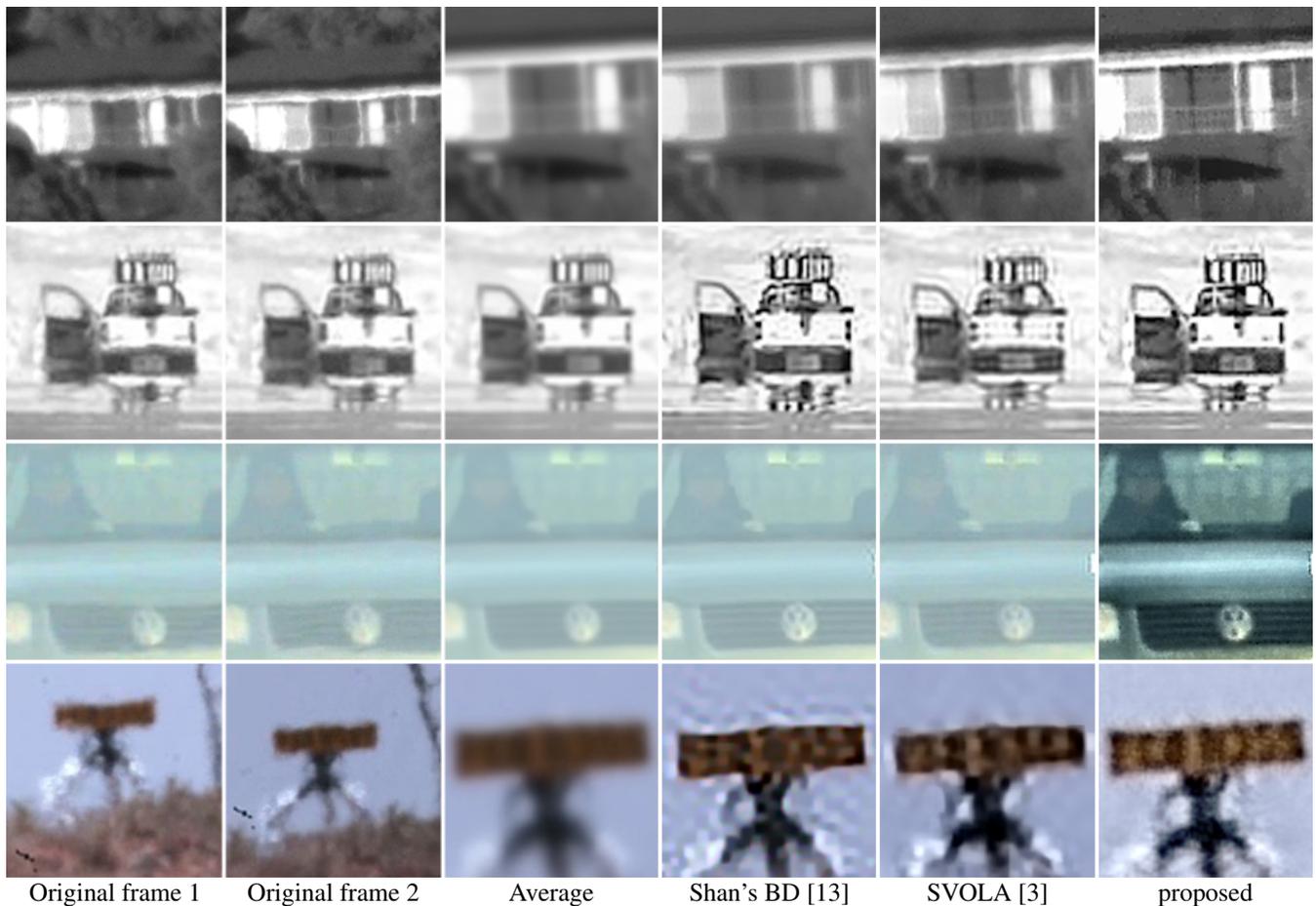
while the other two, *Cold Car* and *Number Plate*, are colour sequences. The colour sequences are converted into greyscale and colour channels, e.g. YCbCr or HSV. Only the greyscale channel is used in our method. The results are subsequently combined with colour channels and converted back to the original colour space. Only the *Number Plate* sequence shows a significant shift of the ROI area between frames, so we have artificially applied spatial shifts of between 1-20 pixels randomly to *Hill House*, *Hot Road* and *Cold Car*.

**Parameter Setting** The DT-CWT is applied with 3 or 4 decomposition levels when the image resolution is smaller or larger than  $128 \times 128$  pixels, respectively. The weights for the frame selection  $w_G$ ,  $w_S$  and  $w_A$  are set to 1. The value  $\tau$  for cleaning the map  $M^{\theta,l}$  is individually set for each subband in each level to  $\tau^{\theta,l} = \overline{|d|}^{\theta,l} - \sigma^{\theta,l}$ , where  $\overline{|d|}^{\theta,l}$  and  $\sigma^{\theta,l}$  are a mean and a standard deviation of the magnitude of highpass subband. The gains  $A_g^{\theta,l}$  are equal for each subband at the same level and are 2.5, 1.8, 1.2 and 1 for  $l = 1, 2, 3$  and 4, respectively. For CLAHE, a window of  $8 \times 8$  pixels is used to compute local histograms which are clipped at 1%.

**Experiment and Discussions** The proposed method is compared with i) an average of registered images, ii) Shan's BD [13] applied to (i), iii) SVOLA by Hirsch *et al.* [3]. To make a fair comparison, these last two methods are applied to the registered images constructed from the object alignment and frame selection. Two original frames which show the extent of distortion and the associated results are shown in Fig. 3 column 1-2 and 3-6, respectively. Although ground truth images are not available, the subjective results clearly show that the proposed method removes atmospheric distortion more efficiently than the other approaches. Shan's blind deconvolution is inefficient for air turbulence since the PSF is assumed to be similar for the entire image, while our method runs a set of homogeneous regions separately. Shan's method also takes four times longer to process than our proposed method, mainly to estimate the PSF. SVOLA subdivides an image into overlapped regions and estimates the PSF separately; as a result, it provides better results compared to Shan's method. However the computational time is even longer and the results are not as sharp as the proposed method. In addition, prior knowledge of PSF size is required for both previous methods.

### 4. CONCLUSIONS

This paper has introduced a new method for mitigating atmospheric distortion in long range surveillance imaging. The improvement of visibility of an ROI in the image sequence is achieved using region-based fusion in the DT-CWT domain. We also propose a simple object alignment method and a new cost function for frame selection to pre-process the distorted sequence. The process is completed with local contrast enhancement to remove haze interference. Experiments with real data show promising results and superior performance



**Fig. 3.** Original images, Average image, and results of Shan's BD [13], SVOLA [3] and proposed method

compared with the existing methods.

## 5. ACKNOWLEDGEMENTS

The authors acknowledge the support of General Dynamics United Kingdom Limited.

## 6. REFERENCES

- [1] Edmund Y. Lam and Joseph W. Goodman, "Iterative statistical approach to blind image deconvolution," *J. Opt. Soc. Am. A*, vol. 17, no. 7, pp. 1177–1184, Jul 2000.
- [2] S. Harmeling, M. Hirsch, S. Sra, and B. Scholkopf, "Online blind image deconvolution for astronomy," in *Proc of IEEE Conf Comp. Photogr.*, 2009.
- [3] M. Hirsch, S. Sra, B. Scholkopf, and S. Harmeling, "Efficient filter flow for space-variant multiframe blind deconvolution," in *Computer Vision and Pattern Recognition (CVPR)*, June 2010, pp. 607–614.
- [4] Dalong Li, "Suppressing atmospheric turbulent motion in video through trajectory smoothing," *Signal Processing*, vol. 89, no. 4, pp. 649–655, 2009.
- [5] P. J. Kent, S. B. Foulkes, J. G. Burnett, S. C. Woods, and A. J. Turner, "Progress towards a real-time active lucky imaging system," in *EMRC DTC Technical Conference*, 2010.
- [6] C. S. Huebner and C. Scheifling, "Software-based mitigation of image degradation due to atmospheric turbulence," in *SPIE conference on Optics in Atmospheric Propagation and Adaptive Systems*, 2010.
- [7] 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. 2005.
- [8] 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.
- [9] 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.
- [10] J. J. Lewis, R. J. O'Callaghan, S. G. Nikolov, D. R. Bull, C. N. Canagarajah, and Essa Basa'ed, "Region-based image fusion using complex wavelets," in *Proc. 7th Inter. Conf. Info. Fusion*, 2004, pp. 555–562.
- [11] 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.
- [12] Artur Loza, David Bull, Nishan Canagarajah, and Alin Achim, "Non-gaussian model-based fusion of noisy images in the wavelet domain," *Comput. Vis. Image Underst.*, vol. 114, pp. 54–65, January 2010.
- [13] Qi Shan, Jiaya Jia, and Aseem Agarwala, "High-quality motion deblurring from a single image," *ACM Transactions on Graphics (SIGGRAPH)*, 2008.