

# Restoration of heat haze in image and video based on DT-CWT image fusion

Özlem Alpergün Tanas<sup>1[0000-0002-8760-3650]</sup>

Ender M. Eksioğlu

**Abstract.** There are some distortion effects on atmosphere that decrease visual quality like out of focused. It is called heat haze, atmospheric turbulence, heat mirage or scintillation. It causes by such reasons that temperature difference, heterogeneous refractive index, or flow of hot air fluctuation. In this work, model-based removal method is used in terms of deep-learning based in MATLAB environment. Some removal steps is applied. DT-CWT based wavelet transform is used for image fusion step. The DT-CWT has the advantage of directional selectivity and shift-invariance properties in accordance with other wavelet transforms. As the input of the removal algorithm, a couple of frames are used to get one cleared image. Then a video is obtained by shifting successive input frames which is the contribution to this work. Camera ready-made algorithm and our removal algorithm for the inputs that taken from different distances is compared by using HVS (human visual system), FR (Full Reference) and NR (No Reference) methods. The inputs is taken from the original heat hazy atmospheric condition from field.

**Keywords:** Heat Haze, Image Fusion, Dual Tree Complex Wavelet Transform.

## 1 Introduction

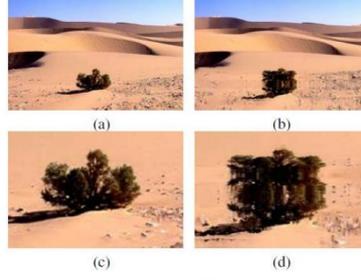
### 1.1 Reason of Heat Haze

Fog, haze or atmospheric turbulence due to temperature variations are some distortion types. The atmospheric distortion can influence the visual quality of video signals during acquisition.

When the temperature difference between the ground and the air increases, the thickness of horizontal layer in the air decreases and the layers move upwards rapidly, leading to faster and greater micro-scale changes in the air's refractive index [1]. So get some unfocused image results like in Fig. 1.

The mathematical expression can be seen from the equation (1). The hazy and un-hazy image vectors, are  $I_{hazy}$  and  $I_{unhazy}$  respectively.  $D$  is a matrix which includes unfocusing elements, and  $e$  is noise.

$$I_{hazy} = DI_{unhazy} + e \quad (1)$$

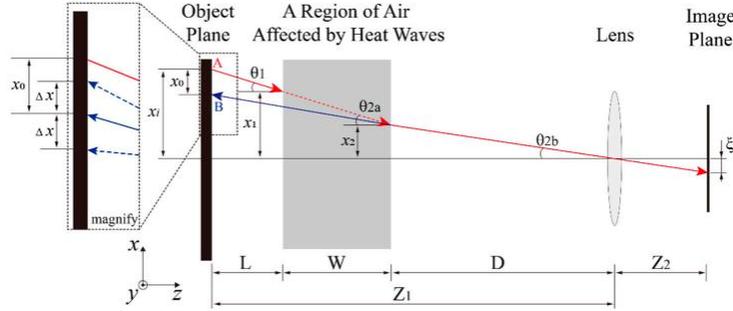


**Fig. 1.** Heat haze effect on desert (a) Non-effected scene; (b) Effected scene; (c) Zoom-in view of (a); (d) Zoom-in view of (b).

**Geometrical visualization of distortions.** The geometrical visualization is located in Fig. 2. According to this Fig. 2, the light emitted from the point A on the object plane appears at the point B in the perspective of the image plane. The difference between A and B, means the distortion  $x_0$  can be expressed as in the equation (2).

$$x_0 = [\tan(\theta_1)(L + W)] - [\tan(\theta_{2b})(L + W)] = (L + W) \frac{1}{n_0} \int_{-\frac{W}{2}}^{\frac{W}{2}} \left( \frac{\partial n'}{\partial x} \right) \partial z \quad (2)$$

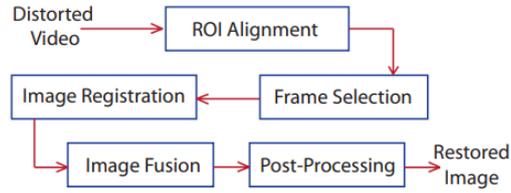
According to equation (2), if the air region affected by heat waves is stable and does not flow, and only the temperature is different from the other region, then  $n_0$  and  $W$  are constant. So the distortion  $x_0$  is also a constant. But if there is a fluidity of the heat waves, the B point will randomly oscillate around  $x_0$ , and based on the main distortion  $x_0$ , a random distortion  $\Delta x$  occurs [2].



**Fig. 2.** Geometrical visualization of heat haze

## 2 Heat Haze Removal Steps

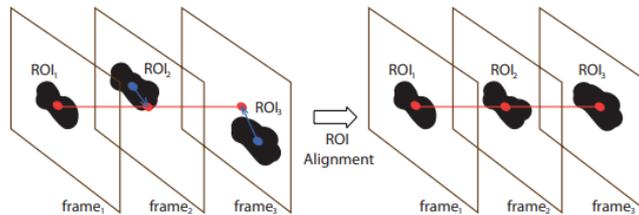
There is a block diagram of removal steps respectively in Fig. 3. There is a misalignment of the ROI (Region of Interest) of the capturing video frames under the effect of heat haze. The distorted video sequence split into its frames to alignment process. Most similar and sequenced video frames are used to generated one restored image.



**Fig. 3.** Heat haze removal steps [1]

### 2.1 ROI Alignment

The detail visualization of ROI alignment process is located in Fig. 4. An erosion process is applied every frames and the areas which is closed to ROI are cleaned. This step is performed iteratively until for isolation. Then all the ROI are aligned by shifting towards the center of the frames.



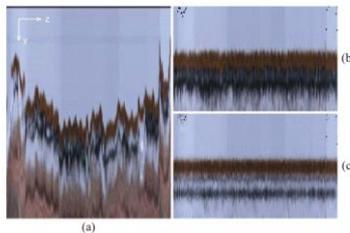
**Fig. 4.** ROI alignment process [1]

### 2.2 Frame Selection

In the frame selection step, a subset of images are carefully selected using these factors: sharpness, intensity similarity and detected ROI size.

### 2.3 Image Registration

In order to clean the unwanted signals, the non-rigid sequence of image is registered in image registration step like in Fig. 5.



**Fig. 5.** yz planes at same x point of each frame. (a) Distorted sequence (b) Aligned sequence corresponding (c) Registered sequence [1]

## 2.4 Image Fusion

Image fusion is a combining of two images into a single image that has the maximum information content like in Fig. 6. DT-CWT (Dual Tree Wavelet Transform) methods is used for its advantage over the other wavelet transform methods in terms of the shift invariance and directional selectivity specifications.



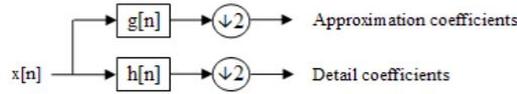
**Fig. 6.** Image fusion example from two input image

**Wavelet Transforms.** Wavelet transformation advantage over other time and frequency transformation is good in time resolution of high frequencies. CoWT, DWT, DT-CWT transforms will be examined in this section to see the difference of them.

*CoWT(Continuos Wavelet Transform).* CoWT represents a signal by rewriting the scale and translation coefficients, which is represented a and b respectively in equation (3), of the wavelets vary continuously.

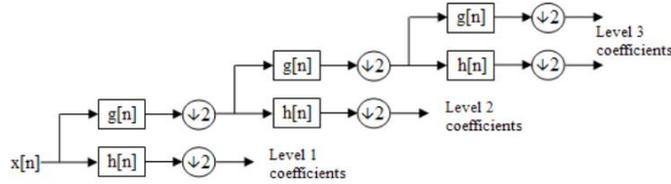
$$X_w(a, b) = \frac{1}{|a|^{\frac{1}{2}}} \int_{-\infty}^{\infty} x(t) \overline{\Psi\left(\frac{t-b}{a}\right)} dt \quad (3)$$

*DWT (Discrete Wavelet Transform).* A DWT is any wavelet transform for which the wavelets are discretely sampled. It captures both frequency and location information (location in time). The samples are passed through a low-pass and high-pass filters and then decimated by a factor of 2 to get approximation and detail coefficient respectively like in Fig. 7.



**Fig. 7.** Approximation and detail coefficients in DWT.

When the approximation coefficients separated with high and low frequency contents and then subsampled, it is visualized as a binary tree and is called as a filter bank. This is the 3-level filter bank in Fig.8.

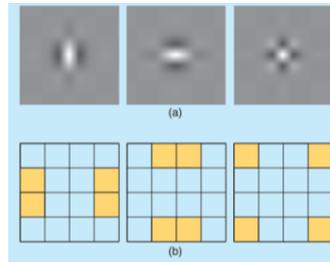


**Fig. 8.** 3-Level filter bank

*2D DWT.* The implementation for the 2-D DWT is consist of these 3 wavelets as defined in equation 4. The wavelet of  $\psi_1(x, y)$  described as a product of the the first dimension of low-pass  $\phi(\cdot)$  and the second dimension of the high-pass (or bandpass) function  $\psi(\cdot)$

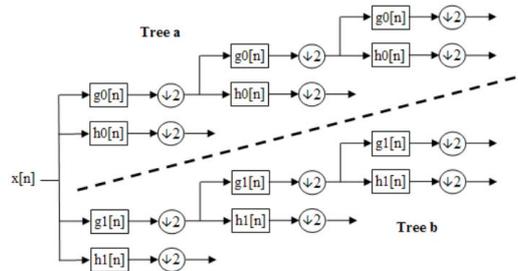
$$\begin{aligned}\psi_1(x, y) &= \phi(x)\psi(y) && \text{(LH wavelet)} \\ \psi_2(x, y) &= \psi(x)\phi(y) && \text{(HL wavelet)} \\ \psi_3(x, y) &= \psi(x)\psi(y) && \text{(HH wavelet)}\end{aligned}\quad (4)$$

There is a space domain visualization for 2D-DWT in Fig. 9.



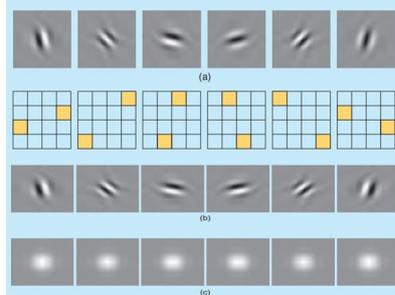
**Fig. 9.** (a) Space domain visualization of the wavelets for LH, HL, HH; (b) 2-D frequency domain visualization of each wavelet.

*DT-CWT (Dual Tree Complex Wavelet Transform).* The DWT arises from real wavelets. However, the DT-CWT is a complex-valued sinusoid signals. A 3-level DT-CWT can be seen in Fig. 10. Each tree of filters are totally real, but for two trees itself they have both a real and imaginary filter for every complex wavelets.



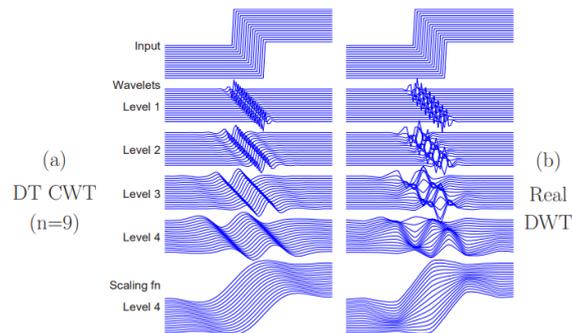
**Fig. 10.** Level 3 DT-CWT.

*2D DT-CWT.* DT-CWT produces oriented wavelets by using 2D-discrete wavelets. The real and imaginary components of every complex wavelet are oriented at identical angles, while the magnitude of their forms are a circular shape like in Fig. 11.



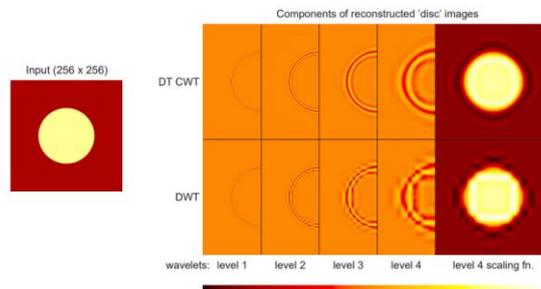
**Fig. 11.** 2D DT-CWT (a) real wavelets; (b) imaginary wavelets; and (c) magnitudes

There is 6-directionally selective filters for real and imaginary parts of DT-CWT, while there was only 3 wavelets in DWT. So, it shows the directionally selection properties of DT-CWT. The shift invariance properties of DT-CWT can be seen from the Fig. 12. For a real wavelet coefficient oscillation pattern there is a distortion by a small shift of the signal around singularities, while for DT-CWT signals, components at each level seems to be similar in shape means shift invariant.



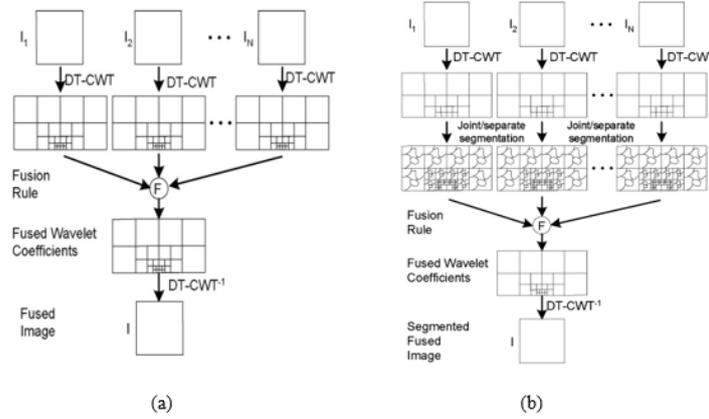
**Fig. 12.** Shift invariance example

There is also shift invariance example for 2D domain in Fig. 13. In DT-CWT, the edge of the disc seems so smooth and continuous means in a good shift invariance. Also, there is six directional sub bands of coefficients have been used as seen from the last level in DT-CWT.



**Fig. 13.** Shift invariance example in 2D domain

**Fusion Method.** There is pixel-based and region-based fusion methods. Pixel-based, combines individual pixels; while the region-based fusion, groups image pixels to form a continuous region. The pixel-based and region-based fusion diagram, located in Fig. 14.



**Fig. 14.** Block diagram of image fusion methods (a)Pixel-based (b) Region-based

The DT-CWT which is expressed as  $\omega$ , is applied to the input images of  $I_1, I_2, \dots, I_N$  as in the equation (5).  $N$  numbered of detail coefficients  $D_n$  and approximation of the frame  $A_n$  are obtained for each of the  $I_n$  image frames. At each level of separation, they use having six selective sub bands properties.

$$[D_n, A_n] = \omega(I_n) \quad (5)$$

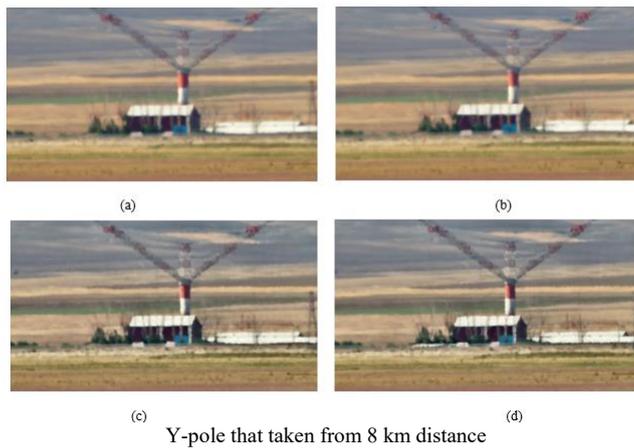
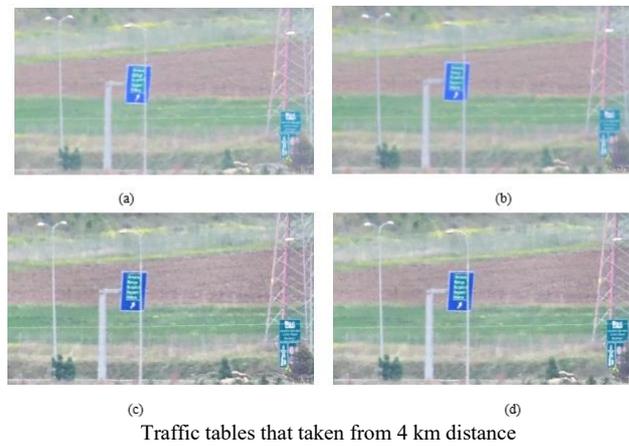
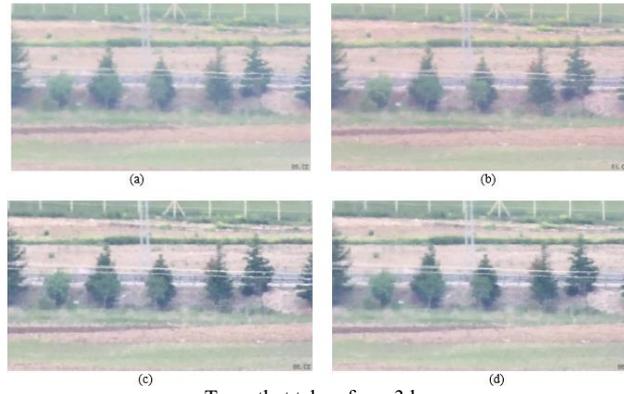
Regions are chosen or eliminated according to the priorities, and it is obtained the wavelets of the reconstructed image at the end of fusion method F. The fusion method  $F$ , generally select the highest priority regions or use a average of the regions to obtain the wavelets of the reconstructed image.

Subsequently, the reconstructed image is derived by applying the inverse transform, denoted as  $\omega^{-1}$ , to the weighted wavelet coefficients obtained from equation (6).

$$F = \omega^{-1}(D_F, A_F) \quad (6)$$

### 3 Results

There is comparison results for camera's ready algorithm and for DT-CWT fusions that include both pixel-based and region-based scenarios. In comparison FR and NR metrics are used.



**Fig. 15.** (a) Original image (with hazy) (b) Restored image with camera algorithm (c) Restored image with pixel-based DT-CWT (d) Restored image with region-based DT-CWT.

As for the FR metrics, GSM, GMSD, SSIM, and MS-SSIM were calculated using reference and restored raw video sequences. The raw video data was applied to the traffic table scene, yielding numerical results found in Table 1. According to the overall findings, the proposed algorithm demonstrates superior performance, particularly in the region-based approach. This is because higher scores across these metrics generally indicate better video quality, with the exception of the GMSD score, which presents an opposing interpretation.

**Table 1.** Quality scores with FR methods for the raw videos of traffic table scene.

FR methods	Camera's Algorithm	Pixel Based DT-CWT	Region Based DT-CWT
GSM	0.8397	0.8831	0.9302
GMSD	0.6140	0.5691	0.4856
SSIM	0.7546	0.8034	0.8792
MSSIM	0.7552	0.7845	0.8817

Table 2 presents the JPEG score as a non-reference (NR) quality metric for both original and restored image frames. In terms of the overall findings, the score for the restored images surpasses that of the original hazy images. When comparing different algorithms, once again, the proposed algorithm emerges as superior, with the region-based approach performing slightly better than the pixel-based one. Analyzing the comparison based on different image scenes reveals that as distances increase, the rate of numeric result improvement diminishes. This trend is particularly notable from the Y-pole scene at an 8 km distance, where the algorithm results closely resemble those of the hazy input frame due to the heightened impact of haze.

**Table 2.** JPEG quality scores for original and restored images.

Images	Original (Heat hazy)	Restored		
		Camera's Algorithm	Pixel based DT-CWT	Region based DT-CWT
Trees	19.1831	22.8547	40.4670	41.3148
Traffic table	16.9042	17.8745	38.1450	39.4282
Y-Pole	26.3428	27.5769	28.3497	29.7787

## 4 Acknowledgement and future works

This work contributes to the literature by showcasing the generation of multiple output frames through the sequential shifting of frames from various inputs. This approach enables continuous processing and the creation of a video composed of multiple output frames. However, due to the mathematical operations involved in this wavelet-based algorithm, the processing time is lengthy. As a future endeavor, there is a plan to transition the software developed in the MATLAB environment to Python or C++ programming languages and enhance its speed using the CUDA library. The ultimate aim is to achieve real-time restoration of live video frames.

## References

1. Anantrasirichai, N., Achim, A., Kingsbury N. G., and Bull, D. R: Atmospheric Turbulence Mitigation Using Complex Wavelet-Based Fusion, in *IEEE Transactions on Image Processing*, 22, 6, 2398-2408, June, doi: 10.1109/TIP.2013.2249078 (2013).
2. Ma, C., Zeng, Z., Zhang, H., and Rui, X.: A Correction Method for Heat Wave Distortion in Digital Image Correlation Measurements Based on Background-Oriented Schlieren. *Applied Sciences*. 9. 3851. doi: 10.3390/app9183851 (2019).
3. Hill, P., Canagarajah N., and Bull, D: Image Fusion using Complex Wavelets, *BMVC*, The University of Bristol, Sep., doi: <https://doi.org/10.5244/c.16.47> (2002).
4. Kingsbury N.G.: Complex wavelets for shift invariant analysis and filtering of signals, *Journal of Applied and Computational Harmonic Analysis*, 10, 3 , 234–253, May, ISSN 1063-5203 (2001).
5. Kingsbury, N.: Dual tree complex wavelets, in *Signal Processing Group*, Dept. of Engineering University of Cambridge, September (2004).
6. Lewis, J. J. , OCallaghan, R. J., Nikolov, S. G., Bull, D. R., and Canagarajah, N: Pixel- and region-based image fusion with complex wavelets, *Information Fusion*, 8, 2, 119–130, Apr., doi: <https://doi.org/10.1016/j.inffus.2005.09.006> (2007).
7. Vijaya durga, R., Kumari, O., Surya prakash M., Dileep kumar, P., and Tirupathi, Y.: Region-Based Image Fusion Using Complex Wavelets., *IOSR journal of electronics and communication engineering*, 9, 1, 23–26, Jan., doi: <https://doi.org/10.9790/2834-09142326> (2014).
8. Pajares, G., and Cruz, J.M.: A wavelet-based image fusion tutorial, *Pattern Recognition*, 37, 9, 1855–1872, Sep., doi: <https://doi.org/10.1016/j.patcog.2004.03.010> (2004).
9. Selesnick, I. W., Baraniuk, R. G. and Kingsbury, N. C.: The dual-tree complex wavelet transform, *IEEE Signal Processing Magazine*, 22, 6, 123–151, Nov., doi: <https://doi.org/10.1109/msp.2005.1550194> (2005).
10. Min, X., Zhai, G., Zhou, J., Farias, M. C. Q. and Bovik, A. C.: Study of Subjective and Objective Quality Assessment of Audio-Visual Signals, in *IEEE Transactions on Image Processing*, 29, 6054-6068, doi: 10.1109/TIP.2020.2988148 (2020).
11. Wang, Z., Sheikh, H. R., and Bovik A. C.: No-reference perceptual quality assessment of JPEG compressed images. in *Proc. IEEE ICIP*, 1, 477–480 (2002).
12. Wavelet transform, [https://en.wikipedia.org/wiki/Wavelet\\_transform](https://en.wikipedia.org/wiki/Wavelet_transform), last accessed 06/11/23