# Processing and Fusion of Thermal and Video Sequences for Terrestrial Long Range Observation Systems


**Leonid P. Yaroslavsky** yaro@eng.tau.ac.il

**Barak Fishbain** fishbain@post.tau.ac.il

**Alex Shteinman** shtainma@post.tau.ac.il

**Shai Gepshtein** shai_g@eng.tau.ac.il

Department of Interdisciplinary Studies
The Iby and Aladar Fleischman Faculty of Engineering,
Tel Aviv University
Ramat Aviv 69978
ISRAEL



**Abstract –** *A two stage fusion image processing system has been developed for visual and thermal range video streams. At the first stage, an intra-channel-inter-frame fusion is applied to each channel images independently, addressing each channel specific limitations. Noise removing processing by means of 3D (spatial-temporal) local DCT adaptive filtering is applied to the thermal channel video sequence. An algorithm that compensates image distortion due to the atmospheric turbulence is applied to the visual range channel video sequence. The second stage is inter-frame intra-channel fusion. At this stage, thermal and visual range channel image frames are fused frame-by-frame using a weighted average scheme with locally adapted weights. Experimental results obtained for both simulated and real life thermal and visual range image sequences are demonstrated*



**Keywords:** Fusion, 3-D DCT, Elastic Registration Method, Sinc-interpolation, Optical Flow, Wiener Filtering, Weighted Average, Noise Estimation.


## 1 Introduction

The development of new imaging sensors has brought with it a need for image processing techniques that can effectively exploit data obtained from different sensors or for multiple images produced by the same sensor. This can be achieved by combining inputs from different sources to obtain a single composite frame with extended or enhanced information content. The goals of such fusion process are to extract all the useful information without artifacts while maintaining reliable and robust system with temporal stability and consistency. A number of works can be found in the literature regarding image fusion [1-9]. Most of the works treat the fusion problem as an "Intra-Frame – Inter-Channel" fusion, hence using single pair of frames, usually from two different sensors, for the fusion process. This paper addresses the problem of temporal and spatial fusion of thermal and video range image sequences for terrestrial Long Range Observation Systems (LOROS). A two-stage fusion algorithm is suggested and implemented. At the first stage, an inter-frame (temporal) intra-channel fusion for noise reduction and atmosphere turbulence distortions compensation is carried out. At the second, inter-channel intra-frame fusion stage, the two channels are combined to generate a final single channel video.

Several classes for digital image improvement and enhancement methods can be found in the literature. One class of methods uses the spatial (intra-frame) redundancy in images for image deblurring and denoising. Among these methods one can mention Wiener and empirical Wiener linear filtering for image denoising and aperture correction [10, 11], local adaptive linear filtering in transform, specifically, in DCT domain [12, 13], and non-linear rank filtering [14, 15]. Another group of methods is intended for processing video sequences and uses the temporal redundancy in an image sequence. In this group one can find moving average integration [10, 16], motions estimation and AC/DC methods.

Observation systems, both thermal and video, produce a stream that is affected by imaging system (detector and electronics) and environmental noise. While combining a signal to a video sequence this noise can be described using 3-D noise models [10, 17, 18]. Therefore if estimation of noise and signal restoration is made in 2-D or 1-D domains, it remains suboptimal and can produce substantial artifacts. In particular,

- Intra-frame (2-D) processing of video sequences results in flickering artifacts when observing frame by frame processed video.
- Inter-frame only processing leaves artifacts due to image blur and fix pattern noise unprocessed. Efficient temporal noise reduction needs a long frame averaging [10] which may affect scanning scenes and moving objects.

This motivates development of 3-D spatial/temporal processing of video sequences. In this paper, we demonstrate two such methods: 3-D local adaptive filtering in DCT domain for image denoising and aperture correction, and elastic (local) and selective image

registration for compensating atmosphere turbulence on steady scenes.

The Thermal Range and the Visual Range Channels have different behavior and feature different image distortions. Visual Range long distant near-earth observations are usually heavily affected by atmospheric turbulence. This causes spatially and temporally random fluctuations in the index of refraction of the atmosphere [19]. Thermal sensors, on the other hand, do not exhibit atmosphere fluctuation sensitivity but quite often suffer from substantial additive noise due to their working conditions and to some other physical characteristics [20]. In this paper, we present two different temporal fusion techniques applied on each channel according to its special characteristics, while both channels are then fused to produce improved results.

The paper is organized as follows. Section 2 discusses intra-channel temporal fusion for thermal channel using 3D de-noising and aperture correction filtering method. Sect. 3 deals with temporal fusion technique for compensating image blur due to atmosphere turbulence. Inter-channel intra-frame data fusion for visual range and thermal images is discussed in Sect. 4. Finally, Sect. 5 combines intra-frame and inter-channel processing to produce a unified fusion mechanism for visual range and thermal video sequences.

## 2 Thermal Range Channel Temporal Fusion Method

Thermal images usually suffer from substantial additive noise and sensor's aperture distortions. When time sequences of thermal images are available, image temporal redundancy offers an additional option for denoising and deblurring of images of still scenes. In [11, 21, 22], a sliding window transform domain 2-D filtering for still image restoration is described. In [13], application of this method for denoising and enhancement of color still images is shown. In this paper, we extend this method for 3-D spatial/temporal denoising of thermal image sequence processing. In Sect. 2.1 we provide a brief outline of the method and in Sect. 2.2 describe its application to processing thermal image sequences and present illustrative examples of processing test and real life images.

### 2.1 Sliding window DCT domain filtering

Let $\mathbf{b}$ be a vector of samples of distorted image to be processed, $\mathbf{a}$ be a vector of samples of undistorted image that is the goal of the processing and let $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ are their corresponding spectra in the domain of an orthogonal transform $\mathbf{T}$:

$$\boldsymbol{\beta} = \mathbf{Tb}\ ;\ \boldsymbol{\alpha} = \mathbf{Ta}\ . \qquad (1)$$

Introduce a scalar filter described by vector $\boldsymbol{\eta}$ of its spectral coefficients

$$\hat{\boldsymbol{\alpha}} = \boldsymbol{\eta} \bullet \boldsymbol{\beta}\ , \qquad (2)$$

where $\hat{\boldsymbol{\alpha}}$ is vector of the filtered image spectral coefficients and $(\bullet)$ denotes element-wise (Hadamard) product of vectors. It is assumed in this formulation that the filtered image $\hat{\mathbf{a}}$ is obtained from its spectrum $\hat{\boldsymbol{\alpha}}$ through transform $\mathbf{T}^{-1}$ inverse to $\mathbf{T}$:

$$\hat{\mathbf{a}} = \mathbf{T}^{-1}\hat{\boldsymbol{\alpha}} \qquad (3)$$

Applying the Mean Square Error (MSE) criterion for evaluation of deviation of the processed image $\hat{\mathbf{a}}$ from the target image $\mathbf{a}$ one can obtain that the optimal filter coefficients $\boldsymbol{\eta}_{opt}$ that minimize MSE are defined as:

$$\boldsymbol{\eta}_{opt} = \mathbf{AV}\left\{\boldsymbol{\alpha} \bullet \boldsymbol{\beta}^{*}\right\} ./ \mathbf{AV}\left\{|\boldsymbol{\beta}|^{2}\right\} \qquad (4)$$

where $\mathbf{AV}\{.\}$ is the averaging operator used in evaluation of MSE. '$./$' is element-wise division of vector components and $|\boldsymbol{\beta}|^{2}$ is element-wise squared module of $\boldsymbol{\beta}$. For a signal distortion model formulated in the transform domain as

$$\boldsymbol{\beta} = \boldsymbol{\lambda} \bullet \alpha + \boldsymbol{\nu}\ , \qquad (5)$$

where $\boldsymbol{\lambda}$ is a vector of transform coefficients that describe image distortions in the imaging system and $\boldsymbol{\nu}$ is a vector that represent random zero mean signal independent system noise, Eq. (4) gives

$$\boldsymbol{\eta}_{opt} = (1 ./ \boldsymbol{\lambda}) \bullet \left(|\boldsymbol{\lambda}|^{2}\, \mathbf{AV}\left\{|\boldsymbol{\alpha}|^{2}\right\}\right) ./ \left(|\boldsymbol{\lambda}|^{2}\, \mathbf{AV}\left\{|\boldsymbol{\alpha}|^{2}\right\} + \mathbf{AV}\left\{|\boldsymbol{\eta}|^{2}\right\}\right) \qquad (6)$$

In image restoration, noise parameters can be extracted from image acquiring system design or can be estimated from observed noisy sequences [14, 21]. Eq. (6) represents a scalar Wiener filter for signal restoration. Its use requires a priory knowledge of signal and noise statistical spectra $\mathbf{AV}\left\{|\boldsymbol{\alpha}|^{2}\right\}$ and $\mathbf{AV}\left\{|\nu|^{2}\right\}$. If these spectra are evaluated from the observed distorted signals, empirical Wiener filter is used:

$$\boldsymbol{\eta} = \max\left\{0, \left(|\boldsymbol{\beta}|^{2} - \mathbf{Thr}\right) ./ |\boldsymbol{\beta}|^{2}\right\} \qquad (7)$$

were $\mathbf{Thr}$ is vector of empirical estimation of noise statistical spectral coefficients. This method of empirical Wiener filtering is sometimes called "soft thresholding". A simplified version of the Empirical Wiener filter is known as "rejecting" filtering or "hard thresholding".

$$\boldsymbol{\eta} = \frac{\left[\mathbf{sign}\left(|\boldsymbol{\beta}|^{2} - \mathbf{Thr}\right) + 1\right]}{2} \qquad (8)$$

In image processing, these filters can be applied globally or locally. In global processing, filters are designed and applied for entire image frames or set of frames. Such filtering is justified if images can be regarded as stationary spatial-wise and time-wise. In reality this obviously is not the case. Local filtering in a window of a finite size sliding both in spatial domain and time-wise is better suited non-stationary nature of images and enables efficient edge preserving image denoising ([10, 11, 13, 21, 22]).

As for the selection of the transform for the filter implementation, Discrete Cosine Transform (DCT) offers one of the best choices. It is very efficient both in terms of image energy compaction capability required for the efficient design of the empirical Wiener filter and in terms

of the computational complexity of image filtering ([11, 13, 22]).

3-D space-temporal sliding window filtering in DCT domain is used in this work for denoising and restoration of thermal image sequences.

## 2.2 Intra-Channel temporal fusion by means of 3D SWDCT filtering

Block diagram of the filtering is shown in Fig. 1. For each position of the cubic window, the DCT transform of the signal volume within the spatia-temporal cube is recursively computed from that of the previous position of the window. The signal spectra coefficients are then non-linearly modified according to Eqs. (7) or (8). The inverse transform need not be computed for all pixels within the cube, since only the central sample of the cube has to be determined in order to form the output signal.

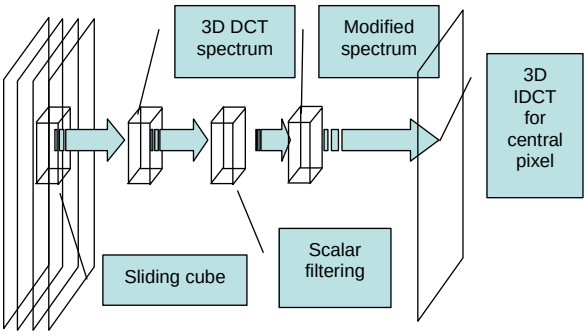



Fig. 1. Sliding cube 3D transform domain filtering

For testing the method, two sets of artificial test movie were generated of images of bars and text with different level of additive Gaussian noise. Examples of noise-less, noisy and denoised test images are shown in Figs. 2 and 3.

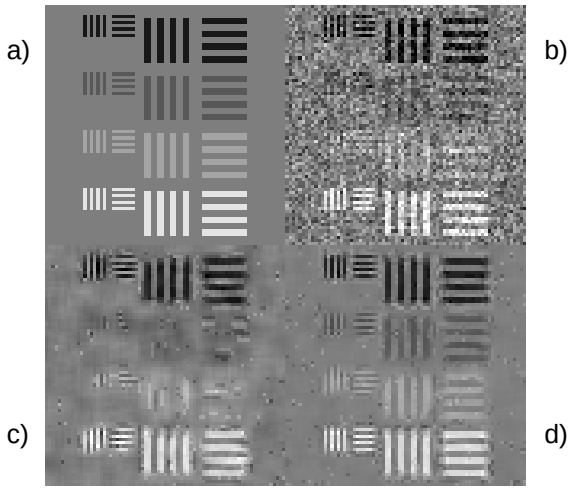



Fig. 2. 2-D and 3-D sliding window DCT domain image denoising. a) - initial test image; b) – noisy test image with additive zero mean Gaussian noise with standard deviation 30 (in image range 0-255); c) result of 2-D image denoising in the window of 5x5 pixels; d) result of 3-D image denoising in the cube of 5x5x5 pixels.

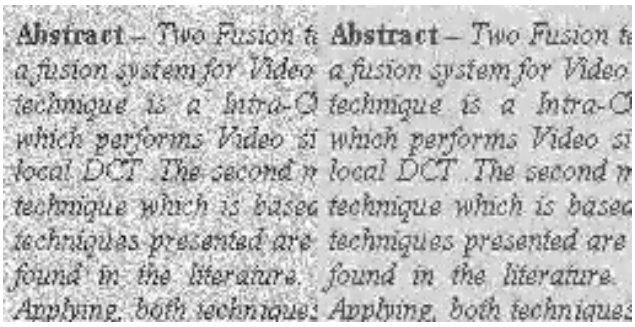



Fig. 3 Image of a text fragment corrupted by additive noise (a) and a result of denoising with 3-D sliding cube 5x5x5 filtering (b).

Corresponding demonstrative movies can be found on authors' web-site ([23]).

Numerical results on noise suppression capability of the filtering obtained for the test images are listed in Table 1. Using these data one can evaluate improvements achieved using 3-D filtering comparing to 2-D filtering of individual image frames.

Table 1. Standard deviation of residual filtering error (RMSE) for 4-bar and text test image sequences

| Block Size | RMSE (Noise STD=30) | | RMSE (Noise STD = 20) | |
|---|---|---|---|---|
| | 4-Bar | Text | 4-Bar | Text |
| 3x3x1 | 19.97 | 24.90 | 12.45 | 18.42 |
| 3x3x3 | 13.90 | 17.88 | 8.86 | 12.40 |
| 5x5x1 | 17.1 | 26.27 | 11.98 | 19.89 |
| 5x5x5 | 12.0 | 16.84 | 7.48 | 11.65 |

Figs. 4 and 5 illustrate noise suppression capability of 3-D sliding window filtering of real life thermal image sequences.

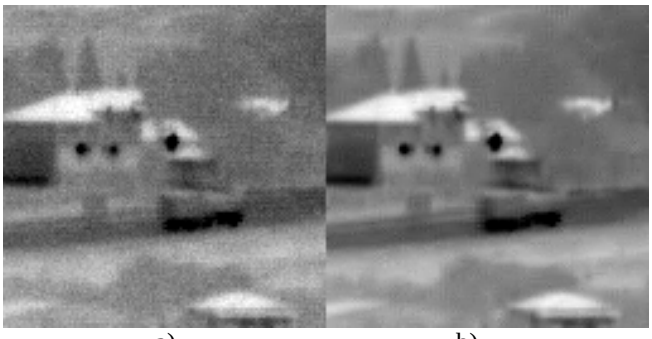
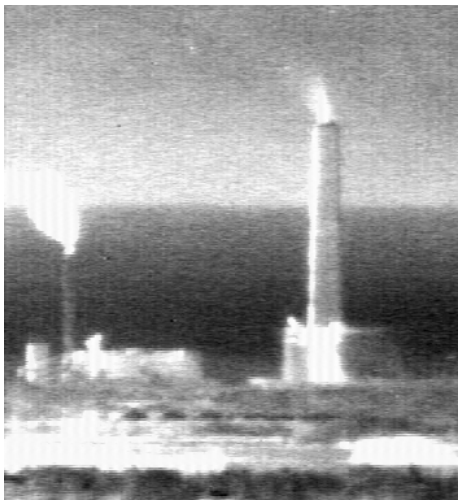



Fig. 4. 3-D sliding 5x5x5 cube denoising thermal video sequence: a) – a frame of original movie; b) – corresponding frame of the processed movie.



a)

b)

Fig. 5. Another example of 3-D sliding 5x5x5 window denoising thermal video sequence: a) a frame of original movie; b) – corresponding frame of the processed movie

## 3. Visual range temporal image fusion for compensating atmospheric turbulent Interferences

Atmospheric turbulence causes a random blur in images due to random fluctuations of the refraction index of the air through which the light propagates. These fluctuations in the optical path length of the propagating light result in phase and amplitude variations of the light's wave front. Unlike astronomical systems, where the entire frame can be modeled by the convolution of the object with a single, though random, point spread function, the long distant near-earth observation have wider field of view and are modeled by convolution with space variant and random point spread functions [24]. This causes small neighborhoods in the image to randomly move in different directions in different frames.

There is a variety of methods for perfecting of turbulent captured images [25-27]. In this work, a local image registration method is used to find the translation vector for each pixel in each frame of the video sequence with respect to a "reference" image derived from the sequence of image frames. These vectors are then used to "inverse-warp" image frames to their "stable" geometry. Using the same method, the areas of real motion of objects in the scene are also detected and the detection results are used for warping back only the static parts of images that do not contain moving objects. In this way a scene is restored where the only moving areas are the real moving objects.

### 3.1 Local image registration method

Assuming there is no real moving objects in the image, the mapping of one turbulent image to a stable "reference" image can be obtained by registration of certain small spatial neighborhood, surrounding each pixel in the image, to the reference image. In this way a field of motion vectors is received. We refer to this method as to local registration. A similar elastic registration method is also described [28]. In it's simplest form, the method assumes that it is sufficient to find only two translation parameters of the translation vector for every pixel.

Let $f(x,y,t)$ be a turbulent source image frame, $\bar{f}(x,y)$ be a target reference image, and $\Delta x$ and $\Delta y$ are translation parameters:

$$f(x,y,t) = \bar{f}(x+dx, y+dy) \tag{9}$$

The translation vector $\vec{\Delta} = [\Delta x, \Delta y]$ is evaluated for a small spatial neighborhood of every pixel through minimization of the mean square differences between the registered areas of the two images:

$$\vec{\Delta}_{\Omega} = \min \sum_{x,y\in\Omega} \left| f(x,y,t) - \bar{f}(x+\Delta x, y+\Delta y) \right|^2 \tag{10}$$

where $\Omega$ denotes a small spatial neighborhood of pixel $(x,y)$. A first-order truncated Taylor series expression of Eq. 10 is:

$$E(dx,dy) \approx \sum_{x,y\in\Omega} [f(x,y,t) - \bar{f}(x,y,t) + \Delta x \cdot f_x(x,y) + \Delta y \cdot f_y(x,y) - \Delta f_t(x,y)]^2 \tag{11}$$

where $f_x(\cdot)$, $f_y(\cdot)$ are the spatial derivatives of $f(\cdot)$, and $\Delta f_t(x,y)$ is temporal difference given by

$$\Delta f_t(x,y) = f(x,y,t) - \bar{f}(x,y). \tag{12}$$

The error function is than approximated as

$$E(\vec{\Delta}) \approx \sum_{x,y\in\Omega} \left| \Delta f_t - [f_x, f_y]^T \cdot [\Delta x, \Delta y] \right|^2 \tag{13}$$

The minimum of the error function corresponds to the values of $\vec{\Delta}$:

$$\vec{\Delta} = \left( \sum_{x,y\in\Omega} [f_x, f_y] \cdot [f_x, f_y]^T \right)^{-1} \cdot \left( \sum_{x,y\in\Omega} [f_x, f_y] \cdot f_t \right) \tag{14}$$

obtained by solving the equation:

$$\frac{\mathbf{dE}(\vec{\Delta})}{\mathbf{d}\vec{\Delta}} = \sum_{\mathbf{x},\mathbf{y}\in\Omega} -2 \cdot [\mathbf{f_x}, \mathbf{f_y}] \cdot \left( \mathbf{f_t} - [\mathbf{f_x}, \mathbf{f_y}]^{\mathbf{T}} \cdot [\mathbf{dx}, \mathbf{dy}] \right) = 0 \tag{15}$$

### 3.2 Generating "reference" image and turbulence compensation

For generating "reference" image for local registration, pixel-wise rank filtering is used [15]. The use of rank smoothing filters such as median and alpha-trimmed mean empirically have shown substantiated in two ways. First, light beam propagating through a turbulent atmosphere will deflect to any point within a certain radius, and the distribution of the deflection has a zero mean which means that the center of this area will be in the same location where the light beam would hit if there were no turbulence present. Therefore, statistically, a pixel real value (if there were no turbulence) would be very close to the mean of the array of the same pixel's values in a long period of time. For the other side, for moving objects that accommodate a pixel for a short period of time, the value of those pixels will be pushed to the tails of the gray level distribution in a long sequence. Therefore rank filter instead of mean filter is required to eliminate from averaging the distribution tales. It is important that the number of the images will be high enough to eliminate the moving objects.

Using a "stationary" scene, and having found the translation vector field, it is now possible to warp each pixel of the turbulent image to its "true" location where it would have been if there had not been any turbulence. This, in general, requires image re-sampling with sub-pixel shifts. We used a warp technique with discrete sinc-interpolation in a moving window in the DCT domain to obtain the value of the intermediate pixels with least mean square error [11, 29]. The resulting image is composed of the interpolated pixels of the turbulent image shifted into their true locations as if there was no turbulence.

For a better compensation, the result of the turbulent compensation can be computed iteratively. For every iteration, it is required to compute translation vector field using, as the source, the set of compensated images found in the previous iteration and the same reference image as the target. Then images obtained in the previous iteration are warped again by discrete sinc-interpolation using the new calculated vector field. It has been found in our experiments, that this process converges very rapidly so only two iterations are sufficient to obtain a near-optimal result.

In the presence, in the scene, of moving object it is necessary to distinguish between real motion and turbulent motion in the image sequence. To this goal, after alignment of turbulent frames to the reference image, the error function is computed again for every pixel in every frame. Obtained arrays of errors contain two types of errors: small errors due to the turbulence mis-compensations and large errors typical for areas in the image where real moving objects appear. Areas occupied by large errors can be easy detected and marked to form a mask for segmentation and extraction of moving objects from initial video sequence.

The proposed algorithm was tested on an artificial video sequence prepared by simulation and on real captured turbulent videos. Sequences consisted of 128 images of turbulent scenes containing moving objects.

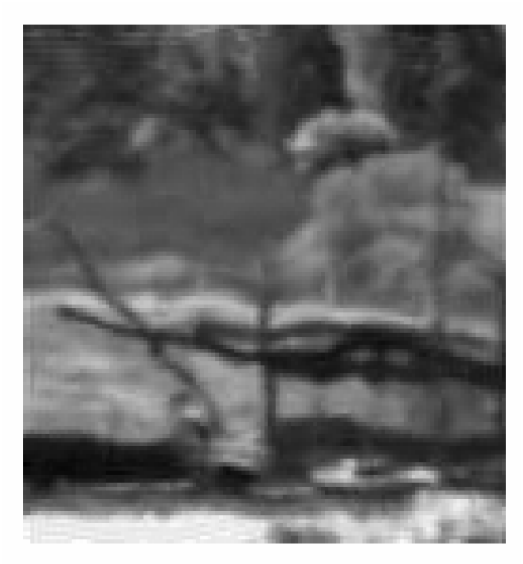


Fig. 6. The turbulent captured image

Fig. 7. The resulting image of the non-turbulent background and unaffected vehicles (compare with Fig.6).

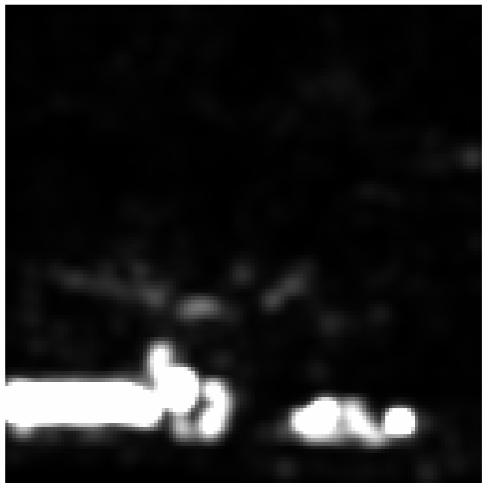


Fig. 8. The error function when the translation vector is substituted

An example of an image frame from a real video sequence captured through turbulent atmosphere is shown in Fig. 6.

In the experiments, a median filter calculated over a sequence of 128 images along the temporal axis was used to generate an un-distorted stationary "reference" image. Error function found for this video sequence is illustrated in form of a weight error function in Fig. 8. The same frame generated by warping only the static objects/background back to their true geometrical location in the scene is shown in Fig. 7. In comparing with image of Fig. 6, the restored image background contains straight lines with no visible geometric distortions while the moving vehicles also appear without any visible artifacts.

## 4. Inter-Channel Intra-Frame Fusion

The Inter-Channel Intra-Frame method is an extension of the method presented by Farooq at el. [1]. The fusion technique consists in weighted pixel-wise averaging of corresponding denoised thermal $I_{IR}$ and turbulence-compensated visual range $I_{visual}$ images:

$$I_{k,l}^{Fused} = \frac{I_{k,l}^{IR}(w_{k,l}^{IR} + w_{k,l}^{IR,N}) + I_{k,l}^{Visual}(w_{k,l}^{Visual} + w_{k,l}^{Visual,N})}{(w_{k,l}^{IR} + w_{k,l}^{IR,N}) + (w_{k,l}^{Visual} + w_{k,l}^{Visual,N})} \quad (16)$$

where $(k,l)$ are pixels' indices, $(w_{k,l}^{IR}, w_{k,l}^{Visual})$ are user defined weights that specify "importance" of the corresponding channel and $(w_{k,l}^{IR,N}, w_{k,l}^{Visual,N})$ are local signal-to-noise ratios in the channels. The use of later assumes that the higher the signal-to-noise ratio in the channel the heavier is the pixel's weight.

For the visual range image weights $w_{k,l}^{Visual}$, the local spatial/time variances of the intensity of the visual range images were suggested in [1]. In our approach, they are computed as:

$$w_{k,l}^{Visual} = \frac{g_1^{Visual} + g_2^{Visual} \cdot \sigma_{(k,l)}^{V}}{\max_{(k,l)}\{g_1^{Visual} + g_2^{Visual} \cdot \sigma_{(k,l)}^{V}\}}, \quad (17)$$

where $g_1^{Visual}$ and $g_2^{Visual}$ are user defined scalars and $\sigma_{(k,l)}^{V}$ are local image variances computed in a running window.

The thermal weights $(w_{k,l}^{IR})$ are determined in the assumption that "importance" of pixels is determined by their contrast with respect to their background defined as a module of difference from an estimate of the "average" neighborhood:

$$w_{k,l}^{IR} = \frac{g_1^{IR} + g_2^{IR} \cdot \left| I_{k,l}^{IR} - \bar{I}_{k,l}^{IR} \right|}{\max_{(k,l)}\left\{ g_1^{IR} + g_2^{IR} \cdot \left| I_{k,l}^{IR} - \bar{I}_{k,l}^{IR} \right| \right\}} \quad (18)$$

where $g_1^{IR}$ and $g_2^{IR}$ are user defined scalars and $\bar{I}_{k,l}^{IR}$ are the "average" neighborhood estimates. In our experiments, two methods for the neighborhood "average", Local Mean and Local Median [15, 21], have shown good results.

For determination of local signal-to-noise ratio in visual range and thermal images, three methods for evaluating noise level in every pixel over pixel's neighborhood were tested: estimation of additive noise variance through local autocorrelation function in a running window, estimation of additive noise variance through evaluation of noise floor in image local spectra in a running window and estimation of impulse noise probability through the histogram of the prediction error [21]. Experiments revealed that no impulse noise was present in the images and that evaluation of additive noise variance through analysis of image local correlation function in running window of 13x13 pixels provided the best results in both visual range and thermal channels.

One frame of the visual image sequence acquired from an observation system is shown in Fig. 9. A frame of the thermal video sequence was shown in Fig. 4(a). Fig. 10 illustrates the importance of weighing fused images according to their local signal-to-noise ratio. Graphs on the figure show row wise average power spectra of images fused without (solid line) and with the weighing (dotted line). One can see from this figure that noise floor in the fused image generated with the weighing is substantially lower.

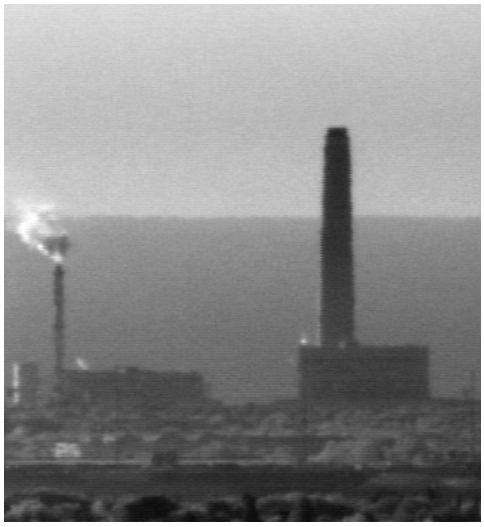


Fig. 9. Visual range image captured from the video sequence

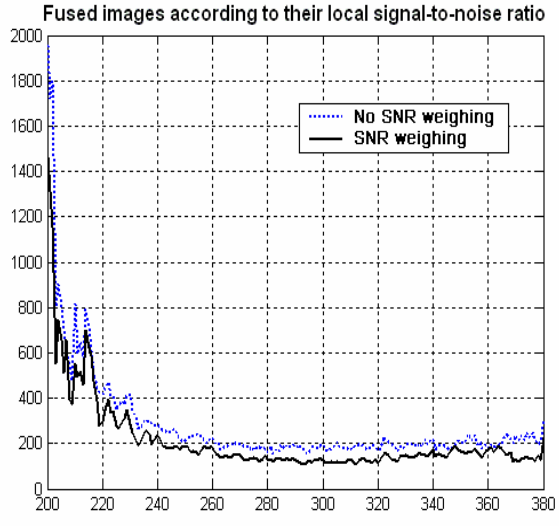



Fig. 10 - Mean Power spectra of fused images without SNR weighing (dashed line) and with SNR weighing (solid line)

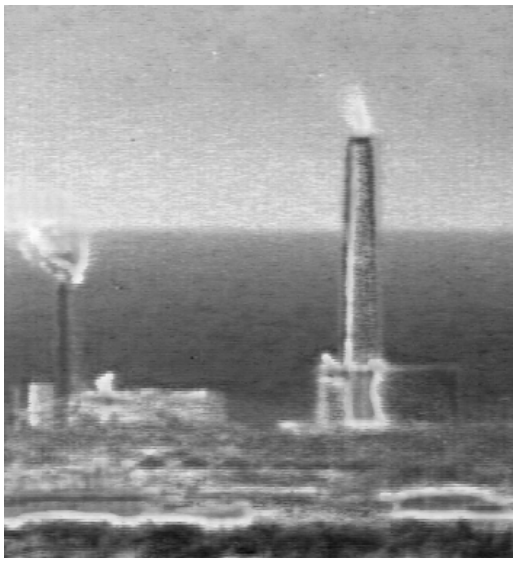


Fig. 11. The Fused Image

Fig. 12 – Fused Image with 3D DCT noise reduction and Turbulence Reduction

Fig. 11 illustrates one frame of the fused image sequence generated using Eq. (17) and Eq. (18) with noise weights evaluated using local correlation functions and with no denoising in the thermal channel. A frame of the final fused image sequence with 3-D SWDCT denoising is shown in Fig. 12. The video sequences can be seen on the authors' web-site ([30]).

## 5. Conclusion

The paper describes a two-phase fusion system for video and thermal streams, which utilizes temporal and spatial fusion techniques (Fig. 13). In Intra-Channel Inter-Frame fusion, compensating atmospheric turbulence in visual range images using local image registration method and image denoising in thermal images using 3-D sliding window filtering in DCT domain are performed. The final fusion is achieved through an Inter-Frame Intra-Channel technique based on the local weighted average method. While each method can stand for itself and has proven good results, the visual and thermal range image fusion system presented here is making use of them all to yield a better system in terms of robustness and visual quality.

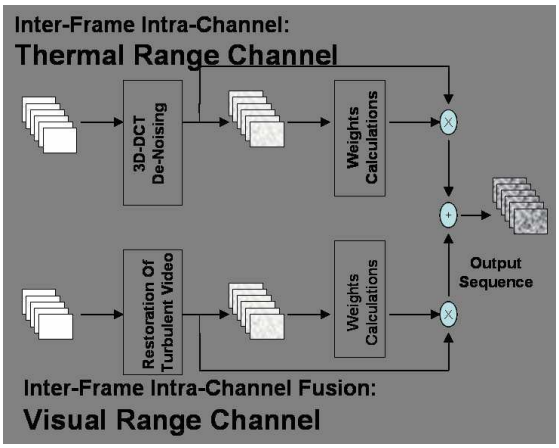



Fig. 13. The image fusion system

## References


[1] E. Lallier and M. Farooq, A Real Time Pixel-Level Based Image Fusion Via Adaptive Weight Averaging, *3rd International Conference on Information Fusion, Fusion 2000*, Volume II, pp. WeC3-3 to WeC3-10, Paris, France, July 2000.

[2] V. Petrovic and C. Xydeas, Computational Efficient Pixel-Level Image Fusion1, Manchester Avionics Research Center (MARC), Department of EE University of Manchester Manchester, UK.

[3] Hassan Ghassemian, Multisensor Image Fusion by Multiscale Filter Banks, *Proceeding of IEEE International Conference on Image Processing ICIP2001*, October 2001.

[4] Terry A Wilson, Steven K. Rogers and Mathew Kabrisky, Perceptual Based Image Fusion for Hyperspectral Data, *IEEE Transaction on Geo science and Remote Sensing*, Vol. 35, No. 4, July 1997.

[5] A. Toet, L.J. van Ruyven, and J. M. Valeton, Merging thermal and visual images by a contrast pyramid, *Opt. Eng.*, Vol. 28, no. 7, pp. 789-792, 1989.

[6] A. Toet, Hierarchical image fusion, *Machine Vision Appl.*, Vol. 3, No. 1, pp. 1-11, 1990.

[7] A. Toet, Multi-scale contrast enhancement with application to image fusion. *Opt. Eng.*, Vol. 31, pp. 1026-1031, May 1992.

[8] P.J. Burt and R. J. Kolczynski, Enhanced image capture through Fusion, *Proc. The 4th Intl. Conf. on Computer Vision*, Berlin Germany, pp. 173-182, May 1993.

[9] H. Li, B. S. Manjunath and K. Mitra, Multisensor image fusion using the wavelet transform, *Graphical models and Image Processing*, Vol. 57, pp. 235-245, May 1995.

[10] Al Bovik, *Handbook of Image and Video Processing*, Academic Press, Canada, 2000

[11] L. Yaroslavsky, *Digital Holography and Digital Image Processing*, Kluwer Scientific Publishers, Boston, 2003
[12] L. Yaroslavsky, K. Egizarian and J. Astola, Transform Domain image restoration methods: review, comparison and interpretation, *Photonics West, Conference 4304, Nonlinear Processing and Pattern Analysis*, San Jose, CA, USA, 2001.
[13] L. Yaroslavsky, Local adaptive image restoration and enhancement with the use of DFT and DCT in a running window, *Proceedings, Wavelet Applications in Signal and Image Processing IV, SPIE Proc. Series, v. 2825*, pp. 1-13, Denver, Colorado, 1996
[14] L. Yaroslavsky, Nonlinear Filters for Image Processing in Neuromorphic Parallel Networks, *Optical Memory and Neural Networks*, v. 12, No. 1, 2003
[15] V. Kim and L. Yaroslavsky, Rank algorithms for picture processing, *Computer Vision, Graphics and Image Processing*, v. 35, p. 234-258, 1986
[16] John L. Miller, *Principles of Infrared Technology*, International Thompson Publishing, New-York, NY, USA, 1994
[17] Gerald C. Holst, *Testing and Evaluation of Infrared Imaging Systems*, second edition, SPIE Press, Bellingham, WA USA, 1998.
[18] Gerald C. Holst, *CCD arrays, Cameras and Displays* second edition, SPIE Press, Bellingham, WA USA, 2001
[19] M. C. Roggermann and B. Welsh, *Imaging Through Turbulence*, CRC Press, Inc, 1996.
[20] Richard D. Hudson Jr., *Infrared System Engineering*, Wiley, Interscience, USA, 1969
[21] L. Yaroslavsky and M. Eden, *Fundamentals of Digital Optics*, Birkhauser, Boston, 1996
[22] L. Yaroslavsky, Local adaptive filtering in transform domain for image restoration, enhancement and target location, *6th Int. Workshop on Digital Image processing and Computer Graphics (DIP-97), SPIE Volume 3346*, pp. 2-17, Vienna, 1997
[23] http://www.eng.tau.ac.il/~yaro/Shtainman/shtainman.htm[1]
[24] W. M. Farmer, *The Atmospheric Filter volume I – Sources*, JCD publishing, 2001.
[25] Suppression of atmospheric turbulence in video using an adaptive control grid interpolation approach, *ICASSP Presentations*, may 2001.
[26] G. Thorpe, A. Lambert and D. Fraser, Atmospheric turbulence visualization through image time-sequence registration, *Proc., International Conference on Pattern Recognition*, vol. 2, p. 1768-1770. 1998
[27] H. van der Elst and J.J.D van Schalkwyk, Modelling and restoring images distorted by atmospheric turbulence, *Proceedings COSMIG*, pp. 162-167, 1992
[28] Senthil Periaswamy, Hany and Farid,, Elastic Registration in the Presence of Intensity Variations, *IEEE Transactions on Medical Imaging*, Volume 22, Number 7, July 2003
[29] L. Yaroslavsky, Boundary effect free and adaptive discrete sinc-interpolatioon, *Applied Optics*, v. 42, No. 20, p.4166-4175, 10 July 2003
[30] http://www.eng.tau.ac.il/~yaro/Fishbain/fishbain.htm[1]

[1] Case Sensitive