RESTORATION OF ATMOSPHERIC TURBULENT VIDEO CONTAINING REAL MOTION USING RANK FILTERING AND ELASTIC IMAGE REGISTRATION

S. Gepshtein, A. Shtainman, B. Fishbain, and L.P. Yaroslavsky

Department of Interdisciplinary Studies, Faculty of Engineering, Tel-Aviv University
Tel-Aviv 69978, Israel
phone: 972 3 6407366, fax: 972 3 6410189, email: shaig@eng.tau.ac.il
web: http://www.eng.tau.ac.il/~yaro/

ABSTRACT

Atmospheric turbulence causes a random blur in images due to random fluctuations of the refraction index of the air through which the light propagates. An algorithm is described that compensates image distortion due to the atmospheric turbulence in video while keeping the real moving objects in the video unharmed. The result of the restoration of the video sequence is a much more "stable" sequence where real moving objects are easier to detect and track. This algorithm is based on Differential Elastic Image Registration which maps each turbulent image to a one good reference image derived from the sequence of the turbulent images. To preserve real motion in the scene, the moving objects are located and the compensation for the distortion of the turbulence is applied only to the static areas of images. The algorithm provides also a smooth transition between the static and the moving areas in the image.

1. INTRODUCTION

Long distant near-earth observations in a hot mid-day are highly affected by atmospheric turbulence that causes spatially and temporally random fluctuations in the index of refraction of the atmosphere [1]. These fluctuations in the optical path length of the propagating light result in phase and amplitude variations of the light's wavefront. 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 modelled by convolution with space variant and random point spread functions [2]. This causes small neighborhoods in the image to randomly move in different directions in different frames. As a result, images captured by optical sensors in the presence of atmospheric turbulence are affected by degradation of resolution and distortion of the image geometry. Watching such video sequences is highly disturbing the eye of the observer since static objects appear to waver in time. Atmospheric turbulent motion appearing throughout the entire image makes it harder for the observer to detect real moving objects in the viewed field.

There is variety of methods for the enhancements of turbulent captured images [3,4,5]. In this paper, a method for a post-processing enhancement of the captured video that stabilizes steady parts of the scene and improves the quality of the image sequence for the observer. A differential image elastic registration method is used to find the translation vector for each pixel in each frame of the video sequence to "inverse-warp" the image to its "true" geometry. Using this same elastic registration method one can also detect the areas of real motion of objects in the scene and use this detection for warping back only the static parts. This way a stable scene is restored where the only moving areas in the scene are the real moving objects.

2. THE ALGORITHM

2.1 Generating geometrically undistorted images with no moving objects

A variety of methods have been proposed for generating one improved image from a sequence of images retrieved by a non-scanning fixed sensor. In [6,7], images are averaged to compensate for small random local displacements, a transfer function is estimated, and a Wiener filter is used for restoration. The algorithm proposed in this paper uses an element-wise rank filtering of each pixel in time sequence to obtain a reference image for elastic registration of steady parts of the scene [8]. The use of rank smoothing filters such as median and alpha-trimmed mean is substantiated in two ways. First, light beam propagating through a turbulent atmosphere will deflect to any point within a certain radius [10], 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, pixel's 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. The second reason for using a rank filter instead of a mean filter is the fact that for moving objects that accommodate a pixel for a short period of time, the value of this pixel will be pushed to the tails of the grey level distribution in a long sequence. The distribution tales will be eliminated when applying the rank filter. It is important that the number of the images will be high enough to eliminate the moving objects. A turbulent degraded image is presented in Fig. 1(a). The local geometric distortions can be easily seen wherever curved lines appear instead of straight lines. A result of a median filter over a sequence of 128 images is shown in Fig. 1(b) .First of all one can see in the figure that local geometrical distortions are

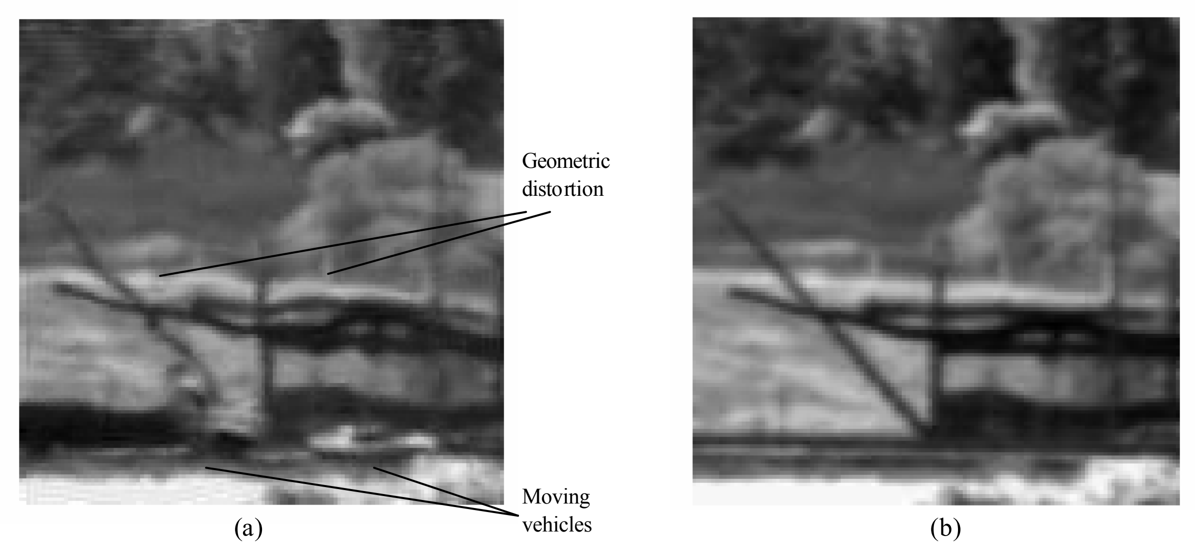


Figure 1: The turbulent captured image (a) and the non-turbulent image obtained by the median filter (b).

fixed. It is even better seen in video sequences. Second, moving objects have disappeared from the image. As can be seen in the figure, the vehicles have dissappeared from the scene leaving an empty road. The non-turbulent image obtained is intended as a reference/target image for computing the translation vectors for each small neighborhood distorted by the turbulence.

2.2 Elastic registration model

Assuming there is no real moving objects in the image, the mapping of one turbulent image to a stable image can be obtained by registrating a spatial neighborhood, surrounding each pixel in the image, to a reference image. In this way a field of motion vectors is received. Each pixel is then mapped to its non-turbulent location. The translation vector is found using an elastic registration method. A similar method is also described in [11]. In it's simplest form, the method assumes that it is sufficient to find only two 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 dx and dy are shift parameters:

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

The translation vector $\vec{\Delta} = (x_0, y_0)$ is computed for each small spatial neighborhood through minimization of the mean square differences between the registrated areas of the two images:

$$\vec{\Delta}_{\Omega 0} = \arg\min \sum_{x,y \in \Omega_0} \left| f(x, y, t) - \bar{f}(x + dx, y + dy) \right|^2$$

where Ω_0 denotes a small spatial neighbourhood of pixel (x_0, y_0). Using a first-order truncated Taylor series expression we get:

$$E(dx, dy) \approx \sum_{x,y \in \Omega} [f(x,y,t) -$$

$$\left(\bar{f}(x_0, y_0, t) + dx \cdot f_x(x_0, y_0) + dy \cdot f_y(x_0, y_0) - \Delta f_t(x_0, y_0)\right)]^2$$

where $f_x(\cdot)$, $f_y(\cdot)$ are the spatial derivative of $f(\cdot)$, and $\Delta f_t(x, y) = f(x, y, t) - \bar{f}(x, y)$.

In vector matrix denotation, the error function is approximated by:

$$E(\vec{\Delta}) \approx \sum_{x,y \in \Omega} \left(\Delta f_t - [f_x, f_y]^T \cdot \vec{\Delta}\right)^2 \qquad (1)$$

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

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

obtained by solving the following equation:

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

Figs. 2 (c,d) illustrate, in form of grey scale images, the spatial distribution of local translation vectors in x and y directions, respectively, obtained for turbulence degraded image shown in Fig. 1 (a) and for a reference image shown in Fig. 1(b). Window size for determination of local translation vectors was 7x7 pixels. This translation vector field provides, for every frame, the image displacements needed to align it with the reference image.

In general, atmospheric turbulence may distort images not only by displacing them but also by rotating neighborhoods. Image frames may also differ in contrast and brightness in comparison to a reference image. One can incorporate displacement and rotation parameters and an explicit change of local contrast and brightness by means of more general model:

$$c \cdot f(x, y, t) + b = \bar{f}(r_1 x + r_2 y + dx, r_3 x + r_4 y + dy)$$

(a) (b) (c) (d)

Figure 2: Translation vector fields from the turbulent-degraded image (a) to the reference image (b) are depicted in (c) and (d) for the x and y components respectively.

as it is described in [11] where dx and dy form a translation vector, r_1 through r_4 form a 2x2 affine matrix and c,b are the parameters of change in image contrast and brightness respectively. Adding these parameters may improve the accuracy of the inverse-warp of the turbulent image.

2.3 Turbulent compensation

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. Absolute values of the translation vector components are in general, non-integer numbers which requires image resampling with sub-pixel shifts. We used a warp technique with discrete sinc-interpolation in moving window in DCT domain to obtain the value of the intermediate pixels with least mean square error [8,9]. The resulting image is composed of the interpolated values 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 in more than one iteration. For a second iteration it is required to compute a translation vector field using the first set of compensated images as the source and the same reference image as the target. The image obtained in the first iteration is warped again by 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.

Bilinear and sinc interpolation were tested on simulated turbulent-degraded images. It was found that in the first iteration the mean square error between the original image and the interpolated turbulent-degraded image is very close using both types of interpolation. When applying a second interation it was found that the MSE converges when sinc-interpolation is applied while it diverges when bilinear interpolation is applied.

2.4 Compensation in the presence of motion

The translation vector $\vec{\Delta}$ is used to warp pixels of static objects and background to their non-turbulent location. At the location occupied by moving objects, $\vec{\Delta}$ will not be a true translation vector because the target image, obtained by the rank filtering, may not contain that moving object. Therefore it is necessary to distinguish between real motion and turbulent motion in the image sequence.

The suggested algorithm detects object's motion in the following way. After alignment of turbulent frames to the reference image, the error function of Eq. (1) is computed again for every pixel. The resulting array of errors contain two types of errors. One type is that of the error of the turbulent mis-compensations due to the inaccuracies of the translation vector and of resampling. This error is small and is bounded to a certain limit. The second type of error are 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.

Compensating the turbulent distortions of only the static areas may leave visible unnatural edges surrounding the moving objects. In order to obtain a smooth transition between the static and the moving areas, the error matrix is truncated at a certain threshold (thr) where any value exceeding this threshold is considered a real motion as following:

$$\widetilde{E}(\mathrm{x},\mathrm{y}) = \begin{cases} 1 & \mathrm{E}(\mathrm{x}_0, y_0) \geq thr \\ \dfrac{\mathrm{E}(\mathrm{x}_0, y_0)}{thr} & \mathrm{E}(\mathrm{x}_0, y_0) < thr \end{cases}$$

A new translation vector field is then calculated for each pixel location(x,y) the new value is given by

$$\vec{\Delta}_2\,(\mathrm{x,y}) = (1 - \widetilde{E}(x,y)) \cdot \overline{\Delta}(x.y)$$

In this way, for no motion $\vec{\Delta}_2$(x,y) remains the same and becomes zero where there is an error exceeding the decided threshold. Since $\widetilde{E}(\vec{\Delta})$ is not a binary number, any value between zero and one will reduce the value of $\vec{\Delta}$ according to the error's size. This way the transition from a moving object to the background in the image will be smooth and made unnoticeable to the eye of the observer.

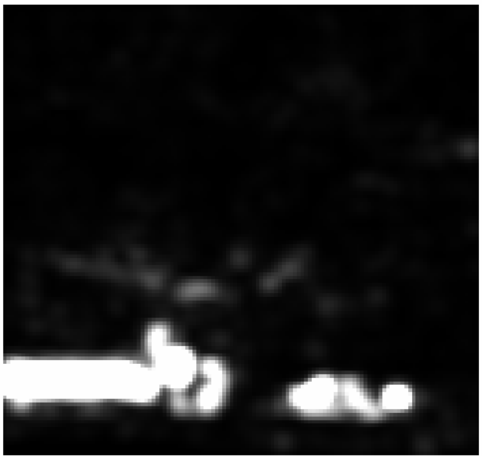
If a second iteration is used for the turbulent compensation step, the new translation vector is calculated in the same manner for a second time but there is no need to compute the error $\tilde{E}(\vec{\Delta})$ again. The error that was found in the first iteration should be used again since it was the one that distinguished between the static and the moving areas:

$$\vec{\Delta}_2\,(x,y)_{2nd\_iter} = (1 - \tilde{E}(\vec{\Delta})_{1st\_iter}) \cdot \vec{\Delta}\ (x,y)_{2nd\_iter}$$

To get the restored un-distorted video, each turbulent image in the video sequence is warped by the matrix $\vec{\Delta}_2\,(x,y)$ calculated for that image using discrete sinc-interpolation for best results.

3. EXPERIMENTAL RESULTS

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. A non-turbulent single frame of a real captured turbulent video was shown in Fig. 1(a). In this experiment, an un-distorted stationary image was calculated using a median filter over each pixel using all 128 images Fig. 1(b). This image served as a reference for computation of the translation vectors from each of the turbulent



(a)

(b)

Figure 3: The error function when the translation vector is substituted (a) and the resulting image of the non-turbulent background and unaffected vehicles (b)

images. As can be easily seen, the vehicles on the road have disappeared.

Finding the translation vectors and solving the error function, we obtain an image of the moving objects' locations in form of a weight error function as it is illustrated in Fig. 3(a). Finally warping only the static objects/background back to their true geometrical location in the scene results in a non-turbulent image retaining the real movement in the field, figure 3(b). As can be compared with the source image (figure 1a), the restored image background contains straight lines with no visible geometric distortions while the moving vehicles appear without any artifacts on sharp edges around them that may appear due to segmentation. A result video can be found on authors website [12].

4. CONCLUSIONS

An algorithm for compensating image blur in a sequence of video frames, obtained through a turbulent atmosphere, is suggested and proved to be effective on a variety of video sequences recorded under turbulence effects. The algorithm is intended to be used for surveillance where the scene is stationary and objects such as cars, people, tanks, etc. happen to pass through the scene. The resulting video scene does not waver as the source and yet the moving objects are left unharmed and easier to detect in a stable background.

REFERENCES

[1] M. C. Roggermann and B. Welsh," Imaging Through Turbulence", CRC Press, Inc, 1996.

[2] W. M. Farmer, "The Atmospheric Filter volume I – Sources", JCD publishing, 2001.

[3] ICASSP Presentations "Suppression of atmospheric turbulence in video using an adaptive control grid interpolation approach" may 2001.

[4] G. Thorpe, A. Lambert and D. Fraser, "Atmospheric turbulence visualization through image time-sequence registration", Proc., International Conference on Pattern Recognition 1998, vol. 2, p. 1768-1770.

[5] H. van der Elst and J.J.D van Schalkwyk, "Modelling and restoring images distorted by atmospheric turbulence", Proceedings COSMIG '92 pp. 162-167.

[6] B. Cohen, V. Avrin, M. Belitsky and I. Dinstein "Generation of a restored image from a video sequence recorded under turbulence effects", SPIE, December 97, vol 36, paper 19027

[7] C. Bondeau, E. Bourennane, "Restoration of images degraded by the atmospheric turbulence", Proceedings of ICSP 1998, vol.2, p. 1056-1059

[8] L.P. Yaroslavsky, "Digital Holography and Digital Image Processing", Kluwer Scientific pub., Boston, 2003.

[9] L.P. Yaroslavsky "Boundary effect free and adaptive discrete sinc-interpolatioon", Applied Optics, 10 July 2003, v. 42, No. 20, p.4166-4175

[10] Y. Glick, A. Baram, H. M. Loebenstein, and Z. Azar, "Restoration of trubulence-degraded images by the most-common mothod", Applied Optics 30(27), 3924-3929 (1991).

[11] Senthil Periaswamy, Hany Farid ", Elastic Registration in the Presence of Intensity Variations", IEEE Transactions on Medical Imaging, July 2003, Volume 22, Number 7.

[12] www.eng.tau.ac.il/~yaro/ShaiGepshtein/movie_no_turb.zip