Sharpening and contrast enhancement of atmospheric turbulence degraded video sequences

Philip Robinson*, Bryn Walters **, Willem Clarke++
School of Electrical Engineering
University of Johannesburg
South Africa
*philipr@uj.ac.za, **bryn.walters@gmail.com, ++willemc@uj.ac.za

Abstract—**Long range imaging systems that capture video through the atmosphere face a major problem in the form of atmospheric turbulence. This turbulence causes a phenomenon called heat shimmer which appears as a blurring and a wavering geometric distortion of the target scene which limits the effective range of the imaging system. We explore an image processing approach to mitigating the blurring effect of this distortion by using a blind deconvolution technique to sharpen the video signal and a dynamic illuminance-reflectance correction technique to improve the signal's contrast. The algorithm is implemented on a Graphics Processing Unit to achieve near real-time performance.**

Keywords-Atmospheric turbulence, heat shimmer, scintillation, terrestrial, blind deconvolution, GPU, illuminance, reflectance

I. INTRODUCTION

The rapid proliferation of digital image capture devices has resulted in the development of technologies that make use of the abundance of still image and video signals produced by these devices [1, 2]. The signals in question can be captured in various situations and will always contain some form of noise or distortion. This can be caused by a variety of factors such as the nature of the image capture device or lens assembly and environmental factors such as lighting, camera jitter or motion in the target scene [2]. The captured signals are not only used by human operators who can cope with these distortions. As image processing technology matures automated systems that use these signals to extract useful data are becoming more common. For these systems degradation in signal quality is a major problem [1, 2, 3].

Modern imaging systems are capable of high magnifications. This class of imaging system is used for long range video surveillance in both the military and civil spaces. In the case where the target scene is being captured from a range over 1 km the effects of atmospheric turbulence become obvious [1, 4].

Turbulence in the atmosphere causes pockets of air of varying temperatures and thus densities to move in a random fashion. This movement is caused by the varying densities of the air pockets, wind and terrain [5]. Light from the target scene must travel through this turbulent atmosphere to reach the imaging system. The varying densities of the air pockets cause this light to be refracted by varying degrees and in a time varying manner. This results in the target scene appearing blurred, washed out and to be wavering or shimmering. This implies that objects in the scene will appear to be moving even when stationary. This effect is dubbed heat shimmer or heat scintillation [1, 3, 4, 5, 6, 7, 8].

Heat shimmer severely limits the effective range of long range imaging systems and as such mitigating the effects of atmospheric turbulence is a major concern when designing these systems.

There are two main schools of approaches to this problem. The first is to make use of a mechanical adaptive optics system to physically compensate for the effects of the atmospheric turbulence on incoming light rays using a system of deformable mirrors [9]. The second approach is to make certain assumptions about the nature of the distortion and make use of an image processing approach to digitally enhance the video signal to attempt to reduce the distortion. A few proposed image processing methods are the direct Discrete Fourier Transform (DFT) solution [4], image registration and fusion [1], Adaptive Control Grid Interpolation [7, 8], Image Time Sequence Registration [10], Neural Network approach based on the Monte Carlo method [11] and the Homomorphic and Power Spectrum approach [12].

This paper approaches the problem of heat shimmer in a similar manner to [7, 8]. The assumption is made that the effects of heat shimmer can be separated into two parts. The first is the blurring and loss of contrast caused by the scattering of incoming light rays and aerosols in the air. The second part is the wavering or geometric distortion that causes the stationary elements of a scene to appear to be moving. This paper will focus on mitigating the first part of the distortion. A blind deconvolution style algorithm is used to sharpen the video and a dynamic luminance-reflectance correction technique is used to improve the contrast while not amplifying any additive noise present. This algorithm is implemented on a Graphics Processing Unit (GPU) to achieve a near real-time performance.

The remainder of the paper will be structured in the following way. Section II will describe the details of the algorithms implementation. Section III will describe the design of the performance experiments. Section IV will show and discuss the experimental results and Section V will be the conclusion.

We would like to thank the Council for Scientific and Industrial Research (CSIR) for sponsoring this work.

II. ALGORITHM DEVELOPMENT

We consider the distortion caused by heat shimmer to be made up of two classes, the first is the blurring and contrast loss caused by light scattering and the second is the geometric distortion. This paper focuses on mitigating the blurring effect of heat shimmer and thus we make use of a primitive stabilization scheme to deal with the geometric distortion. We make the assumption that the geometric distortion is quasi-periodic [7, 8]. This implies that the average of the frames of the video sequence with wavering motion should have a mean displacement of 0. A simple ratio average scheme is used to maintain a running average of the frames in the video sequence as shown in equation 1.

$$g(x,y,t) = \alpha f(x,y) + (1-\alpha)g(x,y,t-1), \quad (1)$$

where *x, y* are the pixel coordinates in an image frame, *t* is the current frame in the sequence, *g(x,y,t)* is the updated ratio frame average, *g(x,y,t-1)* is the previous frame average, *f(x,y)* is the current incoming frame, and α is a scalar value between 0 and 1 which dictates the propotion of the new frame which is added to the average.

We tackle the blurring effect of heat shimmer in two stages. The first part of the algorithm uses a blind deconvolution approach to sharpen the image and the second stage uses a dynamic illuminance-reflectance correction scheme to improve the contrast of the video frame.

A. Blind Deconvolution

The term blind deconvolution was first used and described in [13]. The following model is used to describe the image capture system:

$$g(t) = f(t) * h(t) + n(t), \quad (2)$$

Where *g(t)* is the received frame which contains noise and distortion, *f(t)* is the undistorted scene, *h(t)* is the distortion function which is convolved with the target scene and *n(t)* is the additive noise present in the scene.

In our case the distortion function describes the nature of the blur caused by capturing video through atmospheric turbulence. The additive noise present in the scene is mostly due to the nature of the digital image capture device. The term blind deconvolution refers to the fact that we have no *a priori* information about the nature of the blurring function and have to identify the blur function while only having access to the distorted frame. To achieve this we make use of Hufnagel's model of the blur caused by atmospheric turbulence described in [14]:

$$H(u,v) = e^{-\lambda(u^2+v^2)^{5/6}}, \quad (3)$$

where λ is parameter that controls the intensity of blur and *u* and *v* are the 2 dimensions of the spatial frequencies.

We make use of a Wiener filter to compensate for the blurring effect of the heat shimmer. The Wiener filter is used as it takes into account the power of the additive noise present in the signal to reduce the amplification of that noise in the deconvolved image. Equation (4) describes the Wiener filter [2, 15].

$$H_w = \frac{H^*}{|H|^2 + {S_N}/{S_F}}, \quad (4)$$

where S_N is the power spectrum of the additive noise, S_F is the power spectrum of undistorted the image, H is the model of the distortion and H_w is the Wiener Filter based on the specified model.

To apply the Wiener filter the frame is first transformed into the frequency domain using the Fast Fourier Transform (FFT). The filter is applied in the frequency domain as a convolution operation becomes a simple multiplication [15]. The noise-to-blurred signal ratio is estimated as the difference between the global variance of the whole frame and the average local variance which is assumed to represent the power of the additive noise. This is due to the fact that the noise and the blurred image are uncorrelated. The Noise-to-Signal Ratio (NSR) is thus estimated using the following equation [7, 8]:

$$NSR = \frac{Local\ Variance}{Global\ Variance - Local\ Variance}, \quad (5)$$

To identify the amount of blur present in the image a search space of possible values of λ in equation 3 is specified. The input frame is deconvolved using a Wiener filter constructed from the blur model using a given λ value in the search space. The result must then be evaluated to decide if the best λ value has been found. This is done using a sharpness metric to compare the relative sharpness of deconvolved frames until the value of λ is found that results in the sharpest possible output. In [7, 8] kurtosis is used as the sharpness metric but we found that in images that have large areas of uniform colour this metric is not accurate.

Initially we used a Laplacian operator based metric where the Laplacian operator was applied to the image to accentuate the high frequency information, such as edges, and remove the uniform colour areas. A mean value of this result gave a robust comparative sharpness metric. A problem with this approach is that the metric is applied in the spatial domain. This implies that after each deconvolution the inverse FFT (IFFT) must be applied to convert the result from the frequency domain to the spatial domain. To avoid this computationally expensive approach we use a frequency domain based sharpness metric where the power in the higher frequency bands is used as a measure of the sharpness of the image. Using normalized frequency domain coordinates where 0 is at the DC value and 1 is the highest frequency. We find the power in the frequency band that lies between 0.3 and 0.9.

We use the upper bound of 0.9 to avoid measuring the power of any high frequency noise in the image. This metric allows us to perform the FFT once at the beginning of the algorithm and IFFT only once the optimal value of λ is found.

It was observed that as we traversed the search space for λ the sharpness metric exhibited a single peak and as such once a maxima is found we assume that the best possible value for λ has been located and we have identified the blur present in the current frame. At this stage the IFFT is applied to give the sharpened frame. Figure 1 shows the structure of our blind convolution algorithm.

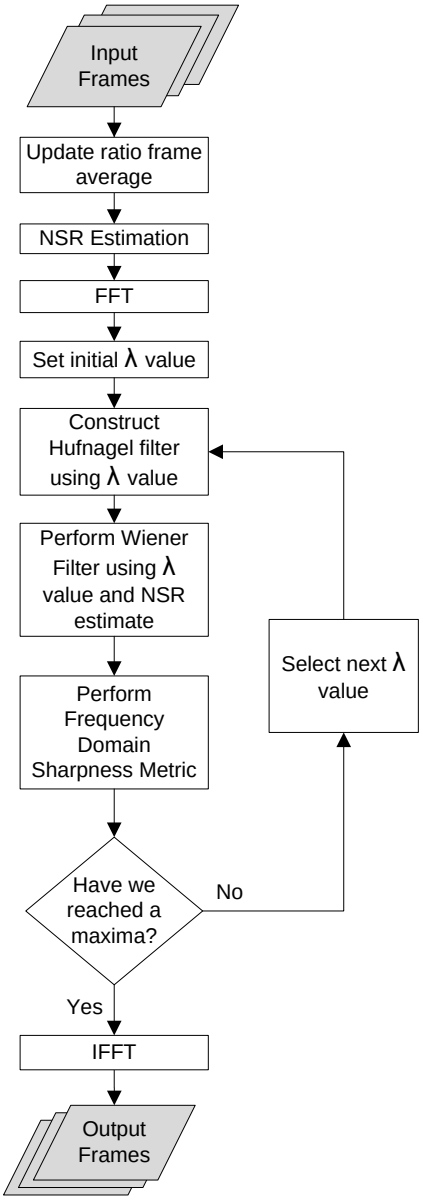

Figure 1: Frame averaging and blind deconvolution algorithm

B. Dynamic Illuminance – Reflectance Correction

Due to the scattering of light and blurring caused by capturing video through the turbulent atmosphere the captured video frames tend to have a low contrast and appear washed out. The standard method for improving contrast is to perform a histogram equalization [2, 16]. This is not ideal in our case as it severely amplifies high frequency additive noise which is present in the original frames and is slightly amplified by the blind deconvolution stage of the algorithm.

We make use of the dynamic illuminance-reflectance adjustment approach presented in [17]. This technique performs dynamic range compression on the low frequency component of an image while preserving the high frequency content so as not to amplify any additive noise present in the image.

The assumption is made that a image *I(x,y)* can be represented as the following product [17]:

$$I(x,y) = L(x,y)R(x,y)\,, \tag{6}$$

where *L(x,y)* is the luminance component of the image and *R(x,y)* is the reflectance component of the image.

The luminance component of the image is assumed to contain the low frequency information and the reflectance component is assumed to contain mostly the high frequency information of the image. The luminance estimate of the image is found by applying a 5x5 discrete Gaussian filter to the input image. The reflectance estimate is then calculated by dividing the input image by the luminance estimate. Once we have the luminance estimate the dynamic range compression is performed using the following sigmoid function [17]:

$$s(v) = \frac{1}{1+e^{v}}, \tag{7}$$

This function is applied to the image in the following steps:

$$L_n{}' = L_n\big(s(V_{max}) - s(V_{min})\big) + s(V_{min})\,, \tag{8}$$

where L_n is the normalized luminance component and V_{min} and V_{max} are chosen based on the mean value of the image as described in [17].

$$L_n{}'' = \ln\left(\frac{1}{L_n{}'} - 1\right), \tag{9}$$

$$L_{n,enh} = \frac{L_n{}'' - V_{min}}{V_{max} - V_{min}}, \tag{10}$$

where $L_{n,enh}$ is the enhanced normalized luminance.

The next step of the scheme is to perform a mid-tone frequency enhancement which is done using equations 11 and 12:

$$E(x,y) = R(x,y)^p = \left(\frac{I_{L(x,y)}}{I(x,y)}\right)^p, \quad (11)$$

where $I_{L(x,y)}$ is the original image filtered with a 10x10 Gaussian filter and p is chosen based on the image's standard deviation as specified in [17].

$$L_{n,enh}{}' = L_{n,enh}{}^{E(x,y)}, \quad (12)$$

where $L_{n,enh}{}'$ is the final fully adjusted luminance estimate. The final step is to reconstruct the image by reintroducing the reflectance component to the adjusted luminance component using equation 6.

This portion of the algorithm is illustrated in the figure 2.

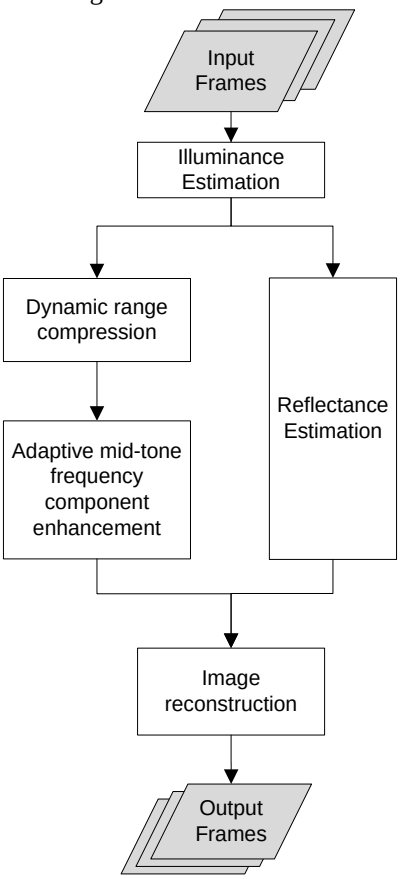

Figure 2: Dynamic Illuminance-Reflectance adjustment algorithm

III. Experimental Design

Once the algorithm was implemented a number of experiments were performed to confirm the algorithm does in fact enhance sharpness and contrast. Firstly, human evaluation was used to compare processed and unprocessed frames from two video sequences that were captured through real atmospheric turbulence for sharpness and contrast improvement. For objective experiments the colour histograms of the processed and unprocessed frames were compared to confirm that colours are evenly distributed throughout the entire range indicating better contrast in the processed frames.

To objectively measure the improvement in sharpness two sharpness metrics were used. The first was a laplacian operator based sharpness metric. This metric applies the laplacian operator to the image which has a high response for high frequency image elements such as edges and a low response for low frequency features such as areas of uniform colour [2, 16]. The mean of the intensities in the resulting image gives a good comparative metric for sharpness. The second sharpness metric used is Shannon Entropy, which measures the average information contained in an image and is described in [16].

The algorithm was implemented on the GPU using the OpenGL API and GLSL shader language to achieve near real-time performance. To compare the GPU implementation to a CPU implementation the algorithm was reimplemented on the CPU using the OpenCV and FFTW libraries. Both GPU and CPU implementation were done using the C++ language using the G++ compiler for consistency. The computational speed of the algorithms were measured by processing a series of videos of increasing resolutions and measuring the average processing time for 200 frames of the same video sequence. The specifications of the computer used to perform these experiments are given in table 1.

CPU	AMD Athlon 7750 Black Edition
Motherboard	Asus M3N-H
RAM	4 GB DDR2 400 MHz
GPU	Nvidia GeForce GTX 260

Table 1: Test PC Specifications

IV. Experimental Results

Figures 3 and 4 show an unprocessed and processed frame from two different video sequences captured through turbulent atmosphere. Visual inspection shows that in each case the high frequency information in the image is enhanced significantly and that the contrast enhancement reveals many details that were not originally visible.

Figure 5 shows the results of the Laplacian sharpness metric for the unprocessed frames and processed frames. It can be seen that the amount of high frequency content in each image has indeed been amplified quite significantly in both cases. Frame B shows a larger improvement due to the larger amount of high frequency components it contains compared to Frame A which has large areas of uniform colour. Figure 7 shows the measured Shannon entropy for each frame. Both frames show an improvement in entropy which means the algorithm restored information that was lost due to the effects of heat shimmer. Frame B shows only a slight increase in entropy because the original frame had a fair amount of high frequency information in it to start with so the blind deconvolution stage of the algorithm did not sharpen the frame very dramatically.

Figure 8 shows the histograms of the unprocessed and processed versions of Frame B. It can be seen that the

unprocessed frame's colours were contained in only a small range of colour values and that the dynamic illuminance-reflectance adjustment stretched the histogram to occupy almost the entire colour range thus increasing the frames contrast significantly.

Figure 6 shows the computational performance results of the full algorithm implemented on the GPU and CPU for a set of different video resolutions. It is apparent from the figure that the GPU implementation runs far faster than the CPU implementation. At the smallest resolution the algorithms run at very similar speeds but as the size of the data increases the CPU experiences an exponential decay in processing times consistent with a serial architecture.

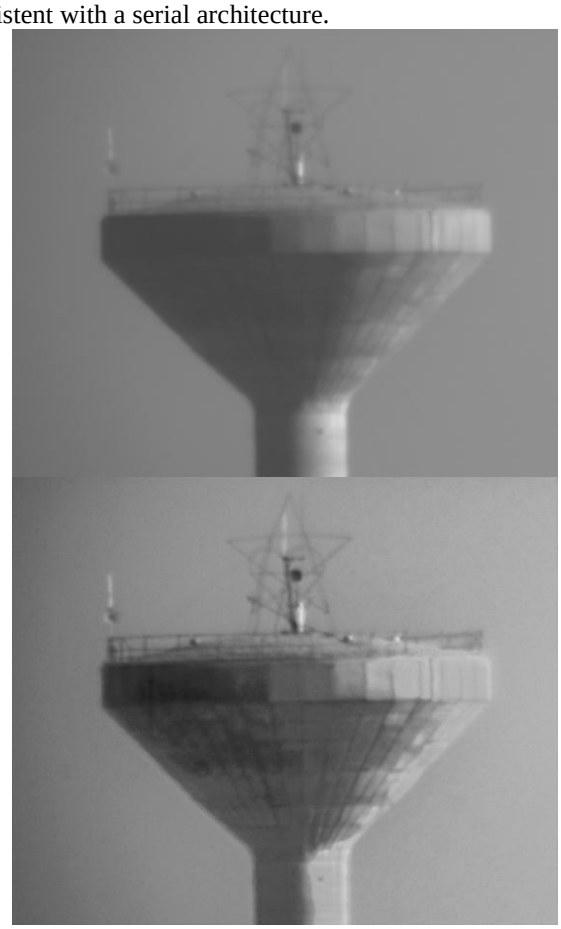

Figure 3: Unprocessed and Processed video frame A

The GPU implementation retains a fairly flat performance profile as the resolution increases and at the largest size runs at two orders of magnitude faster than the CPU implementation.

Figure 4: Unprocessed and Processed frame B

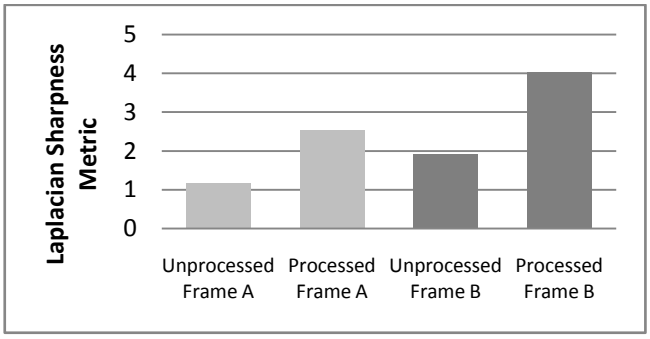

Figure 5: Laplacian sharpness metric results for Frame A

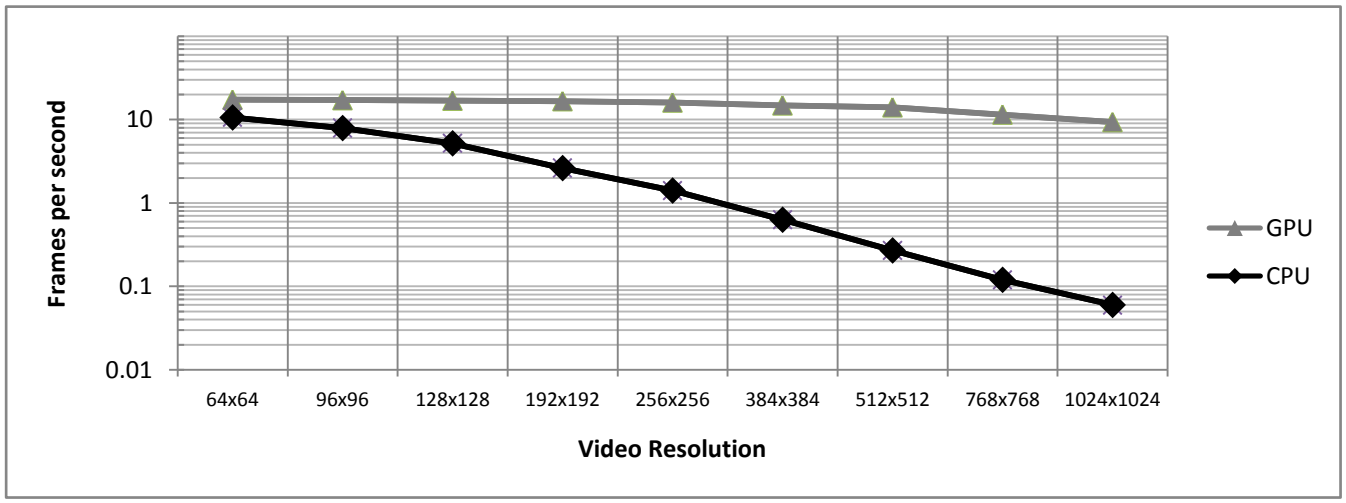

Figure 6: CPU vs GPU computational performance results

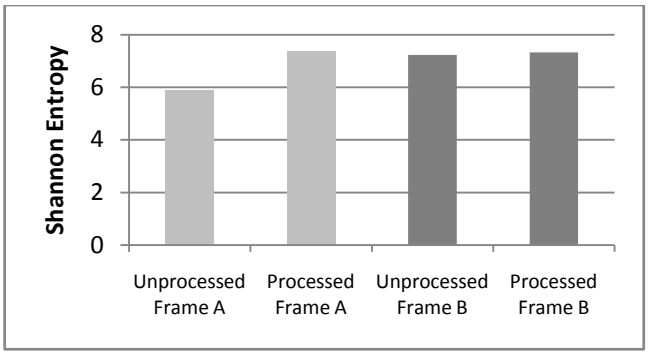

Figure 7: Shannon Entropy results for Frame B

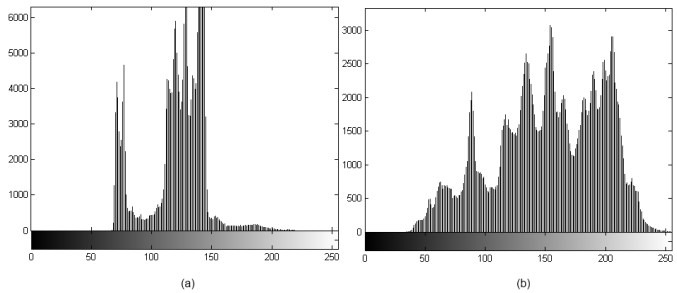

Figure 8: a) Unprocessed Frame A histogram b) Processed frame A histogram

V. CONCLUSION

When long range imaging systems capture video sequences through turbulent atmosphere a phenomenon called heat shimmer is observed. Heat shimmer causes the captured video to appear blurred and washed out. Heat shimmer also causes geometric distortions that make stationary elements of a scene appear to move in a quasi-periodic fashion.

This paper focuses on mitigating the blurring effect and contrast reduction caused by heat shimmer. A primitive stabilisation approach is used to combat the geometric distortions in the video sequences. A blind deconvolution approach is used to identify the intensity of the blurring effect in a given frame and compensate for the distortion without any *a priori* information. A dynamic illuminance-reflectance approach is used to improve the contrast of a frame without amplifying the additive noise present in the frame.

A number of sharpness metrics are used to confirm the effectiveness of the blind deconvolution algorithm at sharpening the image. Histogram analysis is used to confirm the improvement of contrast by the algorithm. The algorithm is implemented on the GPU and the CPU to compare the computational performance increase experienced on the GPU as the resolution of the input video increases.

The algorithm presented in this paper enhances video distorted by heat shimmer resulting in a sharper video with improved contrast and very little increase in additive noise which most sharpening and equalisation approaches experience. The GPU implementation of the algorithm runs at near real-time speeds on a mid-range GPU from the previous generation and it is expected that the algorithm will run at real-time speeds when run on a top-end card from the current generation.

REFERENCES

[1] W. Zhao, L. Bogoni, M. Hansen, "Video Enhancement by Scintillation Removal", *icme*, , IEEE International Conference on Multimedia and Expo (ICME'01), pp. 71, 2001.

[2] A. Bovik, "Handbook of Image and Video Processing", Academic Press, San Diego, 2000.

[3] S.D. Ford[1], B.M. Welsh[1], M.C. Roggemann[2], "Constrained least-squares estimation in deconvolution from wave-front sensing", Optics communications, [1]Department of Electrical and Computing Engineering, Air Force Institute of Technology, [2]Department of Electrical Engineering, Michigan Technological University, 1997.

[4] B.R. Frieden, "An exact, linear solution to the problem of imaging through turbulence", *Optical sciences centre*, University of Arizona, Tucson, AZ 85721, USA, 1997.

[5] J.A. Dutton, *The Ceaseless Wind: An Introduction to the theory of Atmospheric Motion*, Dover Publications, 1986.

[6] G. Barnard. Restoration of Turbulence Degraded Images, Literature study report, Doc. No. QT-RA02-D-P047-001, Thales Advanced Engineering (PTY) LTD., October 1990.

[7] D.H. Frakes, J.W. Monaco, M.J.T. Smith, "Suppression of Atmospheric Turbulence in Video Using an Adaptive Control Grid Interpolation Approach", School of Electrical and Computing Engineering, Georgia Institute of Technology, Atlanta, GA, 2001.

[8] D. Li, R. Mersereau, D. H. Frakes, M. J. T. Smith, "New method for suppressing optical turbulence in video", Proc. EUSIPCO, 2005.

[9] C. Max, "Introduction to Adaptive Optics and its History", American Astronomical Society 197th meeting, 2001.

[10] G. Thorpe[1], A. Lambert[2], D. Fraser[2], "Atmospheric Turbulence Visualization through Image Time-Sequence Registration", [1]*Boeing Australia Limited*, 363 Adelaide Street, Brisbane, QLD 4000, Australia, [2]*School of Electrical Engineering*, University College, University of New South Wales, ADFA Canberra ACT 2600, Australia.

[11] B. Cong, "Encoding Neural Networks to Compute the Atmospheric Point Spread Function", Dept. of Computer Science, California State University, Fullerton, CA 92834 - USA.

[12] I. Goss-Ross, G. Barnard, B. Coetzer. *Restoration of Turbulence Degraded Images*, Algorithm Development, Doc. No. RP-RA02-S-C047-003, Thales Advanced Engineering (PTY) LTD., March 1991.

[13] T.G. Stockham Jr., T.M. Cannon and R.B. Ingebresten, Blind deconvolution through digital signal processing, Proc. IEEE63 (April 1975), 678–692.

[14] R. E. Hufnagel and N. R. Stanley, "Modulation transfer function associated with image transmission through turbulence media," Journal of the Optical Society of America A , vol. 54, pp. 52–61, 1964.

[15] E. C. Ifeacher, B. W. Jervis, Digital Signal Processing: A Practical Approach, Second Edition, Prentice Hall, Pearson Education Ltd. 2002.

[16] R.C. Gonzalez., R.E. Woods, S.L. Eddins, "Digital Image Processing Using MATLAB", Prentice Hall, New Jersey, 2003.

[17] L. Tao, R. Tompkins, V. K. Asari, "An Illuminance-Reflectance Nonlinear Video Enhancement Model for Homeland Security Applications", *Proceedings of the 34th Applied Imagery and Pattern Recognition Workshop*, ISBN: 0-7695-2479-6, October 2005.