# Adaptive Multi-Scale Retinex algorithm for contrast enhancement of real world scenes

Philip E. Robinson and Wing J. Lau
Faculty of Engineering and Built Environment
University of Johannesburg
South Africa
philipr@uj.ac.za, wing.j.lau55@gmail.com

***Abstract*—Contrast enhancement is a classic image restoration technique that traditionally has been performed using forms of histogram equalization. While effective these techniques often introduce unrealistic tonal rendition in real-world scenes. This paper explores the use of Retinex theory to perform contrast enhancement of real-world scenes. We propose an improvement to the Multi-Scale Retinex algorithm which enhances its ability to perform dynamic range compression while not introducing halo artifacts and greying. The algorithm is well suited to be implemented on the GPU and by doing so real-time processing speeds are achieved.**

***Index Terms*—Contrast Enhancement, Retinex, Adaptive, Multi-Scale Retinex, GPU**

## I. Introduction

The human eye is a very complex and amazingly versatile imaging system. It exhibits an enormous dynamic range and can change its sensitivity very rapidly to operate in a large range of light levels; this ability is called brightness adaption. However the range of distinct intensities that the eye can distinguish at any one time is quite small compared to the total range of intensities it can adapt to perceive. This means the eye will struggle to discern very dim intensities when simultaneously exposed to very bright intensities. Unfortunately most artificial imaging systems have a much poorer level of brightness adaption than the human eye and as such can capture a very low dynamic range of intensities [1].

This results in many digital images exhibiting poor contrast either globally or in local regions. Contrast refers to the difference between the highest and lowest intensities used to represent an image. The wider the range of intensity values used to represent the information in an image or area of an image the higher the contrast. Contrast can also describe the distribution of intensity values used to represent the structures in the image. If the occurrence of intensity values are evenly distributed over the entire range of possible values it will be easier for a human viewer to to distinguish differing intensities. This is due to the fact that the various intensity levels will be spread further apart and are thus easier for our eyes to tell apart [1], [2].

There are a number of situations that can result in images exhibiting poor contrast. Some examples include images captured over a long range through the atmosphere where scattering and aerosols in the air result in the representation of the scene only occupying a small portion of the possible intensity values [3]. A second example is scenes with a very high dynamic range where portions of the image are in shadow and another portion of the image contains very bright information; this is otherwise known an High-Dynamic Range images (HDR). A final example is in medical scans where information produced by the detectors is very densely packed into the digital image representation [4].

The literature contains many techniques for contrast enhancement. The simplest is to apply an offset and gain to the image intensities based on the minimum and maximum values found in the image. This technique does improve contrast of most images but it is very sensitive to noise and outliers as a single noisy pixel can be found to be one of the extreme values and drastically perturb the scaling [2].

Histogram equalization quickly became a popular form of contrast enhancement and was first applied to medical scan images. These techniques operate based on the histogram of intensity values of an image. They seek to redistribute the intensities in the image in such a way as to achieve a uniform distribution of intensities across the entire intensity range [4]. Basic histogram equalization considers the histogram of the entire image in a global fashion, and as such struggles in images where a small portion of the image exhibits a drastically different intensity distribution which would then throw off the equalization for the rest of the image. To combat this Adaptive Histogram Equalization (AHE) was developed which performed the same process on a per-pixel basis based only on the pixel's neighbourhood. This approach achieves much higher contrast but amplifies noise, often in an extreme manner [5].

One of the most versatile forms of AHE is Contrast-Limited Adaptive Histogram Equalization (CLAHE) which puts a limit on just how drastically an intensity level can be redistributed. This algorithm works extremely well on medical images and fairly well on most real-world images. It has the added advantage of being relatively simple and as such has been implemented in a real-time system using specialized hardware [6]. While there has been an enormous amount of research done into histogram based contrast enhancement algorithms, such as [7], [8], they have some drawbacks. These algorithms tend to produce unrealistic effects when they are applied to real-world images which is why they have mainly been applied to scientific images like medical, thermal and satellite images. In addition while consumer Graphics Processor Units (GPU) have provided a parallel computing platform that has accelerated the implementation of real-time image processing algorithms the construction of the histogram is awkward on the parallel architecture of the GPU. Efficient implementations of the histogram have been proposed for GPU frameworks like CUDA but for lower level GPU API's like OpenGL the histogram is still costly to compute. This paper explores another approach to contrast enchantment which is better suited to real-world scenes and easily implemented on the GPU.

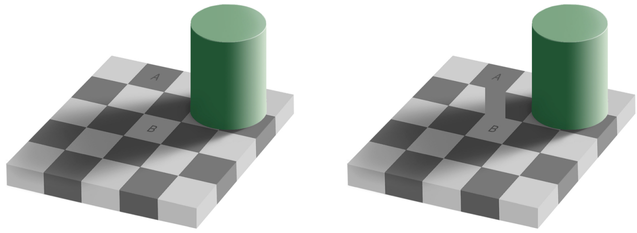
In this paper we are going to make use of Retinex theory to perform contrast enhancement. Retinex theory was first proposed by Land and McCann in [9] to describe a model of how the eye perceives light intensities, which is often at odds with the actual physical intensities the eye experiences [10]. This theory has been greatly expanded for use in image processing since its proposal in papers such as [11]–[14]. This paper aims at furthering this approach which due to its origins in Retinex theory produces very natural looking results and lends itself well to real-time implementation on the GPU.

The remainder of this paper will be structured as follows. Section II will provide a description of Retinex theory and its application to contrast enhancement. Section III will present the proposed algorithm. Section IV will show our results and Section V will conclude the paper.

## II. Overview of Retinex-based contrast enhancement

Retinex theory was developed by Land and McCann to model the disparity they observed between the lightness of various parts of a scene perceived by the human eye and the absolute lightness that was actually incident on the eye. What they found was that the eye does not perceive absolute lightness but rather relative lightness. This means that the eye perceives the variations of relative lightness in local areas in the scene [9], [10]. This phenomenon is what gives the human eye its great dynamic range and is illustrated in the classic optical illusion shown in Fig. 1. While it is difficult to believe, square A and square B in Fig. 1 are the exact same colour. We perceive that square B is a lighter colour because it is surrounded by darker squares and in contrast to its immediate neighbours it is indeed lighter. Square A on the other hand appears to be dark because in contrast to its immediate neighbours it is darker. Our eyes and our brain cannot help but perceive the absolute lightness of square B to be brighter than square A even though we can see that they are identical in the second image.

The second element of Retinex theory that we exploit to achieve contrast enhancement is that our eyes exhibit a logarithmic response to lightness. This is to allow us to differentiate a greater number of dim intensities compared to




(a) (b)

Fig. 1. (a) The Adelson Checker Shadow Illusion [15] (b) Proof that square A and B are identical intensities

bright intensities [1]. This allows us to operate better in dark environments which are far more challenging for our visual system than bright environments. This means that using a logarithmic mapping Retinex based algorithms map intensities using a response curve that appears more natural to our eyes.

Equations 1 and 2 show the basic formulation of the Single Scale Retinex (SSR) scheme.

$$R(x,y) = \frac{\log I(x,y)}{\log[F(x,y) * I(x,y)]} \tag{1}$$

$$R(x,y) = \log I(x,y) - \log[F(x,y) * I(x,y)] \tag{2}$$

where $I(x,y)$ is the 2-dimensional input image, "$*$" denotes the convolution operator, $F(x,y)$ is the surround function, and $R(x,y)$ is the SSR output. $F(x,y)$ is the function that defines the shape and weighting of the averaging kernel used as a measure of the neighbourhood lightness for each pixel [11]. It can be seen that SSR can be considered to be a logarithmic mapping of the ratio of the current pixel intensity to the average intensity around the pixel. In [11] it is shown that the best choice for the surround function is a Gaussian which not only gives the best results but has the added advantage of being a separable kernel. A separable 2D kernel is one that can be expressed as the outer product of 2 vectors. This means that instead of applying the kernel in its 2 dimensional form one can apply each of the constituent vectors. This approach drastically reduces the number of computations required to apply the kernel to an image. Equation 3 describes the Gaussian function.

$$F(x,y) = Ke^{-\frac{x^2+y^2}{2\sigma^2}} \tag{3}$$

where $\sigma$ is the standard deviation that controls the scale of the surround. $K$ is chosen to normalize the kernel such that:

$$\iint F(x,y)\,\mathrm{d}x\,\mathrm{d}y \ = 1. \tag{4}$$

SSR does exhibit a few problems in that if the scale is set too small you get good dynamic range compression but you generate a halo effect around edges. If you set the scale too high you get less dynamic range compression and a greying effect can be seen in more uniform areas. In [12] it is shown that applying the Retinex scheme at only a single scale cannot simultaneously provide good tonal rendition and good dynamic range compression and thus they proposed a Multi-Scale Retinex (MSR) algorithm. This algorithm applies the Retinex technique at several scales and then combines the results using a weighted sum to produce an output as shown in equation 5.

$$R_{MSR}(x,y) = \sum_{n=1}^{N} w_n R_n(x,y) \tag{5}$$

where $R_{MSR}(x,y)$ is the Multi-Scalar Retinex (MSR) output, $R_n(x,y)$ is the output of Single Scale Retinex (SSR) at different scales, and $w_n$ are the weights associated with the different scales. The weights are chosen so that $\sum w_n = 1$, and $N$ designates the number of scale levels used.

The MSR output contains logarithmic values that run from very small negative numbers into the positive domain. As such the final step in the algorithm is to normalize the resulting values to fall between 0 and 1. This is done using a gain/offset scheme as described in equation 6.

$$R_{MSR_i}(x,y) = \alpha[\sum_{n=1}^{N} w_n R_{n_i}(x,y)] - \beta \tag{6}$$

where $\alpha$ is called the gain and $\beta$ is the offset. $\beta$ is based on the minimum value in the image and used to ensure that the minimum value in the final resulting image is 0. The gain $\alpha$ is calculated by dividing 1 by the difference between the maximum and minimum values in the MSR output and scales final resulting image so that its maximum value is 1. These values are calculated globally which means that this approach has a similar problem to a global histogram equalization in that if the image contains areas with drastically different intensity distributions the global $\alpha$ and $\beta$ will not be ideal for all the regions in the image.

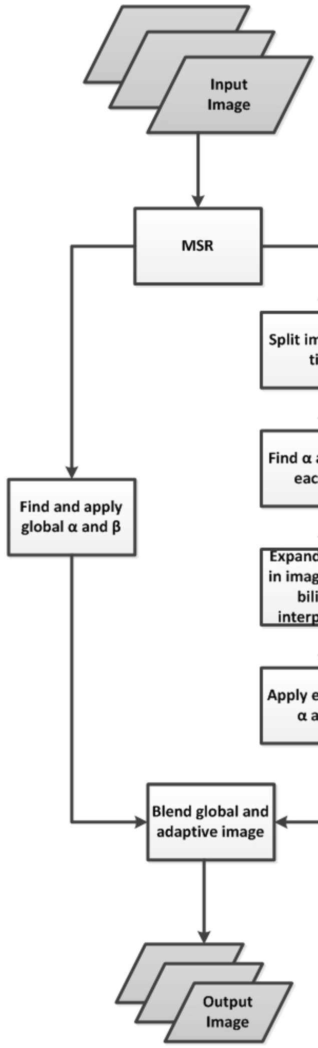



Fig. 2. Overview of proposed algorithm

## III. Proposed Algorithm

In this paper we offer an improvement over the classic formulation of the MSR algorithm. To improve the dynamic range compression of the algorithm without incurring the halo artifacts we propose using an adaptive approach to calculating the gain and offsets for the final stage of the algorithm and to blend these results with the those produced by the global calculation. The overview of our proposed algorithm is illustrated in Fig. 2.

Our approach draws from the adaptive techniques used in CLAHE [6]. The image is firstly divided into a set of tiles. The $\beta$ values are then found for each tile by calculating the minimum intensities. Next the $\alpha$ values are found for each tile by finding the difference between the maximum and minimum intensities. This process produces a 2-dimensional field of $\alpha$ and $\beta$ values the same size as the number of tiles selected. The next step is to expand the field of $\alpha$ and $\beta$ values to be the same size at the image. This is done using bilinear interpolation. This method is chosen because bilinear interpolation is cheap to calculate on the GPU. Once we have expanded the $\alpha$ and $\beta$ fields we will have values for each pixel of the MSR image. We can now apply the $\alpha$ and $\beta$ values to normalize the image. An example of the result of applying the adaptive $\alpha$ and $\beta$ values can be seen in Fig. 3.

As can be seen in Fig. 3 by applying the adaptive $\alpha$ and $\beta$ values we do achieve good dynamic range compression but in tiles where the image intensities are very uniform we end up drastically amplifying the noise in that tile. When calculating the global $\alpha$ and $\beta$ values it is very unlikely that the entire image will be a uniform intensity and as such we will not experience this over-gain. Thus we will not experience the same noise amplification we see when using purely adaptive values for $\alpha$ and $\beta$. Due to this problem we propose blending the outputs of the global gain/offset correction step and the adaptive gain/offset correction step to achieve a compromise between contrast enhancement and noise amplification.

To facilitate the blending of the Global and Adaptive MSR results we have to produce a blend map. We found that the full sized field of $\alpha$ values gives a good indication of how the two MSR images should be blended. Areas that


Fig. 3. Result after applying the adaptive $\alpha$ and $\beta$ values to the MSR image of the $Road$ input image which can be seen in figure Fig.7. The indicated region shows the over-gain problem experienced in areas of the image where the image intensities are very uniform.

require a very large gain usually are areas that are very uniform in intensity and as such areas that should contain a larger portion of the Global MSR output. Areas that required a low $\alpha$ value should contain a larger portion of the Adaptive MSR output. As such our blend map is produced by first normalizing the interpolated field of $\alpha$ values by dividing by the maximum $\alpha$ value which can be seen in Fig. 4.

Fig. 4. Example of a blend map for the $Road$ image

Once we have the normalized blend map we can combine the Adaptive MSR and Global MSR outputs as a weighted sum which can be seen in equation 7.

$$R_{MSR_B} = \phi \times R_{MSR_G} + (1 - \phi) \times R_{MSR_A} \quad (7)$$

where $\phi$ represents the normalised blend map image, $R_{MSR_G}$ represent the Global MSR image, $R_{MSR_A}$ the MSR Adaptive image, and $R_{MSR_B}$ the MSR blended image.

The final design decision we had to undertake was to select the number of scales, size of the scales and the weightings of the scales for the MSR algorithm. In [12] it is shown that 3 scales are sufficient to achieve good tonal rendition and dynamic range compression and this observation was confirmed in our experiments. Jobson et al. suggest standard deviations of 15, 80 and 250 for the scales used to enhance images under a megapixel in size. We found that these values produced good results but needed to be scaled for images of differing sizes for optimal results. It was also noted that a Gaussian kernel with a standard deviation of 250 is very large and almost encompasses an entire image with a VGA resolution. In the interest of reducing the amount of computation required for the algorithm instead of computing the surround function averages of the largest scale we considered them to be the mean value of the entire image. This can be computed efficiently and produces very similar results as using the large scale suggested in [12]. For the two smaller scales we used a basic heuristic, which we based on empirical testing, to choose the scale size based on the input image size. The standard deviation of the surround function for the smallest scale was considered to be 1.5% the size of the width of the image. The second scale was considered to be 5% the size of the width of the image. Finally we had to choose the weighing of scales and we found that while the best results were produced by heavily weighting the largest scale it was critical to have an element of the smaller scales in the algorithm output to enhance the contrast of small image structures. The weights we used for the smallest to largest scales were 0.2, 0.1 and 0.7 respectively. We leave the investigation of what the optimal scales and weightings are for future work.

## IV. Experimental Results

To demonstrate the performance of our proposed algorithm we have selected three images. The first is a HDR image and the final two are images that have been captured through atmospheric turbulence. The proposed algorithm will be compared to four traditional contrast enhancement techniques. Those techniques include Histogram Equalization (HE), Adaptive Histogram Equalization (AHE), Contrast Limited Adaptive Histogram Equalization (CLAHE), and traditional MSR. The results can be seen in Fig. 5, 6 and 7

In Fig. 5 we can see that our proposed Adaptive Multi-Scale Retinex (AMSR) algorithm gives the most pleasing results for this extreme HDR image. Much of the information in the dark areas on the left of the image that were originally hidden is revealed while also providing good contrast in the bright areas of the image. The HE result is very legible but is can be seen that there is saturation in the brightest and darkest areas in the image which is to be expected for a global approach. AHE gives a very strong contrast but is very noisy and unrealistic. CLAHE we found does not cope well with HDR images and even when the Clipping Limit is manually tuned we could not produce an image where neither the dark or light portions of the image were saturated. Global MSR does perform well for this image but as can be seen

AMSR achieves greater contrast, especially in the darkest and brights areas, while retaining realistic tonal rendition.

Fig.6 is a difficult image because it is exhibits a very low contrast and has a large proportion of areas of uniform intensity. In Fig.6 we see that global HE tends to produce unrealistic results and AHE gives strong contrast but is extremely noisy. CLAHE and MSR both produce decent and very similar results but AMSR manages to produce the best tonal rendition especially in the darker area on the left of the tower. For Fig.7 CLAHE produces better results than standard MSR but we can see that the proposed AMSR algorithm produces the best contrast enhancement consistently across the entire image. Again HE produces unrealistic results and AHE is extremely noisy due to the large uniform regions.
It is interesting to look at the histograms of the images in Fig.7 which can be seen in Fig.8. It is apparent that our AMSR algorithm produces the histogram with the smoothest and widest spread without resulting in saturation at the black or white bounds of the histogram. The smooth histogram produced by AMSR captures the same peaks that can be seen in the histogram of the original image and distributes them very neatly across the intensity range resulting in a high contrast output that appears natural to a human viewer. Unfortunately there are no empirically-based metrics in the literature that have been able to objectively and reliably measure the perception of the contrast of complex real world images by a human observer, however work is being performed to develop such a metric based on the survey of a large sample of human observers [16]. In this paper we employed the classic information metric of entropy [1] as an attempt to quantitatively measure the quality gain the algorithms produce, table I shows these results. Firstly we can see the problem with using these sorts of metrics in the results for the AHE outputs. These images are extremely noisy and the metric perceives the noise as large amounts of information even though noise is not perceived as useful to a human observer. We can however see that for the CLAHE, MSR and AMSR results we get a useful comparison. In the HDR image $Shadow$ MSR out performs CLAHE but AMSR gives the most information gain. In $Road$ and $Tower$ we can see that CLAHE produces more information than MSR but AMSR beats CLAHE in both cases.

| | Original | HE | AHE | CLAHE | MSR | AMSR |
|---|---|---|---|---|---|---|
| Tower | 5.95 | 5.4806 | 7.8496 | 6.5591 | 6.3629 | 6.7471 |
| Road | 6.4857 | 5.8594 | 7.993 | 7.3181 | 7.0171 | 7.4123 |
| Shadow | 6.0406 | 5.0049 | 7.8669 | 6.8229 | 7.4627 | 7.7254 |

TABLE I
ENTROPY TEST RESULTS

The AMSR Algorithm was implemented for the GPU using OpenGL. The algorithm was run on a desktop computer with the specifications show in table II. For comparison we used a GPU implementation of the CLAHE algorithm which uses scattering to produce histograms and is discussed in [17] and the source code can be found [18]. The AMSR and CLAHE implementations were run using the $Tower$ and $Road$ images found in Fig.6 and 7. The results are shown in table III. As can been seen the AMSR algorithm runs faster than the CLAHE algorithm by almost an order of magnitude. This is because the AMSR algorithm is based on a series of basic kernel convolutions and does not require the awkward implementation of the histogram that is required in CLAHE.

| | |
|---|---|
| CPU | Intel Core I7-2600k 3.4 GHz Processor |
| RAM | 8 GB DDR3 RAM |
| GPU | nVidia GTX 580 graphics card |

TABLE II
SPECIFICATION OF THE DESKTOP COMPUTER USED IN THE PERFORMANCE TESTS

| | 560x460 resolution | 876x592 resolution |
|---|---|---|
| CLAHE | 30 fps | 14 fps |
| AMSR | 296 fps | 131 fps |

TABLE III
PERFORMANCE TEST RESULTS

## V. CONCLUSION

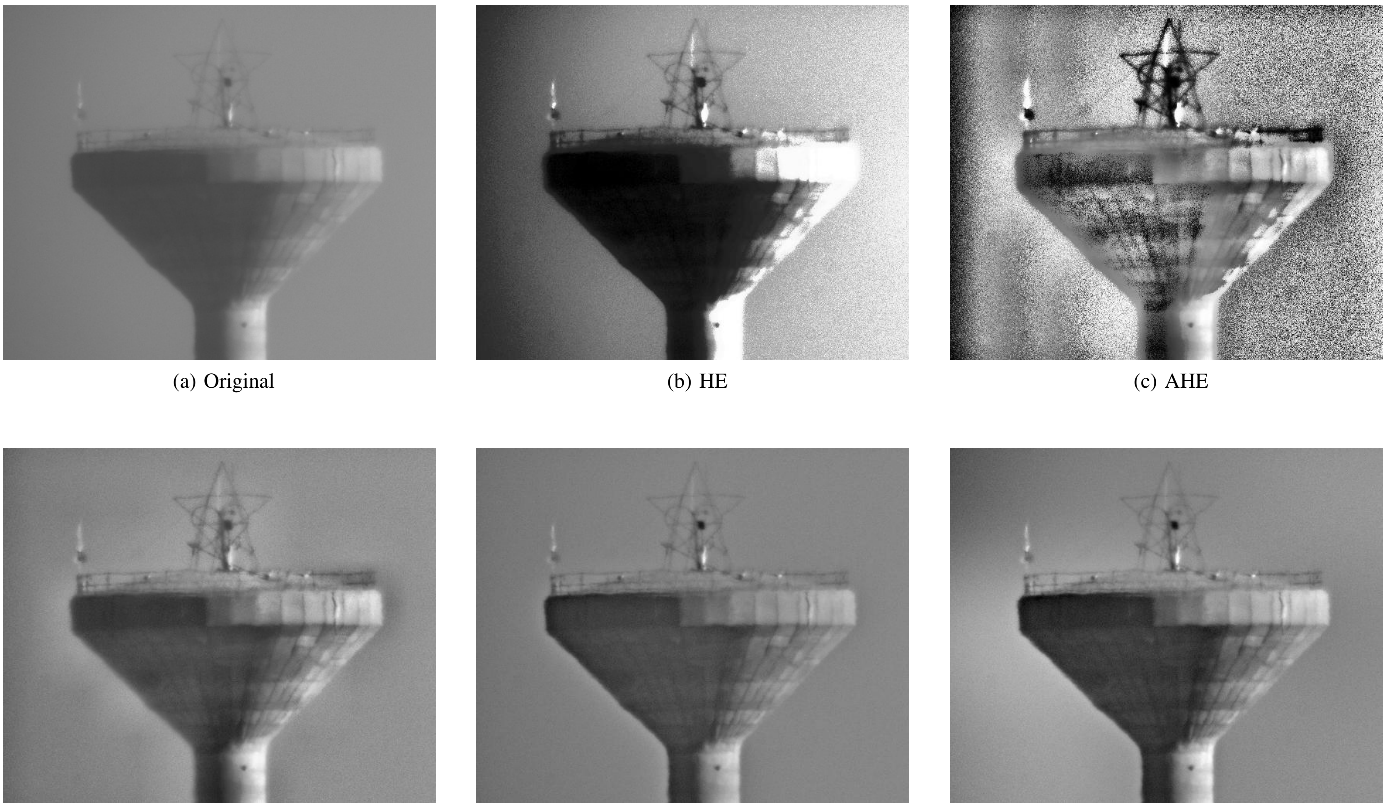
Contrast enhancement is a classic image restoration technique that has been employed to improve the legibility of images and the information they contain since the times of analog image capture. The traditional approach to digital contrast enhancement is to employ a form of histogram equalization. While this approach does improve contrast it often produces an unrealistic and saturated effect which is very apparent when applied to real-world scenes. This paper explores the use of Retinex theory for the purpose of contrast enhancement. An overview of Retinex theory and its use as a digital image processing technique in the form of the Single-Scale and Multi-Scale Retinex algorithms is provided.

This paper proposes an improvement to the traditional global Multi-Scale Retinex algorithm which allows it to improve its dynamic range compression while not producing the traditional artifacts associated with Retinex based methods. The Adaptive Multi-Scale Retinex algorithm makes use of a model of how our eyes naturally perceive scenes and as such the output of the algorithm looks very natural to a human viewer. The experimental results show that for real-world images AMSR produces slightly better results than CLAHE which is currently the most versatile contrast enhancement algorithm in the literature. Our Adaptive Multi-Scale Retinex algorithm is also well suited to implementation on the GPU and achieves speeds around 10 times faster than a GPU implementation of CLAHE as AMSR is based on simple kernel convolutions and does not require the awkward GPU implementation of the histogram.

(a) Original (b) HE (c) AHE

(d) CLAHE (e) MSR (f) proposed

Fig. 5. Contrast enhancement results for the HDR $Shadow$ [19] image



(a) Original (b) HE (c) AHE

(d) CLAHE (e) MSR (f) proposed

Fig. 6. Contrast enhancement results for image $Tower$ which has been captured through atmospheric turbulence

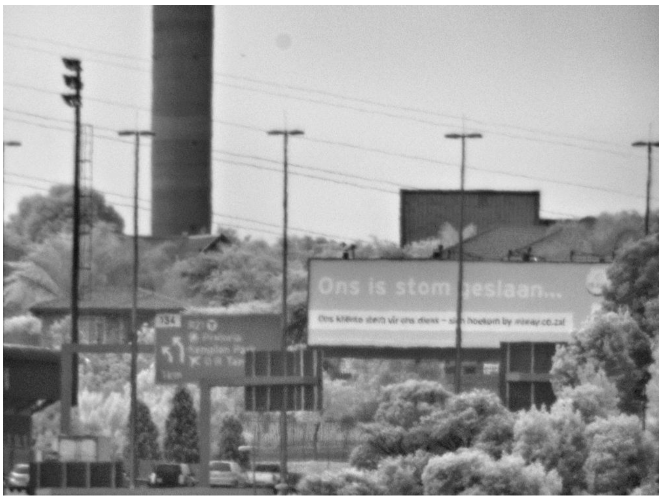
(a) Original

(b) HE

(c) AHE

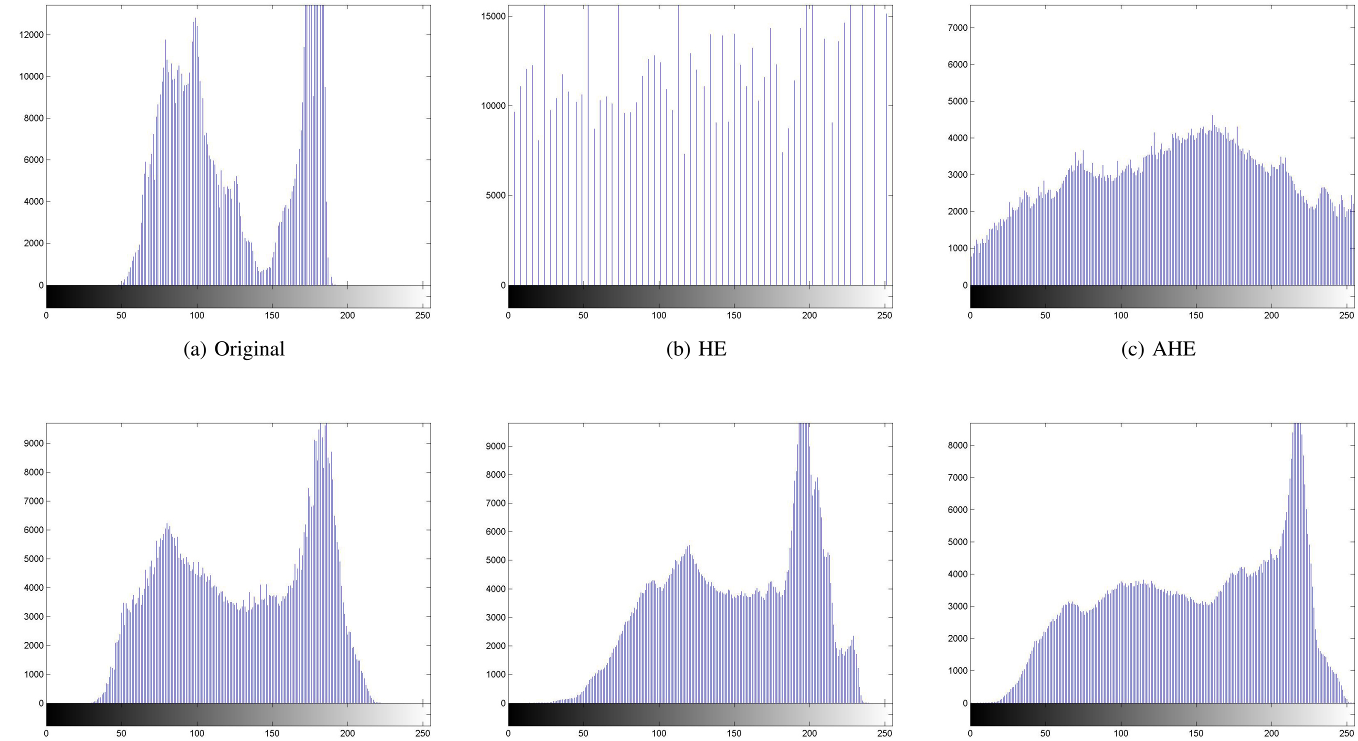
(d) CLAHE

(e) MSR



(f) proposed

Fig. 7. Contrast enhancement results for image *Road* which has been captured through atmospheric turbulence



(a) Original (b) HE (c) AHE

(d) CLAHE (e) MSR (f) proposed

Fig. 8. Histograms for Fig. 7 *Road.*

## References


[1] R. Gonzales and R. Woods, *Digital Image Processing*. Pearson Prentice Hall, third ed., 2008.

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

[3] P. E. Robinson and W. A. Clarke, "Sharpening and contrast enhancement of atmospheric turbulence degraded video sequences," *Proceedings of the Twenty-First Annual Symposium of the Pattern Recognition Association of South Africa*, 2010.

[4] J. Zimmerman, S. Pizer, E. Staab, J. Perry, W. McCartney, and B. Brenton, "An evaluation of the effectiveness of adaptive histogram equalization for contrast enhancement," *Medical Imaging, IEEE Transactions on*, vol. 7, pp. 304 –312, dec 1988.

[5] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing*, vol. 39, pp. 355–368, Sept. 1987.

[6] A. M. Reza, "Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement," *The Journal of VLSI Signal Processing-Systems for Signal, Image, and Video Technology*, vol. 38, pp. 35–44, Aug. 2004.

[7] D. Menotti, L. Najman, J. Facon, and A. de Araujo, "Multi-histogram equalization methods for contrast enhancement and brightness preserving," *Consumer Electronics, IEEE Transactions on*, vol. 53, pp. 1186 –1194, aug. 2007.

[8] Y.-T. Kim, "Contrast enhancement using brightness preserving bi-histogram equalization," *Consumer Electronics, IEEE Transactions on*, vol. 43, pp. 1 –8, feb 1997.

[9] E. H. LAND and J. J. McCANN, "Lightness and retinex theory," *J. Opt. Soc. Am.*, vol. 61, pp. 1–11, Jan 1971.

[10] E. H. Land, "An alternative technique for the computation of the designator in the retinex theory of color vision.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 83, pp. 3078–80, May 1986.

[11] D. J. Jobson, Z. Rahman, and G. a. Woodell, "Properties and performance of a center/surround retinex.," *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, vol. 6, pp. 451–62, Jan. 1997.

[12] D. J. Jobson, Z. Rahman, and G. a. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes.," *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, vol. 6, pp. 965–76, Jan. 1997.

[13] B. Funt, F. Ciurea, and J. Mccann, "Retinex in matlab," in *Journal of Electronic Imaging*, pp. 112–121, 2000.

[14] K. Barnard and B. Funt, "Investigations into multi-scale retinex," in *Color Imaging in Multimedia*, pp. 9–17, Technology, (Wiley, 1999.

[15] E. H. Adelson, "Checker shadow illusion." http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html. Accessed: 21/09/2012.

[16] A. Haun and E. Peli, "Measuring the perceived contrast of natural images," *SID Symposium Digest of Technical Papers, Session 24: Visual Perception (APV)*, pp. 302–304, 2011.

[17] T. Scheuermann and J. Hensley, "Efficient histogram generation using scattering on gpus," *Proceedings of the 2007 ACM Symposium on interactive 3D Graphics and Games*, 2007.

[18] T. Scheuermann, "Uniform and adaptive histogram equalization." http://sebastien.hillaire.free.fr/index.php?option=com_content&view=article&id=59&Itemid=70. Accessed: 11/11/2012.

[19] Z. Doob, "Siena shadow." http://www.photoflavor.com/index.php?id=477. Accessed: 21/09/2012.