

Quantitative analysis of infrared contrast enhancement algorithms

Seth Weith-Glushko¹ and Carl Salvaggio²

Rochester Institute of Technology, 1 Lomb Memorial Drive, Rochester, NY, USA 14623

ABSTRACT

Dynamic range reduction and contrast enhancement are two image-processing methods that are required when developing thermal camera systems. The two methods must be performed in such a way that the high dynamic range imagery output from current sensors are compressed in a pleasing way for display on lower dynamic range monitors. This research examines a quantitative analysis of infrared contrast enhancement algorithms found in literature and developed by the author. Four algorithms were studied, three of which were found in literature and one developed by the author: tail-less plateau equalization (TPE), adaptive plateau equalization (APE), the method according to Aare Mällo (MEAM), and infrared multi-scale retinex (IMSR). TPE and APE are histogram-based methods, requiring the calculation of the probability density of digital counts within an image. MEAM and IMSR are frequency-domain methods, methods that operate on input imagery that has been split into components containing differing spatial frequency content. After a rate of growth analysis and psychophysical trial were performed, MEAM was found to be the best algorithm.

Keywords: CONTRAST ENHANCEMENT, INFRARED, REAL-TIME, THERMAL, ALGORITHM

1. INTRODUCTION

Remote sensing is defined as “the field of study associated with extracting information about an object without coming into physical contact with it.” [10] Remote sensing is important as a process because it allows users to obtain information about phenomena that would be dangerous or impossible for them to detect solely with their senses. The process can be modeled as a chain. In this model, each segment of an imaging system is broken into individual chains: the input link, the processing link, and the display link. A task of the remote sensing professional is to understand how each link in the chain fits together and to avert any problems arising from the interaction between each link. One common problem is data reduction. For example, a detector (input) is able to output pixels that have a dynamic range described by twelve bits. In the same system, the monochrome display (output) is able to output pixels that have a dynamic range of only eight bits. Hence, a procedure aimed at reducing the data must take place in the processing stage to enable the display to work with data from the detector. This procedure must accomplish two goals: reduce the dynamic range of the input image into an image that is acceptable for input by the output system and do this in such a manner that the output image is pleasing to the human observer.

One such procedure that can be employed is dynamic range compression. Dynamic range compression can be defined as the mapping of pixels containing a high dynamic range to pixels that contain a reduced dynamic range. In essence, dynamic range compression is a pixel operator, defining the value of an arbitrarily located pixel in a new image by using the value of the corresponding pixel in the original image. Dynamic range compression has a number of applications in fields such as video telephony [1], radiology [6, 11], and high dynamic range photography [7, 9]. As such, research has been performed with the aim of providing a dynamic range compression algorithm that suits an application area’s needs, such as enhanced image quality, heightened information availability, or real-time processing capability. Although research has been performed, each study has shown a lack of quantitative metrics to describe how well each algorithm performs in terms of image quality or algorithmic efficiency. In most cases, all that is offered is a simple qualitative metric. This aim of this research was to define, and quantitatively evaluate infrared contrast enhancement algorithms for use in a real-time long-wave infrared imager.

2. ALGORITHMS

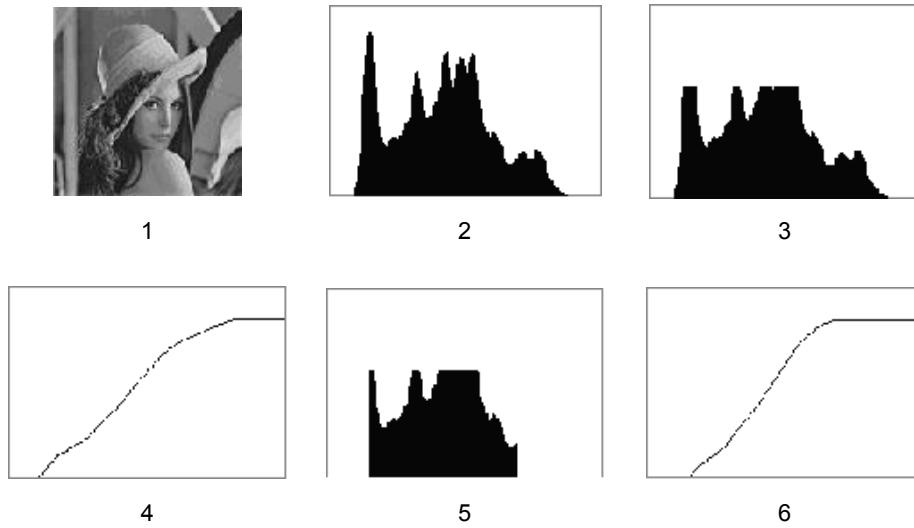
¹ Email: sethwg@sethwg.com

² Email: salvaggio@cis.rit.edu

At the heart of this effort is the development of the infrared contrast enhancement algorithms. In general, infrared contrast enhancement algorithms are designed to take high dynamic range input infrared imagery and output low dynamic range display imagery. Throughout this paper, the input image will be called $f(x,y)$ and the output image will be called $g(x,y)$. In all, there are four algorithms to be studied: tail-less plateau equalization (TPE), adaptive plateau equalization (APE), the method according to Aare Mällo (MEAM), and infrared multi-scale retinex (IMSR).

2.1. Tail-less Plateau Equalization (TPE)

Tail-less plateau equalization is a variation on histogram equalization where a maximum gain parameter, called the plateau, is introduced. The plateau is a clipping value that is applied to a histogram, placing a limit on the number of pixels that can be resident within each histogram bin. The purpose of the plateau is to lessen the chance for excessive contrast enhancement. It does so by making the lookup table more linear, increasing the probability that all possible input pixel values will be present in the output image. An overview of the process can be seen in Figure 5.



1. An image is selected
2. The image's histogram is generated
3. The histogram is clipped to a pre-defined parameter
4. A CDF is created from the clipped histogram
5. Using the clipped CDF, the leading and trailing tails of the histogram are zeroed
6. A CDF is created from the tail-less histogram and is used in the mapping

Figure 1 – Tail-less plateau equalization

To start, the histogram, $H(k)$, of the input image, $f(x,y)$, is calculated. In plateau equalization, a maximum gain parameter, P_{max} , is introduced and a new histogram, $H_p(k)$, is calculated

$$H_p(k) = \begin{cases} H(k) & H(k) \leq P_{max} \\ P_{max} & H(k) > P_{max} \end{cases} \quad (1)$$

Using this modified histogram, the PDF and CDF of $H_p(k)$ is calculated. Next, the “tails” of the modified histogram are eliminated forming a new histogram $H_t(k)$ as described in Equation 2. By removing the tails of the histogram, outlier pixels can be forced into saturation, increasing the contrast in the output image. t_{max} is a value between zero and one-half and represents the percentage of pixels one wishes to remove from the head and tail-end of the histogram. $CDF_p(k)$ represents the CDF of $H_p(k)$.

$$H_t(k) = \begin{cases} H_p(k) & CDF_p(k) \in [t_{max}, 1 - t_{max}] \\ 0 & otherwise \end{cases} \quad (2)$$

The PDF and CDF of the new histogram $H_t(k)$ are calculated and an equalization lookup table is generated and applied to each pixel globally to form the output image $g(x,y)$ [4].

2.2. Adaptive Plateau Equalization (APE)

Adaptive plateau equalization is very similar to tail-less plateau equalization in the sense that a plateau is applied to a histogram before the mapping function is determined. However, instead of having the maximum gain parameter fixed, it is adapted to the current histogram of the scene. First, the histogram $H(k)$ is taken of the scene and its corresponding cumulative density, $CDF(k)$, is determined. Second, the values defined in Table 1 are defined.

Table 1 – Important values for adaptive plateau equalization

Value	Description	Found By
I_{\min}	The greyscale that corresponds to the first histogram bin with a value greater than zero	The first greyscale k where $CDF(k) > 0$
I_{\max}	The greyscale that corresponds to the last histogram bin with a value greater than zero	The last greyscale k where $CDF(k) > 0$
$I_{1\%}$	The greyscale that corresponds to the location in the CDF that is equal to .01	The greyscale k that satisfies $CDF(k) = 0.01$
$I_{99.9\%}$	The greyscale that corresponds to the location in the CDF that is equal to .999	The greyscale k that satisfies $CDF(k) = 0.999$
$I_{99.99\%}$	The greyscale that corresponds to the location in the CDF that is equal to .9999	The greyscale k that satisfies $CDF(k) = 0.9999$
$I_{25\%}$	The greyscale that corresponds to the location in the CDF that is equal to .25	The greyscale k that satisfies $CDF(k) = 0.25$
$I_{75\%}$	The greyscale that corresponds to the location in the CDF that is equal to .75	The greyscale k that satisfies $CDF(k) = 0.75$
$I_{\inf,a}$	The first inflection point of the histogram	See below
$I_{\inf,b}$	The last inflection point of the histogram	See below
η_A	The number of pixels with a value less than $I_{\inf,a}$	$\eta_A = \sum_{i=I_{\min}}^{I_{\inf,a}-1} H(i)$
η_B	The number of pixels with a value greater than $I_{\inf,b}$	$\eta_B = \sum_{i=I_{\inf,b}+1}^{I_{\max}} H(i)$

As one can see from Table 1, almost all of the values needed for the plateau algorithm can be found using the CDF. It should be noted that the values of the CDF are not likely to match the limits specified in Table 1 exactly. Hence, the greyscale value that corresponds to the CDF value that is closest to the target value will be used. $I_{\inf,a}$ and $I_{\inf,b}$ are not derived from the CDF but from the shape of the histogram. These two inflection points are found by applying a moving window sum of width w across the histogram. The low ($I_{\inf,a}$) and high ($I_{\inf,b}$) inflection points correspond to intensities where the moving window sums change from their previous values by a threshold amount based on a fraction of the difference between I_{\min} and I_{\max} , as seen in Equation 3. ΔI represents the threshold amount and ε represents a scalar value that ranges between zero and one.

$$\Delta I = \varepsilon(I_{\max} - I_{\min}) \quad (3)$$

Limitations are placed on the high and low inflection points. These limitations are described in Equations 4 and 5.

$$I_{\inf,a} = \{I_{\inf,a} = k \in I_{\inf,a} < I_{25\%}\} \quad (4)$$

$$I_{\inf,b} = \{I_{\inf,b} = k \in (I_{\inf,b} > I_{75\%}) \& (I_{\inf,b} < \frac{I_{\max} - I_{\min}}{2})\} \quad (5)$$

After the inflection points have been defined, η_A and η_B can be found by using the appropriate entry in Table 1. The maximum gain parameter can be calculated next, but first a number of intermediate values need to be

determined. The first is the ratio of pixels that occupy that central portion of the histogram versus the tails of the histogram. This ratio is defined as Equation 6. X represents the ratio while N represents the total number of pixels in the image.

$$X = \frac{N - (\eta_A + \eta_B)}{\eta_A + \eta_B} \quad (6)$$

After the ratio has been calculated, the nominal plateau value can be computed as

$$P_{nom} = \begin{cases} \frac{X\eta_B}{I_{inf,b} - I_{inf,a}} & I_{99.99\%} - I_{inf,b} > I_{inf,a} - I_{1\%} \\ \frac{X\eta_A}{I_{inf,b} - I_{inf,a}} & I_{99.99\%} - I_{inf,b} \leq I_{inf,a} - I_{1\%} \end{cases} \quad (7)$$

Next, the dynamic range of the scene (R_D) is determined

$$R_D = I_{99.9\%} - I_{1\%} \quad (8)$$

Using this dynamic range metric, a dynamic range adjustment factor can be calculated using Equation 9. F_{DR} represents the dynamic range adjustment factor while L_{max} represents the maximum grey level output after processing.

$$F_{DR} = \begin{cases} 1 - \frac{L_{max}}{R_D} & R_D > L_{max} \\ 1 - \frac{R_D}{L_{max}} & R_D \leq L_{max} \end{cases} \quad (9)$$

In addition to the dynamic range adjustment factor, another adjustment factor, is calculated to create a more natural appearance of extended dark regions whose intensities are greater than $I_{75\%}$. The adjustment factor F_{ED} is computed as

$$F_{ED} = 1 - \frac{I_A - I_{0.1\%}}{I_B - I_{0.1\%}} \quad (10)$$

The actual gain parameter (P_A) that will be used to perform plateau equalization can be calculated using Equation 11. A requirement will be placed upon this gain parameter that the value must be greater than or equal to one [5].

$$P_A = P_{nom} \cdot F_{DR} \cdot F_{ED} \quad (11)$$

Due to the ever-changing nature of infrared imagery, it is possible that the adaptive plateau value can change in value greatly from image to image. To make this algorithm suitable for video, a temporal lowpass infinite impulse response (IIR) filter has been applied to the plateau value. This entails taking a previously calculated plateau value $P_{A,previous}$ and forming a new plateau value $P_{A,filtered}$ using Equation 12. $P_{A,current}$ represents the plateau value calculated for the current image.

$$P_{A,filtered} = \psi P_{A,current} + (1 - \psi) P_{A,previous} \quad (12)$$

where ψ represents a scalar arbitrary set to a value between zero and one. By changing the value of ψ , the amount of a current scene's contribution can be minimized. Hence, flicker can be reduced in the resulting video stream.

Once the gain parameter has been calculated, a new histogram is formed based on the original histogram [5] where $H_p(k)$ represents the new plateau histogram.

$$H_p(k) = \begin{cases} H(k) & H(k) \leq P_F \\ P_A & H(k) > P_F \end{cases} \quad (13)$$

A CDF is calculated using the plateau histogram that defines the mapping function for pixels in the output image $g(x,y)$ [4].

2.3. Method According to Aare Mällo (MEAM)

The method according to Aare Mällo is a departure from performing histogram equalization as the sole dynamic range compression and contrast enhancement instrument. MEAM separates an input image $f(x,y)$ into high spatial frequency and low spatial frequency component images. By operating on the two component images separately,

a higher degree of control over the dynamic range compression and contrast enhancement can be realized. An overview of the algorithm can be seen in Figure 7.

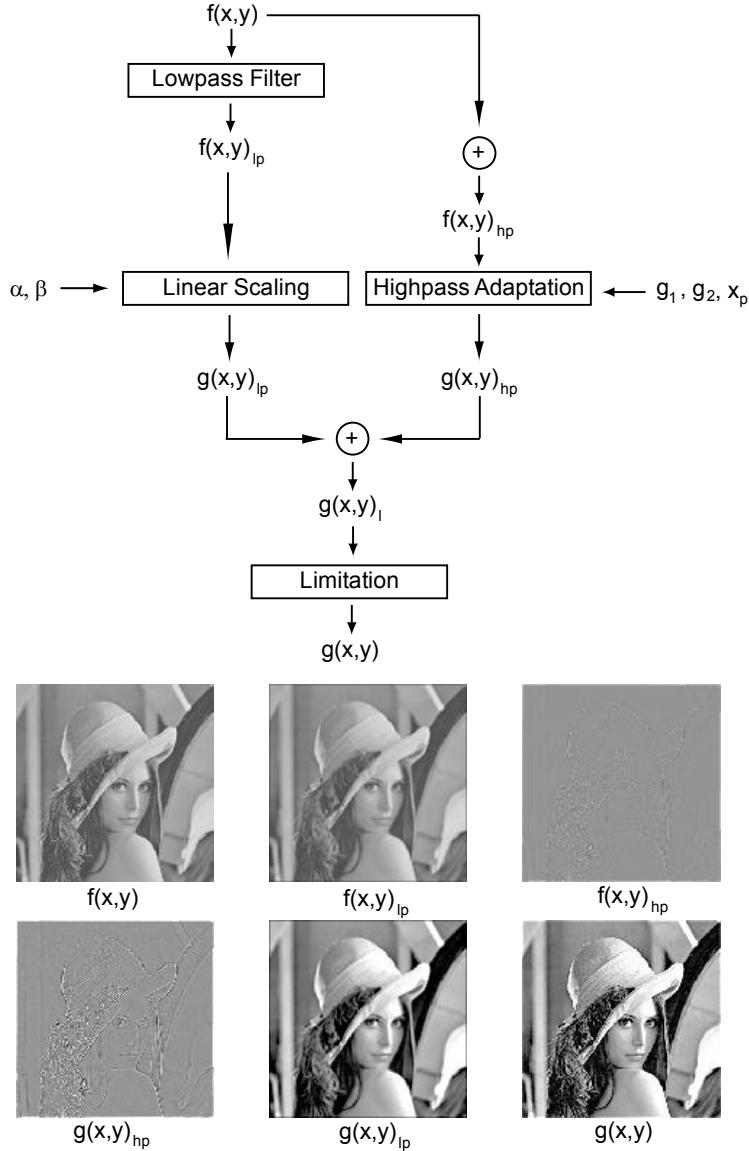


Figure 2 - Overview of MEAM

In this algorithm, the input image $f(x,y)$ is convolved with a lowpass filter, forming a low-pass image ($f(x,y)_{lp}$). This filter is most often a mean value spatial filter. The high-pass image ($f(x,y)_{hp}$) is formed by subtracting the low-pass image from the original image. A gain, determined by Equation 14, is applied to the high-pass image. x_p represents an arbitrary pixel value, above which one gain value (g_2) is applied, and, below which a different gain value (g_1) is applied. This value is specified such that the edge information pixels, which usually have a small value, are enhanced by the first gain parameter while pixels containing noise information, which usually have a larger value, are attenuated by the second gain parameter [3].

$$g(x,y)_{hp} = \begin{cases} g_1 f(x,y)_{hp} & |f(x,y)_{hp}| < x_p \\ g_2 f(x,y)_{hp} & |f(x,y)_{hp}| \geq x_p \end{cases} \quad (14)$$

To the low-pass image, a number of histogram enhancement processes can be performed. For simplicity in this description, a linear scaling will be performed. The transformed low-pass and high-pass images are summed (forming $g(x,y)_l$) and the resultant values are limited between a specified range, forming the final image $g(x,y)$ [5].

2.4. Infrared Multi-Scale Retinex (IMSR)

The infrared multi-scale retinex is an offshoot of a neuro-physiological model called the Retinex calculation. The Retinex calculation “was perceived as a model of the lightness and color perception of human vision.” [8] This model was based on the receptive field structures found within the human visual system. By using a specifically designed spatial filter called a surround, the lateral opponent operation of the human visual system could be mimicked. Later research showed that the model could be applied to existing imagery to enhance it in a way that the eye would find aesthetically pleasing. Expounding on this research, Rahman found that by using this model with a Gaussian surround, existing imagery could be improved. Rahman called this process the “multi-scale retinex.” Through this research, the Retinex calculation has been extended to apply to the thermal infrared modality. The infrared multi-scale retinex can be applied as Equation 15.

$$g(x,y) = F \left(\sum_{s=1}^S W_s \left[G \left(\frac{f(x,y)}{F_s(x,y) * f(x,y)} \right) \right] \right) \quad (15)$$

$f(x,y)$ represents the input image while $g(x,y)$ represents the output image. F and G are normalizing functions:

$$G(z(x,y)) = \rho \frac{z(x,y) - \min(z(x,y))}{\max(z(x,y)) - \min(z(x,y))} \quad (16)$$

$$F(z(x,y)) = L_{\max} \frac{z(x,y) - \min(z(x,y))}{\max(z(x,y)) - \min(z(x,y))} \quad (17)$$

ρ represents the maximum bit depth one wishes to have the ratio image to have. Through experimentation, acceptable values of ρ lie between 2^{10} and 2^{14} . By increasing ρ , one increases the number of bits of relevant edge information and as such, increases the chance of that edge appearing in the final image. S represents the number of surround functions one wishes to use to perform the Retinex calculation, L_{\max} represents the maximum grey value possible in the output image, $I(x,y)$ represents the input image rescaled from zero to one at location (x,y) while W_s represents an arbitrary numerical weight associated with the s^{th} Gaussian surround function defined by F_s . F_s is defined as:

$$F_s(x,y) = \kappa e^{-\frac{(x^2+y^2)}{\sigma_s^2}} \quad (18)$$

σ_s^2 represents the variances of the different scales one uses in the calculations. κ represents a numerical weighting that can be calculated as seen in Equation 19.

$$\kappa = \frac{1}{\sum_x \sum_y F(x,y)} \quad (19)$$

Rahman found that using multiple surrounds is necessary to achieve a balance between dynamic range compression and correct tonal rendition.

3. METHODS

The quantitative analysis was performed in distinct stages. The first stage was the development of engineering code representative of the four infrared contrast enhancement algorithms. The engineering code was developed in the MATLAB development environment. The purpose of this engineering code was to take digital frames that contain 16-bit pixels as input and output 8-bit imagery using an algorithm of choice for image enhancement. This input imagery took the form of single or multiple frames and output images or video, respectively. Additionally, this engineering code has the ability to run a “Big O” notation rate of growth analysis as a measure of algorithmic efficiency. The analysis was run on an arbitrarily resized example infrared image. To facilitate use of the algorithms, a graphical user interface was developed for each algorithm.

The second stage involved the collection of digital frames captured from the infrared imager. To do so, a digital frame grabbing station was constructed. The station consisted of three components: an infrared imager, capturing equipment, and a computer to control the capturing equipment. Using this setup, digital frames containing pixels output from the focal plane array on the camera were generated. The imager was used in a variety of conditions that one can expect it to perform in. For this analysis, three scenes were selected: two artificial and one natural. It should be noted that the artificial scenes contain natural items but it was specifically designed to exhibit an arbitrary range of infrared behavior. To confirm the temperature range in the scene, an infrared thermometer was used to record temperatures of objects with known emissivities.

The first scene involved a frame sequence where the imager captures a scene containing low-contrast objects (the thermal differential across the scene is less than one degree Fahrenheit) and then pans to a scene containing high-contrast objects (the thermal differential across the scene is greater than twenty degrees Fahrenheit). For this scene, a lit hibachi grill was used as the high contrast target while foliage, shortly after a rain event, provided the low-contrast scene elements. The second scene involved a frame sequence where the imager captures a scene containing low-contrast objects and then pans to a scene containing mid-contrast objects (the thermal differential across the scene is greater than one degree but less than twenty degrees Fahrenheit). For this scene, the hibachi grill was doused to provide the mid-contrast target while freshly-wetted foliage provided the low-contrast target. The final scene involved a frame sequence where the imager captured a scene containing natural elements. It should be noted that the thermal differential across this scene was not noted. For this sequence, the imager was pointed at an interstate interchange with full view of objects such as vehicles, roads, light posts, and foliage. Frames were recorded as vehicles passed on the interstate. Visible still imagery from the sequences can be seen in Figure 3.



Figure 3 – Visible imagery of the backyard scene (foliage and grill) and the natural scene

The third stage of the analysis involved the execution of paired-comparison psychophysical trials where test subjects were asked to evaluate processed video from the four different algorithms studied using the three scenes captured, resulting in 18 pairwise comparisons to observe. For this research, the test subjects were separated into two groups: members of the Rochester City Fire Department and students from the Imaging Science program at the Rochester Institute of Technology. Since firefighters are one of the principal users of the infrared imager, their opinions are extremely important to the sponsor of this research. As a basis for comparison, the trial was also open to imaging science students. For each of the 18 videopairs, observers were asked to answer two questions: “Of the two videos, select which one has the best quality” and “Of the two videos, select which one has the most detail.” By asking the observer these specific questions, it was hoped that quantitative evidence would be generated to support the hypothesis that there is no difference between quality and detail between algorithms and computational considerations could be the deciding factor for the comparisons. To facilitate data collection and analysis, a MATLAB program was created to automate the process of data collection and the application of Thurstone’s Case V solution [2]. Using this program, a calculation of interval scale values was generated for each unique combination of image set and question.

4. RESULTS

4.1. Subjective Image Analysis

After image collection, the first analysis that was carried out was a subjective analysis of the processed captured imagery. According to the previously described method, three scenes were collected: a low-to-high contrast scene, a low-to-mid contrast scene, and a natural scene. The low-to-high contrast scene was gathered by placing the infrared camera in a residential backyard that contained a fence, grass, trees, and flowers. The scene also contained a small hibachi grill that contained a small wood fire. As the video stream starts, the camera is facing a tree and some large bushes. As time progressed, the camera panned to the left, allowing the lit grill to enter the field of view. Finally,

the camera panned right, returning to the original scene. Results of infrared contrast enhancement on frames containing the large tree and the lit grill can be seen in Figure 4.

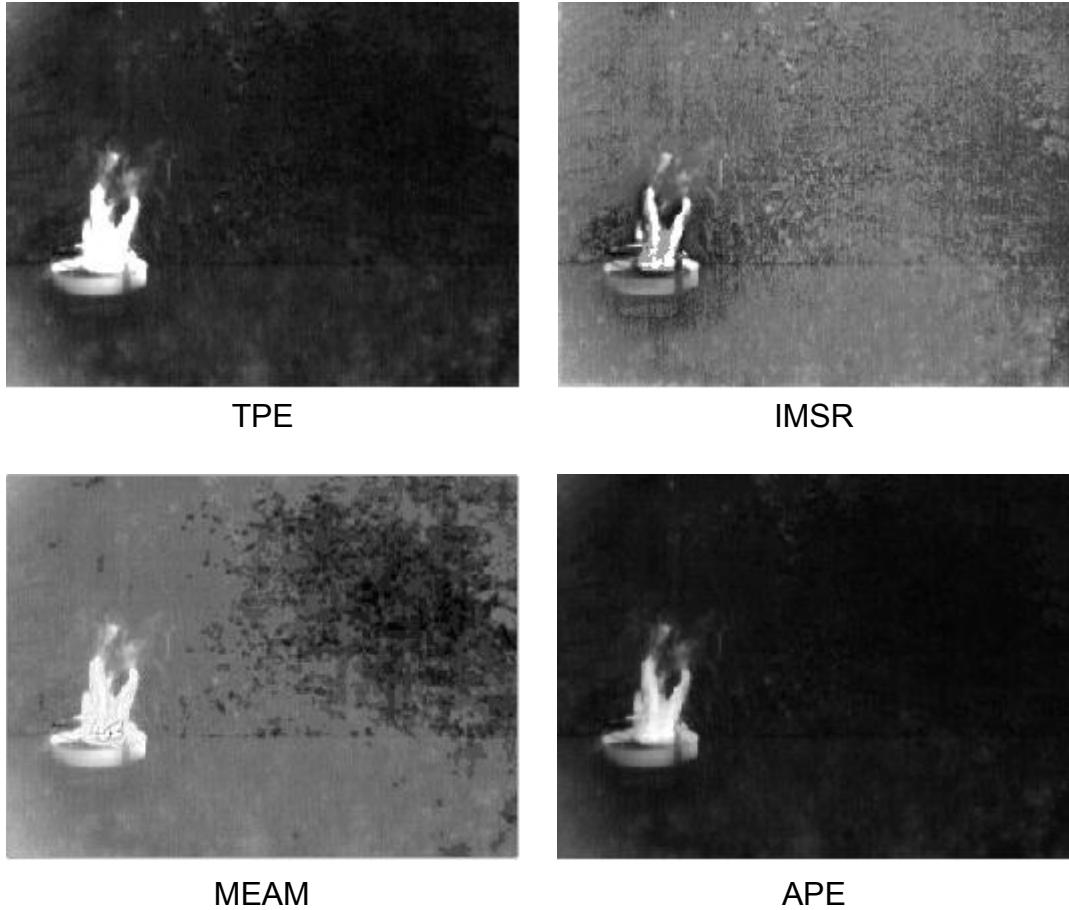


Figure 4 – Processed imagery of the low-to-high contrast scene containing the lit grill

As one can see from Figure 4, IMSR provides an image that shows structure in both the hot fire and the cool background. MEAM shows a bit less detail in the background while retaining a fair amount of detail in the fire. APE loses much of the detail in the background while retaining some in the flames of the fire itself. Finally, TPE loses the most detail in the background and a fair amount in the flames of the fire.

The low-to-mid contrast scene was very similar to the low-to-high contrast scene. Every component of the video stream was the same as the low-to-high contrast video with the exception that the hibachi grill was cooled after dousing the wood fire. Results of infrared contrast enhancement on frames containing the large tree and the unlit grill can be seen in Figure 5.

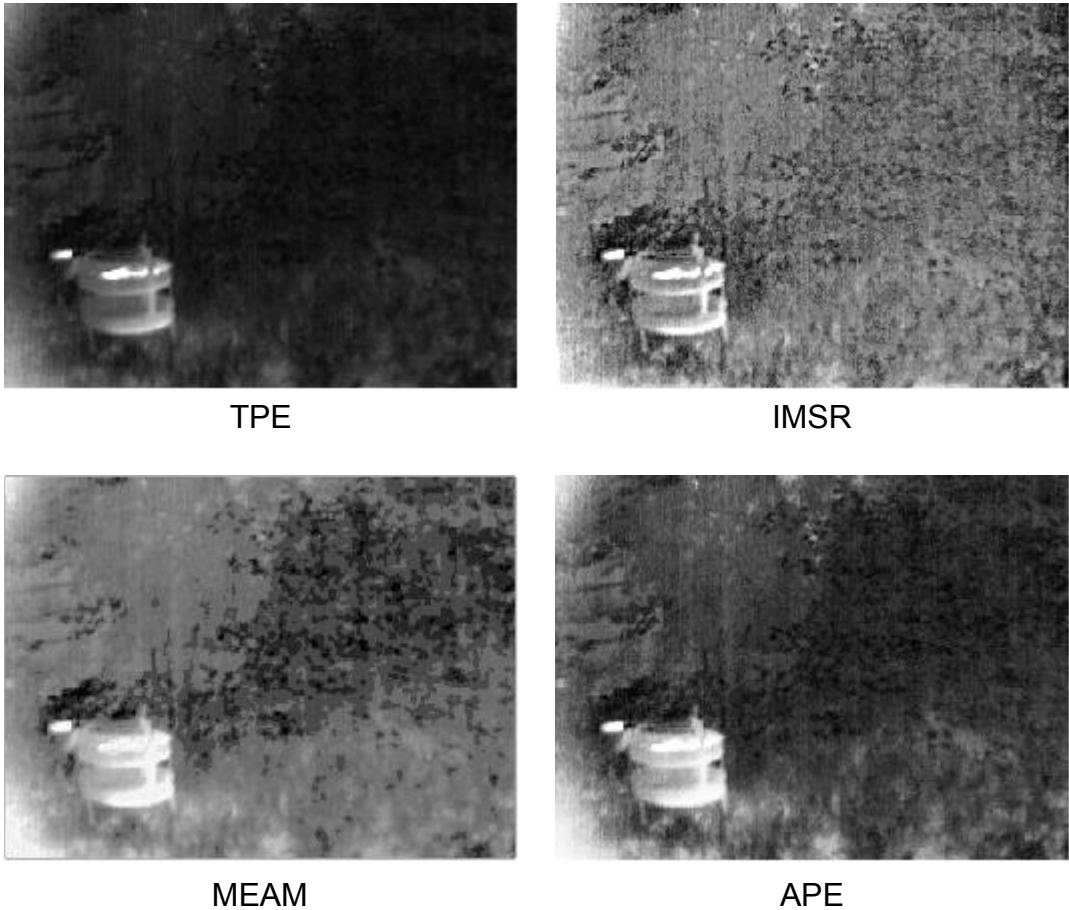


Figure 5 – Processed imagery of the low-to-mid contrast scene containing the unlit grill

As one can see in Figure 5, some differences become apparent when compared to Figure 4. IMSR provided an image that has the most overall contrast; and structure can be seen in the grill, fence, and surrounding foliage. MEAM also has a significant amount of contrast overall but some detail was lost in the grass and tree. APE continues this trend as more detail was lost in the foliage. Finally, TPE loses much of the background contrast.

The natural scene was collected by situating the infrared camera in front of the intersection of East Henrietta Road and Interstate 390 in Brighton, a township located just outside downtown Rochester. While recording a video stream, the camera was kept stationary, allowing the motion of the vehicles on the roadways to provide changing infrared content. The results of image processing can be seen in Figure 6.

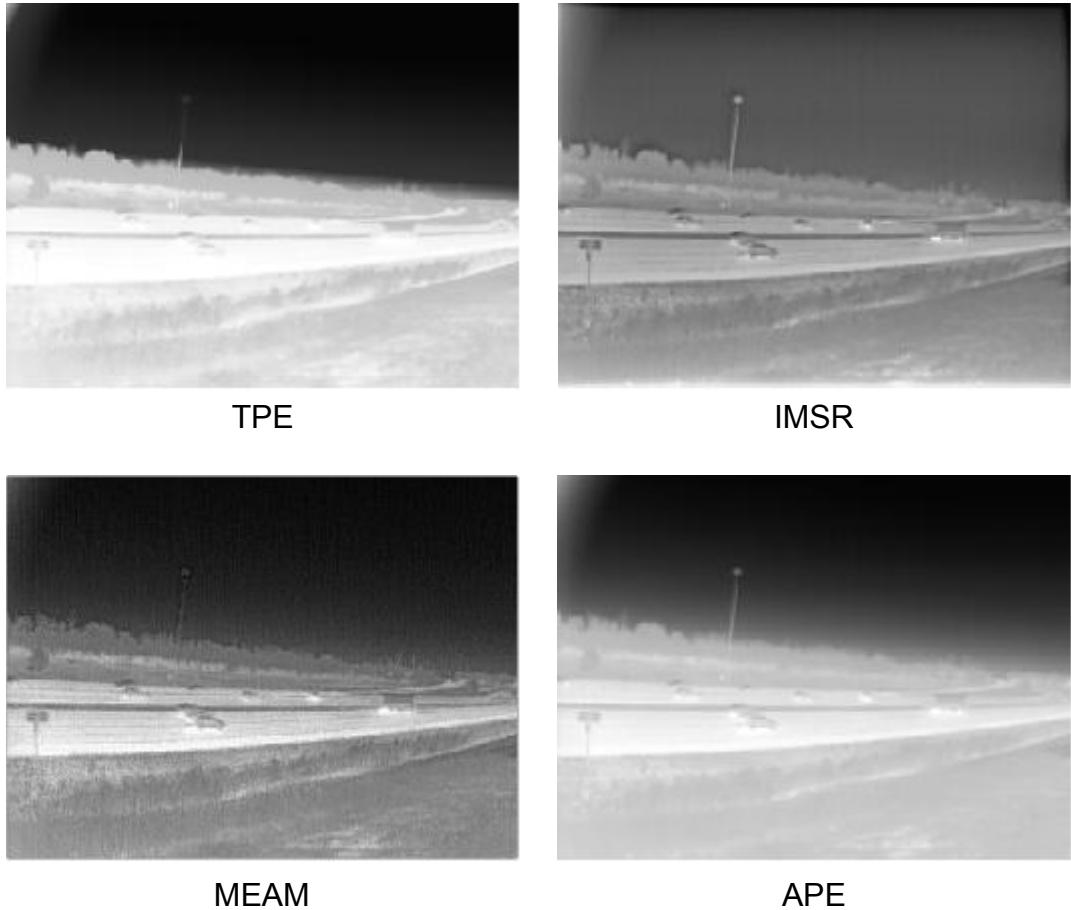


Figure 6 – Processed imagery of the natural scene

As one can see in Figure 6, the reconstructions of the scene vary widely across algorithms. TPE provided an image with stark contrast between the ground and the sky, losing most of the detail between the two fields. In addition, it becomes difficult to detect the tall light pole in the center of the scene. The other three algorithms do not exhibit this behavior. IMSR provides a scene where much detail can be seen in the ground plane at the expense of some of the detail in the sky. MEAM provided a more balanced rendition of the scene while APE provided an image that was slightly better in rendition than TPE.

4.2. Algorithm Analysis

A rate of growth analysis was conducted on each algorithm. This entailed recording the average execution time for each algorithm, repeated an arbitrary number of times, versus the size of the input image to each algorithm. It should be noted that each algorithm was executed on a computer system with swap memory enabled. By measuring the execution time on this system, there is a chance that swap access time may be included in the measurement of execution time if the system did not have enough physical memory to complete the algorithm. Due to no disk accesses being noticed by the investigator during the efficiency test, it is assumed that the swap was not accessed. Hence, the execution times recorded are believed to be for algorithmic execution alone. From the results of the analysis, the APE, TPE, and MEAM algorithms have similar linear operating characteristics that cannot be easily distinguished. That can be explained by the linear nature of each algorithm's components. The IMSR, however, takes on a character that could not be accurately assessed when viewed alone: the rate of growth non-linear. This is easily explained as the Fourier transform is a component of IMSR. As image size increases, the computational load of the Fourier transform increases logarithmically, not linearly. As such, IMSR will have a higher rate of growth since it includes both a logarithmic and linear nature.

4.3. Paired Comparison Testing

A group of firefighters and students were sought to participate in a paired comparison psychophysical trial utilizing the three collected scenes. Members of the Rochester Fire Department and Imaging Science department participated in the trials over the course of three months. Using developed software, observers were presented with exemplar video from each of the four algorithms and were asked to select video that exhibited individually the best quality and most detail.

A quantitative metric called an interval scale can be calculated from the psychophysical data. An interval scale was calculated for each question asked of the observers for the three scenes, resulting in six scales. The first scene to be analyzed was the low-to-high contrast scene. The interval scale ranks APE, MEAM, TPE, and IMSR from best to worst when concerned with quality while the interval scale ranks MEAM, APE, IMSR, and TPE from best to worst when concerned with detail. After comparing the results for the two questions for this scene, it becomes readily apparent that there is a separation between quality and detail.

The second scene to be analyzed was the low-to-mid contrast scene. The ranking from best to worst is APE, MEAM, TPE, and IMSR when concerned with quality. When concerned with detail, the results follow a different pattern: the ranking from best to worst is MEAM, APE, TPE, and IMSR. The final scene to be analyzed was the natural scene. The best to worst ranking when concerned with quality is MEAM, IMSR, APE, and TPE. The best to worst ranking when concerned with detail is IMSR, MEAM, APE, and TPE. Of the three test sets, the results from the natural scene were most consistent.

5. CONCLUSIONS

Dynamic range compression and contrast enhancement are two image processing methods that are important to any person designing an infrared imaging system. Often, the detector in an infrared system has a high dynamic range while the output display device has a much lower dynamic range. As such, an intermediate step must be taken to make these two components of the system compatible in such a way that is pleasing to the human observer. This research attempted to find examples of this intermediary step and quantitatively determine the feasibility and utility of each. By performing a rate of growth analysis on each algorithm, it became possible to compare the resources required by each algorithm in a system-independent fashion. By performing a psychophysical trial, it became possible to use the end user of an infrared system as an objective quantitative metric. Through careful analysis, it became possible to form a decision on which algorithm is the best to use in an infrared system.

The first step in analyzing the results was determining whether there was a difference between quality and detail. This is important because if there is a difference, a unique condition is placed on a system engineer designing the camera, namely the decision of which algorithm to use as the intermediary step becomes complex, where prior to this analysis, if an algorithm was determined to have a high quality, the systems engineer would choose this algorithm. If there is a difference between quality and detail, the system engineer must choose an algorithm that is appropriate to the application. For example, if the end user of an infrared system was a firefighter, their main concern would be whether they could distinguish between human beings or man-made objects. As such, spatial detail would be of primary importance since detail is what differentiates an ambiguous “blob” on the output display from a human being. If the end user of a system is a foot soldier, they would be more concerned with a noiseless output display. This is important because if a soldier sees the quick appearance of random pixels, he may perceive that to be an enemy and take inappropriate action. As such, a noise-free display would be a quality issue. From the results of the psychophysical experiment, the separation between quality and detail is apparent. For the most part, the rankings for the question of detail were different than the true rankings for the question of quality, leading one to believe that quality and detail can be separate metrics.

From the scales, one can also see that the best algorithm seems to be a frequency-based method: MEAM. MEAM was a clear winner when used in a natural scene with a high signal-to-noise ratio. However, more importantly, observers preferred MEAM just as favorably as APE when used in a low signal-to-noise ratio scene such as the artificial scenes. By using an informal ranking, the collated results show that from best to worst, the algorithm of choice is MEAM, APE, IMSR, and TPE. One of the goals of this research was to develop algorithms that were better than the baseline. From this study, that goal has been accomplished. With additional optimization for the hardware it is intended for, the frequency-based methods MEAM and IMSR should prove to be a superior algorithm for infrared contrast enhancement.

6. FUTURE RESEARCH

After completing this research, the investigator found two additional areas of research that needed to be pursued. The first area is a deep exploration of how each input parameter affects the performance of each infrared contrast enhancement algorithm. Due to the nature of the experiment, a single “one size fits all” parameter set was chosen for each algorithm to apply to each scene. In actual usage, it might be more beneficial to have a parameter or parameters that could be changed by an end user to enable the best display of infrared imagery. For example, by applying independent α and β parameters to each of the Gaussian fields in the IMSR algorithm, a smoother image might result.

The second area is in exploration of how small changes to the current algorithms might be beneficial to the algorithm as a whole. For example, the investigator wanted to see if an adaptive attenuation of the high-pass information in the MEAM algorithm would lead to a better overall quality. In theory, by adaptively attenuating the high-pass information, a greater control over the contrast among low and high temperature edges can be achieved.

REFERENCES

1. Blohm, W, “Video dynamic range compression of portrait images by simulated diffuse scene illumination,” *Optical Engineering* **35**, p255-261, 1996.
2. Engeldrum, P.G., *Psychometric scaling: a toolkit for imaging systems development*, Imcotek Press, Winchester, 2000.
3. Enkvist, M. and L. Haglund, “Automatic dynamic range adaptation for image data,” *Proceedings of the SPIE: Visual Information Processing XII*, 5108, p171-180, 2003.
4. Gonzalez, R. C. and R. E. Woods, *Digital image processing*, 2nd Edition, Prentice Hall, Upper Saddle River, 2002.
5. Gruben, J. H., et al, “Scene-based algorithm for improved FLIR performance,” *Proceedings of the SPIE: Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XI*, 4030, p184-195, 2000.
6. Jin, Y., L. Fayad, and A. Laine, “Contrast enhancement by multi-scale adaptive histogram equalization,” *Proceedings of the SPIE: Wavelets: Application in Signal and Image Processing IX*, 4478, p206-213, 2001.
7. Larson, G. W., H. Rushmeier, and C. Piatko, “A visibility matching tone reproduction operator for high dynamic range scenes,” *IEEE Transactions on Visualization and Computer Graphics*, 3, p291-306, 1997.
8. Rahman, Z., D.J. Jobson and G.A. Woodell, “Retinex processing for automatic image enhancement,” *Journal of Electronic Imaging* **13**, p100-110, 2004.
9. Reinhard, E. and K. Devlin, “Dynamic range reduction inspired by photoreceptor physiology,” *IEEE Transactions on Visualization and Computer Graphics*, 11, p13-24, 2005.
10. Schott, J. R., *Remote sensing: the image chain approach*, Oxford University Press, New York, 1997.
11. Tsujii, O., M. T. Freedman, and S. K. Mun, “Anatomic region-based dynamic range compression for chest radiographs using warping transformation of correlated distribution,” *IEEE Transactions on Medical Imaging*, 17, p407-418, 1998.