UNIVERSITY OF KWAZULU-NATAL

MULTIMODAL ENHANCEMENT-FUSION TECHNIQUE FOR NATURAL IMAGES

Rivania Maharaj

Supervisor: Mr Bashan Naidoo

The dissertation is submitted as a fulfilment of the academic requirements for the degree of

Master of Science in Engineering

College of Agriculture, Engineering and Science, School of Engineering

University of KwaZulu-Natal

November 2018

EXAMINER'S COPY

Declaration 1 - Plagiarism

I, Rivania Maharaj, declare that:

1. The research reported in this thesis, except where otherwise indicated, is my original research.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
    a. Their words have been re-written, but the general information attributed to them has been referenced.
    b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed by candidate: _____

As the candidate's supervisor, I agree / do not agree upon the submission of this dissertation.

Signed by supervisor: _____

Date of Submission: 26 November 2018

Declaration 2 - Publications

Details of contribution to publications that form part and/or include research presented in this thesis (include publications in preparation, submitted, in press and published and give details of the contributions of each author to the experimental work and writing of each publication).

Publication 1:

R. Maharaj and B. Naidoo, "An Analysis of Objective and Human Assessments in Contrast Enhancement", 2018.

Submitted to Research India Publications - International Journal of Applied Engineering Research (IJAER) (Accepted), Paper code: 66780, ISSN: 0973-4562 (print version).

Publication 2:

R. Maharaj and B. Naidoo, "Multimodal Enhancement-Fusion Technique for Natural Images", 2018.

Submitted to ELSEVIER- Signal Processing- European Association for Signal Processing (EURASIP) (Under Review), ISSN: 0165-1684 (print version).

These journals are part of the Department of Higher Education and Training (DHET) list of accredited journals.

Signed by candidate: _____

As the candidate's supervisor, I agreed / did not agree upon the submission of these publications.

Signed by supervisor: _____

Acknowledgements

I would like to convey my gratitude to all those who encouraged and accompanied me on this academic journey. Most of all, I would like to express my appreciation to God for blessing me to get this far as well as assisting me in attaining my goal. My experience at UKZN has shaped my journey going forward and made me a strong and independent person.

I extend my sincere thanks and appreciation to my supervisor, Mr Bashan Naidoo, especially for my academic development and growth. His consistent support assisted me to grow. With this support, vision and recommendations my research process took shape. I am appreciative for his constructive feedback. His noble mentorship, words of encouragement, and high prospects made this a rewarding experience. To all staff members in the discipline of Electrical, Electronic and Computer Engineering from the cleaners to the professors, thank you for your assistance and kindness during my studies from undergraduate to postgraduate level. A humble thanks to Telkom, Centre of Excellence, for financially supporting me during this endeavour.

A special thanks to my mother Rekha Maharaj, father Rajen Maharaj and the rest of my family for their tireless efforts and support provided throughout my childhood up to this level of education. I would like to express my gratitude towards my parents for all the sacrifices they have made for me during my life. I would not have been able to achieve this goal without them. Lastly, I would like to convey my sincere appreciation to Mrs J. Tugh for her time taken to proof read my work. To those who participated in my research survey, thank you for your valuable inputs.

Abstract

This dissertation presents a multimodal enhancement-fusion (MEF) technique for natural images. The MEF is expected to contribute value to machine vision applications and personal image collections for the human user. Image enhancement techniques and the metrics that are used to assess their performance are prolific, and each is usually optimised for a specific objective. The MEF proposes a framework that adaptively fuses multiple enhancement objectives into a seamless pipeline. Given a segmented input image and a set of enhancement methods, the MEF applies all the enhancers to the image in parallel. The most appropriate enhancement in each image segment is identified, and finally, the differentially enhanced segments are seamlessly fused. To begin with, this dissertation studies targeted contrast enhancement methods and performance metrics that can be utilised in the proposed MEF. It addresses a selection of objective assessment metrics for contrast-enhanced images and determines their relationship with the subjective assessment of human visual systems. This is to identify which objective metrics best approximate human assessment and may therefore be used as an effective replacement for tedious human assessment surveys. A subsequent human visual assessment survey is conducted on the same dataset to ascertain image quality as perceived by a human observer. The interrelated concepts of naturalness and detail were found to be key motivators of human visual assessment. Findings show that when assessing the quality or accuracy of these methods, no single quantitative metric correlates well with human perception of naturalness and detail, however, a combination of two or more metrics may be used to approximate the complex human visual response.

Thereafter, this dissertation proposes the multimodal enhancer that adaptively selects the optimal enhancer for each image segment. MEF focusses on improving chromatic irregularities such as poor contrast distribution. It deploys a concurrent enhancement pathway that subjects an image to multiple image enhancers in parallel, followed by a fusion algorithm that creates a composite image that combines the strengths of each enhancement path. The study develops a framework for parallel image enhancement, followed by parallel image assessment and selection, leading to final merging of selected regions from the enhanced set. The output combines desirable attributes from each enhancement pathway to produce a result that is superior to each path taken alone. The study showed that the proposed MEF technique performs well for most image types. MEF is subjectively favourable to a human panel and achieves better performance for objective image quality assessment compared to other enhancement methods.

Table of Contents

List of Figures

List of Tables

List of Algorithms

List of Acronyms

2D	Two Dimensional
AEGMM	Automatic image Equalisation and contrast enhancement using Gaussian Mixture Modelling
AELAB	Adaptive Equalisation in LAB Space
AGGD	Asymmetric Generalised Gaussian Distribution
BBHE	Brightness preserving Bi-Histogram Equalisation
BRISQUE	Blind/Reference less Image Spatial Quality Evaluator
CEID	Contrast Enhancement based on Intrinsic Decomposition
CIEBHE	Colour Image Enhancement Based on Histogram Equalisation
CIELAB	CIE Luminance channel, a-channel, b-channel [Colour Space]
CPCQI	Colourfulness-based Patch-based Contrast Quality Index
DR	Dynamic Range
DSIHE	Dualistic Sub Image Histogram Equalisation
EBCM	Edge-Based Contrast Measure
GA	Genetic Algorithm
GGD	Generalised Gaussian Distribution
GM	Gradient Magnitude
GMM	Gaussian Mixture Model
HDR	High Dynamic Range
HDRI	High Dynamic Range Imaging
HE	Histogram Equalisation
HMF	Histogram Modification Framework
HSV	Hue, Saturation, Value [colour space], H-Hue channel, S-Saturation channel, V-Value channel.
HVS	Human Visual System

IQA	Image Quality Assessment
MB	Mean Brightness
MEF	Multimodal Enhancement-Fusion for natural images
MMBEBHE	Minimum Mean Brightness Error Bi-Histogram Equalisation
MSCN	Mean Subtracted Contrast Normalisation
MSE	Mean Square Error
MVG	Multivariate Gaussian Model
NFERM	No-reference Free Energy based Robust Metric
NIQE	Naturalness Image Quality Evaluator
NIQMC	No-reference Image Quality Metric for Contrast distortion
NPEA	Naturalness Preserved Enhancement Algorithm for non-uniform illumination images
NR	No Reference
NSS	Natural Scene Statistics
PC	Phase Congruency
PCQI	Patch-based Contrast Quality Index
PSNR	Peak Signal to Noise Ratio
RGB	Red, Green, Blue [colour space]
RMSHE	Recursive Mean Separate Histogram Equalisation
Y-Cb-Cr	Luma signal, two chroma components (Cb, Cr) [colour spaces]

Chapter 1- Literature review

1.1 Introduction

Image sensors are utilised in digital cameras and in many imaging devices. These sensors are used in industry, multimedia, medical imaging and consumer applications etc. Modern digital imaging sensors have become more widely accessible and less expensive in the past decades [1]. There are many advancements with imaging [1]. Despite advancements with imaging sensors, they seldomly produce ideal raw images. Sensory output is usually subjected to a variety of corrective algorithms before becoming useful. This is especially prevalent in machine vision applications [2]. Various sources of error exist, for example; lens distortion, sensor dynamic range limitations, thermal distortion, etc. These errors can be corrected by using digital image processing techniques.

Digital image processing is the process that utilises computer algorithms to modify a digital image for some purpose. These algorithms are used to enhance images or to extract some useful information. Common examples include filtering, enhancement and fusion techniques. Image enhancement is a significant application of image processing due to its ability to improve visibility and perceivability of poor or distorted images. Distinctive procedures have been proposed to improve the quality of the digital image and to assess the quality of an image. The general image processing domain is very actively researched with multiple high-profile journal publications released in the last year alone. However, there is no "one-enhancement" that will fix all imaging problems. Each enhancement method is optimised for a particular class of problems, and each method is bound to have limitations in general application. Therefore, this dissertation proposes a technique that attempts to mitigate limitations imposed by traditional enhancement methods. This is achieved in two-parts:

1 Existing enhancement methods and associated performance metrics are studied, and
2 A new method that subdivides an image, adaptively enhances each region and seamlessly merges the result is proposed.

It is difficult to improve human perception of an image that correlates well with the direct human perception of a scene. The human visual system (HVS) processes scene illumination nonlinearly. There are some image processing algorithms that are designed to improve an image's illumination and contrast. Some of these techniques operate well on images that have a uniform spatial spreading of grey values, while other images may not, and a loss of clarity of detail and colour may arise [3]. These difficulties are related to illumination. According to [3, 4, 5], enhancement methods are categorised into two groups viz. spatial domain and frequency domain methods. Images may be represented in both spatial or frequency domain. There are many fusion techniques that have been designed for both domains. Image fusion is an application of image processing where images are fused together. It can represent spatial and frequency information.

1.1.1 Image Fusion

The main aim for any image fusion algorithm is to combine all important visual information from various image sources into a resultant image [6]. This resultant image should exhibit improved information and accuracy than any of the image sources without producing artefacts [3]. A good fusion method preserves useful information and does not create any artefacts that can mislead a human observer. It should also be reliable, robust and should not disregard any salient information from the input images. There are different levels of abstraction of information for image fusion: i.e. the pixel level, decision level, signal level and feature level [7]. For signal level fusion, the signals from different sensory sources are fused to produce a new signal. This signal has a better signal to noise ratio. Pixel/ Data fusion merges raw data from many sources into a single resolution data. The output contains more information. Feature level fusion extracts distinctive features from the different data sources. These features are merged into one or more feature maps. Decision level fusion merges the outcome from several algorithms to produce a resultant fused decision.

In addition, there are many fusion techniques that currently exist such as Wavelet form [8], Multiscale transform-based fusion [9], Laplacian pyramid based [10] etc. More examples can be found [8, 11, 12, 13, 14, 15]. Image fusion can be separated into five groups[11]: multisensor fusion, multiview fusion, multitemporal fusion, multifocus fusion and multimodal fusion.

The multisensor fusion integrates captured images of the same scene. These images are captured with different sensors. Multiview fusion integrates captured images of the same scene with numerous pictures with dissimilar views to yield a single image that contains more information. Multifocus fusion integrates captured images from the same scene but with diverse focuses. The multitemporal fusion integrates valuable information from dissimilar images of the equivalent scene at dissimilar time value. Finally, the multimodal fusion, integrates diverse modalities of images captured from the same scene. Multimodal fusion will be a focus of this study.

1.1.2 Multimodal Fusion

Multimodal fusion is a very active area of research with many domains of application. Image fusion and registration consists of three sub-problems [16] namely, the identification and extraction of common features, the determination of corresponding pixels in the image metric, and the determination of registration transformation parameters.

Examples of multimodal fusion may be found in medical imaging where there are different modalities such as computerised tomography, magnetic resonance imaging, ultrasound imaging, etc [17]. These images are merged together to achieve an improved image that contains more

information and better quality. This is achieved by utilising multimodal image fusion. The main goal is to fuse these images to achieve a superior quality image and an image with the most amount of information.

Another example of this type of fusion is a method by Saleem et al [10]. Saleem et al. proposed the "image fusion-based contrast enhancement" that enhances global and local contrast and reduces over-enhancement artefacts though maintaining the original appearance. This method is discussed in Chapter 2 and the model is presented in Appendix C6. The proposed MEF is developed by using the Saleem et al. approach and the methodology of high dynamic range imaging (HDRI).

1.1.3 High Dynamic Range Imaging Synthetisation

In recent years there has been an increasing interest to HDRI since it possesses a superior dynamic range (DR) of intensity values compared to low DR images [18]. To visualise the high dynamic range (HDR) images on standard device, the tone mapping operator is employed. Some examples of these are documented in [19, 20, 21, 22]. For standard imaging devices, only a small subset of the accessible DR of the scene can be taken. This results in over-exposed and under-exposed regions of the attained image. In order to overcome this constraint, different portions of a dynamic range are captured separately and with varied exposure times [23].

The tone mapping operator compresses the DR of HDR images causing loss of information and degradation. Annamária et al. [22] proposed a tone reproduction algorithm which can help the development of difficult to see features and colour content. The author applied the result from their previous work i.e. "gradient based synthesised multiple exposure time HDR image" into their tone reproduction algorithm [24]. The author synthesises resultant image from several registered images of a static scene, taken at different exposures, such that the regions containing the most detail are retained in the final image.

Annamária et al. [24] introduces the concept of the "gradient-based multiple exposure time synthetisation algorithm". The algorithm combines images of different exposures into a single resultant image. The resultant image encodes greater information content than each input image individually. In addition, negligible noise is produced. Each input image is divided into regular small regions. The method separately processes the red, green and blue (RGB) colour planes for each region. The algorithm measures the level of detail across corresponding regions in the registered input image set. This is performed for each colour plane. The registered input region that contains the most information is selected and assigned to the output image. According to Annamária et al. [22], the amount of detail for each region is determined by the sum of the gradient magnitude of the luminance in that image region. The more detail in a region, the greater the totality of gradient values in that region. Once all the input regions are similarly processed and

output regions assigned, the output regions are merged. Thus, producing an image with the most amount of information in each region. The output regions are sourced from variably exposed input images. Therefore, the output image usually has large intensity variations at the region transitions. Smoothing is applied to the image to eliminate the sharp transitions. Monotonically decreasing blending functions are centred over regions and extend beyond the region into the surrounding image. The smoothing function allocates the most weight to the pixel which is positioned on the centre of the measured region. For the remaining pixels in the image, weights that are inversely related to the distances from the centre of the region are allocated. After the weighting is done, the blending of the corresponding colour components is done. The resulting image has high-quality colour, that contains details and colour information while having no discontinuities along region boundaries. Aspects of this approach are adopted in the proposed MEF.

1.1.4 Performance Evaluation

Performance evaluation refers broadly to the measurement of a specific behaviour or outcome of an algorithm. This evaluation emphasises the intrinsic characteristics of an algorithm and assesses the benefits and limitations of an algorithm [25]; and is utilised to quantitatively assess the performance of an image algorithm [25]. To determine whether an algorithm succeeds or fails, the characteristics of success must be well-defined. The process of failure analysis assesses the reason an algorithm fails during testing. The information obtained is then sent to the design process to make further improvements in the algorithm. This process can be complex and difficult in image enhancement application since there is no “ideal” image that can be utilised as a reference image. The purpose of performance metrics is to determine the quality of an image in correlation with human quality assessment [18]. In addition, when evaluating image quality, the quality of image as perceived by the human visual system cannot be satisfactorily correlated with any single quantitative metric. This problem has been highlighted in journal papers such as [25, 26]. In this dissertation it was shown that metrics may be used in combination to better achieve this goal.

Performance metrics for IQA are divided into the categories; subjective IQA and objective IQA. Subjective evaluation refers to human visual inspection while objective evaluation does not involve human assumptions and assessments but involves mathematical models [27]. Objective IQA methods create a mathematical model that repeatably determines the quality of a given image as precisely as possible [28]. This precise or repeatable value is a simulation of the average human assessment [27] and it may lack accuracy.

Subjective IQA is a dependable method for evaluating image quality. It requires human respondents. In most multimedia applications, the end users are average human observers,

therefore, their opinions are necessary [18]. To undertake a test of image quality that is subjective, many global standards are recommended [29, 30, 31]. These standards are designed to facilitate reliable outcomes.

Objective IQA testing is more convenient than human assessment studies, and it is for this reason that it is so commonly used as an estimator of human perception. A common application for objective IQA is image-based quality control systems [32]. This may be achieved by image acquisition, with subsequent objective IQA that is usually integrated into a control loop. It can also be used to benchmark image processing algorithms and to optimise image processing and transmission systems. Objective IQA methods are categorised into three groups depending on the availability of an ideal distortion-free reference image [27].

The first is full-reference IQA, in which the reference image is completely obtainable, for example [32, 33]. In the second group, the reference image is not fully obtainable. This is called reduced reference IQA, examples are [34, 35]. In the third group the reference image is not obtainable. This is known as the "no-reference" or "blind quality assessment" IQA, examples are: [36, 37, 38, 39]. This study uses multiple IQA methods.

1.2 Motivation and Research Objectives

The work presented in this dissertation is aimed at improving existing singular enhancement techniques by introducing a parallel fusion method. The proposed multimodal model is expected to contribute value to machine vision applications as well as personal image collections for the human user. The proposed method identifies and merges desired attributes from parallel enhancement pathways into the resultant image. For this purpose, two papers are presented. The objective of Paper 1 is to address a selection of objective assessment metrics for contrast-enhanced images and determines their relationship with the subjective assessment of human visual systems. Paper 2 proposes the multimodal enhancement-fusion (MEF) technique for natural images.

Paper 1 addresses the problem that has been highlighted in several sources in the literature such as [25, 26]. This problem involves the significant difficulty that arises when finding a suitable evaluation method for an algorithm that provides an objective measurement of performance. The paper addresses a selection of objective assessment metrics for contrast-enhanced images and determines their relationship with the subjective assessment of human visual systems. This is in order to establish which objective metrics best approximate human assessment and may therefore be used as an effective replacement for tedious human assessment surveys. A subsequent human visual assessment survey is conducted on the same dataset to ascertain image quality as perceived by a human observer.

Paper 2 proposes a multimodal enhancement-fusion (MEF) technique for natural images. The MEF contributes value to machine vision applications and personal image collections for the human user. There are various state-of-the-art enhancement techniques that exist for targeted and global enhancement, however, every state-of-the-art method possesses its own favourable and unfavourable characteristics. There is no method that is optimal for all image types. The objective for the MEF is to adaptively select the optimal enhancer (from an available set) for each region. This is performed with the intention to minimise unfavourable characteristics. This technique also focuses on improving chromatic irregularities such as poor contrast distribution and distortions. The MEF proposes a concurrent enhancement pathway that subjects an image to multiple image enhancers in parallel, followed by a fusion algorithm that creates a composite image that combines the contributed strengths of each enhancement path. This study develops a global framework for parallel contrast enhancement, followed by parallel image assessment and region selection, leading to final merging of selected regions from the enhanced set.

1.3 Methodological Approach

The study first critically selects, implements and validates recently published works that are appropriately related to image enhancement and image fusion. Thereafter, the study deploys a theoretical model through analysis, design and synthesis. This stage incorporates novel ideas and comprises of the contribution of this study. Finally, the study assesses and refines the developed model and evaluates against benchmark datasets and published works.

Five state-of-the-art enhancement methods were critically selected. The methods have been assessed using benchmark image datasets. The methods have been replicated in MATLAB [40] and the results were validated. The strengths and weaknesses of the algorithmic output, and the identification of metrics for evaluating these strengths and weaknesses were identified. Thereafter, subjective human surveys and objective image quality assessments were performed.

State-of-the-art image-fusion methods and HDR methods were identified. These methods have been assessed using benchmark image databases. The HDRI method (Annamária et al. [22]) was replicated in MATLAB and then validated by assessing the replication against the author's benchmark dataset. The knowledge obtained is applied to create the proposed MEF framework. Given a segmented input image and a set of enhancement methods, the MEF applies all the enhancers to the image in parallel. The most desired/optimal enhancer in each image segment is identified, and finally, the differentially enhanced set of segments are seamlessly fused, thus creating an output that expresses various strengths across the enhancement methods.

The new MEF was tested along with its constituent singular enhancement methods on a benchmark database. A comparative assessment was performed between the MEF output and

other state-of-the-art methods using established metrics. A human assessment survey was finally used to obtain an authoritative assessment of the proposed MEF relative to the individual enhancement methods.

1.4 Contributions of Included Papers

All the research that is covered in this dissertation is incorporated into the following two papers that are presented in Chapter 2. The details of the included papers are described below:

1.4.1 Paper 1

R. Maharaj and B. Naidoo, "An Analysis of Objective and Human Assessments in Contrast Enhancement".

This paper addresses a selection of objective assessment metrics for contrast-enhanced images and determines their relationship with the subjective assessment of human visual systems. This is to establish which objective metrics best approximate human assessment and may therefore be used as an effective replacement for tedious human assessment surveys. A targeted study of popular contrast enhancement methods and performance metrics is first conducted. Five popular contrast enhancement methods and eight objective performance metrics are chosen. A subsequent human visual assessment survey is conducted on the same dataset to ascertain image quality as perceived by a human observer. The interrelated concepts of naturalness and detail were found to be key motivators of human visual assessment. Findings show that no single quantitative metric correlates well with human perception of naturalness and detail, however, two or more metrics in combination can be used to approximate the complex human response.

1.4.2 Paper 2

R. Maharaj and B. Naidoo, "Multimodal Enhancement-Fusion technique for Natural Images".

Paper 2 proposes a multimodal enhancement-fusion (MEF) technique for natural image. The multimodal enhancer is expected to contribute value to machine vision applications and personal image collections for the human user. The proposed MEF focusses on improving chromatic irregularities such as poor contrast distribution. The multimodal enhancement result is tailored such that it identifies and merges desired attributes from each pathway into the resultant image. It also proposes a concurrent enhancement pathway that subjects an image to multiple image enhancers in parallel, followed by a fusion algorithm that creates a composite image that combines the strengths of each enhancement path. This study develops a global framework for parallel contrast enhancement, followed by parallel image assessment and region selection, leading to final merging of selected regions from the enhanced set.

1.5 Structure of the Dissertation

This dissertation is structured such that the literature review is presented in Chapter 1, Chapter 2 presents Papers 1 and 2. The dissertation is concluded in Chapter 3. The digital copy of the results presented in this dissertation can be found in Appendix A. Summary of MEF algorithm can be found in Appendix B.

The dissertation studies five main types of image enhancement algorithms. These are: -

1. Colour image enhancement based on histogram equalisation [41]
2. Adaptive equalisation in LAB space [42]
3. Contrast enhancement based on intrinsic decomposition [26]
4. Naturalness preserved enhancement algorithm for non-uniform illumination images [43]
5. Automatic image equalisation and contrast enhancement using Gaussian mixture modelling [44].

In addition to the above methods, other enhancement methods are studied. The algorithmic descriptions for the above methods are presented in Appendix C1-C5.

Image fusion methods are also studied. Examples are exposure fusion by Mertens et al. [14], and image fusion-based contrast enhancement by Saleem et al. [10]. More information on Saleem et al. method can be found in Appendix C6. An HDRI method was studied and replicated. The method by Annamária et al. [22] was chosen.

The image processing techniques were evaluated using the following performance metrics. These metrics are carefully studied and implemented:-

1. Mean square error (MSE) [45],
2. Entropy [46],
3. Edge- based contrast measure (EBCM) [44],
4. Naturalness image quality evaluator (NIQE) [47],
5. No-reference free energy based robust metric (NFERM) [48],
6. No-reference image quality metric for contrast distortion (NIQMC) [38],
7. the colourfulness-based PCQI (patch-based contrast quality index [49])(CPCQI) [39],
8. the blind/reference-less image spatial quality evaluator (BRISQUE) [37].

These metrics are explained and implemented in Paper 1. The metrics are also used in Paper 2. In Paper 1 some of the models are demonstrated. The remaining models are presented in Appendix D. In addition to the above performance methods, a human quality assessment survey was done for both papers. The survey for Paper 1 and 2 is presented in Appendix E and F, respectively.

1.6 Future Work

Image enhancement is an active field of research, and iterative improvements are to be expected. The current proposal leaves room for improvement and extension. The following future work is proposed:

- The MEF technique produces better results in comparison to studied enhancements and fusion methods but there may be limitations in real-time applications such as video processing. In future work, the model may be simplified to enhance and improve the efficiency of the algorithm. It may also be further developed to minimise memory usage. The proposed reduction in space and time complexity may be further improved by algorithmic parallelisation for deployment on GPU and multicore CPU platforms.
- Individual image enhancement methods are generally optimised for specific problem types. It is therefore unrealistic to preselect a small set of enhancers for the MEF and expect all problem types to be addressed. If an input image classifier is inserted at the beginning of the processing pipeline, a problem class may be identified. Given a large pool of enhancers, we may then select a tailored subset that is applicable to that image or problem class. In this way, the MEF framework may be automatically targeted at the specific needs of the input image. The MEF will be broadly applicable and specialised at the same time.
- The current MEF uses a detail metric to select the best enhancer for each image segment. This metric can be improved. The research outcome from Paper 1 showed that combined metrics better estimate human perception. In future work, a composite metric may be deployed in the MEF, such that it selects both naturalness and detail as perceived by the human visual system. It is anticipated that this change will produce more pleasing results for the human observer. A specific example is the NFERM metric that uses 23 features to measure naturalness. The features from NFERM can be integrated with features of a detail metric to measure both naturalness and detail.

Finally, more research can be performed to compare the MEF with other enhancement and fusion methods. There is no perfect method for image enhancement therefore further work and research will always be required.

1.7 References

[1] R. L. Khuboni and B. Naidoo, “Adaptive segmentation and patch optimisation for multiview,” University of KwaZulu-Natal, Durban, 2014.

[2] C. Stamatopoulos, C. Fraser and S. Cronk, “Accuracy aspects of utilizing raw imagery in photogrammetic,” *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* vol. B5, pp. 1, 2012.

[3] D. K. Sahu and M.P.Parsai, “Different Image Fusion Techniques: A Critical Review,” *International Journal of Modern Engineering Research (IJMER),* vol. 2, no. 5, pp. 4298-4301, 2012.

[4] R.Gonzalez and P. Wintz, Digital Image Processing, Norwood, MA: Addison-Wesley, 1987.

[5] D. H. Rao and P. P. Panduranga, “A survey on image enhancement techniques: classical spatial filter, neural network, cellular neural network, and fuzzy filter,” in *IEEE International Conference on Industrial Technology*, pp. 2823-2827, 2006.

[6] M. B. A. Haghighata, A. Aghagolzadehab and H. Seyedarabia, “Multi-focus image fusion for visual sensor networks in DCT domain,” *Computers & Electrical Engineering,* vol. 37, no. 5, pp. 789-797, September, 2011.

[7] Y.H. Chen, “Multimodal Image Fusion and Its Applications,” The University of Michigan, 2016.

[8] S.G. Huang, “Wavelet for Image Fusion,” National Taiwan University, 2010.

[9] B. Bavachan and D. Krishnan, “A Survey on Image Fusion Techniques,” *International Journal of research in computer and communication technology,* 2014.

[10] A. Saleem, A. Beghdadi and B. Boashash, “Image fusion-based contrast enhancement,” *EURASIP Journal on Image and Video Processing,* no. 10, 2012.

[11] M. AbouRaya, “Real-time Image Fusion Processing for Astronomical Image,” The University of Toledo, May 2016.

[12] M. Amin-Naji and A. Aghagolzadeh, “Multi-Focus Image Fusion in DCT Domain using Variance and Energy of Laplacian and Correlation Coefficient for Visual Sensor Networks,” *Journal of AI and Data Mining,* vol. 6, no. 2, pp. 233-250, 2016.

[13] A. Rani and G. Kaur, “Image Enhancement using Image Fusion Techniques,” *International Journal of Advanced Research in Computer Science and Software Engineering,* vol. 4, no. 9, 2014.

[14] T. Mertens, J. Kautz and F. V. Reeth, “Exposure Fusion,” in *15th Pacific Conference on Computer Graphics and Applications,* 2007.

[15] R. Rubinstein and A.Brook, “Fusion of differently exposed images - Final Project Report,” Technion, Israel Institute of Technology, 2014.

[16] M. Y. Belkhouche, "Multi-perspective, Multi-modal Image Registration and Fusion," University of North Texas, Denton, Texas, August 2012.

[17] U. B. Mantale and P. V. B. Gaikwad, "Image Fusion of Brain Images using Redundant Discrete Wavelet Transform," *International Journal of Computer Applications,* vol. 74, no. 4, 2013.

[18] P. Mohammadi, A. Ebrahimi-Moghadam and S. Shirani, "Subjective and Objective Quality Assessment of Image: A Survey," Elsevier, 2014.

[19] E. Reinhard, M. Stark, P. Shirley and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Transactions on Graphics,* vol. 21, pp. 267-276, 2002.

[20] G. W. Larson, H. Rushmeier and C. Piatko, "A visibility matching tone reproduction operator for high dynamic range scenes," *IEEE Transactions on Visualization and Computer Graphics,* vol. 3, no. 4, pp. 291-306, 1997.

[21] "Gradient domain high dynamic range compression," *ACM Transactions on Graphics,* vol. 21, pp. 249-256, July 2002.

[22] A. R. Várkonyi-Kóczy, A. Rövid and T. Hashimoto, "Gradient-Based Synthesized Multiple Exposure Time Color HDR Image," *IEEE Transactions On Instrumentation And Measurement,* vol. 57, no. 8, pp. 1779-1785, 2008.

[23] A. Artusi, T. Richter, T. Ebrahimi and R. K. Mantiuk, "High Dynamic Range Imaging Technology," *Journal of Latex class file,* vol. 14, no. 8, 2016.

[24] A. Rövid, A. R. Várkonyi-Kóczy, T. Hashimoto, S. Balogh and Y. Shimodaira, "Gradient based synthesized multiple exposure time HDR image," in *IEEE IMTC*, Warsaw, Poland, May 1–3, 2007.

[25] M. Wirth, M. Fraschini, M. Masek and M. Bruynooghe, "Performance Evaluation in Image Processing," *EURASIP Journal on Applied Signal Processing,* vol. 2006, pp. 1-3, 2006.

[26] H. Yue, J. Yang, X. Sun, F. Wu and C. Hou, "Contrast Enhancement Based on Intrinsic Image Decomposition," *IEEE transactions in image processing,* vol. 26, no. 8 , pp. 3981 - 3994, 8 August 2017.

[27] D. S. Prabha and J. S. Kumar, "Performance Evaluation of Image Segmentation using Objective Methods," *Indian Journal of Science and Technology,* vol. 9, no. 8, 2016.

[28] H. H. Barret, J. L. Denny, R. F. Wanger and K. J. Myers, "Objective assessment of image quality. II. Fisher information, Fourier crosstalk, and figures of merit for task performance," *Journal of the Optical Society of America A,* vol. 12, pp. 834-852, 1995.

[29] BT.500-11, ITU-R recommendations, Methodology for the subjective assessment of the quality of television pictures, Geneva, Switzerland, ITU, 2002.

[30] BT.710-4, ITU-R recommendations, Subjective assessment methods for image quality in high-definition television, Geneva, Switzerland, ITU, 1998.

[31] BT.814-1, ITU-R recommendations, Specification and alignment procedures for setting of brightness and contrast of displays, ITU, 1994.

[32] Z. Wang and A. C. Bovik, *Modern image quality assessment: Synthesis Lectures on Image, Video & Multimedia Processing,* Morgan & Claypool Publishers, 2006.

[33] A. George and S. J. Livingston, "A survey on full reference image quality assessment algorithms," *IJRET: International Journal of Research in Engineering and Technology* , vol. 2, no. 12, 2013.

[34] A. Rehman and Z. Wang, "Reduced-reference image quality assessment by structural similarity estimation," *IEEE Trans. Image Processing,* vol. 21, pp. 3378-3389, Aug. 2012.

[35] D. Yang, Y. Shen, Y. Shen and H. Li, "Reduced-reference image quality assessment using moment method," *International Journal of Electronics,* vol. 103, no. 10, 2016.

[36] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *IEEE Trans. Image Processing,* vol. 18, pp. 717-728, April 2009.

[37] A. Mittal, A. K. Moorthy and A. C. Bovik, "No-Reference Image Quality Assessment in the Spatial Domain," *IEEE Transactions on Image Processing,* vol. 21, no. 1, pp. 4695 - 4708, December 2012.

[38] K. Gu, W. Lin, G. Zhai, X. Yang, W. Zhang and C. W. Chen, "No-Reference Quality Metric of Contrast-Distorted Images Based on Information Maximization," vol. 47, no. 12, 2017.

[39] K. Gu, D. Tao, J. F. Qiao and W. Lin, "Learning a No-Reference Quality Assessment Model of Enhanced Images With Big Data," *IEEE transactions on neural networks and learning systems,* vol. 29, no. 4, pp. 1301-1313, April 2018.

[40] MathWorks, "Matlab installation," Mathworks, [Online]. Available: https://www.mathworks.com/products/matlab.html. [Accessed 20 Decemeber 2017].

[41] S. Arora and K. Kapoor, "Colour image enhancement based on histogram equalization," *Electrical & Computer Engineering: An International Journal (ECIJ),* vol. 4, no. 3, September 2015.

[42] S. Bharal, "L*a*b based contrast limited adaptive histogram equalization for underwater images," *International Journal of Computer Application (2250-1797),* vol. 5, no. 4, June 2015.

[43] S. Wang, J. Zheng, H.-M. Hu and B. Li, “Naturalness Preserved Enhancement Algorithm for non uniform illumination images,” *IEEE Transactions on Image Processing,* vol. 9, no. 9, 2013.

[44] T. Celik and T. Tjahjadi, “Automatic Image Equalization and Contrast Enhancement Using Gaussian Mixture Modeling,” *IEEE Transactions on Image Processing,* vol. 21, no. 1, 2012.

[45] Matlab, “Mean Square Error,” [Online]. Available: https://www.mathworks.com/help/images/ref/immse.html.

[46] S. Rani, A. Kumar and K. Singh, “Illumination based Sub Image Histogram Equalization: A Novel Method of Image Contrast Enhancement,” International Journal of Computer Applications, vol. 119, no. 20, 2015.

[47] A. Mittal, R. Soundararajan and A. C. Bovik, “Making a Completely Blind Image Quality Analyzer,” *IEEE Signal Processing Letters,* vol. 20, no. 3, pp. 209-212, 2012.

[48] K. Gu, G. Zhai, X. Yang and W. Zhang, “Using Free Energy Principle For Blind Image Quality Assessment,” *IEEE Transactions on Multimedia,* vol. 17, no. 1, 2015.

[49] S. Wang, K. Ma, H. Yeganeh, Z. Wang and W. Lin, “A patchstructure representation method for quality assessment of contrast changed images,” *IEEE Signal Processing Letter,* vol. 22, no. 12, pp. 2387-2390, December 2015.

Chapter 2: Included Papers

A. Paper 1: An Analysis of Objective and Human Assessments in Contrast Enhancement

Ms. Rivania Maharaj and Mr. Bashan Naidoo

Submitted to International Journal of Applied Engineering Research (IJAER) (Accepted)

Abstract

This paper addresses a selection of objective assessment metrics for contrast-enhanced images and determines their relationship with the subjective assessment of human visual systems. This is in order to establish which objective metrics best approximate human assessment and may therefore be used as an effective replacement for tedious human assessment surveys. A targeted study of popular contrast enhancement methods and performance metrics is first conducted. Five popular contrast enhancement methods and eight objective performance metrics are chosen. A subsequent human visual assessment survey is conducted on the same dataset to ascertain image quality as perceived by a human observer. The interrelated concepts of naturalness and detail were found to be key motivators of human visual assessment. Findings show that no single quantitative metric correlates well with human perception of naturalness and detail, however, two or more metrics in combination can be used to approximate the complex human response. Three metrics (NIQE, NFERM, BRISQUE) were found to be good estimators of human perception of naturalness; and two metrics (NIQMC and entropy) provided good estimation of human perception of detail.

A.1 Introduction

Performance evaluation refers broadly to a measurement of a specific behaviour of an algorithm. This evaluation emphasises the intrinsic features of an algorithm and evaluates the benefits and limitations of an algorithm [1]. There has been a significant difficulty that arises when finding a suitable evaluation method for an algorithm that provides an objective measurement of performance. Performance metric analysis is, therefore, a significant and useful measurement that is necessary for quantitatively analysing the performance and achievements of an image enhancement model [1]. To determine if an algorithm succeeds or fails, the features of success must be well-defined. It becomes necessary to use failure analysis as a method of determining the reason for an algorithm failure through testing. The information obtained is then sent back into the design stage to produce further improvements in the algorithm. This can be a challenging and complex procedure in image enhancement application since there is no “ideal” image that may be utilised as a reference image. In addition, when evaluating image quality, the image quality and accuracy as perceived by the human visual system cannot be correlated with any single quantitative metric. This problem has been highlighted in several journal papers such as [1, 2]. This paper proposes to address this problem by investigating different image contrast enhancements and the performance metrics.

Before dealing with the problem, the fundamentals of image processing have to be understood. All digital images have an equivalent matrix which stores information concerning each pixel. This

information is in the form of the colour for the pixel and intensity of that colour. This information forms the basis of image processing. Image processing algorithms can remove noise and any type of irregularities or distortions found in an image by making use of the digital computer. Irregularities, noise or distortions can appear in the image either during its formation or during transformation [3]. The goal of image enhancement is to remove these irregularities or distortions without removing any features of the image. One such example is Gaussian noise reduction. Gaussian noise is caused by natural sources and it generally disturbs the gray values in the digital image [4]. This can be solved by using a Wiener filter or the approach proposed in Murugan et al. [5], that identifies noisy pixels and restores those pixels. An important application of image processing is image enhancement because it enhances the visibility and perceivability of poor or distorted images. Several methods of image enhancement have been developed during the last four to five decades. These techniques use mathematical manipulations on the image matrix. As a result of mathematical manipulations, there is an improvement in the image and contrast quality of the original image.

This paper considers 5 different types of contrast enhancement techniques, which are: -

1. Colour image enhancement based on histogram equalisation [6]
2. Adaptive equalisation in LAB space [7]
3. Contrast enhancement based on intrinsic decomposition [2]
4. The Naturalness preserved enhancement algorithm [8]
5. Automatic image equalisation and contrast enhancement using Gaussian mixture modelling [9].

To evaluate each image enhancement technique and measure the image quality, a performance metric is used. This is an important assessment that is necessary to prove the effectiveness and efficiency of the enhancement technique. The purpose of image quality assessment (IQA) metrics is to determine the quality of an image in correlation with human quality assessment [10]. When determining image quality or accuracy, no single metric corresponds well with human assessment [1]. Therefore, choosing the appropriate metric evaluation methodology is important and depends on the intention and objective of the required task. The reason for assessing an algorithm is two-fold, namely, to comprehend its behaviour when exposed with several types of images, and/or to assist in determining the prime parameters for the various image processing applications [11]. The performance metrics are classified as subjective, and objective methods. Subjective evaluation is based on human visual inspection while objective evaluation does not involve human assumptions and assessments but involves mathematical models [12].

Thus, the aim of an objective IQA method is to create a mathematical model which best determines the quality of a given image as precisely as possible [13]. The mathematical model must simulate the quality assessment of an average human observer [12]. The objective IQA methods are categorised into three groups depending on the obtainability of a distortion-free reference image that has perfect quality [12]. The first group is a full-reference IQA in which the reference image is completely obtainable [14]. The second group, where the reference image is not fully obtainable, is called a reduced reference IQA [15]. In the third group, the reference image is not obtainable, which is known as the no-reference IQA [16, 17].

Subjective IQA is a dependable method for evaluating image quality. In most of the multimedia applications, the average human observers are the end users, therefore, making their opinions necessary [10]. In this method, respondents are required to express their view about the quality of a given image. To undertake a test of image quality that is subjective, many global standards are recommended [18, 19, 20]. These standards will provide more efficient results. A visual assessment is conducted to gather information about how the human vision interprets a given image. In this paper, an experiment was done to investigate the quality of images. This included a visual assessment that allowed recipients to score enhanced images according to visual preference, naturalness and details of an image.

The goal of this paper is to address the selection of objective assessment metrics for contrast enhanced images and determine their relationship with the subjective assessment of human visual systems. This is done in order to establish which objective metrics best approximate human assessment and may therefore be used as an effective replacement for tedious human assessment surveys. A visual assessment is done to determine a human's visual response. The human assessment is done to correlate the results of the objective study. This study also identifies objective IQA metrics that can be used to approximate a human assessment and categorises performance metric in terms of the probable application domain by analysing different objective performance metrics. Therefore, the structure of the paper is as follows: The related works is outlined in Section A.2. Section A.2 is divided into two subsections, viz, Section A.2.1 which provides the information for the enhancement algorithms that are implemented and Section A.2.2 which provides information on the metrics used for the performance analysis. Section A.3 explains the experimental method. Section A.4 discusses the results attained from the experiment and Section A.5 concludes this paper.

A.2 Related works

To implement performance metrics on contrast enhancement algorithms, an understanding of the different enhancement techniques and performance metrics are needed. These sections consider key contrast enhancement methods and the metrics that have transitionally been applied to access image performance. Therefore, this section is separated into two sections. Section A.2.1 provides a brief summary of the five enhancements and Section A.2.2 discusses the human assessment and eight objective performance metrics.

A.2.1 Enhancement methods

The different contrast enhancements that are implemented in this study are outlined below. Each enhancement is considered in terms of the goal, previous work, methodology and results from each algorithm.

A.2.1.1 Colour Image Enhancement Based on Histogram Equalisation (CIEBHE)

Goal: Colour image enhancement based on histogram equalisation (CIEBHE) was proposed by Arora et al. [6]. It aims to extend the histogram equalisation approach of gray-level images for colour images [6].

Previous work: The process of remapping pixels intensities of the image within a specific colour plane is called histogram equalisation (HE) [21]. The process distributes the pixel's intensity across the dynamic range (DR) to improve the image contrast. The HE allocates the pixel intensity values of the input image in such a way that the resultant image contains a uniform distribution of intensity, thereby enhancing the contrast [6]. The aim is to increase the brightness of an image which is not the same as the mean brightness (MB) of the original input image. HE is one of many techniques that can be utilised to improve the contrast for a given colour image. Some variants of this approach are:

- Kim et al. [22] proposed the brightness preserving bi-histogram equalisation (BBHE). This enhancement model aims to maintain the MB of an image whilst enhancing image contrast [22]. BBHE separates the histogram of the image into under-exposed and over-exposed sections, according to the mean value. The sub-image histogram of the original image is then equalised independently.
- Wang et al. [23] proposed a dualistic sub image histogram equalisation (DSIHE). DSIHE is utilised for maintaining the illumination of the image based on median value. This method involves separating the original histogram into two sub-image histograms according to their median value. In this way, each sub-image histogram is equalised independently.

- Chen et al. [24] proposed the minimum mean brightness error bi-histogram equalisation (MMBEBHE). MMBEBHE intends to provide maximum brightness preservation and is an extension of BBHE.
- Chen et al. [25] proposed the recursive mean separate histogram equalisation (RMSHE). RMSHE allows an increased level of brightness preservation to bypass unwanted artefacts and an unnatural image enhancement, that occurs because of too much of equalisation whilst improving the contrast of an image.

Method: CIEBHE extends on the HE approach for gray-level images to colour, RGB (red, green and blue) images. The original input image is transformed from RGB to HSV (hue, saturation, value) colour space. It is divided into two sections based on the exposure threshold. Thereafter, histogram equalisation is applied to each section independently. In this method, over-enhancement is controlled by using clipping threshold.

The results: This is a good method for preserving brightness. This method is suitable for under-exposed images. Over-enhancement is controlled by histogram clipping [6]. CIEBHE is effective and efficient for under-exposed images. The findings of the author [6], indicate that CIEBHE produces an enhanced image with maximum entropy and good contrast by decreasing over-enhancement.

A.2.1.2 Adaptive Equalisation in LAB spaces (AELAB)

Goal: AELAB [7] enhancement demonstrates a luminosity preserving contrast enhancing adaptive histogram equalisation technique for colour images.

Previous work: The HE is a common method used for enhancing image contrast where unwanted subject deterioration is frequently occurring. HE works well in RGB colour space and is not suited well for different colour spaces. Several colour models have been created for improving the visual representation of colour images. Colour space is a generalised term used for representing the combination of a colour model plus a mapping function associated with colour images [26]. Colour model and colour spaces are complementary to each other. Colour image contains more accurate information than the gray images. The reason for this is that many colours can be produced by mixing primary colour pigments. Most frequently used colour spaces are RGB [27] and Y-Cb-Cr colour spaces [28]. The selection of the colour model is dependent on the kind of image application and its requirements. The international commission on illumination (CIE) has developed a colour space named CIELAB colour space [29]. The major advantage of the LAB colour space is that it is created to estimate the human visual system (HVS) [30]. For the AELAB enhancement method, adaptive equalisation is applied in LAB spaces. This is similar to Bharal's [7] approach for image enhancement for underwater images. This approach is also similar to, [31] and [32]. This approach is used and applied to all types of images.

Method: The RGB image is transformed into LAB space. The LAB image is then decomposed into the individual components i.e. L, a*, b* components. The image contrast in L component is increased by using the contrast limited adaptive histogram equalisation method [7]. Using CIELAB space before an image enhancement method will increase the aesthetic appurtenance of the images [27]. This enhances the colour feature of the structural components across the different spectral channels of the image [33]. The digital component coefficients are enhanced in lightness component L. The colour information in a* and b* channels are kept unchanged for preserving the colour information. At the end of the process, the LAB colour spaces are converted back to the original RGB colour space to obtain the enhanced image.

The results of this enhancement are much better in relation to previous work composed by other authors/researchers for underwater images [7]. This comparison was performed by the author, Bharal [7]. A comparison between the other methods and the proposed method showed improvement based on criteria like: Mean square error, peak signal to noise ratio, average error and bit error rate.

A.2.1.3 Contrast Enhancement based on Intrinsic Decomposition (CEID)

Goal: CEID was proposed by Yue et al. [2], it seeks to develop a novel intrinsic image decomposition model that is appropriate for contrast image enhancement. This is done by introducing constraints on the reflectance layer and the illumination layer in order to achieve an efficient enhancement.

Previous work: Studies have shown that by altering the decomposed illumination layer the image quality is improved. These layers were altered to enhance under-exposed or over-exposed images. Such models have been proposed in [2, 8, 34, 35, 36]. Different decomposition models have created different illumination layers with distinctive characteristics which can affect the resultant images.

Barrow et al. [37] proposed intrinsic image decomposition. Their aim was to split the image into two layers which are the reflectance layer and the illumination layer. This showed that the amount of reached light is represented by the illumination values. The reflectance values, which is not changed to illumination condition, relates to the intrinsic colour of the image [2]. Intrinsic decomposition was a highly ill-modelled problem. However, there are numerous different inference and subsequent work that have been made to make this a well-modelled problem [38, 39, 40, 41].

Method: To produce an extremely efficient enhancement, CEID propose constraints on the reflectance layer and illumination layer [2]. The reflectance layer is regularised to be piecewise constant. This is done by presenting a weighted $l1$ norm constraint on the neighbouring pixels.

The weighting is in accordance with the colour similarity of an input image. This is done since the illumination information will barely affect the reflectance. A piecewise smoothness constraint is used to regularise the illumination layer. The Split Bregman process is used to resolve the proposed decomposition model. The illumination layer is altered to achieve the enhanced image. Illumination adjustment was introduced to reduce computing complexity and to avoid potential colour artefacts. The decomposition model was implemented along the value channel in HSV space.

An input colour image (RGB image) is transformed into HSV representation. The HSV image is then decomposed into the separate channels (H, S, and V channels). The decomposition model was applied to value (V) channel to decompose the image into two layers, i.e. the illumination layer and reflectance layer. To create an adjusted illumination layer, the Gamma mapping function was adopted. The new enhanced V channel is given by the product of the new illumination layer and the reflectance layer. The mapping function is done globally, whereby the author [2] uses the contrast limited adaptive histogram equalisation [42] to enhance the local contrast of the new V channel further. The final resultant V channel is recombined with the untouched H and S channels. The new HSV image is then changed back into RGB space, to yield a resultant enhanced image.

Results: This enhancement method demonstrates that CEID performs well for a broad range of images. The author's [2] research indicates that the method achieved an improved quality that is comparable to other enhancement methods. There are several limitations to this method. It must be noted that the decomposition model was created for contrast enhancement. Using this model, the result for other image processing applications such as object insertion and surface re-texturing, shall not produce desirable results. Since this enhancement model is created for images, it might cause flickering artefacts if it is applied directly to video enhancement [2].

A.2.1.4 Naturalness Preserved Enhancement Algorithm for non-uniform illumination images (NPEA)

Goal: The NPEA [8] model seeks to preserve the naturalness of an input image while at the same time enhancing its details. This enhancement algorithm is introduced for images that have non-uniform illumination.

Previous work: To preserve the naturalness of an image and enhance its detail, Chen et al. [43] proposed the idea of naturalness preservation for enhancing images. According to Chen et al, the image colour impression must not be altered drastically after the enhancement. He further stated that no additional source of light must be added to the scene and no halo effect be introduced. Also, no blocking effect must be augmented as a result of over-enhancement [43]. Therefore, no

artefact is a simple necessity for visual information fidelity, the global ambiance of an image shall not be modified drastically, and the focal point of the light source shall not be adjusted noticeably [8]. Many examples of natural enhancement models based on the Retinex theory have been proposed in the literature [43, 44, 45]. These algorithms propose to enhance detail while preserving the naturalness in an image. However, these algorithms have a limitation in that they are not desirable for images with non-uniform illuminations. The Retinex-based algorithms effectively improve details and therefore have been universally used. It has been used since the algorithms accept the removal of illumination as a defaulting preference and it cannot restrict the range of reflectance. Therefore, the naturalness of an image that has non-uniform illumination is not effectively maintained [8]. It is important to maintain naturalness for any image enhancement process to produce pleasing perceptual and visual quality. Wang et al. [8] proposes an algorithm to maintain the naturalness of an image while enhancing details for non-uniform illumination images. This method is discussed below.

Method: There are three main contributions made by Wang et al. These contributions are: preservation of naturalness, intensity decomposition and the illumination effect [8]. The first contribution introduces a lightness-order error measurement to objectively evaluate the naturalness preservation. The second contribution is the bright-pass filter, which is used to decompose the image into two layers, namely, reflectance layer and illumination layer. This determines the amount of detail and the naturalness of the image. It also ensures that the reflectance is constraint to the range [0, 1]. In the third contribution, a bi-log transformation is proposed. This transformation maps the illumination to have an equilibrium among detail and naturalness.

There are two constraints that are proposed in the NPEA algorithm. The first constraint relates to detail, which requires the reflectance to be set to a range of [0, 1], by taking into account the property of reflectance [46]. Naturalness is the next constraint, where there should not be drastic change to the relative order of illumination in various local areas.

Results: From the author's results of the experiment [8], it is observed that the NPEA enhance detail and maintain the naturalness for images with non-uniform illumination. It shows that the resultant image is more visually appealing, free of artefacts and maintains naturalness. However, a limitation of this method is that the enhancement model fails to consider the relationship of illumination across various scenes. Therefore, in video applications where the scenes change, flickering can be introduced [8].

A.2.1.5 Automatic image equalisation and contrast Enhancement using Gaussian Mixture Modelling (AEGMM)

Goal: The AEGMM [9] method proposed by Celik et al., automatically segments the contrast domain and adaptively equalises each segment. This aims to make the contrast enhancement more responsive to localised feature in contrast distribution.

Previous work: Many enhancement techniques and methods have been developed in the past years. These techniques and methods can be split into three distinct groups [9]. Group 1 includes techniques which decompose an image into high frequencies and low frequencies signals for operation [47, 48]. Group 2 are transform-based techniques [49] and Group 3 are histogram modification techniques such as [50, 51].

In the histogram modification framework (HMF), the contrast enhancement is handled as an optimisation problem which minimises a cost function. In order to handle noise and black/white stretching, variables are introduced in the optimisation. The HMF can attain various stages of contrast enhancement by utilising diverse adaptive parameters. By manually changing the parameters in accordance with the image content, a better contrast enhancement can be attained.

A parameter-free algorithm is favoured. To create a parameter free algorithm a genetic algorithm (GA) is utilised. The GA is used to obtain a target histogram which will maximise the contrast measurement based on edge information [52]. This approach is called contrast enhancement based on GA. This approach has limitations, such as its dependence on the initialisation and convergence to a local optimum [9].

Method: Celik et al. proposed an adaptive image equalisation algorithm that efficiently improves the human visual quality of various cases of given images. The AEGMM algorithm fits a Gaussian Mixture Model (GMM) to the gray-level distribution intervals at the Gaussian intersection points. The intersection points partition the DR of the image into input intervals of gray-level. To acquire an image where the contrast is equalised; each input interval is equalised in accordance with the dominant Gaussian component as well as the cumulative distribution function of the input interval. The Gaussian components that have low variances are assigned with lower values; likewise, larger values are assigned to Gaussian components that have larger variances [9]. The AEGMM is free of parameter setting for a specific DR of a resultant image. The enhancement may be utilised for a broad variety of images.

Results: It can be noted that, a low contrast image is automatically enhanced in relation to an increment in the DR. It is also observed that image with high contrast is improved, however, this improvement is little. With AEGMM, the colour quality of the wide range input image is enhanced. The quality is enhanced in terms of a few factors such as the colour consistency, developed contrast among foreground and background objects, a larger DR and detail in the image

[9]. The findings of the authors [9] indicate that AEGMM produce a resultant image that is subjectively more visually appealing compared to the original image.

A.2.1 Performance Metric Analysis

The aim of performance metrics is to assess the performance of an image. Both objective and subjective image quality assessments (IQA) are studied. In subjective evaluation, human involvement is a basic necessity, as the evaluation is determined based on the individual's visual inspection of a given image [12]. Objective IQA methods use mathematical models to determine the quality of an image. Objective methods are studied to determine methods that will help determine performance for different types of image enhancements for all types of applications.

A.2.2.1 Subjective Image Quality Assessment Measure: Human Visual assessment

The structure of the survey study was based on the existing survey from [9] and [2]. The survey was carried out face-to-face where a group of respondents was required to fill out a survey form. The reason chosen to carry out a face-to-face session was that previous studies [53] have shown that this method yields higher co-operation, lower refusal rate and higher response quality. The only cost involved was to print the images. To reduce bias, only willing respondents were given the survey. To ensure that the survey was efficient and reliable, global standards were implemented in carrying out this survey as referenced in the report [10].

The survey questionnaire consists of 2 sections. In section 1 of the survey questionnaire, the respondents were asked to refer to the portfolio of images and rate the images according to visual preference. Section 2 of the survey required respondents to choose an image according to detail, natural and unnatural appearance from the set of images in the portfolio.

The survey was conducted among a random sample population. This population representation was unbiased. The standardisation of measurement was the same for every respondent i.e. the same set of questions was asked.

A.2.2.2 Objective Image Quality Assessments Methods

In order to evaluate each image enhancement technique and measure the image quality, a performance metric analysis is done. This analysis is a fundamentally important assessment that is necessary to determine the capability and accuracy of the enhancement technique. The aim of IQA methods is to determine the calibre of input images in relation to human visual system (HVS) [10]. The aim of objective IQA methods is to design mathematical models that can precisely forecast the quality of an image [13]. The performance measures applied in this paper provide some quantitative comparison between different enhancement techniques studied. Each

performance measure provides unique information that is different from another. These metrics will be discussed below.

A.2.2.2.1 Mean Square Error

Mean square error (MSE) is determined by calculating the average squared intensity of the input and output image pixels as indicated in the expression below (1) [7]. MSE is an error metric that assesses image quality [54].

Mathematical expression:

$$MSE = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} e(m,n)^2 \quad (1)$$

Where, $e(m,n)$ is the difference in error between the original (input) and the distorted image (output). The application in this paper, does not use MSE as error metric but rather as a measure of change. This metric can be utilised to assess the change in the enhanced image to the original image or the change between the enhanced image. The smaller the value of MSE, the smaller the change. The higher the value of MSE the larger the change.

A.2.2.2.2 Entropy evaluation

Entropy is used as the Shannon entropy, which contains the maximum information [55]. The Shannon entropy is also referred to as information entropy. Entropy evaluation measures the quality of an image that is determined by estimating the amount of information contained in an image. The larger the entropy score after enhancing an image, the greater the information contained in an image. Therefore, if there is more information present in the enhancement image; it is assumed that image performance is improved and vice versa [12].

Mathematical expression

$$H = -\sum_{l=0}^{L-1} pilog_2 pi \quad (2)$$

Where, pi is the probability of the intensity value l in an image, and H refers to the entropy of the input image. L is defined as the total number of gray levels.

A.2.2.2.3 Edge-Based Contrast Measure

Humankind is more perceptive to edges (contours) [56]. Edge-based contrast measure (EBCM) [9] is created on an observation that sensitivity to an edge (contours) is much more perceived to a human observer [9]. EBCM is anticipated that a resultant image has more edge pixels than the reference/original image. The EBCM measures the intensity of edge pixels in a small region (window) of an image [57].

Mathematical expression: Image *X* as a contrast $c(i,j)$ for a pixel located at (i,j) is therefore expressed as [9]

$$c(i,j) = \frac{|x(i,j) - e(i,j)|}{|x(i,j) + e(i,j)|} \tag{3}$$

The mean edge gray level is given by:

$$e(i,j) = \frac{\sum_{(k,l)\in\mathcal{N}(i,l)} g(k,l)x(k,l)}{\sum_{(k,l)\in\mathcal{N}(i,l)} g(k,l)} \tag{4}$$

The set of all neighbouring pixels of pixel (i,l) is defined as $\mathcal{N}(i,l)$. The magnitude of the image gradient is assessed by utilising the Sobel operator at pixel (k,l) is given as $g(k,l)$ [58]. Image *X* has an EBCM which is calculated as the average contrast value, i.e.,

$$EBCM(X) = \sum_{i=1}^{H}\sum_{j=1}^{W} c(i,j)/HW \tag{5}$$

Where the height of the image is defined as *H* and the width of the image is defined as *W*. If a resultant image *X* of an input image *Y*, if the $EBCM(Y) \geq EBCM(X)$ then the contrast has improved [9].

A.2.2.2.4 Naturalness Image Quality Evaluator

The naturalness image quality evaluator (NIQE) [59] determines the length between the natural scene statistics (NSS)-based features which is computed from the image to the features derived from a database of images that is utilised to train the model. These features are modelled as multi-dimensional Gaussian distribution [60]. NIQE is built along a character-aware collection of statistical features. NIQE is a blind IQA analyser that is utilised to measure deviations. The deviations are from statistical regularities that are noticed in a natural image. This is achieved without training in human-rated distorted images or without introduction to distorted images [61]. Information about predicted distortions or deformities are required for most no-reference (NR) IQA methods. These predicted distortions are in the form of training and correlation in human assessment score [62].

Mathematical expression: The input image is divided into $P \times P$ "patches". A patch is a small area of pixels. The authors use a simple device to preferentially select from amongst a collection of natural patches those that are richest in information and less likely to have been subjected to a limiting distortion. This subset of patches is then used to construct a model of the statistics of the natural image patches. From the coefficients of each patch, certain NSS features are calculated. Only a subset of the patches is used. Let the size of patches be indexed as $d = 1,2,3,\ldots.D$. To calculate the average local deviation field of individual patch indexed *d*, the following approach is directly used [59]:

$$\delta\ (d)\ = \sum_{(i,j)\epsilon patchd} \sum \sigma(i,j) \tag{6}$$

Where, δ denoted as local activity/sharpness. $\sigma(i,j)$ estimates the local contrast. Previous research of NSS based quality assessment of image have shown that the behaviour of the coefficients of natural and distorted form of the images are efficiently captured by the generalised Gaussian distribution (GGD) [63]. The GGD with zero mean is modelled by [17]:

$$f\ (x;\ \alpha,\beta) = \frac{\alpha}{2\beta\gamma\frac{1}{\alpha}} exp\left(-\left(\frac{|x|}{\beta}\right)^{\alpha}\right) \tag{7}$$

The parameters of the GGD at (α, β), are reliably estimated by making use of moment-matching based approach [64]. The quality of the distorted image is depicted as the distance between the NSS feature model and the multivariate Gaussian model (MVG) [59]:

$$D(v_1, v_2, \Sigma_1, \Sigma_2) = \sqrt{(v_1 - v_2)^T (\frac{\Sigma_1 + \Sigma_2}{2})^{-1} (v_1 - v_2)} \tag{8}$$

In the above equation, the term, Σ_1, Σ_2 and v_1, v_2 are denoted as the mean covariance matrices and the mean vectors matrices of the natural MVG model and of the distorted image's MVG model. More information on this deviation can be found in [59]. These NSS features are formed from an assortment of natural (undistorted) images [65]. NIQE IQA, assess the quality or accuracy by determining the distance between the model statistics removed from natural images, and the distorted image [2]. The lower score of NIQE represents improved image qualities.

A.2.2.2.5 No-reference Free Energy Based Robust Metric

For a no-reference IQA, the only input the algorithm accepts is the image that requires the quality to be measured. No reference is required; therefore, it is called, no-reference or objective-blind. No-reference free energy based robust metric (NFERM) [66] is designed by adding human visual system (HVS) inspired features to enhance estimate performance and to reduce the over-all number of features to half [66]. NFERM divides the features used into three different categories. The first category consists of thirteen features. These features are of the free energy and the structural degradation information. This feature is derived from the reduced reference free energy-based distortion metric [67]. It defines the psycho-visual quality/accuracy as the concurrence between the image input and the resultant product of the internal generative model. The reduced reference structural degradation model is used to calculate the structural degradation information [68]. It can be deduced from the free energy theory that the HVS attempts to reduce the concern formed from the internal generative model when observing the input visual stimulus [66]. The author [66] applies a linear autoregressive model to estimate the generative model. This is done to obtain an image that the HVS observes as distorted or deformed [66]. The second category of features consists of six important HVS inspired features, that are calculated from the distorted and

predicted images [67]. The third category consists of four features that ascend from the NSS model. [67].

Mathematical expression for this metric can be found in the journal paper [66]. The NFERM model uses the free energy-based brain theory as well as features that are inspired from HVS [2]. This is used to determine the distortion of images. If the NFERM score is lower the output represents better image quality.

A.2.2.2.6 No-reference Image Quality Metric for Contrast distortion

No-reference image quality metric for contrast distortion (NIQMC) [69] metric generates a total quality estimate of a contrast-distorted image. This is done by joining local and global considerations. Locally, the metric focuses on regions with much information and computes the entropy in these regions [69]. Globally, the histogram of the input image is related to the uniformly distributed histogram that has the most amount of information by the symmetric "Kullback Leibler divergence" [70]. The NR-IQA model is built on the principle that an image that has better quality is more beneficial and contains valued information. The author [69] supposed that the HVS fuses the local and global consideration to observe a visual signal, determining its quality score and salient regions. From this assumption, the NIQMC model tries to envisage the visual quality of an enhanced image [69].

Mathematical expression: The NIQMC is defined as:

$$NIQMC = \frac{Q_L + \gamma Q_G}{1 + \gamma} \tag{8}$$

The constant weight that is utilised for regulating the relative importance among the local and global consideration is defined as γ. The Q_L and Q_G are the local and global quality measures, respectively. More information on this expression can be found in [69] . The NIQMC is used to assess the contrast quality of an image and it provides precise quality predictions for contrast distorted images. Better quality of the image contrast is represented by a higher NIQMC score.

A.2.2.2.7 Colourfulness-based Patch-based Contrast Quality

The colourfulness-based patch-based contrast quality index (CPCQI) [71] quality metric first extracts 17 features of a given image by analysing sharpness, contrast, brightness etc. The metric produces a score of visual quality utilising a regression model. The patch-based contrast quality index (PCQI) metric highly correlates with subjective quality scores on enhancement-relevant databases [72]. However, the metric does not consider the influence of colourfulness. Colourfulness is an important index for image quality assessment. Thus, Gu et al. [71], proposes the colourfulness-based PCQI (CPCQI) metric based on the limitation of PCQI.

Mathematical expression:

$$CPCQI = \frac{1}{M}\sum_{i=1}^{M} Q_{mi}(i).Q_{cc}(i).Q_{sd}(i).Q_{cs}(i) \quad (9)$$

Where Q_{mi} represents the similarity in terms of the mean intensity which is among the original/reference image and distorted images, Q_{cc} represents the contrast adjustment, and Q_{sd} represents the structural distortion. Information about these three terms are reported in journals such as [72]. The number of pixels is represented by M. The similarity of colour saturation is measured by Q_{cs} and is defined as:

$$Q_{cs}(i) = \left(\frac{2ST_1 \cdot ST_2 + \xi}{ST_1^2 + ST_2^2 + \xi}\right)^{\varphi} \quad (10)$$

Where ST_1 represents the colour saturation of the original and ST_2 represents the colour saturation of the distorted images. ξ is an exceedingly small constant number to refrain from division by zero. The fixed pooling index is defined as φ [72]. CPCQI is computed with the original/input image as a reference image. For this measure, higher score of CPCQI signifies better image contrast qualities. CPCQI is a measurement of perceptual distortions from the mean intensity, colour saturation for local patches, signal strength and signal structure [2]. If the CPCQI score is higher than 1, it infers that the output image is enhanced in comparison with the reference image. If the CPCQI score is lower than 1, it means that the detail in the image is not well enhanced, and/or artefacts may be introduced [2].

A.2.2.2.8 Blind/Reference less Image Spatial Quality Evaluator

A blind/reference less image spatial quality evaluator (BRISQUE) is NSS based [73]. This model functions in spatial domain. Distortion such as blocking, blurring or ringing cannot be computed with BRISQUE. BRISQUE is based on features that are derived from an empirical spreading of luminance and product of luminance under an NSS model that is locally normalised. The scene statistics measure the possibility of losses of naturalness of the image. These losses of naturalness are the result of the presence of distortions. BRISQUE is well suited for real time applications since it has very low computational complexity. BRISQUE can be used to identify distortion [73].

The *Mathematical expression* for this metric can be found in the journal paper [73]. BRISQUE function determines the BRISQUE score by utilising a support vector regression model. This model is reliant on a database of images with equivalent differential mean opinion score values [74]. The database has images that have familiar distortions. Some examples of distortions are: compression artefacts, noise and blurring. It also comprises of untouched forms of the distorted images. The input image requires at minimum one of the distortions that the model was trained

for [74]. Scores closer to 0 mean that the image has a good quality performance. From the above theoretical analysis, the experimental method for the experiment is presented.

A.3 Experimental Method

Aim: The aim is to address the selection of objective assessment metrics for contrast enhanced images and determined their relationship with the subjective assessment of human visual systems. This is done in order to establish which objective metrics best approximate human assessment and may therefore be used as an effective replacement for tedious human assessment surveys. When determining the quality of an image, there is no single quantitative metric that corresponds with human assessment. This problem has been highlighted in [1, 2]. The aim of this experiment is to assist with this problem by conducting a human visual and objective assessment. Human visual assessment (subjective evaluation) is used as a performance assessment method in literature [12]. The aim of the human visual assessment is to determine how respondents react to different enhancement methods; in terms of three variables, namely, image quality/preferences, naturalness and details. The goal is to find a relationship between these variables. The visual assessment was conducted using a panel that was asked to rate images according to human preference and choose an image that appeared most natural and another that contains the most detail. The goal of the objective metric experiment is two-fold. The first goal of this experiment is to identify objective IQA metrics that best approximate a human assessment; such that these metrics can be used in place of human assessment. The second goal is to categorise each metric in terms of the probable application domain.

Procedure: This experiment was performed on a PC with 8GB RAM and 2.2GHz CPU. All codes were implemented in MATLAB [75]. Some of the enhancement methods and performance metric codes were provided by the authors in MATLAB [2, 8, 59, 66, 69, 71]. The test images used in the experiment are from the BSD 300 dataset [76], Lossless dataset [77] and the USC-SIPI Image Database [78]. These images are used in various journal papers. The output of enhancements is found in Figure A.1 in the Results and Analysis section. To evaluate these enhancement models, a human assessment and an objective IQA is performed.

A human assessment contains visual assessment and analysis. A visual assessment study was conducted with 40 respondents (20 males and 20 females). The respondents were given a portfolio containing two sections with eight image sets each. Each image set contained six images. The survey questionnaire consisted of two sections. In Section 1 of the survey questionnaire, the respondents were asked to refer to the portfolio of images and score the images according to visual preference and image quality. Section 2 of the survey required respondents to choose an image according to details and natural appearance from the set of images in the portfolio.

The first eight image sets were used in Section 1 of the survey. A respondent is shown six images simultaneously; that is, the original test image is positioned on the top left of the page while the output images from the five enhancement algorithms are positioned randomly after the original test image. Respondents are requested to score the quality of each enhanced image by allocating one of the five numeric scores from 1 to 5 (1, 2, 3, 4, 5). Score "1" represents an annoying enhancement and the image is much worse than the original image i.e. the quality of the image is distorted. Score "3" is given when there is no clear enhancement, which suggests that the enhanced image is similar to the original image. Score "5" suggests that it is a substantial enhancement and the enhanced image has better quality than the original image. Other scores are selected in accordance with how respondents perceived image quality.

Section 2 of the survey determines which images are the most natural and unnatural; as well as the most detailed image in the set. An unnatural image is defined as an image that contains distortions or is perceived as over-enhanced. Section 2 contained eight image sets. Each image set contained the original test image and the output images processed by the five enhancements. These images are different from Section 1. For each image set, the respondent is shown six images at the same time. All the images in this section were placed randomly. The original image is unknown to the respondent. Section 2 of the survey is done to corroborate the results of Section 1 in the survey and to ensure that the respondent's view is accurate.

In addition to the survey, the author conducted a subjective analysis between the enhancement method used in the survey (CIEBHE, AELAB, NPEA, CEID, AEGMM). The information used in this experiment was gathered from the human panel and the authors of the enhancement methods.

All objective performance metrics mentioned in Section 2.1.2 were implemented in MATLAB. The different enhancement algorithms were assessed using these metrics. The images for this test were the same as the images used in the survey. IBM SPSS [79] software was used to capture all data. The results of both experiments are explained in the next Section (Results and Discussion).

A.4 Results and Discussion

This section discusses the results for: enhancement methods, the human assessment and the objective assessment study.

A.4.1 Enhancement results

Figure A.1 shows the output of the five enhancement algorithms that are applied to the original image.

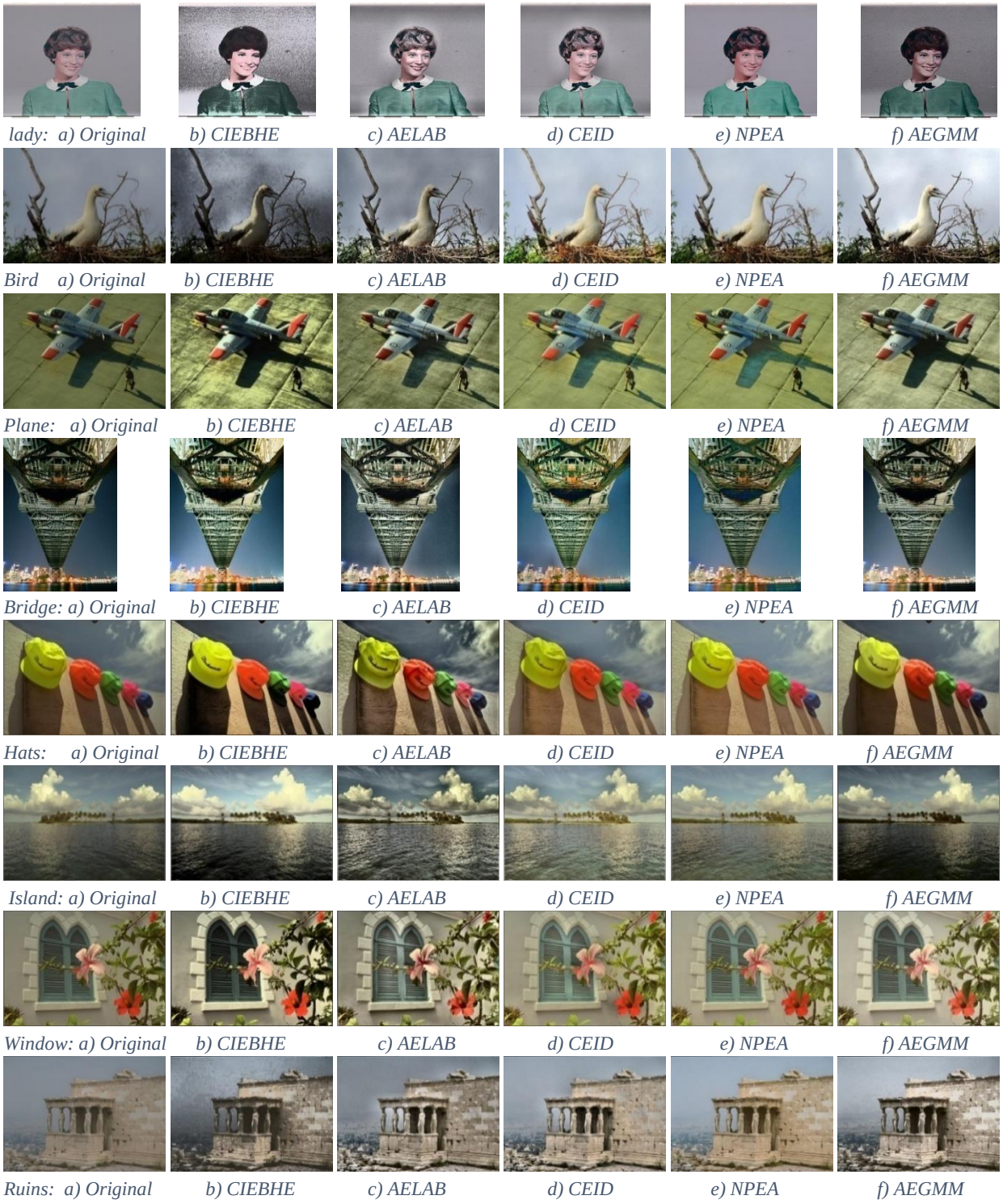


Figure A.1: Output images from five enhancement methods which is applied to the original image of the eight set of test images.

The average execution time consumed for each enhancement is presented in Figure A.2. The average time was calculated using the eight test images in Figure A.1. The implementation of the algorithm was done in MATLAB. From Figure A.2, it can be observed that CIEBHE enhancement

has the fastest execution time in comparison to the other four enhancement algorithms. The CEID algorithm has the slowest execution time.

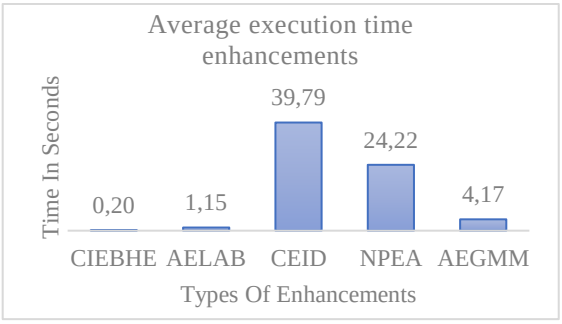


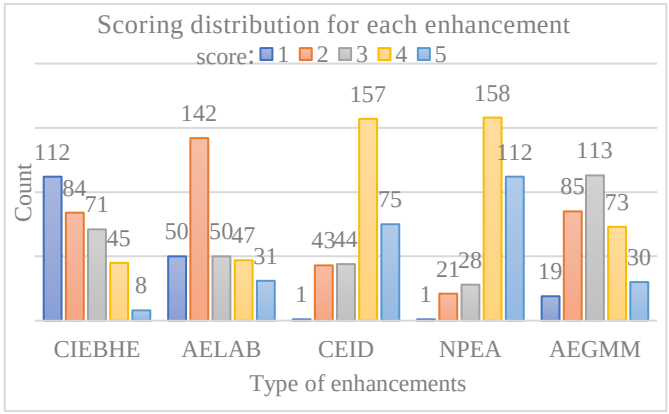
Figure A.2: Average execution time for each of the five enhancement methods in seconds

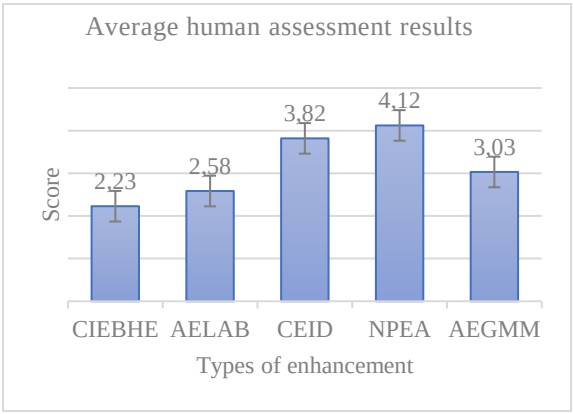
A.4.2 Human Assessment

The human assessment study is broken down into sub sections, i.e. human assessment scores and human perception of enhancements.

a) Human assessment scores

In Section 1 of the survey questionnaire, the respondents were asked to refer to the portfolio of images and rate the images according to visual preference. The score distribution from the survey-Section 1 is summarised in Figure A.3a. The maximum count for each enhancement is 320 (eight image set and 40 participants). The majority of the respondents scored the CIEBHE enhancement between scores 1 and 2. The majority of the scores for the AELAB enhancement images was score 2 and the AEGMM enhancement scored 3, thus the majority found AEGMM to produce an image that is similar to the original image. The most favoured results were given to NPEA and CEID enhanced images. Majority of scores given to the NPEA and CEID were between 4 and 5. All scores are uniformly distributed.



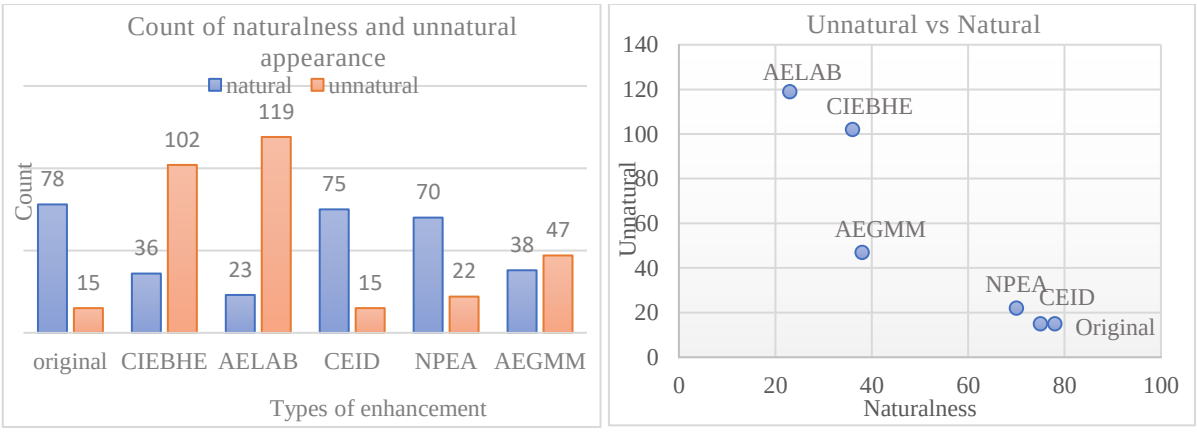


a) *b)*

Figure A.3: a) The graph represents the score distribution for each enhancement for Section 1 of the survey. Score range is from 1 (much worse) -5 (much better). b) The graph represents the average visual score received by each enhancement for Section 1 of the survey

Figure A.3b shows the average score received by each enhancement image. The error bars represent the standard error of measurements for 40 respondents. It can be observed that NPEA achieved the best score. This shows that most respondents favoured NPEA enhanced images; it was the most visually appealing image as compared to the other enhanced images, while CIEBHE was least favoured by participants. CIEBHE was not visually appealing to respondents. Even though AELAB and CIEBHE enhanced images were not favoured by respondents, these enhancements still provide important information and are favoured for other applications [7, 32]. Examples of these applications are underwater images or applications that require a lot of detail. CEID was the second favoured image enhancement followed by AEGMM.

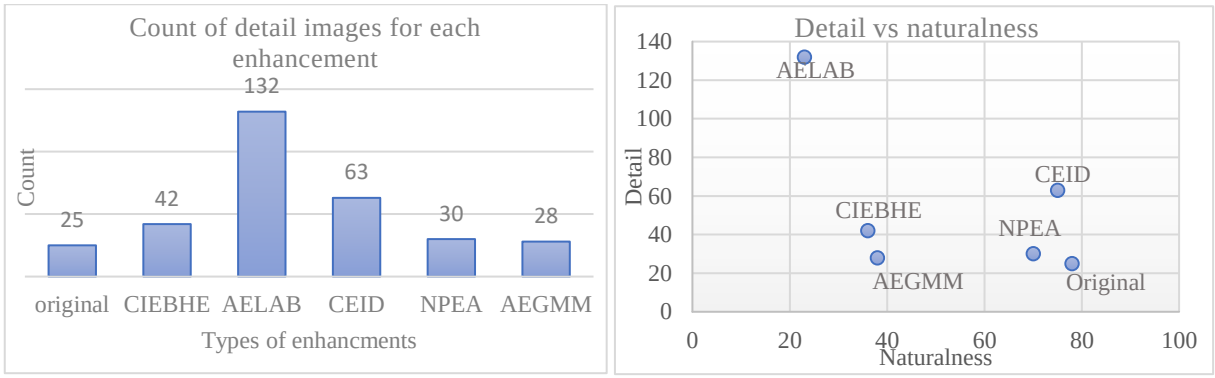
Section 2 of the survey asked the respondents to select images that are most natural, most unnatural and the most detailed image from eight different image sets. The maximum count for natural, unnatural and detail is 320 (eight image sets and 40 respondents). The summarised count distribution are presented in Figures A.4 and A.5. Figure A.4a shows the most natural and most unnatural image response received by respondents. The word unnatural is used to describe images that are over-enhanced and have distortions. CEID enhanced images were identified as the most natural image. The second most natural image chosen by respondents were NPEA enhanced images. There was a slight difference between NPEA and CEID scores. The image that were least chosen to be a natural image were ALEAB enhanced images. AELAB and CIEBHE enhanced images were chosen to be the most unnatural images by respondents. Figure A.4b, shows the graph of natural vs unnatural count; from this it can be observed that the natural and unnatural count are inversely proportional to each other. Enhancements that have the greatest number of response count for natural appearance also have the least response count for the enhancement to appear unnatural and vice versa.



a) *b)*

Figure A.4: a) Respondent's count for natural and unnatural image for each enhancement. This score is from Section 2 of the survey. b) The graph represents the relationship between natural and unnatural response count for each enhancement.

Figure A.5a shows the summary count for image detail for each enhancement. This part of the survey determines the images that appear to contain the most amount of detail. AELAB enhanced images were chosen to have the most detail present in an image. This enhancement received the highest detail count. CEID received the next highest count. The enhancement is known to improve the image while maintaining its naturalness and enhancing its detail. NPEA and AEGMM received the least amount of counts. Detail in NPEA and AEGMM enhancement are not distinct and not as visible compared to other enhancements. It has been observed that there is a relationship between detailed images and natural images. The images chosen to have an unnatural appearance, also contain more detail. Therefore, there is an inverse relationship between an image being perceived as natural and as detailed. This relationship can be seen in Figure A5.b. A possible reason for this is that the detailed images could be over-enhanced and appear to look unnatural.



a) *b)*

Figure A.5: a) The graph represents the respondent's count for detail images for each enhancement method. b) The graph represents the relationship between natural count and detail count.

b) Human perception of enhancements

The following information in this section was collected from the panel of recipients and the author's evaluation of their own enhancement algorithms. To evaluate all these enhancement methods, they are subjectively analysed. The output of these enhancements is found in Figure A.1.

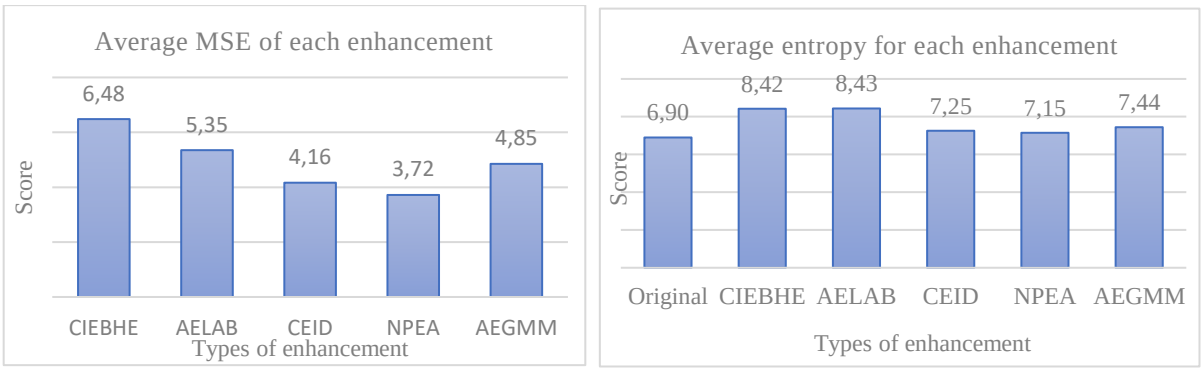
In Figure A.1- 'Island', AELAB enhanced image appears to be over-enhanced and has colour artefacts. This can be observed in the clouds; it also results in shadows in some regions. Even though the image is over-enhanced, the water region of the image shows more detail than other enhancements. It can also be observed that CIEBHE enhanced image also has colour artefacts and results in dark regions. The AEGMM enhanced image produces a dark image in comparison to the original image. CEID produces a natural image with some over-enhanced regions; this over-enhancement can be seen in the clouds. NPEA image produces a natural image and all aspects of these images are clear.

In Figure A.1- ‘bird’ shows that CEID and NPEA enhanced images produce similar output. CEID has more colour components and is more visually preferred. This is because the author integrated colour information of the image with the decomposed reflectance’s layer to produce the colourful reflectance [2]. Both NPEA and CEID images look natural. CIEBHE image enhancement produces a distorted and noisy image. AEGMM enhanced image is similar to the original image. In some areas it is better and brighter than the original and in others it has colour artefacts and results in dark regions. AELAB enhanced image is over-enhanced and looks unnatural.

Overall: Each algorithm is designed for certain applications. Some enhancements are designed to enhance details of an image while some are used to enhance an image for human pleasure. CIEBHE enhances an image by increasing its contrast i.e. by altering its intensity distribution. CIEBHE performs well for uniform bright or dark images [2]. CIEBHE produces unsatisfactory enhancements for non-uniform images. It also generates visible artefacts in certain regions of the image. One such example is the sky region (in Figure A.1 ‘ruin’) which is corrupted by noise. AELAB images show more detail than other enhancements but tend to produce over-enhanced images and generate over-sharpness image results for regions in the images. This method works well for underwater images [7] or applications that require a lot of detail. CEID images can avert the artefacts and efficiently enhance non-uniform illumination images. It attains a favourable balance among detail boosting and noise suppression and naturalness [2]. CEID images perform well for a broad range of images. NPEA enhanced images produce the most natural images. NPEA is good at maintaining the naturalness for non-uniform illumination images and has the intent to enhance the detail of the image but there is loss in detail for certain regions. NPEA demonstrates that the enhancement produces visually pleasing, free of artefact, and natural looking image. AEGMM produces images with a decent amount of enhancements. Some areas tend to remain dark. An example of this can be observed in the image ‘lady’. AEGMM introduces an automatic image enhancement that utilises a GMM of a given input image which produces a visually pleasing resultant image of various types of images.

A.4.3 Objective Image Quality Assessment

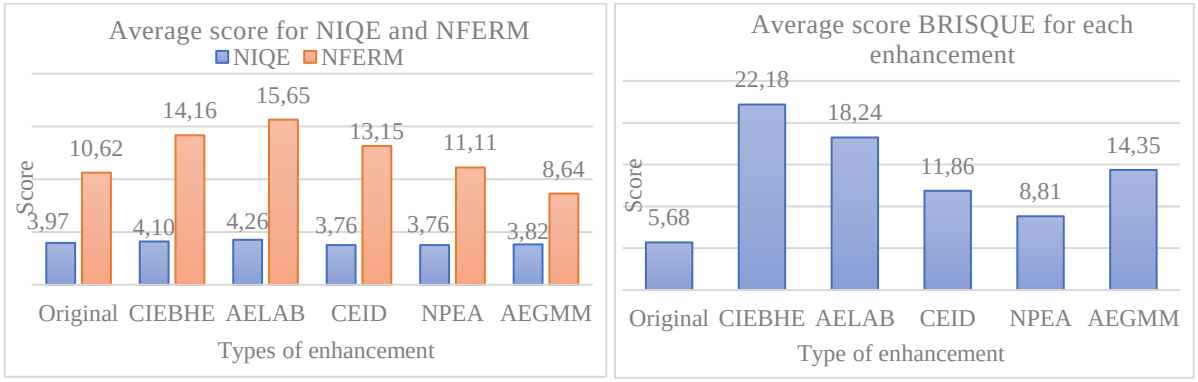
The MSE can be utilised to evaluate the change between the original image and the enhanced image. MSE is not used as an error measure in our application but rather as an indicator for absolute change from the original image. This change can be either good or bad. The lower the value of MSE, the smaller the difference is compared to the original image. From the graph in Figure A.6a, NPEA has the lowest value of MSE and the smallest difference. Therefore, it is assumed to be most similar to the original image



a) *b)*

Figure A.6: a) Graph of the average MSE scores for each of the five enhancement methods. b) Graph of average score for entropy for the enhancement methods and the original image.

Figure A.6b shows the average entropy score of eight image sets. Higher entropy values usually indicate more detail/information. The enhancement that has the highest average value of entropy is AELAB and CIEBHE, which means that this enhancement could have the most amount of information in terms of detail or noise. NPEA has the lowest score of entropy from all the images. This means that NPEA enhancement does not enhance the information as well as the other enhancements. All enhancements have an entropy score greater than the original image, i.e. the enhancements contain more information than the original image



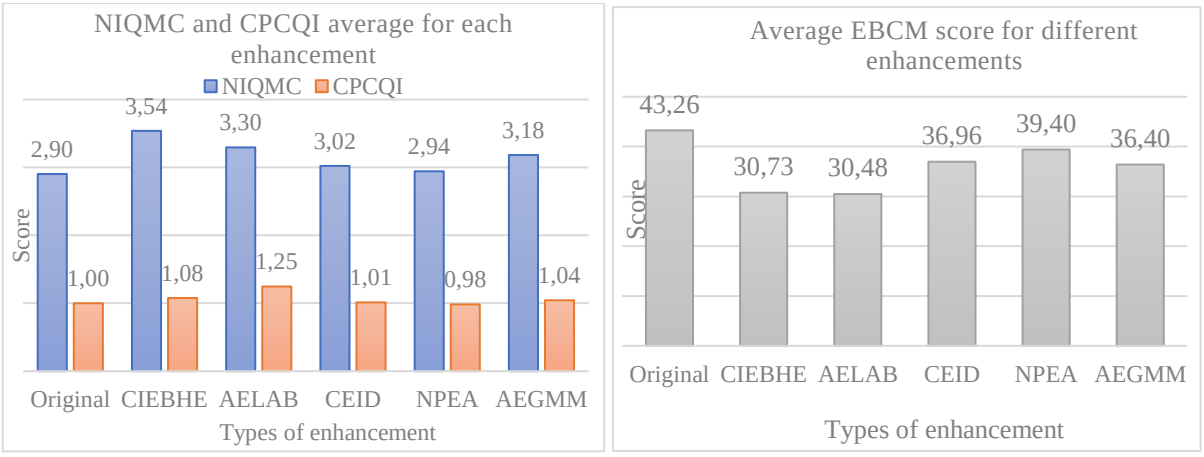
a) *b)*

Figure A.7: a) The graph represents the average score of NIQE and NFERM for enhancement methods and the original image. b) The graph represents the average score for BRISQUE for each enhancement method and the original image.

The NIQE and NFERM are used to assess the quality of an image blindly. For NIQE and NFERM metric, lower values represent better qualities. Figure A.7a shows the average score for NIQE and NFERM metrics for the original images and the resultant images produced by five different contrast enhancement algorithms for all eight images. The image enhancement with the smallest score for both NIQE and NFERM, implies that the resultant image is the most similar to natural images and have the minimum distortion/deformities. The results of enhancements that have larger NIQE and NFERM scores, is due to the over-enhancement and suffer from many artefacts. AEGMM has the lowest score for NFERM, and a low NIQE. CEID and NPEA have the smallest scores for NIQE and a fairly small score for NFERM. The large NFERM scores for AELAB and

CIEBHE could be due to the over-enhancing or the introduction of artefacts and the appearance of dark regions that appear in these images.

Figure A.7b shows the graph of average scores for BRISQUE. The lower the score is, the better the quality of the image. NPEA scores the lowest BRISQUE score from all the enhancements and the CIEBHE scores the highest BRISQUE score. Therefore, NPEA and CEID have best quality performance in this test. From these three metrics it can be assumed that NPEA and CEID have the best quality and the naturalness in resultant image is preserved.



a) *b)*

Figure A.8: a) The graph of the average score for NIQMC and CPCQI for each enhancement and the original image. b) The graph of the average score for ECBM for each enhancement and the original image.

Figure A.8a shows the graph for average scores for NIQMC and CPCQI. NIQMC is used to determine contrast distortions and CPCQI are used to assess the contrast enhancement quality for an image. For both measurements, larger scores represent that the image contrast quality has improved, or it is better than the original [2].

From Figure A8.a, it can be noted that CIEBHE and ALEAB have the larger NIQMC scores. The possible explanation is that the resultant images from these enhancements are over-enhanced, as explained in the Section A.4.2.3. When images are over-enhanced it can be perceived as unnatural. AELAB produces resultant images that are over-enhanced in certain regions, which is not preferable. The NIQMC score for NPEA are smaller since certain details are difficult to observe.

If the score for CPCQI metric is higher than 1 it means that the resultant image is much more enhanced in comparison to the original input image and conversely, if the scores are lower than 1 it means that the details are barely improved or enhanced, or artefacts are produced [2]. CEID and NPEA have the score closest to one, which shows that this method produces a favourable balance between increasing detail and suppressing artefacts.

For the most improved contrast, assume there is an image named X and another image named Y. If the EBCM of X is greater than the EBCM of Y, then the contrast of the image is improved. From Figure A.8b, it is noted that NPEA has the best contrast from all the enhancements. AELAB has the worse contrast in comparison to the other four enhancements.

Analysis of results: Existing objective measurements are reasonable at assessing the specific features of an enhanced image [2]. The selection of a suitable assessment method is reliant on the purpose and objective of the task. Humans are perceptive to artefacts and naturalness; this is not measured well in current objective measures. From the results of the visual assessment, it is observed that there is a relationship between detailed images and naturalness. After some point, the more detailed an image is, the less natural it appears. Objective IQA for image enhancement can be beneficial by mixing effective features to measure artefacts [2]. The metrics studied evaluates specific image characteristics and do not fully determine human preference. However, some do approximate aspects of human perception and serve as effective predictors of these aspects.

From the human visual assessment, it was determined that the most natural enhancement was NPEA. The objective assessment further validated this result. The NIQE, NFERM and BRISQUE metric was used to determine this. The visual assessment determined which enhancement was most detailed. The objective assessment validated this finding. The entropy and NIQMC metric were used to determine the detail/information of an image. The human assessment study determined which image was the most visually pleasing to a human panel. This was not determined with the eight-objective metrics. From the entropy and NIQMC, it was determined that AELAB contains the most detail/information. This result was also found in the human assessment test. The objective metric can determine single characteristics of the HVS. Therefore, two or more metrics are required to replicate a human assessment. Individual metrics assess individual characteristics of an enhancement.

The eight-performance metrics can be classified in terms of the domain application. Entropy and NIQMC metrics can be utilised to assess the amount of information an image contains. NIQE, NFERM and BRISQUE emphasise naturalness of an image and measure the quality of the image while CPCQI, NIQMC and EBCM metrics can be used to measure the contrast of an image.

A.5 Conclusion

Existing objective measurements are reasonable at assessing the specific features of an enhanced image. The selection of assessment metrics is reliant on the purpose and objective of the task. A subjective assessment testing was done to yield the most favoured, natural and detailed images. The results from the visual assessment showed that there is a relationship between detail and

naturalness. The more detailed an image is, the less naturalness is perceived. Images that look natural have moderate detail. The objective metrics cannot determine which image is perceived better, but the human observer can. The objective metrics are useful in evaluating the special characteristics of enhanced images but cannot determine human preference. Significant difficulty arises when finding the suitable evaluation of algorithms that provides an objective measurement of performance. There is no single quantitative metric that corresponds well with the quality of an image as perceived by the HVS when assessing image quality. More than one metric must be used to estimate human perception. This paper helps with choosing metrics that are used for the evaluation of contrast enhanced images. It also categorises the performance metrics in terms of domain application.

A.6 References

[1] M. Wirth, M. Fraschini, M. Masek and M. Bruynooghe, "Performance Evaluation in Image Processing," *EURASIP Journal on Applied Signal Processing,* vol. 2006, pp. 1-3, 2006.

[2] H. Yue, J. Yang, X. Sun, F. Wu and C. Hou, "Contrast Enhancement Based on Intrinsic Image Decomposition," *IEEE transactions in image processing,* vol. 26, no. 8, pp. 3981 - 3994, 8 August 2017.

[3] B. Chitradevi and P.Srimathi, "An Overview on Image Processing Techniques," *International Journal of Innovative Research in Computer and Communication Engineering,* vol. 5, no. 11, November 2014.

[4] A. K. Boyat and B. K. Joshi, "A Review Paper: Noise Models In Digital Image Processing," *Signal & Image Processing : An International Journal (SIPIJ),* vol. 6, pp. 2, 2015.

[5] V. Murugan and R. Balasubramanian, "An Efficient Gaussian Noise Removal Image Enhancement Technique for Gray Scale Images," *International Journal of Computer, Electrical, Automation, Control and Information Engineering,* vol. 9, no. 3, 2015.

[6] S. Arora and K. Kapoor, "Colour image enhancement based on histogram equalization," *Electrical & Computer Engineering: An International Journal (ECIJ),* vol. 4, no. 3, 2015.

[7] S. Bharal, "L*a*b based contrast limited adaptive histogram equalization for underwater images," *International Journal of Computer Application (2250-1797),* vol. 5, no. 4, June 2015.

[8] S. Wang, J. Zheng, H.-M. Hu and B. Li, "Naturalness Preserved Enhancement Algorithm for non-uniform illumination images," *IEEE Transactions on Image Processing,* vol. 9, no. 9, 2013.

[9] T. Celik and T. Tjahjadi, "Automatic Image Equalization and Contrast Enhancement Using Gaussian Mixture Modeling," *IEEE Transactions on Image Processing,* vol. 21, no. 1, 2012.

[10] P. Mohammadi, A. Ebrahimi-Moghadam and S. Shirani, "Subjective and Objective Quality Assessment of Image: A Survey," Elsevier, 2014.

[11] Y. J. Zhang, "Evaluation and comparison of different segmentation algorithms," *Pattern Recognition Letters,* vol. 18, no. 10, pp. 963–974, 1997.

[12] D. S. Prabha and J. S. Kumar, "Performance Evaluation of Image Segmentation using Objective Methods," *Indian Journal of Science and Technology,* vol. 9, no. 8, 2016.

[13] H. H. Barret, J. L. Denny, R. F. Wanger and K. J. Myers, "Objective assessment of image quality. II. Fisher information, Fourier crosstalk, and figures of merit for task performance," *Journal of the Optical Society of America A,* vol. 12, pp. 834-852, 1995.

[14] Z. Wang and A. C. Bovik, "Modern image quality assessment Synthesis Lectures on Image, Video & Multimedia Processing," Morgan & Claypool Publishers, 2006.

[15] A. Rehman and Z. Wang, "Reduced-reference image quality assessment by structural similarity estimation," *IEEE Trans. Image Processing,* vol. 21, pp. 3378-3389, August. 2012.

[16] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)," *IEEE Trans. Image Processing,* vol. 18, pp. 717-728, April 2009.

[17] M. Saad, A. C. Bovik and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Transactions on Image Processing,* vol. 21, no. 8, pp. 3339–3352, 2012.

[18] BT.500-11, ITU-R recommendations, Methodology for the subjective assessment of the quality of television pictures, Geneva, Switzerland: ITU, 2002.

[19] BT.710-4, ITU-R recommendations, Subjective assessment methods for image quality in high-definition television, Geneva, Switzerland: ITU, 1998.

[20] BT.814-1, ITU-R recommendations, Specification and alignment procedures for setting of brightness and contrast of displays: ITU, 1994.

[21] T. Acharya and A. K. Ray, Image Processing: Principles and Applications, Wiley-Interscience, 2005.

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

[23] Y. Wang, Q. Chen and B. Zhang, "Image enhancement based on equal area dualistic subimages histogram equalization method.," *IEEE Transactions on Consumer Electronics,* vol. 45, no. 1 , pp. 68 - 75, 1999.

[24] S.-D. Chen and A. Ramli, "Minimum mean brightness error bi-histogram equalization in contrast enhancement.," *IEEE Transactions on Consumer Electronics,* vol. 49, no. 4, pp. 1310 - 1319 , November 2013.

[25] S.-D. Chen and A. Ramli, "Contrast enhancement using recursive mean-separate histogram equalization for scalable brightness preservation equalization for scalable brightness preservation.," *IEEE Transactions on Consumer Electronics,* vol. 43, no. 4, 2003.

[26] Matlab, "Understanding Color Spaces and Color Space Conversion," Mathworks, [Online]. Available: https://www.mathworks.com/help/images/understanding-color-spaces-and-color-space-conversion.html. [Accessed 27 March 2018].

[27] M. Recky and F. Leberl, "Windows detection using K-means in CIE-LAB color space," in *20th International Conference on Pattern Recognition, ICPR 2010,* Istanbul, Turkey, 2010.

[28] J. M. and S. K. Mitra, "Enhancement of Colour Images by Scaling the DCT Coefficients," vol. 17, no. 10 pp. 1783-1794, 2008.

[29] HunterLab, "Application notes: CIE L*a*b* Color Scale," vol. 8, no. 7, 1996.

[30] H. Mahmoud, L. Abdulhadi, A. Mahmoud and H. Mohammed, "Comparsion of Spectrometer, Camera and Scanner Reproduction of Skin Colour," in *5th Kuala Lumpur International Conference on Biomedical Engineering*, Malaysia, 2011.

[31] G. Jeon, "Colour image enhancement by histogram equalization in heterogeneus colour space," *International Journal of Multimedia and Ubiquitous Engineering,* vol. 9, no. 7, pp. 309-318, 2014.

[32] D. J. Bora, "Importance of Image Enhancement Techniques in Color Image Segmentation: A Comprehensive and Comparative Study," *International Journal of Science and Research,* vol. 15, no. 1, 2017.

[33] A. Alam, M. Abdullah and Ravi Shankar Mishra, "Colour Contrast Enhancement Method by Scaling the DC Coefficients in CIE-LAB Colour Space," *International Journal of Computer Applications,* vol. 97, no. 22, July 2014.

[34] R. Kimmel, M. Elad, D. Shaked, R. Keshet and I. Sobel, "A variational framework for Retinex," *SPIE electronic imaging,* vol. 52, no. 1, pp 7-23, 2003.

[35] X. Fu, Y. Sun, M. LiWang, Y. Huang, X.-P. Zhang and X. Ding, "A novel Retinex based approach for image enhancement with illumination adjustment," *2014 IEEE International*

Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1190-1194, May 2014.

[36] Z. Liang, W. Liu and a. R. Yao, "Contrast enhancement by nonlinear diffusion filtering," *IEEE Transactions on Image Processing* , vol. 25, no. 2, pp. 673-686, February 2016.

[37] H. Barrow and J. Tenenbaum, "Recovering intrinsic scene characteristics from images," *Computer Vision Systems, A Hanson & E. Riseman (Eds.),* pp. 3-26, 1978.

[38] E. Garces, A. Munoz, J. Lopez-Moreno and D. Gutiérrez, "Intrinsic images by clustering," *Computer Graphics Forum,* vol. 31, no. 4, pp. 1415-1425, 2012.

[39] J. Shen, X. Yang, X. Li and Y. Jia, "Intrinsic image decomposition using optimization and user scribbles," *IEEE Transactions on Cybernetics,* vol. 43, no. 2 , pp. 425-436, April, 2013.

[40] S. Bell, K. Bala and N. Snavely, "Intrinsic images in the wild," *ACM Transactions on Graphics,* vol. 33, no. 4, pp. 159, 2014.

[41] J. T. Barron and J. Malik, "Shape, illumination, and reflectance from shading," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 37, no. 8, pp. 1670 - 1687, 2013.

[42] A. M. Reza, "Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement," *Journal of VLSI signal processing systems for signal, image and video technology,* vol. 38, no. 1, pp. 35–44, 2004.

[43] S. Chen and A. Beghdadi, "Natural Rendering of Color Image based on Retinex," *2009 16th IEEE International Conference on Image Processing (ICIP),* pp. 1813-1816, 2009.

[44] S. Chen and A. Beghdadi, "Natural enhancement of color image," *EURASIP J. Image Video Processing,* vol. 2010, pp.1–19, Jan 2010.

[45] B. Li, S. Wang and Y. Geng, "Image enhancement based on Retinex and lightness decomposition," *Proc. IEEE Int. Conference on Image Processing,* pp. 3417–3420, 2011.

[46] E. Hecht, "Optics," MA,USA, Addison Wesley, 2002, pp. 86-149 ch.4.

[47] D. Jobson, Z. Rahman and G. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on Image Processing,* vol. 6, no. 7, pp. 965-976, 1997.

[48] J. Mukherjee and S.Mitra, "Enhancement of color images by scaling the DCT coefficients," *IEEE Trans. Image Process.,* vol. 17, no. 10, pp. 1783–1794, October 2008.

[49] S. Agaian, B. Silver and K. Panetta, "Transform coefficient histogram-based image enhancement algorithms using contrast entropy," *IEEE Transaction on Image Processing,* vol. 16, no. 3, pp. 741-475, March 2007.

[50] C.Wang and Z.Ye, "Brightness preserving histogram equalization with maximum entropy: A variational perspective," *IEEE Transactions on Consumer Electronics,* vol. 51, no. 4, pp. 1326 - 1334, November 2005.

[51] C. Wang, J. Peng and Z. Ye, "Flattest histogram specification with accurate brightness preservation," *IET Image Processing,* vol. 2, no. 5, pp. 249–262, October 2008.

[52] S. Hashemi, S. Kiani, N. Noroozi and M. E. Moghaddam, "An image contrast enhancement method based on genetic algorithm," vol. 31, no. 13, pp. 1816–1824, October 2010.

[53] L. K. Owens, "Introduction to survey research design," 2002.

[54] M. Matlab, "Mean Square Error," [Online]. Available: https://www.mathworks.com/help/images/ref/immse.html. [Accessed 28 March 2018]

[55] S. Rani, A. Kumar and K. Singh, "Illumination based Sub Image Histogram Equalization: A Novel Method of Image Contrast Enhancement," *International Journal of Computer Applications,* vol. 119, no. 20, 2015.

[56] Z. Zhang and P. Xu, "An efficient learning-based bilateral texture filter for structure perserving," in *Next Generation Computer Animation Techniques*, Springer international publishing, 2017.

[57] Shashi, "Matlab," Matlab, 28 February 2012. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/35365-edge-based-contrast-measure-for-image-enhancement-quality-assessment-focused=5221432&tab=function. [Accessed 25 June 2018].

[58] R. C. Gonzalez and R. E.Woods, Digital Image Processing, Upper Saddle River: Prentice-Hall, 2006.

[59] A. Mittal, R. Soundararajan and A. C. Bovik, "Making a Completely Blind Image Quality Analyzer," *IEEE Signal Processing Letters,* vol. 20, no. 3, pp. 209-212, 2012.

[60] Matlab, "Naturalness Image Quality Evaluator (NIQE) mode," Matlab, [Online]. Available: https://www.mathworks.com/help/images/ref/niqemodel.html. [Accessed 10 July 2018].

[61] L. Zhang, L. Zhang and A. C. Bovik, "A Feature-Enriched Completely Blind Image Quality Evaluator," *IEEE Transactions On Image Processing,* pp. 2579 - 2591, 2015.

[62] "Laboratory for Image & Video Engineering," The university of Texas Austin, [Online]. Available: http://live.ece.utexas.edu/research/Quality/nrqa.htm. [Accessed 10 July 2018].

[63] A. K. Moorthy and A. C. Bovik, "Statistics of natural image distortions," in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2010.

[64] K. Sharifi and A. Leon-Garcia, "Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video," *IEEE Trans. Circuits Syst. Video Technology,* vol. 5, no. 1, pp. 52–56, 1995.

[65] S. N. Holambe, U. B. Shinde and P. M. Kshirsagar, "A Brief Review on Blind Image Quality Evaluation Methods," *International Journal of Computer Applications,* vol. 163, no. 6, April 2017.

[66] K. Gu, G. Zhai, X. Yang and W. Zhang, "Using Free Energy Principle For Blind Image Quality Assessment," *IEEE Transactions on Multimedia,* vol. 17, no. 1, 2015.

[67] G. Zhai, X. Wu, X. Yang, W. Lin and W. Zhang, "A psychovisual quality metric in free-energy principle," *IEEE Transactions on Image Processing ,* vol. 21, no. 1, pp. 41–52, 2012.

[68] K. Gu, G. Zhai, X. Yang and W. Zhang, "A new reduced-reference image quality assessment using structural degradation model," vol. 17, no. 1, pp 50-63, 2015.

[69] K. Gu, W. Lin, G. Zhai, X. Yang, W. Zhang and C. W. Chen, "No-Reference Quality Metric of Contrast-Distorted Images Based on Information Maximization," *IEEE Transactions on Cybernetics,* vol. 47, no. 12, 2017.

[70] M. Pedersen, O. Cherepkova and A. Mohammed, "Image Quality Metrics for the Evaluation and Optimization of Capsule Video Endoscopy Enhancement Techniques," Norway, 2015.

[71] K. Gu, D. Tao, J.-F. Qiao and W. Lin, "Learning a No-Reference Quality Assessment Model of Enhanced Images With Big Data," *IEEE transactions on neural networks and learning systems,* vol. 29, no. 4, pp. 1301-1313, April 2018.

[72] S. Wang, K. Ma, H. Yeganeh, Z. Wang and W. Lin, "A patchstructure representation method for quality assessment of contrast changed images," *IEEE Signal Process. Lett.,* vol. 22, no. 12, pp. 2387-2390, December 2015.

[73] A. Mittal, A. K. Moorthy and A. C. Bovik, "No-Reference Image Quality Assessment in the Spatial Domain," *IEEE Transactions on Image Processing ,* vol. 21, no. 1, December 2012.

[74] Mathworks, "brisque," [Online]. Available: https://www.mathworks.com/help/images/ref/brisque.html. [Accessed 01 August 2018].

[75] MathWorks, "Matlab installation," Mathworks, [Online]. Available: https://www.mathworks.com/products/matlab.html. [Accessed 20 December 2017].

[76] " Berkeley Segmentation Dataset and Benchmark," [Online]. Available: https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/. [Accessed May 2018].

[77] Kodak images, "Kodak Lossless True Color Image Suite," [Online]. Available: htttp://r0k.us/graphics/kodak/. [Accessed May 2018].

[78] University of Southern California, "The USC-SIPI Image Database," [Online]. Available: http://sipi.usc.edu/database/. [Accessed May 2018].

[79] IBM SPSS, "IBM SPSS software," IBM, [Online]. Available: https://www.ibm.com/analytics/spss-statistics-software. [Accessed 30 June 2016].

B. Paper 2: Multimodal Enhancement-Fusion technique for Natural Images

Ms. Rivania Maharaj and Mr. Bashan Naidoo

Submitted to ELSEVIER- Signal Processing- European Association for Signal Processing (EURASIP) (Under Review)

Abstract

This paper proposes a multimodal enhancement-fusion (MEF) technique for natural images. The MEF is expected to contribute value to machine vision applications and personal image collections for the human user. This technique identifies and merges desired attributes from parallel enhancement pathways into the resultant image. The proposed MEF focuses on improving chromatic irregularities such as poor contrast distribution. It also proposes a concurrent enhancement pathway that subjects an image to multiple image enhancers in parallel, followed by a fusion algorithm that creates a composite image that combines the strengths of each enhancement path. This study develops a global framework for parallel image enhancement, followed by parallel image assessment and region selection, leading to final merging of selected regions from the enhanced set. The resultant output combines desirable attributes from each enhancement pathway to produce a result that is superior to each path taken alone. The results demonstrate that the proposed MEF technique performs well for most image types. It is subjectively favourable to a human panel and achieves better performance for objective image quality assessment compared to other enhancement methods.

B.1 Introduction

Imaging sensors seldomly produce ideal raw images. A sensory output is usually subjected to a variety of corrective algorithms before becoming useful. This is especially prevalent in machine vision applications [1]. There are various sources of error that exist; for example, lens distortion, sensor dynamic range limitations, thermal distortion, etc. These errors can be corrected by using image processing techniques such as extraction, filtering, image enhancement and fusion techniques. The general image enhancement domain is actively researched. Image contrast enhancement is a conventional and key field of image processing and has been broadly adopted in various applications. Some examples of these applications are; traffic control systems, medical imaging, remote-sensing imagery, daily photo enhancement [2], etc. There are various state-of-the-art enhancement techniques that exist for targeted and global enhancement, however, every state-of-the-art method possesses its own favourable/ unfavourable characteristics. There is no method that is optimal for all image types. This paper designs a multimodal enhancement-fusion (MEF) technique for natural images which proposes to adaptively select the optimal enhancer for every image region. This is done with the intention to overcome unfavourable characteristics. This MEF technique also focuses on improving chromatic irregularities such as poor contrast distribution and distortions. The technique proposes a concurrent enhancement pathway that subjects an image to multiple image enhancers in parallel, followed by a fusion algorithm that creates a composite image that combines the strengths of each enhancement path. The MEF can

be tailored such that it selects and merges desired regions from each pathway into the resultant image. The algorithm is designed to create a resultant image that maintains naturalness and enhances detail.

Human perception and observation work well when developing a visual representation of an image. This visual representation consists of vivid interpretation of colour and the ability to perceive detail across a broad-range of photometric levels as a result of lighting differences [3]. It is difficult to develop an image that relates well with the direct perception of a scene in an environment. The human visual system (HVS) processes scene illumination non-linearly. Some image processing algorithms are designed to improve an image's illumination and contrast.

Image contrast enhancement adjusts the contrast of an image to produce an improved image. Some contrast enhancement techniques operate well on images that have a uniform spatial distribution of gray values while other images may not, and a loss of clarity of detail and colour may arise [4]. According to survey [4, 5, 6], enhancement methods are categorised into two groups, namely, spatial domain methods and frequency domain methods. In spatial domain methods, pixels of an image are manipulated directly to attain the desired result. In frequency domain methods, an image is converted to frequency domain by utilising a Fourier transform and then processing techniques are applied to the image. After applying the processes on the image, the inverse Fourier transform is utilised to convert the image back to the spatial domain and thus the final image is obtained. An image may be represented in both spatial or frequency domain.

There are many fusion techniques that have been designed for both domains. Image fusion is an application of image processing where images are fused together to produce a single image. The main aim of any fusion algorithm is to combine all important visual information from various images into one image. The resultant image contains more information than the individual enhancements and more accuracy, without producing any artefacts. A good fusion method preserves useful information and does not create any artefacts that can mislead a human observer. It should also be reliable, robust and not disregard any salient information from the input images [7]. There are many fusion techniques that currently exist such as Wavelet transform [8], Multiscale transform-based fusion [9], Laplacian pyramid based [10] etc. According to a study conducted by Maharaj et al. [11], performance metrics were categorised in terms of domain application, that could be used to assess enhancement methods. Every enhancement method has some sort of limitation. Therefore, a multimodal enhancer is proposed to adaptively select the optimal enhancer for every image region. The proposed MEF takes parallel enhancement and fuses the desired regions from the different images. The enhancements were chosen from the study [11]. The enhancement algorithms were chosen because they produce a resultant image that maintains naturalness and is visually pleasing.

In this study, the parallel enhancer combines three enhancement pathways, however, any number can be used. These algorithms are discussed in Section B.2. The resultant output of the MEF, combines desirable attributes from each enhancement pathway to produce a result that is superior to each path taken alone. The MEF technique is based on the theoretical modelling of high dynamic range imaging (HDRI) [12] and the concept of Saleem et al. [10]. HDRI covers a collection of image processing methods that allow for a larger dynamic range of exposures than normal image processing techniques [13]. Saleem et al. proposes a fusion based enhancement using Laplacian pyramid decomposition. The proposed method makes use of both approaches by introducing the concept of parallel enhancements and region selection for fusion which will maintain naturalness and enhance detail. In addition, when assessing the quality or accuracy of enhancement methods, there is no single metric that corresponds with human assessment. Therefore, objective and subjective testing was done on the parallel enhancement methods and the proposed MEF technique.

The structure of the paper is as follows. A summary of related works is presented in Section B.2, i.e. parallel enhancement methods and the fusion algorithms. Section B.3 presents the proposed MEF framework. Section B.4 provides the experimental method followed by Section B.5 which shows the results of the experiment and the related analysis. Section B.6 concludes the paper.

B.2 Related works

Several fusion and enhancement models have been proposed in the past. The proposed MEF technique makes use of three powerful enhancement models which were identified because of their naturalness and visual preference in Maharaj et al, study [11]. This section provides a summary of works related to the MEF technique and is followed by a discussion of the three enhancement algorithms that are used to form our parallel pathway of image enhancers.

B.2.1 Image Fusion techniques

Image fusion is a technique that combines vital information from many image sources into a single resultant image [14]. The intent for image fusion is to decrease the quantity of information and to generate an image that is more acceptable and comprehensible for human observation and machine perception [7]. There are different levels of abstraction of information for image fusion: i.e. the pixel level, decision level, signal level and feature level. There are many fusion methods. This section provides a summary for two fusion methods, namely, Saleem et al. method [10] and Mertens et al. method [15].

Saleem et al. [10] proposed the image fusion-based contrast enhancement that balances global and local enhancement requirements. The fusion is processed in a multiresolution manner. It

utilises the Laplacian pyramid decomposition in order to consider the multi-channel aspects in the HVS. This method decomposes the input image into a hierarchy of images in which individual level correlates to a diverse band of frequencies in an image. Thereafter the Gaussian pyramid weight map is computed. This is required for blending. This method demonstrates a fusion approach that achieves a good compromise between different attributes of contrast enhancement. This is done to obtain a pleasing result. The results show that it efficiently enhances the local and global contrast without upsetting the colour equilibrium and it is not appropriate for real-time image processing applications which require a resultant image that has high quality [10].

Exposure fusion was introduced by Mertens et al. [15]. It refers to the process where a registered set of multiple exposures are blended together into a single image. The author [15] introduces a method that instantly fuses many exposures into a single image of high quality and controlled dynamic range (DR), that is optimised for viewing medium. This method produces an improved image but sometimes tends to create an unnatural appearance. This perception is created because the method favours intensity [15]. Another image processing method that requires images of different exposure levels is high dynamic range (HDR) imaging [15]. HDRI is an HDR technique that is utilised in image applications, to produce a superior DR of luminosity than is probable with standard image processing techniques. Other examples of fusion techniques are [8, 16, 17, 18].

HDR imaging techniques are a prime area of focus because of the theoretical importance as well as practical importance [19]. There are many advantages to HDR imaging. These include greater amount of detail, accuracy, and a greater DR. Some examples for the applications of HDRI sensors are in cars, medical imagery, photography, etc. HDR of illumination can create distortions and loss in information when viewing or applying further image processing techniques [12]. There are various HDR methods that exist in literature which address the simple problem of illumination. Each of the methods attempts to condense the HDR of luminance values into a viewable range, and to keep as much data as it can. HDR aims to keep and enhance colour information, since colour information can be valuable. Some examples can be found in [20, 21, 22, 23]. HDRI covers a range of methods which provides a better DR of exposures compared to standard image processing techniques [13].

Annamária et al. [12] addressed the problem of colour image reproduction after HDR of the illumination creates a distorted visual appearance and image contrast in distinct regions. The result of distortions causes a loss in detail as well as a loss in colour information. A new tone reproduction was introduced that helps in the development of difficult-to-see features and improves the visibility of the content in colour images. Pixels of the RGB (Red, Green and Blue) colour components are addressed individually. At the end of this process, the modified RGB

components are blended together, producing an image of high-quality colour HDR, that contains all the important information including detail and colour information.

HDR techniques tend to produce images that have an unnatural appearance [15, 24]. The proposed MEF produces a resultant image that appears natural while enhancing detail. The MEF technique makes use of HDR methodology proposed by Annamária et al. [12] and the approach by Saleem et al. [10] to propose a method that tries to fulfil any limitation posed by the HDR and fusion method. The intention of MEF is to produce a result that has high quality while maintaining naturalness and enhancing detail. This makes it different from current HDR and fusion approaches. Normal HDR algorithms require two or more input images of different exposure levels; the MEF requires one input image of any exposure. The proposed MEF requires less input data than HDR and other fusion methods. It creates a global framework for parallel contrast enhancement, followed by parallel image assessment and region selection, leading to final merging of selected regions from the enhanced set. The resultant output combines desirable attributes from each enhancement pathway to produce a result that is superior to each path taken alone. This technique provides a reliable solution to recover radiance maps from photographs taken with conventional imaging equipment. The parallel enhancer employs three state-of-the-art enhancement methods. These methods are explained in Section B.2.2, B.2.3 and B.2.4 and were chosen because they performed best in a previous study conducted by Maharaj et al. [11].

B.2.2 Contrast Enhancement based on Intrinsic Decomposition (CEID).

Goal: Yue et al. [25] seeks to develop intrinsic image decomposition model that is appropriate for contrast image enhancement. This is done by introducing constraints on the reflectance layer and the illumination layer in order to achieve an efficient enhancement.

Previous work: Studies have shown that by altering the decomposed illumination layer the image quality is improved [25]. These layers were altered to enhance under-exposed or over-exposed images. Such models have been proposed and recorded in the literature [25, 26, 27, 28, 29]. Barrow et al. [30] proposed intrinsic image decomposition. The reflectance values, which are not changed by the illumination condition, relate to the intrinsic colour of the image [25]. Intrinsic decomposition was a highly ill-modelled problem. However, there are numerous different inferences and subsequent work that have been made to make this a well-modelled problem. Such examples are referenced in [31, 32, 33, 34].

Method: To produce an extremely efficient enhancement, CEID [25] proposes constraints on the reflectance layer and the illumination layer [25]. The reflectance layer is regularised to be piecewise constant. This is done by presenting a weighted $l1$ norm constraint on the neighbouring pixels. The weighting is in accordance with the colour similarity of an input image. This is done

since the illumination information will barely affect the reflectance. A piecewise smoothness constraint is used to regularise the illumination layer. The Split Bregman algorithm is used to resolve the proposed decomposition model. The illumination layer is altered to achieve the enhanced image. Illumination adjustment was brought in to lessen computing complexity and to avoid potential colour artefacts. The decomposition model was implemented along the value channel in HSV (hue, saturation and value channel) colour space.

Results: CEID performs well for a broad range of images. The method achieves a much improved and comparable quality in relation to other state-of-the-art enhancement methods. There are several limitations to this method. It must be noted that the decomposition model was created for contrast enhancement. Using this model, the result for other image processing applications such as object insertion and surface re-texturing, shall not produce desirable results. Since this enhancement model is created for images, it might cause flickering artefacts if it is applied directly to video enhancement [25].

B.2.3 Naturalness Preserved Enhancement Algorithm for non-uniform illumination images (NPEA).

Goal: The NPEA [29] model seeks to preserve the naturalness of an input image while at the same time enhancing its details. This enhancement algorithm is introduced for images that have non-uniform illumination.

Previous work: To preserve the naturalness of an image and enhance its detail, Chen et al. [35] proposed the idea of naturalness preservation for enhancing images; they considered that the image colour impression must not be altered drastically after the enhancement. The author further stated that no additional source of light must be added to the scene and no halo effect to be introduced. Also, no blocking effect must be augmented as a result of over-enhancement [35]. Many examples of natural enhancement models on the Retinex theory can be found in journals [35, 36, 37]. These algorithms propose detail enhancement while preserving the naturalness in an image. However, these algorithms have a limitation in that they are not desirable for images with non-uniform illumination. Wang et al. [29] proposed an algorithm to maintain the naturalness of an image while enhancing detail under non-uniform illumination.

Method: There are three main contributions made by Wang et al. These contributions are: preservation of naturalness, intensity decomposition and the illumination effect [29]. The first contribution introduces a lightness-order error measurement to accurately access the naturalness preservation. The second contribution is a bright-pass filter used to decompose an image into two layers, namely, reflectance layer and illumination layer. This determines the amount of detail and the naturalness of the image. It also ensures that the reflectance is constraint to the range [0,1]. In

the third contribution, a bi-log transformation is proposed. This transformation maps the illumination to establish an equilibrium among detail and naturalness.

There are two constraints that are proposed in the NPEA algorithm. The first constraint relates to detail, which requires the reflectance to be set to a range of [0, 1] by considering the property of reflectance [38]. Naturalness is the next constraint, where there should not be a drastic change in the relative order of illumination in various local areas.

Results: It is observed that the NPEA enhances detail and maintains naturalness for images with non-uniform illumination. It shows that the resultant image is more visually appealing, artefact-free and natural. However, a limitation of this method is that the enhancement model fails to consider the temporal relationship of illumination across various scenes. Therefore, in video applications where the scenes change, flickering can be introduced [29].

B.2.4 Automatic image equalisation and contrast Enhancement using Gaussian Mixture Modelling (AEGMM)

Goal: The AEGMM method proposed by Celik et al. [39], automatically segments the contrast domain and adaptively equalises each segment. This aims to make the contrast enhancement more responsive to localised feature in contrast distribution.

Previous work: In the histogram modification framework (HMF), the contrast enhancement is handled as an optimisation problem which minimises a cost function. To manage noise and black/white stretching, variables are introduced in the optimisation. The HMF can attain various stages of contrast enhancement by utilising diverse adaptive parameters. By manually changing the parameters in accordance with the image content, a better contrast enhancement can be attained. A parameter-free algorithm is favoured. To create a parameter free algorithm a genetic algorithm (GA) is utilised. The GA is used to obtain a target histogram which will maximise the contrast measurement based on edge information [40]. This approach is called contrast enhancement based on GA. This approach has a limitation, namely its dependence on initialisation and convergence to a local optimum [39].

Method: Celik et al. proposed an adaptive image equalisation algorithm that efficiently improves the human visual quality of various cases of given images. The AEGMM algorithm fits a Gaussian mixture model (GMM) to the gray-level distribution intervals at the Gaussian intersection points. To acquire an image where the contrast is equalised; each input interval is equalised in accordance with the dominant Gaussian component as well as the cumulative distribution function of the input interval. The Gaussian components that have low variances are assigned with lower values; likewise, larger values are assigned to Gaussian components that have larger variances [39]. In

addition, GMM is employed to assign the components to map the input intervals to the output intervals. The AEGMM algorithm is designed to be free of parameter setting for a given DR of an image.

Results: It is noted that a low contrast image is automatically enhanced in relation to an increment in the DR. It is also observed that image with high contrast is improved, however, this improvement is little. With AEGMM, the colour quality of the wide range input image is enhanced. The quality is enhanced in terms of a few factors such as the colour consistency, higher contrast among foreground and background objects, a bigger DR and detail in the image [39]. The finding of the author [39], indicates that the AEGMM produces results that are more visually pleasing compared to the original image.

B.3 Multimodal Enhancement-Fusion technique for natural images

In this section, an overview of the proposed MEF is presented followed by a detailed mathematical model of the proposed MEF technique.

B.3.1 Overview

The MEF technique creates a concurrent enhancement pathway that subjects an image to multiple image enhancers in parallel, followed by a fusion algorithm. The MEF enhancer adaptively selects the optimal (best) enhancer for each image region. Three different contrast enhancement algorithms are used in the development of the parallel enhancements. These algorithms were chosen from Maharaj et al. [11].

The enhancement algorithms use the RGB and HSV colour spaces. The RGB colours are generally perceived as brighter and much more intense because of the light projected directly into the eye of the human observer [41]. Sometimes it is advantageous to utilise other colour spaces such as HSV, CIELAB, YIQ etc [42]. Therefore, using two different colour spaces may achieve a natural and detailed enhancement. The parallel enhancement uses the following enhancement algorithms:

1 Contrast enhancement based on intrinsic image decomposition (CEID) [25] in HSV plane.
2 Naturalness preserved enhancement algorithm for non-uniform illumination images (NPEA) [29] in RGB plane
3 Automatic image equalisation and contrast enhancement using Gaussian mixture modelling (AEGMM) [39] in RGB plane.

The parallel enhancer can consist of several enhancers. For this experiment, the number of enhancers empirically chosen, has three. These enhancements all have different useful properties.

Therefore, they are fused together using a fusion technique that is based on HDR. The framework of the MEF is depicted in Figure B.1. The modelling for the MEF is adopted from Annamária et al. [12].

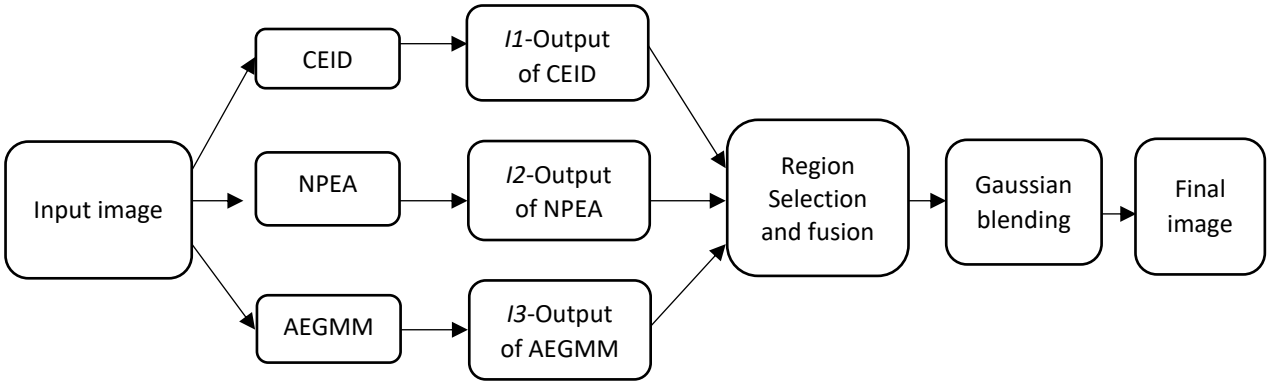


Figure B.1: Framework for the proposed algorithm – The method consists of 3 enhancements in parallel followed by selection and blending.

Image $I1, I2$ and $I3$ are divided into a rectangular grid of regions. The region size was selected empirically. The smaller the region size the greater the computational time, but a better image is obtained. The size was chosen such that there is a balance of computational time and quality. A metric is used to assess detail in each region of each enhanced image; the most detailed region (from $I1, I2$ or $I3$) is mapped into the fused image. Finally, the fused image is blended to remove artefacts at region boundaries.

B.3.2 Mathematically Modelling of the MEF technique

The model requires a single RGB image. The input images are denoted as I_{in}. I_{in} then goes through to each path of the enhancement pathway to produce enhanced image $I1, I2$ and $I3$:

1. Path 1: CEID algorithm is applied to the I_{in} to produce the enhanced image $I1$,
2. Path 2: NPEA algorithm is applied to the I_{in} to produce the enhanced image $I2$,
3. Path 3: AEGMM algorithm is applied to the I_{in} to produce the enhanced image $I3$.

The resultant enhanced image from each of the pathways ($I1, I2$ and $I3$) forms the input images that are required for the fusion. The number of input images chosen for the fusion algorithm is three ($N = 3$). The index k will be the input image such that $1 \leq k \leq N$. Considering a pixel from any enhancement ($I1, I2$ and $I3$) at (x, y). The intensity of the RGB at pixel (x, y) is depicted by $I_R(x, y)$, $I_G(x, y)$ and $I_B(x, y)$, where I_R represents the intensity of the red component, the intensity of the green component is represented by I_G and I_B represents the intensity of the blue component. Considering a 3x3 pixel neighbourhood, centred on (x, y), the neighbourhood pixels are as follows:

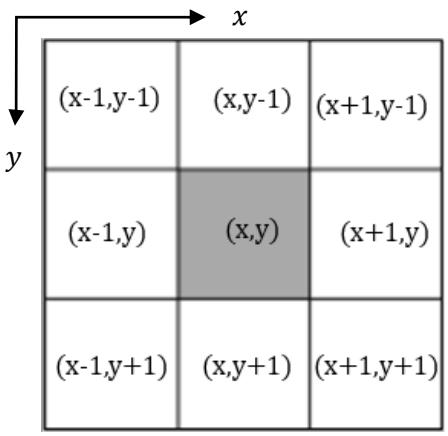


Figure B.2: The 3x3 neighbourhood of (x, y).

The neighbouring pixels are used to compute the gradient of each intensity function in the x direction (horizontal) ΔI_x and in the y direction (vertical) ΔI_y, located at (x, y). Pixel intensity gradients are used to describe the variation in intensity from one pixel to its neighbour.

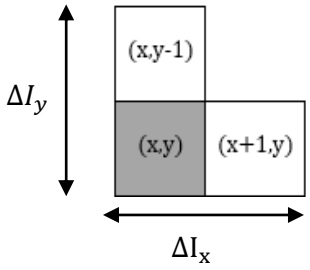


Figure B.3: Representation of the pixel intensity.

Pixel gradients are taken separately from each colour plane of the three enhancements. Consider Red, Green and Blue components from enhanced image $I1$. The gradient for pixel intensity for each component is as follows:

$$\begin{aligned}
\Delta I_X^R &= |I_R(x+1, y) - I_R(x, y)| \\
\Delta I_y^R &= |I_R(x, y-1) - I_R(x, y)| \\
\Delta I_X^G &= |I_G(x+1, y) - I_G(x, y)| \\
\Delta I_y^G &= |I_G(x, y-1) - I_G(x, y)| \\
\Delta I_X^B &= |I_B(x+1, y) - I_B(x, y)| \\
\Delta I_y^B &= |I_B(x, y-1) - I_B(x, y)|
\end{aligned} \tag{1}$$

Where x indicates a horizontal gradient and y represents a vertical gradient. Consider a rectangular region $\boldsymbol{R}$. The width of the region is defined as w and a region height h.

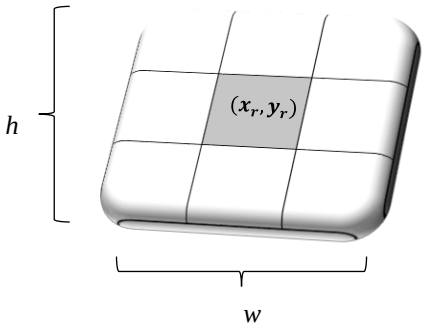


*Figure B.4: Rectangular image region (**R**) with width w and region height h for detail computation.*

A regular grid of regions is applied to image set I_k. Then, a measure of contrast detail is computed for each region. Consider a single region from a single enhanced image. Each colour plane in that region is processed separately. Consider the Red component. The level of Red component detail in the selected region $\boldsymbol{R}$ is given as:

$$S_D^R(\boldsymbol{R}) = \sum_{i=1}^{w}\sum_{j=1}^{h} P\left(max\left(\Delta I_x^R, \Delta I_y^R\right)\right) \tag{2}$$

Where, P is defined as $P(v) = \frac{v}{Imax}$ and is a normalised linear mapping, the maximum intensity, Imax is equal to 255 for 8-bit RGB. The same is done for the Blue (B) and Green (G) components.

$$S_D^B(\boldsymbol{R}) = \sum_{i=1}^{w}\sum_{j=1}^{h} P\left(max\left(\Delta I_x^B, I_y^B\right)\right) \tag{3}$$

$$S_D^G(\boldsymbol{R}) = \sum_{i=1}^{w}\sum_{j=1}^{h} P\left(max\left(\Delta I_x^G, \Delta I_y^G\right)\right) \tag{4}$$

All three components are summed, i.e. the sum of equation (2), (3) and (4), to produce the total level of detail S_D for the region $\boldsymbol{R}$ as follows:

$$S_D(\boldsymbol{R}) = S_D^R(\boldsymbol{R}) + S_D^G(\boldsymbol{R}) + S_D^B(\boldsymbol{R}) \tag{5}$$

In this way, the information density $S_D(\boldsymbol{R})$ is computed for every region $\boldsymbol{R}$, in every enhancement input image. For computational efficiency, the above formulation has been expanded and reformulated as follows.

$$S_D(\boldsymbol{R}) = \sum_{i=1}^{w}\sum_{j=1}^{h} P\{ max\left(\Delta I_x^R, \Delta I_y^R\right) + max\left(\Delta I_x^G, \Delta I_y^G\right) + max\left(\Delta I_x^B, \Delta I_y^B\right)\} \tag{6}$$

The function P(…) is first evaluated for every pixel location in the image, and across all colour-planes as indicated in equation (6). The summation is performed on the region processed. This is done by summing the pixels over i and j which span each identically sized region or grid block. The higher the computed value for S_D, the greater the detail in the analysed regions.

The fusion Synthesis**:** Each enhanced image (*I1*, *I2*, and *I3*) from the enhancement pathway has certain regions that are more desirable than others. The proposed MEF adaptively selects the optimal enhancer for each region. It determines which of the three enhanced images contributes the most desirable information density $S_D(\boldsymbol{R})$ for each region. Each enhanced image is unique; therefore, each region will be different.

Enhancement 1 (CEID) focuses on preserving the naturalness and colourfulness of an image and improving the detail. Enhancement 2 (NPEA) preserves and improves the naturalness of a given image while Enhancement 3 (AEGMM) improves contrast using the GMM to produce a visually pleasing image. The aim is to obtain an enhancement that has a balance of naturalness and detail, therefore, the three enhancements are fused together. Thus, improving feature detection, object recognition, pattern recognition, as well as scene reconstruction. This method may be utilised when lighting conditions are unfavourable.

The $I_k^R(x,y)$, $I_k^G(x,y)$ and $I_k^B(x,y)$ are used to signify the RGB intensity function, where (x,y) are pixel coordinates of the enhanced image with index k. Each enhanced image has regions that have more detail than the equivalent region in the other enhanced images. The aim is to create a resultant image where all three enhancements are combined and contains all the detailed regions in each image without producing any type of noise, irregularities, or distortions, especially at region boundaries.

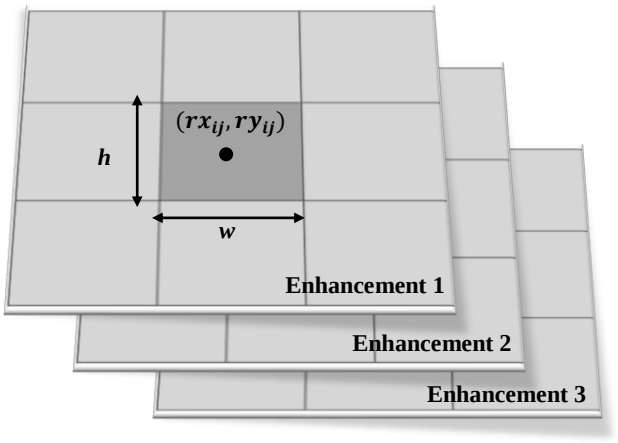


Figure B.5: The image region for the three enhancements.

The output image is to be constituted as the ordered collection of regions derived from the parallel enhancer such that for every region position, only the region with the highest detail is passed through to the output image. During this image reconstitution process, the regions boundaries are blended very carefully. Each enhanced image (*I1*, *I2* and *I3*) is divided to n rows and m columns, that produce a $n \times m$ rectangular image region. The regions are of identical size with height h and width w. This is measured in pixels. This rectangular region of pixels in any image is denoted as $\boldsymbol{R}$. All input images are divided into matching grids of regions, with each element being a region $\boldsymbol{R}$. $\boldsymbol{R}_{ijk}$ is denotes as the region of i^{th} row and j^{th} column in the enhanced image that has index k. The point (rx_{ij}, ry_{ij}) are denoted as the horizontal and vertical points in the region's centre in the i^{th} row and the j^{th} column. Now consider an arbitrary region $\boldsymbol{R}$ (refer to Figure B.5) and assume that $I1$ produces a higher $S_D(\boldsymbol{R})$ than $I2$ and $I3$. Then $\boldsymbol{R}$ will be selected from image $I1$ and inserted into the output image.

Let D denote an $n \times m$ matrix that has the most amount of detail. Each element d_{ij} corresponds to a grid region that stores the image index k of the image that contains the most detailed region at that grid position from the sourced enhancement. Matrix D, therefore, serves as a key or legend for the final fused image, indicating the enhanced image from which every output image region was sourced.

The regions $(\boldsymbol{R}_{ijL})$ contain the most amount of detail and the regions are merged together, in which $L = d_{ij}$, and index $i = i, \ldots, n,\ j = 1, \ldots, m$. The result of fusing the selected regions are three images i.e. the red, the green as well as blue images that form an RGB image. These images are the fused selected regions from the parallel enhancements. The region boundaries are often prominent and must be blended. A 2D Gaussian blending function is employed. Note that the function has a Gaussian numerator and denominator.

$$G_{ij}(x,y) = \frac{e^{-\left(\frac{(x-rx_{ij})^2}{2\sigma_x^2}+\frac{(y-ry_{ij})^2}{2\sigma_y^2}\right)}}{\sum_{p=1}^{m}\sum_{q=1}^{n} e^{-\left(\frac{(x-rx_{pq})^2}{2\sigma_x^2}+\frac{(y-ry_{pq})^2}{2\sigma_y^2}\right)}} \tag{7}$$

Where σ_x, σ_y represents the standard deviations of the 2-D function, and rx_{pq}, ry_{pq} represents the centre co-ordinates of the region at (p, q) on the region grid. The Gaussian function is centred on the region $\boldsymbol{R}_{ij}$ located at (i, j) on the region grid. The pixel (x, y) can be anywhere in the image. From this, the influence of the Gaussian at that pixel location is computed. Furthermore, it is defined that (x, y) is in region $\boldsymbol{R}_{rs}$ i.e. $(x, y) \in \boldsymbol{R}_{rs},\ 1 \leq r \leq n$ and $1 \leq s \leq m$. The Gaussian's influence will span the entire image regardless of the location of the centre. Given the drop-off in the function's output with distance from the centre, a distance threshold may be used,

beyond which the functions influence is set to zero. The distance threshold ε is incorporated into a cutting function $C(x, y)$ that is utilised to limit the range of the Gaussian's influence.

$$\mathrm{C}(x,y) = \begin{cases} 1\,, & where\ (x,y) \in \boldsymbol{R}_{rs} \mid |rx_{rs} - rx_{ij}| \wedge ||ry_{rs} - ry_{ij}| \leq \varepsilon \\ 0 & \text{everywhere else} \end{cases} \tag{8}$$

The cutting function is critical. If not used, blending will not happen. $C(x, y)$ and $G_{ij}(x, y)$ are combined such that the evaluation of weighted pixel intensities across the entire output image is as follows:

$$Iout(x,y) = \langle Iout^{R}(x,y), Iout^{G}(x,y), Iout^{B}(x,y) \rangle \tag{9}$$

$$Iout^{R}(x,y) = \sum_{i=1}^{n}\sum_{j=1}^{m} G_{ij}(x,y)C(x,y)I^{R}_{dij}(x,y) \tag{10}$$

$$Iout^{G}(x,y) = \sum_{i=1}^{n}\sum_{j=1}^{m} G_{ij}(x,y)C(x,y)I^{G}_{dij}(x,y) \tag{11}$$

$$Iout^{B}(x,y) = \sum_{i=1}^{n}\sum_{j=1}^{m} G_{ij}(x,y)C(x,y)I^{B}_{dij}(x,y) \tag{12}$$

The output intensity is a summation of Gaussian blending functions across the grid. Suppose the cutting function is not used, the numerator of $G_{ij}(x, y)$ will be identical to the denominator and the output will be identical to the input. Furthermore, using the cutting function, it will restrict the range of Gaussian summations in the numerator, but the denominator is unaffected. Hence, the output becomes a scaled version of the input.

B.4 Experimental Method

The aim of the experiment is to assess the performance of the proposed MEF. The experiment was conducted in three parts. Firstly, comparisons with an HDR and existing fusion method were done. Secondly, a human assessment survey was conducted which compared four state-of-the-art enhancement algorithms [25, 29, 39, 43] and the MEF result. Lastly, an objective assessment comparison was conducted by using the MEF results, three-enhancement methods and a fusion technique (Saleem et al. [10]). This section discusses the overview, experimental settings, the procedure for the comparison of HDR and fusion software, human assessment and the objective assessments.

B.4.1 Overview

The first experiment subjectively compares a fusion enhancement method and HDR software with the MEF. The MEF adopts the concept of HDR and the fusion method [10]. Therefore, a comparison is done to determine if the MEF achieves a more pleasing result.

The second experiment determines if the proposed method has better image quality and is more visually appealing to a human panel. Since there is no single metric that corresponds well with human preference, a human survey is the best way to determine how a panel responds to the MEF. The assessment consists of four algorithms, three of the algorithms were used in the parallel enhancer for the MEF. The fourth algorithm demonstrates a luminosity preserving contrast enhancing adaptive histogram equalisation technique for colour images. This enhancement is called Adaptive Equalisation applied in LAB space (AELAB). The enhancement algorithm is known for producing images with a lot of detail. Figure B.6 illustrates the output of the four enhancement algorithms and the original image as well as the output image from the proposed MEF. In addition, the author subjectively analyses the MEF and the enhancement techniques used in the survey. This evaluation aims to provide more detail about the appearance of the MEF and enhancements algorithms used in the survey. It compares naturalness, detail and overall appearance for all the algorithms.

The third experiment is an objective test. This is done to compare performance of the MEF technique using objective metrics. Mathematical models are used to determine the quality of enhancement methods and the proposed MEF. An objective study can approximate the image quality perceived by a human observer, therefore, it will be beneficial to determine how the MEF performs.

B.4.2 Global experimental settings

The parameters for the height and width for each region in the image are empirically set as $h = w = 5$, σx and σy are set to 60. ε is chosen, such that the values of the blending function exceed zero over the whole image domain. For enhancement CEID, NPEA and AEGMM; the same settings were used as specified in the journal articles by the authors.

The experiment is performed on a PC with 8G of RAM and 2.2 GHz CPU. All codes were implemented in MATLAB [44]. The test images used in the experiment are from the BSD300 dataset [45], Lossless dataset [46] and the USC-SIPI Image Database [47].

B.4.3 Experiment 1: Comparison with HDR and fusion software

In most image processing applications, the end user is an average human observer. Therefore, it is necessary to conduct subjective testing [19]. The experiment compares HDR software (EasyHDR [48]) with the MEF results and then compares Saleem et al. method [10] with the MEF results.

B.4.4 Experiment 2: Human assessment

The human assessment is done in two parts, namely, the human assessment survey and the analysis. The structure of the survey study was based on the existing survey from [39] and [25]. The survey was carried out face-to-face where a group of respondents was required to fill out a survey form. This method yields higher co-operation, lower refusal rate and higher response quality [49]. To reduce bias, only willing respondents were given the survey. To ensure that the survey was efficient and reliable, global standards were implemented in carrying out this survey as referenced in the report [19]. The survey was conducted among a random sample population. The standardisation of measurement was the same for every respondent i.e. the same set of questions was asked.

The visual assessment study was conducted with 30 respondents (16 males and 14 females). The respondents were given a portfolio containing eight image sets. Each image set contained six images. These images can be found in Figure B.6. A respondent is shown six images simultaneously; that is, the original test image is positioned on the top left of the page while the output images from the five enhancement algorithms are positioned randomly after the original test image. The five enhancements consist of the AELAB, NPEA, CEID, AEGMM and the proposed MEF. The survey questionnaire asked respondents to refer to the portfolio of images and score the images according to the quality of each enhanced image by allocating one of the five numeric scores from 1 to 5 (1, 2, 3, 4, 5). Score “1” represents an annoying enhancement and the image is much worse than the original image i.e. the image quality is distorted. If the image is not clearly enhanced i.e. the enhanced image is similar to the original image, then score “3” is given. Score “5” suggested that it is a substantial enhancement and the enhanced image has better quality than the original image. Other scores are selected in accordance with how respondents perceived image quality. IBM SPSS [50] software was used to capture the data obtained from this survey.

In addition to the scoring of the images, the author conducted a subjective analysis between the enhancement method used in the survey (AELAB, NPEA, CEID, AEGMM) and the proposed MEF. The information and observations used in this experiment was gathered from the human panel, the authors of the enhancement methods and the MEF.

B.4.5 Experiment 3: The Objective Image Quality Assessment

The aim of an objective IQA method is to create a mathematical model which best determines the quality of a given image as precisely as possible. The mathematical model must simulate the quality assessment of an average human observer. There is no objective measure that works like HVS, therefore, several popular metrics are adopted to assess the characteristics of the enhanced images. The following metrics are used in the experiment:

1. Naturalness image quality evaluator (NIQE) [51],
2. No-reference free energy based robust metric (NFERM) [52],
3. Entropy [53],
4. No-reference image quality metric for contrast distortion (NIQMC) [54],
5. The colourfulness-based PCQI (patch-based contrast quality index [55]) (CPCQI) [56],
6. The blind/reference-less image spatial quality evaluator domain (BRISQUE) [57].

For the objective assessment, the MEF technique is compared objectively with enhancements used in the parallel enhancer (CEID, NPEA and AEGMM). These images are shown in Figure B.6. Thereafter, the objective comparison is done with the MEF and Saleem's method. The images are shown in Figure B.8.

The NIQE and NFERM are utilised to assess the quality of an image blindly. NIQE determines the image quality by calculating the distance between the model statistics that was removed from natural images, and the distorted image [58]. The NFERM metric makes use of the free energy-based brain theory as well as the features that are inspired by HVS to calculate the distortion of images. For both NIQE and NFERM, smaller scores represent better quality. BRISQUE is NSS based [58]. This model functions in the spatial domain. BRISQUE is well suited for real time applications since it has very low computational complexity. BRISQUE can be used to identify distortion. The lower the BRISQUE score is, the better the quality of the image. The entropy evaluation measures the amount of information confined in an image. The larger the entropy score after enhancing an image, the greater the information confined in an image. The NIQMC is utilised to assess the contrast quality of an image and it provides precise quality predictions for contrast distorted images. Better quality is represented by a higher NIQMC score. CPCQI is computed with the original/input image as a reference image. CPCQI measures contrast distortions by considering the colourfulness aspect of the image. CPCQI is a measurement of perceptual distortions from the mean intensity, colour saturation in local patches, signal strength, and signal structure. The higher score of CPCQI signifies a better image contrast quality. The results of both subjective and objective experiments are explained in the Results and Analysis section.

B.5 Results and Analysis

The results and analysis from the experiment are presented. This can be divided into three sections. The first section presents the comparison between the state-of-the-art HDR software, Saleem et al. technique and the MEF. The second section presents the results of the human assessment survey and the analysis between the methods. Thirdly, the results and analysis of the objective assessment are presented. Figure B.6 presents the images used for the human assessment (B.5.2) and objective assessment (B.5.3).

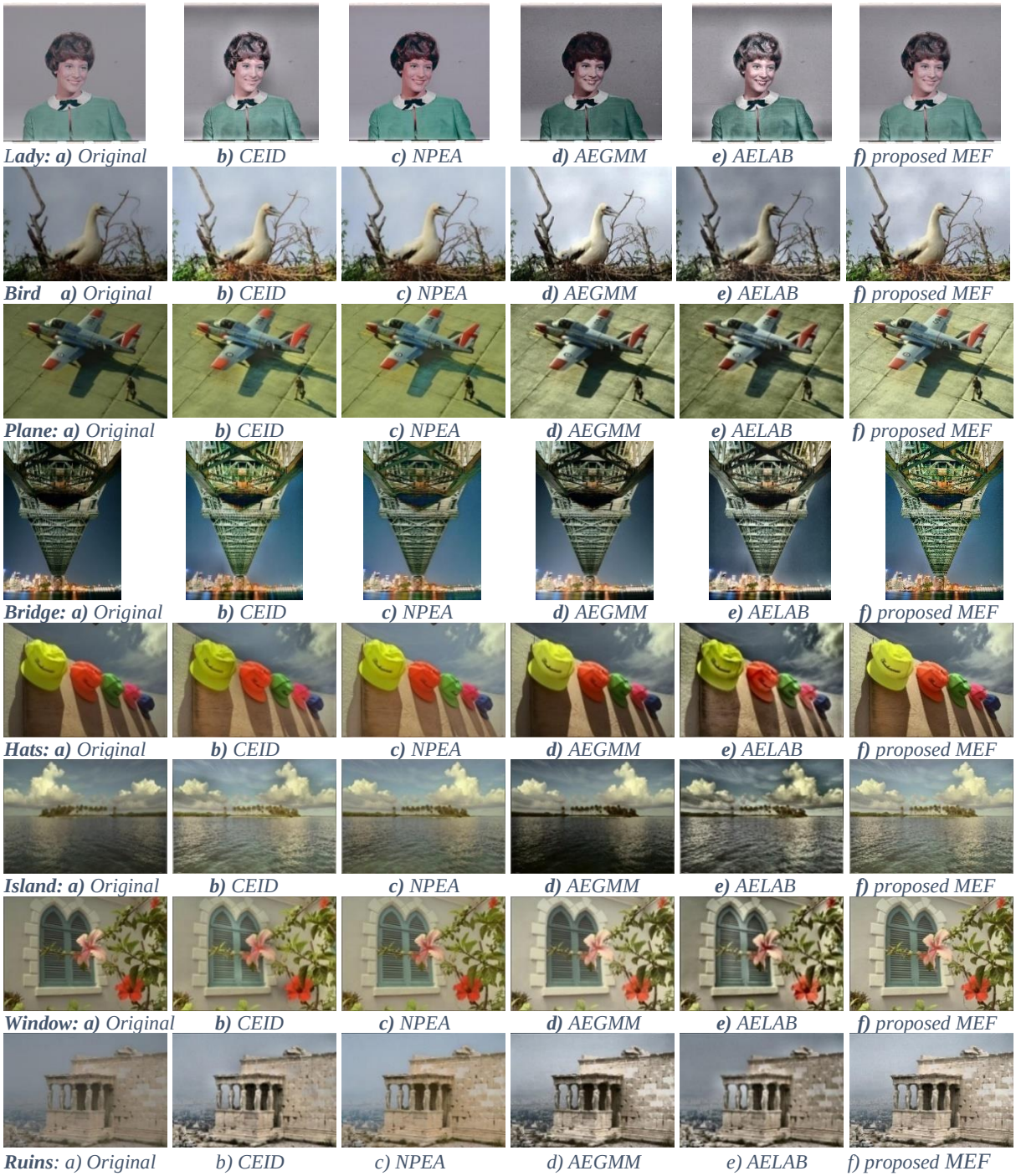


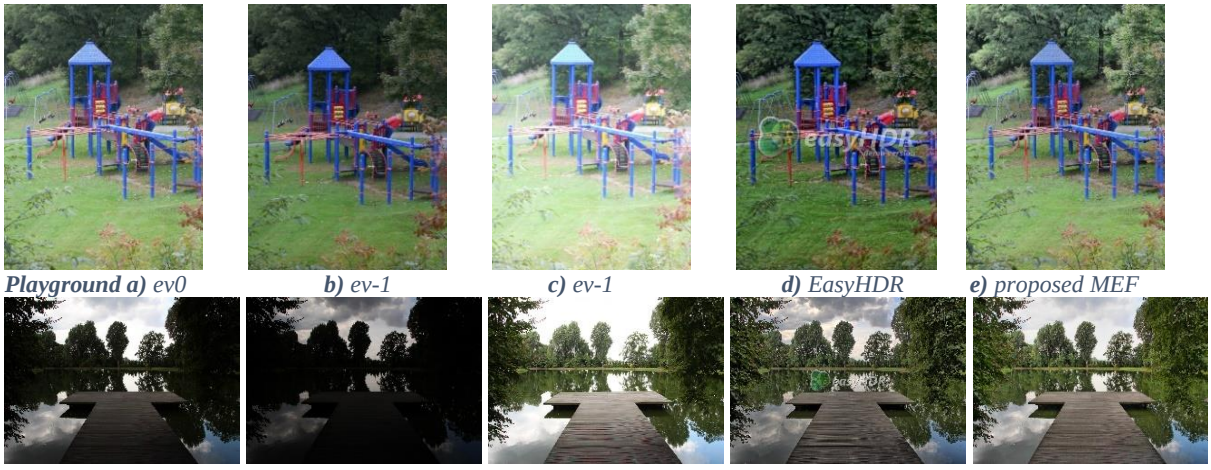
Figure B.6: Output from four enhancement methods and the MEF applied to the set of eight test images.

B.5.1 Experiment 1: Comparison with HDR and Fusion software

First the results from the comparison with the HDR software is presented and then a comparison is made with Saleem et al. method.

a. Results for the HDR software and the MEF technique

Figure B.7 shows the output results from EasyHDR software [48] and the MEF. HDR software requires three input images (EV0, EV-1/-2, EV+1/+2) whereas the MEF model only requires one input image. The EV0 image was used as the input for the MEF.



Playground a) *ev0* ***b)*** *ev-1* ***c)*** *ev-1* ***d)*** *EasyHDR* ***e)*** *proposed MEF*

lake a) *Ev0* ***b)*** *Ev-2* ***c)*** *Ev+2* ***d)*** *EasyHDR* ***e)*** *proposed MEF*

Figure B.7: Output images from the easyHDR software and the MEF are presented. For the MEF algorithm, EV0 was used as the input image.

Analysis of the HDR software: The MEF is designed for image fusion and to improve the images appearance in terms of naturalness, detail and colourfulness. Aspects of HDR methodology were incorporated into the design process of the proposed fusion technique. Therefore, a comparison between the proposed method with the HDR software is done. From Figure B.7, the easyHDR software produces an image that contains a lot of detail in comparison to the other images. This detailed image is composed from images ev-, ev+ and ev0. This software also produces an output that looks unnatural. The clouds in Figure B.7 lake d) have shadows and it looks unnatural. The proposed method contains more detail than ev0, ev- or ev+. The algorithm enhances the detail and produces a more natural and colourful image than the HDR software.

b. Results for Fusion technique and the MEF technique

Figure B.8 shows the output results of Saleem et al. method [10] and the MEF. When designing the current MEF approach, the approach by Saleem et al. was studied. Therefore, a comparison between the MEF and Saleem's method is done.

Bottle a) *original* ***b)*** *Saleem Method* ***c)*** *Proposed MEF*

House a) *original* ***b)*** *Saleem Method* ***c)*** *Proposed MEF*

Lakehouse a) *original* ***b)*** *Saleem Method* ***c)*** *Proposed MEF*

Figure B.8: Fusion resultant images from Saleem's Method [10] and the MEF technique.

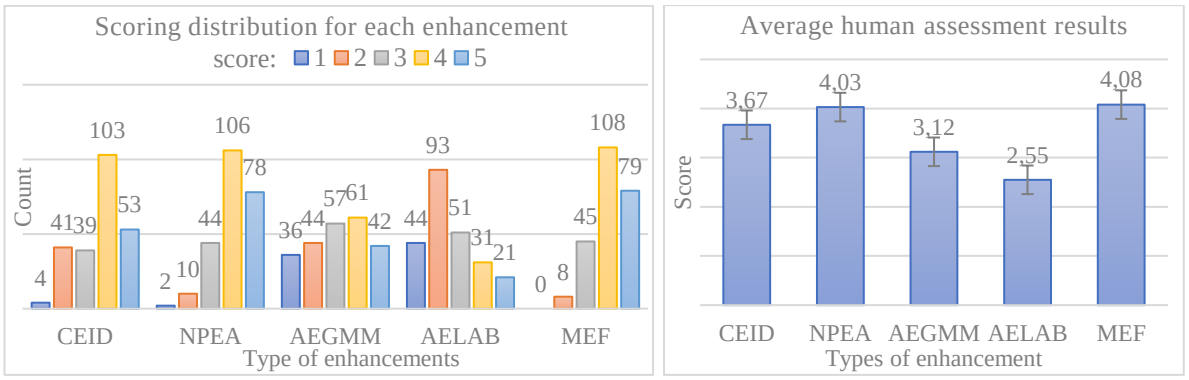
Analysis with Fusion software: Saleem et al. enhancement produces a visually pleasing result. It also produces a natural image. MEF produces an image that is brighter and more colourful. It produces much more detail than Saleem's method, for example Figure B.8 bottle c) the cracks on the wall are much more noticeable. Another example is Figure B.8 house, the art work on the walls are more defined in the MEF in comparison to Saleem's method. The proposed MEF contains more detail and the colours are more vivid and produces a more visually pleasing result. An objective comparison is done at the end of Section B.5.3 for the images presented in Figure B.8.

B.5.2 Experiment 2: Human Assessment

The human assessment results are presented in two parts, namely, the scores from the human assessment survey and the analysis of human perception.

a. Human assessment scores

The survey asked respondents to score the image according to visual preference and image quality. The score distributions from the survey are summarised in Figure B.9a). The maximum count for each enhancement is 240 (eight image set and 30 participants). The majority of the respondents assigned a score of 4 to CEID. While the majority score for AELAB was 2 and AEGMM scored between 3 and 4. The most favoured methods were the MEF and NPEA, with most scores in the range 4 to 5.



a) *b)*

Figure B.9: a) The graph represents the score distribution for each enhancement. Score range is from 1 (much worse) - 5 (much better). ***b)*** *The graph represents the average visual score received by each enhancement for the survey, the error bars represent the standard error of measurements for 30 respondents.*

Figure B.9b shows the average score received by each enhancement method. MEF achieved the best score. This shows that most respondents favoured MEF enhanced images. It was the most visually appealing image compared to the other enhanced images, while AELAB was least favoured by participants. NPEA received the next best score. The average AEGMM enhancement score was 3. Respondents found these images to be similar to the original image.

b. Human perception of enhancement methods

In order to assess the efficiency and the capability of the proposed MEF, a subjective comparison with other enhancement methods (viz. NPEA, CEID, AEGMM and AELAB) was done. The enhancement results are shown in Figure B.6.

In Figure B.6- 'Hats', CEID produces a natural image with some over-enhanced regions. This over-enhancement can be seen in the clouds. NPEA produces a natural image; all aspects of these images are clear but lack detail in some regions, for example in the wall lines where the hats are placed, there is a lack of information compared to the other enhancements. AEGMM produces an image that is more visually appealing than the original image. Dark regions are present in the enhanced image, for example the walls under the hats have shadows. The cloud region in the AEGMM image is more detailed than CEID and NPEA. AELAB contains the most information detail but tends to produce an over-enhanced image. The proposed MEF produces an image with all the best qualities from CEID, NPEA and AEGMM enhanced images. Naturalness is preserved, and the image looks pleasing.

Figure B.6- 'Bird', CEID has more colour components and is visually pleasing. This is because the author integrated colour information with the decomposed reflectance layer to create the colourful reflectance [25]. Both NPEA and CEID images look natural. AEGMM enhanced image is similar to the original image. In some regions of the AEGMM enhanced image, there are regions that are better and brighter than original regions while other regions are darker and less

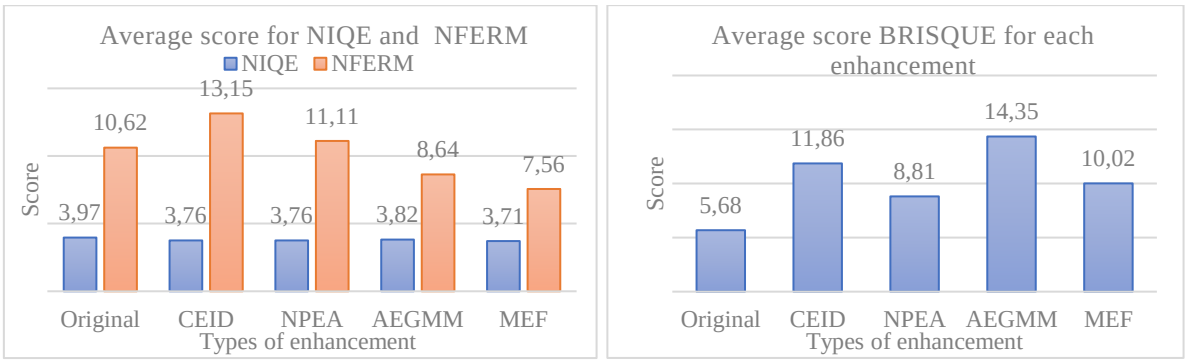
visible than the original image. The cloud region behind the bird is more enhanced for the AEGMM image. The MEF has the most detail and naturalness. It extracts the colourfulness from CEID, the detail from AEGMM and the naturalness from NPEA to produce a superior image.

Summary: CEID images can prevent artefacts and can well manage non-uniform illumination. It also produces a colourful image. It sometimes tends to over-enhance certain regions of an image. NPEA maintains the naturalness of an image that has non-uniform illumination but has detail loss in some areas. AEGMM produces images with a decent amount of enhancements, while some areas tend to remain dark. An example of this can be observed in the background of the image "lady". It also highlights certain detail that CEID and NPEA do not. AELAB produces over-enhanced images and looks unnatural. MEF fuses all the important aspects from each of the three enhancements (CEID, NPEA and AEGMM) to produce the best image. It fuses NPEA naturalness, CEID colourfulness and AEGMM details to produce a superior image. NPEA keeps the relative order of lightness well maintained. The proposed enhanced result is more natural compared to NPEA and the detail loss is addressed. The details of the proposed image are distinct and the whole image looks natural, colourful and sharp.

B.5.3 Experiment 3: Objective Image Quality Assessment

An objective comparison with the MEF and other enhancement methods was done. The objective methods mentioned in Section B4.5 were simulated and the results are presented in this section. The test includes results for the three parallel enhancement algorithms (CEID, NPEA and AEGMM) which were used in addition to the result of MEF and the original image. The image sets used are presented in Figure B.6 and a comparison between MEF and Saleem's method shown in Figure B.8.

The NIQE and NFERM are utilised to evaluate image quality blindly. For the two metrics, smaller scores mean better quality. Figure B.10a shows the graph of the average score for NIQE and NFERM for each image, i.e. the original input image, the three outputs of the parallel enhancements and the proposed MEF. From the graph, it can be noted that the MEF achieves the best (smallest) score for NIQE and NFERM. AEGMM also produce a good score for NIQE and NFERM.

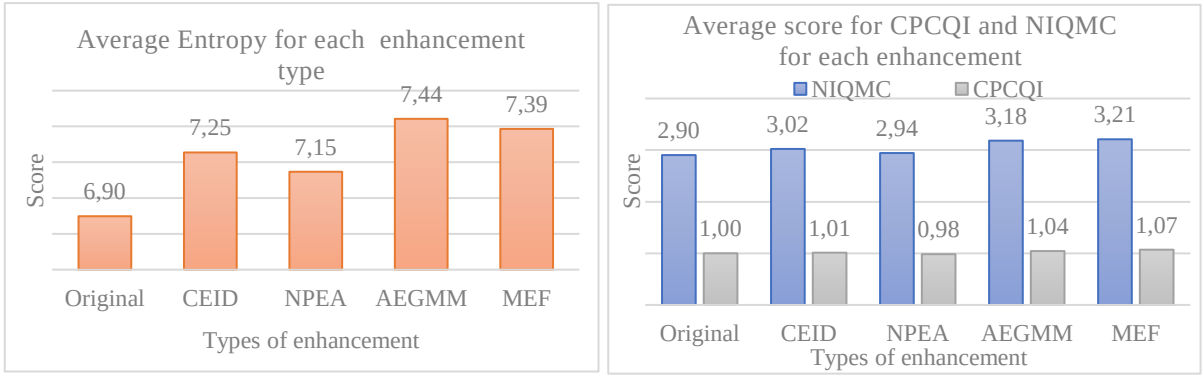


a) *b)*

Figure B.10: a) The graph represents the average score of NIQE and NFERM for each enhancement method and the original image. b) The graph represents the average score for BRISQUE for each enhancement method and the original image.

Figure B.10b shows the graph of average scores for BRISQUE. The lower score indicates better quality. NPEA scores the lowest from all the enhancements and the AEGMM scores the highest. The proposed method has a fairly low score in comparison to the other algorithms. From NIQE, NFERM and BRISQUE metric it can be concluded that the MEF offers the best quality.

Figure B.11a shows the average entropy score of eight image sets. The enhancement with the highest average value of entropy is AEGMM and MEF, which indicates that these enhancements have the most information. All enhancements have an entropy score greater than the original image, i.e. they contain more information than the original image.



a) *b)*

Figure B.11: a) Graph of the average score for entropy for each enhancement method and the original image. ***b)*** *The graph of the average score for NIQMC and CPCQI for each enhancement and the original image.*

The NIQMC metric is used to determine the contrast distortions and the CPCQI metric is used to assess the contrast quality for an image. NIQMC does not require a reference image, while CPCQI scores are computed using the original image as the reference image. For CPCQI metric, if the scores are higher than 1 it means that the result is much more enhanced in comparison to the original input image and conversely, if the scores are lower than 1 it suggests that the detail is barely enhanced, or artefacts are introduced [25]. For the NIQMC metric, larger scores mean less

contrast distortions and that the image has better contrast quality. Figure B.11b shows the graph for average scores for NIQMC and CPCQI. It can be observed that the MEF has the highest score for both CPCQI and NIQMC. It can be concluded that the MEF has the best contrast and detail. The NIQMC metric results for NPEA are smaller since some of the details are hard to notice. Table B.1 shows the summarised scores for the objective assessment for AELAB, NPEA, CEID, AEGMM and the proposed MEF. Figure B.6 images were used.

Table B.1: Tabulated data for the objective assessment for the different enhancement methods. Rows represent the enhancement algorithms and the columns represent the scores for the metric evaluation

	NIQE	NFERM	BRISQUE	Entropy	CPCQI	NIQMC
Original	3.97	10.62	5.68	6.90	1	2.90
CEID	3.76	13.15	11.86	7.25	1.01	3.02
NPEA	3.76	11.11	8.81	7.15	0.98	2.94
AEGMM	3.82	8.64	14.35	7.44	1.04	3.18
MEF	3.71	7.56	10.02	7.39	1.07	3.21

In addition to the performance test for the different contrast enhancement methods, an objective test is done between Saleem's method and the proposed method. The average score was calculated from the simulation of the three images shown in Figure B.8. The results are presented in Table B.2. From the Table it can be seen that the MEF received better objective scores in comparison to Saleem et al.

Table B.2: Tabulated data for the objective assessment for the different fusion methods (Saleem et al. [10]). Rows represent the enhancement algorithms and the columns represent the score for the metric evaluation.

	NIQE	NFERM	BRISQUE	Entropy	CPCQI	NIQMC
Original	3.94	4.29	3.28	3.36	1	3.30
Saleem	4.19	20.64	6.01	3.33	1.03	3.45
MEF	3.13	3.29	1.90	3.72	1.06	3.48

Analysis of objective results: Humans are sensitive to artefacts and naturalness. This is not assessed well in the above objective measures. A visual assessment is required to validate results. Objective IQA for image enhancement can be beneficial by mixing metrics to measure artefacts. Entropy and NIQMC can determine the amount of detail in an image. From these two metrics, it was determined that the MEF contains the most detail. The NIQE, BRISQUE and NFERM were used to determine naturalness of an image. By using the three-objective metric, the results showed that the MEF is the most natural image. CPCQI are utilised to evaluate the contrast quality of the images. The metric showed the proposed method has the best contrast. In addition, the human assessment concluded that the MEF was the most visually appealing image.

B.6 Conclusion

This paper proposes a new multimodal enhancement-fusion (MEF) technique for natural images. The study developed a global framework for parallel contrast enhancement, followed by parallel image assessment and region selection, leading to final merging of selected regions from the enhanced set. The method proposes a concurrent enhancement pathway that subjects an image to multiple image enhancers in parallel, followed by a fusion algorithm that creates a composite image that combines the strengths of each enhancement path. When merging the regions, the RGB colour components are modified separately using the Gaussian blending function. This improves the chromatic irregularities such as poor contrast distribution. The MEF allows for various number of parallel enhancers. This experiment uses three enhancement models for the parallel enhancer. The model fuses the colourfulness improvements from CEID, the naturalness from NPEA and the detail from CEID and AEGMM. The experimental results show that the proposed MEF performs well for most images and achieves better subjective and objective image quality in comparison to other enhancement methods and fusion techniques. Thus, the multimodal enhancer can contribute value to machine vision applications as well as personal image collections for the human user.

B.7 References

[1] C. Stamatopoulos, C. Fraser and S. Cronk, "Accuracy aspects of utilizing raw imagery in photogrammetic," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* vol. B5, pp. 1, 2012.

[2] B. Chitradevi and P.Srimathi, "An Overview on Image Processing Techniques," *International Journal of Innovative Research in Computer and Communication Engineering,* vol. 5, no. 11, 2014.

[3] W. Hao, M. Hea, H. Ge, C.j. Wang and Q.-W. Gao, "Retinex-Like Method for Image Enhancement in Poor Visibility Conditions," *Advanced in Control Engineering and Information Science,* vol. 15, pp. 2798 – 2803, 2011.

[4] D. K. Sahu and M.P.Parsai, "Different Image Fusion Techniques: A Critical Review," *International Journal of Modern Engineering Research (IJMER),* vol. 2, no. 5, pp. 4298-4301, 2012.

[5] R.Gonzalez and P. Wintz, Digital Image Processing, Norwood,MA: Addison-Wesley, 1987.

[6] D. H. Rao and P. P. Panduranga, "A survey on image enhancement techniques: classical spatial filter, neural network, cellular neural network, and fuzzy filter," *IEEE International Conference on Industrial Technology,* pp. 2823-2827, 2006.

[7] R. Suthakar, J. Esther, D.Annapoorani and F. S. Samuel, “Study of Image Fusion-Techniques, Method and Applications,” *International Journal of Computer Science and Mobile Computing,* vol. 3, no. 11, pp. 469-476, 2014.

[8] S.G. Huang, “Wavelet for Image Fusion,” National Taiwan University.

[9] B. Bavachan and D. Krishnan, “A Survey on Image Fusion Techniques,” *International Journal of research in computer and communication technology,* 2014.

[10] A. Saleem, A. Beghdadi and B. Boashash, “Image fusion-based contrast enhancement,” *EURASIP Journal on Image and Video Processing,* no. 10, 2012.

[11] R. Maharaj and B. Naidoo, “An Analysis of Objective and Human Assessments in Contrast Enhancement,” *International Journal of Applied Engineering Research,* paper code: 66780, 2018.

[12] A. R. Várkonyi-Kóczy, A. Rövid and T. Hashimoto, “Gradient-Based Synthesized Multiple Exposure Time Color HDR Image,” *IEEE Transactions on instrumentation and measurement,* vol. 57, no. 8, pp. 1779-1785, 2008.

[13] Y. Li, L. Sharan and E. H. Adelson, “Perceptually based range compression for high dynamic range images,” *Journal of Vision,* vol. 5, no. 8, pp. 598, 2005.

[14] M. B. A. Haghighata, A. Aghagolzadehab and H. Seyedarabia, “Multi-focus image fusion for visual sensor networks in DCT domain,” *Computers & Electrical Engineering,* vol. 37, no. 5, pp. 789-797, 2011.

[15] T. Mertens, J. Kautz and F. V. Reeth, “Exposure Fusion,” in *15th Pacific Conference on Computer Graphics and Applications,* 2007.

[16] P. J. Burt, K. Hanna and R. J. Kolczynski, “Enhanced image capture through fusion,” *Proceedings of the Workshop on Augmented Visual Display Research,* pp. 207–224, 1993.

[17] P. Burt and E. Adelson, “The Laplacian Pyramid as a Compact Image Code,” *IEEE Transactions on Communications,* vol. 31, no. 4, pp. 532 - 540, 1983 .

[18] J. M. Ogden, E. H. Adelson, J. R. Bergen and P. J. Burt, “Pyramid-based computer graphics,” *RCA Engineer,* vol. 30, no. 5, 1985.

[19] P. Mohammadi, A. Ebrahimi-Moghadam and S. Shirani, “Subjective and Objective Quality Assessment of Image: A Survey,” Elsevier, 2014.

[20] E. Reinhard, R. E. Stark, M. Shirley and P. J. Ferwerda, “Photographic tone reproduction for digital images,” in *Proc. 29th annual conference on Computer graphics and interactive techniques,* San Antonio, TX, 2002.

[21] P. E. Debevec and J.Malik, “Recovering high dynamic range radiance maps from photographs,” in *Proceedings of the 24th ACM Annual Conference on Computer Graphics and Interactive techniques (SIGGRAPH ’97)*, Los Angeles, Califonia, USA, 1997.

[22] K. Jacobs, C. Loscos and G. Ward, “Automatic high-dynamic range image generation for dynamic scenes,” *IEEE Computer Graphics and Applications,* vol. 28, no. 2, pp. 84-93, 2008.

[23] A. Tomaszewska and R.Mantiuk, “Image registration for multiexposure high dynamic range image acquisition,” in *Proceedings of the International Conference on Computer Graphics, Visualization and Computer Vision*, Plzen, Czech Republic, 2007.

[24] P. Cyriac, D. Kane and M. Bertalmío, “Perceptual Dynamic Range for In-Camera Image Processing,” Department of Information and Communication Technologies Universitat Pompeu Fabra, Barcelona, Spain, 2015.

[25] H. Yue, J. Yang, X. Sun, F. Wu and C. Hou, “Contrast Enhancement Based on Intrinsic Image Decomposition,” *IEEE transactions in image processing,* vol. 26, no. 8, pp. 3981 - 3994, 2017.

[26] R. Kimmel, M. Elad, D. Shaked, R. Keshet and I. Sobel, “A variational framework for Retinex,” *SPIE electronic imaging,* vol. 52, no. 1, pp. 7-23, 2003.

[27] X. Fu, Y. Sun, M. LiWang, Y. Huang, X.-P. Zhang and X. Ding, “A novel Retinex based approach for image enhancement with illumination adjustment,” *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* pp. 1190-1194, 2014.

[28] X. Fu, Y. Liao, D. Zeng, Y. Huang, X.-P. Zhang and X. Ding, “A probabilistic method for image enhancement with simultaneous illumination and reflectance estimation,” *IEEE Transactions on Image Processing,* vol. 24, no. 12 , pp. 673-686, 2016.

[29] S. Wang, J. Zheng, H.-M. Hu and B. Li, “Naturalness Preserved Enhancement Algorithm for non-uniform illumination images,” *IEEE Transactions on Image Processing,* vol. 9, no. 9, 2013.

[30] H. Barrow and J. Tenenbaum, “Recovering intrinsic scene characteristics from images,” *Computer Vision Systems, A Hanson & E. Riseman (Eds.),* pp. 3-26, 1978.

[31] E. Garces, A. Munoz, J. Lopez-Moreno and D. Gutiérrez, “Intrinsic images by clustering,” *Computer Graphics Forum,* vol. 31, no. 4, pp. 1415-1425, 2012.

[32] J. Shen, X. Yang, X. Li and Y. Jia, “Intrinsic image decomposition using optimization and user scribbles,” *IEEE Transactions on Cybernetics,* vol. 43, no. 2, pp. 425-436, 2013.

[33] S. Bell, K. Bala and N. Snavely, “Intrinsic images in the wild,” *ACM Transactions on Graphics,* vol. 33, no. 4, pp. 159, 2014.

[34] J. T. Barron and J. Malik, “Shape, illumination, and reflectance from shading,” *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 37, no. 8, pp. 1670 - 1687, 2013.

[35] S. Chen and A. Beghdadi, “Natural Rendering of Color Image based on Retinex,” *2009 16th IEEE International Conference on Image Processing (ICIP),* pp. 1813-1816, 2009.

[36] S. Chen and A. Beghdadi, “Natural enhancement of color image,” *EURASIP J. Image Video Process.,* vol. 2010, pp. 1-19, 2010.

[37] B. Li, S. Wang and Y. Geng, “Image enhancement based on Retinex and lightness decomposition,” *Proc. IEEE Int. Conf. Image Process,* pp. 3417–3420, 2011.

[38] E. Hecht, “Optics,” MA,USA, Addison Wesley, 2002, pp. 86-149 .

[39] T. Celik and T. Tjahjadi, “Automatic Image Equalization and Contrast Enhancement Using Gaussian Mixture Modeling,” *IEEE Transactions on Image Processing,* vol. 21, no. 1, 2012.

[40] S. Hashemi, S. Kiani, N. Noroozi and M. E. Moghaddam, “An image contrast enhancement method based on genetic algorithm,” vol. 31, no. 13, pp. 1816–1824, 2010.

[41] Chezian, N. PM and D. R.Manicka, “Various Colour Spaces And Colour Space Conversion Algorithms,” *Journal of Global Research in Computer Science,* vol. 4, no. 1, 2013.

[42] A. Ford and A. Roberts, “Colour Space Conversions,” 1998.

[43] S. Bharal, “L, a*, b* based contrast limited adaptive histogram equalization,” *Electrical & Computer Engineering: An International Journal (ECIJ),* vol. 4, no. 3, 2015.

[44] MathWorks, “Matlab installation,” Mathworks, [Online]. Available: https://www.mathworks.com/products/matlab.html. [Accessed 20 December 2017].

[45] Berkeley, “The Berkeley Segmentation Dataset and Benchmark,” [Online]. Available: https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/BSDS300/html/dataset/images.html. [Accessed May 2018].

[46] Kodak, “Kodak Lossless True Color Image Suite,” 2013. [Online]. Available: htttp://r0k.us/graphics/kodak/. [Accessed May 2018].

[47] University of Southern California, “The USC-SIPI Image Database,” [Online]. Available: http://sipi.usc.edu/database/. [Accessed May 2018].

[48] Easy HDR, “High Dynamic Range photography made easy,” easyHDR, 2006. [Online]. Available: https://www.easyhdr.com/. [Accessed 20 August 2018].

[49] L. K. Owens, “Introduction to survey research design,” Survey Research Lab Webinar Series, 2002.

[50] IBM SPSS, “IBM SPSS software,” IBM, 2017. [Online]. Available: https://www.ibm.com/analytics/spss-statistics-software. [Accessed 30 June 2018].

[51] A. Mittal, R. Soundararajan and A. C. Bovik, “Making a Completely Blind Image Quality Analyzer,” *IEEE Signal Processing Letters,* vol. 20, no. 3, pp. 209-212, 2012.

[52] K. Gu, G. Zhai, X. Yang and W. Zhang, “Using Free Energy Principle For Blind Image Quality Assessment,” *IEEE Transactions on Multimedia,* vol. 17, no. 1, 2015.

[53] D. S. Prabha and J. S. Kumar, “Performance Evaluation of Image Segmentation using Objective Methods,” Indian Journal of Science and Technology, vol. 9, no. 8, 2016.

[54] K. Gu, W. Lin, G. Zhai, X. Yang, W. Zhang and C. W. Chen, “No-Reference Quality Metric of Contrast-Distorted Images Based on Information Maximization,” vol. 47, no. 12, 2017.

[55] S. Wang, K. Ma, H. Yeganeh, Z. Wang and W. Lin, “A patchstructure representation method for quality assessment of contrast changed images,” *IEEE Signal Processing. Letter,* vol. 22, no. 12, pp. 2387-2390, 2015.

[56] K. Gu, D. Tao, J. F. Qiao and W. Lin, “Learning a No-Reference Quality Assessment Model of Enhanced Images With Big Data,” *IEEE transactions on neural networks and learning systems,* vol. 29, no. 4, pp. 1301-1313, 2018.

[57] A. Mittal, A. K. Moorthy and A. C. Bovik, “No-Reference Image Quality Assessment in the Spatial Domain,” *IEEE Transactions on Image Processing,* vol. 21, no. 1, pp. 4695 - 4708, 2012.

[58] K. Deb and M. V, “No-reference image quality measure for images with multiple distortions using multi-method fusion,” Image Analysis and Stereology, vol. 37, pp. 105-117, 2018.

Chapter 3: Conclusion

Conclusion

Existing objective metrics reviewed in this study are reasonable at assessing specific features of an enhanced image. When selecting any such metric, the purpose and objective of the task for which it is employed must be carefully considered. Paper 1 of this dissertation addressed a selection of objective assessment metrics. The relationship with the subjective human assessment of contrast enhanced images was interrogated. In this study, the human assessment was regarded as the assessment benchmark, or ideal assessment. The experiment then identified which of the objective metrics best approximated human assessment and could therefore be used as an effective replacement for typically tedious human assessment surveys. The results from the visual assessment showed that there is a relationship between detail and naturalness. Beyond a certain level, the more detailed an image is, the less naturalness is perceived. Images that look natural have moderate detail in general. It was shown that objective metrics are useful in evaluating the special characteristics of enhanced images, but they are individually unable to determine human preference. Despite the finding that no single quantitative metric under investigation correlates well with human perception; two or more metrics in combination were able to approximate the complex human response. Three metrics (NIQE, NFERM, BRISQUE) were found to be good estimators of human perception of naturalness; and two metrics (NIQMC and entropy) provided good estimation of human perception of detail.

This information was used in the multimodal enhancement-fusion (MEF) of natural images. The MEF proposed a framework that adaptively fuses multiple enhancement objectives into a seamless pipeline. Given a segmented input image and a set of enhancement methods, the MEF applied all the enhancers to the image in parallel. The most appropriate enhancement in each image segment was identified, and finally, the differentially enhanced segments were seamlessly fused. Paper 2 developed this global framework for parallel contrast enhancement, followed by parallel image assessment and region selection, leading to final merging of selected regions from the enhanced set. The MEF is tailored such that it selects and merges desired attributes from each pathway into the resultant image. The method proposes a concurrent enhancement pathway that subjects an image to multiple image enhancers in parallel, followed by a fusion algorithm that creates a composite image that combines the strengths of each enhancement path. The enhancement path is made up of the NPEA, CEID and AEGMM algorithms. The model fuses the colourfulness improvement from CEID, the naturalness from NPEA and the detail from CEID and AEGMM. The MEF also improves the chromatic irregularities such as poor contrast distribution. This is achieved during the merger of image regions. Once the optimal/best regions are selected from the enhancement set, the RGB colour components are modified separately using the Gaussian blending function. Thus, the MEF combines desirable attributes from each enhancement pathway to produce a result that is superior to each path taken alone. Experimental

results show that the proposed MEF performs well for most images and achieves better subjectively and objectively assessed image quality in comparison to other enhancement and fusion methods, namely, NPEA, AEGMM, CEID, AELAB and Saleem's method.

In conclusion, the work presented in this dissertation provides detailed insight into human perception, image enhancement and assessment techniques as well as a new approach for contrast image enhancement. This approach contributes value to machine vision applications as well as personal image collections for the human user.

Appendices

Appendix A: Digital copy of the results

All the results presented in this dissertation have been captured and saved in a digital format for viewing later on. The following items can be found in the digital copy (i.e. DVD):

1. The digital copy of the *Turn-it-in* report.
2. The digital copy of the final dissertation.
3. The MEF results presented in this dissertation. This includes images used in the report as well as data from survey and the objective tests which were captured on SPSS software.
4. The source code for the MEF algorithm.
5. The source code for the individual enhancements and the metrics used.
6. The images used in the surveys as well as the copy of Survey 1 and 2 are provided.

The DVD is attached

Appendix B: Summary of MEF Algorithm

Refer to the DVD for the code and implementation.

Algorithm 1: The proposed MEF algorithm

Input = RGB image

Output= Enhanced RGB image

Stage 1: Initialisation of enhancement pathways.

An image is loaded into the framework.

- The CEID algorithm is applied to the original input image to produce *I1*;
- The NPEA algorithm is applied to the original input image to produce *I2* and
- AEGMM algorithm is applied to the original input image to produce *I3*. These enhancements form the parallel enhancer.

Stage 2: Initialisation of the MEF.

- The region width and height (*rw* and *rh*) was empirically selected to be 5.
- The sigma value for the 2-D Gaussian and the maximum intensity value was defined.
- Then the dimension of the image pixel from the first colour plane of image *I1* was determined.
- Next, the number of regions that will fit the image horizontally and vertically must be determined as well as the footprint size.
- Stack the enhanced image (*I1*, *I2* and *I3)* planes all into a single data collection metrics matrix such that, the first three planes come from image 1, the next three planes from image 2, the next three planes from image 3. This will form the parallel enhancement pathway.

Stage 3: Compute enhanced image densities information.

The image is split into its colour components. Thereafter, compute the intensity of each colour component to produce $I_R(x,y)$, $I_G(x,y)$ and $I_B(x,y)$. A function is created to compute the intensities.

Stage 4: Compute region and pixel selection matrix.

The first step is to create a region selection matrix that indicates in each grid location, which image region contributes the most information. This is achieved by a vectorised logical comparison of the region densities matrix produced in the previous stages. The second step is to generate a graphic representation of region selection matrix. Lastly, create a Pixel Selection matrix. Pixel Selection matrix is the region selection data expanded to full pixel resolution.

Stage 5: Produce the raw (stitched) MEF image.

Producing the raw fused image (without smoothing), is done by simply using the pixel selection matrix to inform the assignment of pixels from one of the enhanced input images to the resultant image. Furthermore, the assignment is conducted on each colour plane individually. These colour

planes are stored to different planes of the data collection matrix. The new three colour planes form the final raw (stitched) image.

Stage 6: Produce the blended MEF image.

The Gaussian blending function is implemented.

$$G_{ij}(x,y) = \frac{e^{-\left(\frac{(x-rx_{ij})^2}{2\sigma_x^2}+\frac{(y-ry_{ij})^2}{2\sigma_y^2}\right)}}{\sum_{p=1}^{m}\sum_{q=1}^{n} e^{-\left(\frac{(x-rx_{pq})^2}{2\sigma_x^2}+\frac{(y-ry_{pq})^2}{2\sigma_y^2}\right)}}$$

The constant denominator matrix for the Gaussian Blending is created first. Every pixel location is scanned through and the pixel weighting factor is computed. This is done by creating a grid of region centre-pixel co-ordinates and then scanned across the full image region, a single pixel at a time. The evaluation $G(x.y).* C(x,y)$ for every (x,y) is done to obtain the weighting factor matrix *'W'*. $C(x,y)$ is the cutting function.

$$C(x,y) = \begin{cases} 1\,, & where\ (x,y) \in R_{rs} \mid |rx_{rs} - rx_{ij}| \wedge ||ry_{rs} - ry_{ij}| \leq \varepsilon \\ 0 & \text{everywhere else} \end{cases}$$

Now the cutting function $C(x,y)$ will be computed. The cutting function is critical. If not used, blending will not happen. $C(x,y)$ and $G_{ij}(x,y)$ are combined to evaluate the weighted pixel intensities across the entire image. The output intensity is a summation of Gaussian blending functions across the grid. Suppose the cutting function is not used, the numerator will be identical to the denominator and the output will be identical to the input. Furthermore, using the cutting function, it will restrict the range of Gaussian summations in the numerator, but the denominator is unaffected. Hence, the output becomes a scaled version of the input. After the blending the final image is achieved.

Appendix C: Algorithmic description of Contrast Enhancements

Appendix C provides an algorithmic description for the five enhancements as well the fusion technique used in the experiment. These enhancements are:

1. Colour image enhancement based on histogram equalisation [1].
2. Adaptive equalisation in LAB space [2].
3. Contrast-enhancement based on intrinsic decomposition [3].
4. Naturalness preserved enhancement algorithm for non-uniform illumination images [4].
5. Automatic image equalisation and contrast enhancement using Gaussian mixture modelling [5].
6. Image fusion-based contrast enhancement [6].

Appendix C1: Algorithmic description of Colour Image Enhancement Based on Histogram Equalisation

Algorithm 2: Colour image enhancement based on histogram equalisation

Input = RGB image

Output= Enhanced RGB image

Stage 1: Conversion

Convert the RGB image into HSV colour image. Decompose image into separate channels and compute the histogram of the V image.

Stage 2: Exposure threshold applied to V channel

Compute the exposure threshold by using the following formula:

$$exposure = \frac{1}{L}\frac{\sum_{k=1}^{L} h(k)k}{\sum_{k=1}^{L} h(k)} \quad \text{(C1.1)}$$

L denotes the number of gray levels and $h(k)$ is the histogram for the image obtained in stage 1.

Stage 3: clipping of the V channel histogram

The clipping threshold has to be computed.

$$T_c = \frac{1}{L}\sum_{k=1}^{L} h(k), \qquad h_c(k) = T_c \ \ for\ h(k) \geq T_c \quad \text{(C1.2)}$$

T_c is defined as the clipping threshold and $h_c(k)$ is the clipped histogram.

Stage 4: Dividing the Clipped Histogram

Utilising the exposure threshold, the parameter X_m has to be found:

$$X_m = L(1 - exposure) \quad \text{(C1.3)}$$

Based on X_m the histogram is divided into sub histograms. It is divided into under-exposed and over-exposed histograms.

Stage 5: Apply Histogram to sub images of the V image

The probability density function and the cumulative density function (CDF) are calculated for both sub-images. Then the histogram is applied independently to each sub-image.

Stage 6: Recombining of sub- images

After independent equalisation, the transfer function for the under-exposed image is given by:

$$F_u = X_m * C_u \quad \text{(C1.4)}$$

C_u is the CDF for the under-exposed images. The transfer function for the over-exposed image is given by

$$F_o = (X_m + 1) + (L - X_m + 1)C_o \quad \text{(C1.5)}$$

C_o , is the CDF of the over-exposed image. F_u and F_o are multiplied to achieve the combined histogram.

Stage 6: Final image

The V channel is recombined with the H and S channels to produce the HSV image. The image is then converted back into RGB to yield the final image.

Appendix C2: Algorithmic description of Adaptive Equalisation in LAB space

Algorithm 3: Adaptive equalisation in LAB space

Input = RGB image Output= Enhanced RGB image
Stage 1: Conversion Transform the RGB image into LAB colour image. Decompose image into separate channels.
Stage 2: Equalisation Normalise the luminosity (L) channel and then apply "contrast-limited adaptive histogram equalisation" to the normalised L channel.
Stage 3: Enhanced image Bring the L to the non-normalised form. Recombine the L channel with the untouched a* and b* channels to produce the resultant image.

Appendix C3: Algorithmic description of Contrast Enhancement based on Intrinsic Decomposition

The RGB image is converted into HSV image and decomposed into separate H, S and V images. The value (*V*) image can be decomposed into a reflection layer (*R*) and illumination layer (*L*).

$$V = L \cdot R \tag{C3.1}$$

Where '·' represents pointwise multiplication. The CEID model has two constraints. The constraints are:

- The neighbouring pixels with alike colours must have the equivalent reflectance
- and the neighbouring pixels should hold the similar or comparable illumination.

The author formulated the intrinsic decomposition model as a minimisation problem. The minimisation problem is of an energy function as depicted in equation C3.2. The following variables are utilised to the vector model of the V, L and *R* viz. v, l and r. Minimisation equations are defined as:

$$min_{l,r} E(l,r) = E_r(r) + \mu E_l(l) + \theta E_d(v; l, r) + \beta E_0(l, l_0) \tag{C3.2}$$
$$\text{such that } 0 \leq r \leq 1$$

The μ, θ, and β are the weighting parameters. First and second expression of the equation are the regularised reflectance layer and illumination layer. The third expression of the equation ensures the reliability of decomposition. A $\ell 2$-norm penalty ($\|v - l \cdot r\|_2^2$) is used to tolerate noise. The final expression of the equation is utilised to restrict the value of the illumination. The L_o is the chromatic normalisation value which is defined as: $\sqrt{{I_r}^2 + {I_g}^2 + {I_b}^2}$. The components I_r, I_g and I_b are the intensity of the red, green and blue colour components, respectively. The reflectance layer is constrained to E_r (r) term, which is piecewise constant. The value of reflectance at pixel i is given by r_i and $\mathcal{N}$ (i) is the neighbourhood of pixel i.

$$E_r(r) = \sum_i \sum_{j \in \mathcal{N}(i)} w_{ij} \|r_i - r_j\|_1 \tag{C3.3}$$

w_{ij} refers to the measurement of the similarity between chromatic value at pixel i and pixel j. The w_{ij} function is expressed in the following equation:

$$w_{ij} = exp\left(-\frac{\|f_i - f_j\|_2^2}{2\sigma^2}\right) \tag{C3.4}$$

The f_i and f_j terms are the value of the pixel i and j in the LAB colour plane and is expressed as $f_i = [\tau li, ai, bi]^T$, $f_j = [\tau lj, aj, bj]^T$. The variable τ is constrained to $\tau < 1$. This is done

to ensure the reduction of the impact on illumination variations on colour similarity quantity. The author assumed that the input image contained u pixels, and the pixels consist of n neighbouring pixel pairs. The authors create a matrix of dimension $n \times u$, such that $M = \{m_{ij}\}$. $m_{ki} = w_{ij}$ and $m_{ki} = -w_{ij}$, if pixels i and pixel j are neighbours that creates the k^{th} neighbouring pair. The term $E_r(r)$ can be written as:

$$E_r(r) = \|Mr\|_1 \tag{C3.5}$$

The $E_l(l)$ expression implies that the illumination must be locally smooth utilising isotropic total variation.

$$E_l(l) = \|D_x l\|_2^2 + \left\|D_y l\right\|_2^2 \tag{C3.6}$$

The term D_x and D_y are matrix that represent the derivative operator in the horizontal (x) and vertical (y) direction. Thus, the new energy function is defined as:

$$min\|Mr\|_1 + \mu(\|D_x\, l\|_2^2 + \left\|D_y l\right\|_2^2) + \theta\|v - l.r\|_2^2 + \beta\|l - l_o\|_2^2 \tag{C3.7}$$
$$\text{Such that } 0 \le r \le 1$$

Equation C3.7 can be solved using Split Bregman algorithm. The equation is transformed into an optimisation problem.

$$(l, r) = \; argmin_{l,r,h}\, \|h\|_1 + \mu(\|D_x\, l\|_2^2 + \left\|D_y l\right\|_2^2) + \theta\|v - l.r\|_2^2 + \beta\|l - l_o\|_2^2 \tag{C3.8}$$
$$\text{Such that } Mr = h, \quad 0 \le r \le 1$$

The equation is separated in r, l and h sub-problems, thus minimising one variable at a time and fixing the remaining variables.

Algorithm 4: The CEID algorithm

Input = RGB image

Output= Enhanced RGB image

Stage 1: Initialisation

An image is opened from a folder.

1.1. Initialise the illumination layer to $L^0 = \sqrt{{I_r}^2 + {I_g}^2 + {I_b}^2}$, h^0, b^0=*0* and Set μ = *5, σ = 1, β = 50, τ = 0.5, θ = 80, γ = 2.2, λ = 120.*

1.2 Determine the weighted matrix $w_{ij} = exp\left(-\frac{\|f_i - f_j\|_2^2}{2\sigma^2}\right)$

Stage 2: Converting to HSV

Convert RGB image into HSV Image and then it is decomposed into the individual channels.

Stage 3: While loop:

The V channel is further decomposed into reflectance and illumination using the while loop.

While $\left\|r^k - r^{k-1}\right\|_2^2 > \epsilon$, do:

3.1 Solve the r-problem to achieve r^{k+1} by using the following equation:

$$r^{k+1} = A^{-1}z \tag{C3.9}$$

$$A = \theta I + \lambda M^T M \tag{C3.10}$$

$$z = \theta\left(\frac{v}{l^k}\right) + \lambda M^T(h^k - b^k) \tag{C3.11}$$

Where, *I* is the identity matrix. A preconditional conjugate gradient (PCG) is used to solve r^{k+1}. Since matrix *A* is symmetric and positively defined. The variable *r* is constricted to $0 \leq r \leq 1$. The reflectance is projected as $r^{k+1} = min(max(r^{k+1}, 0), 1)$.

3.2 Solve the *l*-problem to achieve l^{k+1} by using the following equation:

$$\left(L^{k+1}\right) = argmin_L \quad \mu(\|d_x * L\|_2^2 + \left\|d_y * L\right\|_2^2) + \theta\|V - L.R\|_2^2 + \beta\|L - L^0\|_2^2 \tag{C3.12}$$

$$L^{k+1} = F^{-1}\left(\frac{F\left(\frac{\theta V}{R^{k+1}} + \beta Lo\right)}{F(\theta+\beta) + \mu\left(F*(d_x)F(d_x) + F*\left(d_y\right)F\left(d_y\right)\right)}\right) \tag{C3.13}$$

$$d_x = [-1, 1]$$

$$d_y = [-1, 1]^T$$

The '*' signifies complex conjugate and *F* is the fast Fourier transform algorithm. The multiplication and division operations are performed element-wise.

3.3 Solve the *h*-problem to achieve h^{k+1} by using the following equation:

$$\left(h^{k+1}\right) = argmin_h \quad \left\|h^k\right\|_1 + \lambda\left\|h^k - Mr - b^k\right\|_2^2 \tag{C3.14}$$

L1 norm minimisation problems can be obtained:

$$h^{k+1} = soft\left(Mr^{k+1} + b^k, {}^1/_\lambda\right) \tag{C3.15}$$

Where Mr=h and

$$b^{k+1} = b^k - (h^{k+1} - Mr^{k+1}) \tag{C3.16}$$

The $soft()$ function is the shrinkage operator and b^{k+1} is updated accordingly.

Stage 4: Illumination Adjustment

The output from the loop will result in the new illumination and reflectance. The illumination is further adjusted. The gamma function is adopted.

$$Lnew = 255 \times \left(\frac{L}{255}\right)^{\frac{1}{\gamma}} \quad \text{(C3.17)}$$

Stage 5: Combining the channels

To get the new *V* channel, $V = Lnew.* R$. The *V* channel is further enhanced by contrast limited adaptive equalisation to get the enhanced *V* channel. The new *V* channel is recombined with the untouched *S* and *H* channels. The HSV image is converted to the RGB image to produce the enhanced image.

Appendix C4: Algorithmic description of Natural Preserved Enhancement Algorithm for non-uniform illumination images

The author proposes the bright pass filter that constricts the reflectance to [0,1]. The bright pass filter (BPF) is the average of the adjacent pixels. These pixels are positively weighted and is associated to the frequency $Q(k,l)$. $Q(k,l)$ represents the local mean and win refers to the window size.

$$Q(k,l) = \frac{(\sum_{i=l-win}^{i=l+win} Q'(k,i))}{(2 \cdot win + 1)} \tag{C4.1}$$

Where,

$$Q'(k,l) = \sum_{x=1}^{m} \sum_{y=1}^{n} NN_{k,l}(x,y) \tag{C4.2}$$

$NN_{k,l}(x,y)$ specifies the amount of its neighbours of value l. The number of pixels in the height and width are denoted as m and n, respectively. The frequency $Q'(k,l)$ for the pixel value of k and l are neighbours everywhere in the image is defined in equation C4.2. The $BPF(\cdot)$ is defined as:

$$BPF(G(x,y)) = \frac{1}{W(x,y)} \sum_{(i,j)\epsilon\Omega} (Q(G(x,y), G(i,j)) \cdot U(G(i,j), G(x,y)) \cdot G(i,j)) \tag{C4.3}$$

The $G(x,y)$ is the set of neighbour pixels. The variable Ω refers to the local patch centred at the co-ordinates (x,y). Note that the patch is a small area of pixels. This is sometimes referred to as a window. The patch size is set to 15x15. The unit step function is denoted as $U(x,y)$. The $W(x,y)$, is the normalising factor such that the summation of the pixel weight is one.

$$U(x,y) = \begin{cases} 1 \; for \; x \geq y \\ 0 \; , else \end{cases} \tag{C4.4}$$

$$W(x,y) = \sum_{(i,j)\epsilon\Omega} (Q(G(x,y), G(i,j)) \cdot U(G(i,j), G(x,y))) \tag{C4.5}$$

The BPF is utilised in the image decomposition. The intensity $L(x,y)$ is obtained using:

$$L(x,y) = max_{c\epsilon[r,g,b]} I^c(x,y) \tag{C4.6}$$

Where $I^c(x,y)$ represents the lightness of the colour channel *C*. The BPF is refined to yield the illumination layer.

$$L_r(x,y) = \frac{1}{W(x,y)} \sum_{(i,j)\epsilon\Omega} (Q(L(x,y).L(i,j)) \cdot U(L(i,j), L(x,y)) \cdot L(i,j)) \tag{C4.7}$$

The reflectance $R(x,y)$ is attained by extracting the illumination layer.

$$R^c(x,y) = I^c(x,y)/L_r(x,y) \quad \text{(C4.8)}$$

The illumination mapping is done by utilising the bi-log transformation (BLT). The mapped illumination is obtained using the BLT.

$$L_m(x,y) = cf^{-1}[cL(L_r(x,y))] \quad for\ v = 0,1,2,\ldots.,L-1. \quad \text{(C4.9)}$$

$$cL(v) = \sum_{k=0}^{v} mp(k) = \frac{\sum_{i=0}^{m}\sum_{j=0}^{n} L_{lg}(i,j)\cdot U(v, L_r(i,j))}{\sum_{i=0}^{m}\sum_{j=0}^{n} L_{lg}(i,j)} \quad \text{(C4.10)}$$

$$L_{lg}(x,y) = log(L_r(x,y) + \epsilon) \quad \text{(C4.11)}$$

The log shape is given by C4.11. The parameter ϵ is a small constant and is assigned to 1. The weight histogram is defined as $mp(n)$. The $cL(v)$ represents the cumulative density function of the weighted histogram. Refer to Journal [4] for more information. The combination of the reflectance and mapped illumination for the resultant enhanced image is:

$$EI^c(x,y) = R^c(x,y) \times L_m(x,y) \quad \text{(C4.12)}$$

Algorithm 5: The NPEA

Input = RGB image Output= Enhanced RGB image
Stage 1: Image Decomposition First the image is broken down into reflectance and illumination using the BPF. This is done by utilising equation C4.3, C4.7 and C4.8.
Stage 2: Illumination Transformation The illumination is treated by utilising the bi-log transformation. This is done using equation C4.9.
Stage 3: Combination of reflectance and mapped illumination The mapping should not supress any details. The resultant image is attained by producing the reflectance and the mapping illumination (Equation C4.12).

Appendix C5: Algorithmic description of Automatic image Equalisation and contrast enhancement using Gaussian Mixture Modelling

Assume an input image, X has an image height of H and an image width of W such that $X = \{x(i,j)|1 \leq i \leq H, 1 \leq j \leq W\}$. The dynamic range (DR) of the image X is given as $[x_d, x_u]$. $x_d < x_u$ and $x_d, x_u \in \mathbb{R}$ such that $x(i,j) \in [x_d, x_u]$. Assume the algorithm yield an enhanced image Y, the height and width are defined H and W respectively. The size of image X and Y is HxW. The DR of the image Y is given as $[y_d, y_u]$. $y_d < y_u$ and $y_d, y_u \in \mathbb{R}$ such that $y(i,j) \in [y_d, y_u]$.

The GMM data distribution is in relation to the linear mixture of different Gaussian distribution that has diverse parameters. Every component of the Gaussian distribution consists of non-similar; mean, weight and standard deviation in the mixture module. The gray level distribution or spread of the input image is denoted as $p(x)$. The distribution is a linear combination of N Gaussian function utilising a GMM.

$$p(x) = \sum_{n=1}^{N} P(w_n) p(x|w_n) \tag{C5.1}$$

w_n is specific Gaussian n. The $p(x|w_n)$ represents the nth component density and the $P(w_n)$ signifies the previous probability of the data point produced by w_n.

$$p(x|w_n) = \frac{1}{\sqrt{2\pi\sigma_{w_n}^2}} exp^{\left(-\frac{(x-\mu_{w_n})^2}{2\sigma_{w_n}^2}\right)} \tag{C5.2}$$

Where μ_{w_n} is the mean or position of the Gaussian's peak, $\sigma_{w_n}^2$ is the variance (breadth and height relationship) therefore σ_{w_n} is the standard deviation (breadth and height relationship). The GGM is parameterised by:

$$\theta = \{P(w_n), \mu_{w_n}, \sigma_{w_n}^2\}_{n=1}^{N} \tag{C5.3}$$

Maximum-Likelihood techniques like Expectation-Maximisation can be used to estimate θ for a best fit to any given distribution. Maximum-Likelihood estimation techniques can adjust the statistical model such that the model best fits the data. For the likelihood, assuming that all data points $\boldsymbol{X} = \{x_1, x_2, \ldots, x_{HxW}\}$ are independent. The likelihood of $\boldsymbol{X}$ given by θ is:

$$\mathcal{L}(\boldsymbol{X}; \theta) = \prod_{\forall k} p(x_k; \theta) \tag{C5.4}$$

The distribution parameter is θ. The goal is to find $\hat{\theta}$ that maximises the likelihood.

$$\hat{\theta} = arg_{\theta} max\, \mathcal{L}(\boldsymbol{X}; \theta) \tag{C5.5}$$

The next step is the partitioning which is based on intersection points. The Figueriedo- Jain algorithm is used to simultaneously optimise the number of Gaussian components, as well as the means and variances. Having done this, the optimised mixture model is parameterised as:

$$\theta = \theta' = \{P(w_n), \mu_{w_n}, \sigma^2_{w_n}\}_{n=1}^{N} \quad \text{(C5.6)}$$

The intersection points of the above Gaussians lie within the dynamic range $[x_d, x_u]$ are now computed. The intersection of w_m and w_n may be found by solving:

$$P(w_m)p(x|w_m) = P(w_n)p(x|w_n) \quad \text{(C5.7)}$$

Now the equation has to be expanded and simplified to achieve the intersection points. These intersection points effectively partition the domain into intervals. Only the significant intersections x_s (those between dominant Gaussians) are considered. Furthermore, only significant intersections that are within the dynamic range of x are considered. A set of intervals is achieved.

$$[x_d, x_u] = [{x_s}^{(1)}, {x_s}^{(2)}] \cup [{x_s}^{(2)}, {x_s}^{(3)}]\,] \cup \cdots \cup [{x_s}^{(k-2)}], {x_s}^{(k-1)}]\,] \cup [{x_s}^{(k-1)}, {x_s}^{(k)}] \quad \text{(C5.8)}$$

Subinterval $[{x_s}^{(k)}, {x_s}^{(k+1)}]$ is denoted by a Gaussian component w_k which is dominant with respect to all remaining components in that interval. Mapping of intervals from input intervals to output intervals has to be computed. The intervals $[{x_s}^{(k)}, {x_s}^{(k+1)}]$ where $k = 1,2, \ldots . K-1$ is mapped onto the DR of the resultant image. *K* is the total number of intervals and *N* is the total number of Gaussian mixture models. An intensity transformation is used to transform the partitioned input dynamic range $[x_d, x_u]$ into the desired output dynamic range $[y_d, y_u]$. The weight α_k is defined as:

$$\alpha_k = \frac{\alpha_{w_k}^{\gamma}}{\sum_{i=1}^{N} \alpha_{w_k}^{\gamma}} \frac{F(x_s^{(k+1)} - F(x_s^{(k)})}{\sum_{is=1}^{K-1} F(x_s^{(i+1)}) - F(x_s^{(i)})} \quad \text{(C5.9)}$$

The first term alters the illumination of the equalised image. The variable $\gamma \in [0,1]$ is the illumination (brightness) constant. The γ is set to 0.5. Using, α_k the input interval $[{x_s}^{(k)}, {x_s}^{(k+1)}]$ is mapped onto the output interval $[{y_s}^{(k)}, {y_s}^{(k+1)}]$*:*

$$y^{(k)} = y_d + (y_u - y_d) \sum_{i=1}^{k-1} \alpha_i$$
$$y^{(k+1)} = y^{(k)} + \alpha_k (y_u - y_d) \quad \text{(C5.10)}$$

The mapping of Gaussians from input to output: The Gaussian distribution w_k with parameters μ_k and σ_k is defined in the input range $[x_d, x_u]$. The distribution is then transformed to the output range $[y_d, y_u]$,where it is represented as transformed distribution ${w_k}'$ with parameters ${\mu_k}'$ and ${\sigma_k}'$.

$$y = \sum_{i=1}^{N} ((\frac{x - \mu_{w_i}}{\sigma_{w_i}})\, \sigma_{w_i'} + \mu_{w_i'}) P_{w_i} \tag{C5.11}$$

$$\mu_{w_k'} = \frac{y^{k+1}(x_s^{(k)} - \mu_{w_k}) - y^{k}(x_s^{(k+1)} - \mu_{w_k})}{x_s{}^{(k)} - x_s{}^{(k+1)}} \tag{C5.12}$$

$$\sigma_{w_k'} = \frac{(y^{(k)} - \mu_{w_k'})}{(x_s^{(k)} - \mu_{w_k})} \sigma_{w_k} \tag{C5.13}$$

Given that $y^{(k)}$, $y^{(k+1)}$, $x_s^{(k)}$, $x_s^{(k+1)}$ and μ_k are known, $\mu_{w_k'}$ may be computed. Furthermore, given that σ_{w_k} and $\mu_{w_k'}$ are known, $\sigma_{w_k'}$ ′ may be computed.

Algorithm 6: The AEGMM algorithm

The algorithm for this enhancement was provided by Mr B. Naidoo

Input = RGB image

Output= Enhanced RGB image

Stage 1: Initialisation of enhancement

1.1 Load the input image.

1.2 Set all necessary parameters.

1.3 Convert image to greyscale if necessary.

1.4 Convert to CIELAB colour space, apply the colour transform and extract each component from Lab space.

Stage 2: Compute the FJ algorithm

Stage 3: Compute the Gaussian mixture intersection points

Compute the intersection points of all Gaussian components. Filtering is then required. The filter is applied to remove those points outside the dynamic range and those points that are intersection of non-dominant Gaussian in the given interval.

Stage 4: Weight the input intervals

Define a matrix to describe the dynamic range and its ordered subintervals moving from x_d to x_u.

Stage 5: Compute the output intervals

Compute the output intervals $[y(i), y(k)]$. The output dynamic range is set to $[0, 255]$.

Stage 6: Linear pixel transform

Utilise a dedicated linear pixel transform in each interval. The entire transform is piece-wise linear.

Stage 7: Output image

Produce the mapped image. Save in grayscale form. Convert image back to RGB.

Appendix C6: Algorithmic description of Image fusion-based contrast enhancement

This method was proposed by Saleem et al. [6]. The method is to fuse the input images. The fusion model requires three input images. It fuses the images as a weighted blending of the input images. The images are merged by calculating the weighted average for all the pixel.

$$F_{i,j} = \sum_{k=1}^{N} \widehat{W}_{i,j,k} I_{i,j,k} \tag{C6.1}$$

I_k, are the kth input images and $\widehat{W}_k$ are the k^{th} weight map. $F_{i,j}$ represents the resultant image. The weighted blending function is depicted in equation C6.2.

$$\widehat{W}_{i,j,k} = [\sum_{k\prime=1}^{N} W_{i,j,k}\prime]^{-1} \quad W_{i,j,k} \tag{C6.2}$$

The images are decomposed into a hierarchy of images. This is done using the Laplacian pyramid decomposition of the original image. It decomposes the images into levels that correlate to the diverse bands of image frequencies. Next, the Gaussian pyramid of the weighted maps are calculated. This is required for blending. Blending is done for each level individually. *N* is the number of input images.

$$L\{F\}_{i,j}^{l} = \sum_{k=1}^{N} G\{\widehat{W}\}_{i,j,k}^{l} L\{I\}_{i,j,k}^{l} \tag{C6.3}$$

L{F}, is the Laplacian pyramid while *G{F}* is the Gaussian pyramid. The $L\{F\}^{l}$ is collapsed to acquire the merged image *F*. The framework is depicted in Figure 1.

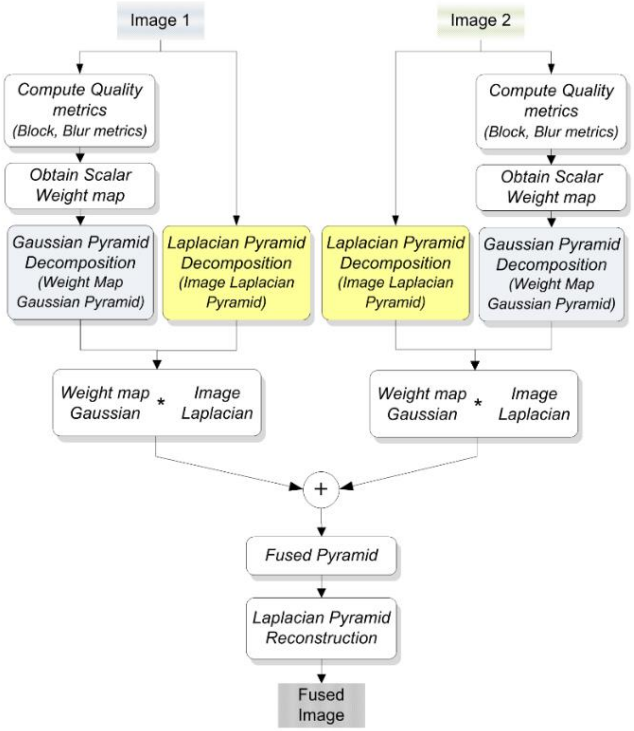


Figure 1: The framework for the image fusion-based contrast enhancement. The image was sourced from the journal article [6].

Appendix D: Algorithmic description for Performance Metrics

The following metrics are carefully studied and implemented:

1. Mean square error (MSE) [7],
2. Entropy [8],
3. Edge- based contrast measure (EBCM) [5],
4. Naturalness image quality evaluator (NIQE) [9],
5. The no-reference free energy based robust metric (NFERM) [10],
6. The no-reference image quality metric for contrast distortions (NIQMC) [11],
7. The colourfulness-based PCQI (patch-based contrast quality index [12]) (CPCQI) [13]
8. The blind/reference-less image spatial quality evaluator (BRISQUE) [14].

The metric NIQE, NFERM, NIQMC, CPCQI and BRISQUE were trained on the BSD300 database [15], Lossless database [16], USC-SIPI Image Database [17], LIVE IQA database [18] etc. All images used for the evaluation can be found in these databases. The mathematical expressions for MSE, entropy, NIQE, NIQMC and CPCQI were explained in Paper 1. The MSE function was provided by MathWorks [7]. The code for NIQE, NIQMC, CPCQI and BRISQUE were provided by the author as mentioned in Paper 1. Therefore, Appendix D provides the remaining mathematical models i.e. for NFERM and BRISQUE and the algorithms for entropy, EBCM.

Appendix D1: Algorithmic description of Entropy Metric

The entropy is expressed in D1.1. Where, pi is the probability of intensity value l in an image, and *H* is denoted as the entropy of the input image. *L* is defined as the total number of gray level.

$$H = -\sum_{l}^{L-1} pilog_2 pi \quad \text{(D1.1)}$$

Algorithm 7: The Entropy Metric

Input = RGB image that will be evaluated Output= Score of information
Stage 1: Initialisation Load image and then normalise image to a range of [0,1].
Stage 2: function "entropy" 2.1 Assume I in the range [0,1]. 2.2 Create histogram of image. 2.3 Remove zero entries that would cause log2 to be undefined. 2.4 Normalise histogram to unity. 2.5 Apply entropy's definition *H = -sum(p.*log2(p)).*
Stage 3: Pass image through entropy function and receive the score.

Appendix D2: Algorithmic description of Edge-Based Contrast Measure

For the EBCM we apply the following equations on MATLAB. The input image is denoted as *X*:

$$c(i,j) = \frac{|x(i,j) - e(i,j)|}{|x(i,j) + e(i,j)|} \tag{D2.1}$$

The mean edge gray level is given as:

$$e(i,j) = \frac{\sum_{(k,l)\in\mathcal{N}(i,l)} g(k,l)x(k,l)}{\sum_{(k,l)\in\mathcal{N}(i,l)} g(k,l)} \tag{D2.2}$$

$\mathcal{N}(i,l)$ signifies the set of neighbouring pixels of pixel (i,l). The magnitude of the image gradient projected utilising the Sobel operators at pixel (k,l) is defined as $g(k,l)$ [19]. The EBCM for image *X* is calculated as the average contrast value, that is:

$$EBCM(X) = \sum_{i=1}^{H}\sum_{j=1}^{W} c(i,j)/HW \tag{D2.3}$$

Algorithm 8: The Edge- based contrast measure metric

Input = RGB image to be evaluated

Output= Score of EBCM

Stage 1: Initialisation

Load image

The input image matrix (not normalised 0-255)

Stage 2: Function "entropy"

2.1 Input is presumed to be not normalised.

2.2 Get size of input image (Height * Width).

2.3 Obtain the Sobel kernel and normalise the image before processing.

2.4 Execute Sobel filter on normalised image.

2.5 Get $g(k,l)$ by multiplying the normalised image and the executed Sobel filter. This is then numerator of mean gray level equation.

2.6 Obtain average kernel, default size is 3x3 and execute image average filter on the Sobel filtered image.

2.7 Now calculate equation D2.2 to achieve the final mean gray level.

2.8 Then, calculate equation D2.1 to achieve the sum across all pixels of contrast image.

2.9 The equation D2.3 is used to obtain the EBCM result. This result is passed out.

Appendix D3: Algorithmic description of No-reference Free Energy based Robust Metric

The NFERM divides the features into three groups of features;

Group one of features:

Group one features are made up of 13 features ($f_{01} - f_{13}$). These features are of free energy and structural degradation information.

$$S_a(I) = E\left(\frac{\sigma_{(\mu_I\bar{\mu}_I)} + C1}{\sigma_{(\mu_I)}\sigma_{(\bar{\mu}_I)} + C1}\right) \tag{D3.1}$$

$$S_b(I) = E\left(\frac{\sigma_{(\sigma_I\bar{\sigma}_I)} + C1}{\sigma_{(\sigma_I)}\sigma_{(\bar{\sigma}_I)} + C1}\right) \tag{D3.2}$$

The $S_a(I)$ and $S_b(I)$ are the structural degradation for the input image (I). The teams μ_I and σ_I are the local mean and variance of the input image by a 2D circularly symmetric Gaussian weighted function. The term $\bar{\mu}_I$ and $\bar{\sigma}_I$ has the similar definitions except the utilisation of the impulse function in its place of the Gaussian weighting function. The function $E(\cdot)$ is a direct average pooling. $\sigma_{(\mu_I\bar{\mu}_I)}$ and $\sigma_{(\sigma_I\bar{\sigma}_I)}$ is represented by the local covariance. $C1$ refers to a minor constant that prevent the denominator from being zero or undefined. $S_a(I)$ is modified to retain diverse kinds of distortions:

$$\hat{S}_a(I) = \begin{cases} -S_a(I)\ if\ F(I) > T \\ S_a(I)\ \text{otherwise} \end{cases} \tag{D3.3}$$

where T is given as 5 conferring to the observation and $F(I)$ is the free energy approximation of the image I. $\hat{S}_b(I)$ is adapted similarly as $S_b(I)$. The linear dependence amongst the free energy feature and the structural degradation information offers a chance to characterise distorted images without original image information. $\hat{S}_s$ is the structural degradation information. The linear regression model is defined as:

$$\begin{aligned} F(I_r) &= \alpha_s \cdot \hat{S}_s(I_r) + \beta_s \\ F(I_r) &= \theta_s \cdot \check{S}_s(I_r) + \varphi_s \end{aligned} \tag{D3.4}$$

where α_s, β_s, θ_s and φ_s are attained from the least square method. The 12 features are defined as

$$\begin{cases} f_{01} - f_{06}: \hat{S}_s S_s, & s = \{a_1, a_3, a_5, b_1, b_2, b_5\} \\ f_{07} - f_{12}: \check{S}_s\, S_s, & s = \{a_1, a_3, a_5, b_1, b_2, b_5\} \end{cases} \tag{D3.5}$$

Where,

$$\begin{aligned} \hat{S}_s S_s &= F(I_d) - \left(\alpha_s \cdot \hat{S}_s(I_r) + \beta_s\right) \\ \check{S}_s\, S_s &= F(I_d) - (\theta_s \cdot \check{S}_s(I_r) + \varphi_s) \end{aligned} \tag{D3.6}$$

Furthermore, the NFERM corresponds efficiently with human ratings with regards to noisy and blurred images. This is used as the feature f_{13}.

Group two of features:

The second group consists of 6 features, $f_{14} - f_{19}$ is influenced by the free energy theory. This demonstrates that the HVS tries to recognise as well as comprehend a visual response. This is done by the reduction of the vagueness created on the internal generative model. Feature 14 is computed as the Peak Signal to Noise Ratio (PSNR) amongst the distorted image $\mathrm{I_d}$ and the predicted version I_p

$$f_{14} = 10log_{10}\left(\frac{255^2}{\frac{1}{M}\sum_{i=1}^{M}\left[I_d(i) - I_p(i)\right]^2}\right) \tag{D3.7}$$

M is defined as the number of pixels in the image. The luminance similarity is correlated to PSNR, the author chose contrast and structural similarities amongst I_d and I_p to be features $f_{15} - f_{16}$:

$$f_{15} = E\left(\frac{2\sigma_{(I_d)}\sigma_{(I_p)} + 2C_1}{\sigma^2_{(I_d)} + \sigma^2_{(I_p)} + 2C_1}\right) \tag{D3.8}$$

$$f_{16} = E\left(\frac{\sigma_{(I_dI_p)} + C_1}{\sigma_{(I_d)}\sigma_{(I_p)} + C_1}\right) \tag{D3.9}$$

$E(\cdot)$ is to calculate the mean or anticipation value. Physiological and psychophysical studies suggest the phase congruency *(PC)* model offers unpretentious yet biologically plausible model of how the HVS senses and recognises features of an image [20, 21]. Therefore, feature f_{17} is set as:

$$f_{17} = E(PC_m) = E\{max[PC(I_d), PC(I_p)]\} \tag{D3.10}$$

PC is defined and broadly implemented in [21]. The gradient magnitude (*GM*) is defined as $GM = \sqrt{GM_x^2 + GM_y^2}$, where, GM_x and GM_y are partial derivatives of the image in the horizontal (*x*) direction and vertical (*y*) directions utilising the Scharr operator (a gradient operator). The *GM* is the eighteenth feature f_{18}:

$$f_{18} = E\left(GM_{map}\right) = E\left(\frac{2GM(I_d)\cdot GM\left(I_p\right) + C_2}{GM(I_d)^2 + GM\left(I_p\right)^2 + C_2}\right) \tag{D3.11}$$

The salient regions (e.g. PC_m) have a greater effect on HVS when assessing the quality of an image. The *PC* component and *GM* components which are weighted by PC_m are combined to obtain the feature f_{19}:

$$f_{19} = \frac{E(GM_{map} \cdot PC_{map} \cdot PC_m)}{E(PC_m)} \tag{D3.12}$$

$$PC_{map} = \frac{2PC(I_d) \cdot PC(I_p) + C_3}{PC(I_d)^2 + PC\left(I_p\right)^2 + C_3} \tag{D3.13}$$

The variables C_2 and C_3 are similar to C_1. They are fixed constants.

Group three of features:

The third group consists of four features $f_{20} - f_{23}$ which ascend from the natural scene statistics (NSS) model. The author expresses the GGD as:

$$f\ (x;\ \alpha, \alpha^2) = \frac{\alpha}{2\beta\Gamma(\frac{1}{\alpha})} exp\left(-\left(\frac{|x|}{\beta}\right)^{\alpha}\right) \tag{D3.14}$$

$$\beta = \sigma \sqrt{\frac{\Gamma\left(\frac{1}{\alpha}\right)}{\Gamma\left(\frac{3}{\alpha}\right)}} \tag{D3.15}$$

$$\Gamma(a) = \int_0^{\infty} t^{a-1} e^{-t}\, dt \tag{D3.16}$$

The gamma function is denoted as $\Gamma(.)$. The parameter α influences the shape and structure of the GGD, whilst the σ^2 depicts the variance of the distribution. For the NFERM, the zero mean distribution is chosen owing to the generally symmetric distribution of the mean subtracted contrast normalisation (MSCN) coefficients. The model is deployed to fit the MSCN empirical distributions from distorted images and undistorted ones. In each image, the author estimates two pairs of parameters ($\alpha,\ \sigma^2$). This parameter is from a GGD fit of the MSCN coefficients at two scales. This creates the last group of features. The algorithm is now demonstrated.

Algorithm 9: The NFERM metric

Input: An image that needs to be evaluated.
Output: A quality score of the image. Higher value represents a lower quality.
Stage 1: Initialisation
Load an image, then create a function called NFERM. Pass input image to function.
Stage 2: Computing features
3.1 Convert image into gray image and convert to double
3.2 Compute the Free energy equation
3.3 Compute features from group one
3.4 Compute features from group two

3.5 Compute features from group three

3.6 Integrate the features

Stage 3: The Score

Once features are integrated the score is produced.

Appendix D4: Algorithmic description of Blind/Reference-less Image Spatial Quality Evaluator.

BRISQUE is a no-reference image quality assessor that operates in the spatial domain. The BRISQUE model utilises the NSS model of locally normalised luminance coefficients. The BRISQUE model enumerates naturalness. The model presents a statistic model of pairwise products of neighbouring luminance values. The claim made by the authors [14] is that the characterising locally normalised luminance coefficients is sufficient to quantify naturalness and quantify the amount of distortions. The NSS in spatial domain has to be determined. The GGD with zero mean is expressed as equation D4.1:

$$f\ (x;\ \alpha, \alpha^2) = \frac{\alpha}{2\beta\Gamma(\frac{1}{\alpha})} exp\left(-\left(\frac{|x|}{\beta}\right)^{\alpha}\right) \tag{D4.1}$$

$$\beta = \sigma\sqrt{\frac{\Gamma\left(\frac{1}{\alpha}\right)}{\Gamma\left(\frac{3}{\alpha}\right)}} \tag{D4.2}$$

$$\Gamma(a) = \int_0^{\infty} t^{a-1}e^{-t}\ dt\ \ a > 0 \tag{D4.3}$$

The shape of the distribution is controlled by the parameter α. The variance is controlled by parameter α^2. The theory is that the MSCN coefficients are symmetric and they have distinctive statistical properties which are altered by distortion. Measuring these alternations makes it probable to envisage the nature of distortion disturbing an image and its perceptual quality. The zero mean distribution is chosen because the MSCN coefficients are symmetric. For the Gaussian coefficient model, and the assumption that the MSCN coefficients are zero mean and unit variance, these products follow the distribution in the non-appearance of distortion [22].

$$f(x, \rho) = \frac{exp\left(\frac{|x|\rho}{1-\rho^2}\right)\ K_0\left(\frac{|x|}{1-\rho^2}\right)}{\pi\sqrt{1-\rho^2}} \tag{D4.4}$$

The asymmetric probability density function is defined as f. The variable, ρ signifies the correlation coefficient of adjacent coefficients. The K_0 is the adapted Bessel function of the second kind. The author implements the general asymmetric generalised Gaussian distribution (AGGD) model [23]. The AGGD with zero mode is expressed by:

$$f(x;v;\sigma_l^2;\sigma_r^2) = \begin{cases} \frac{v}{(\beta_l+\beta_r)\Gamma\left(\frac{1}{v}\right)} exp(-\left(\frac{-x}{\beta_l}\right)^v) & x < 0 \\ \frac{v}{(\beta_l+\beta_r)\Gamma\left(\frac{1}{v}\right)} exp(-\left(\frac{x}{\beta_r}\right)^v) & x \geq 0 \end{cases} \quad \text{(D4.5)}$$

Where $\beta_l = \sigma_l \sqrt{\frac{\Gamma\left(\frac{1}{v}\right)}{\Gamma\left(\frac{3}{v}\right)}}$ and $\beta_r = \sigma_r \sqrt{\frac{\Gamma\left(\frac{1}{v}\right)}{\Gamma\left(\frac{3}{v}\right)}}$

The shape of the distribution is controlled by the parameter v controls. The spread is controlled by the parameters σ_l^2 and σ_r^2 on the separate sides of the model. The author adopts an asymmetric generalised Gaussian distribution (AGGD). The parameters $(\eta, v, \sigma_l^2, \sigma_r^2)$ of the best AGGD fit are removed where η is expressed as:

$$\eta = (\beta_l - \beta_r) \frac{\Gamma\left(\frac{2}{v}\right)}{\Gamma\left(\frac{1}{v}\right)} \quad \text{(D4.6)}$$

Each paired product, consisting of sixteen parameters are calculated, producing the next set of features. Images are naturally multiscale and distortions affect image structure. Studies showed that incorporating multiscale information when evaluating the quality assessment methods perform better and correlate with the HSV [24, 25].

Appendix E: Survey for Paper 1

An analysis of objective and human assessments in contrast enhancement- Questionnaire

As part of my master's research at the University of KwaZulu Natal, I am conducting a survey that investigates the quality of images. This visual assessment will allow the recipient to score enhanced images according to how they perceive it. This survey is to be taken willingly and respondents are not obliged to participate. Your attitudes and opinions are critical to the success of the study. The value of your time is recognised and your efforts are sincerely appreciated. This survey should take 10-15 minutes to complete. Thank you for your time and responses.

Demographic Data

Age: ☐ Gender: ☐ Male ☐ Female

Any visual disability: ☐

If so, provide details: _____

On a scale from 1 to 10, where 1 is poor vision and 10 is excellent vision, how would you rate your vision? ☐

Image processing involves manipulation of a digitalised image, normally to improve the quality of the image. Digital image processing technique can be applied in many different fields such as object detection and matching, background subtraction in video, traffic control systems, locating objects in face recognition, iris recognition, medical imaging, etc. Digital image processing addresses challenges and issues such as loss of image quality and to enhance degraded images. This experiment aims to assess image enhancements methods. The survey will be conducted in two parts/sections:

Instructions:

Section 1: Enhancement Rating

You are given 8 sets of 6 images. Each set contains the original image and five enhanced images placed next to it. You must compare and evaluate each image on a scale from score 1 to 5. The scores indicate how you perceive the enhancement quality and visual preference, where:

1-very poor (the enhancement is much worse)

2- poor (worse)

3- the image is the same

4- good (Image has improved)

5- excellent (the image is much better than original)

Section 2: Best image

You are given 8 sets of 6 images and you will be required to choose the most natural/unnatural/detail image. The original image is included in the set. The set is mixed randomly.

Section 1: Enhancement Rating

1. **Image set 1**

Please tick the score you choose to give

Image \ Rating	1	2	3	4	5
2	☐	☐	☐	☐	☐
3	☐	☐	☐	☐	☐
4	☐	☐	☐	☐	☐
5	☐	☐	☐	☐	☐
6	☐	☐	☐	☐	☐

2. **Image set 2**

Please tick the score you choose to give

Image \ Rating	1	2	3	4	5
2	☐	☐	☐	☐	☐
3	☐	☐	☐	☐	☐
4	☐	☐	☐	☐	☐
5	☐	☐	☐	☐	☐
6	☐	☐	☐	☐	☐

3. **Image set 3**

Please tick the score you choose to give

Image \ Rating	1	2	3	4	5
2	☐	☐	☐	☐	☐
3	☐	☐	☐	☐	☐
4	☐	☐	☐	☐	☐
5	☐	☐	☐	☐	☐
6	☐	☐	☐	☐	☐

4. Image set 4

Please tick the score you choose to give

Image \ Rating	1	2	3	4	5
2	☐	☐	☐	☐	☐
3	☐	☐	☐	☐	☐
4	☐	☐	☐	☐	☐
5	☐	☐	☐	☐	☐
6	☐	☐	☐	☐	☐

5. Image set 5

Please tick the score you choose to give

Image \ Rating	1	2	3	4	5
2	☐	☐	☐	☐	☐
3	☐	☐	☐	☐	☐
4	☐	☐	☐	☐	☐
5	☐	☐	☐	☐	☐
6	☐	☐	☐	☐	☐

6. Image set 6

Please tick the score you choose to give

Image \ Rating	1	2	3	4	5
2	☐	☐	☐	☐	☐
3	☐	☐	☐	☐	☐
4	☐	☐	☐	☐	☐
5	☐	☐	☐	☐	☐
6	☐	☐	☐	☐	☐

7. Image set 7

Please tick the score you choose to give

Image \ Rating	1	2	3	4	5
2	☐	☐	☐	☐	☐
3	☐	☐	☐	☐	☐
4	☐	☐	☐	☐	☐
5	☐	☐	☐	☐	☐
6	☐	☐	☐	☐	☐

8. Image set 8

Please tick the score you choose to give

Image \ Rating	1	2	3	4	5
2	☐	☐	☐	☐	☐
3	☐	☐	☐	☐	☐
4	☐	☐	☐	☐	☐
5	☐	☐	☐	☐	☐
6	☐	☐	☐	☐	☐

Section 2: The best image

1. Which image looks the most natural and which one looks least natural?
   Please tick the most natural image and add a cross to the least natural image.

Image set	a	b	c	d	e	f
9	☐	☐	☐	☐	☐	☐
10	☐	☐	☐	☐	☐	☐
11	☐	☐	☐	☐	☐	☐
12	☐	☐	☐	☐	☐	☐
13	☐	☐	☐	☐	☐	☐
14	☐	☐	☐	☐	☐	☐
15	☐	☐	☐	☐	☐	☐
16	☐	☐	☐	☐	☐	☐

2. Which image has the most detail?
   Please tick the rating you choose to give.

Image	a	b	c	d	e	f
9	☐	☐	☐	☐	☐	☐
10	☐	☐	☐	☐	☐	☐
11	☐	☐	☐	☐	☐	☐
12	☐	☐	☐	☐	☐	☐
13	☐	☐	☐	☐	☐	☐
14	☐	☐	☐	☐	☐	☐
15	☐	☐	☐	☐	☐	☐
16	☐	☐	☐	☐	☐	☐

Thank you for your time.

Appendix F: Survey for Paper 2

Multimodal Enhancement-Fusion Technique for Natural Images – Questionnaire

As part of my master's research at the University of KwaZulu Natal, I am conducting a survey that investigates the quality of images. This visual assessment will allow the recipient to score enhanced images according to how they perceive it. This survey is to be taken willingly and respondents are not obliged to participate. Your attitudes and opinions are critical to the success of the study. The value of your time is recognised and your efforts are sincerely appreciated. This survey should take 10-15 minutes to complete. Thank you for your time and responses.

Demographic Data

Age: ☐ Gender: ☐ Male ☐ Female

Any visual disability: ☐

If so, provide details: _____

On a scale from 1 to 10, where 1 is poor vision and 10 is excellent vision, how would you rate your vision? ☐

Image processing involves manipulation of a digitalised image, normally to improve the quality of the image. Digital Image Processing technique can be applied in many different fields such as object detection and matching, background subtraction in video, traffic control systems, locating objects in face recognition, iris recognition, medical imaging, etc. Digital Image Processing addresses challenges and issues such as loss of image quality and to enhance degraded images. This experiment aims to assess image enhancements methods.

Instructions:

Enhancement Rating

You are given 8 sets of 6 images. Each set contains the original image and five enhanced images placed next to it. You must compare and evaluate each image on a scale from score 1 to 5. The scores indicate how you perceive the enhancement quality and visual preference, where:

1-very poor (the enhancement is much worse)

2- poor (worse)

3- the image is the same

4- good (Image has improved)

5- excellent (the image is much better than original)

1. **Image set 1**

Please tick the score you choose to give

Image \ Rating	1	2	3	4	5
2	☐	☐	☐	☐	☐
3	☐	☐	☐	☐	☐
4	☐	☐	☐	☐	☐
5	☐	☐	☐	☐	☐
6	☐	☐	☐	☐	☐

2. **Image set 2**

Please tick the score you choose to give

Image \ Rating	1	2	3	4	5
2	☐	☐	☐	☐	☐
3	☐	☐	☐	☐	☐
4	☐	☐	☐	☐	☐
5	☐	☐	☐	☐	☐
6	☐	☐	☐	☐	☐

3. **Image set 3**

Please tick the score you choose to give

Image \ Rating	1	2	3	4	5
2	☐	☐	☐	☐	☐
3	☐	☐	☐	☐	☐
4	☐	☐	☐	☐	☐
5	☐	☐	☐	☐	☐
6	☐	☐	☐	☐	☐

4. Image set 4

Please tick the score you choose to give

Image \ Rating	1	2	3	4	5
2	☐	☐	☐	☐	☐
3	☐	☐	☐	☐	☐
4	☐	☐	☐	☐	☐
5	☐	☐	☐	☐	☐
6	☐	☐	☐	☐	☐

5. Image set 5

Please tick the score you choose to give

Image \ Rating	1	2	3	4	5
2	☐	☐	☐	☐	☐
3	☐	☐	☐	☐	☐
4	☐	☐	☐	☐	☐
5	☐	☐	☐	☐	☐
6	☐	☐	☐	☐	☐

6. Image set 6

Please tick the score you choose to give

Image \ Rating	1	2	3	4	5
2	☐	☐	☐	☐	☐
3	☐	☐	☐	☐	☐
4	☐	☐	☐	☐	☐
5	☐	☐	☐	☐	☐
6	☐	☐	☐	☐	☐

7. Image set 7

Please tick the score you choose to give

Image \ Rating	1	2	3	4	5
2	☐	☐	☐	☐	☐
3	☐	☐	☐	☐	☐
4	☐	☐	☐	☐	☐
5	☐	☐	☐	☐	☐
6	☐	☐	☐	☐	☐

8. Image set 8

Please tick the score you choose to give

Image \ Rating	1	2	3	4	5
2	☐	☐	☐	☐	☐
3	☐	☐	☐	☐	☐
4	☐	☐	☐	☐	☐
5	☐	☐	☐	☐	☐
6	☐	☐	☐	☐	☐

References

[1] S. Arora and K. Kapoor, "Colour image enhancement based on histogram equalization," *Electrical & Computer Engineering: An International Journal (ECIJ),* vol. 4, no. 3, September 2015.

[2] S. bharal, "L*a*b based contrast limited adaptive histogram equalization for underwater images," *International Journal of Computer Application (2250-1797),* vol. 5, no. 4, June 2015.

[3] H. Yue, J. Yang, X. Sun, F. Wu and C. Hou, "Contrast Enhancement Based on Intrinsic Image Decomposition," *IEEE transactions in image processing,* vol. 26, no. 8, pp. 3981 - 3994, 8 August 2017.

[4] S. Wang, J. Zheng, H.-M. Hu and B. Li, "Naturalness Preserved Enhancement Algorithm for non uniform illumination images," *IEEE Transactions on Image Processing,* vol. 9, no. 9, 2013.

[5] T. Celik and T. Tjahjadi, "Automatic Image Equalization and Contrast Enhancement Using Gaussian Mixture Modeling," *IEEE Transactions on Image Processing,* vol. 21, no. 1, 2012.

[6] A. Saleem, A. Beghdadi and B. Boashash, "Image fusion-based contrast enhancement," *EURASIP Journal on Image and Video Processing,* no. 10, 2012.

[7] M. Matlab, "Mean Square Error," [Online]. Available: https://www.mathworks.com/help/images/ref/immse.html.

[8] S. Rani, A. Kumar and K. Singh, "Illumination based Sub Image Histogram Equalization: A Novel Method of Image Contrast Enhancement," *International Journal of Computer Applications,* vol. 119, no. 20, 2015.

[9] A. Mittal, R. Soundararajan and A. C. Bovik, "Making a Completely Blind Image Quality Analyzer," *IEEE Signal Processing Letters,* vol. 20, no. 3, pp. 209-212, 2012.

[10] K. Gu, G. Zhai, X. Yang and W. Zhang, "Using Free Energy Principle For Blind Image Quality Assessment," *IEEE Transactions on Multimedia,* vol. 17, no. 1, 2015.

[11] K. Gu, W. Lin, G. Zhai, X. Yang, W. Zhang and C. W. Chen, "No-Reference Quality Metric of Contrast-Distorted Images Based on Information Maximization," vol. 47, no. 12, 2017.

[12] S. Wang, K. Ma, H. Yeganeh, Z. Wang and W. Lin, "A patchstructure representation method for quality assessment of contrast changed images," *IEEE Signal Process. Lett.,* vol. 22, no. 12, pp. 2387-2390, December 2015.

[13] K. Gu, D. Tao, J. F. Qiao and W. Lin, “Learning a No-Reference Quality Assessment Model of Enhanced Images With Big Data,” *IEEE transactions on neural networks and learning systems*, vol. 29, no. 4, pp. 1301-1313, April 2018.

[14] A. Mittal, A. K. Moorthy and A. C. Bovik, “No-Reference Image Quality Assessment in the Spatial Domain,” *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 4695 - 4708, December 2012.

[15] Berkeley, “The Berkeley Segmentation Dataset and Benchmark,” [Online]. Available: https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/BSDS300/html/dataset/images.html. [Accessed May 2018].

[16] Kodak, “Kodak Lossless True Color Image Suite,” 2013. [Online]. Available: htttp://r0k.us/graphics/kodak/. [Accessed May 2018].

[17] University of Southern California, “The USC-SIPI Image Database,” [Online]. Available: http://sipi.usc.edu/database/. [Accessed May 2018].

[18] University of Texas, “LIVE Image Quality Assessment Database,” [Online]. Available: http://live.ece.utexas.edu/research/quality/subjective.htm. [Accessed May 2018].

[19] R. C. Gonzalez and R. E.Woods, Digital Image Processing, Upper Saddle River: Prentice-Hall, 2006.

[20] M. C. Morrone, J. Ross, D. C. Burr and R. Owens, “Mach bands are phase dependent,” *Nature, International Journal of Science*, vol. 324, pp. 250-253, Nov. 1986.

[21] P. Kovesi, “Image features from phase congruency,” *Videre: J. Comp. Vis. Res.*, vol. 1, no. 3, pp. 1-26, 1999.

[22] A. H. Nuttal, “Accurate efficient evaluation of cumulative or exceedance probability distributions directly from characteristic functions,” Naval Underwater Systems Center, New London, 1983.

[23] N. E. Lasmar, Y. Stitou and Y. Berthoumieu, “Multiscale skewed heavy tailed model for texture analysis in Processing” *IEEE Int. Conf. Image processing*, pp. 2281–2284, Nov. 2009.

[24] M. Saad, A. C. Bovik and C. Charrier, “Blind image quality assessment: A natural scene statistics approach in the DCT domain,” *IEEE Trans. Image Processing*, vol. 21, no. 8, pp. 3339–3352, Aug. 2012.

[25] Z. Wang, E. P. Simoncelli and A. C. Bovik, “Multiscale structural similarity for image quality assessment,” *Proc. Asilomar Conf. Signals Systems Computer*, vol. 2, pp. 1398–1402, 2003.