**“Comparing Image Quality Assessment Algorithms with Human Perception for Optimising Image Compression for ExoMars”**

CS39440 Major Project

Author: Charlie Curtis ([chc73@aber.ac.uk](mailto:chc73@aber.ac.uk))

Supervisor: Dr. Helen Miles ([hem23@aber.ac.uk](mailto:hem23@aber.ac.uk))

3rd March 2022

Version 1.0 (Draft)

This report is submitted as partial fulfilment of a BSc degree in

Computer Science (G400)

Department of Computer Science

Aberystwyth University

Aberystwyth

Ceredigion

SY23 3DB

Wales, UK

Declaration of originality

I confirm that:

* This submission is my own work, except where clearly indicated.
* I understand that there are severe penalties for Unacceptable Academic Practice, which can lead to loss of marks or even the withholding of a degree.
* I have read the regulations on Unacceptable Academic Practice from the University’s Academic Registry (AR) and the relevant sections of the current Student Handbook of the Department of Computer Science.
* In submitting this work, I understand and agree to abide by the University’s regulations governing these issues.

Name …………………………………………

Date ……………………………………………

Consent to share this work

By including my name below, I hereby agree to this project's report and technical work being made available to other students and academic staff of the Aberystwyth Computer Science Department.

Name …………………………………………

Date ……………………………………………

Acknowledgements

I am grateful to…

I’d like to thank…

Abstract

Include an abstract for your project. This should be approximately 300 words.

The abstract is an overview of the work you have done. Highlight the purpose of the work and the key outcomes of the work.

Contents

Table of Contents

[2 Background, Analysis & Process 5](#_Toc98490675)

[2.1 Background 5](#_Toc98490676)

# 

# **Background, Analysis & Process**

## **Background**

A sensible approach to begin seemed to be to conduct some preliminary research into image compression. This included not only image compression algorithms used in a variety of disciplines including medicine, graphic design and military applications, but also, the theory behind image compression, i.e., lossy versus lossless compression, techniques used in image compression, included but not limited to; quantization, discrete cosine transform (DCT), fractal compression – and performance comparisons of some compression algorithms.

The goal of this was to try and establish what sort of numerical operations these algorithms were performing, and through this, trying to build an understanding of how the quality assessment algorithms might begin to calculate numerical quality in a general sense. This was very useful for understanding metrics such as Mean Square Error (MSE), Root Mean Square Error (RMSE), and Peak Signal to Noise Ratio (PSNR).

During this preliminary research, a grouping system of image quality assessment (IQA) methods was identified. Methods are grouped into Full Reference (FR), Reduced Reference (RR), and No Reference (NR) IQA methods. They refer to if a ‘perfect’ quality - original image is used as part of the algorithm to compare to the compressed version. For this project, FR seemed the most appropriate since there *are* original images to be accessed and so the more comprehensive algorithms can be exercised as part of the study, in an effort to provide as accurate a set of results as possible.

Initial research into image compression in a medical setting proved particularly insightful. It’s hypothesised that since diagnosis and prognosis of disease have immense consequences - primarily to the quality of life and life-expectancy of humans - the quality of images used as part of these processes are vital, in ensuring the maximum amount of accurate information is portrayed to doctors and nurses dealing with patients. This is useful as it provides an upper ‘boundary’ or guideline, as to what may be considered an ‘unsafe’ level of compression due to the renunciation of various details within an image. In the context of PanCam, this may be a small mineral vein, or water droplet for instance.

This research of course has its limitations, however. Because of the high standard of modern technology and the lack of necessity to send these images over huge distances such as those involved with inter-planetary communication. Compression algorithms are rarely applied in a medical scenario since the risk of missing information is too great to be endured. High quality images are taken to facilitate the work of medical professionals, the most common form of compression in medicine is lossless compression, the reversible process means details need not be lost. Because of this, there isn’t much to be learnt about specifically IQA, since images aren’t routinely compressed in a lossy format, and as such, are always of maximum quality. Given the ultimate judge of image quality is the human visual system (HVS) and subjective study by humans, it can be concluded that - given misdiagnosis accounts for about 10% of adverse events in hospitals in the UK [2] – image quality is ‘high enough’, according to the HVS in 90% of cases.

One motivation for this project was having the opportunity to produce something that could potentially have real world implications towards the operational effectiveness of the European Space Agencies (ESA) Rosalind Franklin rover. Moreover, developing a tailored script that’s sole purpose is to facilitate the gathering, processing and analysis of the numerical aspects of IQA.

## **Analysis**

The project brief had great scope, allowing for numerous possible approaches to the question of “Assessing image quality for ExoMars.” After it was proposed to conduct a subjective study, using human participants to observe varying levels of compression on a series of images, it was later decided upon that it would best suit this project to compare the findings from this study, which was eventually done through other avenues, with the varying scores of different IQAs through the means of statistical analysis.

Hence, the title of the project.

This posed a few key considerations. Which IQAs should be used? Simple, accessible algorithms like the ones mentioned previously could act as controls however wouldn’t provide much substance to the investigation in isolation since they are strictly numerical metrics that don’t (as will be explored later in the report) don’t attempt to replicate any aspects of the HVS, moreover, the quality ratings these algorithms provide, don’t always correlate to that provided by the HVS or a Mean Opinion Score (MOS) derived from asking human participants to rate the quality of a given image.

# **References**

1.

2. Neale G, Hogan H, Sevdalis N. Misdiagnosis: analysis based on case record review with proposals aimed to improve diagnostic processes. *Clin Med (Lond)*. 2011;11(4):317-321. doi:10.7861/clinmedicine.11-4-317