A MULTI-FOCUS IMAGE FUSION TECHNIQUE FOR RGB COLOR IMAGES USING DISCRETE COSINE TRANSFORM (DCT).

SAFA RIYADH WAHEED

A dissertation submitted in partial fulfillment of the requirements for the award of the degree of Master of Science (Computer Science)

Faculty of Computing
Universiti Teknologi Malaysia

JAN 2014
This dissertation is dedicated to my family for their endless support and encouragement.
ACKNOWLEDGEMENTS

All praise is to Allah and my peace and blessing of Allah be upon our prophet, Muhammad and upon all his family and companions. In particular, I wish to express my sincere appreciation to my thesis supervisor, PROF. DR. GHAZALI SULONG, and to all my friends for encouragement, guidance, critics, advices and supports to complete this research. I really appreciate his ethics and great deal of respect with his students, which is similar to brothers in the same family.

In addition, I am extremely grateful to my father for unlimited support and encouragement during this research. I would like to thanks my beloved family: mother, fiancée, brother, and sisters who I am always beholden to them for their everlasting patience, support, encouragement, sacrifice, and love which have been devoted to me sincerely. I could endure for being away and eventually my master program came to fruition. For that, I ask Allah to bless all of them.

Besides, I would like to thank the authority of Universiti Teknologi Malaysia (UTM) and Faculty of Computing (FK) for providing me with a good environment and facilities such as computer laboratory to complete this project with software which I need during process.
Abstract

Image fusion is a process that combines information from multiple images of the same scene into a single image. In this study, a new image fusion technique is proposed using the Discrete Cosine Transform (DCT). Each of the source images may represent a partial view of the scene, and contains both ‘relevant’ and ‘irrelevant’ data. Since the target of image fusion technique is to combine the source images into a single composite image which contains a more accurate description of the scene than any of the individual sources. Moreover, the fused image has gotten the best possible quality without introduction of distortion or loss of information. DCT algorithm is considered efficient in image fusion. The proposed scheme is performed in five steps: (1) RGB colour image (input image) is split into three channels R, G, and B for both images. (2) DCT algorithm is applied to each channel (R, G and B) for image1 and image2 respectively. (3) The variance values are computed for the corresponding 8x8 blocks of each channel. (4) Each block of R (image 1) is compared with each block of R (image2) based on the variance value and then the block having maximum variance value is selected to be the block in the new image. This process is repeated for all channels of image 1 and image 2. (5) Inverse discrete cosine transform (IDCT) is applied on each fused channel to convert it from coefficient values to pixel values, and then combined all the channels to generate the fused image. The proposed technique can potentially solve the problem of undesirable side effects like blurring or blocking artifacts from reducing the quality of the resultant image in image fusion. The proposed approach is evaluated using three measurement units: the average of Qabf, standard deviation and peak Signal Noise Rate (PSNR). The experimental results of this proposed technique achieved good results as compared to previous studies.
Abstrak

# TABLE OF CONTENT

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DECLARATION</td>
<td>ii</td>
<td></td>
</tr>
<tr>
<td>DEDICATION</td>
<td>iii</td>
<td></td>
</tr>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td>iv</td>
<td></td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>v</td>
<td></td>
</tr>
<tr>
<td>ABSTRAK</td>
<td>vi</td>
<td></td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>vii</td>
<td></td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
<td></td>
</tr>
</tbody>
</table>

## 1 INTRODUCTION

1.1 Introduction 1
1.2 Problem Background 3
1.3 Problem Statement 5
1.3.1 Research Question 6
1.5 Aim of The Study 6
1.6 Objectives of The Study 7
1.7 Scope of The Study 7
1.8 Thesis organization 8

## 2 LITERATURE REVIEW

2.1 Introduction 9
2.2 RGB Color Images 9
2.2.1 Color spaces 10
2.2.1.1 RGB 11
2.2.1.2 RGB to grey-scale image conversion 14
2.2.1.3 YCbCr Color Space 15
### 2.3 Purpose of image fusion

#### 2.4 Image fusion technologies overview

#### 2.5 Image fusion technologies

<table>
<thead>
<tr>
<th>2.5.1</th>
<th>Fusion using Principle Component Analysis (PCA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5.2</td>
<td>Fusion using HIS color space</td>
</tr>
<tr>
<td>2.5.3</td>
<td>Fusion using Laplacian pyramid method</td>
</tr>
<tr>
<td>2.5.4</td>
<td>Fusion using gradient pyramid method</td>
</tr>
<tr>
<td>2.5.5</td>
<td>Fusion using filter-subtract-decimate (FSD) pyramid method</td>
</tr>
<tr>
<td>2.5.6</td>
<td>Fusion using discrete wavelet transforms (DWT)</td>
</tr>
<tr>
<td>2.5.7</td>
<td>Multi-sensor and multi-resolution image fusion using the linear mixing model</td>
</tr>
<tr>
<td>2.5.8</td>
<td>Image fusion schemes using ICA bases</td>
</tr>
<tr>
<td>2.5.9</td>
<td>Statistical modeling for wavelet-domain image fusion</td>
</tr>
<tr>
<td>2.5.10</td>
<td>Bayesian methods for image fusion</td>
</tr>
<tr>
<td>2.5.11</td>
<td>Multidimensional fusion by image fusion</td>
</tr>
<tr>
<td>2.5.12</td>
<td>Fusion of multispectral and images as an optimization problem</td>
</tr>
<tr>
<td>2.5.13</td>
<td>Fusion of edge maps using statistical approaches</td>
</tr>
<tr>
<td>2.5.14</td>
<td>Region-based multi-focus image fusion</td>
</tr>
<tr>
<td>2.5.15</td>
<td>Concepts of image fusion in remote sensing applications</td>
</tr>
<tr>
<td>2.5.16</td>
<td>Pixel-level image fusion metrics</td>
</tr>
<tr>
<td>2.5.17</td>
<td>Objectively adaptive image fusion</td>
</tr>
<tr>
<td>2.5.18</td>
<td>Performance evaluation of image fusion techniques</td>
</tr>
</tbody>
</table>
3 RESEARCH METHODOLOGY

3.1 Introduction 53
3.2 General framework 54
3.3 Discrete Cosine Transform (DCT) 58
  3.3.1 The One-Dimensional DCT 58
  3.3.2 The Two Dimensional DCT 59
3.4 Material And Method 60
3.5 Performance measure 61
3.6 Comparing both fusion method 64
3.7 Robustness Testing 65
3.8 Summary 65

4 INITIAL RESULT

4.1 Introduction 66
4.2 Dataset 66
4.3 Implementation Result 70
4.4 summary 83

5 CONCLUSION

5.1 summary 84
5.2 Contribution 85
5.3 Future Work 85

6 REFERENCES 86
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>The RGB color cube</td>
<td>10</td>
</tr>
<tr>
<td>2.2</td>
<td>The RGB channels</td>
<td>12</td>
</tr>
<tr>
<td>2.3</td>
<td>RGB color space as a 3-D cube</td>
<td>13</td>
</tr>
<tr>
<td>2.4</td>
<td>An example of RGB color image (left) to grey-scale image (right) conversion</td>
<td>14</td>
</tr>
<tr>
<td>2.5</td>
<td>The conversion formula between digital RGB data and YCbCr</td>
<td>16</td>
</tr>
<tr>
<td>2.6</td>
<td>Image fusion for an IKONOS scene for Cairo, Egypt: (a) multispectral low resolution input image, (b) panchromatic high resolution input image, and (c) fused image of IHS</td>
<td>17</td>
</tr>
<tr>
<td>2.7</td>
<td>CCD visual images with the (a) right and (b) left clocks out of focus, respectively; (c) the resulting fused image from (a) and (b) with the two clocks in focus</td>
<td>18</td>
</tr>
<tr>
<td>2.8</td>
<td>Images captured from the (a) visual sensor and (b) infrared sensor, respectively; (c) the resulting fused image from (a) and (b) with all interesting details in focus</td>
<td>19</td>
</tr>
<tr>
<td>2.9</td>
<td>(a) MRI and (b) PET images; (c) fused image from (a) and (b)</td>
<td>20</td>
</tr>
<tr>
<td>2.10</td>
<td>Image fusion block scheme of different abstraction levels</td>
<td>21</td>
</tr>
<tr>
<td>2.11</td>
<td>The image fusion classification level</td>
<td>23</td>
</tr>
<tr>
<td>2.12</td>
<td>The image fusion system using ICA/ Topographical ICA bases</td>
<td>28</td>
</tr>
<tr>
<td>2.13</td>
<td>Fusion of multispectral and panchromatic images as an optimization problem</td>
<td>32</td>
</tr>
<tr>
<td>2.14</td>
<td>The subjective and objective image fusion process evaluation</td>
<td>35</td>
</tr>
<tr>
<td>2.15</td>
<td>Block diagrams of generic fusion schemes where the input images have identical (a)</td>
<td></td>
</tr>
</tbody>
</table>
and different resolutions (b)

2.16 Objective method for image quality assessment 42

2.17 Methods for assessing image quality 46

3.1 Set 1. Image 1 (a) focus on left, Image 2 (b) focus on right 54

3.2 General framework of the proposed technique 55

3.3 DCT fusion process for selected channel 60

4.1 Qabf Evaluation using RGB and YCbCr 78

4.2 Qabf Evaluation levels using RGB and YCbCr 78

4.3 Standard Deviation Evaluation using RGB and YCbCr 79

4.4 Standard Deviation Evaluation levels using RGB and YCbCr 79

4.5 PSNR for fused images RGB and YCbCr 81

4.6 Mean Square Error for fused images by using RGB and YCbCr 81

4.7 Structural Content for fused images by using RGB and YCbCr 82

4.8 Normalized Cross-Correlation for fused images by using RGB and YCbCr 82
## LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE NO.</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>comparison of different image fusion techniques</td>
<td>43</td>
</tr>
<tr>
<td>4.1</td>
<td>Data set Images</td>
<td>67</td>
</tr>
<tr>
<td>4.2</td>
<td>present the 19 dataset couple images and the results</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td>those applied by using RGB color space model</td>
<td></td>
</tr>
<tr>
<td>4.3</td>
<td>present the 19 dataset couple images and the results</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>those applied by using YCbCr color space model</td>
<td></td>
</tr>
<tr>
<td>4.4</td>
<td>showing values of PSNR, NCC, MSE, AD and Structural Content</td>
<td>80</td>
</tr>
</tbody>
</table>
Rapid development of the technique of sensors, micro-electronics, and communications requires more attention on information fusion. Several situations in image processing require high spatial and high spectral resolution in a single image. For example, the traffic monitoring system, satellite image system, and long range sensor fusion system are all used in the image processing. However, the majority of available equipment is not capable of providing this type of data convincing. The sensor in the surveillance system can only provide the scenery view in a narrow depth for a particular focus, yet the demand of the application of the system requires a clear view with a high depth of the field. Thus, the image fusion provides the possibility of combining different sources of information.

Besides, four applications motivate this thesis includes: (1) Multi-sensor image fusion, (2) Medical image fusion, (3) Surveillance System, and (4) Aerial and Satellite imaging. The particular applications discussed include medical diagnosis, remote sensing, surveillance systems, biometric systems, and image quality assessment.

In all sensor networks, every sensor can observe the environment, produce and transfer data. Visual sensor networks (VSN) is the term used in the literature to refer to a system with a large number of cameras geographically spread at monitoring
points (Patricio et al., 2006). In VSN, sensors are cameras which can record either still images or video sequences. Therefore, the processing of output information is related to image processing and machine vision subjects.

A distinguished feature of visual sensors or cameras is the great amount of generating data. This characteristic requires more local processing resources to deliver only the useful information represented in a conceptualized level (Castanedo et al., 2008). Image fusion is generally defined as the process of combining multiple source images into a smaller set of images, usually a single one that contain a more accurate description of the scene than any of the individual source images. The goal of image fusion, besides reducing the amount of data in network transmissions is to create new images that are more informative and suitable for both visual perception and further computer processing (Drajic et al., 2007). Due to the limited depth of focus on optical lenses, only the objects at a particular depth in the scene are in focus and those in front of or behind the focus plane will be blurred. In VSN we have the opportunity of extending the depth of focus using multiple cameras.

So far, several researches have been focused on image fusions which are performed on the images in the spatial domain (Lewis et al., 2007; Yang et al., 2008; Roux et al., 2008; Zaveri et al., 2009; Arif et al., 2012). The algorithms based on multi-scale decompositions are more popular. The basic idea is to perform a multi-scale transform on each source image and then integrate all of these decomposition coefficients to produce a composite representation. These methods combine the source images by monitoring a quantity that is called the activity level measure. The activity level determines the quality of each source image. The integration can be carried out either by choosing the coefficients with larger activity level or a weighted average of the coefficients. The fused image is finally reconstructed by performing the inverse multi-scale transforms. Examples of this approach include Laplacian, gradient, morphological pyramids, and discrete cosine transform (DCT).
In general, most of the spatial domain image fusion methods are complex and time-consuming, which made it not suitable for real-time applications. In VSN, especially the links between nodes, which are wireless consumed energy for data processing is much less than the consumed energy for communication. Therefore, in most of the sensor network systems, data are compressed before transmission to the other nodes. In VSN, images are compressed in camera nodes and then transmitted to the fusion agent, and afterwards the compressed fused image will be saved or transmitted to an upper node. Hence, when the source images are coded in Joint Photographic Experts Group (JPEG) standard, as a prevalent standard or when the fused image will be saved or transmitted in JPEG format, the fusion methods which are applied in DCT domain will be more efficient.

1.2 Problem Background

The success of fusion images acquired different modalities or an instrument is of great importance in many applications as medical imaging, microscopic imaging, remote sensing, computer vision and robotics. Image fusion can be defined as the process by which several image features are combined together to form a single image. Image fusion can be performed at different levels of the information representation. Four different levels can be distinguished according to (Abidi and Gonzalez, 1992) i.e. signal, pixel, feature and symbolic levels. To date, the results of image fusion from remote sensing and medical imaging are primarily intended for presentation to a human observer for easier and enhanced interpretation. Therefore, the perception of the fused image is of paramount importance when evaluating different fusion schemes.
Some generic requirements can be imposed on the fusion result:

a) The fused image should preserve closely as possible for all relevant information contained in the input images.

b) The fusion process should not introduce any artifacts or inconsistencies, which can distract or mislead the human observer or any subsequent image processing steps Rockinger.

c) In the fused image, irrelevant features and noise should be suppressed to a maximum extent.

When fusion is done at pixel level, the input images are combined without any preprocessing. Pixel level fusion algorithms vary from very simple image averaging to very complex e.g. principal component analysis (PCA), pyramid based image fusion and wavelet transform fusion.

Several approaches to pixel level fusion can be distinguished based on whether the images are fused in the spatial domain or in a transform domain. After the fused image is generated it may be processed further to extract some features of interest. In image fusing, two or more images which are not more suitable for analysis are taken and these are converted into a single image in which noise is eliminated, but retaining the important features of each of the source images is more suitable for exploration. Fusion of color images can be carried in two ways:

a) First convert color images into gray images and apply fusion for the gray images and then convert fused gray image back to color images.

b) Decompose the color image into three components then apply fusion to red components of the input images individually and then fuse green components and then blue (or simultaneously). Afterward, the fusion combines all the three colors to form the final image.
But the problem involved in the process (a) is while converting image first from color to gray before fusion, gray to color after fusion the final image color characteristics may change because of the approximations involved in this process. So, a direct fusion of color images is preferred where the color of the images has the most importance.

1.3 Problem Statement

In recent years, many image fusion techniques and algorithms have been exploited and they have been successfully used in the fusion process. More recently, the development of DCT theory was implemented for DCT multi-scale decomposition to take the place of the pyramid decomposition for image fusion.

Image fusion is a process that combines information from multiple images of the same scene into a single image. In this study, a new image fusion technique is proposed using the Discrete Cosine Transform (DCT). Each of the source images may represent a partial view of the scene, and contain both ‘relevant’ and ‘irrelevant’ data. So the goal of an image fusion algorithm is to combine the source images into a single composite image which contains a more accurate description of the scene than any of the individual sources. Moreover, the fused image got the best possible quality without the introduction of distortion or loss of information.

The Discrete Cosine transform DCT fusion scheme offers several advantages over similar pyramid based fusion schemes when it comes to image fusion:

(a) The DCT transform provides directional information while the pyramid representation doesn’t introduce any spatial orientation in the decomposition process.
In pyramid based image fusion, the fused images often contain blocking effects in the regions where the input images are significantly different. No such artifacts are observed in similar DCT based fusion results.

Images generated by DCT image fusion have better signal-to-noise ratios (PSNR) than images generated by a pyramid image fusion when the same fusion rules are used.

1.3.1 Research Question

This Study will focus on the following questions:

1- How can multi-focus image fusion on color image at pixel level of the input images?
2- How can discrete cosine transform technique (DCT) be proposed as a new multi-focus fusion technique?
3- Is there any difference if a different color space model was used to get better technique for image fusion?
4- How can we evaluate the proposed technique using standard dataset image?

1.4 Aim of the Study

The aim of this research is to implement an Image Fusion Technique for Color Images using Discrete Cosine Transform by different color space model for improving the image fusion from the input images.
1.5 Objectives of the Study

The objectives of this research are:

1. To propose a new fusion technique using discrete cosine transform technique (DCT) based on RGB color space model.
2. To propose a new fusion technique using discrete cosine transform technique (DCT) based on YCbCr color space model.
3. To compare and evaluate the proposed method using standard and referenced dataset.

1.6 Scope of the Study

In order to achieve the objectives stated above, the scope of this study is limited to the following:

1. The suggested technique applied by MATLAB language version on windows 7 milieu.
2. To focus on RGB input multi-focus images.
3. To focus on RGB and YCbCr color space model.
4. The database aimed to be used in this study is a standard RGB picture using in most popular search in the field of image fusion.
1.7 Thesis organization

The dissertation is organized as follows: In the first chapter provide the introduction, the problem background, problem statements, aim of the study, objective of this study and research question to the research. The chapter two also provides a simple overview introduction about the image fusion technique specifically themulti-focus images, basic principles of image fusion and image fusion techniques whilethe chapter three makesprovision for theproposed technique methodology explanation. The chapter four discussed a result of the methodology provided in chapter three. Finally, the overall conclusions of the thesis and recommendations for future works are in chapter five.


Xydeas, C. and Petrovic, V. 2006. Pixel-level image fusion metrics. Imaging Science and Biomedical Engineering. University of Manchester, Oxford Road, Manchester, UK.


