ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

MSc THESIS

Sultan Sevgi TURGUT

INVESTIGATION OF MULTI-FOCUS IMAGE FUSION

DEPARTMENT OF COMPUTER ENGINEERING

ADANA-2020

ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

INVESTIGATION OF MULTI-FOCUS IMAGE FUSION

Sultan Sevgi TURGUT

MSc THESIS

DEPARTMENT OF COMPUTER ENGINEERING

We certify that the thesis titled above was received and approved the award of degree of the Master of Science by the board of jury on

…………………………….	…………………………………….	…………………………………….....
Assoc.Prof.Dr. Mustafa ORAL	Assoc. Prof. Dr. Zekeriya TÜFEKÇİ	Asst. Prof. Dr. Esra SARAÇ EŞSİZ
SUPERVISOR	MEMBER	MEMBER

This MSc thesis is written at the Computer Engineering Department of the Institute of Natural and Applied Sciences of Çukurova University.
Registration Number:

Prof. Dr. Mustafa GÖK
Director
Institute of Natural and Applied Science

ABSTRACT

MSc THESIS

INVESTIGATION OF MULTI-FOCUS IMAGE FUSION

Sultan Sevgi TURGUT

ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF COMPUTER ENGINEERING

Supervisor : Assoc. Prof. Dr. Mustafa ORAL
Year: 2020, Pages: 105
Jury : Assoc. Prof. Dr. Mustafa ORAL
: Assoc. Prof. Dr. Zekeriya TÜFEKÇİ
: Asst. Prof. Dr. Esra SARAÇ EŞSİZ

In this study, a study has been carried out on the multi-focus image fusion that is the sub-branch of the image processing area. The aim of this thesis is to propose a new and effective multi-focus image fusion method that takes advantage of the successful aspects of present methods and improves the missing aspects. The goal of multi-focus image fusion is to provide a completely focused image by distinguishing between focused and unfocused pixels. Based on the literature review, it has been noticed that the importance of edge information in distinguishing pixels is emphasized. It has been concluded that the difference between focused/ non-focused pixels can be revealed more clearly with gradient images rather than the original images. The energy of Gradient (EOG) and standard deviation functions have been used as the focus measurement on gradient images. In order to minimize the error, the majority voting method is applied with different window sizes for the final fusion. The proposed method has been compared with seventeen different new and traditional multi-focus image fusion techniques both visually and objectively. Six different quantitative metrics have been used for objective evaluation. It has been observed that the proposed method is promising according to visual evaluation and 83.3% success has been achieved by being first in five out of six metrics according to objective evaluation.

Keywords: Image processing, multi-focus image fusion, dithering, the energy of gradient, objective metrics

ÖZ

YÜKSEK LİSANS TEZİ

ÇOK ODAKLI GÖRÜNTÜ FÜZYONUNUN İNCELENMESİ

Sultan Sevgi TURGUT

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

Danışman : Doç. Dr. Mustafa ORAL
Yıl: 2020, Sayfa: 105
Jüri : Doç. Dr. Mustafa ORAL
: Doç. Dr. Zekeriya TÜFEKÇİ
: Dr. Öğr. Üyesi. Esra SARAÇ EŞSİZ

Bu çalışmada görüntü işleme alanının alt dalı olan çok odaklı görüntü füzyonu konusu üzerine bir çalışma gerçekleştirilmiştir. Tezin amacı, mevcut çok odaklı görüntü füzyon yöntemlerinin başarılı yönlerinden faydalanan ve eksik yönlerini tamamlayan yeni ve etkili bir Çok Odaklı Görüntü Füzyon yöntemi önermektir. Çok odaklı görüntü füzyonunun hedefi, odaklı ve odak-dışı pikselleri ayırt ederek tamamen odaklı bir görüntü sunabilmektir. Yapılan literatür taramasıyla, pikselleri ayırt etmede kenar bilgisinin önemine dikkat çekildiği fark edilmiştir. Bu yaklaşımla, orijinal kaynak resimlerden gradyan resimler elde edilmiş ve odaklı/odak-dışı pikseller arasındaki farkın ortaya çıkması sağlanmıştır. Energy of Gradient (EOG) ve Standart Sapma fonksiyonları yardımıyla gradyan resimler üzerinde odak ölçümü yapılmıştır. Son olarak oluşan hatayı minimize etmek için, çoğunlukla seçim (Majority Voting) yöntemi farklı pencere boyutları ile uygulanmış ve nihai görüntüye ulaşılmıştır. Önerilen yöntem yeni ve geleneksel yöntemlerden oluşan on yedi farklı çok odaklı görüntü füzyon teknikleriyle hem görsel hem de objektif açılardan karşılaştırılmıştır. Objektif değerlendirme için altı farklı niceliksel metrik kullanılmıştır. Önerilen yöntemin görsel değerlendirmeye göre umut vaat edici olduğu, objektif değerlendirmeye göre altı metrikten beşinde birinci olarak %83.3 oranında başarı sağladığı gözlenmiştir.

Anahtar kelimeler: Görüntü işleme, çok odaklı görüntü füzyonu, objektif metrikler

EXTENDED ABSTRACT

Image fusion is an important branch of image processing. The fusion field has four different types; multi-focus(mFF), multi-view (mVF), multi-modal(mMF) and multi-temporal (mTF). mFF fuses different views of the same scene while mMF fuses images from different modalities to reveal principal information. mTF allows observing alternation of a scene over time. mFF fuses images with different focal lengths to produce a fully focused image.

In this thesis, multi-focus image fusion is examined. Due to the limited depth of field feature of the optical lenses, it allows focusing on only one region at a time. Therefore, it is sometimes impossible to get fully focused images. Many methods have been developed to solve this problem.

The aim of this thesis is to examine the present methods in the literature in detail and propose a new method based on inferences. The proposed method consists of three steps; preprocessing, initial fusion, decision fusion.

When the literature is examined, it is stated that details are not lost in focused areas while it is difficult to capture the details in out of focus areas. A powerful way to distinguish details is to view edge information. For this reason, preprocesses have been applied to enrich the information of the original images and to gain the edge information. Local Histogram Equalization has been applied to make image details more dominant. In order to obtain edge information, the Halftoning Method has been used, unlike traditional edge detection methods. As a result of preprocessing, gradient images of source images are obtained.

In the Initial Fusion step, Energy of Gradient and Standard Deviation functions are used as the focus measurement method on gradient images. Images are divided into blocks and blocks that have higher energy are selected. When the energy function is insufficient, the standard deviation values of the respective blocks are compared and the one with high-values are selected.

In the Decision Fusion step, with the ensemble learning approach, majority voting is performed according to different window sizes and it is aimed to minimize the errors that occurred in the previous step.

In order to prove the success of the proposed method, visual and objective comparisons have been made with 17 different methods. In the visual comparison, the fused images of each method and the error images have been examined and interpreted. In the objective evaluation, six different metrics (Q_{NMI}, Q_{NCIE}, Q_{Y}, Q_{CB}, $Q_{AB/F}$, Q_{PWW}) measuring different characteristics features of images have been implemented. Thus, a quantitative comparison could be performed. According to these evaluations, it has been proved that the proposed method produces visually successful fused images and objectively outperforms the other methods in five out of six metrics.

As a result, it has been believed that developing a new method is necessary due to some short comings of present fusion methods such as difficulty in capturing the details, inability to work correctly in moving areas, and long-running time or requiring huge datasets. The proposed method has proven to stand out against state-of-art and traditional methods.

GENİŞLETİLMİŞ ÖZET

Görüntü füzyonu, görüntü işlemenin önemli bir dalıdır. Füzyon alanının dört farklı türü vardır; çoklu-görünüm görüntü füzyonu, çoklu-modalite görüntü füzyonu, çoklu-zaman görüntü füzyonu ve çok-odaklı görüntü füzyonu. Çoklu-görünüm görüntü füzyon teknikleri aynı sahnenin farklı görünümlerini birleştirirken, çoklu-modalite görüntü füzyonu temel bilgileri ortaya çıkarmak için farklı modaliteler ile elde edilen görüntüleri birleştirir. Çoklu-zaman görüntü füzyonu bir sahnenin zaman içinde değişimini gözlemlemeye yarar. Çok-odaklı görüntü füzyonu ise, tamamen odaklanmış bir görüntü oluşturmak için farklı odak uzunluklarına sahip görüntüleri birleştirir.

Bu tezde çok-odaklı görüntü füzyonu incelenmiştir. Optik lenslerin sınırlı alan derinliği özelliği nedeniyle, aynı anda yalnızca bir uzaklığa odaklanabilirler. Bu nedenle, tamamen net görüntüler elde etmek imkansızdır. Bu sorunu çözmek için birçok yöntem geliştirilmiştir.

Bu tezin amacı literatürdeki mevcut yöntemleri detaylı olarak incelemek ve çıkarımlara dayalı yeni bir yöntem sunmaktır. Önerilen yöntem üç adımdan oluşur; ön işleme, başlangıç füzyonu, karar füzyonu.

Literatür çalışmaları incelendiğinde, odaklı resim bölgelerinde detayların kaybolmadığı, odaklı olmayan bölgelerde ise detayların yakalanmasının zor olduğu belirtilmektedir. Ayrıntıları elde etmenin güçlü bir yolu, kenar bilgisinden yararlanmaktır. Bu nedenle, orijinal görüntülerdeki detayları daha belirgin hale getirmek ve kenar bilgisini elde etmek için ön işlemler uygulanmıştır. Görüntü ayrıntılarını daha fazla ortaya çıkarmak için Yerel Histogram Eşitlemesi (Local Histogram Equalization) uygulanmıştır. Kenar bilgisi elde etmek için, geleneksel kenar algılama yöntemlerinden farklı olarak Halftoning-Inverse Halftoning Dönüşüm Yöntemleri kullanılmıştır. Ön işleme adımının bir sonucu olarak, kaynak görüntülerin gradyan görüntüleri elde edilir.

Başlangıç Füzyonu adımında, net resim piksellerinin seçilmesi için, gradyan görüntüler üzerinde uygulanacak odak ölçüm yöntemlerine karar verilmelidir. Odak ölçün yöntemi olarak Gradyan Enerjisi (Energy of Gradient) ve Standart Sapma fonksiyonları kullanılmıştır. Görüntüler blok tabanlı işlenir ve yüksek enerjiye sahip piksel bloklarının seçilmesi sağlanır. Bloklar arasında seçim yapmak için enerji fonksiyonu yetersiz kaldığında, ilgili blokların standart sapma değerleri karşılaştırılır ve yüksek satandart sapma değerine sahip olan bloklar seçilerek başlangıç füzyon işlemi tamamlanır.

Karar Füzyonu adımında, Topluluk Öğrenimi (Ensemble Learning) yaklaşımıyla, Çoğunluk Oyu (Majority Voting) yöntemi farklı pencere boyutları kullanılmasıyla gerçekleştirilmiştir ve başlangıç füzyonu adımında meydana gelen hataların en aza indirgenmesi hedeflenmiştir.

Önerilen yöntemin başarısını kanıtlamak için on yedi farklı yöntemle görsel ve objektif karşılaştırmalar yapılmıştır. Görsel karşılaştırmalarda, her yöntemin sonuç görüntüleri ve hata görüntüleri incelenmiş ve yorumlanmıştır. Objektif değerlendirmede, görüntülerin farklı karakteristik özelliklerini ölçen altı farklı metrik (QNMI, QNCIE, QY, QCB, QAB / F, QPWW) kullanılmıştır. Böylece, nicel bir karşılaştırma yapılabilmiştir. Bu değerlendirmelere göre, önerilen yöntemin diğer yöntemlere kıyasla, görsel olarak daha başarılı sonuç görüntüler ürettiği ve altı metrik arasından beşinde objektif olarak en iyi performansı gösterdiği kanıtlanmıştır.

Sonuç olarak, mevcut füzyon yöntemlerinin resim ayrıntılarını yeterince yakalayamaması, hareketli resim bölgelerinde düzgün çalışamaması, uzun çalışma sürelerine sahip olmaları veya büyük veri kümeleri gerektirmeleri gibi çeşitli eksiklikleri nedeniyle yeni bir yöntem geliştirmenin gerekli olduğuna kanaat getirilmiştir. Önerilen yöntemin yeni ve geleneksel yöntemlere karşı öne çıktığı kanıtlanmıştır.

ACKNOWLEDGMENT

Firstly, I would like to express my gratitude to my supervisor Assoc. Prof. Dr. Mustafa ORAL, for his unwavering support, guidance, and encouragement.

I would also like to thank Selma Ayşe ÖZEL who has supported me throughout my Master's degree.

I would like to thank members of Master thesis jury; Assoc. Prof. Dr. Zekeriya TÜFEKÇİ, Assist. Prof. Dr.Esra SARAÇ EŞSİZ, for their suggestions.

Finally, I would like to thank my family for their endless supports and beliefs.

CONTENTS **PAGE**

LIST OF TABLES **PAGE**

LIST OF FIGURES **PAGE**

ABBREVIATIONS

AB/F	: Xyedas-Petrovic Metric
ANN	: Artificial Neural Network
CBF	: Cross Bileteral Filter
CDF	: Cumulative Density Function
CNN	: Convolutional Neural Network
CSR	: Convolutional Sparse Representation
DCT	: Discrete Cosine Transform
DWT	: Discrete Wavelet Transform
ELM	: Extreme Learning Machine
EOG	: Energy of Gradient
GF	: Guided Filtering
GHE	: Global Histogram Equalization
HVS	: Human Vision System
IF	: Image Fusion
LHE	: Local Histogram Equalization
LP-PCNN	: Laplacian Pyramid Transformation and Pulse Coupled Neural Network
Min/Max	: Minimum/Maximum
mFF	: multi-Focus Fusion
MFIF	: Multi Focus Image Fusion
mMF	: multi-Modal Fusion
MSSF	: Multi-Scale Image Decomposition And Saliency Detection
mTF	: multi-Temporal Fusion
mVF	: multi-View Fusion
MWGF	: Multi-Scale Weighted Gradient-Based
NCIE	: Nonlinear Correlation Information Entropy Metric
NMI	: Normalized Mutual Information

NSCT	: Nonsubsampled Contourlet Transform
PCA	: Principal Component Analysis
PNN	: Probabilistic Neural Network
RBFN	: Radial Basis Function Network
Std	: Standard Deviation
UC	: Unique Color

1. INTRODUCTION

Image fusion (IF) is a branch of image processing. It is a process of combining two or more images to generate new features or enhance the available information. Image Fusion is divided into four categories according to source image types; multi-view fusion (mVF), multi-modal fusion (mMF), multi-temporal fusion (mTF), multi-focus fusion (mFF) (Bedi and Khandelwal, 2013).

mVF merges multiple images which are instances of the same scene, but they are taken from a different point of view. These images that have different views are combined to improve the high resolution and quality. mMF aims to reveal principal information. Source images occur of different modalities of the same scene such like images that are taken from different devices. This method is generally used for medical purposes. The mTF provides to view alternation along the time. Generally, it is used for the detection of the world and environmental changes (Yuhendra and Sumantyo, 2017). mFF fuses different images that have the same scene but having different parts of focus and these images combination has a much clear view than others. A completely clear image is quite useful for almost all applications (Bavachan and Krishnan, 2014). In mTF, images have the same scene, but they are acquired at different times. This study concentrates on the mFF.

The depth of field property of the optical lens causes the view to have different depth of field because it can focus on only one region at a time. There are three main factors to control the depth of field; the aperture, the distance to the subject, and the focal length of the lens. As the aperture increases, the depth of field decreases, that is, the depth of field and aperture form a reciprocal proportion. When the distance to the subject increases, likewise the depth of field increases. The final factor is the focal distance of the lens that is the distance between the lens and image sensor, when the focal length decreases, the depth of field increases. In this way, images with different focal points are created. These generated images are fused by using several algorithms. These algorithms aim to select pixels from the images to obtain a fused image whose each pixel is as well focused as possible (Hariharan and 2011).

Image fusion has three different levels; pixel level, feature level, and decision level. Pixel level is basic-level fusion, the processing is performed directly on the pixels. Comparing to the other levels, this level is the simplest in terms of the operation of the algorithm. Feature level is an intermediate level of fusion, this approach extracts features such as edges, segments or direction from source images and uses these features in the fusion procedure. The aim of using features at this level is to achieve comparable features, and thus different source images can be fused at a similar point. Decision level is a high level of fusion, source images are identified processed to understanding images by several algorithms and outputs are combined with decision rules. Fusion Algorithms are implemented with these levels (Kolekar and Shelkikar, 2015).

A Common approach to categorize image fusion algorithms can be defined in two domains; spatial domain and transformation domain. Spatial domain image fusion methods use the most straightforward approaches that are based on the source image's pixel values. These algorithms can be applied to either single-pixel or a block of pixels. All methods in the spatial domain perform operations directly on the pixel, but the source image can also be divided into blocks. Pixel-based image fusion has performed a pixel-wise operation, whereas block-based image fusion is performed an operation to a group of pixels at a time. The pixel-based approach is more sensitive and has low fault tolerance. In the block-based approach, block size selection must be done carefully as there is no fixed block size for each image. Block size effects the creation of fused images extortionately (Siddiqui et al., 2011). The spatial domain does not require transformation, it is more tolerant of registration errors and less affected by noise. However, the spatial domain fusion causes spectral distortion and spatial degradation in the fused image. This domain includes algorithms such as Averaging, Minimum/Maximum Selection, Principal Component Analysis, Unique Color, and Standard Deviation, etc. (Balachander and Dhanasekaran, 2016). In the frequency domain, before the image fusion, the source images are transformed into a different scale domain. After all the operations are done, the fused image is obtained by applying the inverse transformation process. Algorithms for this domain are IF based on Pyramid Based Algorithms, Discrete Cosine Transform, Discrete Wavelet Transform, Nonsubsampled Contourlet Transform, Convolutional Sparse Representation, Cross Bilateral Filter, Guided Filtering, Multi-Scale Image Decomposition, And Saliency Detection, Multi-Scale Weighted Gradient-Based, etc. Transform domain-based algorithms may be time-consuming and registration is a crucial operation because even small changes can be degraded the fused image (Yang et al.,2008). In recent years deep learning has been used for image fusion to overcome fusion techniques' problems like spatial degradation and registration. Several neural networks have been used in the literature for this approach, Probabilistic Neural Network and Radial Basis Function Network (Li et al., 2002), Deep Convolutional Neural Network (Liu et al., 2017), Extreme Learning Machine (Yang et al., 2017), etc.

Image fusion can be the solution for countless purposes and can be used in different areas. Images that have a high-quality vision and without losing information are an important necessity in almost every system. Thus, image fusion is a substantial procedure.

In this study, our objective is to examine present mFF techniques and carry out comparative work to demonstrate the success rates of the methods. It is, also, aimed to be able to propose a new method with a new perspective taking advantage of the existing methods. Additionally, we attempt to improve this method by considering both the implementation and performance aspects. We try to develop a method by examining not only the advantages but also the disadvantages of the present methods. The proposed method should be shift-invariant, resistant to misregistered images, and be able to capture details in focus. It should not need a large dataset and training process. This is one of the reasons we develop our method in the spatial domain.

Our method consists of three steps; preliminary operations, initial fusion, and decision fusion. Preliminary operations; these are the processes to highlight important details by eliminating the noise in the image. To highlight these details, two features were selected in this study; contrast and edge information. As a result of pre-processing, a gradient image is obtained by processing the highlighted features. In the initial fusion step, a focus measurement method on the gradient image is applied with a sliding window. The Energy of Gradient (EOG) function is activated here. When the EOG function is insufficient, a second focus measurement function is enabled. It has been decided that the second focus measurement function is the standard deviation. The initial fused image is obtained by fusing blocks with high measurement values of source images. In order to minimize the error which occurs in the initial fusion step, one more chance is given to the fusion system and the decision fusion step is applied. A decision fusion mechanism is developed with machine learning approaches. Just as in ensemble learning with a majority voting system, voting is performed by implementing different block sizes. As a result, it is hoped that a fully focused image is obtained.

This thesis is organized as follows; Chapter 2 examines present traditional and state-of-art methods in the literature review. Chapter 3 details existing mFF methods, presents the proposed method and, informs about evaluation metrics and experimental usages. Chapter 4 includes results; visual and objective evaluations and discussions. Chapter 5 concludes the thesis and recommendations for future studies.

2. LITERATURE REVIEW

Multi-Focus Image Fusion Algorithms can be divided into two groups according to the domains used for fusion process; spatial domain and transformation domain methods. The spatial domain does not require transformation, it is more tolerant of registration errors and less affected by noise. However, the spatial domain fusion causes spectral distortion and spatial degradation in the fused image. (Balachander and Dhanasekaran, 2016). In the transformation domain, before the image fusion, the source images are transformed into a different scale domain. After all the operations are done, the fused image is obtained by applying the inverse transformation process. Transform domain-based algorithms may be time-consuming and registration is a crucial operation because even small changes can be degraded the fused image (Yang et al., 2008).

Morris and Rajesh have proposed three well-known spatial domain methods; Averaging, Min/Max selection and Principal Component Analysis (PCA(Morris, Rajesh, 2015)). Averaging and Min/Max selection methods have basic implementation and understanding. While in Averaging method, averaging pixels of source images is used to create fused image, in Min/Max methods, pixels with minimum or maximum intensities are selected to achieve same goal. The main use of PCA is to reduce high dimensional data with correlated variables into a low dimension with more variant structure without loss of information.

Unique Color (Oral and Kartal, 2016) and Standard Deviation (Jassim, 2013) Methods are good examples of successful spatial algorithms. These two methods have a similar perspective. Both methods are based on the principle of diversity as a method of detecting focused pixels.

Naidu and Gawari et al. have proposed important transform domain methods: DCT and DWT (Naidu, 2012; Gawari and Lalitha, 2014). These methods transform data from spatial domain to frequency domain. In image processing, obtaining pixels' frequencies knowledge is useful where distinguishing details and outlines in the images are important. DCT uses only cosine component for decomposition so it has a cost-effective approach. On the other hand, DWT can diminish spatial distortion.

Another example of traditional transform domain algorithms is pyramid-based ones; Laplacian Pyramid (Burt, 1981), Morphological Pyramid (Ramac et.al., 1998),Ratio Pyramid (Toet, 1989). These methods try to achieve details of images by means of lowpass and bandpass filtering. Despite their simple structure, they are considered as promising algorithms.

Zhang et al. have developed Nonsubsampled Contourlet Transform as a different type of transform domain method (Zhang and Guo, 2009). They have focused two properties on their algorithms; directionality and anisotropy. NSCT is a shift-invariance method and so it can be robust for misregistered images.

He et al. have developed a new approach to transformation methods with the Guided Filter Method (He et al., 2012). They have made multi-scale transformations by filtering instead of transform operations. Guided Filter serves to reveal gradient information and eliminate noise in the image.

Liu et al. contribute to the literature with the Convolutional Sparse Representation that performs transformation by filtering (Liu et al., 2016). This method has been developed in order to obtain details more successfully and find solutions to misregistered images.

Another example of filtering methods is the Cross Bilateral Filtering that is proposed by Kumar (Kumar, 2012). The difference of CBF is that it uses both similarities in color space and proximity in geometric space. This procedure processes the details and outlines of an image separately.

Bavirisetti and Dhuli aim to detect saliency information and thus get rid of noises by using the average filter (Bavirisetti and Dhuli, 2016). It provides performance gain with its simple filtering structure.

One of the outstanding studies by processing gradient information has been developed by Zhou and Wang (Zhou and Wang, 2014). Multi-Scale Weighted

Gradient-Based Method concentrates on anisotropic blur and misregistration problems. to the algorithm obtains the borders by processing the structural gradient information. It can be considered one of the most successful methods.

Liu et al. are researchers who developed the fusion algorithm benefit from deep learning methods (Liu et al.., 2017). They aim to perform both feature extraction and classification operations together using Convolutional Neural Network. They have inspired other studies in the literature with their successful results. However, the requirement for large data sets and the training process are the disadvantages of their method.

Another example of deep learning model, Pulse Coupled Neural Network (Eckhorn et al., 1990) has attracted a lot of attention over time and has been developed by various researchers; Dual-Channel PCNN (Wang and Ma, 2007), Multi-Channel PCNN (Wang and Ma, 2008). The PCNN model based fusion method (Wang et al.,2010) is combined with existing fusion methods to achieve successful results. However, researchers could not avoid the performance burden of combining more than one method.

3. MATERIALS AND METHODS

3.1. Image Fusion Algorithms

3.1.1. Averaging Method

This simple and early method (Morris and Rajesh, 2015) creates a fused image from source images with averaging of their pixels. Averaging operation is made by taking pixels in the order. Implementation and understanding of this method are easy. But the disadvantage is that although the effect of pixels that have a high depth of field is higher, the fused image inevitably includes noises. Equation (1) shows generation fused image's pixel that is in x and y coordinates, I is the source image and n is the number of source images, F is the fused image.

$$F(x,y) = \frac{\left(\sum_{i=1}^{n} I_i(x,y)\right)}{n} \tag{3.1}$$

3.1.2. Maximum/Minimum Selection Method

Minimum and Maximum selection methods (Morris and Rajesh, 2015) are examples of old but cost-effective fusion algorithms. The fused image is created by selecting the maximum or minimum intensity of the pixel from among source images' pixels. In this method, the resulting image that is composed of maximum pixels will be a lighter color than it should be and vice versa, the resulting image formed by the minimum pixels becomes darker. Equation (2) demonstrates the minimum selection while Equation (3) demonstrates the maximum selection, I is the source image, x and y coordinates of the image pixel, F is the fused image.

$$F(x,y) = \begin{cases} I_i, & l_i(x,y) < l_j(x,y) \\ l_j, & otherwise \end{cases} \tag{3.2}$$

$$F(x,y) = \begin{cases} I_i, & l_i(x,y) > l_j(x,y) \\ l_j, & otherwise \end{cases} \tag{3.3}$$

3.1.3. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) (Morris and Rajesh, 2015; Masthanaiah and Kumar, 2014) was established in 1901 by Karl Pearson and developed by Harold Hotelling in the 1930s. PCA is a spatial domain and a pixel-level algorithm. It is one of the most commonly used methods for image fusion. It is also used as a downscaling method and application requiring data relationships.

PCA is a statistical and vector space transform method. It transforms image space into eigenspace to reduce dimension. The purpose of this method is to keep the descriptive properties by decreasing the dependency structure and correlation, additionally represent the original image in a smaller size without loss of information. Thereby redundant information is diminished, and important variables are revealed which is called principal components.

Fusion operation is as follows; first, matrices of the source images converted to column vectors. Correlation analysis is applied to column vectors and a covariance matrix which contains the variance elements along the diagonal and the covariance elements in the remaining part are obtained. Then eigenvalues and eigenvectors are extracted from the covariance matrix. Principal components are normalized according to the bigger eigenvalue. Finally, principal components are multiplied with source images' pixels. These multiplication results for each source image are summed and the fused image is achieved. Figure 3.1 illustrates the PCA steps. In Figure 3.1, F is the fused image, I1 and I2 are source images, P1 is the principal component of I1 and P2 is the principal component of I2.

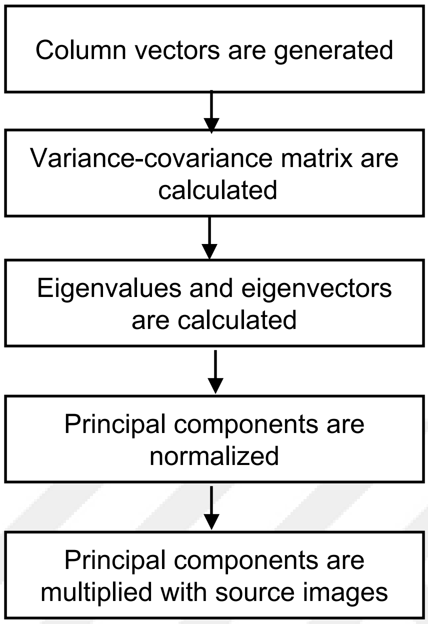


Figure 3.1. Principal Component Analysis Steps.

3.1.4. Unique Color

The Unique Color method which is introduced by Oral and Kartal is a pixel-based algorithm with a very simple structure (Oral and Kartal, 2016). It is based on an approach that image with clear view contains a wide range of colors, whereas blurred image has few color diversity and these colors are close to each other. The color diversity of source images is observed and used to decide whether the view is clear or not. While the fused image is being created, the source image which has more unique colors is selected.

This method is performed on the image blocks. Firstly, the block size is determined. The block size is selected 64x64 in their study. Then the number of colors in the corresponding blocks of each source image is calculated. The reason for the big block size is to prevent the blocking effect from occurring in the fused image. Blocks that have more color variation are selected and the fused image consists of these selected blocks.

Although it is a simple algorithm, it is a promising method in terms of both success and time complexity. On the other hand, it has a disadvantage, blurred parts of images that have sharp edges can have more colors than focused parts due to the semitones and defocusing. Thus, the fused image can have blurred parts because of this situation. This method can be illustrated as follows;

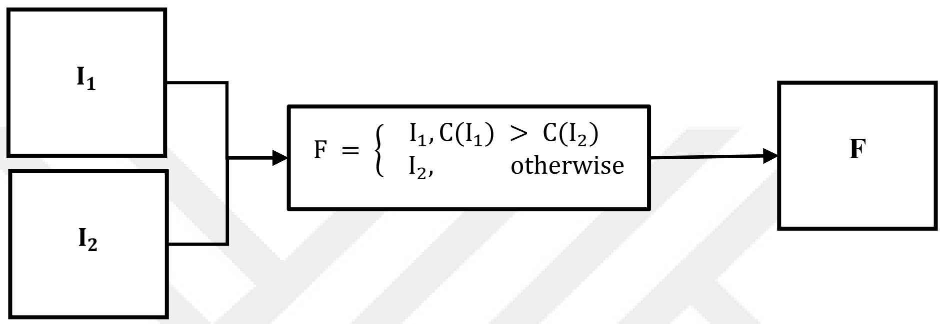


Figure 3.2. Unique Color procedure.

Where I_1, I_2, and F are source images and fused image respectively. C(I) is the number of unique color image blocks.

3.1.5. Standard Deviation Method

The Standard Deviation Method (Jassim, 2013) is a measure of the average dispersion from the mean value for each value within a dataset. When these values are similar, the standard deviation is smaller. On the contrary, if these values are different from each other, the standard deviation is larger. This approach can be used for image fusion by means of pixels variations. Standard deviation on the focused part of an image is bigger than blurred part of the image because focused parts have details and different colors but blurred parts generally viewed close to each other. In this method, source images are divided into blocks and the block that has high standard deviation among source images is selected and placed into the fused image standard deviation formula is given in Equation (4), N represents the number of processing elements, x_i is the i^{th} element and μ is the mean of elements.

Figure 3.3 shows block selection with respect to this method. I_1, I_2, and F are source and fused images.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(xi-\mu)^2} \tag{3.4}$$

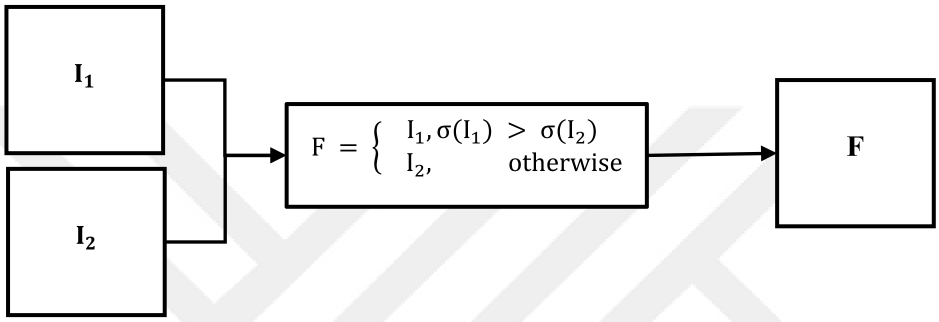


Figure 3.3. Standard Deviation procedure.

3.1.6. Discrete Cosine Transform (DCT)

Discrete Cosine Transform (Naidu, 2012; Mishra and Palkar, 2015) method provides the transformation from the spatial area to the frequency area. In image processing, in order to distinguish details and outlines in the image, obtaining pixels' frequency knowledge is useful and DCT is a proper procedure for working with frequencies. It uses only the cosine component for the transformation, so it provides a fast and non-complex solution. Inverse Discrete Cosine Transform (IDCT) turns the frequencies back to their original form.

In the image fusion, DCT is applied to identical blocks of source images and DCT coefficients are achieved. These coefficients are processed with different techniques to form the coefficients of the fused image. At last, the fused image is obtained by applying the inverse DCT to coefficients. Naidu proposed six different DCT fusion techniques; DCTav, DCTma, DCTah, DCTch, DCTe, and DCTcm. He has demonstrated that DCTe and DCTma are the best of these methods. Another important issue is the block size, 8x8 is known as the most appropriate size from previous studies. General cosine transform formula is given in Equation (5) and inverse DCT formula is given in Equation (6). $F(k_1, k_2)$ is the 2D DCT of MxN source image, f(x,y) is pixel values of the source image, $0 \leq k_1 \geq M-1$ and $0 \leq k_2 \geq N-1$ values in the frequency domain are equivalent $0 \leq x \geq M-1$ and $0 \leq y \geq M-1$ index of image pixels.

$$F(k_1, k_2) = \alpha(k_1)\alpha(k_2) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \cos\left(\frac{\pi(2x+1)k_1}{2M}\right) \cos\left(\frac{\pi(2y+1)k_2}{2N}\right),$$

Where $\alpha(k_1) = \begin{cases} \frac{1}{\sqrt{M}}, & k_1 = 0 \\ \sqrt{\frac{2}{M}}, & 1 \leq k_1 \leq M-1 \end{cases}$

and $\alpha(k_2) = \begin{cases} \frac{1}{\sqrt{N}}, & k_2 = 0 \\ \sqrt{\frac{2}{N}}, & 1 \leq k_2 \leq N-1 \end{cases}$ (3.5)

$$f(x,y) = \sum_{k_1=0}^{M-1} \sum_{k_2=0}^{N-1} \alpha(k_1)\alpha(k_2)\, F(k_1, k_2)\, cos\left(\frac{\pi(2x+1)k_1}{2M}\right) cos\left(\frac{\pi(2y+1)k_2}{2N}\right) \quad (3.6)$$

DC coefficient is the D(0,0) value in the block, and it is calculated by using all values in the block which is AC coefficients, and AC coefficients constitute the rest all coefficients. All these coefficients that are obtained from different source images are merged with a predefined method. Several methods are found; DCTav, DCTmx, DCTah, DCTch, DCTe, and DCTcm etc. The fused image is obtained by averaging the frequency components of the source image in the DCTav method. In DCTma, the largest magnitude coefficients are selected from the source images in order to form the fused image. For DCTah and DCTch, take average both DC coefficients and lowest AC components, the other half of AC coefficients are selected according to the absolute maximum value in DCTah, the greatness of contrast value in DCTch. In DCTcm the AC coefficients are selected based on a contrast measure and DC components are averaged. Finally, in DCTe, AC components corresponding to the frequency band having the largest energy level are chosen and DC components are averaged. Equations (7)-(12) gives DCT operations, X_f is the fused coefficient, X_1 and X_2 are source images' coefficients.

DCT av

$$X_f(k_1,k_2) = 0.5X_1(k_1,k_2) + 0.5X_2(k_1,k_2) \text{ where } k_1,k_2: 0,1,\dots,N-1 \quad (3.7)$$

DCT mx

$$X_f(0,0) = 0.5X_1(0,0) + 0.5X_2(0,0)$$

$$X_f(k_1,k_2) = \begin{cases} X_1(k_1,k_2), & |X_1(k_1,k_2)| \geq |X_2(k_1,k_2)| \\ X_2(k_1,k_2), & |X_1(k_1,k_2)| < |X_2(k_1,k_2)| \end{cases} \text{ where } k_1,k_2: 1,\dots,N-1 \quad (2.8)$$

DCT e

$$X_f(0,0) = 0.5X_1(0,0) + 0.5X_2(0,0)$$

$$X_f(k_1,k_2) = \begin{cases} X_1(k_1,k_2), & E_{j1} \geq E_{j2} \\ X_2(k_1,k_2), & E_{j1} < E_{j2} \end{cases} \text{ where } \substack{k_1,k_2:0,1,\dots,N-1 \\ j=k_1+k_2} \quad (2.9)$$

DCT cm

$$X_f(0,0) = 0.5X_1(0,0) + 0.5X_2(0,0) \quad (2.10)$$

$$X_f(k_1,k_2) = \begin{cases} X_1(k_1,k_2), & C_1(k_1,k_2) \geq C_2(k_1,k_2) \\ X_2(k_1,k_2), & C_1(k_1,k_2) < C_2(k_1,k_2) \end{cases} \text{ where } k_1,k_2: 1,\dots,N-1$$

DCT ah

$$X_f(k_1,k_2) = 0.5X_1(k_1,k_2) + 0.5X_2(k_1,k_2)$$

$$\text{where } k_1,k_2: 0,1,\dots,0.5N-1$$

$$X_f(k_1,k_2) = \begin{cases} X_1(k_1,k_2), & |X_1(k_1,k_2)| \geq |X_2(k_1,k_2)| \\ X_2(k_1,k_2), & |X_1(k_1,k_2)| < |X_2(k_1,k_2)| \end{cases}$$

$$\text{where } k_1,k_2: 0.5N, 0.5+1, 0.5N+2,\dots,N-1 \quad (2.11)$$

DCT ch

$X_f(k_1,k_2) = 0.5X_1(k_1,k_2) + 0.5X_2(k_1,k_2)$ where $k_1, k_2: 0,1,\dots,0.5N-1$

$$X_f(k_1,k_2) = \begin{cases} X_1(k_1,k_2), & C_1(k_1,k_2) \geq C_2(k_1,k_2) \\ X_2(k_1,k_2), & C_1(k_1,k_2) < C_2(k_1,k_2) \end{cases}$$ where

$k_1, k_2: 0.5N, 0.5+1, 0.5N+2, \dots, N-1$ (2.12)

3.1.7. Discrete Wavelet Transform

Another widely used transformation method in image fusion applications is the Discrete Wavelet Transform (Gawari and Lalitha, 2014; Yang, 2011). It operates with the frequency information of the pixels as in DCT. But DWT can reduce spatial distortion.

Firstly, the source image's signals are decomposed to obtain wavelet coefficients. A fusion method is implemented, and the fused coefficients are generated. By applying IDWT, the fused image is obtained.

DWT decomposes an image into four frequency bands; HH, HL, LH, LL. The low-frequency components of an image give the outline, while the high-frequency components give details of the image. To separate these components, the decomposition step is implemented with high and low-pass filters and down-sampling operations. Several fusion methods can be used for the creation of fused wavelet coefficients. One of them is averaging, source images are decomposed, and their coefficients are obtained. The source images are processed according to the same bands, for example, an average of the first image's HH band and second image's HH band generates the fused image's HH band. Another fusion method is maximum selection, the source images are compared according to the same bands, such as the first image's HH band is compared to the second image's HH band. The element with the maximum frequency band is selected.

DWT has some disadvantages like shift-invariance and lacking directional selectivity. Figure 3.4 and Figure 3.5 show these operations, I_1, I_2, and F represent source images and fused images, respectively. ↓2 symbol means that down-sampling by 2.

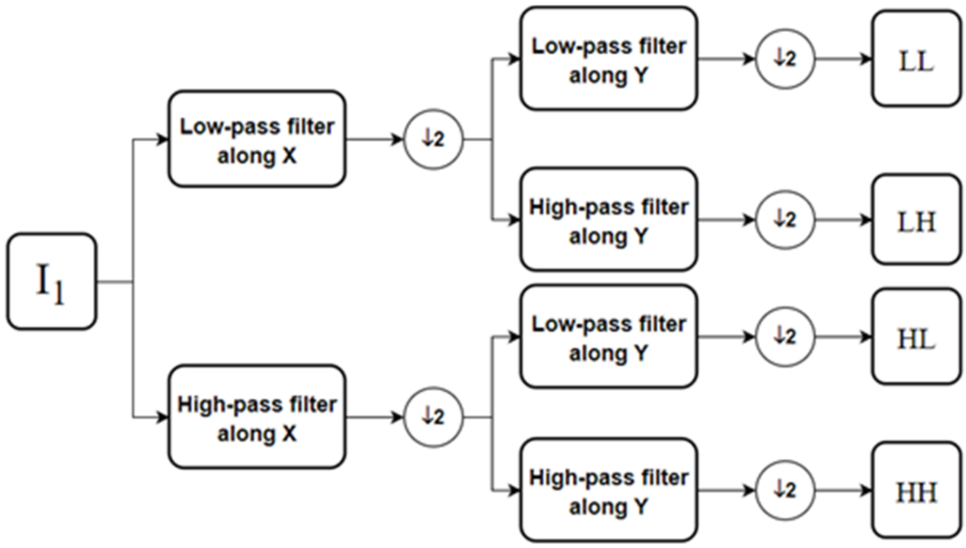


Figure 3.4. Illustration of Discrete Wavelet Transform decomposition.

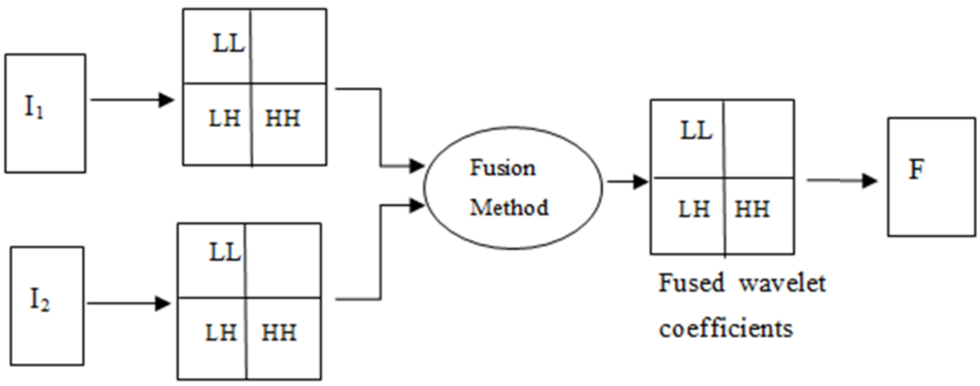


Figure 3.5. DWT fusion steps.

3.1.8. Dual-Tree Complex Wavelet Transform Based Fusion Method

DWT has some disadvantages, such as lack of directional selectivity and shift-invariance. Dual Tree Complex Wavelet Transform (DTCWT) (Kingsbury, 1998) which is advanced DWT is developed to deal with DWT's disadvantages.

DTCWT applies the decomposition process to the signal at the same time twice. This is provided by using two parallel trees. One of the trees decomposes a real sample of the signal while the other decomposes the imaginary sample of the signal. It achieves shift and directional invariance. Other operations are the same as the DWT. The only disadvantage of this method is computational cost. Figure 3.6 illustrates the DTCWT process, I_1 is an image, ↓2 is the down sampling operation.

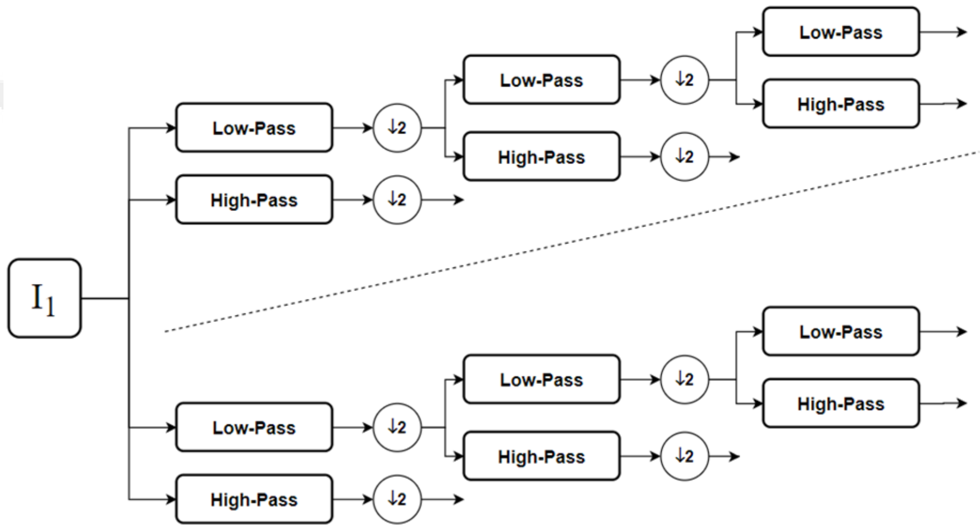


Figure 3.6. Dual Tree Complex Wavelet Transform decomposition process.

3.1.9. Pyramid Based Fusion Methods

Pyramid algorithms (Sravani and Chandra, 2015) are performed in three steps by generating pyramids of an input image with applying Pyramid Transform, combining created pyramids together and creating the fused image using the Inverse Pyramid Transform. Pyramid transform is constituted of an image with lowpass or bandpass filtering. Lowpass filter alleviates high-frequency information, bandpass filter alleviates too low and too high-frequency information. This step is applied to a predefined number of times. Size and resolution are reduced before each repetition. When pyramid transform operation is finished, created pyramids are merged. The inverse pyramid transform is applied with the transpose of the filter that is used in the pyramid transform step. Several pyramid-based methods exist; Laplacian Pyramid, Morphological Pyramid, Ratio Pyramid, etc.

3.1.9.1. Laplacian Pyramid Based Fusion

Laplacian Pyramid (Xiaopeng et al.,2015; Wang and Chang, 2011) is generated by using the Gaussian Pyramid. The Gaussian Pyramid consists of different levels of an image. The base level is the original image and the other levels are created by using lowpass filter and doing down sampling that means decreasing the size. The Gaussian function can be used to form the resulting consecutive levels of the Gaussian Pyramid. Levels of the Laplacian Pyramid are obtained by taking the difference between levels of the Gaussian Pyramid. A level is subtracted from its higher level and generated the band-pass filtered image. But each level has a different size so the higher level which has a smaller size of the relevant level must be expanded to the same size as its. The last level of the Gaussian Pyramid hasn't got a subsequent level thus it is taken directly to the last level of the Laplacian Pyramid. After forming the Laplacian Pyramids of source images, each level is fused with different techniques such as maximum, minimum selection or averaging methods to generation fused pyramid. Then Inverse Laplacian Pyramid Transform is implemented to the fused pyramid and the fused image is obtained. Equation (13) contains the Gaussian Function and Gaussian Pyramid transactions. Equation (14) gives the Laplacian Pyramid Procedure. Figure 3.7 shows the creation of the Laplacian Pyramid from the Gaussian Pyramid. In the equations, i and j are coordination of pixels, g and L represents the Gaussian and Laplacian Pyramids respectively. Levels are specified with l.

$$w(i,j) = \frac{1}{(\sqrt{2\pi}\,\sigma)^2} e^{-\frac{i^2+j^2}{2\sigma^2}}$$

$$g_l(x,y) = \sum_x \sum_y w(i,j) g_{l-1}(2x+i, 2y+j) \tag{3.13}$$

$$L_l = \begin{cases} g_l - expand(g_{l+1}) \ , & l < L \\ expand(g_l) \ , & l = L \end{cases} \tag{3.14}$$

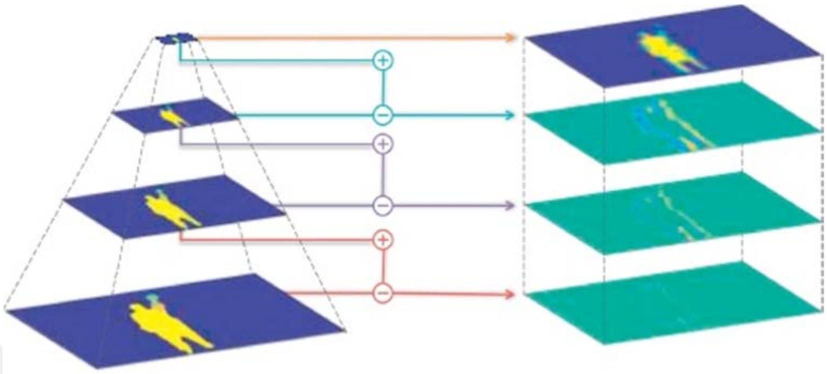


Figure 3.7. Generation Laplacian Pyramid from Gaussian Pyramid (Xiaopeng et al., 2015.).

3.1.9.2. Morphological Pyramid Based Fusion

Lowpass and bandpass filters cause some changes in the image like shifting the exact location. Morphological Pyramid (Ramac et.al., 1998) uses morphological filters to deal with composed deterioration in the Laplacian Pyramid Method. These filters can be applied with two methods; image opening and image closing. The opening and closing operations are performed by respective application of erosion and dilation to the image. First erosion and then dilation is applied in the opening operation. In the closing operation, the first dilation and then erosion is applied. Erosion and dilation operations are implemented with the same structural element. The Morphological Pyramid is formed from the Gaussian Pyramid, like the Laplacian Pyramid. The structural element is applied to the levels of the Gaussian Pyramid and the remaining operations are the same as in the Laplacian Pyramid Process.

3.1.9.3. Ratio of Lowpass Pyramid

Ratio Pyramid (Toet, 1989) starts with generating Gaussian Pyramid like Laplacian Pyramid but as a difference from Laplacian, this pyramid is created by the computing ratio between consecutive levels of Gaussian Pyramid. This approach's purpose is to protect important details in the image. Other operations are the same as Laplacian Pyramid; the fused pyramid is created by applying maximum selection, minimum selection or averaging method. The fused image has more details. Equation (15) shows generating L level Ratio Pyramid, R is the Ratio Pyramid, g is the Gaussian Pyramid, l is the number of levels. Ratio Pyramid can be called Contrast Pyramid because of luminance contrast. Equation (16) gives the relation between luminance contrast and ratio; C is the luminance contrast and I is the unit vector. The differences between Ratio and Gaussian Pyramid is shown in Figure 3.8. Ratio Pyramid makes clear details than Gaussian.

$$R_l = \begin{cases} \dfrac{g_l}{expand(g_{l+1})} & , \ l < L \\ expand(g_l) & , \ l = L \end{cases} \tag{3.15}$$

$$C_l = \frac{g_l}{g_{l+1}} - I \quad \text{so} \quad R_l = C_l + I \tag{3.16}$$

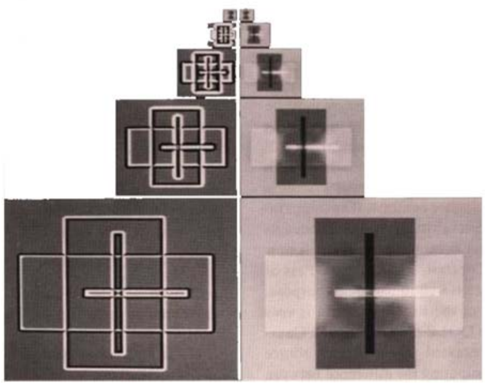


Figure 3.8. The left side is the Ratio Pyramid, the right side of the Gaussian Pyramid (Toet, 1989).

3.1.10. Multi-focus Image Fusion Using The Nonsubsampled Contourlet Transform (NSCT)

Contourlet Transform (CT) (Candès and Donoho, 1999; Do and Vetterli, 2005) is a two-dimensional transform method that has been proposed relatively recently and gains advantages over previous transform methods in some perspectives. CT attempts to achieve the geometrical details of the image. The most important features that distinguish this method; directionality and anisotropy. Unfortunately, CT that has superior features cannot provide shift-invariance. As a result of this imperfection, Nonsubsampled Contourlet Transform (NSCT) (Cunha et.al, 2005), a shift-invariance transform method, was introduced by Cunha et al. NSCT is performed with Nonsubsampled Filter Banks (NSFB) and Nonsubsampled Directional Filter Banks (NSDFB). NSCT is like the CT method except for shift-variance. NSFB is used for acquiring a Nonsubsampled Pyramid (NSP). NSCT operations flow is given in Figure 3.9.

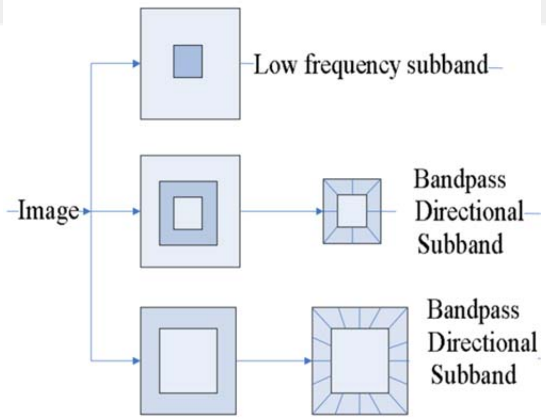


Figure 3.9. Nonsubsampled Contourlet Transform (Cunha, Zhou, Do, 2005).

Zhang and Guo proposed a new multi-focus image fusion method based on NSCT in 2009 (Zhang and Guo, 2009). Their fusion method is roughly processed in that way, NSCT is applied to source images and coefficients are obtained.

Thereafter, average of lowpass sub-band coefficients is constructed the fused lowpass sub-band. Fused directional sub-bands are constituted by selecting absolute maximum coefficients of source images' directional sub-bands. According to the study, the NSCT has been able to produce better results, however, it sacrifices the cost against the performance.

3.1.11. Image Fusion with Convolutional Sparse Representation (CSR)

Sparse representation (SR) expresses the data with as few as possible linear combinations of elements. This means that it aims to reveal the most powerful structure from the data. These elements that are also known as "atoms" are stored in a specific dictionary. Sparse representation has been extensively investigated in many popular and substantial areas, from image processing to signal processing and machine learning (Zhang and 2015). This subject was first used by Yang and Li (Yang and Li, 2010) in the multi-focus image fusion domain. In their study, the superiority of the developed method has been demonstrated and a solution to the shift-variance problem, which is an important detail in image fusion, has been proposed.

Inspired by SR based fusion studies, Liu and Chen et al. (Liu et al., 2016) developed a new method for image fusion by means of the Convolutional Sparse Representation (CSR) Method. They have tried to ascertain the deficiencies of SR-based fusion methods. According to their conclusions, first, many SR based fusion methods may lose details of images while creating the fused image and the second one, many methods cannot achieve successful results for misregistered images. A robust fusion algorithm should work well not only for unproblematic images but also for images that include various problems such as misregistration. The CSR overcomes the losing details problem by processing the whole image instead of examining the patches (Wohlberg, 2015). Concerning the second problem, the shift-invariant property of the CSR method ensures resistance to misregistered images and obtains successful results.

As to the experiments' results, it is proved that CSR can accomplish superiority over the SR-based methods. CSR based fusion study can be considered as an improved SR method version in terms of identifying and eliminating the weaknesses of SR-based fusion methods.

3.1.12. Image Fusion based on Pixel Significance Using Cross Bilateral Filter (CBF)

Kumar has developed a new image fusion method based on Cross Bilateral Filtering (Kumar, 2012). Bilateral Filtering applies a low-pass filter to the image to normalize the correlation among pixels. This process is performed by using neighboring pixels and reduces the information in the image while highlighting the details. The difference of Cross Bilateral Filter is that it uses both similarities in color space and proximity in geometric space. While one of the source images is used to form the filter, the other one is used to convolving the filter.

First, the original source images are percolated by the Cross Bilateral Filter. Then detail images are generated by taking differences between original and filtered images to gain weights. After the weights of each detail image are gained the final stage proceeds. In the final stage, a pixel-based approach is used to determine each pixel of the fused image by the weighted average method. Figure 3.10 describes the outline of the algorithm. A and B are the source images, wt_a and wt_b are the weights of A_{CBF} and B_{CBF} filtered results respectively. Then fused image F(i,j) is created based on pixel-level.

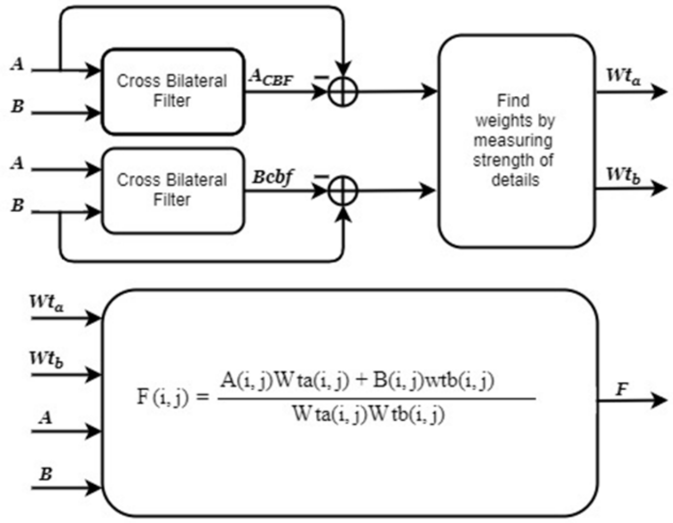


Figure 3. 10. Outline of the algorithm (Kumar, 2012).

3.1.13. Image Fusion with Guided Filtering (GF)

Before developing a fusion algorithm, it is necessary to make generalizable determinations. Afterward, algorithms are realized and tried to be strengthened with these determinations. With this approach, the necessity of preserving the details during the fusion process has been decided by the researchers. It is concluded that the way to protect the details is to capture the edges of images. Because strong edges are not lost by blurring. One of the ways in which image processing methods is the edge detection filters for preserving details. Recent researches show that, Guided Filter (GF) promises highly effective results among other edge detection filters He et al., 2012).

Smoothing in GF clears noises and artifacts yet strong details still remain. This property of GF helps to discriminate blurred and strongly focused pixels on the image.GF is an efficient filter because it doesn't damage the saliencies, for example, does not cause rings around the edges. In addition, it can be applied quickly by means of the feature independent of the filter size. GF has two important parameters r (kernel size) and ε that need to be decided, it can be expressed as in Equation (17). I_1 is the source image and I_2 is the guide image that may be source image itself or different. When the whole image is taken into consideration, it is logically concluded that the variance is high in the regions that contain the details, while the values of the other regions are similar and therefore the variance of these regions is low. Thus, the edges can be distinguished as regions that have a high variance. GF parameter ε determines the limit of blurriness. Regions that have much lower variance values than ε ($\sigma^2 \ll \varepsilon$) are smooth as a result of filtering, while large values remain clear. Figure 3.11 illustrates the effects of these two parameters on the input image.

$$S = G_{r,\varepsilon}(I_1, I_2) \tag{3.17}$$

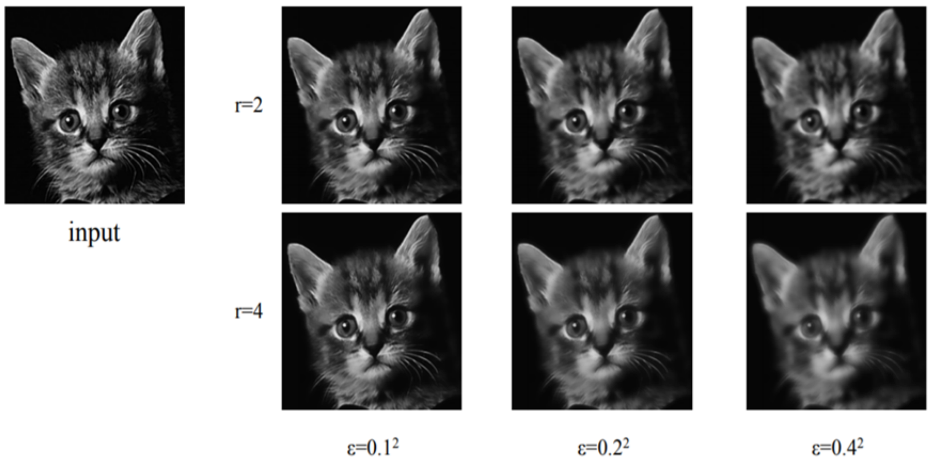


Figure 3.11. Effects of parameters in Guided Filtering (He et al.,2012).

Guided Filter was first used in the image fusion algorithm developed by Li et al. (Li et al.,2013). A promising approach has emerged in terms of performance and the fused image. GF is served as a two-scale transform process, thus provided less complex transformation. While GFF protects essential information of the source images like brightness, color, and details, it is observed that other methods create artifacts in the fused image and lose the light that causes to lose important details. In addition to these, GFF may also be robust to misregistered images. Finally, it is determined that this algorithm does not have high computation cost.

3.1.14. Multi-focus Image Fusion Using Multi-Scale Image Decomposition and Saliency Detection (MSSF)

Bavirisetti and Dhuli have developed a novel fusion method that is focused on saliency detection and multi-scale decomposition. Multi-scale methods decompose images into coefficients and these coefficients employ in the fusion operations rather the entire pixels. Coefficients are generally obtained by applying transformation operation to the images. However, this algorithm brings a different approach, decomposition is carried out by using the average filter instead of applying the transformation. In this way, the working complexity is reduced. The multi-scale decomposition process is supported by saliency detection in their study. Saliency detection allows obtaining remarkable parts, objects, and details of the image. A saliency detection method should be able to obtain prominent features and disregard noises in the picture. In this study, the saliency detection method is applied as in (Bavirisetti and Dhuli, 2016) reference since it works effectively even in challenging images. Thus, a new saliency detection based multi-scale fusion algorithm is formulated.

This algorithm is compared with seven different fusion methods. They are compared with three approaches; visual, quantitative and computational time. Benchmark metrics indicate that this study gets the upper hand on other compared methods. In addition to this, according to the computational time, this method is the shortest one among others.

3.1.15. Multi-Scale Weighted Gradient-Based Fusion for Multi-Focus Images (MWGF)

Zhou and Wang (Zhou and Wang, 2014) have developed a new method for solving major problems where present fusion algorithms fail to satisfy. These problems are defined as anisotropic blur and misregistration by researchers. MWGF method consists of two parts; First, the coarsely focused and out-of-focus regions are settled, and a structural gradient image is created. Then the multi-scale focus model is used to calculate the gradient weights. Second step help to identify the border of the focused region. Figure 3.12 shows maps of the initial (a) and final (b) results of this method.

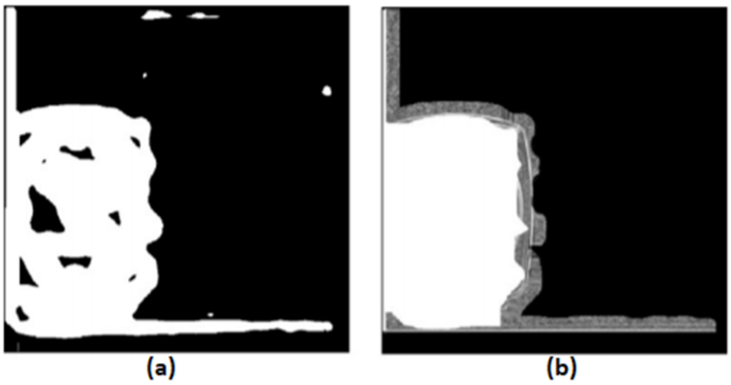


Figure 3.12. The map results of MWGF (Zhou and Wang, 2014).

According to the study, it has been stated that the developed method solves the targeted problems and beat the comparison methods with the helping of multi-scale fusion idea.

3.1.16. Multi-Focus Image Fusion Based on Wavelet Transform and Adaptive Block (DWTDE)

The DWTDE method has proposed by Liu and Wang in 2013. According to the study, it is stated that it is a promising method based on both visual and quantitative results. Furthermore, it has been suggested that DWTDE is a cost-effective method and has a high level of fusion quality.

This algorithm can be summarized as follows. The algorithm is based on the DWT. The image pixels are transferred to the frequency domain by applying DWT. The distinctive aspect of this study is the use of an adaptive block. The size of the adaptive block is optimized by the Differential Evolution Algorithm (Storn and Price, 1995). Then, this adaptive block helps to fuse low-frequency coefficients. As a result of the fusion of low-frequency coefficients by the adaptive block, a source label map is obtained. High-frequency coefficients fuse by using this source label map and wavelet energy method. Finally, high and low-frequency fusion results are reconstituted to achieve final fusion.

3.1.17. Fusion with Probabilistic Neural Network and Radial Basis Function Network

This study (Li et al,2002) is one of the first studies to use an Artificial Neural Network (ANN) models in multi-focus image fusion applications. They have used two neural network models; Probabilistic Neural Network (PNN) (Specht, 1990) and Radial Basis Function Network (RBFN) (Moody and Darken, 1989).

In this approach, the neural network is fed with images' features rather than direct images. Features that are considered important and distinctive are identified for the fusion process and these features are extracted from the source images. In other words, the properties of the images are trained, and the network is expected to be able to distinguish between the characteristics of the focused pixels and the defocused pixels.

First, the two images are divided into blocks, and each image block is processed as a pair. Feature extraction is performed on each block and three properties are obtained; spatial frequency, visibility, and edge information. These features are kept as feature vectors. After that, they are subtracted from each other and the resultant vector is given to the artificial neural network for training. If the first image's block is focused than the second image, the output of ANN is 1, otherwise, it is 0. After the training is done, the test phase is carried out by using the trained network. The fused image is formed as shown in Equation (18). A confirmation step is performed to make the fused image more accurate. If a part of the fused image is mostly taken from the same source image, it is ensured that the next block is taken from the same image again.

$$\text{Fused image block} = \begin{cases} 1^{\text{st}}\ \text{block},\ \text{output} > 0.5 \\ 2^{\text{nd}}\ block,\ \text{otherwise} \end{cases} \tag{3.18}$$

Both network models were trained with 60 image pairs separately and it is observed that they produced close results.

3.1.18. Multi-Focus Image Fusion with A Deep Convolutional Neural Network (CNN-Based)

Liu et al. (Liu et al.., 2017) developed a spatial domain image fusion study based on deep learning. In this study, a Convolutional Neural Network Model (Lecun et al., 1998) is used as a focus measure technique for the first time. The most important reason for choosing the CNN model is that this model can perform feature extraction and classification operations together. In this way, there is no need for manual operations or many workloads for these processes.

Before the fusion stages, the CNN training process is carried out with a huge dataset. This dataset comprises several images that are divided into the 16x16 size of pieces. These pieces are grouped as focused-blurred pairs. Eventually, 1000000 pairs are formed, and training is achieved. The algorithm consists of four basic steps; focus detection, initial segmentation, consistency verification, and fusion. In the focus detection step, a pre-trained CNN classifies two source images. The network constitutes a score map that includes numbers between 0 and 1. These scores make it possible to determine the focused patches of source images' pairs. Once the map is acquired, initial segmentation and consistency verification are applied to increase the robustness of the map. In the initial segmentation stage, the map is converted to the binary map with applying thresholding to classify the map's scores. Binary Map BM(x,y) is obtained from the map M(x,y) according to an empirically selected threshold value τ by using Equation (19):

$$BM(x,y) = \begin{cases} 1, M(x,y) > \tau \\ 0,\ otherwise \end{cases} \tag{3.19}$$

If the pixel values of the map are greater than τ , these pixels are assumed to correspond with the first source image otherwise, they correspond with the second source image. Following this, Small Region Removal and Guided Image Filtering (He et al.,2013) are applied for consistency verification. Consequently, the final decision map is multiplied with source images and the fused image is obtained. These steps are completely summarised in Figure 3.13.

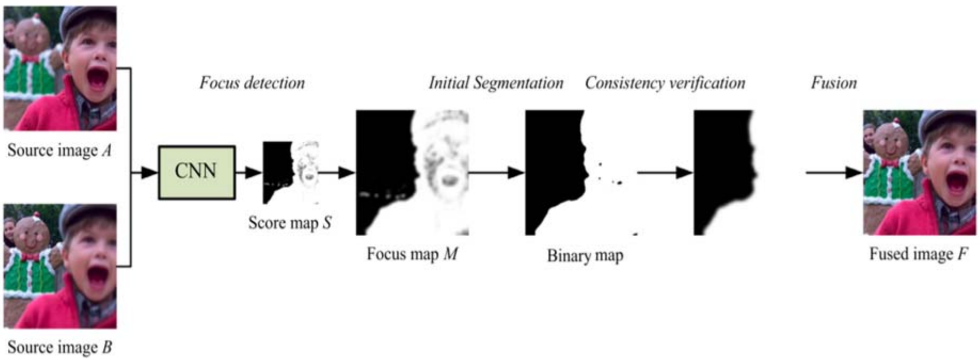


Figure 3.13. Summary of CNN fusion steps (Liu et al., 2017).

CNN based fusion technique can produce successfully fused images and compete with the other fusion algorithms. On the other hand, it has several disadvantages; computational time is not short and memory consumption is high due to the large dataset.

3.1.19. Laplacian Pyramid Transformation and Pulse Coupled Neural Network (LP-PCNN)

The Pulse-Coupled Neural Network Model (Eckhorn et al., 1990) is invented in accordance with neurophysiological discoveries of the cat's visual cortex. Good sides of the PCNN model are pulse synchronization and global coupling (Yan et al., 2017). However, this model involves an excessive number of parameters and high computational time. PCNN based image fusion method (Huang and Jing, 2007) was first introduced by Huang et al. Following this, the PCNN model is modified such as Dual-Channel PCNN (Wang and Ma, 2007), Multi-Channel PCNN (Wang and Ma, 2008) and these models subject to different image fusion studies (Wang et al.,2010). In some studies, the PCNN method is combined with different current fusion methods to acquire better fusion studies; PCNN-GF (Wang et al.,2016), PCNN-NSCT (Gayathri and Deepa, 2016), PCNN-NSST (Jin et al., 2016). One of them is the combining of PCNN and Laplacian Pyramid (Mingrui et al.,2015).

Pyramid transform (Andelson et al., 1984) discriminates an image to different levels with lowpass or bandpass filtering. Whereas the lowpass filter serves to alleviate high-frequency information, the bandpass filter serves to alleviates too low and too high-frequency information. Laplacian Pyramid is generated by using the Gaussian Pyramid. The Gaussian Pyramid consists of different levels of an image. The base-level includes the original image and the other levels are created from the former level by applying a lowpass filter and down sampling operation that performs decreasing the size. Levels of Laplacian Pyramid (L_l) are obtained by taking the difference between levels of the Gaussian Pyramid (G_l). Notice that, when applying the difference, the Gaussian levels need to be the same size, therefore these levels are expanded by the up sampling. These operations are given in Equations (20) and (21).

$$G_l = [w * G_{l-1}]_{\downarrow 2} \tag{3.20}$$

$$L_l = G_l - 4w * [G_{l+1}]_{\uparrow 2} \tag{3.21}$$

w is the lowpass filter and ↑ , ↓ symbols refer to down sampling and up sampling operations. The Laplacian Pyramid is obtained.

The fusion process is carried out in the following order; first of all, source images are decomposed into the Laplacian Pyramids' levels L(x,y). Afterward, these pyramids are fed to the PCNN. As a result of the PCNN, frequency maps of each level are achieved. For using as a focus measure, local entropy H(x,y) values are calculated for each map. Then, the fusion result is obtained by applying special operations based on entropy values of maps. These operations are summarized in Equation (22):

$$Fused(x,y) = \begin{cases} \frac{L_1(x,y)+L_2(x,y)}{2}, |H_1(x,y) + H_2(x,y)| \leq 0.015 \\ L_1(x,y),\ |H_1(x,y) + H_2(x,y)| > 0.015 \ and\ H_1(x,y) > H_2(x,y) \\ L_2(x,y),\ |H_1(x,y) + H_2(x,y)| > 0.015 \ and\ H_1(x,y) < H_2(x,y) \end{cases} \tag{3.22}$$

As a result, it has been ascertained that the better-fused images are obtained by means of combining these well-working methods, but it should be noted that combining the two methods increases the cost.

3.1.20. Fusion with Extreme Learning Machine (ELM-Based)

Extreme Learning Machine (ELM) (Huang et al., 2004) is a simpler network according to others, it has few parameters and small structure. It is proposed as a fusion method (Yang et al.,2017) and it has been argued that it gives better results than many methods, including the CNN fusion method. This method is performed two steps, in the first step is generating the initial fused image and it is like the study of fusion with the artificial neural network (Li et al., 2002). Source images are processed with pixel pairs. Three different properties of each pixel are extracted; texture feature, gradient, and the local visual contrast. If the feature difference is bigger or equal to 0, the output is 1, otherwise, it is 0. For the fused image, when the ELM network's output is greater than 0.5, the first image’s pixel is taken, contrarily second image’s pixel is taken. After the whole pixels are processed, the verification procedure is applied to the fused image. By the 3x3 window, if five or more pixels are taken from one image, other pixels that are in the window are selected from the same image again. Then the second step is constructed, Root Mean Square Error (RMSE) is calculated between source images and initial fused image. According to the smallest RMSE values, fused image pixels are decided. To further increase accuracy, morphological opening and closing operations are applied to recover the small number of faulty pixels. Consequently, the final fused image is created.

3.2. Histogram Equalization

Image enhancement is a crucial operation for image processing applications. It is generally seen as a front-end processing. Improving image information empowers the appearance of key features of the image, simplifies operations, and makes results more effective. One of the image enhancement methods is histogram equalization. The histogram is a graphical representation of the distribution of gray level in the image. The X-axis shows gray levels, while the Y-values depict the total number of pixels’ occurrences at that level in the image. In the histogram, values trapped between certain levels cause distribution distortion. In other words, derangement occurs when the values that can be distributed in the [0-255] interval are aligned within a small value range. So, the contrast information of such images is low. The uniform distribution is generated by spreading the high occurrence values to the low areas. Thus, the contrast value of details is increased by means of a histogram equalization technique.

Histogram equalization is generally examined into two way as Global (GHE) and Local Histogram Equalization (LHE). In GHE, all image information is processed together. The aim of it to rearrange the density to all grey levels for reaching an enhanced image. Thus, it provides an overall contrast enhancement. Although this application is proper for general improvement, the GHE cannot maintain some characteristics of local regions.

The sequence of operations is as follows; First of all, the histogram of the image is created. Then, the Cumulative Density Function (CDF) is calculated. The cumulative density is the magnitude that includes the values obtained by the sum of the previous and each value of the histogram itself. By using these cumulative density values, new gray level values are achieved as in Equations (23 and 24).

$$p_n = \frac{number\ of\ pixels\ in\ the\ n^{th} gray\ level}{total\ number\ of\ pixels} \tag{3.23}$$

$$g_k = floor\big((L-1)\sum_{n=0}^{k} p_n\big)\,,\ k = 0,1,2,\ldots,L-1 \tag{3.24}$$

where n represents the gray level, L indicates the number of possible gray levels.

Because of GHE takes the whole picture into consideration, it causes loss of local information such as brightness. In order to find a solution to this problem, the LHE technique has been developed. The difference of the LHE method is that the image is processed based on the blocks, instead of processing all pixels at the same time. It scans the image with the user-specified window size and transactions only work on pixels within the window.. Figure 3.14 shows the histogram of original image before appplying local histogram equalization. Figure 3.15 is the result of LHE Figure 3.26 displays the original image (left) and the equalized image (right). As seen in the resulting image, the contrast of the details of the image is enhanced with this method successfully (Gonzalez and Woods, 2008; Zhu et al., 1999; Kong et al., 2013).

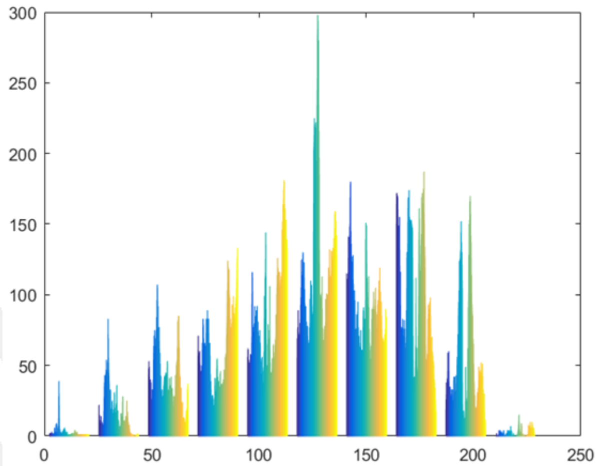


Figure 3.14. Histogram of the image, before equalization.

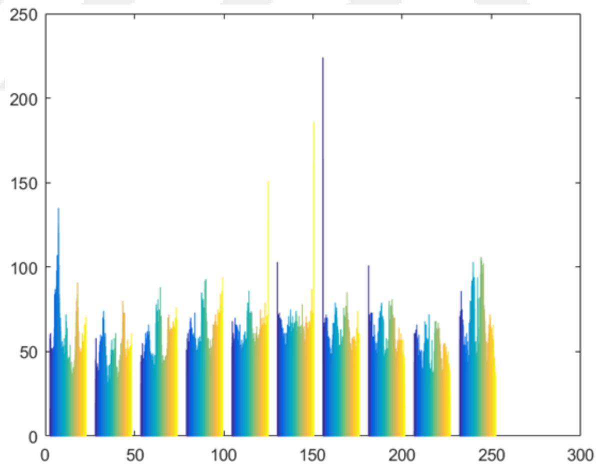


Figure 3.15. Histogram of the image, after equalization.

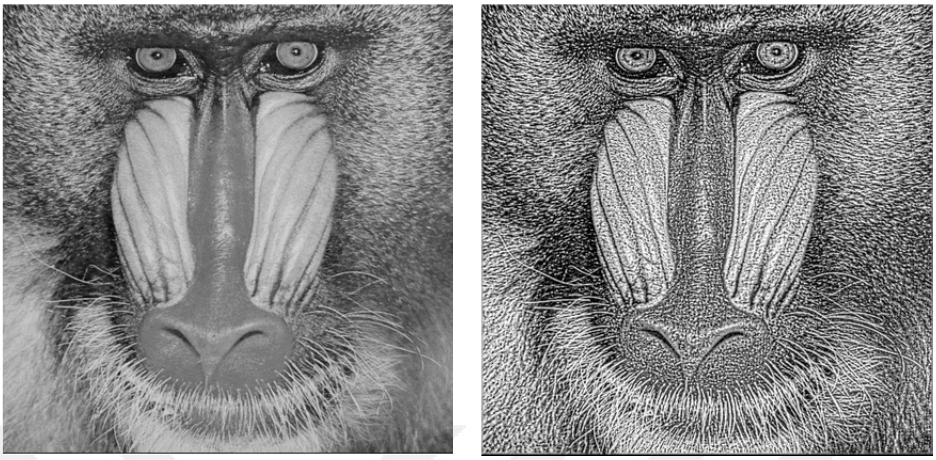


Figure 3.16. Results of Local Histogram Equalization.

3.3. Halftoning

Digital Image Halftoning or Dithering (Ulichney, 1987) provides to converting a continuous-tone image to a binary-valued (black and white) image. This technique places the black and white microdots in appropriate regions to generate a bi-level image that is visually similar to a continuous image. Due to the low pass frequency property, the human eyes perceive the gray-toned patches, instead of the microdots. Thus, halftoning create an optical illusion and a small size image that has similar visual to the original image is obtained. So it may be considered as image compression techniques. However,. halftoning techniques are generally used in printing, facsimile (FAX), scanner, computer graphics and in different image processing and electronic applications. Grayscale images may contain 256 gray levels, on the other hand, traditional printers have one color ink In order to reproduce continous tone images on printed surface, by using single color ink, we should find a magical way. So as a solution, the halftoning method creates an illusion with black dots on the white paper to achieve gray shades (Sindhu, 2013). Halftoning algorithms can be categorized as in Figure 3.17.

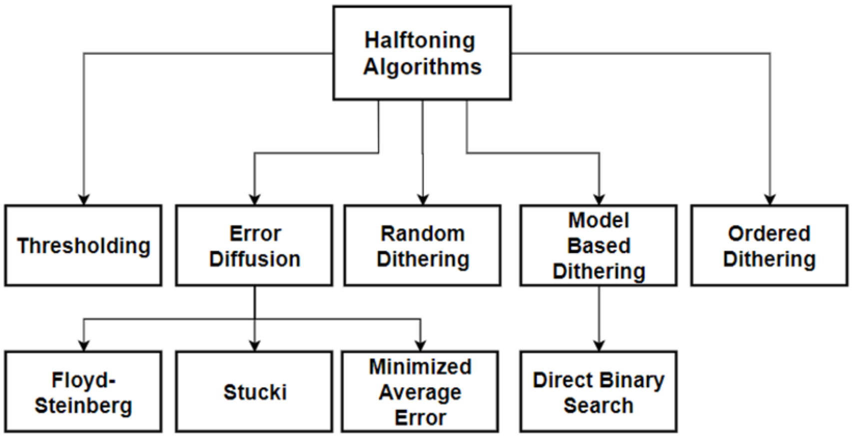


Figure 3.17. Halftoning Algorithms.

3.3.1. Thresholding (Average Dithering)

A fixed variable is determined and it is used as a threshold value then applied to the continuous image. This threshold value scales the image to ‘0’ and ‘1’. The pixels that are higher than the threshold value take ‘1’ and the remaining values take ‘0’. Eventually, a bi-level image is achieved. As seen, this algorithm, which has a very clear approach, is unfortunately not sufficiently effective. It causes artificial artifacts (Silva et al., 2000). Figure 3.18 shows the result of the thresholding application, the left image is the original one, and the right one is the dithered image.

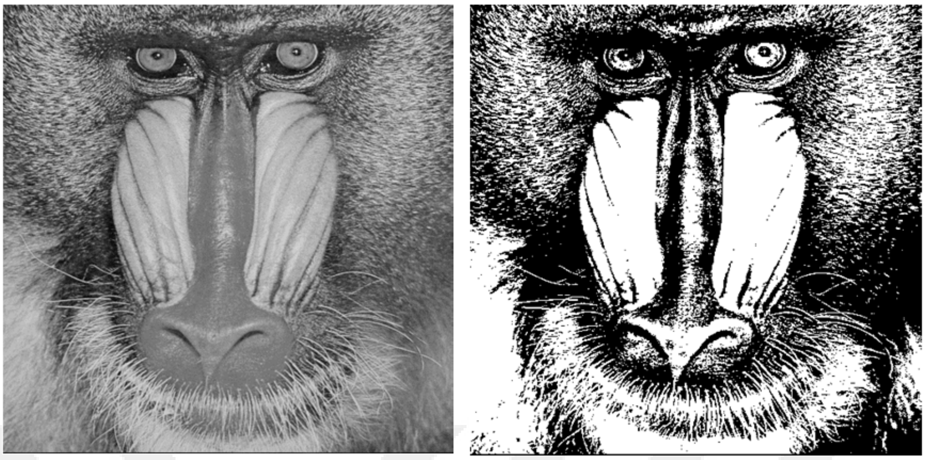


Figure 3.18. Threshold dithering application.

3.3.2. Random Dithering (White Noise Dithering)

This algorithm type has been developed to overcome the drawbacks of the threshold dithering method. Instead of a fixed threshold value, a random threshold value is generated for each pixel and the pixel value is compared to it. If the pixel value is greater than the threshold it is assigned to ‘1’, otherwise, ‘0’ is assigned. In this method, too much noise is added to the image while protecting from artificial artifacts (Ulichney, 1988). Random Dithering application is shown in Figure 3.19, the left image is the original one, and the other one is the Random dithered image.

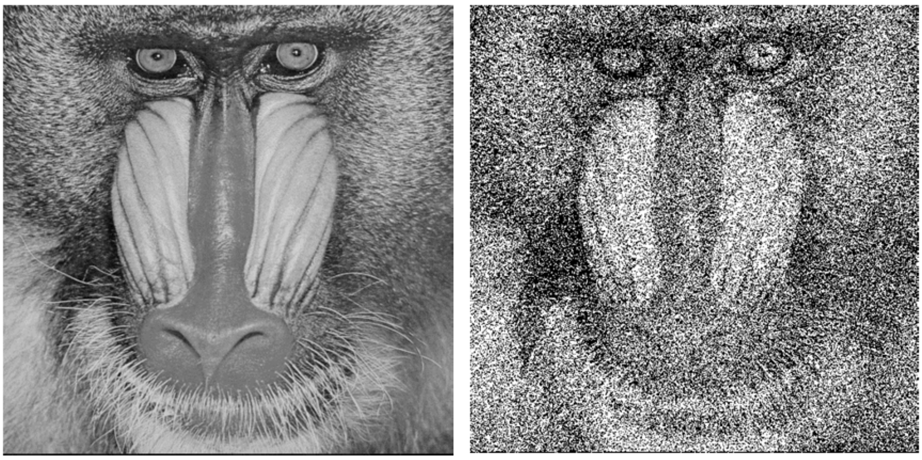


Figure 3.19. Random Dithering application.

3.3.3. Ordered Dithering

In this method, a matrix (screen) is created and each pixel is compared to the corresponding value in the matrix. If the pixel value is less than the matrix's value it is assigned to '0', otherwise, it is assigned to '1'. The most appropriate matrix must be created. The algorithm is usually implemented in two ways; clustered-dot and dispersed-dot dithering. The structure of the matrices applied by the two methods is different. While low-frequency values are in the center in the clustered-dot method, the low values are appeared disorderly in the dispersed-dot method. These methods protect the sharpness of the original image. Figure 3.20 is illustrated by clustered-dot and dispersed-dot methods. Figure 3.21 is the result of ordered dithering on a sample image, the left image is the original one, the right image is the order dithered image (Sullivan et al.,1993).

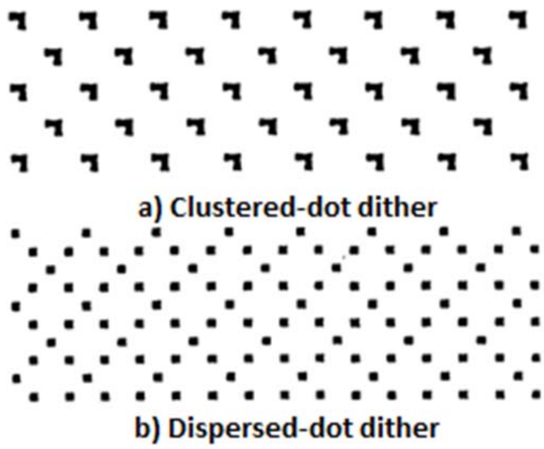


Figure 3.20. Ordered Dithering Method's illustration (Sullivan et al.,1993).

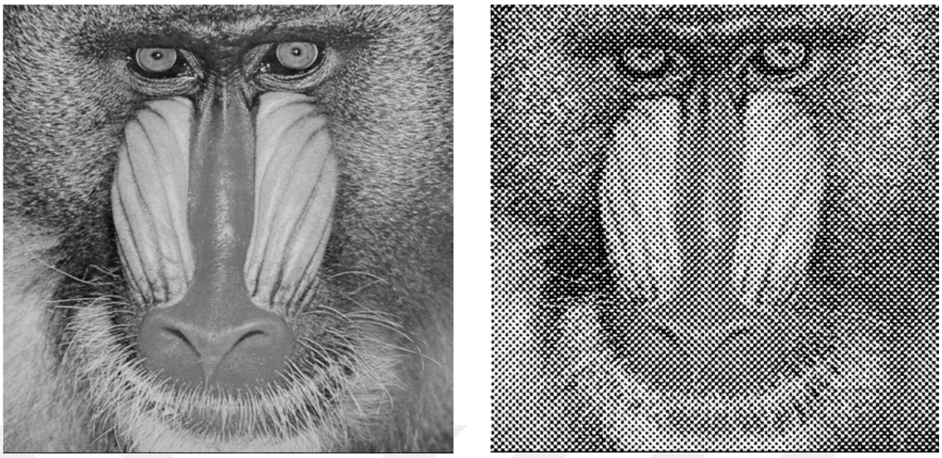


Figure 3.21. Ordered Dithering application.

3.3.4. Error Diffusion

One of the most popular halftoning techniques is the error diffusion method. This study field dates back to 1976 to compose halftoned images from continuous images. Its simple algorithm structure produces surprisingly effective results. The error diffusion method processes on a pixel and diffuses the error to its neighbors. Thus, it tries to reduce the total error by distributing each pixel's error.

The algorithm actually works as follows; the picture is reshaped with pure black(0) and pure white(255) dots. When deciding on the new value of the pixel, its proximity to black and white is determined and assigned to the value it is close to. Figure 3.22 illustrates an example of four pixels image piece.

30	120
96	150

→

0	0
0	255

Figure 3.22. An example of new values selection in Error Diffusion.

However, in this way, we can't get accent color illusion and we lose the details of the image. For this, the error diffusion algorithm has provided an effective approach. The working principle of the algorithm is as follows; operations start to top-left of the image. First of all, the new value of the pixel is decided to be black or white as in Equation (25). For equations, $0 < p_{i,j} \leq M$ refers to the related pixel of the continuous image and M is the maximum value that the pixel can take. $P'_{i,j}$ is a symbol of the new value of the pixel. Then the error is calculated based on the new value of the pixel as in Equation (26). And this error is distributed to the neighbors of the pixel. This operation is shown in Equation (27). The error is only distributed to unprocessed neighbors. For performing error propagation, the movement to the related pixel's neighbors is provided by adding x and y parameters. Finally, the alpha $\alpha_{x,y}$ parameter is the diffusion matrix coefficient corresponding to the neighbor pixel. Error diffusion to the neighbors is shown in Figure 3.23. Black circles represent processed pixels, while empty circles are unprocessed pixels. Thus X that is the related pixel distributes the error to the only empty circles. All processes are summarized in Figure 3.24.

$$p'_{i,j} = \begin{cases} 1, p_{i,j} \geq M/2 \\ 0, otherwise \end{cases} \tag{3.25}$$

$$E_{i,j} = p_{i,j} - p'_{i,j} \tag{3.26}$$

$$p'_{i+x,j+y} = p_{i+x,j+y} + E_{i,j} * \alpha_k \tag{3.27}$$

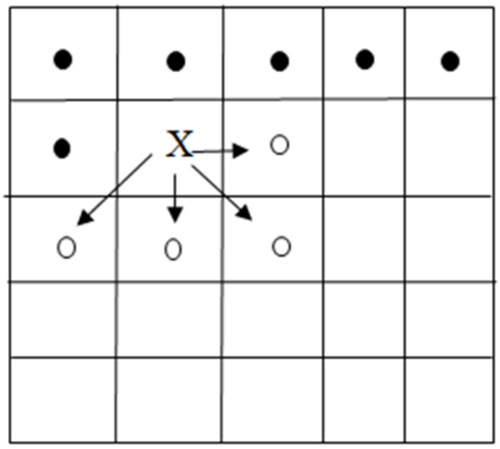


Figure 3.23. Illustration of error diffusion to the neighbours

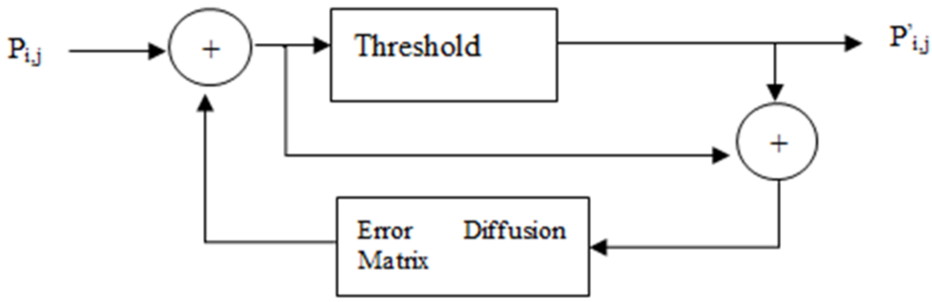


Figure 3.24. Illustration of error diffusion to the neighbours.

Error Diffusion has been investigated by several researchers, especially Floyd (Floyd and Steinberg, 1976), Ninke (Jarvis et al.,1976), and Stucki et al (Stucki, 1981). These researchers have presented different diffusion matrices. Although the Floyd-Steinberg method is a little worse than the other two methods, it is the most preferred error diffusion method because it has easier computation than other algorithms.

3.3.4.1. Floyd-Steinberg

Floyd-Steinberg Error Diffusion method is firstly presented in 1975 (Floyd and Steinberg, 1976). They have introduced a small matrix as the diffusion matrix.

This feature makes easier error propagation. The error is scattered to the four neighbors. These neighbors are multiplied with the diffusion matrix's weights. Figure 3.25 (a) shows a source image and Figure 3.25 (b) shows the Floyd-Steinberg diffusion matrix. For example, error diffusion is implemented for a neighbor as in Equation (28). Generally, Floyd-Steinberg is preferred among researchers by means of efficiency and simplicity. Figure 3.26 shows the implementation of this algorithm, the left image is the original one and the other one is the dithered image.

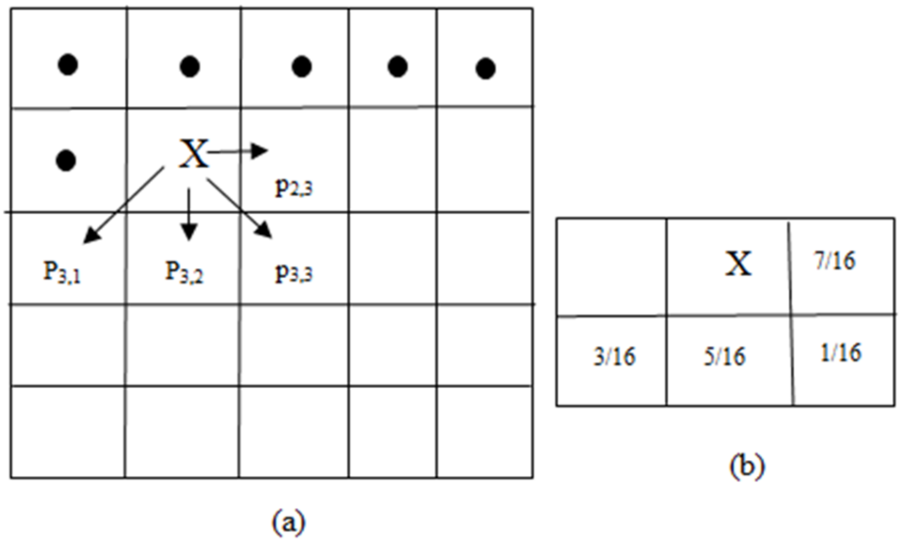


Figure 3.25. Floyd-Steinberg Diffusion matrix and a sample source image.

$$p'_{2,3} = p_{2,3} + E_{2,2} * \frac{7}{16} \quad (3.28)$$

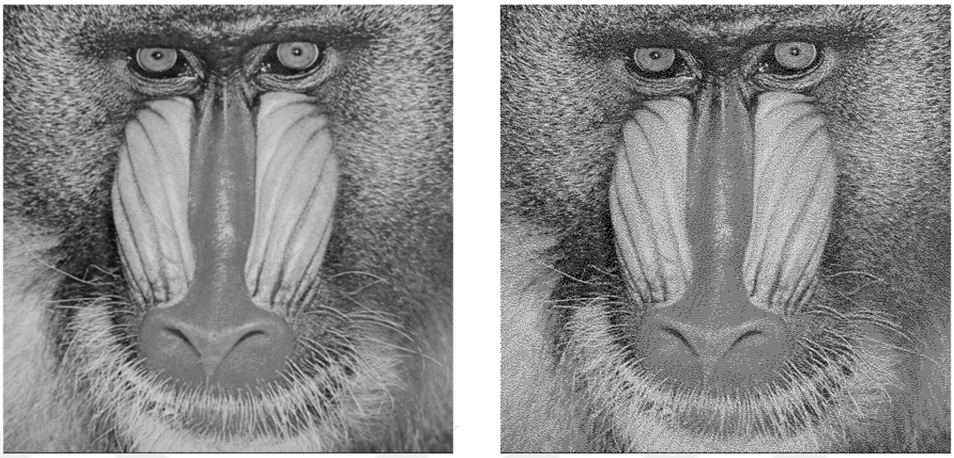


Figure 3.26. Floyd-Steinberg implementation.

3.3.4.2. Minimized Average Error

Jarvis, Judice, and Ninke proposed an error diffusion algorithm in 1975 after the Floyd-Steinberg Algorithm (Jarvis et al.,1976). It is a more complicated filter than the Floyd-Steinberg Algorithm. The error diffuses to twelve neighbors instead of four neighbors. Figure 3.27 shows the filter of this algorithm and its result image is given in Figure 3.28 (right-hand side).

●	●	●	●	●
●	●	X	7/48	5/48
3/48	5/48	7/48	5/48	3/48
1/48	3/48	5/48	3/48	1/48

Figure 3.27. Minimized Error Algorithm Matrix.

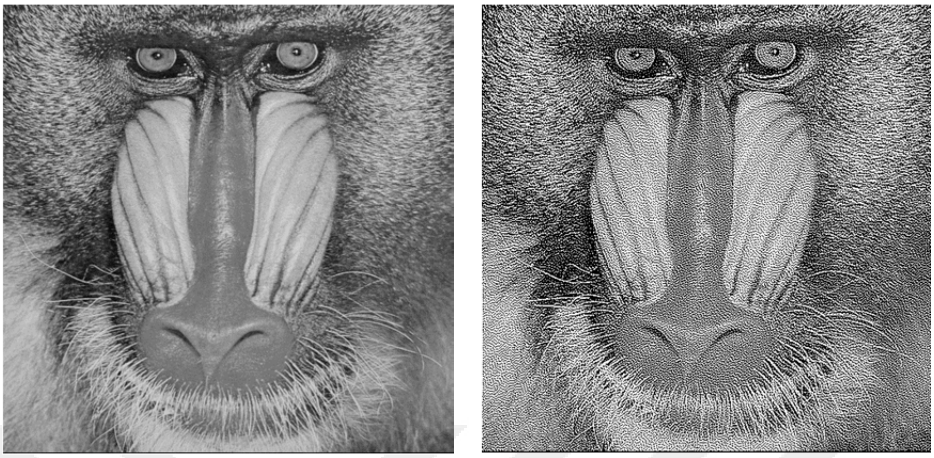


Figure 3.28. Minimized Error Algorithm implementation.

3.3.4.3. Stucki

Stucki dithering (Stucki, 1981) produces similar results to Floy-Steinberg but generates fewer noises. Generally, it is preferred because of its fast working principle. However, this algorithm exhausts the processor, so it is not considered as an effective algorithm for major projects. In addition, it is suitable for applications that require the diminishing of color depth. Figure 3.29 mimics the matrix of the Stucki Algorithm. The implementation result is given on the right side of Figure 3.30.

●	●	●	●	●
●	●	X	8/42	4/42
2/42	4/42	8/42	4/42	2/42
1/42	2/42	4/42	2/42	1/42

Figure 3.29. Stucki matrix.

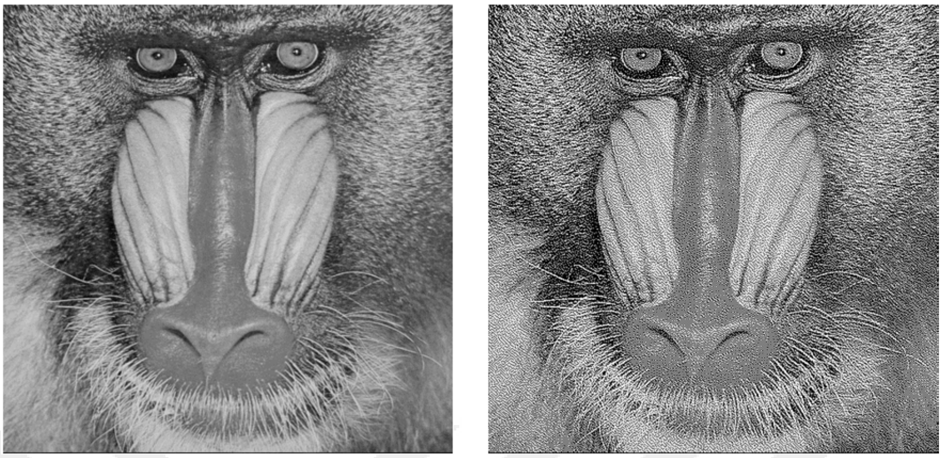


Figure 3.30. Implementation of Stucki.

3.3.5. Model-Based Halftoning

In the algorithms described above, approaches are presented independently of the perception of the human eye. The purpose of model-based halftoning is to obtain more successful halftoned images by utilizing models that mimic the Human Vision System (HVS). It considers both printer and human eye systems together. There are many models in this category in the literature, one of which is explained in sub-titles.

3.3.5.1. Direct Binary Search

In error diffusion methods, each pixel is processed only once and is not returned to the processed pixel. Direct Binary Search Halftoning (Allebach and Lin, 1996) is an iterative procedure. The pixels continue to be processed until a certain number of iterations or errors can no longer be reduced. All of the selections, such as the HVS model, algorithm parameters, processing order of pixels, all influence the success of the method. Toggle action updates the value of the pixel being processed to ‘0’ or ‘1’. Swap action exchanges the pixel being processed with its eight neighbors. At each iteration, the error gradually decreases and the best result is achieved when the error can no longer be minimized. Figure 3.31 illustrates toggle and swap actions.

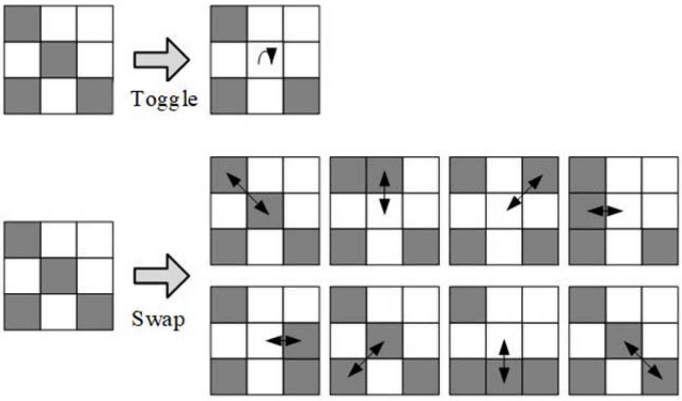


Figure 3.31. Toggle and swap operations (Sankar, 2019-http://imageprocessing-sankarsrin.blogspot.com/2019/05/direct-binarysearch-halftoning-dbs.html).

DBS algorithm works very slowly, even if it produces very successful images. Therefore, it exploits the hardware and becomes difficult to choose in big projects. Here is the implementation's result of DBS Halftoning in Figure 3.32. The image that is on the left is the original one, the right one is the result of DBS.

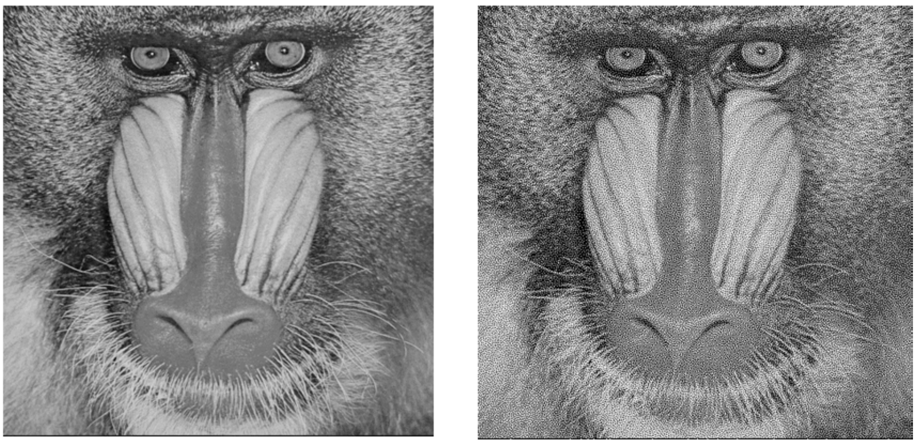


Figure 3.32. Implementation of DBS Halftoning.

3.4. Inverse Halftoning

Halftoned images are not proper to be used in many image processing applications such as compression, scaling, etc. For these situations, continuous-tone images are used. The inverse halftoning methods converts the binary image that is obtained by halftoning algorithms to continuous form. The major problem of inverse halftoning, there can be more than one grayscale version of a halftoned image. Even if the applied halftoning algorithm is known, a single fixed image cannot be obtained. Another problem is that inverse halftoning does not attract as much attention as halftoning among researchers. In the literature, the original images are commonly considered lowpass. But by applying a lowpass filter to halftoned images, we cannot get the original image directly, because we can also lose the important detail information at the same time. Therefore, the inverse halftoned process is not very easy, and the multiple features of the image need to be examined (Me, 2000).

Inverse Halftoning methods are divided into two categories; Dithering Model-Based Inverse (DMI) and Not based on the Dithering Model Inverse (NDMI). The methods in the DMI category implement inverse halftoning according to the halftoning algorithm of the corresponding halftoned image. That is, these algorithms have pre-knowledge and produce solutions depend on this information. NDMI algorithms are not interested in how the halftone image is generated or which algorithms are part of the halftone procedure. These methods view inverse halftoning as a filtering method or removing dithering noise. Since these algorithms develop inverse halftoning methods without prior knowledge, their performance is restricted. If information is available, better results can be accomplished by using this information. However, NDMI algorithms have high speed and less complexity while sacrificing performance. There are various inverse halftoning methods in the literature; Look Up Table based (Me, 2000), Minimum Mean Square Error (MMSE) and Maximum a Posteriori Probability (MAP) based (Wong, 1995), Decision Tree-based (Kim and Queiroz, 2003), etc.

3.5. Focus Measure

The key point of a fusion method is the focus measure. An appropriate focus measure is used to distinguish focused image blocks from out-of-focused image blocks. Succesful focus measures have some rule of a thumb:

- Invariability according to image structure.
- Resistant to noises and artifacts.
- Unimodal; the measurement must have only one peak value.
- Efficient computational time and complexity

The focus measurement and fusion application should be in compliance with each other. So, we have concentrated on high-frequency information (gradients) of the images. Correspondingly, gradient-based focus measurements are mainly examined. In the following, Energy of Image Gradient and Laplacian, Brenner's Gradient, Tenengrad and its variant have been presented as focus measures. They are formulated as in Equations (29)-(33). (Shah, 2017; Tian, 2011).

- The Energy of Image Gradient (EOG)

$$F_{EG} = \sum_r \sum_c (I_r^2 + I_c^2) \tag{3.29}$$

where I_r and I_c correspond to gradient information at r and c positions in the image.

- Energy of Laplacian

$$F_{EL} = \sum_r \sum_c (\nabla^2 I(r,c))^2 \tag{3.30}$$

$$where, \nabla^2 I(r,c) = \sqrt{I_r^2 + I_c^2}$$

where gradients I_r and I_c is obtained by convolving the Laplacian kernel.

$$\text{Laplacian Kernel} = \begin{pmatrix} -1 & -4 & -1 \\ -4 & 20 & -4 \\ -1 & -4 & -1 \end{pmatrix}$$

- Brenner's Gradient

$$F_{BGR} = \sum_r \sum_c |I(r,c) - I(r+2,c)|^2 \tag{3.31}$$

$$\text{In case } |I(r,c) - I(r+2c)|^2 \geq 0$$

- Tenengrad

Equation (32) defines the difference of a pixel with its neighbor at a distance of two units vertically / horizontally.Tenengrad

$$F_{TNG} = \sum_r \sum_c \left(G_x(r,c)^2 + G_y(r,c)^2\right) \tag{3.32}$$

In the equation, G_x(x-direction) and G_y (y-direction) are gradient information that is produced by using the Sobel filter.

$$Sobel\ Filter\ G_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}$$

$$Sobel\ Filter\ G_y = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}$$

- Tenengrad Variance

$$F_{VGR} = \sum_r \sum_c (G(r,c) - \bar{G})^2 \tag{3.33}$$

$$G = \sqrt{G_x^2 + G_y^2}$$

In Equation (33) $\bar{G}$ is the mean of G (Gradient information).

3.6. Metrics

A procedure to prove the success of a method is objective evaluation. Various metrics are available for this purpose. The common feature of these metrics is that they do not need a reference picture. The fully focused image is called as a reference image, and many test images do not have a reference. Therefore, non-reference metrics have been developed. The common feature of the metrics that are used in this study is they do not require a reference picture. Metrics are examined in four categories; information theory-based, image structural similarity-based, human perception-based, image feature-based. In this study, six metrics are selected to make assessments from these categories. These;

i. Information Theory-Based:
   Normalized Mutual Information (Q_{NMI}),
   Nonlinear Correlation Information Entropy Metric (Q_{NCIE}).
ii. Image Feature Based:
   Xyedas-Petrovic Metric ($Q_{AB/F}$).
   Peng-Wei Wang Metric (Q_{PWW}).
iii. Image Structural Similarity Based:
   Yang Metric (Q_Y).
iv. Human Perception Based:
   Chen-Blum Metric (Q_{CB}),

3.6.1. Mutual Information (NMI)

Mutual Information (MI) (Qu et al., 2002) determines the similarity relations between two discrete variables. This is described as Kullback–Leibler Divergence (relative entropy) that is the distance concerning the probability distribution of two variables in mathematics. In the image fusion, MI is used as a metric to gain performance information about algorithms. Kullback–Leibler Divergence equation expansion is outlined for two images as follows:

$$I_{AB}(a,b) = \sum_{a,b} p_{AB}(a,b) \log \frac{p_{AB}(a,b)}{p_A(a)p_B(b)} \tag{3.34}$$

In this equation, A and B are two images and a,b are their related pixels. While $p_{AB}(a,b)$ is the joint distribution, $p_A(a)$ and $p_B(b)$ are the marginal probabilities and their multiplication is the distribution associated including total independence.

Joint distribution is achieved by using the normalized joint histogram (entropy) of images. Joint distribution is defined as:

$$p_{AB}(a,b) = \frac{1}{2NM} h_{AB}(a,b) \tag{3.35}$$

$$h_{AB}(a,b) = \sum_{n=1}^{N} \sum_{m=1}^{M} h_{AB}(I_A(m,n), I_B(m,n)) \tag{3.36}$$

where N and M values are the size of images. $I_A(m,n)$ and $I_B(m,n)$ are intensities at m and n coordination.

In the fusion process, the amount of information obtained is taken into consideration. That is, the mutual information value is extracted between the fused image and the input images, and these values are summed to get the total information in the fused image. For performance comparison, the image including the highest MI value is considered more successful and knowledgeable. Kullback–Leibler Divergence is used for obtaining the joint distribution of fused and source images, then these values are added. These operations can be summarized as in Equations (37)-(39).

$$I_{FA}(f,a) = \sum_{f,a} p_{FA}(f,a) \log \frac{p_{FA}(f,a)}{p_F(f)p_A(a)} \tag{3.37}$$

$$I_{FB}(f,b) = \sum_{f,b} p_{FB}(f,b) \log \frac{p_{FB}(f,b)}{p_F(f)p_B(b)} \quad (3.38)$$

$$M_F^{AB} = I_{FA}(f,a) + I_{FB}(f,b) \quad (3.39)$$

For the above equations, while A and B are the source images, F is the fused image. a,b,f values refer the related pixels' positions.

Hossny et al. (Hossny et al., 2008) distinguish that the MI values $I_{FA}(f,a)$ and $I_{FB}(f,b)$ that are obtained for each input image is not calculated on the same scale. Thus they update the MI equation by this means:

$$M_F^{AB} = \frac{I_{FA}(f,a)}{H(f) + H(a)} + \frac{I_{FB}(f,b)}{H(f) + H(b)} \quad (3.40)$$

The new version of the MI is named as Normalized Mutual Information (NMI).

3.6.2. Nonlinear Correlation Information Entropy Metric (NCIE)

A Nonlinear Correlation Coefficient (NCC) (Wang et al., 2008; Liu et al., 2012) defines the relationship between two discrete values of the multivariable data set by means of creating a nonlinear correlation matrix. NCC is similar to Mutual Information (MI) and it produces a number that is in the interval [0,1], bigger values indicates strong relationship. It is formulated as:

$$NCC(A,B) = H'(A) + H'(B) - H'(A,B) \quad (3.41)$$

Where H' values are the reversed entropy, A and B are two discrete variables and b is the number of elements. Reversed entropy equations are given in Equations (42)-(44).

$$H'(A,B) = -\sum_{i=1}^{b}\sum_{j=1}^{b} h_{AB}(i,j) log_b h_{AB}(i,j) \tag{3.42}$$

$$H'(A) = -\sum_{i=1}^{b} h_A(i) log_b h_A(i) \tag{3.43}$$

$$H'(B) = -\sum_{i=1}^{b} h_B(i) log_b h_B(i) \tag{3.44}$$

Let R be the correlation matrix and each matrix element refers to an NCC value. If we take image fusion into consideration, the correlation matrix of the source images and the fused image can be shown below. A, B and F refer to source images and fused image respectively. Since the correlation of the same variables in diagonal elements is examined, the similarity is exactly the same and these elements take the value of '1'.

$$R = \begin{pmatrix} NCC_{AA} & NCC_{AB} & NCC_{AF} \\ NCC_{BA} & NCC_{BB} & NCC_{BF} \\ NCC_{FA} & NCC_{FB} & NCC_{FF} \end{pmatrix} = \begin{pmatrix} 1 & NCC_{AB} & NCC_{AF} \\ NCC_{BA} & 1 & NCC_{BF} \\ NCC_{FA} & NCC_{FB} & 1 \end{pmatrix}$$

The NCIE of the R matrix is calculated as:

$$NCIE = 1 + \sum_{i=1}^{3} \frac{\lambda_i}{3} log_b \frac{\lambda_i}{3} \tag{3.45}$$

Eigen values of the matrix R is represented by λ_i, (i= 1,2,3).

3.6.3. Xydeas and Petrovic Metric

Xydeas and Petrovic (Xydeas and Petrović, 2000; Haghighat et al., 2011) have introduced a fusion metric that measures pixel-based performance. This metric calculates the amount of edge information transferred from the source images to the resulting image. So it performs visual comprehension just like the human eye. When performing on edge operation, it disregards regional information. Images that have more edge information transferred, in other words, high metric values, are considered successful. Sobel edge detection method is made use of to acquire strength g(i,j) and orientation α(i,j) information. i and j are the symbols of pixel coordinates. The Sobel detector results are represented by s_A^x in horizontal and s_A^y in vertical. The relative orientation and relative strength equations obtained from a source image are shown below:

$$g_A(i,j) = \sqrt{s_A^x(i,j)^2 + s_A^y(i,j)^2} \tag{3.46}$$

$$\alpha_A(i,j) = tan^{-1}\left(\frac{s_A^y(i,j)^2}{s_A^x(i,j)^2}\right) \tag{3.47}$$

The relative strength G^{AF}(i,j) and orientation A^{AF}(i,j) of source image A corresponding with the fused image F are given in Equations (48 and 49).

$$G^{AF}(i,j) = \begin{cases} \dfrac{g_F(i,j)}{g_A(i,j)}, & if\ g_A(i,j) > g_F(i,j) \\ \dfrac{g_A(i,j)}{g_F(i,j)}, & otherwise \end{cases} \tag{3.48}$$

$$A^{AF}(i,j) = \frac{\left||\alpha_A(i,j) - \alpha_F(i,j)| - \pi/2\right|}{\pi/2} \tag{3.49}$$

Then edge strength and orientation results are obtained. Preservation values of the edge strength $Q_g^{AF}(i,j)$ and orientation $Q_\alpha^{AF}(i,j)$ indicate the amount of representation source image A in the fused image F. These equations are given in Equations (50 and 51).

$$Q_G^{AF}(i,j) = \frac{\Gamma_G}{1 + e^{K_g(G^{AF}(i,j) - \sigma_g)}} \tag{3.50}$$

$$Q_\alpha^{AF}(i,j) = \frac{\Gamma_\alpha}{1 + e^{K_\alpha(A^{AF}(i,j) - \sigma_\alpha)}} \tag{3.51}$$

Where the values Γ_G, K_g, σ_g and Γ_α, K_α, σ_α define the precise shape of the sigmoid function. Edge preservation values are shown in Equation (52). This result is in interval [0,1]. While '0' means that the source image's edge information is never transmitted to the fused image, while '1' means that complete edge information is associated with the source image to the fused image.

$$Q^{AF}(i,j) = Q_g^{AF}(i,j) Q_\alpha^{AF}(i,j) \tag{3.52}$$

In the image fusion, edge preservation values of two source images $Q^{AF}(i,j)$ and $Q^{BF}(i,j)$ with respect to the fused image are calculated and normalized with weights $w^A(i,j)$ and $w^B(i,j)$ as follows:

$$Q^{AB/F} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{M} Q^{AF}(i,j) w^A(i,j) + Q^{BF}(i,j) w^B(i,j)}{\sum_{i=1}^{N} \sum_{j=1}^{M} (w^A(i,j) + w^B(i,j))} \tag{3.53}$$

Where weights are computed by $w^A(i,j) = [g_A(i,j)]^L$ and $w^B(i,j) = [g_B(i,j)]^L$. L is a constant value in the formulas.

3.6.4. Peng-Wei Wang Metric (PWW)

Wang et al. (Wang and Liu, 2008) have discerned the importance of saliency information in the image fusion process, and developed a metric that is based on edge information. However, this metric examines the saliency information on different scales of images. It uses the DWT method to decompose the image into scales. DWT is a transformation method, it transforms the image from spatial level to the frequency level. DWT yields four different frequency bands; HH, HL, LH, LL as a result of decomposition. In this metric calculation, edge preservation values are computed for each scale. The edge preservation vector formula can be given as in Equation (54). for a scale l.

$$\overrightarrow{\varphi_l}(i,j) = \Psi_{H_l}(i,j)\vec{h} + \Psi_{V_l}(i,j)\vec{v} + \Psi_{D_l}(i,j)\vec{d} \tag{3.54}$$

In Equation (54), i and j indicate that coordinate points and $\vec{h}$, $\vec{v}$, $\vec{d}$ values are the unit vector with respect to horizontal, vertical and diagonal. $\Psi_{H_l}(i,j)$ values of each direction are calculated separately between a source image A and fused image F. These equations are given in Equations (55)-(57).

$$\Psi_{H_l}^{AF}(i,j) = \exp\left(-\left|LH_l^A(i,j) - LH_l^F(i,j)\right|\right) \tag{3.55}$$

$$\Psi_{V_l}^{AF}(i,j) = \exp\left(-\left|HL_l^A(i,j) - HL_l^F(i,j)\right|\right) \tag{3.56}$$

$$\Psi_{D_l}^{AF}(i,j) = \exp\left(-\left|HH_l^A(i,j) - HH_l^F(i,j)\right|\right) \tag{3.57}$$

Last value of the edge preservation value EP is calculated as:

$$EP_l^{AF} = \frac{\Psi_{H_l}(i,j) + \Psi_{V_l}(i,j) + \Psi_{D_l}(i,j)}{3} \tag{3.58}$$

Then the normalized metric value after the fused image is processed with all input images is:

$$Q_l^{AB/F} \frac{\sum_i \sum_j \left(EP_l^{AF}(i,j) w_l^A(i,j) + EP_l^{BF}(i,j) w_l^B(i,j)\right)}{\sum_i \sum_j \left(w_l^A(i,j) + w_l^B(i,j)\right)} \tag{3.59}$$

$$w_l^A(i,j) = LH_{A_l}^2(i,j) + HL_{A_l}^2(i,j) + HH_{A_l}^2(i,j) \tag{3.60}$$

$$w_l^B(i,j) = LH_{B_l}^2(i,j) + HL_{B_l}^2(i,j) + HH_{B_l}^2(i,j) \tag{3.61}$$

The final metric is measured with an exponential value α_i, this value is used to adjust different relative significance to the different scales.

$$Q_{EP} = \prod_{l=1}^{N} \left(Q_l^{AB/F}\right)^{\alpha_l} \tag{3.62}$$

3.6.5. Yang Metric

Yang et al. (Yang et al., 2008) have proposed a quality metric that measures structural similarity. The approach of the metric is the same as in Piella, but a threshold value is used in this metric. With this threshold, it is postulated that the distinction between noise and real information in images is made. This threshold value is taken as 0.75 in the article. The metric's result is decided as follows:

$$\begin{cases} \lambda(w) SSIM(A,F|w) + \left(1-\lambda(w)\right) SSIM(B,F|w), & SSIM(A,B|w) \geq 0.75 \\ \max\{SSIM(A,F|w), SSIM(B,F|w)\} & , \quad otherwise \end{cases} \tag{3.63}$$

This study improved the similarity measure with a simple approach without disturbing its basic structure.

3.6.6. Chen Blum Metric

The authors proposed a new metric in 2009 (Chen and Blum, 2009). Image processing applications try to examine images as if they were being processed by the human eye. The human visual system is sensitive to contrast. With this approach, they remark to the contrast values of the images. The metric calculation consists of multiple stages.

All transactions consist of five main categories. First of all, one of the Contrast Sensitivity Filters (CSF), such as Mannos & Sakrison, Barton and Dog, etc., is applied to the images that are subjected to the Fourier transform as in the previous subtopic. $I_A(x,y)$ is the resulting image of Fourier Transform. S(r) is one of the CSFs. And the filtered result is the $\widetilde{I_A}(x,y)$ of an image A in (x,y) coordinates. Here is the Equation (64):

$$\widetilde{I_A}(x,y) = I_A(x,y)S(r) \tag{3.64}$$

In the second step, the local contrast maps of the filtered images are found. There are several computational methods in the literature, including Root Mean Square (RMS) contrast measure (Raj and Geisler, 2005), a local-band measure (Alenezi, 2018). After obtaining the contrast maps C_A', C_B', C_F' of source images A, B, and fused image F, it is measured how much contrast information is transmitted from the source image to the fused image, in other words, how much of this information can be preserved. According to Equation (65)., masked maps are generated. Then, the degree of preservation is measured as in Equation (66).

$$C_A' = \frac{k(C_A)^p}{h(C_A)^q + Z} \tag{3.65}$$

$$Q_{AF}(x,y) = \begin{cases} \dfrac{C_A'(x,y)}{C_F'(x,y)}, & C_A'(x,y) < C_F'(x,y) \\ \dfrac{C_F'(x,y)}{C_A'(x,y)}, & otherwise \end{cases} \tag{3.66}$$

Where k, p, h, q, and Z are static parameters. It produces a value between ‘0’ and ‘1’, while ‘1’ indicates that all information of A is transferred to the fused image, ‘0’ means that no information is carried. Finally, the global metric value is calculated for each pixel as follows:

$$Q_c(x, y) = \lambda_A(x, y)Q_{AF}(x, y) + \lambda_B(x, y)Q_{BF}(x, y) \tag{3.67}$$

In order to reach the last metric value, the values in the above equation are averaged.

3.7. Proposed Method

3.7.1. Overview

In order to create more robust and accurate fused images, preprocessing of source images can be functional. These processes may be beneficial to make easier decisions and achieve more accurate results. We decided to use two features that are highly emphasized in the literature and in our prelimanary works; image contrast enhancement and edge information.

Contrast enhancement as a pre-processing operation has been applied to the source images to make the image details more prominent. To accomplish this, we use histogram equalization method. Histogram equalization ensures regular distribution of image information and makes the image more convenient to capture details in the latter processes.

Since using only contrast-enhanced source images for image fusion is a straightforward approach, we make use of another important feature; edge information. As a natural result of edge detection, trivial information will be discarded and the remaining information, the details, can be treated as a summary of the original source image. There are numerous edge detectors such as Canny, Sobel, Laplacian, LOG, etc. The fusion algorithms that use edge information, employ one of these traditional edge detectors to extract edge information.

Contrary to the current literature, we propose to employ an unconventional method to extract gradient information. The gradient information is conventionally extracted from local information preserved in an image. However, we get the gradient information from a pair of images that are the source image and its counterpart that is obtained by applying Halftoning-Inverse Halftoning (H-IH)transform consecutively.

Surprisingly, the gradient information that is calculated by taking the diference between the source image and its H-IH tranformed counterpart is richer for fusion purposes than the one that is provided by conventional edge detectors. Figure 3.33 illustrates the parts of the proposed method described so far.

Detection of in focuse and out of focus pixels is achieved by applying Energy of Gradient as focus measure function. Higher values of gradient energy are the indicator of in focus pixels. However, in some circumstances, the result of the focus measurement of the compared pixel blocks may be equal. In such a case, the secondary focus measure method is determined to avoid losing information. Standard deviation is selected as the second focus measure.

The success of the proposed method relies on the correct decision fusion mechanism. In some cases, it is possible to make the wrong decisions in the selection mechanism. Many practical studies in the literature implement a decision fusion operation to minimize this error. Decision fusion operation makes the proposed algorithm more reliable and more robust against errors. In Figure 3.34, the flow of all steps is given for a better understanding of the proposed method. The details of the proposed algorithm is provided in the following sub-sections.

Figure 3.33. A sample for the generation of gradient images.

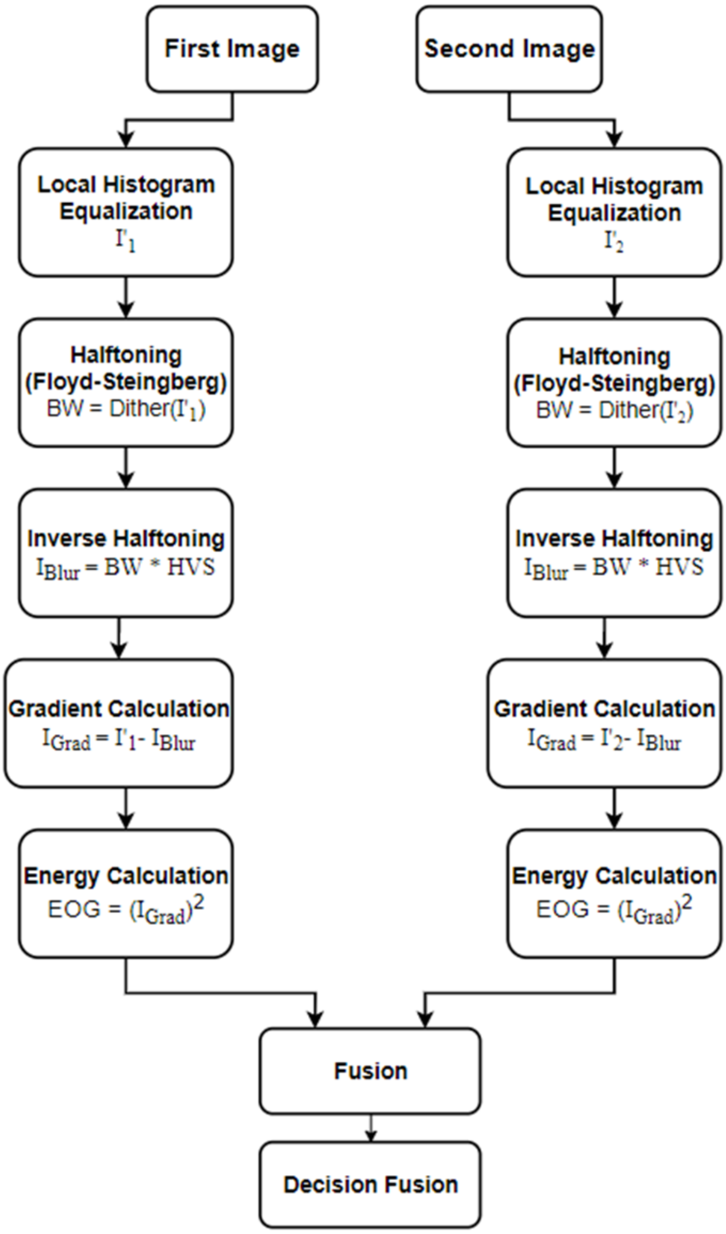


Figure 3.34. The flow of the proposed method.

3.7.2. Local Histogram Equalization

In the fusion process, we can apply to the pre-processing to get more effective results. In this study, we provide image enhancement with histogram equalization. This enables the contrast of the image to be increased and the salients are revealed. Due to the disadvantages of global histogram equalization as described in section 3.2, we have preferred local histogram equalization. For the implementation, the window size has been decided empirically. In other words, window sizes in different ranges are tried and the size that gives the best results is determined. An exemplary result of histogram equalization is given in Figure 3.35. As can be seen from Figure 3.35, the contrast enhancement makes the details become clearer and the contours of the image appear. Thus, while the focused pixels become more prominent, the out of-focus ones continued to have less sharpness. Simply the original image (I) is LHE treated as in Equation (68) and I' is obtained.

$$I' = LHE(I) \tag{3.68}$$

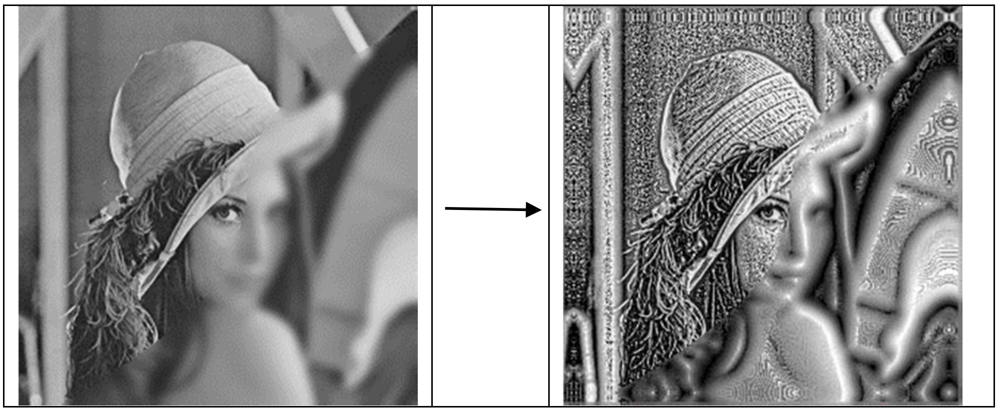


Figure 3.35. The result of Local Histogram Equalization.

3.7.3. Halftoning

Gradient information, which is regarded as another important feature for the fusion process, is obtained by the implementation of the halftoning. Among the halftoning methods, we preferred the Floyd-Steinberg Error Diffusion Method which is an effective and highly preferred method.. As explained in the general description, when the difference between the inverse halftoned image and the original one is taken, the gradient image is obtained. On the basis of these approaches, the first step of acquiring the gradient image is dithering. Floyd-Steinberg Error Diffusion is applied to the image (I') and halftoned image (BW) is obtained. It is illustrated in Equation (69). Figure 3.36 shows the halftoned image (BW) after applying to halftone.

$$BW = FloydSteinberg(I') \tag{3.69}$$

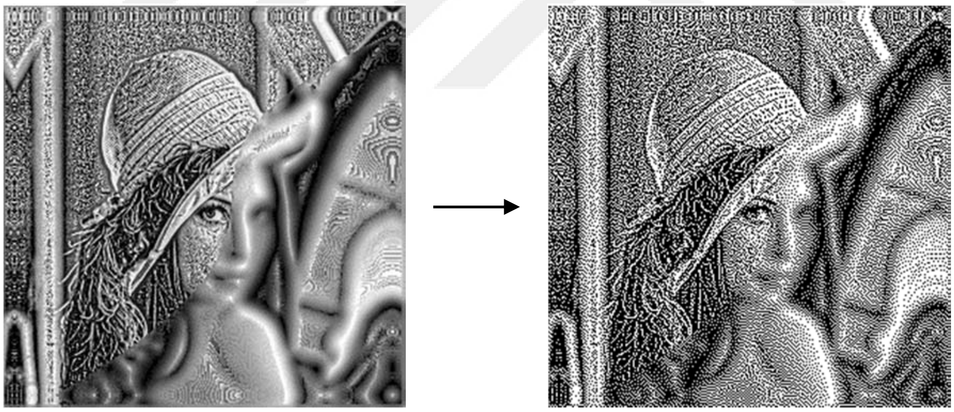


Figure 3.36. The result of the Floyd-Steinberg Error Diffusion.

3.7.4. Inverse Halftoning

We proceed with the inverse halftoning technique to access the edge information. Although we try to capture the closest version to the original source image, it is very difficult to capture the image's details.

As described in the relevant section, the inverse halftoning technique can also be seen as a filtering process. However, in order to do this filtering process, a suitable kernel must be created. Here, instead of theoretically observations, we can focus on the behavior of the real human eye. We use the Human Vision System (HVS) for this. As the distance to images increases, noise and artifacts become indistinguishable along with the details in the image. In fact, eliminating noises is intended to be performed in inverse halftoning. For this purpose, we have performed inverse halftoning with the HVS filter model. There are various HVS models. One of the most similar models of the human visual system is the Gaussian HVS model. Gaussian HVS is described by spatial frequency response also called a Modulation Transfer Function (MTF). Its formula is given in Equation (70). The inversed image is generated from the convolution of the dithered image by the HVS filter.

The process can be summarized as follows; firstly, the dithered image (BW) is generated by Floyd-Steinberg. Then the inverse halftoning method is applied and an inversed image (Iblur) that is close to the original image is achieved as in Equation (71). Inverse Halftoning example is given in Figure 3.37. Finally, when the difference between the image (I') and inverse halftoned (Iblur) imageis taken, the gradient image (IGrad) is calculated as in Equation (72), and presented in Figure 3.38.

$$h = e^{-(x^2+y^2)/2\sigma^2} \quad (3.70)$$

$$IBlur = BW * h \quad (3.71)$$

$$IGrad = I' - Iblur \quad (3.72)$$

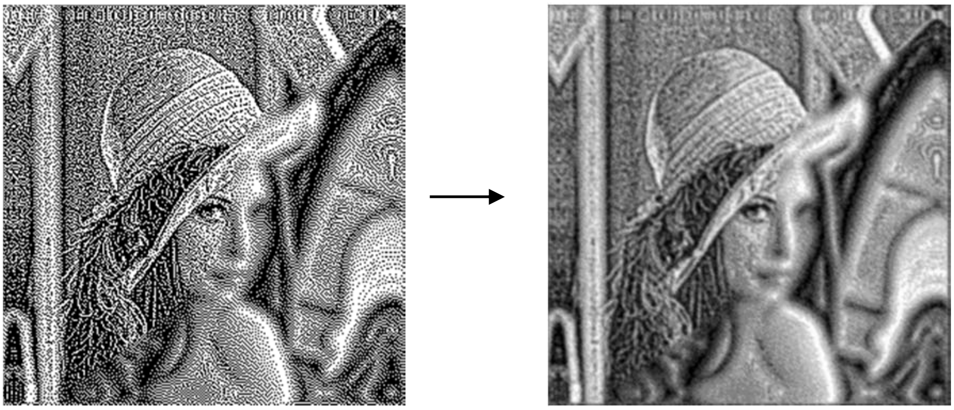


Figure 3.37. The result of the Inverse halftoning.

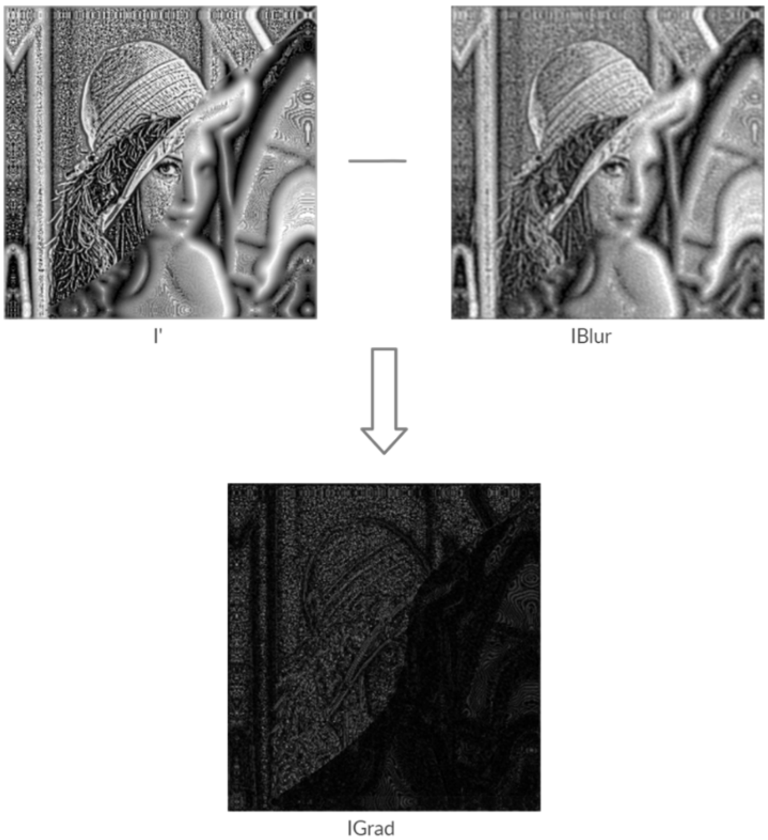


Figure 3.38. The creation of the gradient image.

3.7.5. Focus Measure

After obtaining the gradient image of each source image, the pre-processing of the images is completed. The next step is to perform focus measurement in order to perform the fusion process. There are various methods and functions for focus measurement. The methods from the literature that are already described in detail in the literature review section are selected to be used in fusion process. Generally, preferred functions are given in section 3.5 with their formulas. In this study, we utilize the Energy of Gradient Function (EOG). Energy evaluates how harmonious the gradient and image data. It is specified as the measure of the gradient. The gradient images are processed block by block. The EOG of the respective block of each source images is calculated and the block that has the most energy is directly placed in the fused image. The illustration of the fused gradient image creation is given in Figure 3.39. Please note that colors are inverted to allow better viewing for the printed document. In the equation, i and j are coordinates of source images I_1 and I_2.

$$Fused(i,j)\begin{cases} \mathrm{I}_1(i,j), EOG(I_1) > EOG(I_2) \\ \mathrm{I}_2(i,j), \qquad otherwise \end{cases}$$

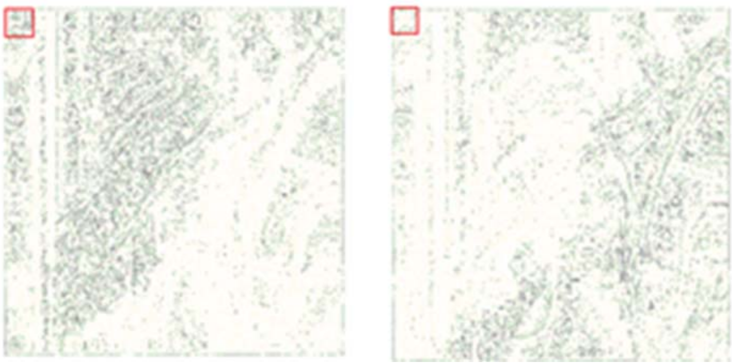


Figure 3.39. The creation of the fused gradient image according to the focus measure function.

3.7.6. Second Focus Measure

When pre-processing is performed to make the gradient information in the images more pronounced, loss of information may occur in the images. In this case, the energy function can be zero when focus measurement is performed with the energy function. Another possible situation is that the energy values in the collated blocks of the images are the same. In such cases, an alternative focus measurement method is required.

When we investigate the focus measurement methods, we can give some thought to the standard deviation value. In fact, the standard deviation is a measure of how values are spread. As the distribution rate of the data increases, the deviation also increases. Such distributions result in more detailed images. Therefore we can relate standard deviation with focus measure. The loss of detail in blurred images indicates that the standard deviation of blurred parts is lower than the focused ones. Since standard deviation is a distinctive metric, it has been considered as an alternative focus measurement in this study. As we cannot obtain energy from gradient images, we approve that it is a good way to use the original images in the second focus measurement method. Instead of using the raw image, the augmented version of the raw image which is produced by applying the local histogram equalization method on the raw image is used.

3.7.7. Decision Fusion

The image blocks are compared by using the focus measurement function and the block that will be included in the fused image is selected. But it is possible to make a wrong choice. Therefore, a step we call decision fusion, which is supported by studies in the literature, is added to the algorithm. Decision fusion is an added step to correct wrong choices. The result makes the fused image more accurate and powerful.

Before making decision fusion, a fused map is generated from the indices that correspond to source images pixels. The decision mechanism on this map is applied. In order to accomplish decision making, we applied three different filters with 2x2, 4x4, 8x8 window sizes. Each filter has been used to apply a pooling operation. After all three filters have been applied, the majority voting (Sewell, 2011) has been used to combine the results of filter processing.

4. RESULTS AND DISCUSSIONS

4.1. Tests

In order to evaluate the success of the proposed method, seventeen different methods are compared. Among these methods, relatively state of art ones; Traditional for MSSF, LP-PCNN, NSCT, CSR, CBF, DWTDE, GF, MWGF, UC, Std, and traditional ones; PCA, Averaging, DCT, DWT, Laplacian Pyramid, Ratio Pyramid are used. These methods are explained in detail under the literature section.

Thirty three image pairs are used to examine these methods. twenty one pairs of these images are taken from the Lytro Dataset (Nejati et al., 2015), and the others consisted of commonly used images for fusion studies. Samples of these test images are given in Figure 4.1. In order to make the experiment proper for all methods, the examination is carried out by reducing images to and 256x256 size.

Evaluation is performed in two ways; visual and objective. In the visual evaluation, samples of fused images are given and their error images that are formed by subtracting fused and source images are presented. Visual evaluation between methods is conducted using fused and error images. In objective evaluation, the comparison of methods is performed quantitatively by using six different metrics that can evaluate different characteristics of images.

In addition, the intermediate results are summarized and illustrated with all the steps of the proposed method.

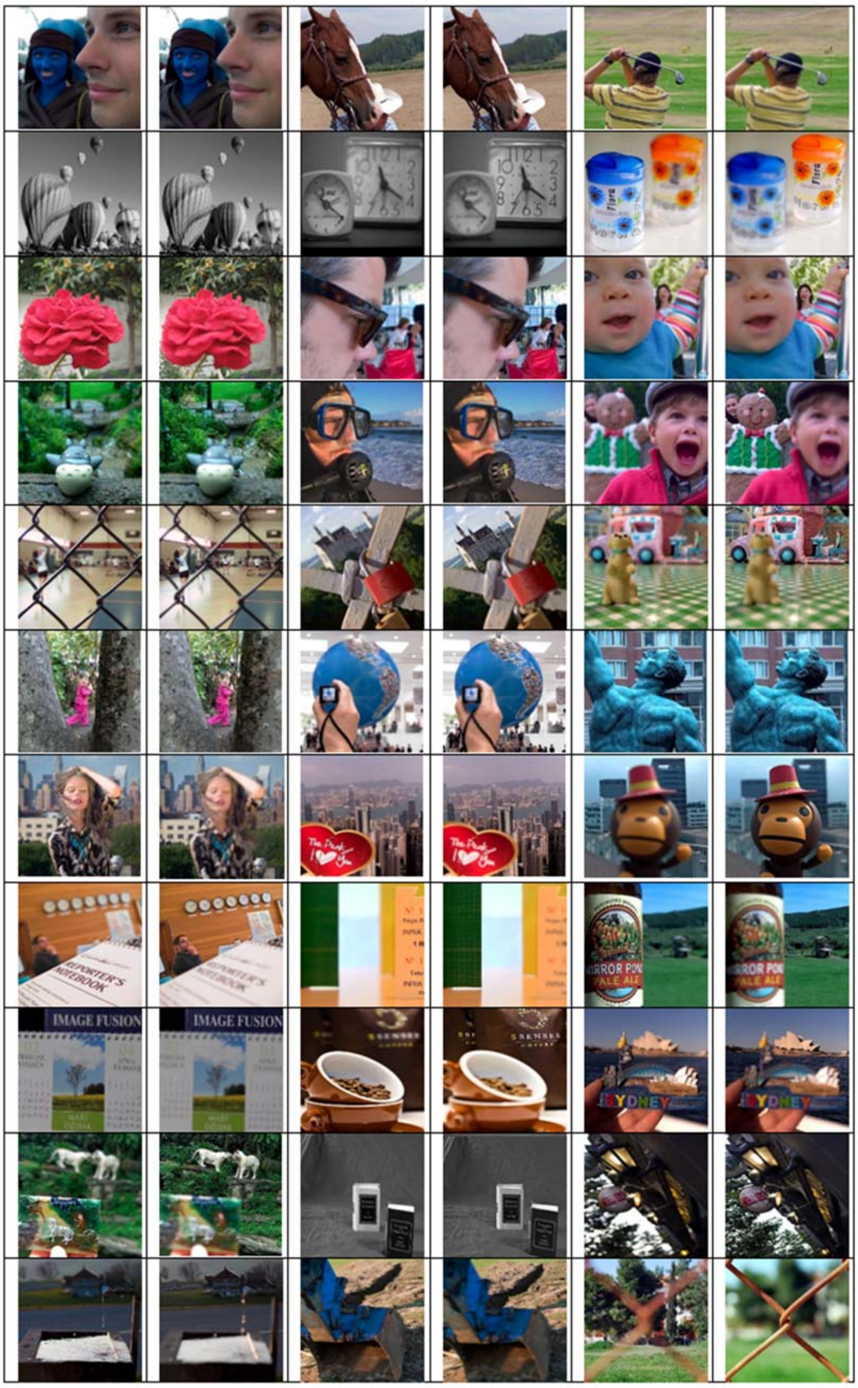


Figure 4.1. Samples of test images.

4.1.1. Intermediate Results of the Proposed Method

In Figure 4.2, the results of each step are shown on the pictures selected over the test set of algorithm steps detailed in the previous sections. The first column shows the result of the Local Histogram Equalization that provides to reveal the image details more clearly and to enhance the image. The gradient image that is generated by differing the inverse halftoned and the original images, is shown in the second column. In this image, it is seen that while the focused pixels are clear and distinguishable, unfocused pixels are lost. Therefore, blocks containing excessive information can be called focused blocks. The focus measurement function is applied to the gradient image and the initial fusion process is carried according to the selected test images' blocks. The initial fusion map can be seen in the third column. The focus measurement method may make the wrong decision for some pixels, thus causing the blurry pixels to appear in the fused image. For this reason, to make the image more robust and reliable, the decision fusion operation is conducted in many studies (Liu et al., 2017; Yang, 2017). In the decision fusion step, the majority voting method is performed so that incorrectly grouped pixels assort with around them. The decision maps of the samples are given in the fourth column. Finally, the fused images can be seen in the fifth column.

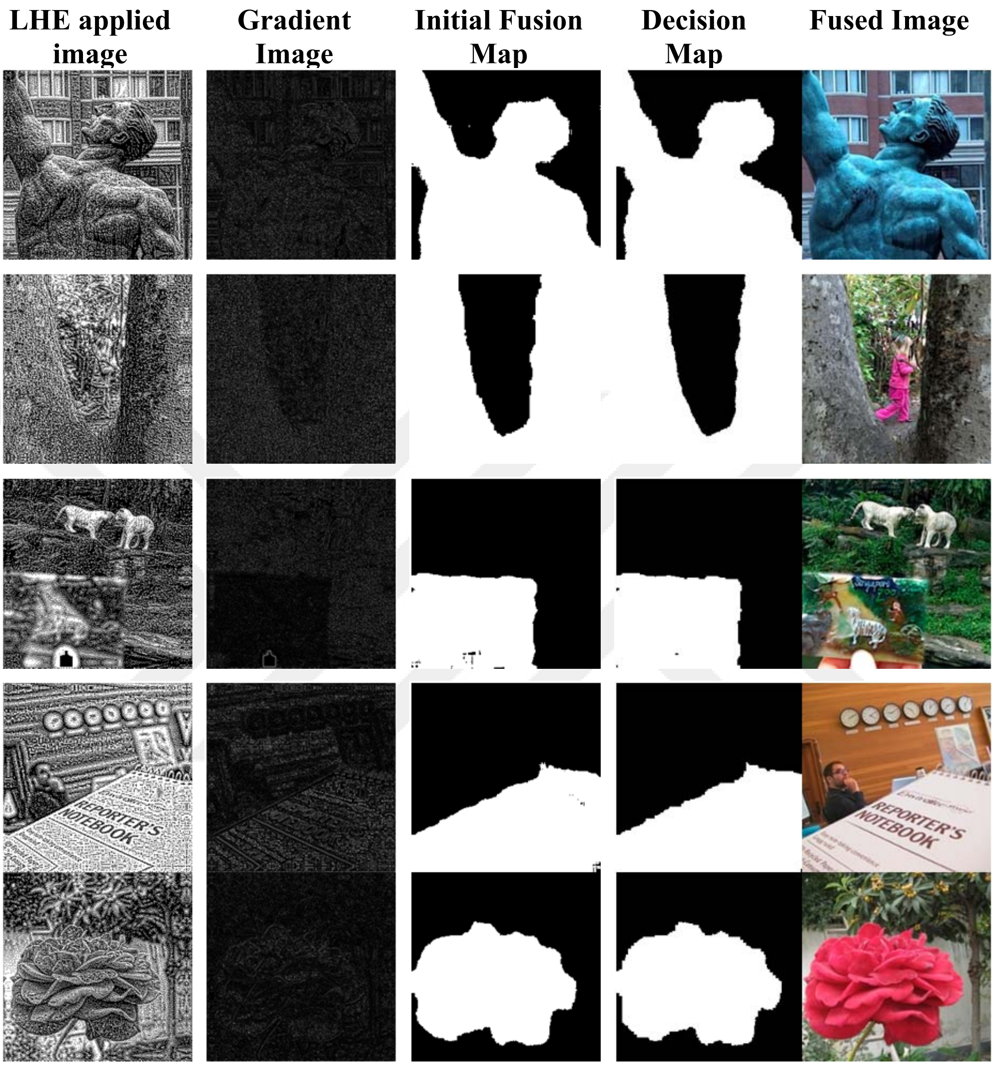


Figure 4.2. The intermediate results of test images.

4.2. Visual Comparisons

First of all, we compare the proposed method with its counterparts visually. Two sample test images are used to make the visual comparisons. Many inferences are made by using the fused images of all methods and the error images obtained by subtracting the first source image from the fused images.

The first sample is Flora Image, fused images are shown in Figure 4.3 and error images are displayed in Figure 4.4. The methods are examined in details separately.

In MSSF, the fused image has never close to the correct result and this method produces too many artifacts.

GF, DWT, PCA and Averaging methods' error image results indicate that these methods engender hazy images. Defective pixels appear in almost all of the fused image so these methods underperform compared to others.

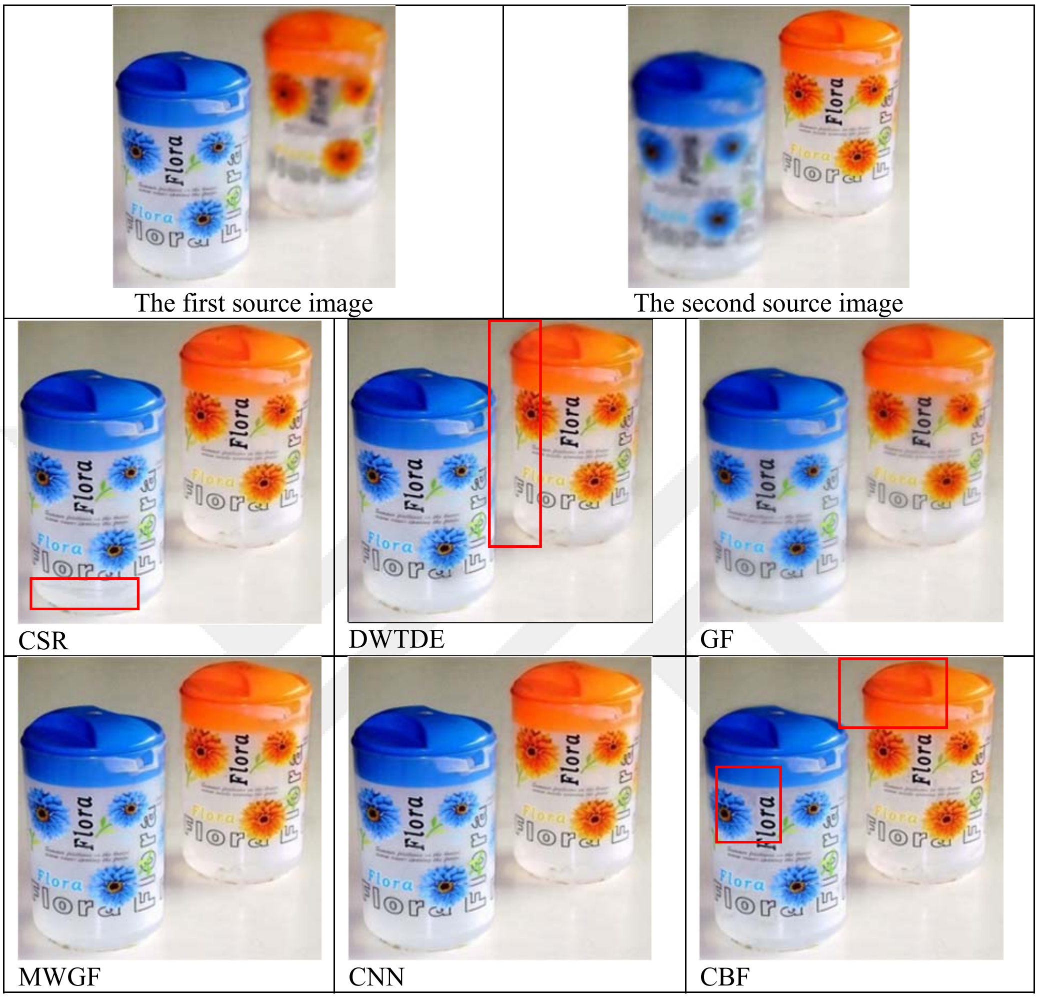
In Standard Deviation, Unique Color, and DCT methods, blocking-effects, and spatial distortions have widely appeared in their results. Blocking-effects are the appearance of block boundaries in the fused image due to the wrong selection of pixel blocks. The blocking effect is prominent because of the incorrect functioning of the selection mechanism of these methods.

In the fused images of CBF and DWTDE methods, local errors have been observed. Error examples are marked with red rectangles in the resulting images. Although CBF and DWTDE methods work correctly on text, they have caused artifacts in the images. CBF is unable to capture the border of the front object. DWTDE, like the CBF method, causes bright shadows and makes distortions at the border of the object.

In spite of the fact that the Laplacian Pyramid, Ratio Pyramid, and NSCT methods can capture the outlines, these methods lead to similar unsuccessful results with less qualified images. Erroneous pixels can be seen along with the whole of the objects on the fused images. As can be understood from their error images, the fore object is visible in the form of a silhouette due to failures. The Ratio Pyramid method has incorrect results on some specific local parts of the image where text pattern exists and this method has less promising results than the other two methods.

LP-PCNN method also works incorrectly in the local regions and caused fluctuation, as can be seen in the error image.The CSR method, on the other hand, can produce well-focused images. But it still cannot be considered to work perfectly, as it produces some errors. Because the fused image of the CSR method has errors around the edges and under the "Flora" text.

The fused images on both CNN, MWGF, and the proposed method are pretty much similar. Composed images that are generated by these methods are successful and outperforms the other methods. In the fused image of CNN method, artifacts are clearly visible on the right edge of the front bottle. The result of MWGF is visually similar to the proposed method's result with small differences. These differences can be seen at the bottom side of the back bottle. It can be said that the proposed method gives a more comprehensive result.


The first source image
The second source image
CSR
DWTDE
GF
MWGF
CNN
CBF

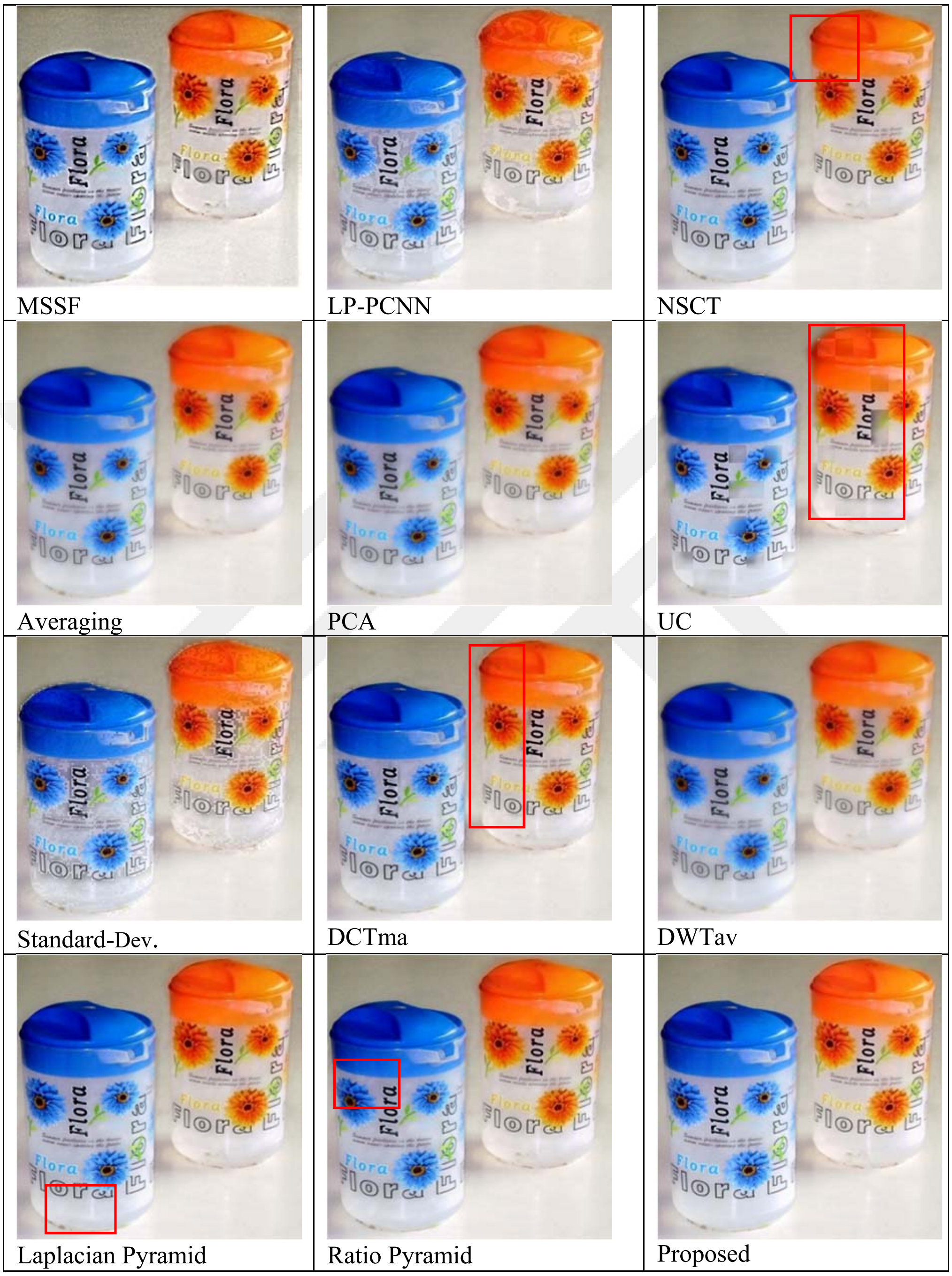


Figure 4.3. Fused images of all methods for Flora Image.


CSR


DWTDE


GF


MWGF


CNN


CBF


MSSF


LP-PCNN


NSCT


Averaging


PCA


UC

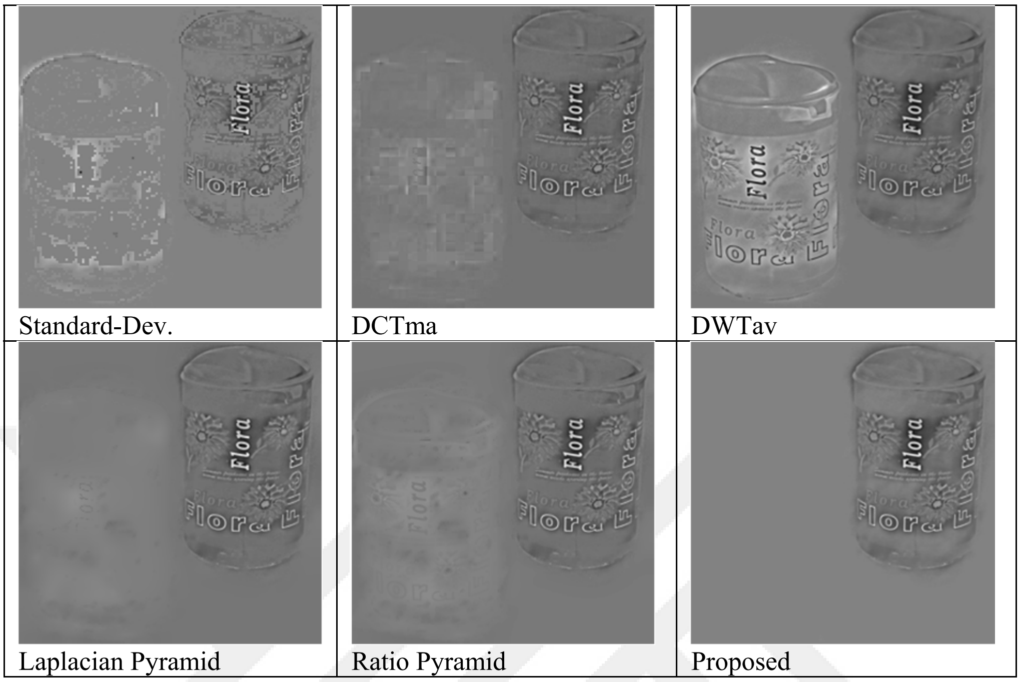


Figure 4.4. The error images of all methods for Flora Image.

Heart Image is the second example. Fused images of all methods are demonstrated in Figure 4.5 and error images are located in Figure 4.6.

The fused image of the MMSF method doesn't work correctly in the almost whole image and it can be said that this method gives the worst results compared to other methods. In addition, the errors are substantially high on the pixels. Another non-promising result is the DWTDE method. This method works incorrectly for both the border and the texts of the heart object and it produces noises. But the error is not high as much as in the MSSF method.

DWT, Averaging, GF, and PCA methods have hardly captured the details of images. When the error images are reviewed, the details of the images are obviously visible. This leads to unclear fused images.

The CBF method produces an unclear image, as methods mentioned above, it does not work correctly for both the borders and the texts. However, it does not generate excessive noise according to these methods. Nevertheless, it presents an inadequate result.

Since the UC, Std and DCT methods are block-based algorithms, they have caused the blocking effect in the image as appeared in the Flora Image. Although Std gives better results than the other two methods, it fails to compose a well-performed fused image due to the wrong selection.

When the fused and error images of Laplacian Pyramid and NSCT methods are viewed, although they perform well for the texts, they can not distinguish patterns on the heart shape. Therefore, fused images of these methods seem to be inadequate. As for the fused image of the Ratio Pyramid method, it has appeared that the error image has faulty pixels on the text and the frame of the heart shape. The resulting image is close to a formal successful outcome. But, it cannot be one of the best-performing methods.

LP-PCNN method causes noises and the details of the heart shape can't be specified as seen in the error images. CSR largely fails at the boundaries of the heart object. CSR crank success up a notch but it doesn't seem to be successful enough.

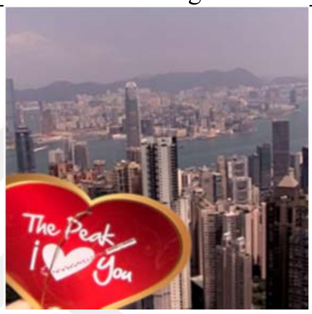
As in the flora image, MWGF, CNN, and the proposed method yielded the best performance results for this sample image. In the error image of the MWGF method, composed errors within the border of the heart shape cause to form shadows. This affects the quality of the fused image. CNN and the proposed method produce nearly impeccable images except for the frame of the heart shape. According to the error images, it can be seen clearly the proposed method has fewer errors. The proposed method has produced more precise fused image.

The first source image

The second source image

CSR

DWTDE



GF

MWGF

CNN

CBF

MSSF

LP-PCNN

NSCT

Averaging

PCA

UC

Standard-Dev.

DCTma

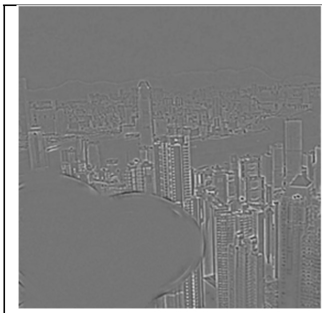
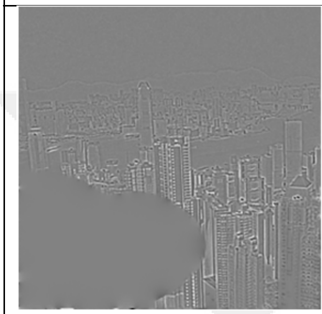
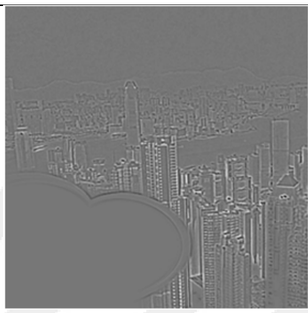
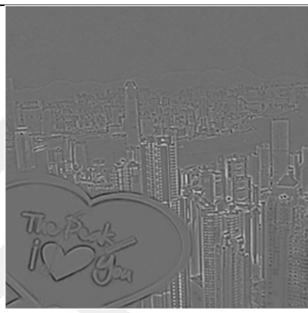
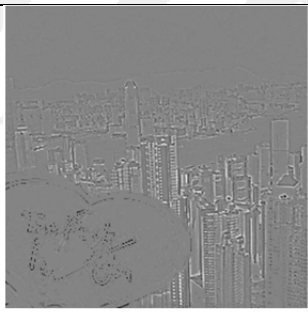
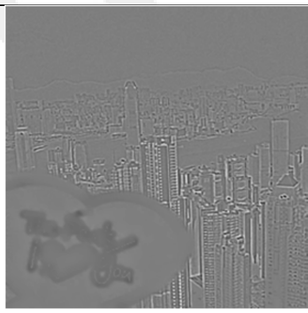
DWTav

Laplacian Pyramid

Ratio Pyramid

Proposed

Figure 4.5. Fused images of all methods for Heart Image.

 CSR	 DWTDE	 GF
 MWGF	 CNN	 CBF
 MSSF	 LP-PCNN	 NSCT
 Averaging	 PCA	 UC

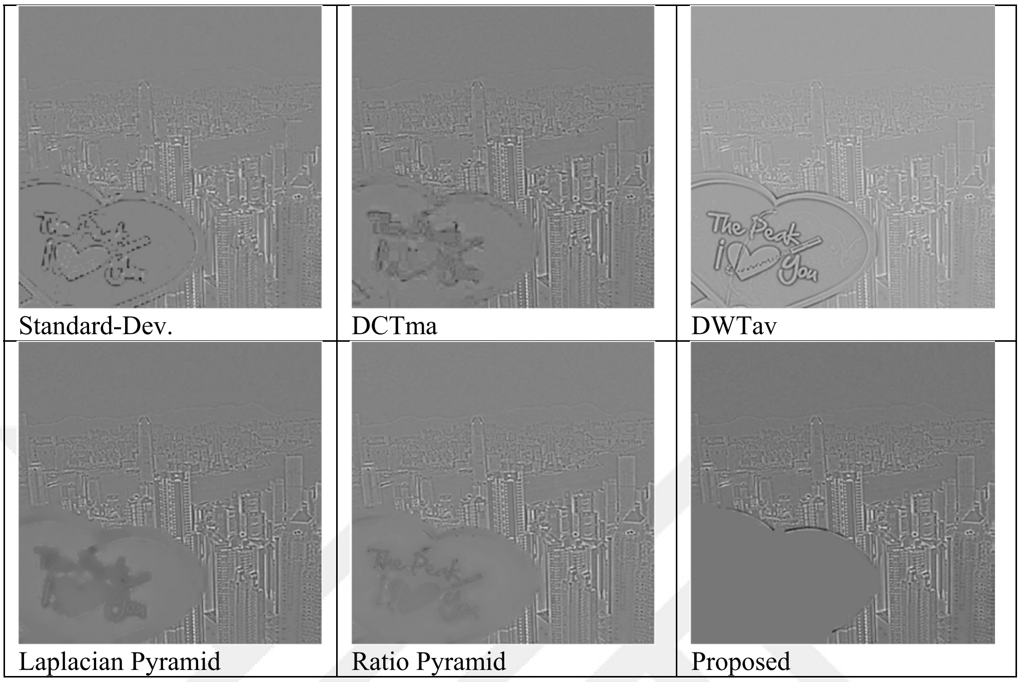


Figure 4.6. The error images of all methods for Heart Image.

It is difficult to obtain registered multi-focus images. External factors such as camera motion and wind cause the formation of misregistered images. In the past, these misregistered images were first registered and then fused. But nowadays, newly developed algorithms are expected to be resistant to these misregistered images. Therefore, test datasets include these images.

The results of an example of misregistered images are examined, fused and error images for all methods are given in Figure 4.7 and Figure 4.8.

Among all methods, MWGF, CSR, and CNN stand out with the proposed method. The common feature of these methods is that they are shift-invariant. In this way, they are robust for misregistered images. They can produce meaningful results for unknown regions that are caused by misregistration. These methods make the most mistake at the border of the front-object. Because the most affected area from the misregistration is the border of this object. It can be said that MWGF and the proposed method have both the most successful and the closest results.

While MWGF generates errors on all sides of the object in front, the proposed method produces more intense errors on two sides. While the errors caused by the MWGF method are softer, the proposed method is sharper. Based on all this, it is seen that the proposed method produces successful results even though the image is not registered. So, it can be seen that the proposed method is robust to anisotropic blur.


The first source image


The second source image


CSR


DWTDE


GF


MWGF


CNN


CBF


MSSF


LP-PCNN


NSCT

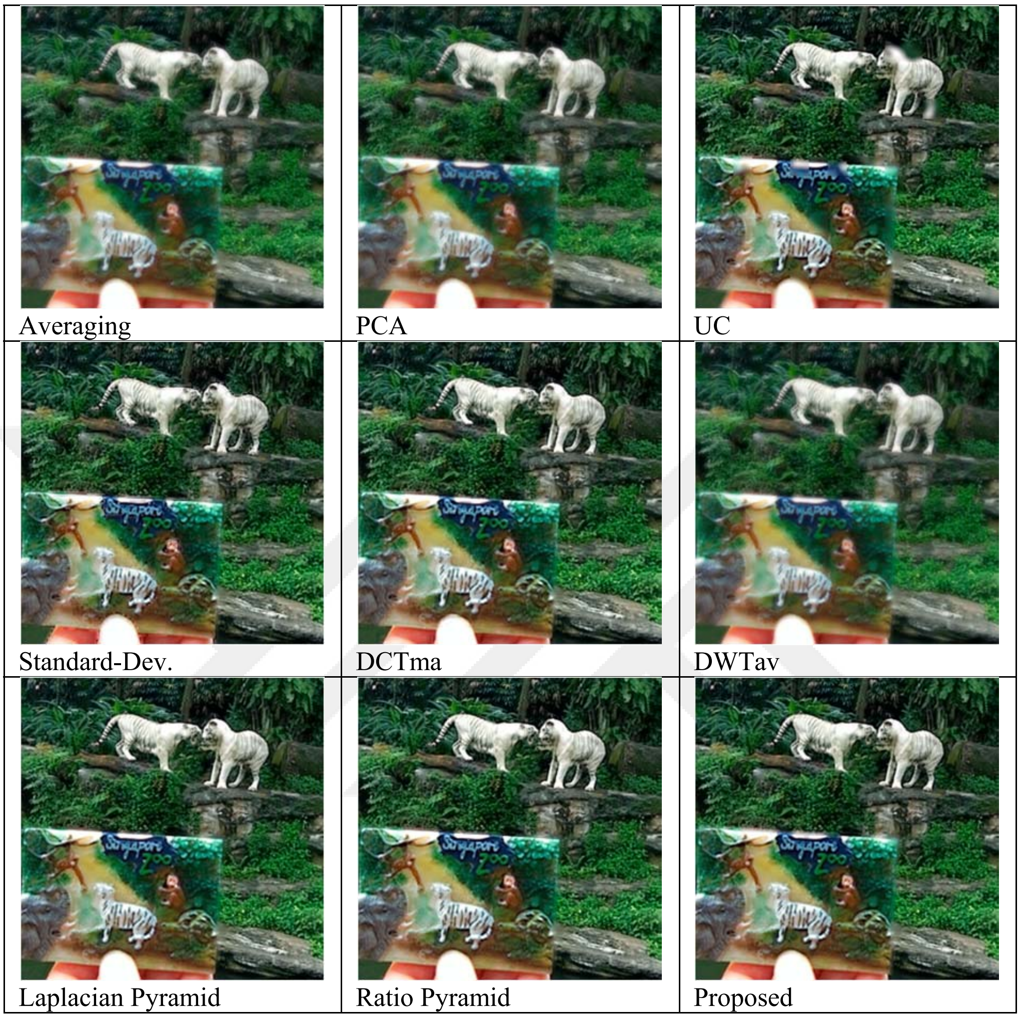
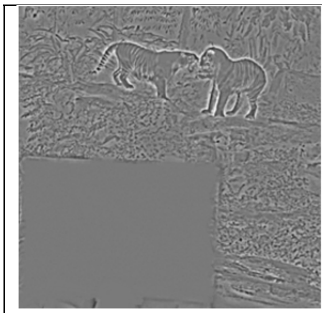
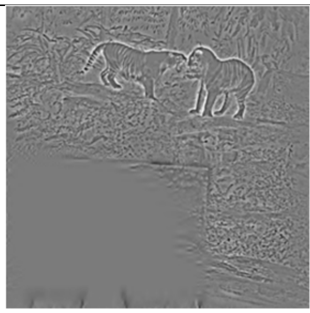
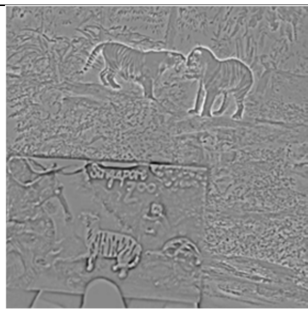
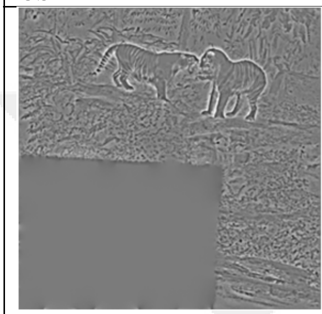
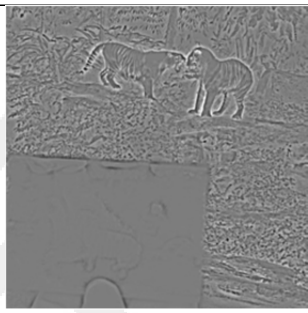
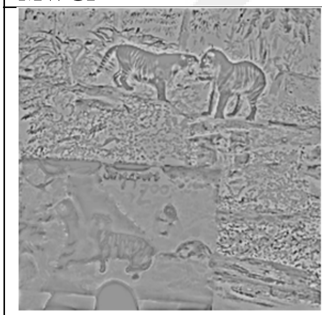
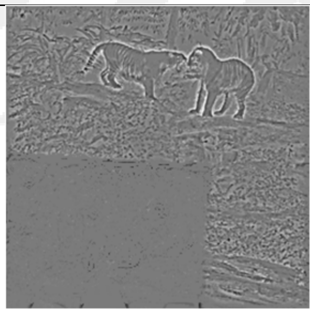
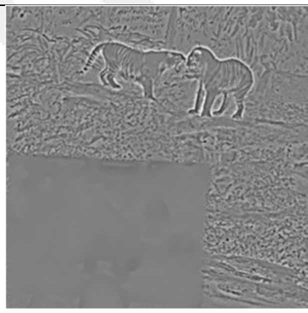
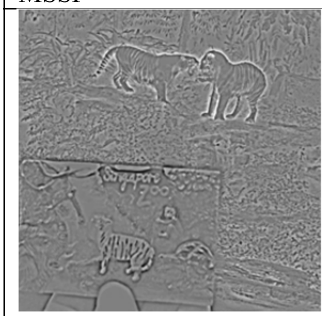
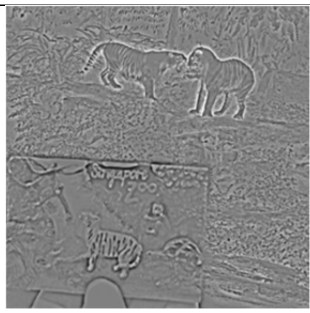
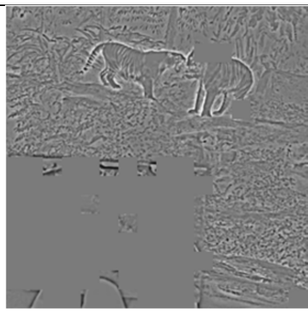
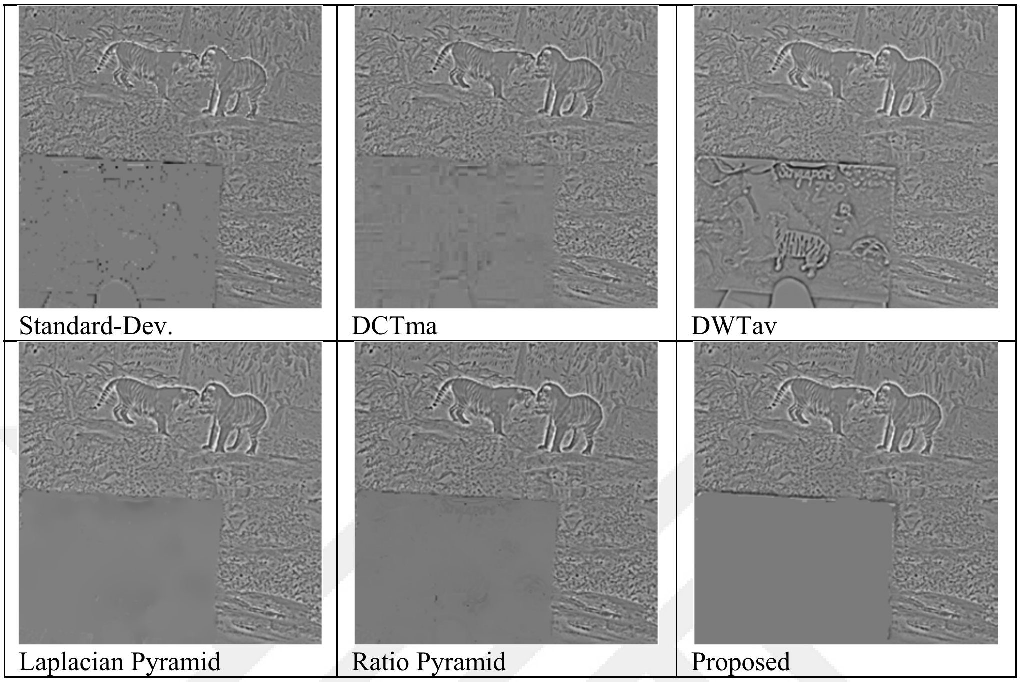


Figure 4.7: Fused images of all methods for a misregistered image.

 CSR	 DWTDE	 GF
 MWGF	 CNN	 CBF
 MSSF	 LP-PCNN	 NSCT
 Averaging	 PCA	 UC



Standard-Dev. | DCTma | DWTav
Laplacian Pyramid | Ratio Pyramid | Proposed

Figure 4.8. The error images of all methods for a misregistered image.

4.3. Objective Comparison

Objective evaluations have been performed by using six well-known metrics; Normalized Mutual Information (NMI), Chen Blum, Peng-Wei Wang Metric (PWW), Xydeas and Petrovic Metric (Qabf), Nonlinear Correlation Information Entropy Metric (NCIE), and Yang Metric. The average results are obtained by taking the results of each image separately and considering the total number of images. These average results are shown in Table 4.1 and the metrics are described in detail previously.

According to the objective results; apart from the Chen-Blum metric, the proposed method is superior to the other methods for the five metrics. In general, when the metric results are analyzed, it is seen that the proposed method and CNN method outperform other algorithms.

MSSF, DCT, DWT, and GF methods have failed for objective evaluation similar to visual outcomes. MSSF produces the worst results for all metrics and DCT is insufficient and leads to a blocking effect. DWT is not successful in motion areas for misregistered images because it is not shift-invariant. GF is not good for capturing the details on the borders of the objects.

Averaging, PCA, and NSCT methods have followed these methods and failed to produce satisfying results. The NSCT method shows an increase in the Chen-Blum metric, but this is not regarded as a great success according to other methods. Because like most of the fusion methods, it can not be successful in motion areas.

DWTDE, CSR, and CBF produce similar results and approach intermediate level success.

The results of UC, Std, LP-PCNN, CNN, and MWGF methods are similar to the proposed method. Although the UC and Std methods give reasonable results for objective evaluation, they are visually failed because they cause the blocking effect. The fused images are imperfect imagery such as the examples given in the visual evaluation section. LP-PCNN is also a method that cannot be considered the most successful method when both visual and objective evaluations are taken into account.

The proposed method's objective evaluations in all of the metrics (except Chen-Blum) are superior to its rivals'. CNN and MWGF have accomplished the closest results to the proposed method. Although the MWGF's success in visual evaluations is close to the proposed method, its objective evaluations is even worse than CNN. CNN's objective evaluations reveal that it is almost as successful as the proposed method in almost every metric. However, CNN's objective evaluation with Chen and Blum metric is slightly better than the proposed method with a small fraction. However, we should bear in mind that the proposed method in other metrics is superior against CNN and others.

Table 4.1. The experimental metrics' results.

	NMI	NCIE	Qabf	PWW	Yang	Chen-Blum
Proposed	**40,92764**	**28,03751**	**23,73657**	**77,00543**	**32,24719**	26,65982
CSR	33,53968	27,53004	20,96648	36,34235	30,24071	25,30441
DWTDE	33,51523	27,56438	21,52673	50,92057	30,82045	25,34033
GF	30,31577	27,36552	20,10963	11,86925	29,00172	23,58846
MWGF	36,63057	27,73089	23,09858	66,20133	32,15072	26,22907
CNN	38,48345	27,85032	23,66498	66,07441	32,23399	**26,69238**
CBF	33,85731	27,56256	23,20546	31,14054	31,49232	25,4419
MSSF	20,21257	26,97973	13,02013	5,654819	25,01741	18,40857
LP-PCNN	39,35598	27,85958	20,81643	34,93953	29,54071	24,40773
NSCT	31,91414	27,47147	22,63006	49,61661	31,40976	25,54019
Averaging	30,35054	27,36756	20,15235	11,94378	28,99407	23,6964
PCA	30,35054	27,36756	20,15235	11,94378	28,99407	23,6964
UC	40,14773	27,96131	23,56802	69,36395	31,77149	25,81093
Std	39,60735	27,90306	21,92388	64,37132	29,8204	24,06993
DCTma	29,1399	27,3275	20,45215	35,30037	30,1604	23,98114
DWTav	30,26303	27,36288	19,44294	11,82511	28,5054	23,58327
Laplacian	32,46495	27,50144	22,83046	59,45482	31,41163	25,54083
Ratio	32,71287	27,51203	22,59878	38,23684	31,30143	25,48736

5. CONCLUSIONS

Completely focused images are necessary for many fields such as medicine, biology, computer science and, etc. Due to the limited depth of field feature of optical lenses, it is difficult to obtain a fully focused image. Researchers suggest multi-focus image fusion methods into this problem. The multi-focus image fusion methods combine the imagery of the same scene taken at different focal lengths and aim to achieve a fully focused fused image.

The proposed method can capture details, perform well in moving areas, and work effectively. Local histogram equalization uses to enhance the images. The dithering method helps to capture the details. The Energy of Gradient function takes a part in as a focus measure. Thanks to the majority voting approach, the fused image has been strengthened.

The proposed method is compared with seventeen different methods to prove its performance. Experiments are performed using thirty three pairs of images. The methods are examined and compared visually and objectively.

According to the visual results, it is seen that

- The proposed method produces near-perfect fused images.
- The CNN and MWGF methods have produced the most similar results visually to the proposed method.
- UC and Std methods produced the excessive blocking effect in the fused images
- CBF has a medium success in visiual evaluations. This method is unable to capture the outlines of most of the images.
- MWGF is very close to the proposed method, but it has difficulties to capture the borders of the images.

Six different quantitative metrics are used for objective evaluation. Each metric focuses on different features in the images for evaluations. The conclusions from the objective evaluations are as follows;

- The proposed method is the most successful method in five of the six metrics.
- CNN is the most successful method with a slight difference in the Chen-Blum metric.
- CNN, UC, MWGF, Std, LP-PCNN methods also presents have prominent results. Although UC and Std methods have simple structures, they have achieved good objective results.
- CBF has a medium success in objective evaluations too.

Furthermore, CNN is the method that approaches the proposed method most visually and objectively. CNN is the closest method to the proposed method in visually and objective aspects. However, the CNN method has important performance defects such as requiring a training period and a large dataset for training.

In future improvements, the time and resource cost of the proposed method can be investigated to achieve better performance. In addition, the decision fusion step can also be improved for smooth-edged and flawless fusion results. Finally, the proposed method can be generalized for other fusion methods.

REFERENCES

Alenezi, F., 2018. Perceptual Local Contrast Enhancement and Global Variance Minimization of Medical Images for Improved Fusion. International Journal of Imaging Science and Engineering (IJISE), 10(3): 1- 8.

Allebach, J., Lin, Q., 1996. FM screen design using DBS algorithm. in Proc. IEEE International Conference on Image Processing (ICIP ’96), Lausanne, Switzerland, 1, 549–552.

Andelson, E.H., Bergen, J.R., Burt, P.J., Ogden, J.M., 1984. Pyramid methods in image processing. RCA Engineer, 33-41.

Balachander B., Dhanasekaran D., 2016. Comparative Study Of Image Fusion Techniques In Spatial And Transform Domain. ARPN Journal of Engineering and Applied Sciences, 11(9): 5779-5783.

Bavachan, B., Krishnan, P., 2014. A Survey on Image Fusion Techniques. IJRCCT, 48-52: 2278-5841.

Bavirisetti, D.P., Dhuli, R., 2016. Multi-focus image fusion using multiscale image decomposition and saliency detection. Ain Shams Engineering Journal, 9(4):1103-1117.

Bedi, S.S., Khandelwal, R., 2013. Comprehensive and Comparative Study of Image Fusion Techniques. International Journal of Soft Computing and Engineering (IJSCE), 3(1): 300-304.

Burt, P. J. 1981. Fast filter transform for image processing.Computer Graphics and Image Processing. 16: 20-51.

Candès, E. J., Donoho, D. L., 1999. Curvelets – a surprisingly effective nonadaptive representation for objects with edges, Saint-Malo: Vanderbilt University Press.

Chen, Y., Blum, R. S., 2009. A new automated quality assessment algorithm for image fusion. Image and Vision Computing, 27: 1421–1432.

Cunha, A.L., Zhou, J., Do, M.N., 2005. Nonsubsampled Contourlet Transform: Filter Design And Applicatıons In Denoising. IEEE Transactions on Image Processing, 15(10):3089 – 3101

Do, M. N., Vetterli, M., 2005. The contourlet transform: an efficient directional multiresolution image representation, IEEE Transactions on Image Processing, 14(12): 2091–2106.

Eckhorn, R., Reitboeck, H. J., Arndt, M., Dicke, P., 1990. Feature Linking via Synchronization among Distributed Assemblies: Simulations of Results from Cat Visual Cortex. Neural Computation 2: 293-307.

Floyd, R.W., Steinberg, L., 1976. An adaptive algorithm for spatial grey scale. Proc. Society of Information Display, 17, 75077.

Gawari, N., Lalitha Y.S., 2014. Comparative study of PCA,DCT & DWT based image fusion techniques. International Journal of Emerging Research in Management &Technology, 3(5).

Gayathri, N., Deepa, P. L., 2016. Multi-focus Color Image Fusion using NSCT and PCNN. International Conference on Communication Systems and Networks (ComNet), 173-178.

Gonzalez, R. C., Woods, R. E., 2008. Digital Image Processing, Pearson International Edition, Third Edition: 976p.

Haghighat, M. B. A., Aghagolzadeh, , A., Seyedarabi, H., 2011. A nonreference image fusion metric based on mutual information of image features. Computers and Electrical Engineering, 37, 744–756.

Hariharan, H., 2011. Extending Depth of Field via Multifocus Fusion. The University of Tennessee Doctorate dissertation, Knoxville, 157p. (unpublished).

He, K., Sun, J., Tang, X., 2012. Guided Image Filtering. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(6):1397-1409.

Hossny, M., Nahavandi, S., Creighton D., 2008. Comments on ‘Information measure for performance of image fusion'. Electronics Letters, 44(18).

Huang, G., B., Zhu, Q-Y., Siew, C. K., 2004. Extreme learning machine: A new learning scheme of feedforward neural networks. Neural Networks, 2004. Proceedings. IEEE International Joint Conference, Budapest,Hungary, 2: 985-990.

Huang, W., Jing, Z., 2007. Multi-focus image fusion using pulse coupled neural network. Pattern Recognition Letters, 28:1123–1132.

Jassim, F. A., 2013. Correcting Multi-focus Images via Simple Standard Deviation for Image Fusion. I.J. Image, Graphics and Signal Processing, 12: 56-61.

Jin, X., Nie, R., Zhou, D., Wang, Q., He K., 2016. Multifocus Color Image Fusion Based on NSST and PCNN. Journal of Sensors, 1-12.

Jarvis, J. F., Judice, C. N., Ninke, W. H., 1976. A survey of techniques for the display of continuous tone pictures on bilevel displays. Computer Graphics and Image Processing, 5(1), 13-40.

Kim, H. Y., Queiroz, R. L., 2003. Inverse halftoning by decision tree learning. Conference Paper in Proceedings / ICIP International Conference on Image Processing, 2: 1-5.

Kingsbury, N., 1998. The Dual-Tree Complex Wavelet Transform: A New Efficient Tool For Image Restoration And Enhancement. Proceedings of the European Signal Processing Conference, Scotland.

Kolekar, N.B., Shelkikar, R.P., 2015. Decision level based Image Fusion using Wavelet Transform and Support Vector Machine. 4(12): 54-58, 2347-3878.

Kong, N.S.P., Ibrahim, H., Hoo, S.C., 2013. A Literature Review on Histogram Equalization and Its Variations for Digital Image Enhancement. International Journal of Innovation, Management and Technology, 4(4): 386-389.

Kumar, B. K. S., 2013. Image fusion based on pixel significance using cross bilateral filter. Signal, Springer - Image and Video Processing 9: 1193–1204.

Lecun, Y., Haffner, P., Bottou, L., Bengio, Y., 1998. Object Recognition with Gradient-Based Learning.

Li, S., Kwok, J. T., Wang, Y. (2002). Multifocus image fusion using artificial neural networks. Pattern Recognition Letters, 23(8): 985-997.

Li, S., Kang, X., Hu, J. (2013). Image Fusion with Guided Filtering. IEEE Transactions On Image Processing, 22(7): 2864-2875.

Liu, Y., Wang, Z., 2013. Multi-focus image fusion based on wavelet transform and adaptive block. Journal of Image and Graphics, 18(11): 1435-1444.

Liu, Y., Chen, X., Ward, R.K., Wang, Z.J., 2016. Image Fusion With Convolutional Sparse Representation. IEEE Signal Processing Letters, 23(12):1882–1886.

Liu, Y., Chen, X., Peng, H., Wang, Z., 2017. Multi-focus image fusion with a deep convolutional neural network. Information Fusion, 36: 191-207.

Liu, Z., Blasch, E., Xue, Z., Zhao, J., Laganie`re, R., 2012. Objective Assessment of Multiresolution Image Fusion Algorithms for Context Enhancement in Night Vision: A Comparative Study. IEEE Transactions On Pattern Analysis And Machine Intelligence, 34(1): 94-109.

Masthanaıah, M., Kumar, P. J. S., 2014. Development of Image Fusion Algorithms by Integrating PCA, Wavelet and Curvelet Transforms. International Journal of Innovative Research in Computer and Communication Engineering, 2(4): 162-172.

Me, M., 2000. Look Up Table (Lut) Inverse Halftoning. IEEE International Symposium on Circuits and Systems (ISCAS), May 28-31, Geneva, Switzerland, 517-520.

Mingrui, C., Junyi, Y., Guanghui, C., 2015. Multi-focus Image Fusion Algorithm using LP Transformation and PCNN. 237-241.

Mishra, D., Palkar, B., 2015. Image Fusion Techniques: A Review. International Journal of Computer Applications, 130(9): 7-13.

Moody, J., Darken, C. J., 1989. Fast learning in networks of locally-tuned processing units. Neural Computation, 1: 281-294.

Morris, C., Rajesh, R.S., 2015. Modified Primitive Image Fusion Techniques For The Spatial Domain. Informatol, 48: 71-77.

Multi-focus image contest in 2015. http://www.pxleyes.com/photography-contest/19726/multi-focus.html.

Naidu, V. P. S., 2012. Discrete Cosine Transform based Image Fusion Techniques. Journal of Communication, Navigation and Signal Processing, 1(1): 35-45.

Nair, M. S., 2020. Histogram Equalization and Local Histogram Equalization(https://www.mathworks.com/matlabcentral/fileexchange/13729-histogram-equalization-and-local-histogram-equalization), MATLAB Central File Exchange. Retrieved January 16, 2020.

Nejati, M., Samavi, S., Shirani, S., 2014. Multi-focus Image Fusion Using Dictionary-Based Sparse Representation. Information Fusion, 25:72-84.

Oral, M., Kartal, S., 2016. Görüntü Birleştirme İçin Bölgesel Renk Yoğunluğuna Dayalı Yeni Bir Yaklaşım. Eeb Elektrik-Elektronik ve Bilgisayar Sempozyumu, Türkiye, 67- 70.

Qu, G., Zhang, D., Yan, P. 2002. Information measure for performance of image fusion. Electronics Letters, 38(7): 313-315.

Raj, R. G., Geisler, W. S., Frazor, R. A., Bovik, A. C., 2005. Natural Contrast Statıstics And The Selection Of Visual Fixations. IEEE International Conference on Image Processing, Genova, Italy.

Ramac, L. C., Uner, M. K., Varshney, P. K., Alford, M. G. Ferris, D. D., 1998. Morphological filters and wavelet-based image fusion for concealed weapons detection. Sensor Fusion: Architectures, Algorithms, and Applications II, 3376: 110-119.

Sankar, 2019. http://imageprocessing-sankarsrin.blogspot.com/2019/05/direct-binarysearch-halftoning-dbs.html.

Sewell, M., 2011. Ensemble Learning. Dept. Comput. Sci., UCL, London, U.K., Tech. Rep. RN/11/02: 12.

Shah, M. I., Mishra, S., Sarkar, M., Rout, C., 2017. Identification of Robust Focus Measure Functions for the Automated Capturing of Focused Images from Ziehl-Neelsen Stained Sputum Smear Microscopy Slide. International Society for Advancement of Cytometry, 91(8), 800-809.

Siddiqui, A. B., Jaffar, M. A., Hussain A., 2011. Block-Based Pixel Level Multi-Focus Image Fusion Using Particle Swarm Optimization. International Journal of Innovative Computing, Information and Control, 7(7): 3583–3596.

Silva, A. C., Lucena, P. S, Figuerola, W. B. (2000). Average Dithering. Image Based Artistic Dithering, Visgraf Lab.

Sindhu, V. (2013). A Review on Image Halftone Processing. International Journal of Advances in Engineering Sciences, 3(2): 19-23.

Specht, D. F., 1990. Probabilistic Neural Networks. Neural Networks, 3(1): 109-118.

Sravani, N., Chandra, B. S., 2015. Digital Camera Image Fusion Algorithm Using Laplacian Pyramid. International Journal of Computer Science and Mobile Computing, 4(7): 43-49.

Storn, R., Price, K., 1995. Differential Evolution: A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Spaces. Technical Report TR-95-012, International Computer Science Institute, Berkeley.

Stucki, P., 1981. MECCA - a multiple error correcting computation algorithm for bi-level image hard copy reproduction. Research report RZ1060, IBM Research Laboratory, Switzerland.

Sullivan, J., Miller, R., Pios, G., 1993. Image Halftoning Using a Visual Model in Error Diffusion. J. Opt. Soc. Am. 10(8), 1714-1724.

Tian, J., Chen L., Ma L., Yu, W., 2011. Multi-focus image fusion using a bilateral gradient-based sharpness criterion. Optics Communications, 284: 80–87.

Toet, A., 1989. Image fusion by a ratio of low pass pyramid. Pattern Recognition Letters, 9: 245-253.

Ulichney, R. A., 1987. Digital Halftoning. MIT Press, United States, 362p.

Ulichney, R. A., 1988. Dithering with Blue Noise. IEEE, 76(1):56-79.

Wang, Z., Ma, Y., 2007. Dual-channel PCNN and Its Application in the Field of Image Fusion. Third International Conference on Natural Computation.

Wang, Z., Ma, Y., 2008. Medical image fusion using m-PCNN. Information Fusion 9(2):176-185.

Wang, Z., Liu, B., 2008. A Novel Image Fusion Metric Based on MultiScale Analysis. IEEE 9th International Conference on Signal Processing, Beijing, China, 965-968.

Wang, Z., Shen, Y., Jin, J., 2008. Performance evaluation of image fusion techniques. Image Fusion: Algorithms and Applications, 1st ed.; Stathaki, T., Ed.; Elsevier: London, UK. 19: 469–492.

Wang, Z., Ma, Y., Gu, J., 2010. Multi-focus image fusion using PCNN. Pattern Recognition, 43:2003–2016.

Wang, Z., Chang, F., 2011. A Multi-focus Image Fusion Method Based on Laplacian Pyramid. Journal Of Computers, 6(12): 2559-2566.

Wang, Z., Wang, S., Zhu, Y., 2016. Multi-focus Image Fusion Based on the Improved PCNN and Guided Filter. Neural Processing Letters, 45(1): 75–94.

Wohlberg, B., 2015. Efficient Algorithms for Convolutional Sparse Representations. IEEE, 25(1): 301-315.

Wong, P. W., 1995. Inverse Halftoning and Kernel Estimation for Error Diffusion. IEEE Transactions On Image Processing, 4(4), 486-498.

Xiaopeng, J., Jun, C., T., Xinyu, W., Wei, F., 2017. The spatial Laplacian and temporal energy pyramid representation for human action recognition using depth sequences. Knowledge-Based Systems, 122: 64-74.

Xydeas, C. S., Petrović, V. (2000). Objective image fusion performance measure. Electronics Letters, 36(4): 308 – 309.

Yan, T., Liu F., Chen B., 2017. Microscopy image fusion algorithm based on saliency analysis and adaptive m-pulse-coupled neural network in nonsubsampled contourlet transform domain. International Journal of Distributed Sensor Networks, 13(6).

Yang, B., Li, S., 2010. Multifocus Image Fusion and Restoration With Sparse Representation. IEEE Transactions On Instrumentation And Measurement, 59(4):884-889.

Yang, C., Zhang, J., Wang, X., Liu, X., 2008. A novel similarity based quality metric for image fusion. Information Fusion, 9: 156–160.

Yang, J., Ma, Y., Yao, W., Lu, W.T., 2008. A Spatial Domain And Frequency Domain Integrated Approach To Fusion Multifocus Images. XXXVII, Part-B7: 1147-1150.

Yang, Y., 2011. A Novel DWT Based Multi-focus Image Fusion Method. Procedia Engineering, 24: 177-181.

Yang, Y., Yang, M., Huang, S., Que, Y., Ding, M., Sun, J., 2017. Multifocus Image Fusion Based on Extreme Learning Machine and Human Visual System. IEEE, 5: 6989 – 7000.

Yuhendra, Sumantyo, J. T. S., 2017. Assessment of Multi-Temporal Image Fusion for Remote Sensing Application. Advanced Science Engineering Information Technology, 7(3): 778-784.

Zhang, Q., Guo, B., 2009. Multifocus image fusion using the nonsubsampled contourlet transform. Signal Processing, 89:1334–1346.

Zhang, Z., Xu, Y., 2015. A survey of sparse representation: algorithms and applications. IEEE Access, 3:490-530.

Zhou, Z., Li, S., Wang, B., 2014. Multi-scale weighted gradient-based fusion for multi-focus images. Information Fusion, 20: 60-72.

Zhu, H., Chan, Francis H. Y., Lam, F. K., 1999. Image Contrast Enhancement by Constrained Local Histogram Equalization. Computer Vision and Image Understanding, 73(2): 281-290.

CURRICULUM VITAE

Sultan Sevgi TURGUT was born in Burdur, in 1994. She completed the elementary school education at Burdur, Turkey. She graduated from the Isparta Gülkent Anatolian High School in 2012. She graduated from the Department of Computer Engineering, Çukurova University, Adana, in 2017. She started the academy at Yıldız Technical University in 2019 as a Research Assistant.