AN EVALUATION OF GHOST REMOVAL ALGORITHMS FOR EXPOSURE FUSION

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

TUĞSER KUTLU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

DECEMBER 2015

Approval of the thesis:

AN EVALUATION OF GHOST REMOVAL ALGORITHMS FOR EXPOSURE FUSION

submitted by **TUĞSER KUTLU** in partial fulfillment of the requirements for the degree of **Master of Science in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. M. Gülbin Dural Ünver _____
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Gönül Turhan Sayan _____
Head of Department, **Electrical and Electronics Eng.**

Prof. Dr. Gözde Bozdağı Akar _____
Supervisor, **Elec. and Electronics Eng. Dept., METU**

Examining Committee Members:

Prof. Dr. Uğur Halıcı _____
Electrical and Electronics Engineering Dept., METU

Prof. Dr. Gözde Bozdağı Akar _____
Electrical and Electronics Engineering Dept., METU

Assoc. Prof. Dr. Ahmet Oğuz Akyüz _____
Computer Engineering Dept., METU

Assist. Prof. Dr. Sevinç Figen Öktem _____
Electrical and Electronics Engineering Dept., METU

Assist. Prof. Dr. Aykut Erdem _____
Computer Engineering Dept., Hacettepe University

Date: 01.12.2015

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name : Tuğser KUTLU

Signature :

ABSTRACT

AN EVALUATION OF GHOST REMOVAL ALGORITHMS FOR EXPOSURE FUSION

Kutlu, Tuğser

M.S., Department of Electrical and Electronics Engineering

Supervisor: Prof. Dr. Gözde Bozdağı Akar

December 2015, 97 pages

In high dynamic range imaging (HDR), the goal is to capture a scene with a higher dynamic range than the camera capable of capturing with a single exposure. Similar to HDR, exposure fusion is a process that takes multiple images and combines them to create a single dynamically enhanced image by only keeping the properly exposed elements. When using multiple images, local motion of objects can influence the quality of the final image in such a way that local motion of objects causes a ghost artifact. In this thesis, an exposure fusion algorithm is implemented and the techniques used for ghost removal are discussed and compared in terms of quality and computational time. We also propose modifications to the existing algorithms to improve the image quality while decreasing the computational time.

Keywords: High Dynamic Range Imaging, Exposure Fusion, ghost detection, ghost removal

ÖZ

POZLAMA FÜZYONU İÇİN GÖLGELEME GİDERME ALGORİTMALARININ İNCELENMESİ

Kutlu, Tuğser

Yüksek Lisans, Elektrik-Elektronik Mühendisliği Bölümü

Tez Yöneticisi: Prof. Dr. Gözde Bozdağı Akar

Aralık 2015, 97 sayfa

Yüksek dinamik aralığa sahip sahneler, farklı pozlandırma süreleri ile çekilmiş görüntülerin birleştirilmesiyle elde edilir. Ancak, resim çekme sırasında, ortamda bir hareketli nesne bulunur ise, hayalet imgeleri belirir. Bu tezde, görüntüleri birleştirmek için pozlama füzyonu üzerinde çalışılmıştır, çünkü pozlama füzyonu işleminde bütün iyileştirmeler düşük dinamik aralıkta gerçekleştirilmekte ve giriş resimleri ile ilgili herhangi bir bilgiye ihtiyaç duymamaktadır. Buna ek olarak hayalet olgusuz yüksek dinamik aralıklı resimler için kullanılan algoritmalar, hesaplama zamanı ve kalite olarak incelenmiştir. Ayrıca bu çözümlere geliştirmeler önerilmiştir.

Anahtar Kelimeler: Yüksek Dinamik Aralıkta Görüntüleme, Pozlama Füzyonu, hayalet saptama, hayalet kaldırma

To Mommy …

ACKNOWLEDGMENTS

I would like to thank my advisor Prof. Gözde Bozdağı Akar for her guidance and support. I have learned a lot from her, not only theoretically, but also how to be an engineer in practice. Having the opportunity to observe her approach to any kind of problems is one of the greatest benefits that I gained during my master studies.

I would also like to thank my family and friends for their support and encouragement. They were always there with their smiling faces to give me hope. They made me feel safe and loved which provided me the strength to carry on.

I also express my sincere gratitude to my managers and my colleagues from ASELSAN INC. and ASELSAN INC. itself, for their initiative ideas and guidance that helped to construct this work.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

TABLE OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

1D	One Dimensional
2D	Two Dimensional
CCD	Charge-coupled Device
CDF	Cumulative Distribution Function
CMOS	Complementary Metal-oxide Semiconductor
CRF	Camera Response Function
DC	Direct Current
EF	Exposure Fusion
HDR	High Dynamic Range
HDRI	High Dynamic Range Imaging
JPEG	Joint Photographic Expert Group
MD	Multiscale Decomposition
MTB	Median Threshold Map
ZNCC	Zero Mean Normalized Cross Correlation

CHAPTER 1

INTRODUCTION

The main aim of digital cameras is to form an image of a scene similar to the scene as observed by the human visual system. With the advancement of digital camera technology, this goal is achieved to a great extent. However, not all problems are solved in capturing the scene as close as possible to reality. In [1], Eastman Kodak Company describes the sources of noise for the CCD image sensors; and underlines the performance of the digital camera is limited by CCD.

Although, the digital cameras have been improved in terms of performance and quality, photography is still having a problem with a wide range of radiance variations in the real world. The scenes in the real world comprise of harsh lightening conditions that cause shadows (underexposed regions) or highlights (overexposed regions) in digitally captured images. The reason is that the dynamic range of camera sensors is not high enough to capture the dynamic range of the scenes. Dynamic range of the scene is defined as the ratio of radiances between the brightest and the darkest points in the scene. For example, the radiance range of the real world reaches up to 1:500000 but today's camera sensors have pixel depth ranging from 8-bit to 14-bit which correspond 256 to 16384 digital values [2].

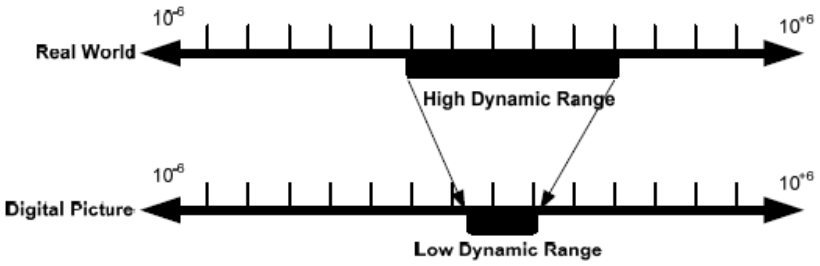


Figure 1-1 Mapping from High Dynamic Range to Low Dynamic Range [3]

To visualize the scene as realistic as possible, dynamic range enhancement methods are presented. For this purpose, some improvements are made in software as well as hardware. For example, Toshiba Company have presented a new type of CMOS image sensor that is capable of simultaneously capturing an image with different exposure times and merges them into a single image [4]. However, production of this type of sensors and their control integrated circuits are expensive, and also on the market these types of sensors are not common. Therefore, software solutions are presented to form an image of a scene similar as the scene observed by the human visual system. These solutions are for the commercial cameras. Mainly two approaches are used in software. Firstly, High Dynamic Range (HDR) imaging techniques is used to extract the actual radiance map of the scene and tone-mapping is used to display it for the display devices, secondly, fusion of the input images with different weights are presented to create "HDR-like" image [5].

In the software, post-processing techniques are used to increase the dynamic range of the scene. However, because of the limitations of the digital camera, dynamic range cannot be fully extracted from one image. The most common method for dynamically enhanced image generation is based on the combination of multiple distinct exposures. The motivation behind this technique is that different exposures capture different dynamic range characteristics of the scene.

However, moving objects in the scene cause unwanted artifacts in HDR or HDR-like images. Classical image enhancement methods cannot eliminate these. For this purpose, ghost removal algorithms have been developed. Mainly, these algorithms are intended to be used with the image enhancement methods. Ghost detection and ghost removal are the two main steps of these algorithms. One can achieve ghost-free high dynamic range with the help of these ghost removal algorithms.

1.1 Scope of this thesis

In this thesis, an exposure fusion algorithm is implemented and the techniques used for ghost removal are discussed and compared in terms of quality with different image sets. In terms of computational time, ghost removal algorithms are compared, and improvements are given.

The main goals of this thesis are given as follows:

- Describing the principles of exposure fusion algorithms to create an enhanced image
- Performing detailed analysis of ghost removal algorithms with different image sets
- Comparison of the performance of the ghost removal algorithms in terms of quality and computational time

Common datasets found in the literature are used for performance evaluation [6] [7]. All images have been aligned; thus, registration of these images is out of the scope of this thesis. Image sets are shown in Appendix A.

1.2 Outline of the Thesis

In Chapter 2, a literature review on image enhancement methods and ghost removal algorithms is presented. HDR image algorithms and exposure fusion algorithms are explained. Ghost removal algorithms found in the literature and their techniques are briefly explained and compared.

In Chapter 3, detailed explanations of analyzed algorithms are given. The theory behind each algorithm is explained.

In Chapter 4, the performance evaluation of ghost removal algorithms is presented. This chapter mainly focuses on the comparison of ghost removal algorithms.

In Chapter 5, conclusion is given and possible future directions are discussed.

CHAPTER 2

LITERATURE REVIEW

Many natural scenes have a dynamic range that is larger than the dynamic range of a camera's image sensor. A popular approach to producing an image without under- and over-exposed areas is to capture several input images with varying exposure settings, and later merge them into a single high-quality dynamically enhanced image.

The main limitation for the combination of the multiple exposures is the requirement of the static scene when taking images. Certainly, any movement in the scene causes the noticeable ghosting artifacts in resulting image. This problem cannot be avoided in outdoor environments which contain movements, such as automobiles, people and nature movement due to wind.

This chapter is composed of different literature review sections: 1) HDR imaging and tone mapping 2) fusion in the image domain 3) methods used for ghost removal.

In the first section, HDR imaging and tone mapping are described, in the second section, exposure fusion techniques are described. In the subsection of this section, exposure fusion technique used in this thesis is detailed and experimental results are shown. In the third section, overview of various techniques that have been proposed for ghost removal is given.

2.1 HDR Imaging and Tone Mapping

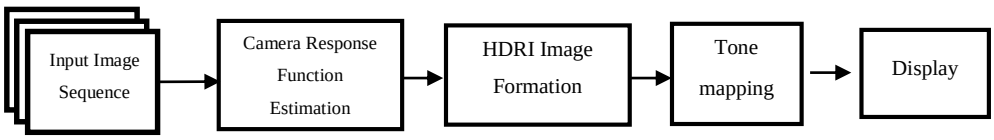


Figure 2-1 Scheme of Fusion in Radiance Domain

The HDR image generation in radiance domain consists of three steps [8]: firstly, estimation of camera response function (CRF) is conducted to bring pixel brightness values into radiance domain. However, it is not an easy task because CRF depends on camera characteristics, and it is not same for all cameras. Since CRF is not provided by the manufacturers, different methods are proposed for its estimation using series of different exposed LDR images [8] [9]. Secondly, the sensor irradiance is calculated by using CRF and the exposure times of the input image. This operation can be explained mathematically as follows:

$$L(i,j) \approx E(i,j) * \Delta t_k = T^{-1}\big(I_k(i,j)\big)\ where\ k = 1,2,\dots,N \qquad (2.1)$$

$$I_k(i,j) = T(E(i,j) * \Delta t_k)\ where\ k = 1,2,\dots,N \qquad (2.2)$$

where $I_k(i,j)$ is the intensity value of the pixel at *(i,j)*, and Δt_k is the exposure time of k^{th} image. The sensor irradiance is denoted by (i,j) .

The calculated image represents an approximation of the sensor irradiance, and it is called as HDR image or HDR radiance map. The sensor irradiance has been calculated, but the scene radiance has to be found. Most of the time, it is assumed that scene radiance $L(i,j)$ is proportional to sensor irradiance $E(i,j)$. This assumption is valid because manufactured lenses include designs to compensate the nonlinearity between the scene radiance and sensor irradiance. By the help of these designs, the relation between the radiance and the irradiance becomes almost constant.

Finally, tone mapping operators are used to make HDR radiance maps displayable on common low dynamic range monitors [8] [10].

Methods that combine the images in radiance domain, produce a true HDR radiance map. For this purpose, accurate estimation of the CRF, which is sensitive to image noise and misalignment, is needed for this purpose.

2.2 Fusion in the Image Domain

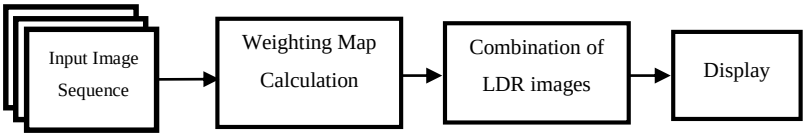


Figure 2-2 Scheme of Fusion in Image Domain

Alternative methods combine multiple exposures directly without the knowledge of the camera response function [11] [12]. The fusion methods combine LDR images by conserving only the best parts of each exposure. These parts could be pixel, feature or symbol. The final enhanced image is obtained as a weighted average of pixel values across exposures. Note that, fusion algorithms create an "HDR-like" image which can be displayable on common display devices without any mapping. Although, it does not display actual radiance map of the scene, features, colors and dynamic range of the scene are conserved in the final enhanced image.

The choice of the weighting function is crucial to get accurate results. The main goal of image fusion can be defined as; transferring salient information in the input images to the fused image by preserving the details of the input images and not causing any artifacts in the fused image. For this purpose, gradient information [13], entropy measure [12] and bilateral filter [14] are used.

Introducing artifacts and losing important features of the input images are the main problem of the selection of the weight function. In this thesis, the technique proposed in [15] is implemented because, firstly, it is a pixel level fusion algorithm. This helps to eliminate the problematic pixels at the final enhanced image by using a binary mask. Secondly, it uses multi-dimensional blending technique. This technique helps to eliminate the artifacts. Finally, estimation of the CRF is not required. This reduces computational cost compared to tone mapping. In the following section, the implemented exposure fusion algorithm is explained.

2.2.1 Exposure Fusion Algorithm (EF)

The exposure fusion algorithm is the work of Mertens, Kautz and Reeth [15]. This algorithm uses Laplacian and Gaussian pyramids with three quality measures. These measures determine the weight of the pixel in the final HDR image. The methodology can be explained symbolically as follows:

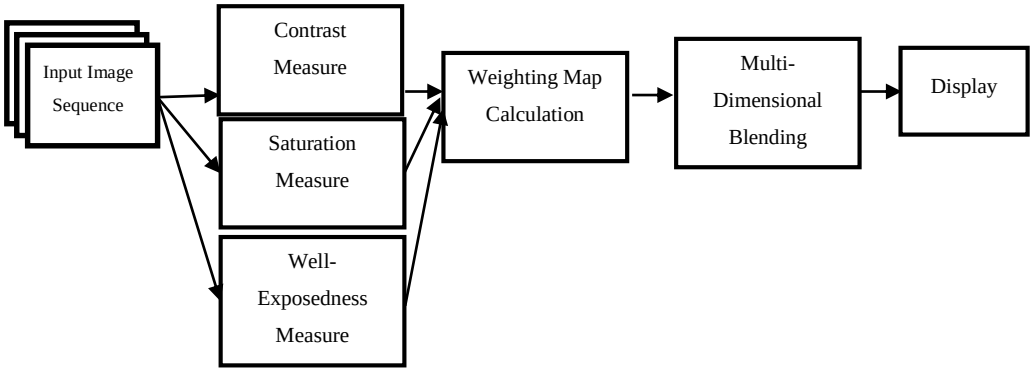


Figure 2-3 Scheme of Exposure Fusion

As it has been mentioned previously, three different quality measures are used for this method. These measures are explained in details in following sections. For visualization of the quality measures and the enhanced image, image set in Appendix A.1 is used.

Contrast

Contrast is the difference in luminescence that makes object distinguishable. It is one of the important characteristics of an image to get the salient features, such as textures and edges. In the exposure fusion algorithm, contrast measures are conducted as follows:

$$C_k = conv\left(I_{mono,k}, H_{Laplacian}\right) \tag{2.3}$$

where, $conv(.)$ is the convolution operation, k is the number of the input, I_{mono} is the grayscale converted image of the input and $H_{Laplacian}$ is the Laplacian kernel, defined as:

$$H_{Laplacian} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (2.4)$$

By using Laplacian kernel, sum of differences over the nearest neighbors of the central pixel are calculated. If there is an edge, or rapid changes, the resulting C_k has more values. Contrast measures of input scene are seen in Figure 2-4.

Weights of the edges are dominant in the contrast measures of each input. Selecting these dominant pixel improves the performance of the final image, since keeping edge information is perceptually important to human visual system.

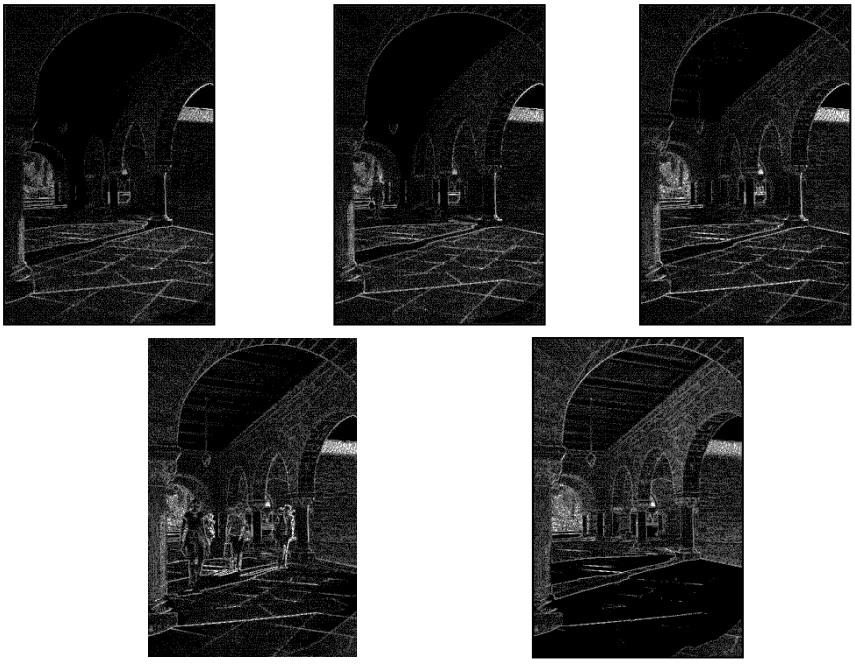


Figure 2-4 Contrast Measure of the Input Set 1

Saturation

The saturation of a color is determined by a combination of light intensity and how much it is distributed across the spectrum of different wavelengths. In other words, most saturated color is achieved by using only one wavelength. In RGB domain, most saturated color of a pixel is achieved such a way that, one or two of the color channels has the most value while others are the least. Saturation measure is conducted by computing standard deviation of the Red, Green and Blue channel of a pixel in the input image. This computation is done as follows:

$$a_{i,j,k} = \frac{R_{i,j,k} + G_{i,j,k} + B_{i,j,k}}{3} \tag{2.5}$$

$$S_{i,j,k} = \sqrt{\frac{(R_{i,j,k} - a_{i,j,k})^2 + (G_{i,j,k} - a_{i,j,k})^2 + (B_{i,j,k} - a_{i,j,k})^2}{3}} \tag{2.6}$$

where $R_{i,j,k}$, $G_{i,j,k}$ and $B_{i,j,k}$ are the color values of the pixel at (i,j) in kth input image, red, green and blue, respectively. $a_{i,j,k}$ is mean of the color values, and $S_{i,j,k}$ is the saturation value at that pixel.

The saturation measure of the input scene is seen in Figure 2-5. Under-exposed and over-exposed pixels tend to get lowest values because the color information of these areas are not recovered by the camera.

Figure 2-5 Saturation Measure of the Input Set 1

Saturation measure is the key measure to keep the color information as much as possible in the final image by taking only most saturated pixels. Main goal of this measure is the selection of the best color of the pixel in the final image.

Well-Exposedness

Over-exposed and under-exposed pixels are the undesirable regions in an image, because, pixels in that region are not exposed enough to show the sufficient information in that scene. Well-exposedness measure excludes the pixels of that kind. Calculation of this measure is given as follows:

$$\mathrm{E}_{i,j,k} = \exp\left(-\frac{(I_{i,j,k} - 0.5)^2}{2\sigma^2}\right) \tag{2.7}$$

where, $I_{i,j,k}$ is the intensity value at the pixel (i, j) in kth input image. σ is a constant which is equal to 0.2.

This measure is a Gaussian filter, therefore, it weights each pixel based on how close it is to 0.5. Importance of this measure is the selecting the pixels by looking its exposed value. In this method, intensity map is used to get the exposed value of a pixel.

The well-exposedness measure of the input set is seen in Figure 2-6.

Figure 2-6 Well-Exposedness Measure of the Input Set 1

As stated previously, this measure excludes the under-exposed and over-exposed regions. For this input set, ceiling is the under-exposed area and sky is the over-

exposed areas. Weights of these regions are relatively low compared to other regions. This circumstance is seen for each input image.

Fusion of the Images

For each input image, a weight map is calculated by using these quality measures. The calculation of the weight maps is given as follows:

$$W_{i,j,k} = C_{i,j,k}^{wC} * S_{i,j,k}^{wS} * E_{i,j,k}^{wE} \tag{2.8}$$

where, *wC*, *wS* and *wE* are used to control the contribution of each quality measure to the weight map. In this thesis, these exponents are selected as unity, because, the scope of this thesis does not cover the effect of changing contributions of each quality measure.

Until now, weights along each pixel of *N* input images are calculated. To obtain a consistent result, the values of the *N* weight maps are normalized such that they sum to one at each pixel as follows:

$$\widehat{W}_{i,j,k} = \left[\sum_{k=1}^{N} W_{i,j,k}\right]^{-1} * W_{i,j,k} \tag{2.9}$$

where, $\widehat{W}_{i,j,k}$ is the normalized weight of the pixel in the *k-th* image. Weights of each pixel in each input set can be seen in Figure 2-7.

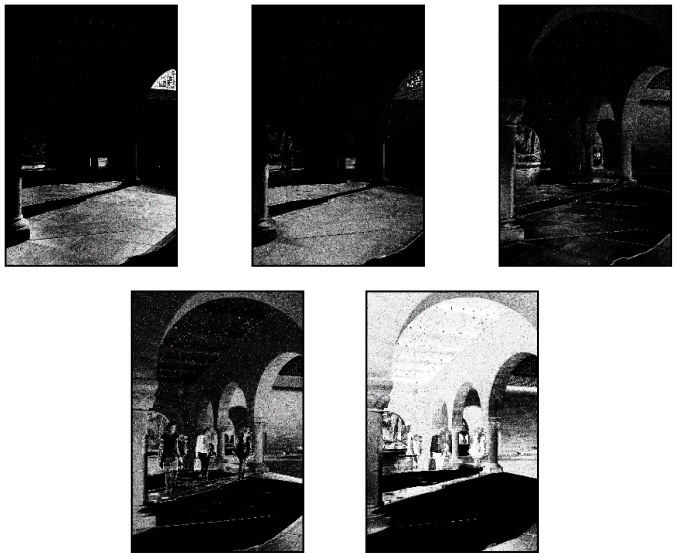
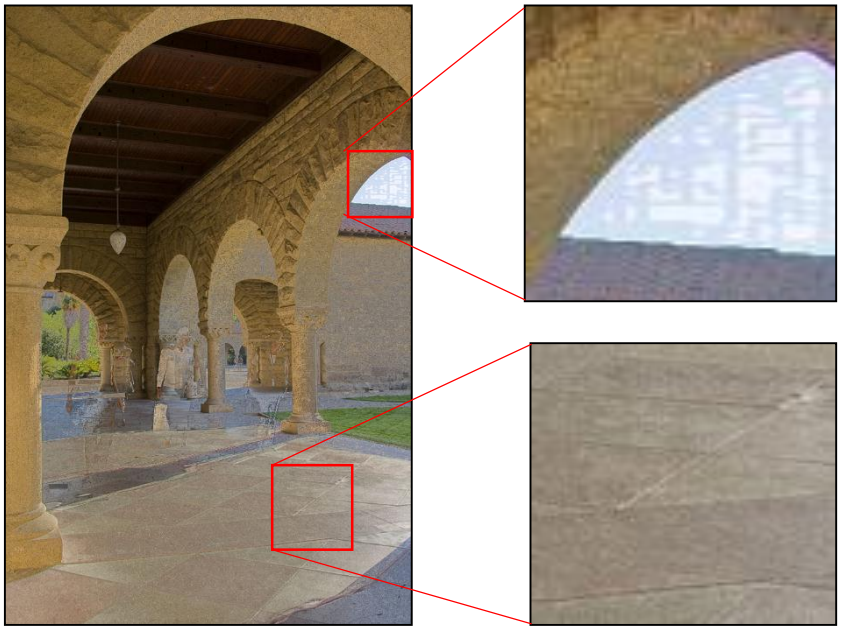


Figure 2-7 Normalized Weights of Each Input Image

For fusion of the input images, weighted blending technique is not feasible; because of rapid changes in weights. These changes are degraded the final image such a way that disturbing seams will appear. This problem can be seen in Figure 2-8.



(a) (b)

Figure 2-8 Seams with Weighted Blending, (a) Final Image, (b) Detailed Regions

Seam problem is avoided by using multi-dimensional blending technique. This problem is avoided by applying a Gaussian filter to the weight map prior to the blending operation. On the other hand, this filtering causes artifacts in the fused image. However, these artifacts are not as dominant as seams. Therefore, multi-dimensional blending improves the quality of the image. The fusion operation is defined as follows:

$$L\{R\}_{i,j}^{l} = \sum_{k=1}^{N} G\{W\}_{i,j,k}^{l} * L\{I\}_{i,j,k}^{l} \tag{2.10}$$

where, $L\{\}$ is the Laplacian pyramid with level l, $G\{\}$ is the Gaussian pyramid with level l. R is the final image, I is the input image, and k is the number of input image.

In a Gaussian pyramid, subsequent images are weighted down using Gaussian average and down sampling. It creates a stack of successive smaller images with each pixel containing a local average that corresponds to a pixel neighborhood on

a lower level of the pyramid. Each value within *level (n)* is computed as a weighted average of values in *level (n-1)* within a 5-by-5 window. The window is specified as follows:

$$a = [0.0625\ 0.25\ 0.375\ 0.25\ 0.0625] \tag{2.11}$$

$$H_{5x5} = a^T * a \tag{2.12}$$

This window is separable, normalized and symmetric [16]. Therefore, it is easy to implement, and neighborhood pixels contribute Gaussian-like distribution at filtering.

Laplacian pyramid generates a sequence of error images. Each image is the difference between two levels of Gaussian pyramid as follows:

$$L_l = g_l - EXPAND\ (g_{l+1}) \tag{2.13}$$

where, *EXPAND(.)* is the interpolation operator. While Gaussian pyramid is set of low-pass filtered images, Laplacian pyramid is set of band-pass filtered version of the image. The scale of the Laplacian operator doubles from level to level of the pyramid, while the center frequency of the passband is reduced by an octave.

Multi-dimensional blending can be summarized in Figure 2-9.

Multi-dimensional blending operation provides consistent results because these methods blend features rather than pixel intensities. Therefore, the resultant image does not have any blending artifacts or sharp transitions. Finally, the fused image is calculated by inverse Laplace Pyramid. The resulting image by using input set 1 is seen Figure 2-10.

Using Laplacian and the Gaussian pyramids have satisfying result in terms of seamless blending, this result is expected since multi-dimensional blending of multiple images, guided by each level of pyramid weight measure, is the process of combining to create the appearance of partial or full transparency. It should be noted that the processing is longer as the level increasing, because decomposition and composition of the image by using pyramids requires more time. However, filtering artifacts are unnoticeable when the level of the pyramids are increasing.

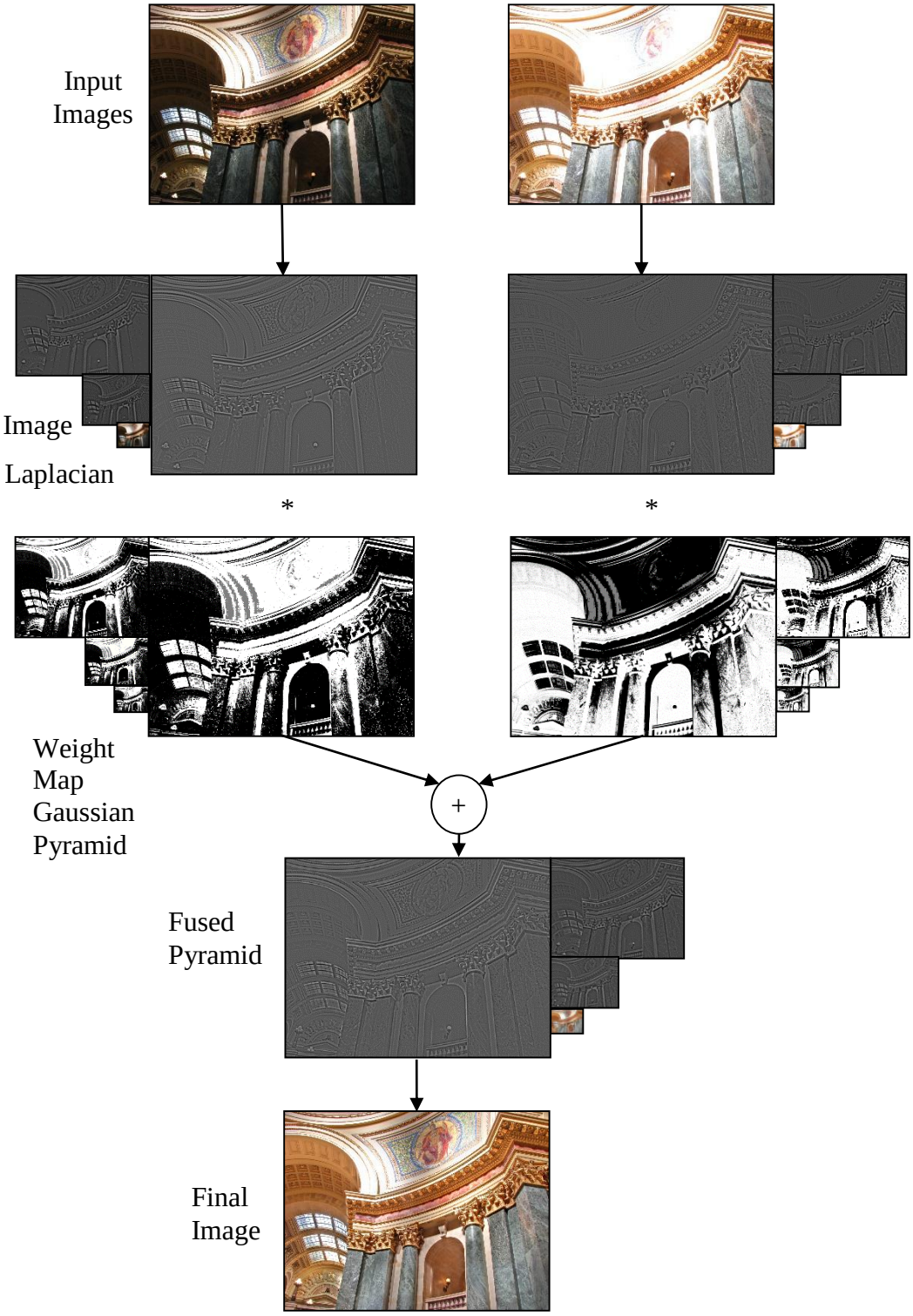


Figure 2-9 Modelling of Multi-Dimensional Blending

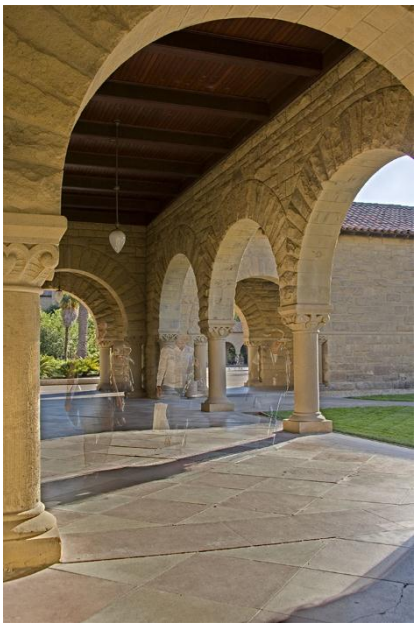


Figure 2-10 The Final Enhanced Image by Using Multi-Dimensional Blending

Exposure fusion algorithm creates an “HDR-like” image which can be displayable on common display devices without any mapping. Although, it does not display actual radiance map of the scene, features, colors and dynamic range of the scene are conserved in the final enhanced image.

2.3 Ghost Removal Algorithms

The main limitation of the multiple input fusion technique is the requirement of a complete static scene when capturing the images. Indeed, any object movement in the scene can cause ghosting artifacts in the resulting dynamically enhanced image. The ghosting problem is a severe limitation of the multiple input fusion technique since motion can hardly be avoided in outdoor environment.

In the fusion of images, several problems occur due to global and local motions of an object or camera. Global motion is the cause of misalignment between input images, it occurs when the images are captured by a hand-held camera. Even though this misalignment is eliminated by image registration algorithms, the ghost effect by local motion of an object still remains, influencing the quality of the final enhanced image. In this thesis, image registration algorithms are out of the scope.

Mainly, ghost detection methods are based on detection of motion, present in the input exposure sequence. Detection of motion is done in various ways [6].

Variance measure is used to detect the ghost regions. First of all, camera response function is estimated and then computation of radiance map is conducted. To make a comparison between images, variance of radiance values at each spatial location is generated. Regions affected by motion are tending to have high variance. Therefore, thresholding is applied to variance image to get the ghost regions.

After alignment of input images, Grosch [17] suggests that deviation between predicted intensity value of a pixel and its actual intensity can be used to detect either ghost pixel or non-ghost pixel. Actual intensity of a pixel is given as follows:

$$I'_{i,j,k} = f\left(\frac{\Delta t_k}{\Delta t_l} f^{-1}(I_{i,j,l})\right) \tag{2.14}$$

where, $f(.)$ denotes the camera response function, $I_{i,j}$ is the color of the pixel at position (i, j), taken with exposure time Δt_l. Δt_k is the exposure time of second image.

For each consecutive pair of input images, pixels which show significant difference between predicted value and actual value are labelled as ghost pixels based on thresholding as follows:

$$|I'_{i,j,k} - I_{i,j,l}| < \varepsilon \tag{2.15}$$

The user parameter ε depends on the amount of noise in the camera images. This threshold is to find the significant difference between the two colors. After then, pixel is labelled as ghost if this rule is violated.

Jacobs et al. [18] defines two types of motion; high contrast and low contrast motion. High contrast motion is caused when motion object is different from background and variance based ghost detection is used to detect this kind of motions. For high contrast motion, it uses approach similar to Grosch [17]. On the contrary, low contrast motion is caused when background and moving object are similar in color and intensity, hence; variance map is not able to detect this kind of motion. In order to detect low contrast motion, entropy based ghost detection is presented. First, a local neighborhood based entropy map is computed and then

Uncertainty Image (UI) is then derived from the weighted difference of the precomputed entropy images as follows:

$$UI(x) = \sum_{n=1}^{N} H(x_n) + \sum_{n=1, n\neq m}^{N} \left(\sum_{m=1}^{N} H(x_n|x_m) \right) \quad (2.16)$$

where, x_n small segments of the input image. $H(x_n)$ is per segmented entropy, and $H(x_n|x_m)$ is the conditional entropy between matching segments from different input images are calculated using the histogram and joined histogram of the segments [19]. Then, ghost regions are found based on thresholding.

Gallo et al [20] uses an approach to generate ghost-free HDR images in such a way that firstly, the input image with least saturated regions is selected as reference image. Secondly, in order to determine if a patch is affected by ghost, log intensities of the patch are plotted against the related patch in the reference image. Then, using RANSAC procedure, best fit line through the plot is obtained. Distance threshold is used to determine the percentage of outliners. Finally, if this percentage of the patch is greater than predefined threshold, the patch is labelled as ghost region. Sidibe et al. [7] use a similar approach to generate ghost-free HDR images. However, their algorithm is based on pixels rather than working with image patch processing.

Heo et al. [21] uses joint probability densities to estimate the global relationship between different exposures coarsely and using energy minimization based on graph cuts methods. Ghost map for each exposure is created based on a threshold by comparing the joint probability densities. However, ghost regions estimated by thresholding are noisy. The use of graph-cuts method helps to overcome the noise.

Bogoni [22] proposed a ghost detection method based on the assumption of "ghost regions are mainly occurred due to the moving object in the scene". If the motion between exposures is found, it is possible to estimate the ghost map. For this reason, Lucas and Kanade [23] algorithm is used to detect dense local motion. After that, all input images are correctly aligned by warping the image. Finally, differences in warped images are labelled as ghost regions.

Khan et al. [24] propose a kernel density estimation method that iteratively estimates the probability that a pixel belongs to the static part of the scene, and then, weighting function [8] used in the HDR image generation equation is adjusted by using this information. In addition, Pedone and Heikkilä [25] propose similar iterative method to compute the weighting function. However, instead of using kernel density estimation, they propose an iterative method to propagate influences of pixels that have low chances to belong to the static part of the scene through an image-guided energy minimization approach.

Among the algorithms found in the literature, there are mainly two approaches in order to get the luminous intensity of the scene, one is the grayscale image converted by using RGB channel [26] [27], other is the L channel in CIELAB color space [28]. In addition, detection of the ghost region is carried out in pixel level [7] or patch level [29]. In the literature, the comparison of images that are taken under different exposure settings by effectively removing most of the illumination differences between images is carried out with different approaches. Median threshold is used to reveal image features while removing intensity differences between different exposures [11], on the contrary, instead of using only median value, classification the intensity values into multi-levels is used [27]. In addition, histogram matching method [26] is used to remove intensity differences between different exposures. These algorithms are selected, because they cover the whole representative of ghost removal algorithms. Detailed explanations of these algorithms are given in the next chapter.

CHAPTER 3

GHOST REMOVAL ALGORITHMS

3.1 Introduction

The main problem with exposure fusion technique is that ghost artifacts are created in the final image when a local or global motion is present in the scene. Image quality is degraded in the enhanced image unless ghost artifact are eliminated. This problem can be seen in Figure 3-1. In order to reduce the effect of this problem, ghost reduction algorithms are presented.

Figure 3-1 Ghost Artifacts in the Enhanced Image

This section is organized as follows, 1) The details of the methods used in this thesis are given 2) The improvements are proposed, and the comparison of the methods are provided.

3.2 Methods Implemented for Ghost Removal

Ghost removal methods are developed to overcome the ghost artifacts in dynamic scenes. Mainly these algorithms use two step strategies; firstly, regions affected by the ghost are detected, secondly ghost artifacts are removed. The main problem is the detection of the ghost region. In the literature, different ghost detection techniques are presented and they differ based on how they approach the ghost removal problem. The groups of approaches are given as follows [6]:

- Registration methods that are used for the registration of the input exposures.
- Moving object removal methods are used to remove ghost artifacts in the final enhanced image by estimating static background.
- Moving object selection methods are used to detect the inconsistencies in the intensities of pixels or patches, and remove of these regions.
- Moving object registration methods is used to recover the ghost pixels by searching for the best matching region in the input exposures.
- Video deghosting methods are used in HDR videos, they remove the ghost artifacts by the using temporal information of videos.

In this thesis, mainly moving object selection methods are analyzed. This approach is suitable for exposure fusion algorithms, because, resulted ghost maps are used to mask out the pixels that are selected as ghost. In addition, they try to maximize dynamic range by using many exposure as possible for each dynamic region.

In this section, detailed explanation of analyzed algorithms is given.

3.2.1 Pixel order relation [ALG-1]

Order relation between each LDR images are investigated in this method [7]. Assuming that linear response function of the imaging system and brightness value (Z) is related to the scene radiance value (L):

$$Z = L * \frac{cos^4 \emptyset}{h^2} * E \tag{3.1}$$

where, E is the exposure value of the image, $\emptyset$ is the angle of principle ray and h is the focal length of the camera. E can be modeled as:

$$E = \frac{\pi d^2}{4} * t \tag{3.2}$$

where, d is the aperture size of the camera lens and t is the exposure duration.

Specifically, on the assumption that camera response function is a monotonic function, pixel values should be in the increasing order from low to high exposures, in the non-ghost region.

$$I_{i,j,k} < I_{i,j,k+1} \tag{3.3}$$

where, *(i,j)* is the pixel location, *k* is the number of LDR image. If this rule is violated, that pixel is stated as ghost region.

This algorithm is based on the increasing pixel intensity with increasing exposure times. However, characteristic of an object that in motion determines the ghost labelling. For example, a moving white ball cannot be detected as ghost region with background of pure black because intensity level of the white ball is always larger than the background, and this algorithm fails to detect it. In order to overcome this problem, improvement is applied in such a way that rapid change in the intensity level of a pixel is also considered as ghost region. The expression of this is given as follows:

$$|I_{i,j,k+1} - \left(\frac{e_{k+1}}{e_k}\right) I_{i,j,k}| > 0.1 \tag{3.4}$$

where, e_{k+1} and e_k are the exposure times of the *(k+1)th* and *k-th* input image, respectively. Please note that exposure time of *(k+1)-th* image is greater than *k-th* image. Linear assumption is valid for the input sequences, since linear camera response function is used to convert RAW images into 8-bit LDR images [6].

This rule helps same pixels in each image to be calibrated with using exposure times. Linear scaling between pixel intensity levels of input images are used for thresholding. This threshold value of 0.1 is obtained experimentally. If the calibrated value of a pixel is larger than this threshold, ghost region is detected. Otherwise, it is not labelled as ghost region. This approach helps to reduce the

dependency of the intensity level between ghost pixel and its background, such a way that, rapid change in intensity difference is also considered as ghost region.

Pixel order relation is the simplest method among other methods. There is no need of calibration among input images, only pixelwise comparison is done by looking intensity image of the inputs. As it is said, this algorithm is in the group of moving object removal, and it is selected to evaluate the performance of a ghost removal algorithm belong to this group.

3.2.2 Histogram Based Ghost Detection [ALG-2]

This method [26] is based on utilization of photometrically calibration function; this is simply a transform by using histogram of the reference image. By using this calibration, all input images are mapped to a common domain by histogram matching, and then ghost probability is found by the pixelwise difference between the reference image and the corresponding image:

$$\mathrm{M_{i,j,k}} = \exp\left(-\frac{(I_{i,j,ref} - I'_{i,j,k})^2}{2 * c * \sigma^2}\right) \tag{3.5}$$

where, $I'_{i,j,k}$is the modified image that is created by using histogram matching towards the reference image, $\mathrm{M_{i,j,k}}$ is the non-ghostness probability of the k-th image at *(i,j)* pixel, *c* is the threshold controlling constant, and σ is the noise level of the image.

As it is expected, the more intensity difference occurs between reference and the input image, the resulting M is diverging to the 0. Therefore, by multiplying M with weights coming from the exposure fusion, the difference regions are masked, and at the final multi-dimensional blending operation, these regions are less effective, also, it should be noted that each entity of M has a value different from 1. Therefore, each weight of pixel is degraded by using M, whether it is a ghost pixel or not.

Input mage, whose mean of well-exposedness quality measure is maximum, is selected as reference image and its M matrix is not taken into account.

These M values for other LDR images are multiplied by the weights coming from the exposure fusion. After that, multi-dimensional blending is done to get the final enhanced image.

3.2.3 Bitmap Movement Detection [ALG-3]

In this approach [11], in order to make a comparison of the LDR images, firstly, median threshold bitmaps (MTB) [30] of each LDR image are found. This algorithm is effective to remove the illumination differences, due to the exposure time, for each image. When, a stack of MTB is obtained, a comparison is conducted such a way that, firstly, for a static scene, it is expected that each pixel preserves its bit value across all bitmaps, secondly, if the value changes in a pixel, there was movement underneath it. By using this assumption, a ghost map is created. However, it is susceptible to noise. To avoid this problem, series of morphological operations are conducted to that ghost map.

After finding the noise-free ghost map, it is converted to the cluster map, where each cluster has a unique label. Clustering is done by using Connected Component labelling [31]. By using these unique labels of clusters, at the final enhanced image, these clusters pixel are selected by comparing the well-exposedness weight of each input LDR image. Input image, whose maximum average of the well-exposedness has the largest value in the cluster, is used. Pixel data of that cluster area is imported from that image. Then, multi-resolution blending technique is used as in the case of exposure fusion.

This algorithm does not need a reference image. Only one ghost map is created using median threshold maps of each input image and the detected ghost regions are filled from the best exposed input image for each individual ghost region. Therefore, it uses single image to fill the pixels in the final image. This approach degrades the dynamic range of the enhanced image, because, using only one input image is against the nature of exposure fusion.

3.2.4 An Exposure Fusion Approach without Ghost for Dynamic Scenes [ALG-4]

Method of acquiring ghost map [28] is conducted in *CIELAB* color domain. For this purpose, over-exposure image is selected as reference, and after all input images are converted into *CIELAB* color space, the average brightness of each input LDR is calculated by using the *L* component of the image, as follows:

$$l_k = \frac{\sum_{x=1}^{H} \sum_{y=1}^{W} I_{i,j,k}}{H * W} \tag{3.6}$$

where, l_k is the average brightness of *k-th* input LDR image, *H* is the height of the image, *W* is the width of the image, $I_{i,j,k}$ is the actual brightness value of the *(i,j)* pixel in the *k-th* image. Each l_k is compared to l_{ref} such a way that, Δl_k, difference of the average brightness values, is found by using following formula:

$$\Delta l_k = l_{ref} - l_k \tag{3.7}$$

Then, modified brightness value for each pixel is found. These values help to find the ghost map by comparing the modified brightness values of each pixel in the LDR images. Simply, by adding Δl_k value to the each of the pixel in *k-th* brightness image, unified brightness image is found. After then, $\Delta I_{i,j,k}$, brightness differences of each pixel is found as follows:

$$\Delta I_{i,j,k} = |I^*_{i,j,k} - I_{i,j,ref}| \tag{3.8}$$

where, $I^*_{i,j,k}$ is the brightness value of the pixel *(i,j)* in the unified brightness image of the *k-th* input image. Finally, ghost map is found by thresholding $\Delta I_{i,j,k}$ value, thresholding function is as follows:

$$M_{x,y,k} = \begin{cases} 0, & \Delta I_{i,j,k} \geq T_{i,j,k} \\ 1, & \Delta I_{i,j,k} < T_{i,j,k} \end{cases} \tag{3.9}$$

where, $M_{i,j,k}$ is the ghost map pixel *(i,j)* for the *k-th* image and $T_{i,j,k}$, threshold value is defined as:

$$T_{i,j,k} = |\Delta l_k|^{\beta} + \left(\frac{I_{i,j,k} - 50}{\sigma}\right)^2 \tag{3.10}$$

β, tolerance factor which represents degree of correlation, is selected as 0.7 and σ, scaling factor, is selected as 15. This function is an adaptive threshold function which is calculated for each pixel value.

After creating ghost maps for each input image, in order to avoid false detections and pepper noises, erosion and dilation operations are done with square kernel sizes 5x5, 9x9, respectively. For each input image (except reference image), $M'_{i,j,k}$ noise-free ghost map is found. This ghost map acts as a mask for the final weights at the exposure fusion algorithm.

While other algorithms use RGB color domain, this algorithm uses *CIELAB* color domain. *L* component of this domain represents lightness of the scene. Therefore, it is useful to get the intensity values of the scene by using this component. In addition to that, this color domain is device independent. This means that the colors are defined independent of their nature of creation or the device they are displayed one. In addition, calibration of each input image to reference image is done in a simple manner, adding the average brightness difference between them. This reduces the complexity compared to other algorithms. Moreover, this algorithm uses an adaptive threshold for each pixel to find ghost regions, while others use general threshold.

3.2.5 Ghost Detection and Removal Based on Super-Pixel Grouping in Exposure Fusion [ALG-5]

Superpixel grouping is used for segmenting of an image into homogeneous regions based on color, intensity and texture details. Instead of comparing pixelwise difference of an input image to the reference image, Simple Linear Iterative Clustering (SLIC) method [32] is used for splitting the reference image and then each splitting area is compared to other input images in order to find ghost regions [29].

SLIC is one of the segmentation algorithm. It has a different distance measurement that enables compactness and regularity in the superpixel shapes. SLIC generates superpixels by clustering pixels based on their color similarity and proximity in the image plane. CIELAB color domain is used in this algorithm, because it is

considered as perpetually uniform for small color distances. The clustering procedure begins with an initialization step. For an image with *N* pixels and desired number of approximately equally-sized superpixels *K*, the approximate size of each superpixel is therefore *N/K* pixels. Initial cluster centers are sampled on a regular grid spaced $S = \sqrt{N/K}$ pixel apart as follows:

$$C_i = [l_i \; a_i \; b_i \; x_i \; y_i] \tag{3.11}$$

Next, in the assignment step, each pixel *i* is associated with the nearest cluster center whose search region overlaps its location, this brings speeding up the algorithm by reducing the number of distance measurements. Then, distance measurement is done as follows:

$$D = \sqrt{d_c^2 + (\frac{d_s}{S})^2 \, m^2} \tag{3.12}$$

where,

$$d_c = \sqrt{(l_j - l_i)^2 + (a_j - a_i)^2 + (b_j - b_i)^2} \tag{3.13}$$

$$d_s = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \tag{3.14}$$

where, d_c is the color distance measure, d_s is the spatial distance measure, *m* is the constant to regularize the weight the relative importance between color similarity and spatial proximity, thus when *m* is large spatial proximity is more important and the resulting superpixels are more compact.

In this thesis, *m* is selected as 1, so that, the resulting superpixels adhere more tightly to image boundaries, but have less regular size and shape, and superpixel size is selected as 20-pixel.

In order to get the ghost regions, comparison is done by using the normalized cross correlation with zero mean for each splitting area. Normalization must be done because of the reducing the effect of exposure times of the input images. Zero mean normalized cross correlation (ZNCC) between the rest of images and the reference

one is presented. ZNCC is a measure factor of comparison of two super-pixels computed as:

$$ZNCC_x = \frac{\sum_{(i,j)\epsilon S_x} D_{ref}(i,j) D_{input}(i,j)}{\sqrt{\sum_{(i,j)\epsilon S_x} D_{ref}(i,j)^2} \cdot \sqrt{\sum_{(x,y)\epsilon S_x} D_{input}(i,j)^2}} \quad (3.15)$$

where,

$$D_{input}(i,j) = I_{input}(i,j) - \mu(I_{input}(i,j)) \quad (3.16)$$

$$D_r(i,j) = I_r(i,j) - \mu(I_r(i,j)) \quad (3.17)$$

where; I_r stands for reference image, I_i stands for *i-th* exposure image, S_x represents for the *x-th* super-pixel, μ is the average operator.

After this calculation for each superpixel, a threshold, *t=0.55*, is applied to determine whether that super-pixel is ghost region or not. By using this threshold binary maps are created for each input image except reference one a follows:

$$M_x(i,j) = \begin{cases} 1, & when\ ZNCC > t \\ 0, & otherwise \end{cases} \quad (3.18)$$

These binary maps are used for masked out the ghost areas in related input image. Then, modified weights are normalized, and exposure fusion algorithm is used for creating the final enhanced image.

Instead of pixelwise comparison between input images, this algorithm uses the superpixel groups. This approach is different from other algorithms. In addition, global calibration of the input images is not implemented. Calibration is done within the each superpixel cluster.

3.2.6 Zero Mean Cross Correlation comparison between 4x4 pixel-groups [ALG-6]

For this algorithm, an improved version of ALG-5 is used. Instead of using SLIC method, whose complexity is high, a 2D 4x4 pixel window is used for grouping the pixels. Without any segmentation method, reference image is divided into groups as follows:

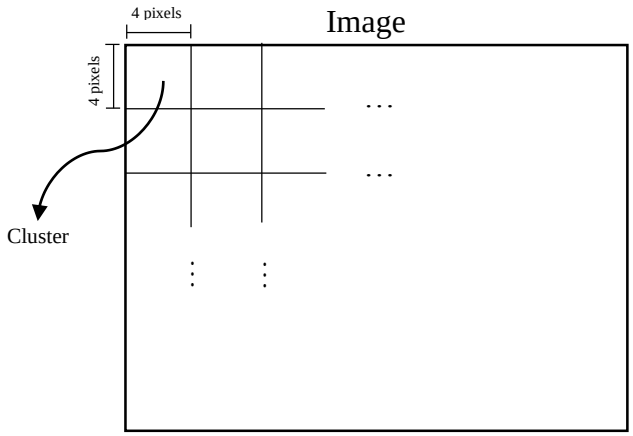


Figure 3-2 Grouping the Pixels with 4x4 Window

After this grouping, same ZNCC method, and thresholding is used for detection of the ghost areas. However, at the edges of the ghost areas, grouping results as false detections. In order to overcome this situation morphological operations are done.

After getting final binary images, which represent ghost areas, is used for mask the related weights of input image. Then, normalization of the weights is done before using the exposure fusion algorithm to get the enhanced image.

Using 4x4 pixel-groups without a segmentation helps to speed up the ALG-5. This algorithm is used in order to compare it with the ALG-5.

3.2.7 Improved Histogram Based Ghost Removal in Exposure Fusion for High Dynamic Range Images [ALG-7]

Multi-level threshold map is used to reduce the effect of the varying exposures at the LDR images [27]. Unlike [11], multi-level threshold map has more than one threshold. These levels are labelled as rank of a pixel. Rank of a pixel is found according to the cumulative distribution function of intensity map of the input image. If median is selected as threshold, then its level is specified as 2, it is similar to median threshold map [11]. Then, multi-level is designated to be used as multi thresholding based on cumulative distribution function. After finding each rank of pixels, it is normalized as follows:

$$\hat{r}_{i,j,k} = round\left(\frac{r_{i,j,k} - 1}{R_k - 1}\right), \qquad 0 \le \hat{r}_{i,k} \le 2^N - 1 \tag{3.19}$$

where, R_k is the last rank in *k-th* exposure image, $r_{i,j,k}$ is rank of pixel (*i,j)* in *k-th* exposure image, $\hat{r}_{i,j,k}$ is the normalized value of $r_{i,j,k}$, finally *N* is the threshold levels. Higher this value is, better performance at the final enhanced image is. However, increasing the value of *N* is resulted in higher computational load.

After then, differences of each image rank to the reference image ranks are computed as follows;

$$d_{i,j,k} = | \hat{r}_{i,j,ref} - \hat{r}_{i,j,k} | \tag{3.20}$$

where, subscript *ref* represents the reference image. With the help of $d_{i,j,k}$, ghost map is found for each input image such a way that:

$$M_{i,j,k} = \begin{cases} 1, & for\ k = ref \\ 0, & for\ d_{i,j,k} \geq T, k \neq ref \\ 1, & otherwise \end{cases} \tag{3.21}$$

where, *T* is the threshold value. Non-ideal intensity changes result in the wrong detection of ghost. To avoid this, morphological operations are applied to ghost maps. Firstly, erosion operation and then dilation operation are conducted to remove isolated pixel, and to connect disconnected pixels. Finally, holes surrounded by 1-value pixels are filled.

Therefore, final ghost map $M_{i,j,k}$ is found, and this ghost map is acted as a mask to the weight map of the exposure fusion.

In this thesis, rank is selected as 255. When the rank is selected as 255, cumulative distribution function of all input images are linearized, excluding the values of 0 and 255. In addition, histograms tend to have equally distributed. Therefore, this algorithm making a global calibration without a reference image. However, when thresholding, it needs a reference image. It differs from other algorithm in the calibration process.

3.2.8 Majority Voting [ALG-8]

ALG-8 is a proposed method. Majority voting is conducted in such a way that if four of seven algorithms are labelling a pixel in their ghost maps, this pixel is selected as ghost region. Otherwise, this pixel is not labelled as ghost region.

The main aim of using majority voting is to analyze the overall performance of the implemented ghost removal algorithms. By using majority voting, number of false detected pixels are reduced, because, implemented algorithms do not have same reasoning to get the false detection. The analyzes about this are given in Chapter 4.

As it is explained before ALG-1 and ALG-3 do not need a reference image, also ALG-3 creates one ghost map, not for all input images. However, in order to be convenient, reference image is selected for these algorithms, and comparison is done such a way that, for ALG-1 all images are compared to the reference image with same assumptions, and ghost map of each input image is created. For ALG-3, median threshold map of reference image and other inputs are compared by using exclusive-or operation, and ghost map of each input image is created.

3.3 Improvements and Grouping of the Methods

In this section, improvements that are done for each algorithm are given, and classification of the reviewed algorithms is provided.

For ALG-1, exposure time of the images are taken into account in order to find the abrupt change in the intensity of the pixel between reference image and input image. This leads to detect ghost pixels if the background lowly exposed.

ALG-2, and ALG-7 are adapted so that number of input images is limitless. Dynamic range is improved if the number of input images is increased. Therefore, in the final enhanced image, pixel weights are obtained from differently exposed input images.

ALG-6 is created to reduce the complexity of ALG-5. SLIC method is an exhaustive segmentation method, because of this, 4x4 blocks without looking the characteristics of the pixel in the groups are created in order to detect the ghost regions. Moreover, morphological operation is conducted to minimize the false detections.

None of the algorithms gives the logical explanation about the selection of the reference image. For this purpose, in the input set, the image whose average well-exposedness measure is maximum is selected as reference image. This gives a brief information about the histogram. Well-exposedness measure is used to keep

intensities that are not near zero (underexposed) or one (overexposed). Therefore, by looking at just the raw intensities within a channel, reference image has the maximum average of well-exposedness value.

Grouping of these methods are done in terms of photometrically calibration of the input images, threshold tuning, reference image selection, grouping. This overview gives the brief introduction to the algorithms. Classification of implemented ghost removal algorithms in this thesis are given in Figure 3-3.

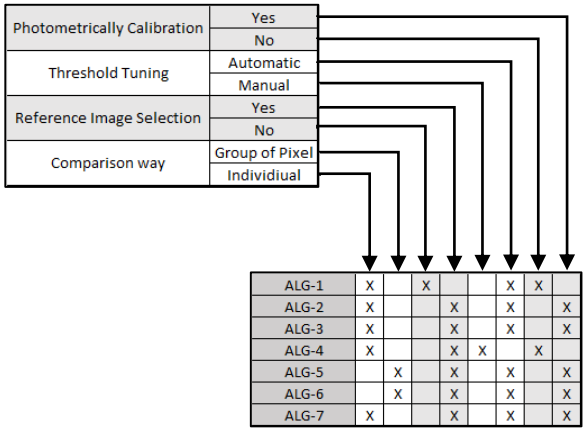


Figure 3-3 Classification of Ghost Detection and Removal Methods

Among these algorithms, ALG-1 is selected as a simple baseline algorithm, ALG-2 is selected as it creates probability map instead of binary map to represent ghost regions, ALG-3 is selected as the representative of the category which aims to completely eliminate moving objects, ALG-4 is selected as using different color domain and adaptive thresholding, ALG-5 is selected as segment the pixel by using SLIC method, ALG-6 is selected as grouping the pixel without any state-of-art segmentation, ALG-7 is selected as it uses histogram equalization to calibrate the input image.

CHAPTER 4

COMPARISON AND DISCUSSION

4.1 Introduction

Until now, the implemented algorithms have been explained by giving the advantages and disadvantages of the algorithms with respect to the authors of the algorithms. However, most of these studies do not include a detailed comparison between proposed and former algorithms. Therefore, in this chapter detailed outputs of each algorithm are presented and related discussions are made. Also detailed explanations of the ghost maps that are created by each algorithm are given.

For simplicity, acronyms of each of the ghost reduction technique and exposure fusion technique are used. The implemented algorithms are named as in Table 4-1.

Table 4-1 Acronyms and Description of the Algorithms

Acronyms	Description of the Algorithms
EF	"Exposure Fusion", *Tom Mertens, Jan Kautz, Frank Van Reeth* [15]
ALG-1	"Ghost Detection and Removal in High Dynamic Range Images", *Desire Sidibe, William Puech, Oliver Strauss* [7]
ALG-2	"A Simple Ghost-Free Exposure Fusion for Embedded HDR Imaging", *Y.-S. Moon, Y.-M. Tai, J. H. Cha and S.-H. Lee* [26]
ALG-3	"Bitmap Movement Detection: HDR for Dynamic Scenes", *F. Pece and J. Kautz* [11]
ALG-4	"An Exposure Fusion Approach Without Ghost for Dynamic Scenes", *C. Wang and C. Tu* [28]
ALG-5	"Ghost detection and removal based on super-pixel grouping in exposure fusion", *S. Jiang, X. Zhihai, L. Qi, C. Yueting and F. Huajun* [29]
ALG-6	"4x4 Pixel Grouping and ghost removal using ZNCC algorithm"
ALG-7	"Improved Histogram Based Ghost Removal in Exposure Fusion for High Dynamic Range Images", *D.-K. Lee, R.-H. Park and S. Chang* [27]
ALG-8	Majority Voting

All of algorithms are implemented in MATLAB environment, except exposure fusion algorithm, of which implementation is obtained from Mertens et al. [15].

For all the experiments, image sets give in [6] and [33] are used. The image sets are given in Appendix A. There are twelve datasets. The image sets have different characteristics, such as, object motion type and magnitude, object motion color and histogram distribution. In addition, ground truth of the ghost regions are extracted by hand for 1st input set, in order to find the percentages of the false and correct detection for each algorithm. These binary images contain the data of the pixel movement.

Since, there are a large number of ghost maps that are outputted for each algorithm, all the ghost maps for each algorithm are not given. However, related discussions are made for each algorithm in their respective sections.

This section is organized as follows, 1) summary of the ghost removal algorithms are presented 2) for each algorithm, related experimental results and discussions are shown 3) comparison of the computation times are given and percentages of false and correct detections are shown using ground truth 4) objective evaluation with the final enhanced images are shown and related comparisons are given.

4.2 Summary of Ghost Removal Algorithms

ALG-1 compares the intensity values, if the monotonic behavior of a pixel is lost between input images with increasing exposure time; this pixel is labelled as ghost region. In addition, rapid change in the intensity values is also labelled as ghost region.

ALG-2 makes comparison between the reference image and the input images. This comparison is done by making histogram of the input images equal to the reference image. Exponential function is used to find the ghost probability of each pixel to create the ghost map.

ALG-3 creates ghost map by comparing the median threshold maps, ghost map for each input image is created by comparing the binary images. Finally, morphological operations are done in order to minimize the effect of noise.

ALG-4 uses L channel in the *CIELAB* domain to make the comparison instead of RGB domain. After selecting the reference image, others are calibrated to the reference image by adding the difference of the mean values at the L channel. Then, thresholding is applied and ghost map for each input are created.

ALG-5 uses a rank based comparison between the intensity images such that intensity images divided into regions using cumulative distribution function. After that a ranked image is created, if there is a difference between the ranked-intensity image and the reference image, these areas are specified as ghost regions.

ALG-6 is done such a way that a reference image is segmented by super-pixel grouping. For this purpose, SLIC method is used. After that, normalized cross correlation with zero mean is computed to classify the ghost regions and these regions are extracted from the initial LDR images.

ALG-7 is an improvement version of the ALG-7. Instead of using SLIC method, 4x4 pixel grouping is used to group the pixel. Within these groups, normalized cross correlation with zero mean is used to detect the ghost regions.

ALG-8 is using majority voting such that if five of seven algorithms are labelling a pixel in their ghost maps, this pixel is selected as ghost region. Otherwise, this pixel is not labelled as ghost region. For this purpose, since ALG-2 uses the ghost probability values, a threshold of 0.25 is used to convert that image to the binary image. In other words, if the ghost probability of a pixel is smaller than 0.25, it is labelled as a ghost in the ghost map. For ALG-1, all images are compared to the reference image with same assumptions, and ghost map of each input image is created. For ALG-3, median threshold map of reference image and other inputs are compared by using exclusive-or operation, and ghost map of each input image is created.

In order to compare the algorithms, reference image of a dataset is chosen the same for all algorithms, and image whose well-exposedness measure has maximum average is selected as the reference image.

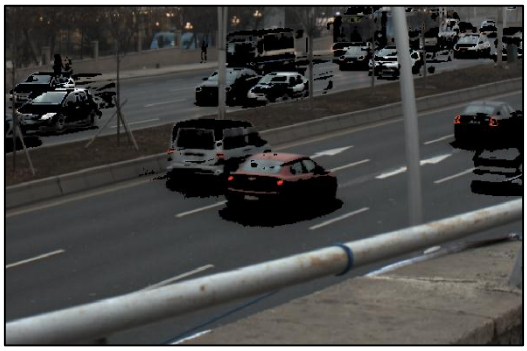
4.3 Experimental Results for the Ghost Removal Algorithms

Until now, the algorithms have been explained individually without giving any results. By using different image sets, we aim to compare the algorithms and see their weaknesses and strenghts with respect to the other algorithms.

4.3.1 ALG-1

As it is said, ALG-1 compares pixel intensities. The main advantage of this algorithm is that its complexity is lower than the other algorithms, because, there is not any calibration between input images before creating the ghost map. However, comparing only the intensity values without any preprocessing causes some problems, such as, if intensity level of an object in motion is the same or slightly larger than the background, ghost detection cannot be achieved accurately. This problem can be seen in Figure 4-1. In this figure, black regions show the ghost pixels. As a conclusion, this algorithm is not useful for these kind of input set.

(a)



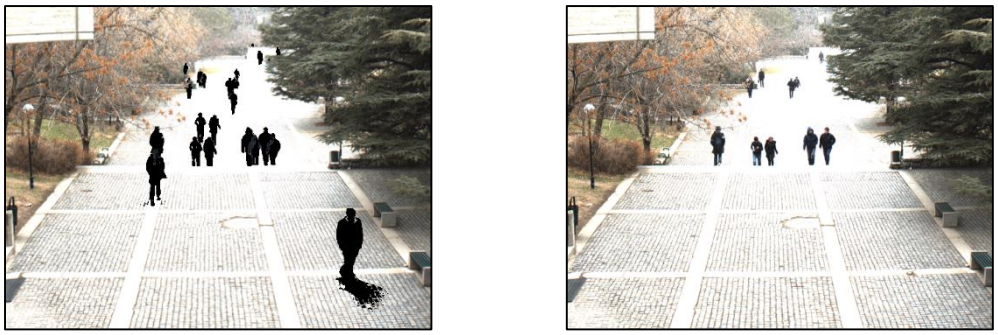
(b)

Figure 4-1 Ghost maps and the Related Part of Image of the (a) Input Set 1, (b) Input Set 4

On the other hand, this algorithm succeeds to extract the ghost regions, if the background of the object in motion contains highly exposed pixels. Because, the difference of the intensity levels between background and the object in motion is more distinct. Successfully extraction of this kind of areas are seen in Figure 4-2, black regions in Figure 4-2 (a) show the ghost regions. As can be seen in Figure 4-2 (b), background contains highly exposed pixels, so that ghost map are created

without any deficiency. Note that, ghost regions which are not derived from this input image are result of the reference image.

Morphological operations are not implemented to reduce the noise or false detection. However, it is seen that the percentage of the non-ghost regions labelled as ghost is very low for this algorithm. This is expected because assumption of the increasing in the intensity with increase in exposure time is always valid if the photographs are taken in consecutive times with a functioning camera. Therefore, disadvantages of the morphological operations are not seen in this algorithm, such as annihilation of the small movement detection. This advantage is depicted in the Figure 4-3.



(a) (b)

Figure 4-2 (a) Ghost Map and Related Part of 9th Input Image of Input Set 12, (b) Related Part of the 9th Input Image of Input Set 12 without Ghost Map

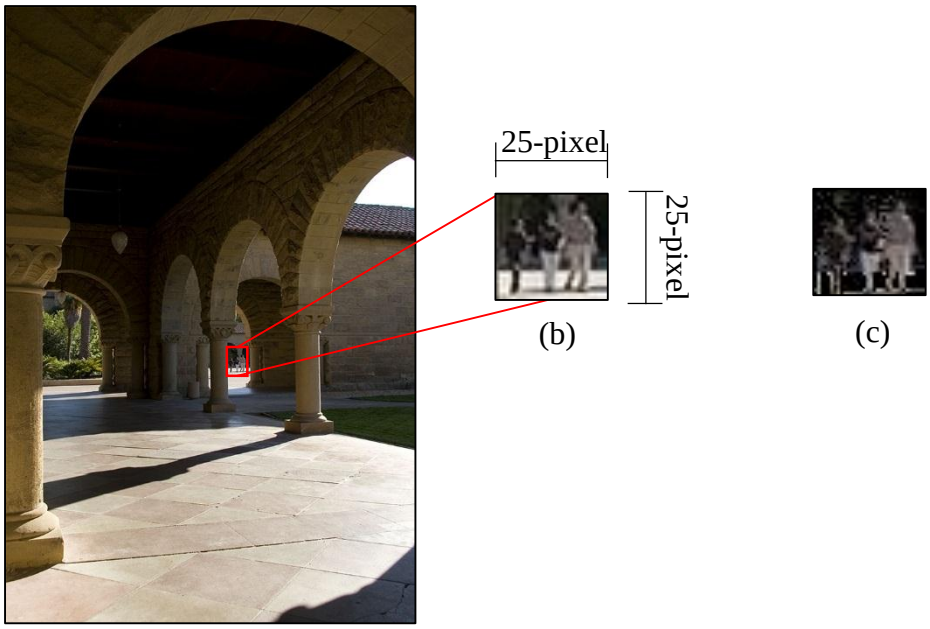


Figure 4-3 (a) 3rd Input Image of the Input Set 1, (b) Region where Small Movement Occurs, (c) Detected Ghost in this Region

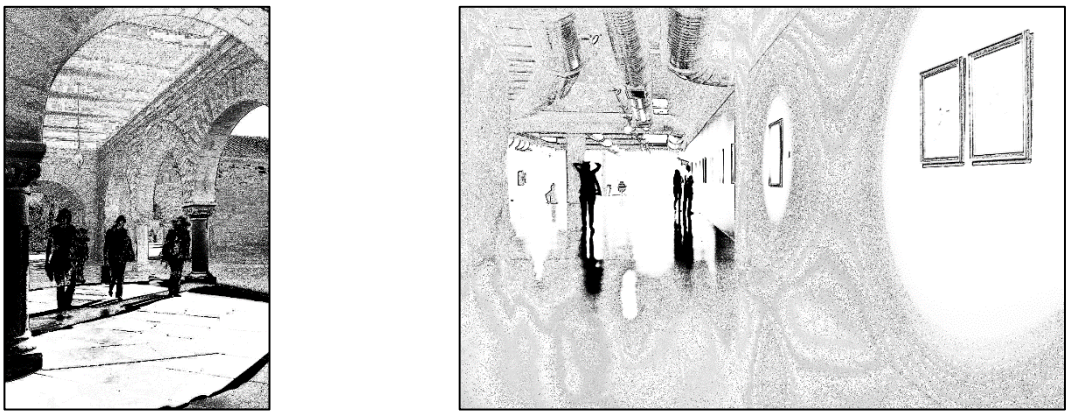
In Figure 4-3 (c), black regions show the non-ghost areas. The kernel sizes used for the dilation and erosion of the ghost map affect the final results and a good balance between the erosion kernel size and the detected ghost size is required. Although, erosion operation helps to reduce the noise and false detection, it excludes the small ghost regions, such a way that, if the size of the detected ghost regions is smaller than the size of the kernel, ghost region is removed. Therefore, erosion operation reduces the performance of ghost detection in such cases. As a conclusion, it is seen that this algorithm is good to find the ghost regions where small objects are in movement.

To sum up, this algorithm is depending on intensity level of the object in motion and intensity level of the background of the object. These determine the quality of this algorithm. On the other hand, complexity of this algorithm is very low. In addition, for small movement detection, it has an advantage.

4.3.2 ALG-2

ALG-2 uses an approach to reduce the effect of the exposure difference in such a way that after the selection of the reference image, other exposure images are

photometrically calibrated toward the reference image, to do this the histogram matching between the reference image and their individual exposure image is used. Then non-ghostness probabilities of each pixel is found with exponential function, as explained before. Since exponential function is used, all the pixels have the probability of belonging to the ghost region. This degrades the final enhanced image, because, when weights are calculated according to the probabilities, the pixel weights are reduced, therefore, dynamic range of the final enhanced image are reduced. In addition, since pixelwise differences are calculated, there are lots of noise in the ghost map. These problems are depicted in Figure 4-4.



(a) (b)

Figure 4-4 (a) Ghost Map of the 4th Image of the Input Set 1, (b) Ghost Map of the 5th Image of the Input Set 7

The main source of noise is salt-and-pepper noise which is brief sudden intensity spikes. This can be seen in Figure 4-4 (b), non-ghost pixels are labelled as ghost with mainly 1-pixel long. It should be noted that there is not preprocessing to eliminate noise.

With increasing exposure time, features in small size of the scene cannot be obtained, this is mainly because of the capturing sensor characteristic. Since, representing scene with a one pixel is achieved by taking average of the neighboring sensor nodes, the scene characteristics in the vicinity of the feature are also important. For example, in Figure 4-5, two of the images from the input set 12 are shown. The red squares on each image show the same region. However, the feature information of these red squares are not same for each input image. Highly exposed background degrades the Figure 4-5 (b), because features in the red square are not

visible. This causes a problem for this algorithm, because, when comparing these two input images, the intensity difference in the red square is so high that these pixels has high ghost probability. However, it is a false detection. The related ghost map of 6th image of input set 12 is shown in Figure 4-6, note that for this set 8th image is selected as the reference image. As a summary, the static regions in front of the highly exposed background have high ghost probability in this algorithm, because of the capturing device characteristics.

(a) (b)

Figure 4-5 (a) 6th and (b) 8th Images of Input Set 12

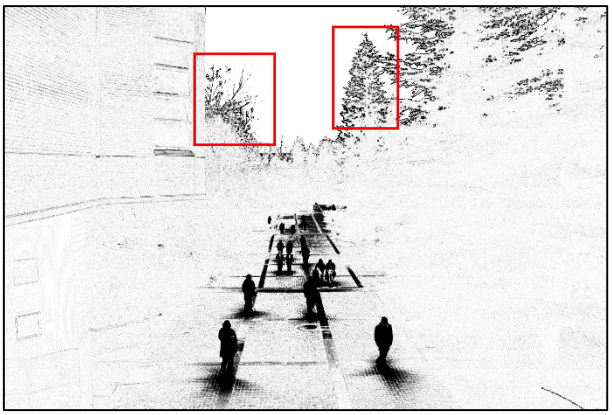


Figure 4-6 Ghost Map of 6th Input Image of Input Set 12

Exponential function characteristics is seen in Figure 4-7 with image noise level of 2, note that difference values are changing from 0 to 20. In this figure, it is seen that intensity difference of 10 correspond to nearly 0.08, which means that a pixel whose difference of 10 is extracted out from the final enhanced image. Because of this, another problem occurs in the process of histogram matching toward the reference

image. This problem mainly occurs at the pixel which possess saturated values at the reference image but not in the other images. At the histogram matching method, unsaturated pixels become more saturated to match the histograms. However, limited histogram of reference image causes problems, histogram matching toward a limited histogram, make a wide histogram to shrink and lose actual intensity information of the image. For example, in Figure 4-8 (a), input image has a wide histogram, however, after equalization, maximum value is seen at 255. Therefore, the well-exposed pixels are become more saturated after equalization. The probabilities have more values at these point, because histogram equalization is not good at the highly exposed pixel. One of the example is given by using input set 8.

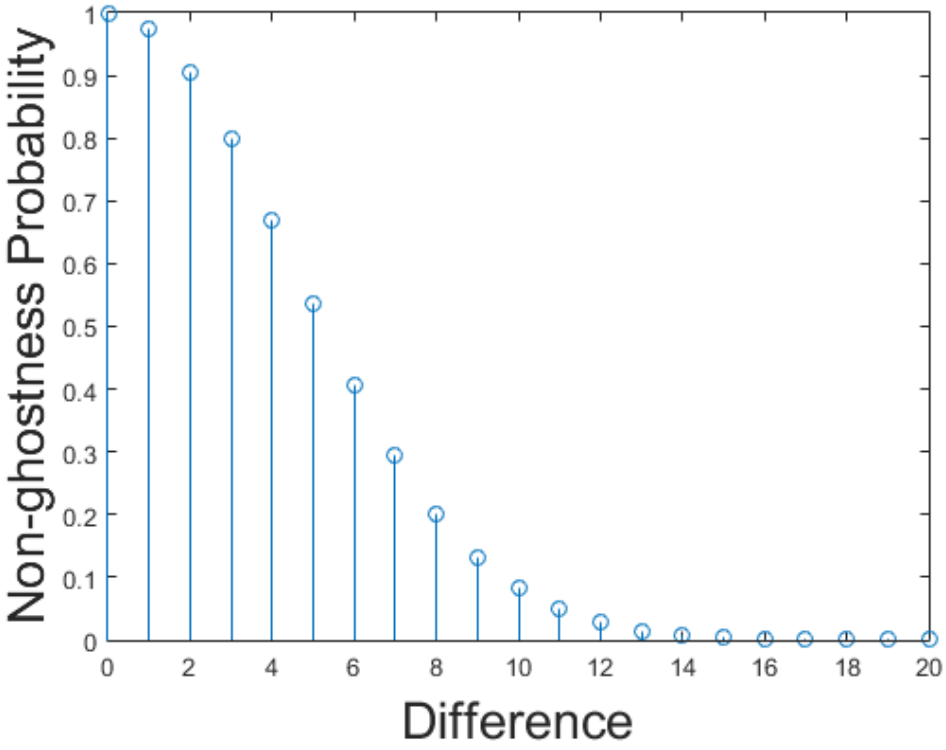
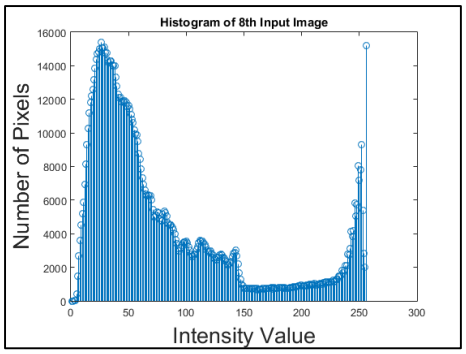
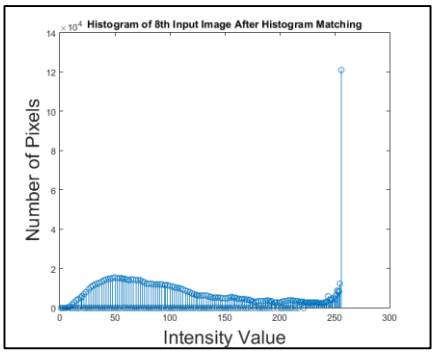


Figure 4-7 Exponential Function used for ALG-2



(a)



(b)

Figure 4-8 Histograms of 8th Input Image of Input Set 8, (a) Actual, (b) After Histogram Matching toward Reference Image

Pixelwise operation without morphological operations makes this algorithm to find the small movement detection, like ALG-1. In addition, histogram matching method reduce the dependency of the intensity level of the object in motion, it is used to reduce the effect of exposure time, thus, all the inputs possess same characteristic as if they are taken at the same exposure value. However, mapping between histograms can cause some problems. Firstly, some bins in the matched histogram have 0 pixel, such as making a wide histogram to the limited histogram or vice versa. Secondly, a pixel in the object in motion can have the same intensity with reference image, after histogram matching, but, these problems have limited effects because multi-dimensional operation in exposure fusion is greatly reduce these kind of pixels.

To sum up, histogram matching is good for eliminating the effects coming from the exposure time. Also pixelwise difference are useful to detect small movement detection. However, reference image characteristics, such as, number of saturated or unsaturated pixels, limited histogram has effect on the creating the ghost map. In addition to that, all pixels have non-ghostness probability which reduce the actual dynamic range of the final enhanced image.

4.3.3 ALG-3

ALG-3 is using binary images to get the ghost areas. Instead of using histogram equalization, median threshold map is used. This reduces the complexity of the algorithm, compared to histogram equalization as in ALG-7.

There are limited false detections of under- and over-exposed areas, because of using median threshold map, these regions are in different bins independently from exposure time of input image.

Ghost regions could not be fully found, if intensity values of these regions are related to the background. Reason of this problem is because of the median threshold map. If, background and moving pixels are in the same histogram bin, ghost detection could not be achieved. Therefore, this algorithm is dependent on the background characteristics. For example, areas in red rectangles in Figure 4-9 show the deficient detection because of the above reasoning.

Figure 4-9 4th Input of the Input Set 1 and ALG-3 Ghost Map

Morphological operations are used in this algorithm. Although, it reduces effect of the noise and false detections, erosion operation leads to exclude small movement detection in ghost map. This is a disadvantage of this algorithm.

To sum up, ALG-3 has low complexity and low percentage of false detections in under- and over-exposed pixels, however, background intensity level affects the ghost labelling process, so it is dependent on the object and background intensity levels, like ALG-1.

4.3.4 ALG-4

ALG-4 is using L channel in *CIELAB* domain is used to get the intensity level of the scene, because *CIELAB* is an absolute reference space for color. Whereas RGB are relative spaces and can have various different definitions. L component in *CIELAB* domain specifies intensity, whose levels are in between [0 100]. Unlike others, adaptive threshold is used for each pixel. Adaptive threshold assignment has mainly two advantage, firstly, calculation of the deviation from mean brightness value ($(\frac{l_{i,j,k}-50}{\sigma})^2$) results in larger threshold value for the pixels that brightness values are near 0 or 100 (underexposed or overexposed), secondly, exposure difference between the reference image and the related input image is taken into account ($|\Delta l_k|^{\beta}$). These calculations greatly reduce the number of unnecessary labelling. Correctly classification of non-ghost regions improves the performance of the final enhanced image in terms of dynamic range, because more pixels contributed at the fusion. Calibration is done by adding a DC level as explained before. However, this causes problems if the reference image is dominated by highly exposed pixels, because, this kind of reference image does not show the characteristics of the scene. These are at the limits of the capturing device, thus; average value of the L channel is high. Therefore, adaptive threshold values β and σ should be selected according to the input set. In this thesis, adaptive selection of these values are not implemented, this is done in future work. Effect of these values are shown in Figure 4-10. For this input set, percentage of pixels that are in [90, 100] is nearly %38. Note that, for this input set, reference image is 9th image.

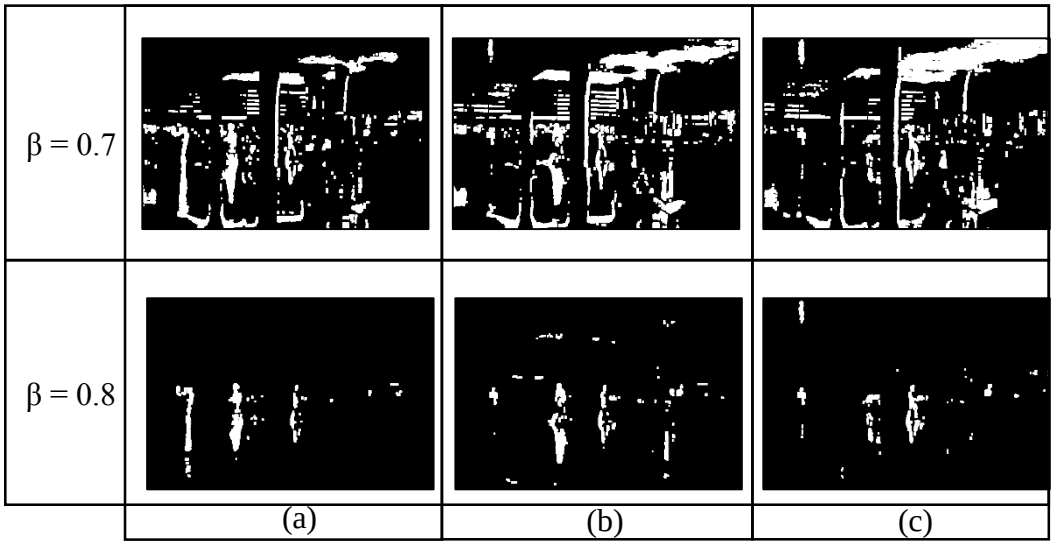


Figure 4-10 Effect of β Values used in Adaptive Threshold in Image Set 2, (a) Ghost Maps of 6th Image, (b) 7th Image, (c) 8th Image

In Figure 4-10, it is seen that increasing β greatly reduces false detections. Increasing β results in increasing of $|\Delta l_k|^{\beta}$. Therefore, it is more robust at the highly exposed areas. When adaptive thresholding, these pixels are not labelled as ghost. However, β should be carefully selected, because, actual ghost regions may not be detected if this threshold is high. Therefore, in terms of false detections, increasing β reduces the effect of the highly exposed pixels in the reference image.

Morphological operations are used in this algorithm. Although, it reduces the effect of the noise coming from rapid change in pixel values, erosion operation leads to exclude small movement detection in ghost map. This is a disadvantage of this algorithm.

Adding DC level in order to calibrate is not sufficient if background luminescence level is close to the actual ghost pixel. Normalization of the image in order to reduce the effect of exposure time cannot be satisfied by adding a DC level, because capturing device limits this approach. In theory, the luminescence value captured by the sensor is linearly dependent on exposure time. If the exposure time is doubled, then intensity value of captured pixel is doubled. However, this assumption is not valid for capturing devices. That is why, adding a DC value does not normalize the images.

To sum up, using adaptive threshold for each pixel is good for the minimization of the false detection, however, adaptive threshold values should be carefully selected for each input set, since it is dependent on the average intensity value of the reference image.

4.3.5 ALG-5

Unlike comparing pixel values, super-pixel groups are used for comparison in ALG-5. First of all, it groups the pixels by its similarity, using SLIC method. After that, ZNCC (Zero Mean Normalized Cross Correlation) values for each group of pixels in each input image are compared with the reference image, before thresholding.

The main problem of ZNCC calculation arises when deviation in values of the super pixel at the reference image is very low. This problem leads to unnecessary labelling such a way that ZNCC calculation does not give reasonable outputs. Consider the case where the superpixel in reference image is following:

$$I_{ref} = \begin{bmatrix} 150 & 149 & 149 \\ 149 & 149 & 148 \\ 150 & 150 & 150 \end{bmatrix}$$

And, same superpixel in an input image:

$$I_{input} = \begin{bmatrix} 129 & 128 & 128 \\ 129 & 128 & 128 \\ 128 & 128 & 128 \end{bmatrix}$$

ZNCC calculation of these two matrix is 0.3162, since it is below the threshold value this superpixel is labelled as ghost. However, there is not any ghost pixel, they are the regions with smooth changing in luminescence. Average of the pixels in the pixel groups are close to the values of the pixel such that nominator of the ZNCC is in the vicinity of 0. As a conclusion, using pixel groups with ZNCC gives such problems when average of pixel groups very close to the pixel values. This problem can be seen in the Figure 4-11. The seams in the red square is labelled as ghost in most of the ghost maps for this set.

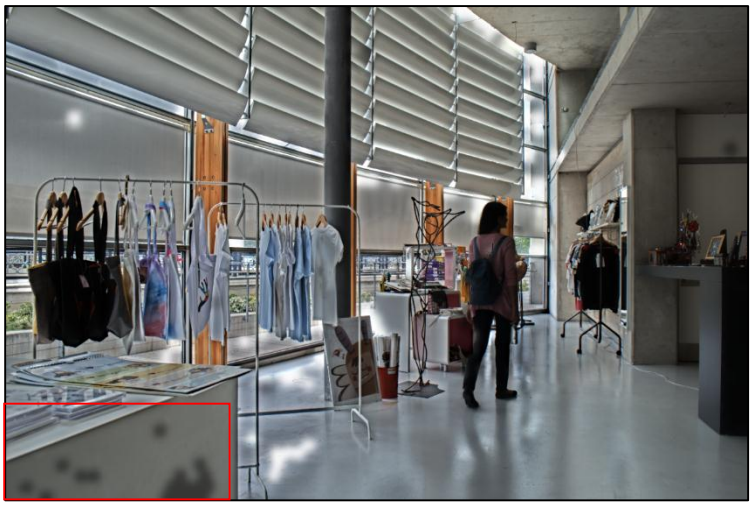
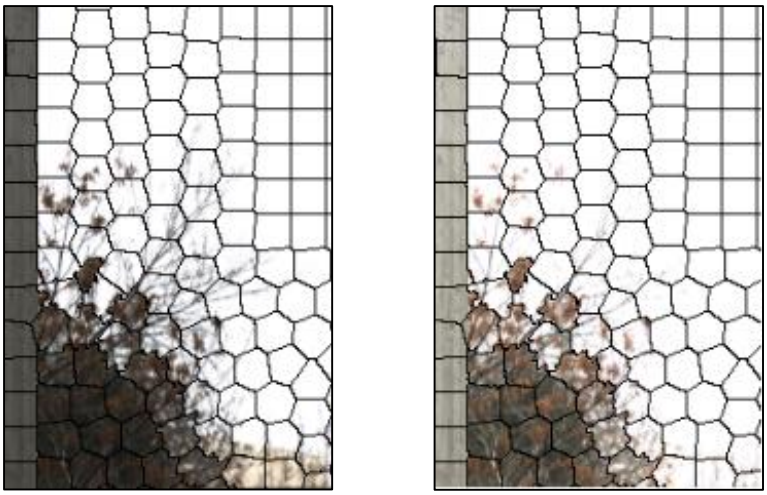


Figure 4-11 Final Enhanced Image of ALG-5

SLIC method is used for segmentation of the image by comparing the salient features and color information. In order to segment all features in a scene, reference image should be properly exposed. However, due to the capturing device characteristics, high luminance at the background effects the foreground object. This problem can be seen in Figure 4-12. Branches of the tree cannot be properly exposed. Since, ALG-5 is comparing group of pixels, these regions have resulted as ghost region such that, in the low-exposures these branches are more distinct, and when comparing ZNCC values the weight of the pixels of the branches are weighted more. Therefore, there is the detection of the ghost. It is concluded that to make the group of pixel, all the salient information should be properly exposed.

In addition to that, ZNCC calculation of the highest exposed pixelgroup, whose pixel intensity levels are 255 for all pixels, is not logical, because, average value of this kind of pixelgroups are equal to the pixel values, so that, ZNCC value is undefined. Therefore, reference selection of this algorithm is important to avoid these kind of problems.



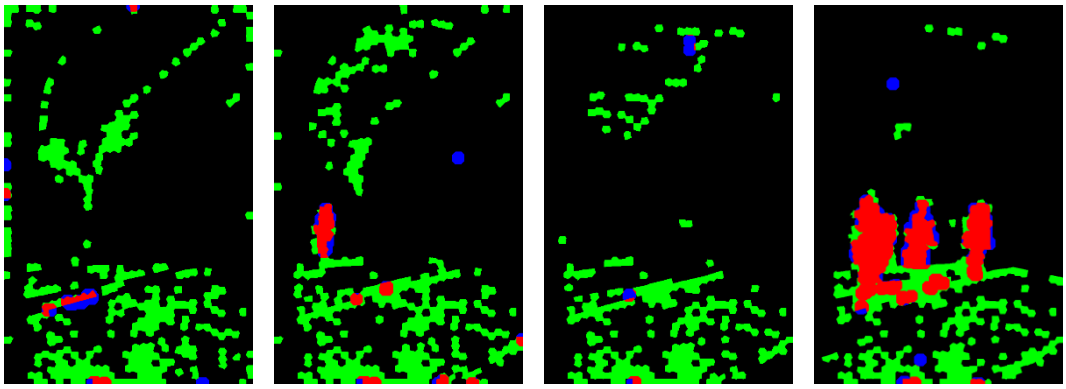
(a) (b)

Figure 4-12 Lost in Information due to the Luminescence of Background, (a) Part of the 7th Input Image of Image Set 12 with Superpixel Groups, (b) Reference Image

To sum up, unlike pixelwise comparison, superpixel grouping is used in this algorithm and this makes the complexity of this algorithm higher than others. Problems due to the grouping and ZNCC calculation makes this algorithm be dependent of the reference image and input set. High percentage of the highly exposed pixel in the reference image effects badly for this algorithm.

4.3.6 ALG-6

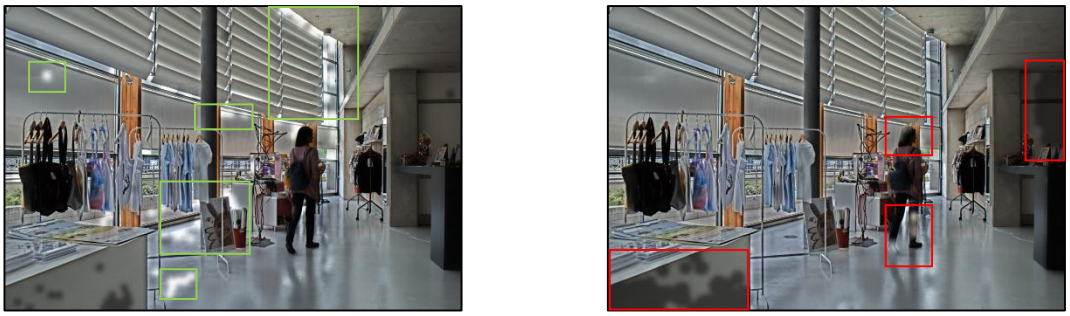
4x4 pixel groups are used for ALG-6, it has less complexity compared to ALG-5 because of the SLIC method used in grouping. In addition, in order to avoid the false detections caused by ZNCC calculations, morphological operations are used. For example, there is an improvement in input set 1, by the means of reducing false detection while conserving actual ghost regions. The improvement can be seen in the Figure 4-13 where red areas represents areas of intersection, green areas represents areas in ALG-5 but not in ALG-6, and blue areas represents areas in ALG-6 but not in ALG-5. According to the differences, it is clearly to say that output of ALG-5 has a lot of false detection than ALG-6. Under- and over-exposed areas are detected as ghost regions. This problem is avoidable by using morphological operations.



(a) (b) (c) (d))

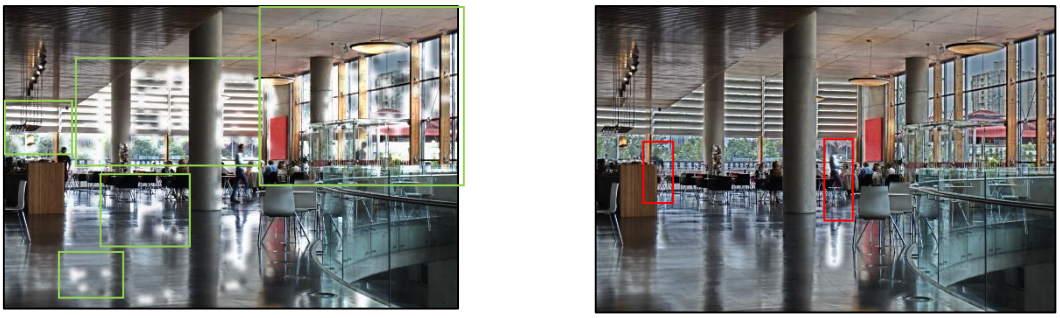
Figure 4-13 Differences between Ghost Maps of ALG- 5 and ALG-6 for Input Set 1, (a) Differences in 1st Image, (b) 2nd Image, (c) 3rd Image and (d) 4th Image

Although, using morphological operations reduces percentage of false detection, erosion operation leads to exclude small movement detection in ghost map. This is a disadvantage of this algorithm. Morphological operation is help to reduce the number of seams that are created in the ALG-5, however, for the large group of unwanted seams, it is not effective. These situations can be seen in Figure 4-14 and Figure 4-15. Green rectangles show the improved areas; red rectangles show the worsened areas due to the morphological operations.



(a) (b)

Figure 4-14 Final Enhanced Image of (a) ALG-5 and (b) ALG-6 for Input Set 11



(a) (b)

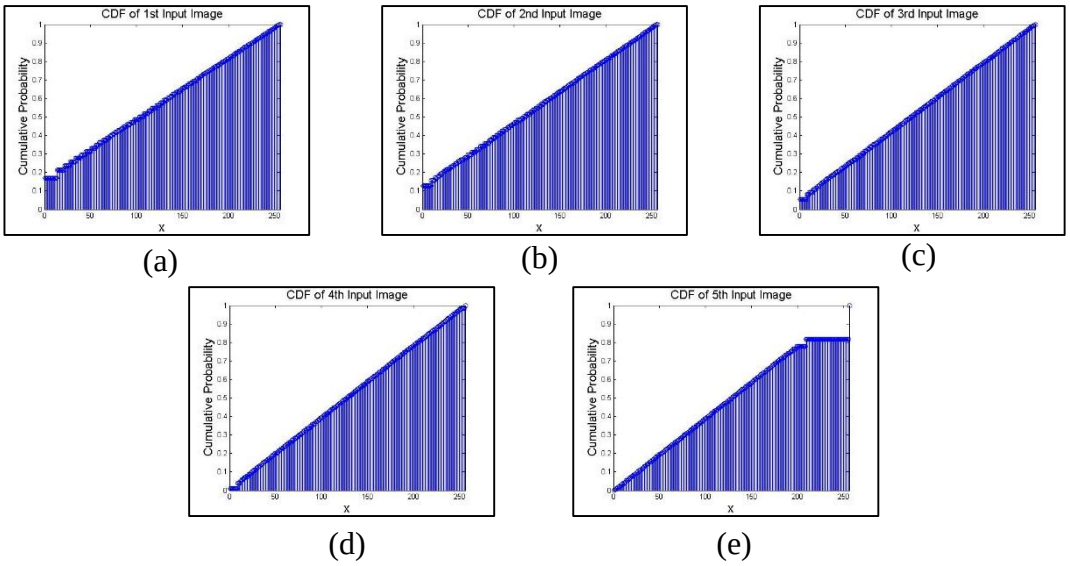
Figure 4-15 Final Enhanced Image of (a) ALG-5 and (b) ALG-6 for Input Set 2

To sum up, unlike segmentation, 4x4 grouping is used in this algorithm and this makes the complexity of this algorithm lower than ALG-5. Seams due to the SLIC grouping and ZNCC calculation are slightly reduced by using this algorithm.

4.3.7 ALG-7

Normalization of the image in order to reduce the effect of exposure time is obtained by the histogram equalization in ALG-7. In other words, cumulative distribution function of each input set is linearized for each input image. In order to find the ghost pixels, difference between histogram equalized image and histogram equalized reference image is calculated. After that, thresholding is applied to the difference measures.

However, depending on the number of the least exposed and most exposed pixels, CDF equalization behaves differently. Example of this situation is given in Figure 4-16, in this figure image set 1 is used. CDF of reference image (5^{th} image) has abrupt change at level of 256, and between 208 to 256, there is not any differences in the CDF. The problem is that when finding the difference between each input to the reference image, these places are resulted higher than expected. Therefore, applying thresholding to these places are resulted in non-ghost pixels labelled as ghost. Therefore, percentage of non-ghost labelling is larger than expected.



(a) (b) (c)

(d) (e)

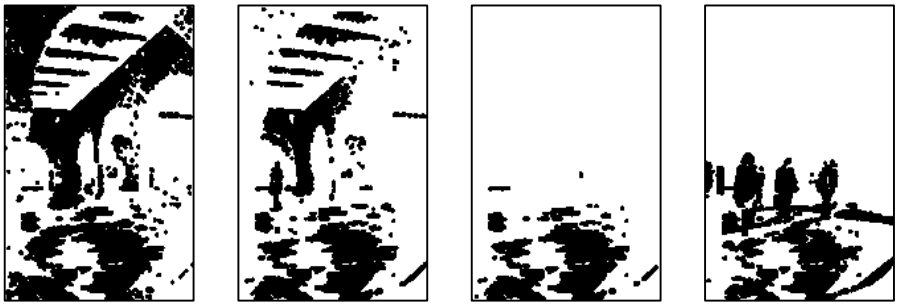
Figure 4-16 CDF of Input Set 1 after CDF Equalization, (a) 1st Image, (b) 2nd Image, (c) 3rd Image, (d) 4th Image and (e) 5th Image

As it is concluded from Figure 4-16, linearization of the CDF is successfully achieved for 4th image, however, since main ghost regions are in that image, reference image is selected as 3rd image for comparison. The resulted enhance images when reference is 5th image and 3rd image is seen in Figure 4-17. It is concluded that reference image selection should be done by comparing the CDF of input set.

(a) (b)

Figure 4-17 Final Enhanced Image of ALG-7, when (a) 5th Image as Reference, and (b) 3rd Image as Reference

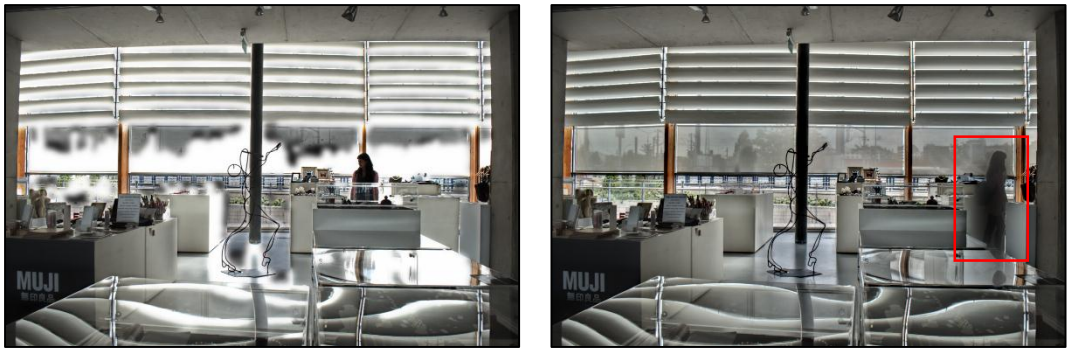
Resultant ghost maps by using input set 1 can be seen Figure 4-18. As expected, due to the deficiency of linearization in the CDF of reference image, there are high percentage non-ghost labelling.



(a) (b) (c) (d)

Figure 4-18 Ghost Maps of ALG-7 when Input Set 1 is used, (a) 1st Image, (b) 2nd Image, (c) 3rd Image, (d) 4th Image

Selection of the reference image can be adaptively applied. Such a way that, number of least exposed pixels and most exposed pixels in an input image is found, and input image that has the minimum number of these pixels is selected as reference. The improvements in the enhanced images are seen in the Figure 4-19 and Figure 4-20.



(a) (b)

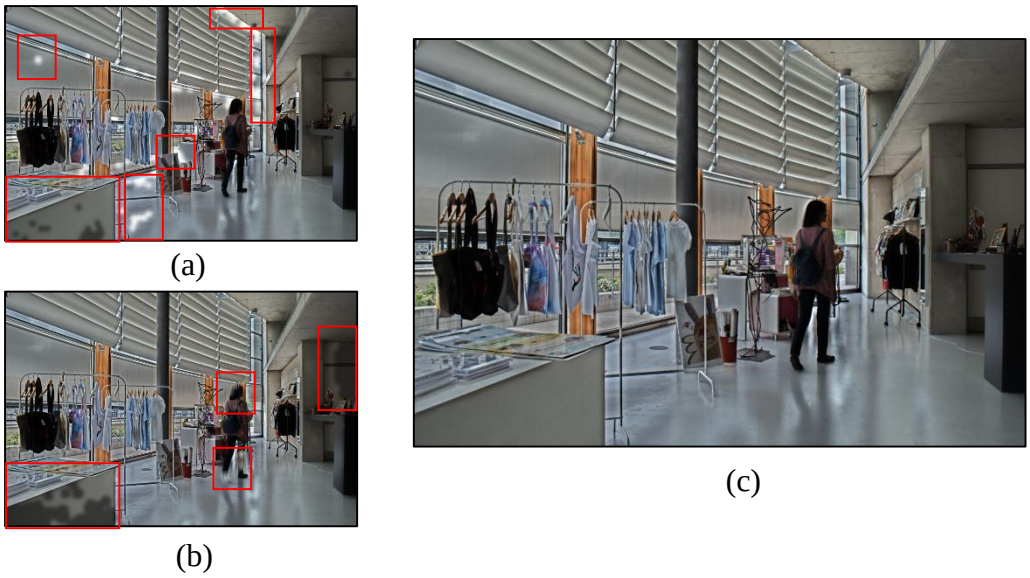
Figure 4-19 Final Enhanced Images of Input Set 10, when Reference Image Selection by Evaluating, (a) Well-Exposedness, (b) Percentage of Saturated and Unsaturated Pixel

(a) (b)

Figure 4-20 Final Enhanced Images of Input Set 4, when Reference Image Selection by Evaluating, (a) Well-Exposedness, (b) Percentage of Saturated and Unsaturated Pixel

In the red squares, part of the ghost regions is shown. Blending into the final image is not satisfied at these areas. Pixel information of these areas are from the reference image, because, in all ghost map these parts are selected as ghost. Since, adaptive selection makes the reference image be the low contrast than before, these blending problems are arisen. Therefore, reference selection by comparing number of least exposed pixels and most exposed pixels reduces the number of the false detection, however, degradation in ghost regions is obtained in the final enhanced image.

ALG-5, and ALG-6 have illogical outputs when ZNCC equation is conducted in the smooth regions, ALG-7 does not show this problem. Because, pixelwise comparison and thresholding make this algorithm be robust in these region. The problematic regions of ALG-5 and ALG-6, and the resulted regions of ALG-7 are given in Figure 4-21.



(a)

(b)

(c)

Figure 4-21 Final Enhanced Images of (a) ALG-5, (b) ALG-6 and (c) ALG-7

In the red rectangles, unwanted seams are given for ALG-5 and ALG-6, the reason behind these seams have been explained in their respective section. The enhanced image of ALG-7 has satisfied results at these regions, because histogram equalization is done globally. Using pixel groups with ZNCC calculation is not logical for such input set where smooth regions creates the majority of the scene.

As a consequence, histogram equalization is robust choice to eliminate the effects of the exposure value of each image, however, when thresholding, reference image characteristics effects the unwanted ghost labelling.

4.3.8 ALG-8

ALG-8 is a proposed method. Majority voting is conducted in such a way that if five of seven algorithms are labelling a pixel in their ghost maps, this pixel is selected as ghost region. Otherwise, this pixel is not labelled as ghost region.

As it is explained before that ALG-1 and ALG-3 do not need a reference image, also ALG-3 creates one ghost map, not for all input images. However, in order to be convenient in this subsection, reference image is selected for these algorithms, and comparison is done such a way that, for ALG-1 all images are compared to the reference image with same assumptions, and ghost map of each input image is created. For ALG-3, median threshold map of reference image and other inputs are

compared by using exclusive-or operation, and ghost map of each input image is created.

Majority voting is used to detect the overall performance. Aim of this algorithm is combining all the ghost maps created by each algorithm. For this purpose, a color map is created. Different color codes are given to the algorithms, such a way that, if all the algorithms label same pixel as ghost, its color is pure white, in other words, red, green and blue channel values of this pixel are equal 255. If none of the algorithms label a pixel as ghost, its color is pure black.

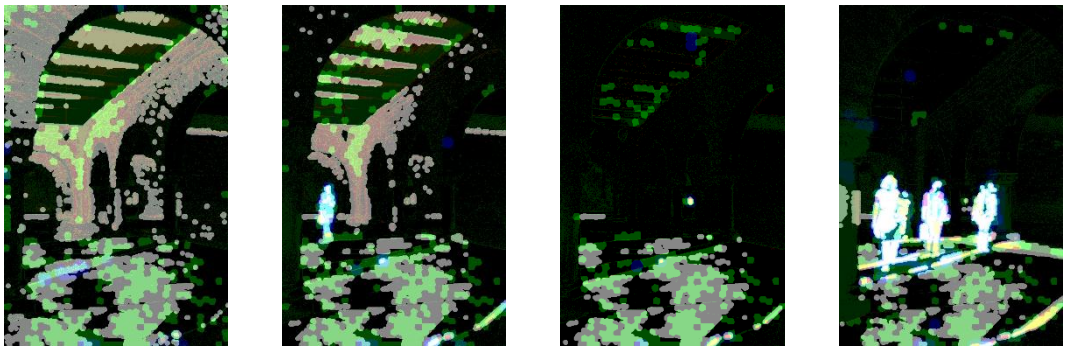
For each algorithm, different labels are set as following:

Table 4-2 Algorithm Labelling

	Red Channel	***Green Channel***	***Blue Channel***
ALG-1	40	0	0
ALG-2	0	40	0
ALG-3	0	0	40
ALG-4	80	0	0
ALG-5	0	80	0
ALG-6	0	0	80
ALG-7	135	135	135

For example, if a pixel is selected as ghost pixel by both ALG-1 and ALG-2, its value is [40 40 0] and its color is yellowish. Addition is conducted, if more than one algorithm finds a pixel as ghost. Since ALG-8 is used in order to compare each algorithm, all the ghost color maps are shown, and related discussions are given.

Ghost color maps of the input set 1 is seen in Figure 4-22.

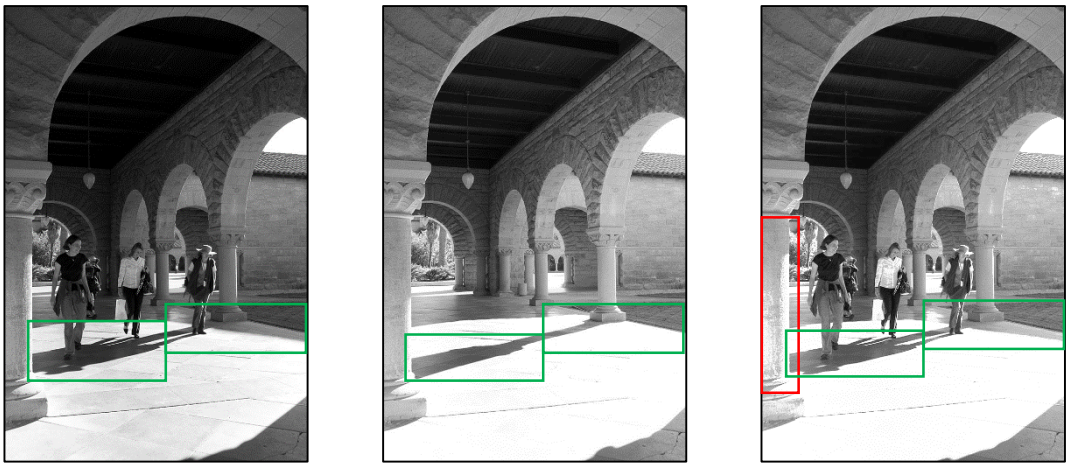


(a) (b) (c) (d)

Figure 4-22 Ghost Color Maps of Input Set 1, Outputs of (a) 1st Image, 2nd (b) Image, (c) 3rd Image, (d) 4th Image

Low exposed areas are detected as ghost region for ALG-1, ALG-2, ALG-5 and ALG-7, especially in the first color map. For ALG-1, when pixels are multiplied with the ratio of exposure values, amplification of the noise occurs. Luminescence value captured by the capturing device is so low that linearity is lost in those areas, thus, sudden changes of the intensity levels are occurred. Therefore, ALG-1 is prone to noise when ratio of exposure value increases. For ALG-2, it is said that histogram matching method is not correctly achieved when the lowly exposed pixel and highly exposed pixel are dominated in the image. Limited histogram of first input image causes failure when mapping of the pixels. Therefore, intensity difference is so high that false detection occurs in these regions. As it is said, ZNCC calculation in ALG-5 fails when smooth changing areas. False detections occur in floor of the scene is because of the reference image selection. In the reference image, these areas are highly exposed and change in the pixel intensities are very low. Therefore, ZNCC calculation gives illogical results. For ALG-7, histogram equalization is not able to linearize CDF function because of the high number of the high exposed pixels, this problem is explained previously.

ALG-2 finds false detection in the fourth color map. Main problem is the change in the number of mostly exposed pixel drastically. This is explained by giving visual aid. In Figure 4-23, grayscale image of 4th input, reference grayscale image and histogram matched grayscale image of 4th input are seen.



(a) (b) (c)

Figure 4-23 Grayscale Image of (a) 4th Input, (b) Reference Grayscale Image, (c) Histogram Matched Grayscale Image of 4th Input

Regions in green rectangles of Figure 4-23 (a) show the regions where low exposed pixels in 4th image become highly exposed pixel in reference image, because of the shadow change and object motion. When mapping these pixel to the reference image, these pixels are not going to the highly exposed pixels, as can be seen in Figure 4-23 (c). Therefore, well-exposed pixels in 4th image are mapped to the over-exposed pixels when histogram matching. As can be seen in the red region, well-exposed pixels become highly exposed, and then to find the ghost regions, these regions are resulted as ghost. This problem is not seen in 3rd image as expected, because the number of highly exposed pixels are mostly conserved between 3rd image and reference image. This can be seen in Figure 4-24.

(a) (b)

Figure 4-24 Grayscale Image of (a) 3rd Input, (b) Histogram Matched Grayscale Image of 3rd Input

False detected regions of ALG-3 are because of the median value. The pixels near median are resulted in different bins in each image so that they are labelled as ghost pixels. However, these false detections are not so much compared to other algorithms, this number is greatly reduced by the dilation operation. ALG-4 has successful results, because this set contains mostly well-exposed pixels. Therefore, using DC level to reduce the effect of exposure value is succeed. Using morphological operations in ALG-6 reduce the false detections. As it has given, illogical outputs of ZNCC calculation increase the false detections. However, using dilation operation reduce the number of false detected regions.

Another comparison is done with the input set 12, because the characteristic of this image set is different than input set 1; firstly, ghost regions are smaller, secondly, dynamic range of the scene is larger and finally object movement occurs mainly in front of the high-exposed background. Ghost color maps of input set 12 are seen in Figure 4-25 for each input image. Reference image of this input set is 8th image.

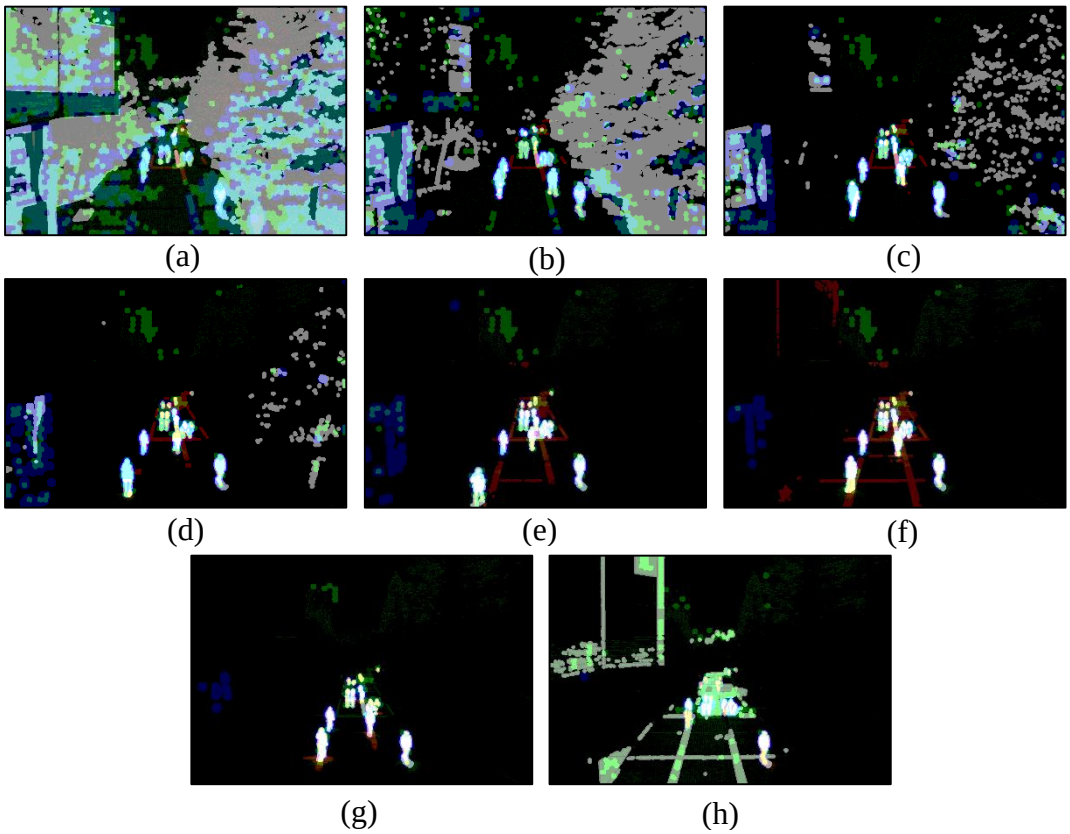


Figure 4-25 Ghost Color Maps of Input Set 12, Outputs of (a) 1st Image, (b) 2nd Image, (c) 3rd Image, (d) 4th Image, (e) 5th Image, (f) 6th Image, (g) 7th Image, (h) 9th Image

Previous explanations hold for this input set, about finding false detection for ALG-1, ALG-2, ALG-5 and ALG-7. These false detections are mainly increased with increasing exposure difference between the reference and input image.

As it is explained before, segmentation results false detections for ALG-5. This is explained in the respective section of ALG-5.

Red pixels in Figure 4-25 (f) represent false detections of ALG-4. These regions are overly exposed in reference image, while in 6th image they are well exposed. Calculation of adaptive threshold is failed to these regions, such a way that, second term, $((\frac{l_{i,j,k}-50}{\sigma})^2)$, gives nothing but a low value. However, difference of these pixel with the reference image is large, because they are highly exposed in that region. Therefore, these regions are labelled as ghost.

These comparisons have been done with different input set. Problems that is observed are stated. For other input set, these problems occur, however, for simplicity, selected input sets are used. For others, color maps of the all input sets can be seen in the Appendix B.

As a conclusion, exposure value difference increases the number of false detections of ALG-1 because of noise amplification. Other algorithms do not show this problem because, firstly, morphological operations help to eliminate this kind of regions, secondly, histogram equalization with threshold and histogram matching methods reduces the noise effect. On the other hand, reference selection is main issue for histogram related calibration. Characteristics of the reference image is greatly affect the performance of these approaches. In addition, segmentation of the reference image is logical when all the salient features are distinguishable. Adaptive threshold selection is dependent on the values; values should be carefully selected with respect to the scene.

4.4 Comparison with respect to Computational Time and Ground Truth

In this subsection, percentages of non-ghost and ghost labelling are given by using input set 1. This approach is conducted by comparing ground truth and ghost maps of each algorithm. This ground truth and related input set can be seen in Figure 4-26. Note that, ground truth is a binary image where black regions represent a

moving object on a static background. Most of the motion occurs in the 4th image. In addition, there are smaller size of moving object in the 3rd image. Moreover, there is not any moving object in the 1st and 5th image. In this section, these ground truths are used to compare the performance of the algorithms for this input set.

In addition to that, computational times of each algorithm are given for input set 1. Computational time of the algorithm is only dependent on the size of input images, because none of the algorithms uses any iterative or adaptive approach.

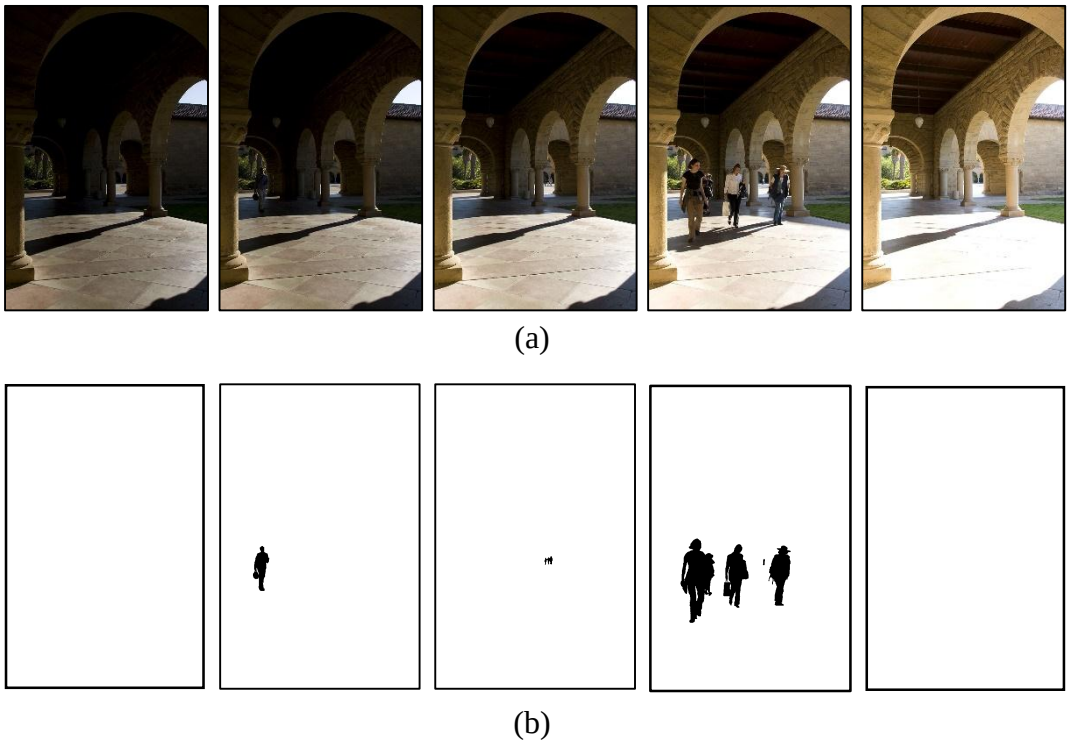


Figure 4-26 (a) Input Set 1, (b) Ground Truth Extracted

Reference image is the 5th image for this set. Because, its average well-exposedness measure is 0.1920. Average well-exposedness measures of others are 0.1792, 0.1526, 0.0687 and 0.1588, for 1st, 2nd, 3rd and 4th image, respectively.

The number of ghost pixel and non-ghost pixel of this dataset is given in Table 4-3.

Table 4-3 Number of Ghost and Non-ghost Pixels in the Input Set 1

	1st Image	2nd Image	3rd Image	4th Image
Number of Ghost Pixels	0	4447	428	37277
Number of Non-ghost Pixels	688128	683681	687700	650851

Table 4-4 shows a binary classification of the implemented algorithms. For this reason, true-positive (TP), true-negative (TN), false-positive (FP) and false-negative (FN) classifications are used. TP is a correct result where ghost pixels are correctly detected as ghost. TN is also a correct result where non-ghost pixel is correctly detected as non-ghost. However, FP is an error in which non-ghost pixel is labelled as ghost. FN is an error where ghost pixel is not labelled as ghost pixel. In Table 4-4, the number of pixels for each classification is seen. It should be noted that 5th image is selected as reference image by selecting the image which has the most average value of well-exposedness measure. In Table 4-4, each input image except reference image is classified as explained before.

In terms of non-ghost labelling, ALG-7 is the worst, as explained in its respective section, CDF linearization degrades the non-ghost labelling if the reference image has a great number of highest exposed pixels. Using majority voting for ALG-8 is greatly decreased the non-ghost labelling, this is expected because there is not any common reasoning between the algorithms in terms of false detections, for example, ALG-2 is prone to the exponential function, on the other hand, ALG-5 gives false detections because of the ZNCC calculation.

Table 4-4 Binary Classification of the Implemented Algorithms

		1st Image	2nd Image	3rd Image	4th Image
ALG-1	***TP***	0	3982	311	28507
	TN	653176	643794	636043	627264
	FP	34952	39887	51657	23587
	FN	0	465	117	8770
ALG-2	***TP***	0	4108	376	35216
	TN	483920	522332	617993	513414
	FP	204208	161349	69707	137437
	FN	0	339	52	2061
ALG-3	***TP***	0	4346	344	32613
	TN	684076	674533	682709	599142
	FP	4052	9148	4991	51709
	FN		101	84	4664
ALG-4	***TP***	0	2831	153	32803
	TN	687526	680803	686236	630234
	FP	602	2878	1464	20617
	FN		1616	275	4474
ALG-5	***TP***	0	4277	328	36143
	TN	578310	573044	607443	558478
	FP	109818	110637	80257	92373
	FN		170	100	1134
ALG-6	***TP***	0	4084	0	35477
	TN	679050	674193	684033	623724
	FP	9078	9488	3667	27127
	FN		363	428	1800
ALG-7	***TP***	0	4263	205	35349
	TN	358636	488976	594524	537289
	FP	329492	194705	93176	113562
	FN		184	223	1928
ALG-8	***TP***	0	3785	114	31887
	TN	687803	681306	686872	638814
	FP	325	2375	828	12037
	FN		662	314	5390

In terms of ghost labelling, ALG-6 loses information because of the morphological operations, in Image 3. However, for the large sizes of the movement it is satisfactory. ALG-5 properly detects object motions in 4th Image, superpixel grouping helps to achieve for this input set. On the other hand, ALG-1 is failed to extract ghost regions, because, the intensity level of the background of the object in motion is related to the intensity level of the object.

Computational time of each algorithm is given in Table 4-5. It should be noted that computational time ALG-8 is equal to the summation of computational times of others. Therefore, for comparison, ALG-8 is not taken into account.

Table 4-5 Comparison of the Computation times

Algorithm	**Time (sec)**
ALG-1	0.231
ALG-2	0.364
ALG-3	2.571
ALG-4	1.972
ALG-5	150.643
ALG-6	31.035
ALG-7	3.267
ALG-8	189.488

ALG-1 has the minimum computation time. There is not any morphological operation which increases the complexity of the implemented software code. Only pixelwise differences are computed.

In addition, there is gap between computation times of ALG-2 and ALG-3. This situation is result of the morphological operations done in ghost map of each input. While, exponential function is used for ALG-2, ALG-3 is implemented with dilation and erosion operations. Therefore, it is concluded that complexity of algorithm is greatly increased with morphological operations.

ALG-5 and ALG-6 has the maximum two of the computation time. This is mainly because of partitioning image to group of pixel and computing zero mean normalized cross correlation for each partition. This greatly increases the computation time. ALG-5 has the most computation time. Unlike ALG-6, using SLIC method to partition the image makes huge load of computation.

As a conclusion, grouping methods are greatly increased the complexity of the algorithms, while, global image calibrations, such as histogram matching, histogram equalization have reasonable computation times.

4.5 Visual Results

The goal of this subsection is to give some visual results that have been obtained from the algorithms. As it has been mentioned before, there are twelve image sets, and each image set consists of differently exposed images. There are eight ghost removal algorithms. Therefore, total number of outputs is nearly one hundred. It is not possible to show all the images, so some of the images are given in this subsection. The outputs of the best and the worst performed ghost removal algorithms with respect to the area of unwanted seams, conserving the dynamic range, success of ghost removal are given for each image set. In the following figures, the image on the left belongs to the algorithm performs best, the middle one belongs to the worst algorithm, and right one belongs to the exposure fusion output without any ghost removal algorithms. The figures are given as follows:

(a) (b) (c)

Figure 4-27 Enhanced Images of Image Set 1, (a) ALG-6, (b) ALG-1, (c) EF

(a) (b) (c)

Figure 4-28 Enhanced Images of Image Set 2, (a) ALG-4, (b) ALG-5, (c) EF

(a) (b) (c)

Figure 4-29 Enhanced Images of Image Set 3, (a) ALG-5, (b) ALG-1, (c) EF

(a) (b) (c)

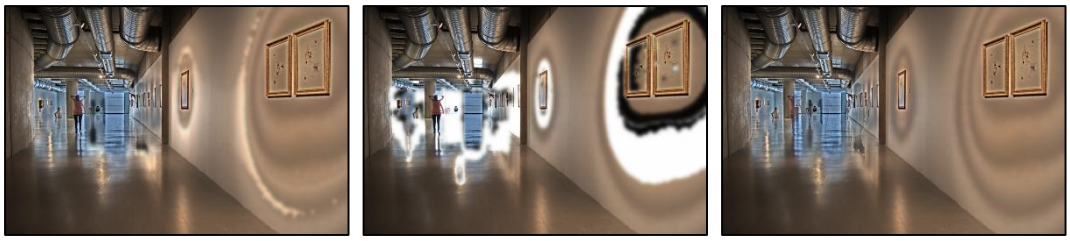
Figure 4-30 Enhanced Images of Image Set 4, (a) ALG-2, (b) ALG-7, (c) EF

(a) (b) (c)

Figure 4-31 Enhanced Images of Image Set 5, (a) ALG-2, (b) ALG-5, (c) EF

(a) (b) (c)

Figure 4-32 Enhanced Images of Image Set 6, (a) ALG-1, (b) ALG-4, (c) EF

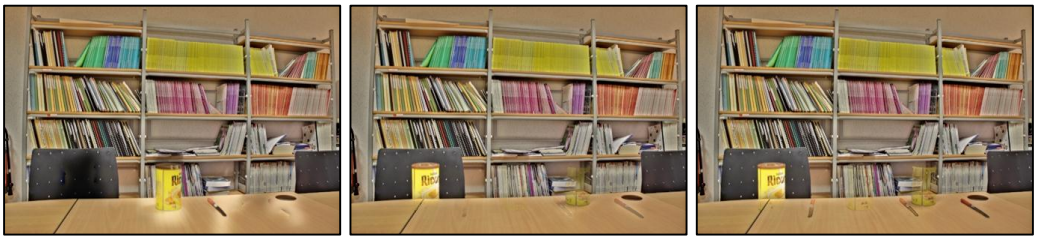


(a) (b) (c)

Figure 4-33 Enhanced Images of Image Set 7, (a) ALG-4, (b) ALG-7, (c) EF

(a) (b) (c)

Figure 4-34 Enhanced Images of Image Set 8, (a) ALG-2, (b) ALG-6, (c) EF



(a) (b) (c)

Figure 4-35 Enhanced Images of Image Set 9, (a) ALG-5, (b) ALG-1, (c) EF

(a) (b) (c)

Figure 4-36 Enhanced Images of Image Set 10, (a) ALG-4, (b) ALG-5, (c) EF

(a) (b) (c)

Figure 4-37 Enhanced Images of Image Set 11, (a) ALG-7, (b) ALG-5, (c) EF

(a) (b) (c)

Figure 4-38 Enhanced Images of Image Set 12, (a) ALG-4, (b) ALG-6, (c) EF

Using exposure fusion to make enhanced image creates seams. These regions mainly occur at the edges where information of one region is selected from the lowly-exposed image and the other is selected from highly exposed image. This is expected because multi-dimensional blending technique uses the low pass filter and downsampled images in each level thus when increasing levels, edge is not conserved. On the other hand, decreasing the level of the blending results in unrealistic images with noise, it has been explained previously. The evaluation of the performance of the exposure fusion algorithm is out of the scope of this thesis.

Until now, ghost removal algorithms are analyzed with the input sets where there is movement in the scene at the time of capture. Ideally, if there is not any movement in the scene at the time of capture, ghost removal algorithms should not find any ghost regions. However, this is not seen in the algorithms. For this purpose, image set 13 is used to evaluate the performance of the algorithms with static scene. This image set does not contain any movement; it contains a color checker with twenty-four different colors. Since, there are large number of ghost maps for each algorithm, a binary classification table is used to show the performance of the algorithms. This classification is done such a way that, false-positive (FP) is an error in which algorithm improperly indicates presence of a ghost pixel when in

reality there is not, true-negative (TN) is a correct result in which algorithm properly indicates presence of a non-ghost pixel. Since, there is not any movement in this set, true-positive (TP) and false-negative (FN) classifications are not used. For image set 13, classification table is shown in Table 4-6. It should be noted that reference image for this set is selected as 8^{th} image.

Table 4-6 False and True Detection Performance of the Image Set 13

		1st Image	2nd Image	3rd Image	4th Image	5th Image	6th Image	7th Image	9th Image	10th Image
ALG-1	***TN***	324035	459694	615460	719683	720893	720896	720886	720784	720896
	FP	396861	261202	105436	1213	3	0	10	112	0
ALG-2	***TN***	157646	456104	563607	553762	552134	586093	573761	448866	312689
	FP	563250	264792	157289	167134	168762	134803	147135	272030	408207
ALG-3	***TN***	665473	715154	683389	719006	719951	719951	719006	679594	679897
	FP	55423	5742	37507	1890	945	945	1890	41302	40999
ALG-4	***TN***	720896	720896	720896	720896	720896	720896	720896	720535	720896
	FP	0	0	0	0	0	0	0	361	0
ALG-5	***TN***	280412	347801	392243	435308	463097	481639	476711	547680	591907
	FP	440484	373095	328653	285588	257799	239257	244185	173216	128989
ALG-6	***TN***	294788	202468	101102	134253	219305	259692	247778	545999	638692
	FP	426108	518428	619794	586643	501591	461204	473118	174897	82204
ALG-7	***TN***	155909	182557	674614	680792	699459	699662	699833	689188	441521
	FP	564987	538339	46282	40104	21437	21234	21063	31708	279375
ALG-8	***TN***	383025	488424	686983	720896	720896	720896	720896	720896	706576
	FP	337871	232472	33913	0	0	0	0	0	14320

As it is concluded from the table, ALG-4 has the lowest number of false positive pixels. Adding DC value to calibrate the images and using adaptive threshold helps to reduce the number of false detected pixels. Increasing in the difference in exposure value between input image and the reference image results in increasing false detected pixels. Number of improperly exposed pixels degrades the performance of ghost removal algorithms, such a way that, calibration toward reference image is not satisfied at these regions. On the other hand, these regions are not contributed to the final enhanced image because, the weights calculated in exposure fusion are negligible compared to the properly exposed pixel. Therefore, degradation in the final enhanced image is not noticeable. Since, image set 13 contains color checker board, each color cell shows the characteristics of smooth regions. Therefore, ZNCC comparison results a great number of false positives for

ALG-5 and ALG-6. These algorithms are failed to label the non-ghost regions at these kind of regions, as explained before. Final enhanced image of these algorithms has the most degradation in the dynamic range compared to others. Using majority voting reduces number of the false positives for the images where under- and over-exposed pixels are limited. Therefore, this algorithm is better choice to get the actual ghost region and to reduce the number of false detected pixels. Ghost color map of input set 13 can be seen in the Appendix B, Figure B-13.

As a conclusion, there is not any best algorithm for each input set. It has been explained before, reference image selection effects the performance of the output enhanced image. Best in one dataset could be worst in other because of the number false detected regions. These regions are mainly created from the algorithms that use calibration of the images, such as histogram matching and equalization, zero mean cross correlation. However, these algorithms are successful to extract the ghost regions, by minimizing the effects of the exposure difference. False detections are not main problem for other algorithms, however, extracting the actual ghost area, these algorithms are dependent on the scene characteristics. Therefore, for the enhanced outputs, global calibration methods have satisfactory results with selection of the proper reference image.

As a summary, implemented algorithms are summarized in Table 4-7, in terms of the differences in features. Performances of the algorithms with different characteristics of scene is given in Table 4-8.

Table 4-7 Features of the Algorithms

		Features							
		Color Domain		**Morphological Operation**		**Threshold**		**Representation of Ghost Regions**	
		RGB	CIELAB	Yes	No	Yes	No	Pixel Based	Patched Based
ALGORITHMS	ALG-1	X			X		X	X	
	ALG-2	X			X		X	X	
	ALG-3	X		X			X	X	
	ALG-4		X	X		X		X	
	ALG-5	X			X	X			X
	ALG-6	X		X		X			X
	ALG-7	X		X		X		X	

Table 4-8 Performance of the Algorithms in Terms of Scene Characteristics

		Performance of Algorithms				
		Big Size of Motion	**Small Size of Motion**	**Smooth Regions**	**Texture Regions**	**Effects of Parameter Selection**
ALGORITHMS	ALG-1	Bad	Good	Satisfactory	Bad	Good
	ALG-2	Satisfactory	Good	Bad	Satisfactory	Good
	ALG-3	Satisfactory	Bad	Good	Satisfactory	Good
	ALG-4	Satisfactory	Bad	Satisfactory	Satisfactory	Bad
	ALG-5	Good	Satisfactory	Bad	Good	Satisfactory
	ALG-6	Good	Bad	Satisfactory	Good	Satisfactory
	ALG-7	Good	Good	Satisfactory	Satisfactory	Good
	ALG-8	Good	Satisfactory	Good	Good	Good

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1 Summary and Conclusions

In this thesis, ghost removal methods are studied by using input sets. The main methods used in image enhancement methods are introduced briefly, and detailed information about ghost removal is presented. Different approaches of ghost removal algorithms are implemented, and the performance of the algorithms is compared with different input sets.

The results are compared with twelve image sets. It is observed that for each input set algorithms provides different results because of the characteristics of the scene.

Firstly, background of the object in motion effects the performance of ghost removal. Sidibe et al. [7] fails to extract all the ghost regions because of the similarity of the intensity level between ghost regions and background. It is concluded that morphological operation minimizes the number of deficient ghost detections and false detections. Kautz et al. [11] provides better result with morphological operations than Sidibe et al. [7]. Secondly, number of the under-exposed and over-exposed pixels in the reference image effects the performance of ghost detection. It is seen that, false detection occurs if the reference image contains highly exposed pixel, such as Lee et al. [27], and Jiyang et al. [29]. When the comparison is done to find the ghost pixels, both of the algorithms give the false detection because of the methods used to normalize the image. Thirdly, finding non-ghostness probability for each pixel is giving false detection. Moon et al. [26] have false detected pixels, with increasing exposure difference between reference image and input image.

Based on the outputs of image sets, it is seen that global calibration of the input images, such as histogram matching [26], histogram equalization [27] provides better results in terms of finding the actual ghost regions, because, the methods are minimize the effect of exposure value difference. In addition to that, using segmentation [29] improves the ghost detection, such a way that, object in motion is segmented, and related comparison can be done. However, reference image selection is done properly, because it effects the performance of the comparison. Although, using different color domain than RGB [28] does not make impact about finding the ghost regions, adaptive thresholding is suitable with proper selection of the values, because each pixel in each image behaves differently when exposure value is increased.

Based on computational times, it is seen that segmentation of the image increases the computational time, in some input set, improvements using segmentation is not so much. In addition, global calibration of the images is faster than segmentation, as well as, they have satisfying results in terms of ghost labelling.

As a conclusion, it is difficult to select the best algorithm for all the image sets. However, with proper selection of the reference image, algorithms that provides global calibration to minimize the effect of the exposure value have better results than others, because after the calibration all the input images, ghost pixel extraction can be easily done.

5.2 Future Work

Due to the time and computational cost of the subjective test, the performance of all the algorithms cannot be evaluated. To give whole comparison of the ghost removal algorithms, advanced and comprehensive subjective test can be performed. The results of this test will give the detailed information about the performance of the ghost removal algorithms.

The most appropriate reference image for an algorithm can be determined by analyzing the characteristics of input scene. Therefore, performance of an algorithm is evaluated with respect to the reference image selection.

A detailed study in ghost removal in HDRI algorithms can be performed. The subjective tests can be repeated for these algorithms. Additionally, the comparison of ghost removal in HDRI algorithms and ghost removal in exposure fusion algorithms can be performed to show the advantages and disadvantages of each method.

REFERENCES

[1] E. K. C. -. I. S. Solutions, "CCD Image Sensor Noise Sources," 8 August 2001. [Online]. Available: https://www.uni-muenster.de/imperia/md/content/ziv/multimedia/downloads/kodak___noise_sources.pdf. [Accessed 8 August 2015].

[2] S. -J. Youm, C. Won-Ho and K.-S. Hong, "High Dynamic Range Video through Fusion of Exposure-Controlled Frames," in *Machine Vision Applications*, Tsukuba Science City, Japan, 2005.

[3] D. Eftaiha, "Light & Photography: Exposure and Tonal Range Considerations," Envato Pty Ltd., 2011 February 23. [Online]. Available: http://photography.tutsplus.com/articles/light-photography-exposure-and-tonal-range-considerations--photo-5685. [Accessed 7 August 2015].

[4] Toshiba Corporation, "Imagesensor Technology: HDR," [Online]. Available: http://toshiba.semicon-storage.com/ap-en/product/sensor/cmos-sensor/tech-hdr.html. [Accessed 8 August 2015].

[5] N. Gelfand, A. Adams, S. H. Park and K. Pulli, "Multi-exposure Imaging on Mobile Devices," in *18th ACM International Conference on Multimedia*, Firenze, Italy, 2010.

[6] O. T. Tursun, A. O. Akyüz, E. Aykut and E. Erdem, "The State of the Art in HDR Deghosting: A Survey and Evaluation," *EUROGRAPHICS,* vol. 34, no. 2, pp. 683-707, 2015.

[7] D. Sidibe, W. Puech and O. Strauss, "Ghost Detection and Removal in High Dynamic Range Images," in *Proceedings of International Conference on Multimedia and Expo-ICME,* 2009.

[8] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, Los Angeles, CA, USA, 1997.

[9] N. S. K. Mitsunaga T., "Radiometric self calibration," in *IEEE International Conference on Computer Vision and Pattern Recognition*, 1999.

[10] E. Reinhard, G. Ward, S. Pattanik and M. Debevec, High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting, Morgan Kaufmann, 2005.

[11] F. Pece and J. Kautz, "Bitmap Movement Detection: HDR for Dynamic Scenes," in *Proceedings of Visual Media Production (CVMP)*, 2010.

[12] A. A. Goshtasby, "Fusion of multi-exposure images," *Image and Vision Computing*, vol. 23, pp. 611-618, 2005.

[13] C. W.-K. Zhang W., "Gradient-directed composition of multiexposure images," in *IEEE International Conference on Computer Vision and Pattern Recognition*, 2010.

[14] C. S. Raman S., "Bilateral filter based compositing for variable exposure photographs," in *Eurographics Conference*, 2009.

[15] T. Mertens, J. Kautz and F. Van Reeth, "Exposure Fusion," *Pacific Graphics*, pp. 382-390, 2007.

[16] P. J. Burt and E. H. Adelson, "The Laplacian Pyramid as a Compact Image Code," *IEEE Transactions on Communications*, Vols. COM-31, pp. 532-540, 1983.

[17] T. Grosch, "Fast and Robust High Dynamic Range Image Generation with Camera and Object Movement," in *Vision, Modeling and Visualization Conference*, Aachen, Germany, 2006.

[18] K. Jacobs, G. Ward and C. Loscos, "Automatic HDRI Generation of Dynamic Environments," *IEEE Computer Graphics and Applications*, vol. 2, no. 28, pp. 84-93, 2008.

[19] W. W. M. Viola P., "Alignment by maximization of mutual information," in *Proceedings of the Fifth International Conference on Computer Vision*, 1995.

[20] O. Gallo, N. Gelfand, W.-C. Chen, M. Tico and K. Pulli, "Artifact-free High Dynamic Range Imaging," in *IEEE International Conference on Computational Photography (ICCP)*, San Fransisco, CA, USA, 2009.

[21] Y. S. Heo, K. M. Lee, S. U. Lee, Y. Moon and J. Cha, "Ghost-Free High Dynamic Range Imaging," in *10th Asian Conference on Computer Vision—ACCV*, Queenstown, New Zealand, 2010.

[22] L. Bogoni, "Extending dynamic range of monochrome and color images through fusion," in *15th International Conference on Pattern Recognition*, Barcelona, 2000.

[23] B. Lucas and K. Takeo, "An iterative image registration technique with an application to stereo vision," in *7th International Joint Conference on Artificial Intelligence*, Vancouver, Canada, 1981.

[24] E. A. Khan, A. O. Akyüz and E. Reinhard, "Ghost removal in high dynamic range images," in *IEEE International Conference on Image Processing*, Atalanta, GA, USA, 2006.

[25] M. Pedone and J. Heikkilä, "Constrain propagation for ghost removal in high dynamic range images," in *International Conference on Computer Vision Theory and Applications—VISAPP*, Funchal, 2008.

[26] Y.-S. Moon, Y.-M. Tai, J. H. Cha and S.-H. Lee, "A Simple Ghost-Free Exposure Fusion for Embedded HDR Imaging," in *2012 IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, 2012.

[27] D.-K. Lee, R.-H. Park and S. Chang, "Improved Histogram Based Ghost Removal in Exposure Fusion for High Dynamic Range Images," in *2011 IEEE 15th International Symposium on Consumer Electronics (ISCE)*, Singapore, 2011.

[28] C. Wang and C. Tu, "An Exposure Fusion Approach Without Ghost for Dynamic Scenes," in *Image and Signal Processing (CISP), 2013 6th International Congress on*, Hangzhou, 2013.

[29] S. Jiang, X. Zhihai, L. Qi, C. Yueting and F. Huajun, "Ghost detection and removal based on super-pixel grouping in exposure fusion," in *7th International Symposium on Advanced Optical Manufacturing and Testing Texhnologies: Optoelectronics Materials and Devices for Sensing and Imaging*, 2014.

[30] G. Ward, "Fast, Robust Image Registration for Compositing High-Dynamic Range Photographs from Handheld Exposures," *Journal of Graphics Tools*, vol. 8, no. 2, pp. 17-30, 2003.

[31] R. M. Haralick and L. G. Shapiro, Computer and Robot Vision, Boston, MA:: Addison-Wesley, 1992.

[32] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua and Süsstrunk, "SLIC Superpixels Compared to State-of-the-Art Superpixel Methods," *IEEE Transcations on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2281-2274, 2012.

[33] A. Srikantha and D. Sidibe, "Ghost Detection and Removal for High Dynamic Range Images: Recent Advances," *Signal Processing: Image Communication*, vol. 27, no. 1, pp. 650-662, 2012.

[34] Ö. Serkan, *Image Dynamic Range Enhancement (Master's Thesis, Middle East Technical University)*, Ankara, Turkey, 2011.

[35] D. Moughamian, "32-bit Imaging with HDR Toning in Photoshop CS5," Pearson, 28 February 2011. [Online]. Available:

http://www.adobepress.com/articles/article.asp?p=1682884&seqNum=2. [Accessed 9 August 2015].

[36] A. Jaehyun, S. H. Lee, J. G. Kuk and N. I. Cho, " A multi-exposure image fusion algorithm without ghost effect," in *Acoustics, Speech and Signal Processing (ICASSP)*, 2011.

APPENDIX A

IMAGE SETS

Figure A-1 Image Set 1 [1024x680] [1/500 1/250 1/125 1/60 1/30]

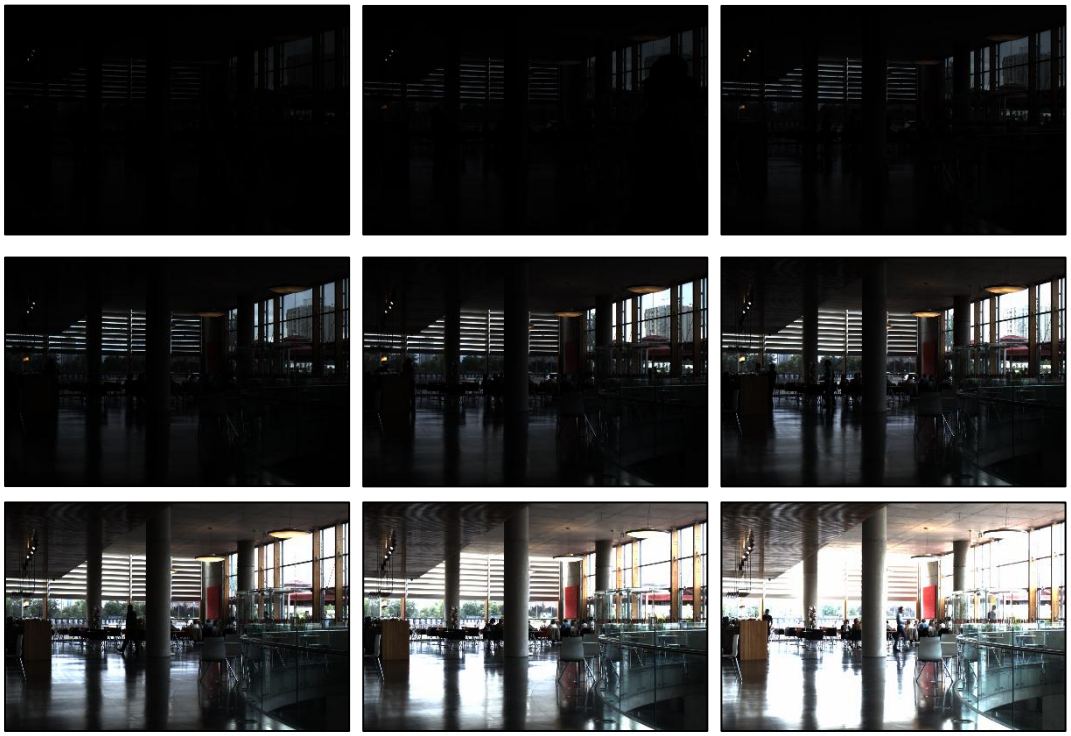


Figure A-2 Image Set 2 [824x1240] [1/4000 1/2000 1/1000 1/500 1/250 1/125 1/60 1/30 1/15]

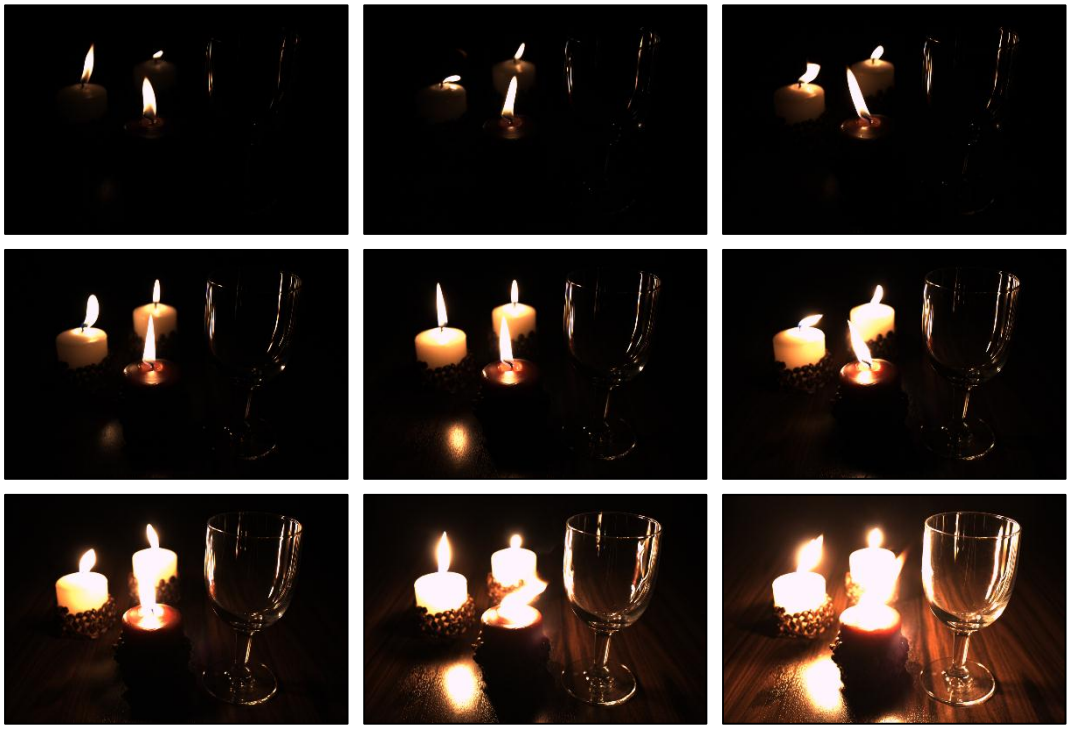


Figure A-3 Image Set 3 [872x1304] [1/2049 1/1025 1/512 1/256 1/128 1/64 1/32 1/16 1/8]

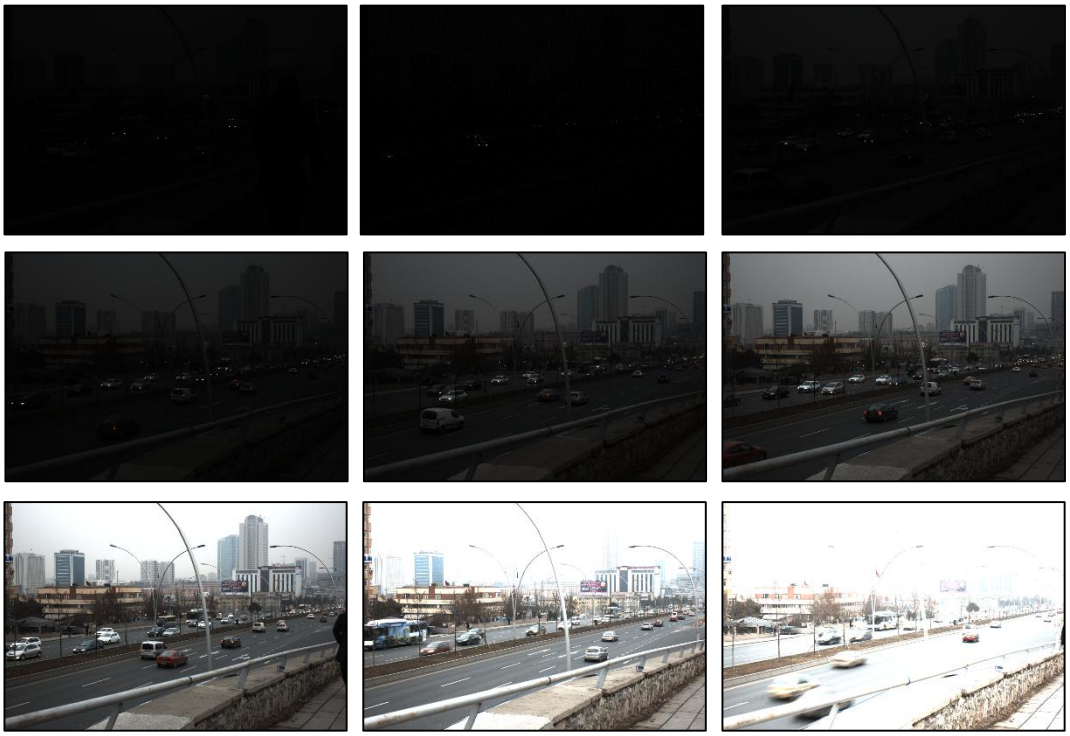


Figure A-4 Image Set 4 [872x1304] [1/4098 1/2049 1/1025 1/512 1/256 1/128 1/64 1/32 1/16]

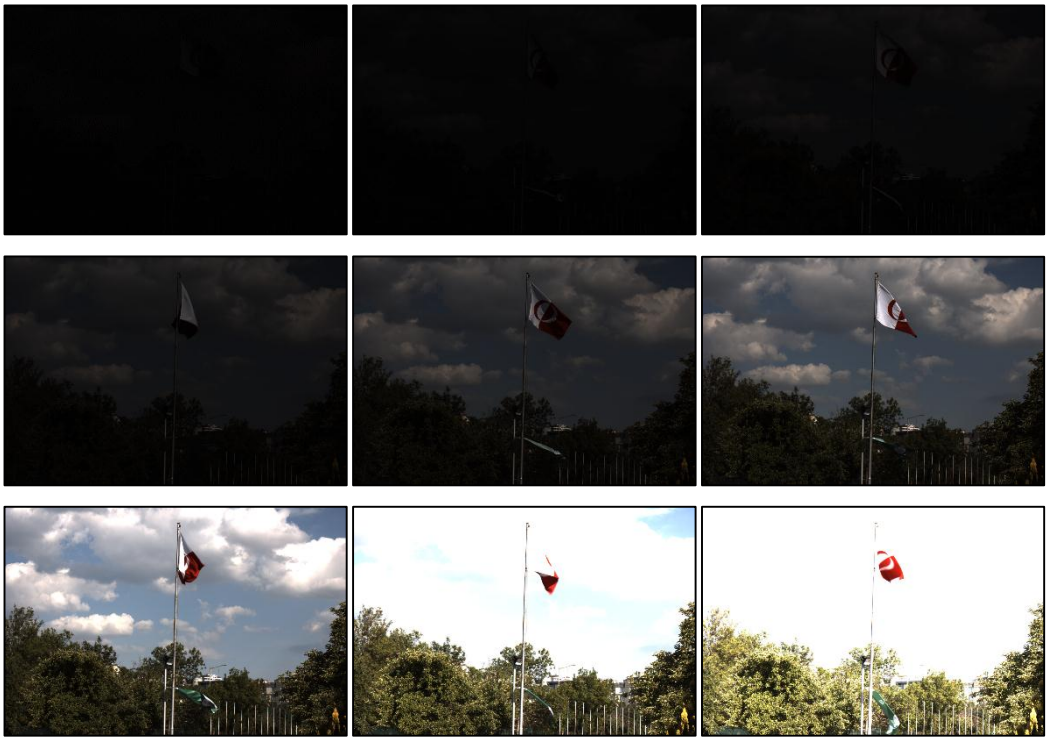


Figure A-5 Image Set 5 [872x1304] [1/4098 1/2049 1/1025 1/512 1/256 1/128 1/64 1/32 1/16]

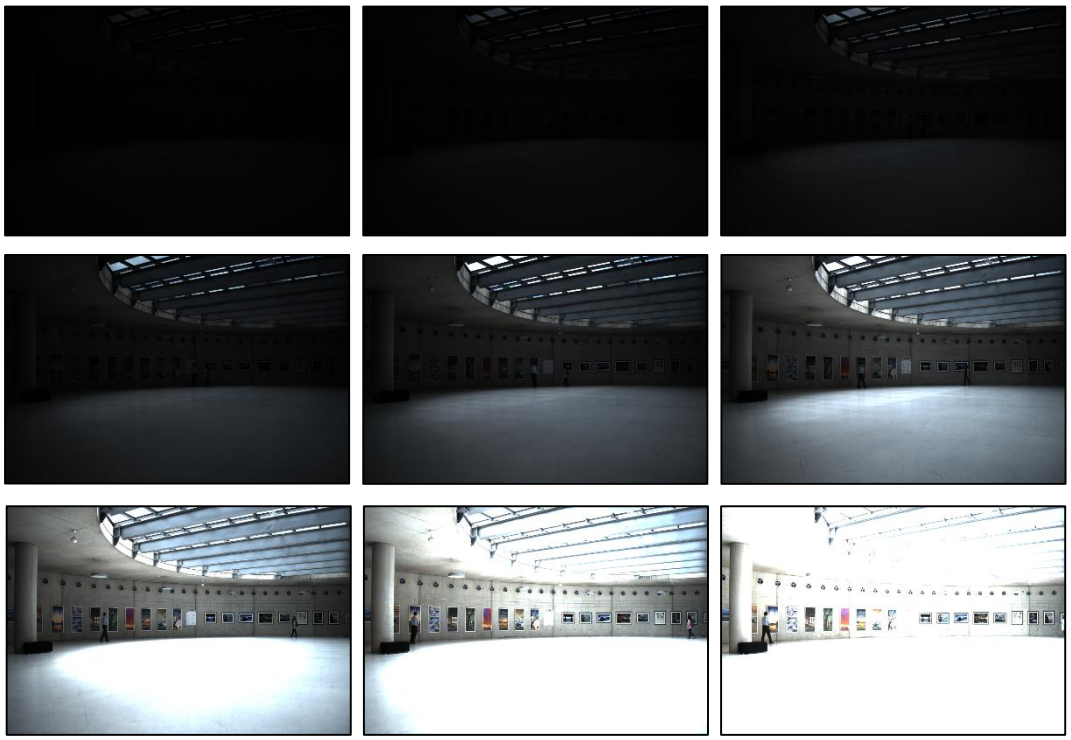


Figure A-6 Image Set 6 [824x1240] [1/4000 1/2000 1/1000 1/500 1/250 1/125 1/60 1/30 1/15]

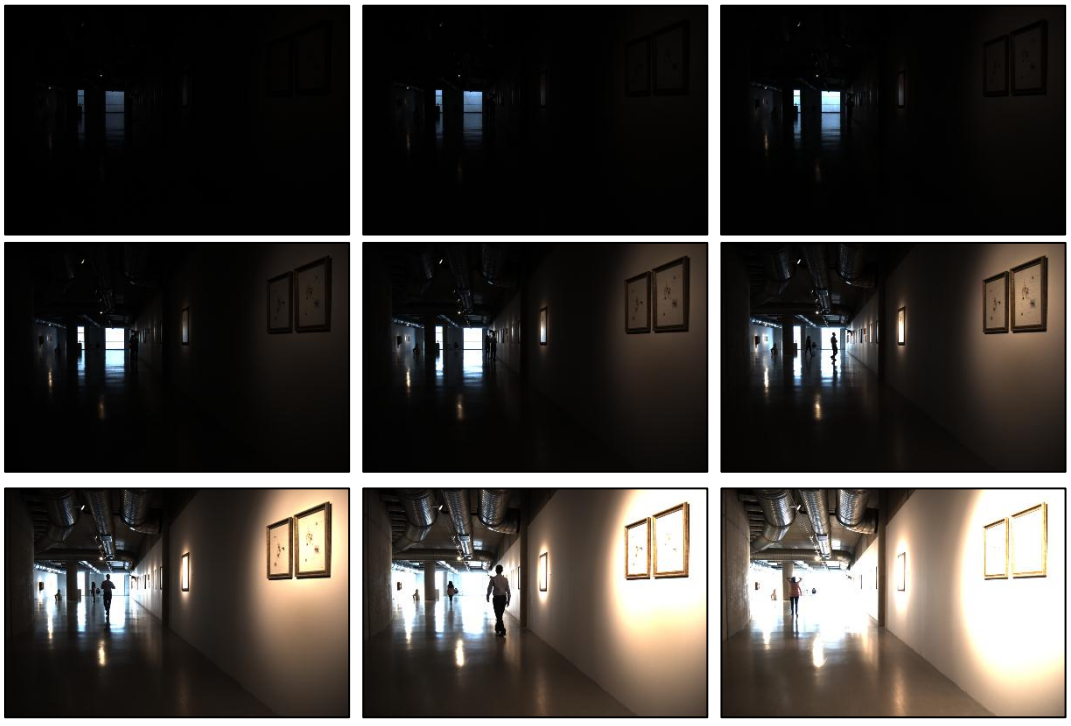


Figure A-7 Image Set 7 [824x1240] [1/4000 1/2000 1/1000 1/500 1/250 1/125 1/60 1/30 1/15]

Figure A-8 Image Set 8 [872x1304] [1/1580 1/790 1/395 1/197 1/99 1/49 1/25 1/12 1/6]

Figure A-9 Image Set 9 [304x448] [1/50 1/30 1/20 1/13 1/8 1/5 1/3]

Figure A-10 Image Set 10 [824x1240] [1/4000 1/2000 1/1000 1/500 1/250 1/125 1/60 1/30 1/15]

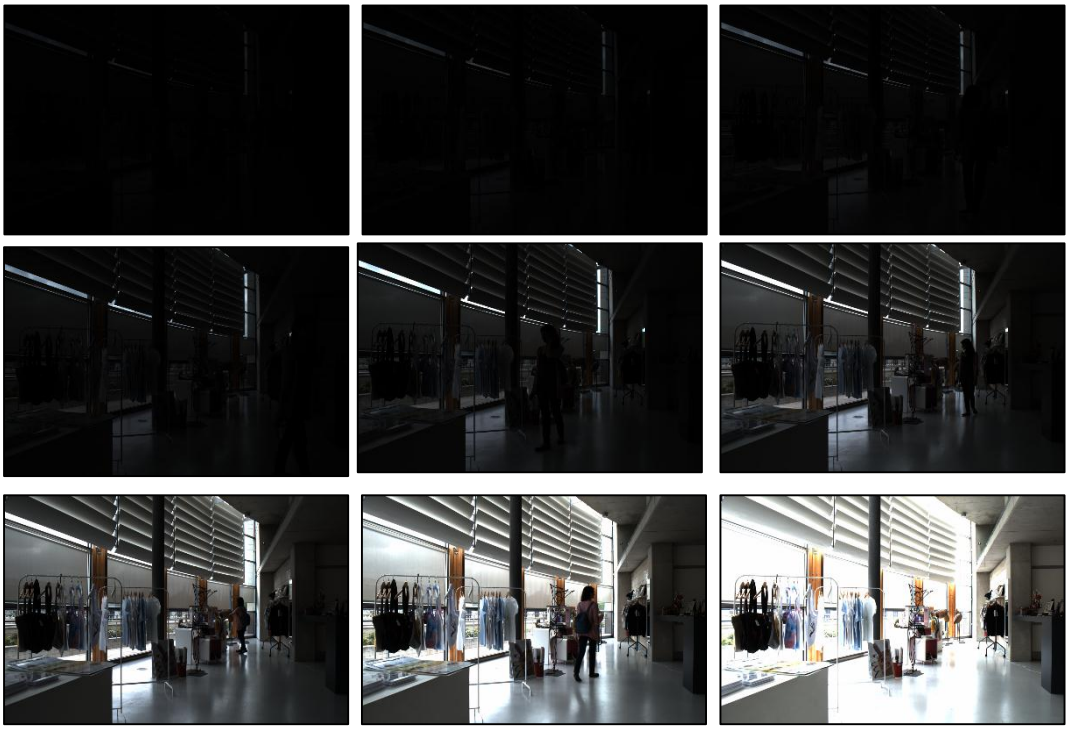


Figure A-11 Image Set 11 [824x1240] [1/4000 1/2000 1/1000 1/500 1/250 1/125 1/60 1/30 1/15]

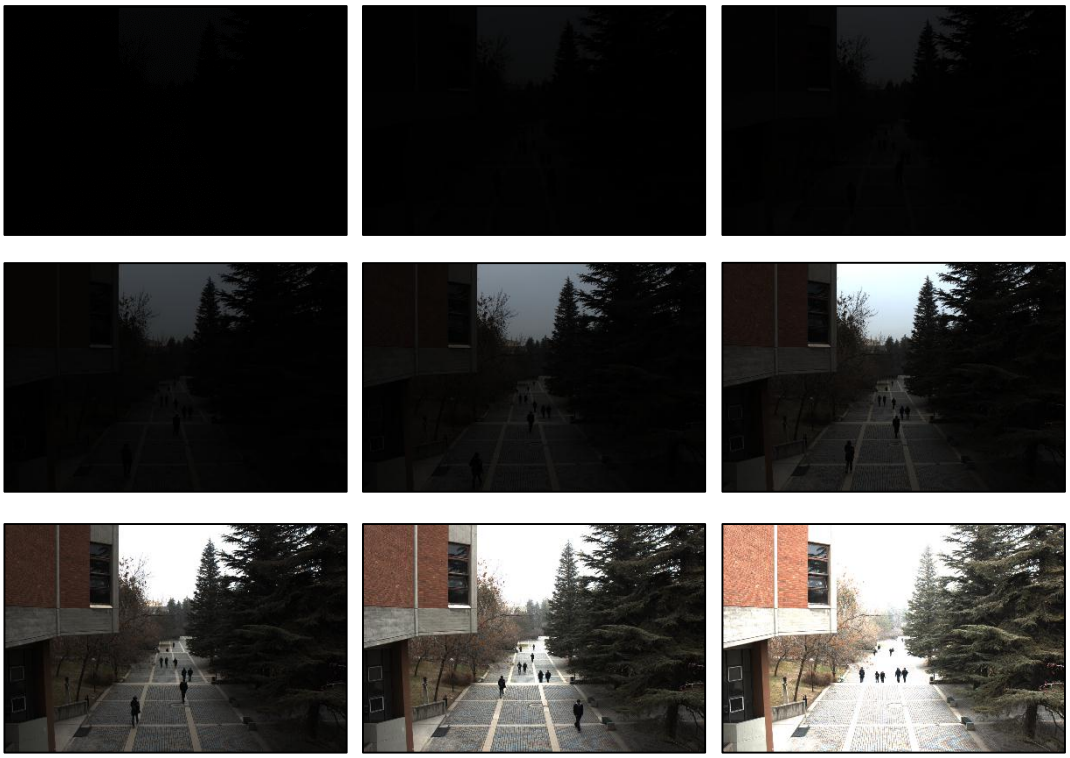


Figure A-12 Image Set 12 [872x1304] [1/4098 1/2049 1/1025 1/512 1/256 1/128 1/64 1/32 1/16]

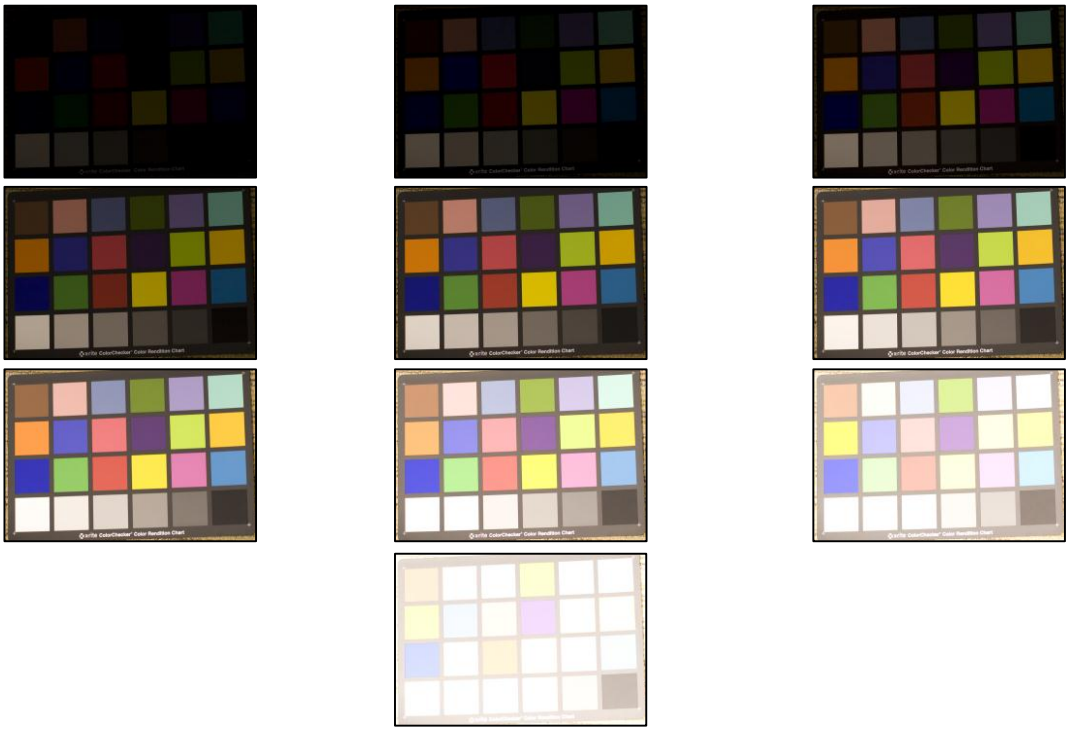
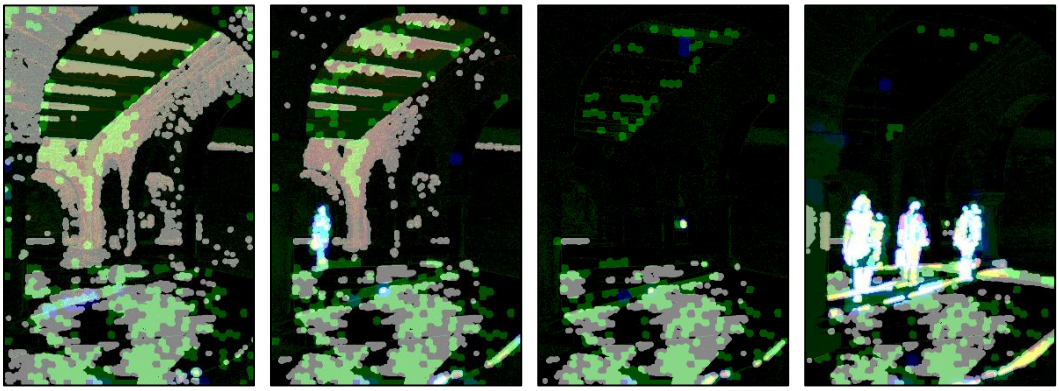


Figure A-13 Image Set 13 [698x1024] [1/250 1/125 1/60 1/30 1/15 1/8 1/5 1/2.5 1/1.25 1/0.625]

APPENDIX B

GHOST COLOR MAP OF EACH INPUT SET



(a) (b) (c) (d)

Figure B-1 Ghost Color Maps of ALG-8 for Image Set 1 when Reference Image is 5th Image, Ghost Color Maps of (a) 1st Image, (b) 2nd Image, (c) 3rd Image and (d) 4th Image

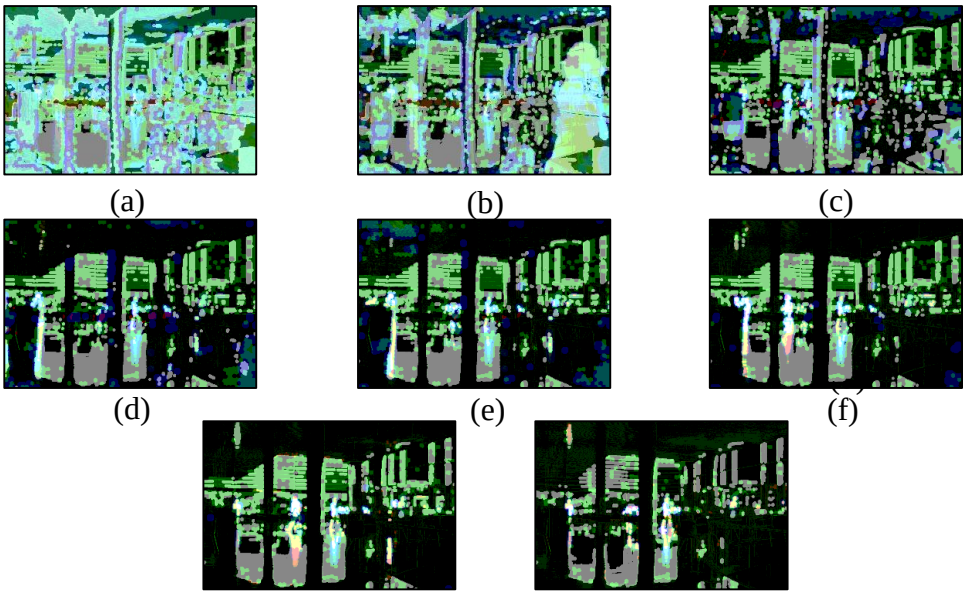


Figure B-2 Ghost Color Maps of ALG-8 for Image Set 2 when Reference Image is 9th Image, Ghost Color Maps of (a) 1st Image, (b) 2nd Image, (c) 3rd Image, (d) 4th Image, (e) 5th Image, (f) 6th Image, (g) 7th Image, (h) 8th Image

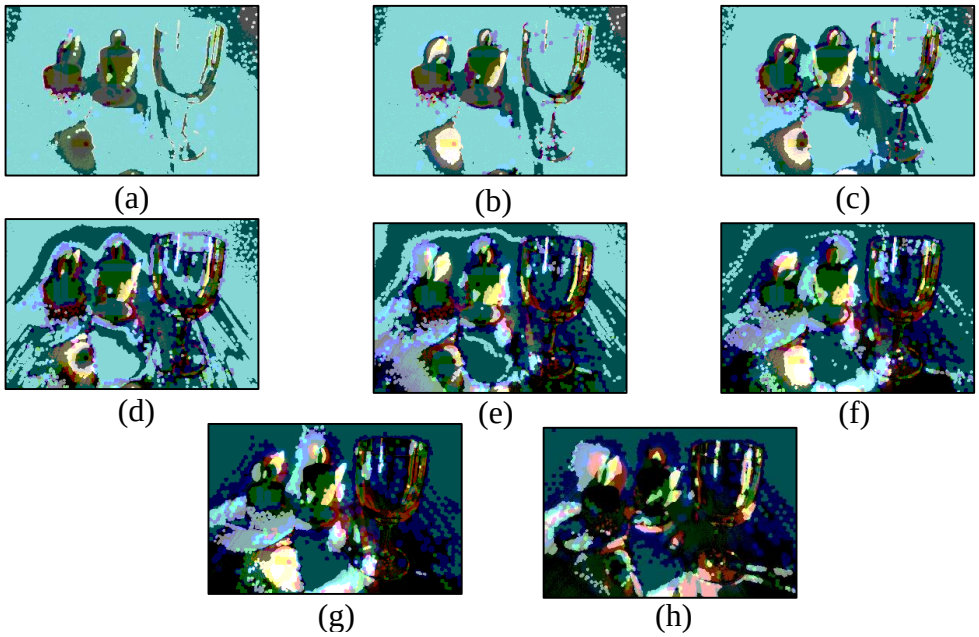


Figure B-3 Ghost Color Maps of ALG-8 for Image Set 3 when Reference Image is 9th Image, Ghost Color Maps of (a) 1st Image, (b) 2nd Image, (c) 3rd Image, (d) 4th Image, (e) 5th Image, (f) 6th Image, (g) 7th Image, (h) 8th Image

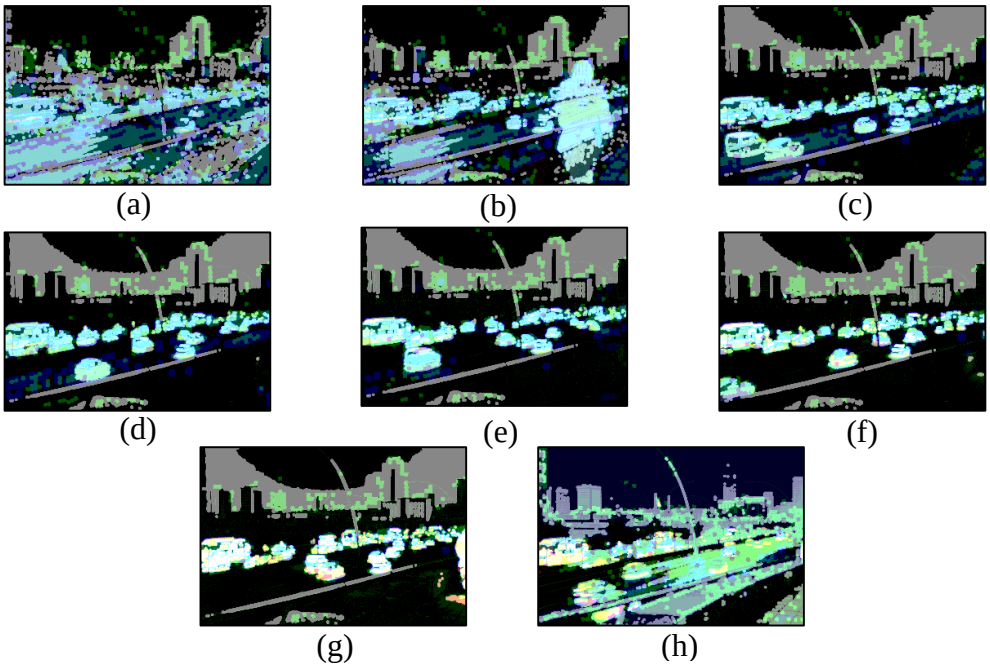


Figure B-4 Ghost Color Maps of ALG-8 for Image Set 4 when Reference Image is 8th Image, Ghost Color Maps of (a) 1st Image, (b) 2nd Image, (c) 3rd Image, (d) 4th Image, (e) 5th Image, (f) 6th Image, (g) 7th Image, (h) 9th Image

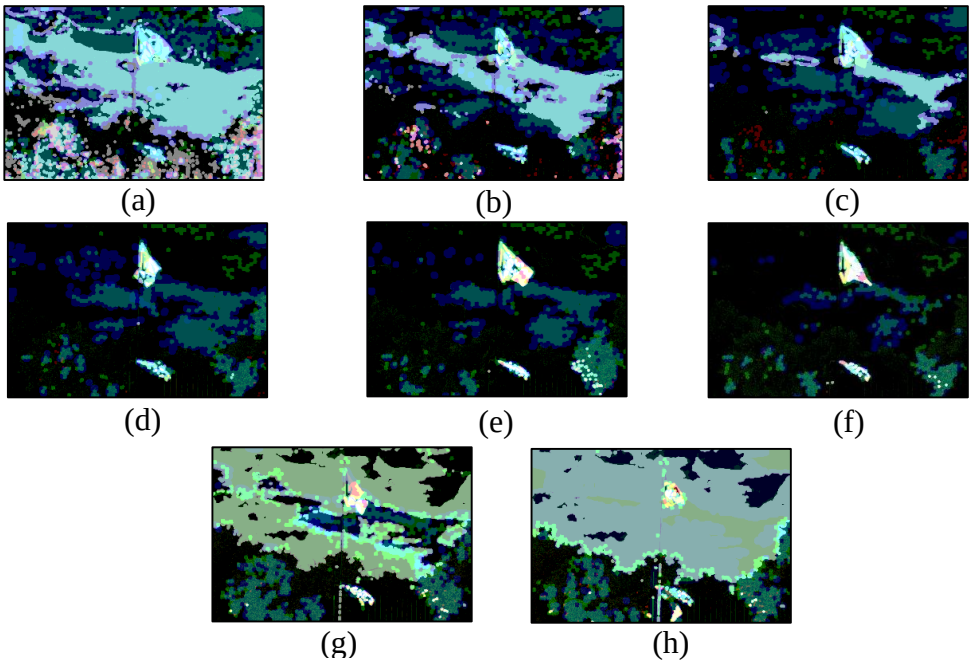


Figure B-5 Ghost Color Maps of ALG-8 for Image Set 5 when Reference Image is 7th Image, Ghost Color Maps of (a) 1st Image, (b) 2nd Image, (c) 3rd Image, (d) 4th Image, (e) 5th Image, (f) 6th Image, (g) 8th Image, (h) 9th Image

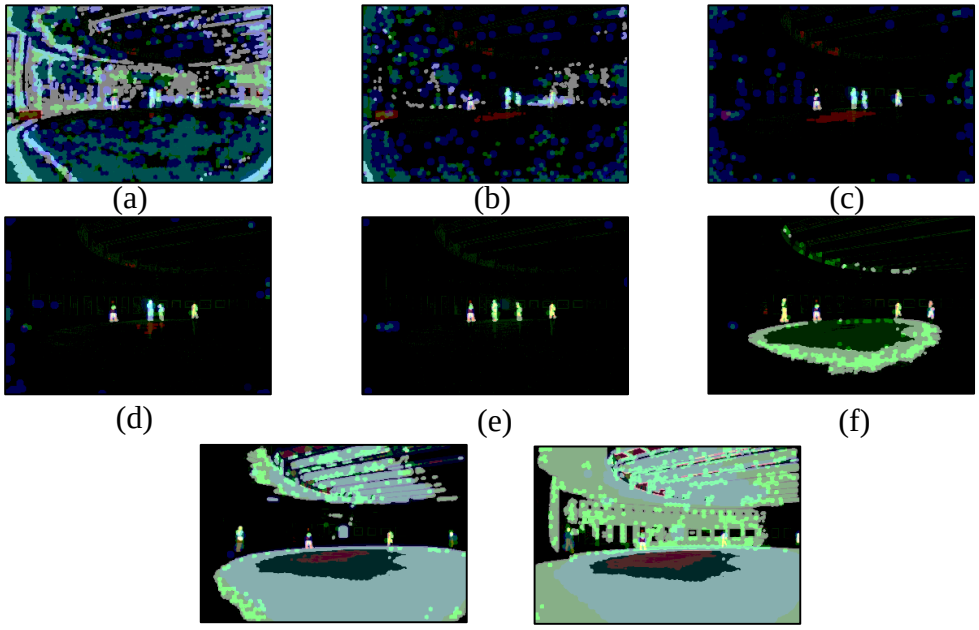


Figure B-6 Ghost Color Maps of ALG-8 for Image Set 6 when Reference Image is 7th Image, Ghost Color Maps of (a) 1st Image, (b) 2nd Image, (c) 3rd Image, (d) 4th Image, (e) 5th Image, (f) 6th Image, (g) 8th Image, (h) 9th Image

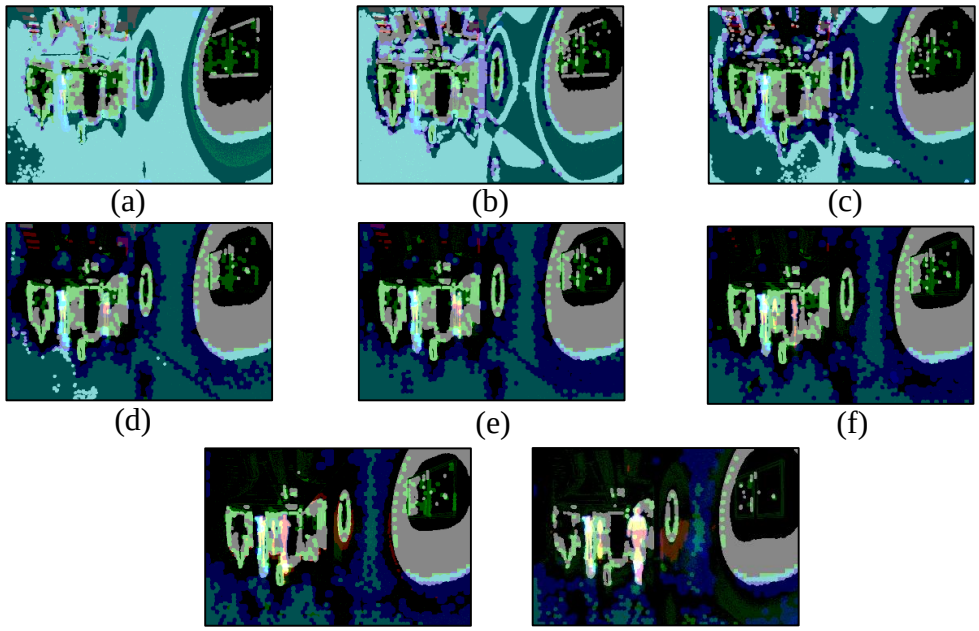


Figure B-7 Ghost Color Maps of ALG-8 for Image Set 7 when Reference Image is 9th Image, Ghost Color Maps of (a) 1st Image, (b) 2nd Image, (c) 3rd Image, (d) 4th Image, (e) 5th Image, (f) 6th Image, (g) 7th Image, (h) 8th Image

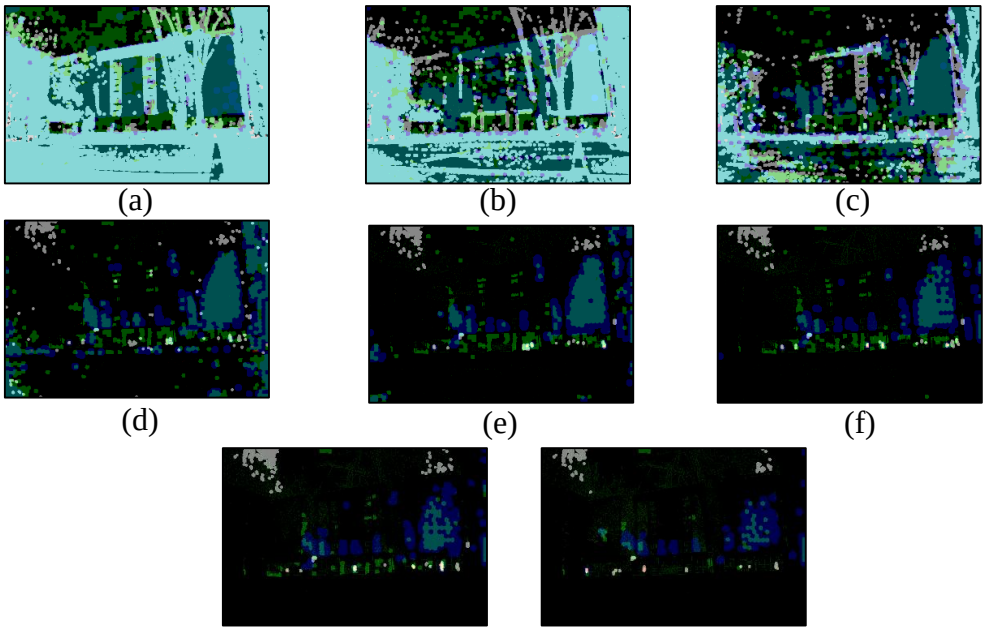


Figure B-8 Ghost Color Maps of ALG-8 for Image Set 8 when Reference Image is 9th Image, Ghost Color Maps of (a) 1st Image, (b) 2nd Image, (c) 3rd Image, (d) 4th Image, (e) 5th Image, (f) 6th Image, (g) 7th Image, (h) 8th Image

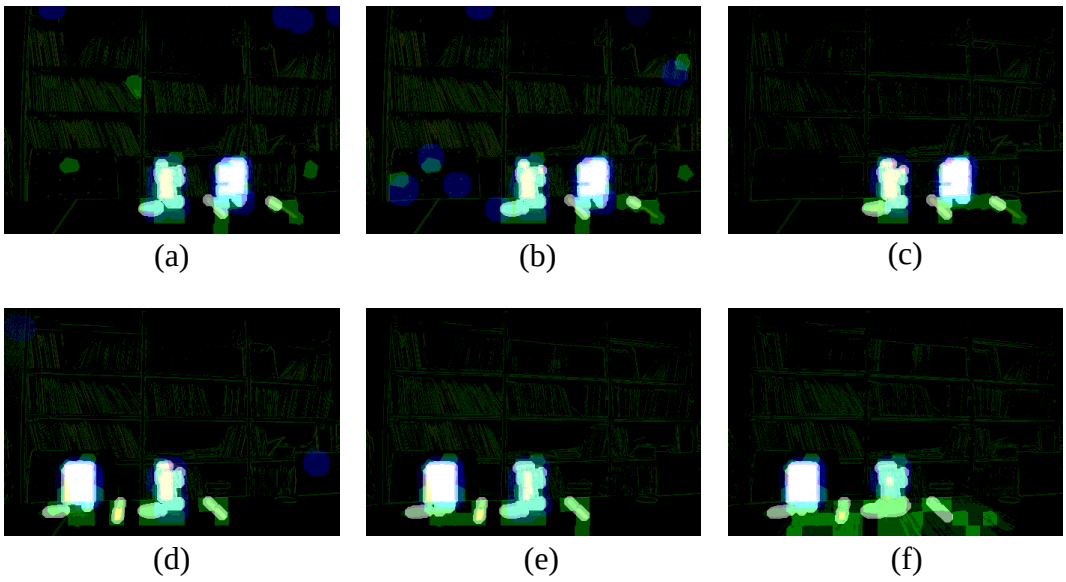


Figure B-9 Ghost Color Maps of ALG-8 for Image Set 9 when Reference Image is 4th Image, Ghost Color Maps of (a) 1st Image, (b) 2nd Image, (c) 3rd Image, (d) 5th Image, (e) 6th Image, (f) 7th Image

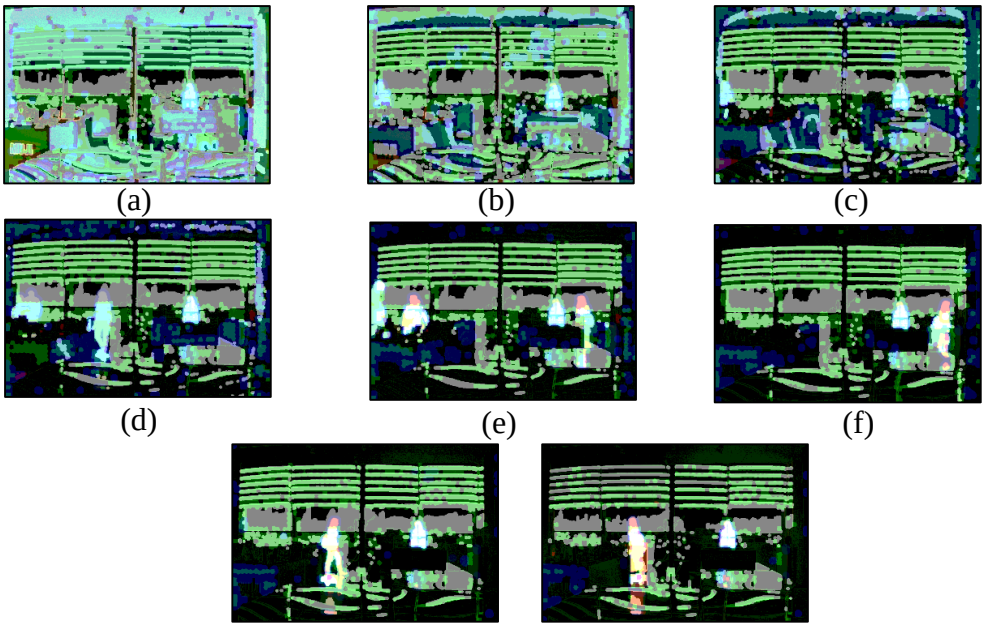


Figure B-10 Ghost Color Maps of ALG-8 for Image Set 10 when Reference Image is 9th Image, Ghost Color Maps of (a) 1st Image, (b) 2nd Image, (c) 3rd Image, (d) 4th Image, (e) 5th Image, (f) 6th Image, (g) 7th Image, (h) 8th Image

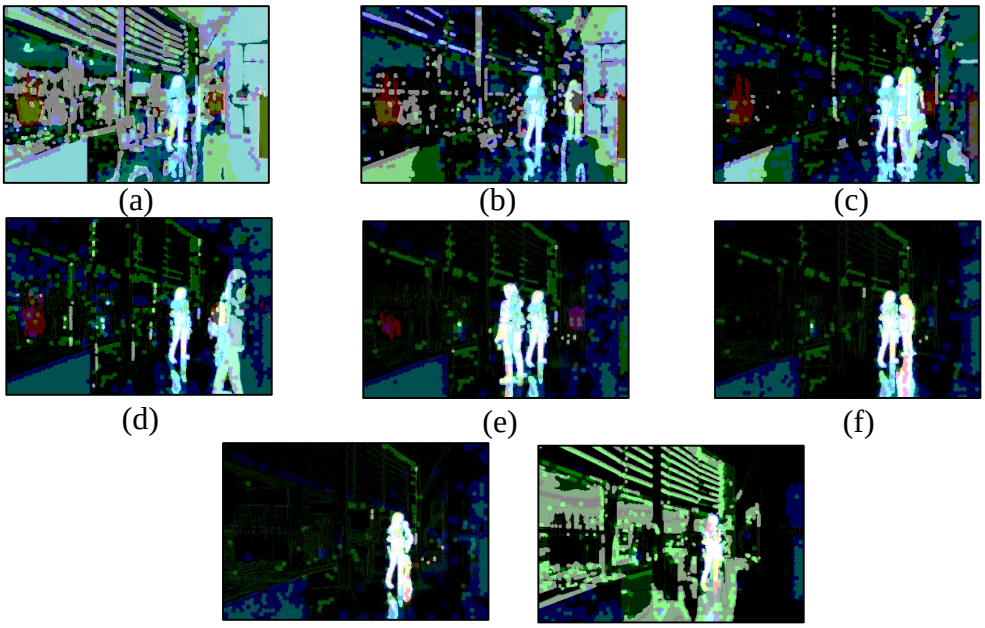


Figure B-11 Ghost Color Maps of ALG-8 for Image Set 11 when Reference Image is 8th Image, Ghost Color Maps of (a) 1st Image, (b) 2nd Image, (c) 3rd Image, (d) 4th Image, (e) 5th Image, (f) 6th Image, (g) 7th Image, (h) 9th Image

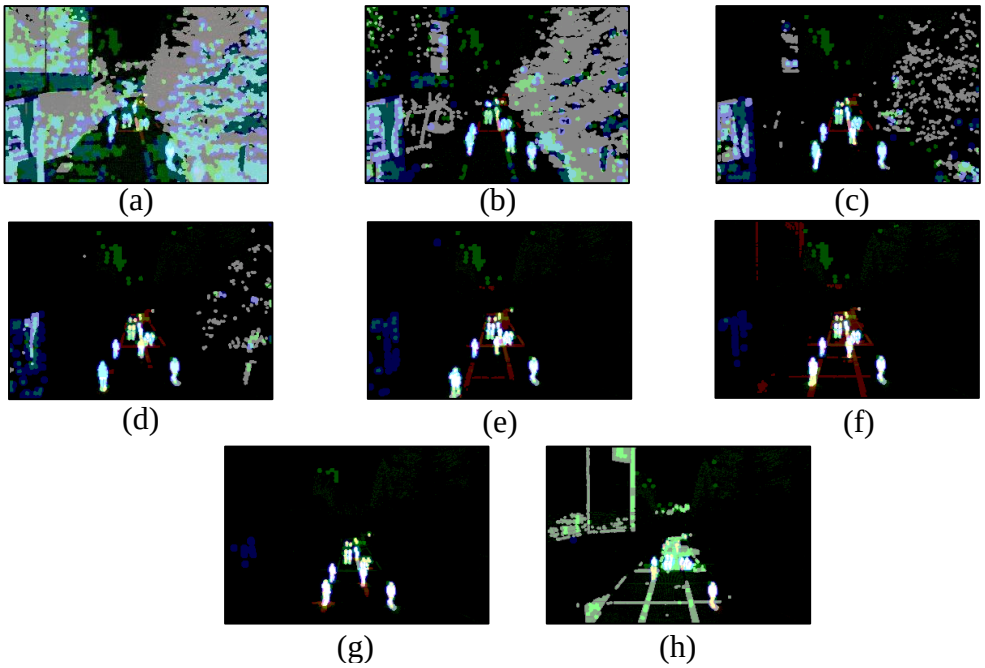


Figure B-12 Ghost Color Maps of ALG-8 for Image Set 12 when Reference Image is 8th Image, Ghost Color Maps of (a) 1st Image, (b) 2nd Image, (c) 3rd Image, (d) 4th Image, (e) 5th Image, (f) 6th Image, (g) 7th Image, (h) 9th Image

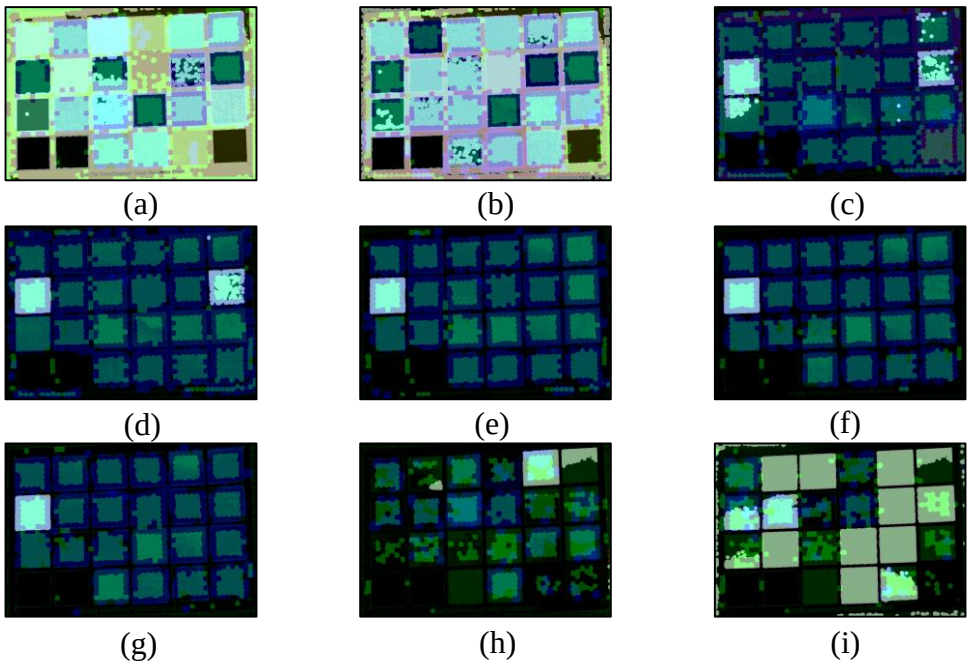


Figure B-13 Ghost Color Maps of ALG-8 for Image Set 13 when Reference Image is 8th Image, Ghost Color Maps of (a) 1st Image, (b) 2nd Image, (c) 3rd Image, (d) 4th Image, (e) 5th Image, (f) 6th Image, (g) 7th Image, (h) 9th Image, (i) 10th Image