Efficient Image Retargeting for High Dynamic Range Scenes

Submitted in partial fulfillment of the requirements
of the degree of

Master of Technology

by

Govind Salvi
(Roll No. PG201172008)

Supervisor(s):

Dr. Shanmuganathan Raman
Dr. Puneet Sharma

Center of Excellence in Information and Communication Technology
INDIAN INSTITUTE OF TECHNOLOGY JODHPUR
2013

This thesis is dedicated to my parents, who have always believed and supported me in all my endeavors, and to my entire family whom I love and cherish with all my heart.

Thesis Approval

The thesis entitled

Efficient Image Retargeting for High Dynamic Range Scenes

by

Govind Salvi
(Roll No. PG201172008)

is approved for the degree of

Master of Technology

_____	_____
Supervisor - 1	Supervisor - 2
_____	_____
Examiner - 1	Examiner - 2

Date: _____

Place: _____

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

_____

(Signature)

Govind Salvi
(Name of the student)

PG201172008
(Roll No.)

Date: _____

Abstract

MOST of the real world scenes have a very high dynamic range (HDR). The mobile phone cameras and the digital cameras available in market are limited in their capability in both the range and spatial resolutions. Same argument can be posed about the limited dynamic range display devices which further differ in spatial resolution and aspect ratio. In this thesis work, we address the problem of displaying the high contrast low dynamic range (LDR) image of a HDR scene in a display device which has different spatial resolution compared to that of the capturing device. We want to achieve this task while preserving the salient scene contents.

The optimal solution proposed in this work can be employed with any camera which has the ability to shoot multiple differently exposed images of a scene. Further, the proposed solutions provide the flexibility in the depiction of entire contrast of the HDR scene as a LDR image with user specified spatial resolution. This task is achieved through an optimized content aware retargeting framework which preserves the salient features apart from a novel algorithm to combine multi-exposure images. We show that the proposed approach performs exceedingly well in the generation of high contrast LDR image of varying spatial resolutions.

Contents

List of Tables

List of Figures

Nomenclature

HDR High Dynamic Range

LDR Low Dynamic Range

CRF Camera Response Function

HVS Human Visual System

AEB Auto Exposure Bracketing

Chapter 1

Introduction

REAL world scenes have a very high dynamic range (HDR). An example of such a HDR scene is the one which has both brightly and poorly lit regions. This implies that the range of brightness levels present in the scene is very high. The term ”scene” describe the natural or artificial environment which become the topic for an image. Human visual system (HVS) can visualize all the brightness levels of the scene through visual adaptation. Even analog cameras can capture a major percentage of the brightness levels. The digital capturing devices such as mobile phone cameras and digital cameras can not capture the entire HDR of a given scene. Similarly, the digital cameras are also limited in terms of their spatial resolution as witnessed by the limited spatial resolution in various digital imaging sensor architectures. In other words, digital cameras have limited range and spatial resolutions which are caused primarily due to the limitations posed by the imaging sensor design. It is highly challenging to capture all the brightness levels of a HDR scene as it require one to expose for longer duration.

A digital image can be realize as a two dimensional array (matrix), where the dimensions corresponds to vertical and horizontal size of the image. Each entry in this matrix is called as pixel value. This pixel value defined the color at the respective pixel location. A pixel color depends upon the amount of light (irradiance) coming towards its location. These irradiance values are different for different pixel locations. Dynamic range of an image is defined as the ratio of maximum irradiance value to minimum irradiance value present in the image. Therefore an image with high dynamic range is capable of displaying both poorly and brightly lit regions of the scene.

The images captured with low cost digital cameras available in the market do not have a high dynamic range. This limited dynamic range is caused mainly due to the limited well

capacity of a given sensor element. The dynamic range of the image can be enhanced by the HDR imaging techniques which rely on the capture of multi-exposure low dynamic range (LDR) images of the scene [1]. These approaches recover the camera response function (CRF) of the imaging system and employ it to linearize the intensity values. The HDR image of the scene is created by composting the linearized intensity values. The generated HDR image is tone mapped into a high contrast LDR image compatible with a given digital display device. Alternately, the high contrast LDR image of the scene can be directly generated without the knowledge of CRF ([2],[3]).

The spatial resolution of the image can either be reduced or enhanced by employing super-resolution algorithms [4]. These techniques perform resolution change through efficient interpolation without detecting the salient contents of the scene. Not all the regions in an image are equally important, generally a natural image contains objects of interest as well as some background details. While moving from one resolution to other we want that the important (salient) regions in the image should be minimally affected. Changing the spatial resolution of an image is known as image retargeting. The image retargeting approaches which have recently been developed enable one to change the spatial resolution of the image while preserving the contents of the scene which are important [5]. Retargeting has become the state-of-the-art approach when one wants to modify the spatial resolution of a given image.

Consider a set of multi-exposure images of a scene captured using traditional technique such as Auto Exposure Bracketing (AEB). The problem we would like to address is whether we can achieve a degree of flexibility in both the spatial and dynamic range resolutions. We assume that we are given a set of multi-exposure images corresponding to a static scene or a dynamic scene. One of the solutions to this problem is to first generate a HDR image using traditional algorithms and then perform spatial resizing either by super resolution or by image retargeting. The question to be addressed while employing such a solution is to judge whether this solution is the optimal one, or can we arrive at a better solution. This work is primarily focused on exploring the possibilities of alternate better solutions to this challenging problem.

1.1 Thesis Objectives

The main objective of the thesis is to search for an optimal solution to achieve a flexible range (contrast) and spatial resolution, given a set of muti-exposure LDR images of a static scene. we

have developed an algorithm which tends towards achieving an optimal solution to this problem. We show that the proposed approach performs far better than the traditional solution and leads to the generation of a high contrast LDR image with a provision to adapt the size of the image compatible with a given display device. The key contributions of the proposed novel approach are listed below.

1. Flexible content aware spatial retargeting of an image corresponding to a HDR scene,
2. Depiction of high contrast information of the scene within the user specified spatial resolution,
3. Achieving high quality LDR images without any visible artifacts even in the case of dynamic scenes, and
4. Assumption : No knowledge of exposure times, scene information, and CRF.

The main objective behind the thesis is to find an efficient way to merge multiple differently exposed images of a static scene as well as dynamic scene, into a high contrast LDR image with flexible spatial resolution. This task is achieved by an efficient algorithm which performs this task while reducing the loss in contrast and reducing any artifacts in the final LDR image.

1.2 Motivation

The recent work for spatial as well as dynamic range improvement from a set of multi-exposure images requires one to capture multi-exposure images with subpixel shifts [6]. This approach is a combination of the traditional HDR imaging and super-resolution approaches posed in a unified optimization framework. Therefore, this method does not enable one to perform content aware resizing though it helps in improving the dynamic range and the spatial resolution. Existing methods on simultaneous improvement of spatial resolution and dynamic range do not take into consideration, the content present in the scene ([7], [8], [6]).

The primary motivation behind this work is to generate a high contrast LDR image corresponding to a given HDR scene with flexible content aware image resizing capability. This application is quite useful in the present scenario as we have digital display devices which have different spatial resolutions and aspect ratios but can only display LDR content. Examples of

such display devices include Apple iPad, netbooks, smartphones, and tablets. The obvious solution to this problem as discussed earlier is to fuse the multi-exposure images and then to retarget the resultant image spatially in order to make it compatible with a given display device. This work is an attempt to probe for alternate efficient solutions for this problem and show how such solutions can indeed be better than the obvious solution in terms of image contrast and lesser artifacts incurred.

The main objective behind this work is to find an efficient way to merge multiple differently exposed images of a static scene or a dynamic scene into a high contrast LDR image with flexible spatial resolution. This task is achieved by an efficient algorithm while reducing the loss in contrast and reducing any ghosting artifacts in the final LDR image.

1.3 Thesis Organization

The thesis is organized as follows.

Chapter 2, presents the review of prior relevant work, which are key to the discussions later on. This chapter addresses the work done by the researches in order to eliminate the main problem associated with digital cameras (i) limitation in dynamic range and (ii) limitation in spatial resolution. However all the solutions defined so far are concerned only on either improving the dynamic range or get a flexibility in spacial resolutions, No single effort is made till now in order to get the flexibility in both dynamic range and spacial resolution.

Chapter 3, presents the proposed algorithm for simultaneous contrast and content-aware spatial retargeting in detail. This chapter shows that the solution achieved in this work are remarkable compare to the the most general solution of the problem. The solutions presented in this chapter are limited only for static scenes, however real world scenes are not static. For dynamic scenes we have to modify the approach.

Chapter 4, discusses the possible solution one can apply for getting a flexibility in range and spatial resolution for dynamic scenes. In this chapter we have extended our approach discussed in chapter 3 so that it can work for dynamic scenes as well.

Chapter 5, presents the results corresponding to various aspects of the proposed solution. Results are shown both visually as well as analytically for the proposed approach. Tools from different disciplines are used in order to quantify the results.

Chapter 6, discusses the possible add-ones one can add to the present work. This chapter

gives different new research concepts which can be added to improve the similar problems or to generate entirely new ideas, in order to eliminate the limitation in range and spatial resolution of the digital cameras.

Chapter 7, presents the future improvement or advancement in the present work. The detail conclusion of the problem discussed is also presented. The interdisciplinary aspect of the present work is discussed in brief.

Chapter 2

Previous Work

In recent times, creation of images which depict all the brightness levels in a natural scene has been a topic of great interest. Various research groups have been working on this topic and have proposed various solutions to this challenging problem [1]. Image resizing (super-resolution) is a different problem in which one attempts to change the spatial resolution of the given image. Super-resolution can be achieved by using multiple images of the same scene with sub-pixel shifts [4].

In this chapter, we will explain all the scientific efforts made over the years, by different research groups, in order to remove the limitation of dynamic range and spacial resolution from digital images.

2.1 Limitation in Dynamic range

A bracketed exposure sequence, which spans the entire dynamic range of the real world scene, comprises of a set of LDR images that are shot with a digital camera. The CRF(Camera response function) should be recovered in order to linearize the intensities. The HDR image can be generated by compositing these multi-exposure images in linearised intensity domain ([9], [10], [11]). The HDR images can be displayed in specialized HDR displays [12]. However, for visualizing the generated HDR image in common LDR displays we need to perform tone reproduction operation. Many different tone mapping operators have been proposed in recent years with various performance levels for different scenes ([13], [14]).

(a) Exposure Sequence

(b)

Figure 2.1: Top row: Input exposure sequence of LDR images, Bottom row: Exposure fused high contrast LDR image. Image courtesy: T. Mertens

2.1.1 Exposure Fusion

Contrary to this, exposure fusion approaches relieve us the need of intermediate HDR image generation and tone mapping operation ([2], [3]). Exposure fusion involves compositing the different Laplacian pyramid levels of the multi-exposure images with appropriate weights in order to reduce saturation and enhance contrast ([2],[15]). Exposure fusion choses the best region out of each image from the exposure stack, in the final high contrast LDR image. The key parameters it uses to decides which regions are best for output image are:

- Contrast
- Saturation
- Well Exposedness

Contrast is determined by applying laplacian filter on the gray scale images corresponding to each input image. Through contrast information it gives a higher weight to pixels which are part of edges and texture. Similarly we want saturated regions in the final image. Both underexposed and overexposed region do not contain any information. Parameter saturation takes cares of it. It takes standard deviation in R,G and B color channel, to calculate saturation amount. Well-Exposedness gives a higher weight to the pixel having values near to 0.5 (take normalize image

intensities 0-1). Weights are assign according to gaussion $exp(-\frac{(i-0.5)^2}{2\times 0.2^2})$. If $I_{ij,k}$ represents pixel value at (i, j) in image k. where $1 \leq k \leq n$, n is the total images in exposure stack.

$$W_{ij,k} = (C_{ij,k})^{\omega_C} \times (S_{ij,k})^{\omega_S} \times (E_{ij,k})^{\omega_E} \tag{2.1}$$

Here C, E and S are contrast, well-exposedness and saturation, ω is weighted assign to that particular parameter that is if ω_E is zero, well exposedness will not play any role in determining the final weight to a pixel at location (i, j).

The actual weight assign to a pixel value is normalize and calculated as

$$\overline{W_{ij,k}} = \sum_{t=1}^{n} (W_{ij,t})^{-1} W_{ij,k}$$

intensity at location (i, j) in final image R is given by equation (2.2).

$$R_{i,j} = \sum k = 1n\overline{W_{ij,k}} I_{ij,k} \tag{2.2}$$

Similar approach can be further used for merging flash/no-flash images to get the best information out of both the images and create a better image ([16], [17]). Dynamic scenes captured with the help of multi-exposure images lead to artifacts which requires appropriate deghosting prior to compositing [18]. Recently, researchers have turned their attention to reconstruct a HDR image of a non-static scene with the knowledge of CRF [19] and without the knowledge of CRF ([20], [21], [22]). The generation of a HDR image from a set of multi-exposure images when both the camera and scene change has been addressed in the recent works ([6], [23], [24]).

2.2 Limitation in Spatial Resolution

Content aware resizing should be done in a way that minimizes the amount of important information we lose during resizing operation. Approaches such as face detectors and visual saliency map detectors can be used to achieve this task ([25], [26]). After creating a visual saliency map, image can be cropped to capture the most salient regions in the image. These methods are based on the conventional technique of either cropping or removal of columns and rows.

These methods are often constrained by the limit to which a given image can be resized. Resizing an image beyond a critical factor generates a high degree of artifacts. Recently, methods have been proposed by which this critical factor can be improved [27]. Changing the spatial

resolution of the image is also important and used enormously in texture synthesis, where goal is to generate a large textured image from a small textured image [28]. However the solution obtain texture synthesis can not be extended to natural scenes directly as they follow complex statistics and geometry. A natural image may have multiple different regions of importance and sometimes user interaction is exploited to specify the regions which are of greater importance [29].

2.2.1 Optimal Seam Carving

Image retargeting is a much better automatic approach which has been widely used for content aware resizing [5]. The first popular implementation of image retargeting, Seam Carving, involves the identification of minimum energy seams which have to be removed (image downsizing) or added (image enlarging) so that there is minimum loss of information or in other words the retargeting operation should effect the object of interest minimally. An efficient energy metric based on gradient measure serves as the energy function. Optimal seam carving can alternately use different types of energy functions such as gradient magnitude, entropy, visual saliency, eye-gaze movement, and more. The removal or insertion of seams can be done in such a way as to make it compatible with the resolution and aspect ratio of the display device. Seam carving can be extended to perform video retargeting ([30], [31]). An overview of the different types of image retargeting approaches can be found in the recent tutorial [32].

Figure 2.2: Left: Vertical and horizontal minimum energy seam found by seam carving algorithm, Top-right: Image enlarging by minimum energy column insertion, Bottom-right: enlarging through seam carving algorithm. Image Courtesy: Shai Avidan

Figure 2.2 Left: shows the one vertical and one horizontal minimum energy seam found in the test image. Figure also shows the enlargement of the image by insertion of minimum

energy column and the other with insertion of minimum energy seam. One can easily see that enlargement with seam carving give better results, (preserve salient features better).

There is always a trade-off between the spatial resolution and the range resolution of an imaging sensor. A typical example is the assorted pixels which use multiple sensor elements with different sensitivities to create a HDR image [33]. Here, we sacrifice some spatial resolution to gain more dynamic range. The size of the sensor element can not be made smaller than a particular size due to noise and limited well capacity ([34], [35], [36]). These studies on the imaging sensor emphasize the need for creating a new application with flexible spatial and range resolutions.

We shall present the basic algorithm behind the proposed solution for static scenes in the next Chapter.

Chapter 3

Proposed Approach

THIS Chapter presents various approaches for efficient retargeting of a HDR scene. The proposed algorithm uses a set of LDR images having different exposure times. Input images are registered LDR images of the same static scene. Consider the problem of efficient retargeting of a HDR scene. Let us assume that a set of LDR images having different exposure times are given. We also assume that the LDR images are perfectly aligned while the scene can either be static or dynamic. Let $I_1, I_2, I_3, \ldots, I_n$ be the set of n given LDR images each of dimension $M \times N$. We shall use magnitude of gradient corresponding to a given image as the energy metric. Equation (3.1) represents the magnitude of gradient at a given pixel location (x, y).

$$\chi(x, y) = \left|\frac{\partial I(x, y)}{\partial x}\right| + \left|\frac{\partial I(x, y)}{\partial y}\right| \tag{3.1}$$

This energy matrix can be employed to generate a cumulative matrix, which would enable us to find the minimum energy seams (seams with least importance) in the LDR images. Consider a cumulative matrix C, in case of vertical seams, first row of it will be same as energy matrix, for the rest it is computed as, $C(x, y) = \chi(x, y) + \min\{C(x-1, y-1), C(x-1, y), C(x-1, y+1)\}$ for a pixel location (x, y). In order to get the minimum energy seam we find the lowest cumulative value in the last row of generated matrix C. With this minimum value we backtrack the procedure. Figure 3.1(a) shows an example of this energy matrix and the corresponding cumulative matrix is shown in Figure 3.1(b), Figure 3.1(b) also shows a minimum energy vertical seam. Consider a natural image in Figure 3.1(c) which shows a minimum energy vertical seam. Similar minimum energy seams can be calculated in the horizontal direc-

18	21	5	8
23	30	6	0
16	12	9	24
15	23	15	14
9	17	15	8

(a)

18	21	5	8
41	35	11	5
51	23	14	29
38	37	29	28
46	46	43	36

(b)

(c)

Figure 3.1: Finding Minimum Energy Seams: (a) Energy matrix given in equation (3.1), (b) Cumulative matrix corresponding to the energy function and the calculated minimum energy vertical seam, (c) Minimum Energy seam calculated by the algorithm in a natural image. Image courtesy: Jim Krause

tion as well. We shall start the discussion with a general approach for retargeting LDR images corresponding to a HDR scene. We shall assume that we do not know the exposure times and the CRF in our work. Let us formulate an approach for retargeting a static HDR scene and then extend the approach to a dynamic HDR scene.

3.1 Direct Approach

The general approach for resizing an image corresponding to a HDR scene is to take multiple LDR images of the scene with different exposure times and subjecting them to exposure fusion [2]. This approach results in an image having much higher contrast than the individual LDR images. We apply optimal seam carving on this high contrast image to generate the resized high contrast LDR image of the scene [5].

Performing exposure fusion before resizing deprives us of the information presents in the original LDR images while resizing. Hence, the contrast of the final high contrast LDR image depends mainly on the exposure fusion algorithm. We shall now explore alternate approaches to arrive at a better solution to this problem. Suppose we have multiple images of the same scene with different exposure times, we can obtain multiple seams with least energy measures corresponding to each of the LDR images. We shall consider how removing or adding seams with minimum energy before applying exposure fusion may result in a better high contrast LDR image of the scene.

Figure 3.2: Data Set: Beach, Top Row: Bracketed LDR exposure sequence, Middle Row: Energy function corresponding to each LDR image, Last Row: Vertical seams detected by the proposed algorithm. Notice that on each image the minimum energy seams found by the algorithm are different.

3.2 Statistical Approach

The major drawback with the direct approach is that it increases the contrast first (by exposure fusion) and then on the resultant image retargeting is performed. However by doing this operation, we are including least important seams in the exposure fusion process. On the other hand these least important seam may be removed during the retargeting process. Alternatively for a given energy matrix, we can find the cumulative energy matrix for each individual LDR image. With the help of this cumulative energy matrix, we can find seams with minimum energy in each of the LDR images. Notice that the seams found by the algorithm need not be the same on each image (see bottom row of Figure 3.2).

Suppose we are given a set of n LDR images corresponding to a static HDR scene. For a given LDR image I_r, c_r denotes the computed minimum energy seam and e_r denote the corresponding cumulative energy. We obtain the minimum energy seams, and corresponding cumulative energy measures for each image. We shall pose the problem to be a decision problem. We want to make a decision as to which seam has to be chosen for retargeting among these minimum energy seams. One option is to pick the seam having the least energy among these minimum energy seams.

$$e_k = \min\{e_r : r = 1, 2, \ldots, n\}$$

In this case, seam c_k will be chosen as the minimum energy seam and will be considered

for retargeting operation. Instead one can also choose the seam having cumulative energy value which is the median of all the cumulative energy values corresponding to the minimum energy seams found in different images. In either case, the accuracy of the result solely depends upon the scene statistics and geometry.

It may be noted that for $r \neq k$, seam c_k might not represent the seam with minimum energy in the image I_r. Thus by deleting or adding the seam c_k, we can not guarantee to add or delete the seam with minimum energy in the image I_r. As the input images are perfectly aligned in order to maintain the consistency in coordinates, the same minimum energy seam has to be added or deleted in each LDR image.

We can further improve this method by making sure that the minimum total energy seam should be added or removed from each LDR image every iterations. Consider two images with minimum energy seams located at different coordinates. We can sum up the cumulative seam energy of the minimum energy seam with the cumulative seam energy of the seam present in the same coordinates in the other image. In this process, we would get two measures corresponding to the minimum energy seams in each image. We shall now choose the minimum out of these two measures and remove the seams which lead to that measure from each of the images.

We shall now generalize this process for n LDR images. Let s_{ir} be the seam having the same coordinates in image i of the minimum energy seam in image r. If the seam c_r is deleted from each of the LDR images, the total energy added or removed is given by the equation (3.2).

$$\Phi_r = \sum_{i=1}^{n} \phi(s_{ir}), 1 \leq r \leq n \tag{3.2}$$

where, $\phi(s_{ir})$ denotes the energy of the seam s_{ir} and Φ_r denotes the cumulative energy if the seam c_r is chosen for retargeting. Let Φ_k be the minimum cumulative energy.

$$\Phi_k = \min\{\Phi_r : 1 \leq r \leq n\}$$

In this case, the total amount of energy removed or added will be Φ_k and desired seam to be deleted or added will be c_k. Using this method, we can get better results compared to the previous approaches.

However while adopting the statistical approach discussed, the seam having the least energy need not be the one among the candidate low energy seams. This does not guarantee the removal or the addition of the desired least energy seam. This is due to the fact that while calculating the total minimum energy we are only concerned about the energy of the candidate low

energy seams in each LDR image. Other possible seams which could have lead to a much better solution to the problem have been discarded. This shows that the statistical approach does not lead to the optimal solution.

3.3 Aggregate Energy Matrix Approach

Instead of finding energy matrices for individual LDR images separately, we can think of a modified energy matrix which is a function of all the energy matrices corresponding to LDR images. We shall call this modified energy matrix as an aggregate energy matrix, as it indicates the aggregate energy of a pixel across all the images. Through this aggregate energy matrix a cumulative matrix is generated. Any seam indicated as minimum energy seam by this cumulative matrix will be of least importance. This criterion is necessary because it guarantees that we will not lose important information during the retargeting process.

For example, if we are taking the magnitude of gradient as our energy matrix (for individual LDR images) then the aggregate energy matrix will be a function of the gradients of individual images. Consider a function h which takes the energy matrices of individual LDR images and creates an aggregate energy matrix as defined by equation (3.3).

$$\chi = h(\chi_1, \chi_2, \chi_3, \ldots, \chi_n) \tag{3.3}$$

where $\chi_1, \chi_2, \chi_3, \ldots, \chi_n$ are the energy matrices for the images $I_1, I_2, I_3, \ldots, I_n$ respectively. Let us first assume this function to be a linear combination of the gradient magnitude of individual LDR images.

$$\chi = \sum_{r=1}^{n} \alpha_r \chi_r, 0 \leq \alpha_r \leq 1 \tag{3.4}$$

where $\sum_{r=1}^{n} \alpha_r = 1$, parameter α_r corresponds to the weight given to the energy matrix of image I_r. Through this aggregate energy matrix χ, our algorithm generates a cumulative energy matrix which enables us to calculate the minimum energy seams. The weight parameter α_r should be chosen in such a way that the regions which are underexposed or overexposed in the LDR images should get lesser weight compared to the other regions. The average energy of a pixel in a given image could be used as a weighting parameter. In this case α_r would be the

average energy per pixel in image I_r and is defined as $\alpha_r = \sum_{x=1}^{M}\sum_{y=1}^{N}\chi(x,y)$.

This weighting parameter has an important role in making the decision regarding which seam needs to be added or deleted in image retargeting. We further try to calculate this weighting parameter using some other image characteristics. Laplacian of an image calculates the second order derivative along both the spatial directions (horizontal as well as vertical). Laplacian operator can also serve as an edge detector kernel. In order to make sure that the final resized high contrast LDR image contains the edges information, we can give weight to the energy matrices of individual LDR images through the Laplacian operator. However as the sharpness of the edges are different for the LDR images, we need to modify this Laplacian before using it as a weighting parameter.

We calculate the weighted Laplacian for each LDR input image and then perform an element wise multiplication of this matrix with the energy matrix of each image and then calculate the sum. This matrix will now work as the desired aggregate energy matrix.

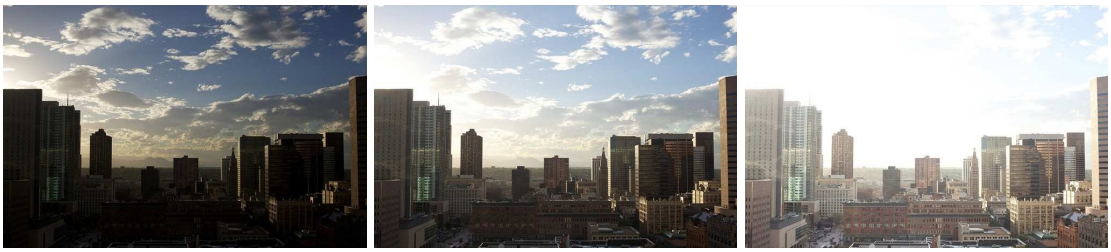
$$\chi(x,y) = \sum_{i=1}^{n} L_i^*(x,y)\chi_i(x,y) \tag{3.5}$$

$$\text{where,}\quad L_i^*(x,y) = \frac{L_i(x,y)}{\sum_{i=1}^{n} L_i(x,y)}$$

Here both multiplication and division are performed element wise.

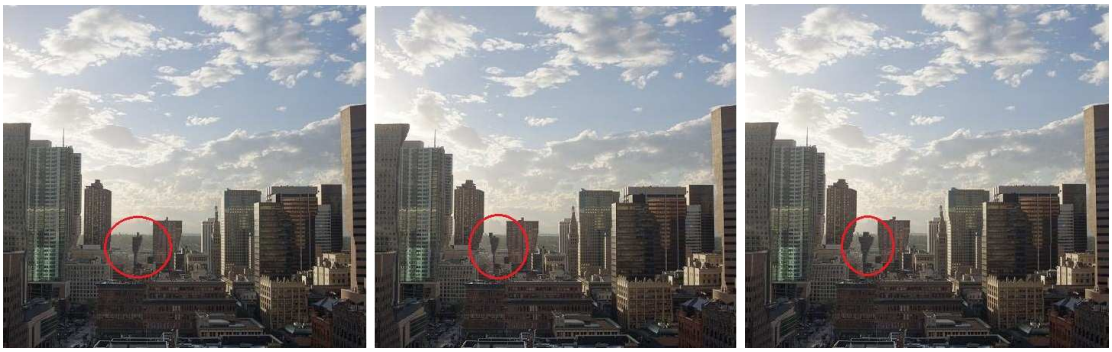
With this approach (aggregate energy matrix with weighted Laplacian as a weighting parameter), the final image is not only loses (or adds in the case of enlarging images) minimum energy but also the output resized high contrast LDR image will be of better quality than the direct approach as well as statistical approach.

Figure 3.3 shows the results we got using various approaches discussed. The detail discussion about the results will be give in Chapter 5.



(a)

(b) (c)



(d) (e) (f)

Figure 3.3: Data Set: Building, Results obtained when reducing the aspect ratio of the image (30%). Top row: (a) Exposure sequence, Middle row: (b) Exposure fused image and (c) its energy distribution, Bottom row: Results with (d) Direct Approach, (e) Statistical Approach. (f) Aggregate Energy Matrix Approach (weighted using energy per pixel). Marked region shows how different approaches affect the salient feature in that region. Image courtesy: Jim Krause

Chapter 4

Towards the Dynamic Scenes

PREVIOUS Chapter presents various approaches for content aware resizing of high dynamic range scenes. A major drawback of these approaches is that they only works for static image sets. Where the input image set is differ only in exposure time (camera shutter speed). But in real situations capturing a static scene with sufficient number of different exposure time settings is very difficult. The image set in which images are differ not only in exposure time, but also have some object movement are called dynamic scenes. Applying exposure fusion on a image set having dynamic scene, will results into ghosting effect in the resultant high contrast LDR image. Figure 4.1 shows the exposure sequence for a dynamic scene and Figure 4.1(b) shows the resultant image with ghosting artifacts after applying exposure fusion.

Generation of a high contrast LDR image for a dynamic LDR image set is a challenging problem in computational photography. Over the years various techniques have been proposed by the researchers to solve this problem efficiently ([19], [21]).

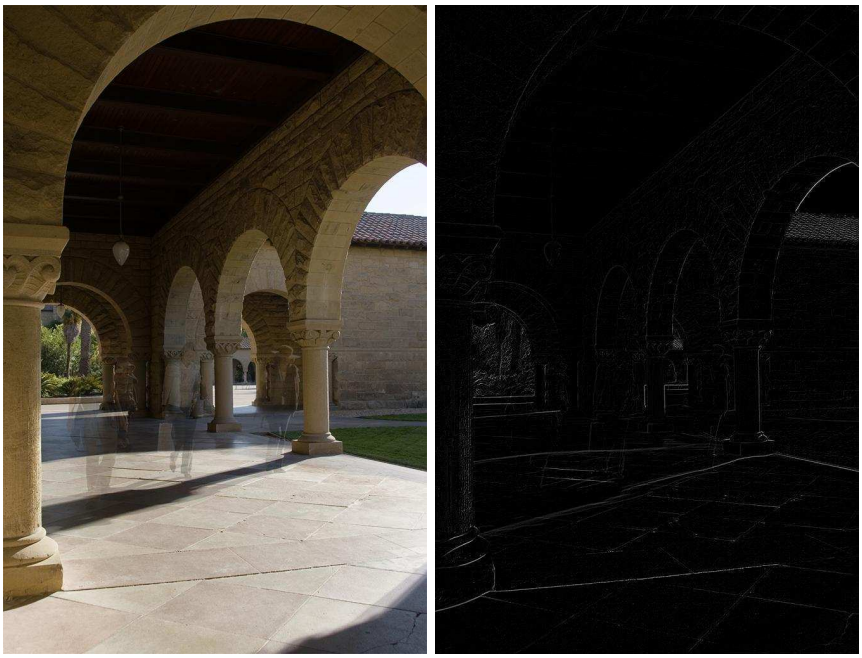
4.1 Generating static images from the LDR images of a dynamic HDR scene

For dynamic scenes, before applying any of the image retargeting approaches discussed in Chapter (3), we need to create static images (corresponding to a reference image) of the LDR image set. The two dimensional image formation can be defined by the equation (4.1).

$$I(x, y) = f(E(x, y)\Delta t) \tag{4.1}$$

where $I(x, y)$ is the pixel intensity and $E(x, y)$ is the image irrediance at the pixel location

(a)



(b)

Figure 4.1: (a) Exposure sequence of a dynamic scene, notice that in some images there are object movements. (Image Courtesy: Orazio Gallo, NVidia Research). (b) Result of applying exposure fusion on the given exposure sequence. As the scene is not static the results have ghosting artifacts and energy matrix of the output image is also shown.

(x, y), and Δt is the exposure time. Given a set of LDR images of a dynamic scenes, this equation can be modified and rewritten as equation (4.2). Even in the case of dynamic scenes, we assume that the LDR images are perfectly aligned. The change in the images is assumed to be only due to the movement of objects.

$$I_k(x, y) = f(E_k(x, y)\Delta t_k) \tag{4.2}$$

where $I_k(x, y)$ represents the LDR image having the exposure time Δt_k. In this equation, one can observe that the intensity at pixel location (x, y), varies not only due to change in exposure time but also due to the change in irradiance $E_k(x, y)$.

Consider that we have two images I_1 and I_2 of the same scene, captured using two different exposure times Δt_1 and Δt_2. We assume that appreciable number of pixels in the scene are static. Then one can define the relation between pixel intensities in the static regions of the scene by equation (4.3) [37].

$$log f^{-1}(I_2(x, y)) = log f^{-1}(I_1(x, y)) + log\left(\frac{\Delta t_2}{\Delta t_1}\right) \tag{4.3}$$

The intensity values in these two images can be related through a function known as camparametric function [38]. This comparametric function can be defined by equation (4.4)

$$I_2(x, y) = CM_{2,1}(I_1(x, y)) \tag{4.4}$$

where $CM_{2,1}$ defines the relation between the intensity values of the images I_1 and I_2. This comparametric function is uniquely defined for a pair of LDR images. Let us choose one image from the input exposure sequence as a reference image. Comparametric function is defined between the reference image and any other LDR image. For estimating the comparametric function, weighted variance measure is used. This weighted variance matrix calculates the variation in the pixel intensity for all pixel locations across the exposure sequence [21].

4.1.1 Classification of static pixels

Assume that if a particular pixel is static over all the images then the weighted variance of its intensity values across all the images in exposure stack will be very less. This is due to the fact that only exposure time is changing and the scene is static. However, if the pixel corresponds to a dynamic region of the scene then both the exposure time and the irradiance values are

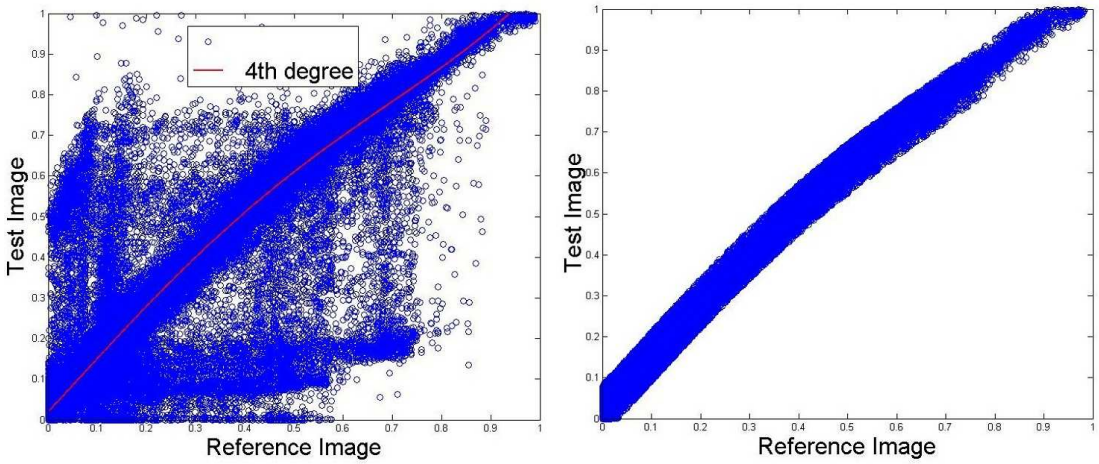
changing. Therefore the weighted variance would be high for such a pixel locations. The weighted variance can be calculated for all pixels in a given image set by equation (4.5). Here the initial process we adopt is similar to [21].

$$V(x,y) = \frac{\sum_{k=1}^{n} \varepsilon_k(x,y) I_k^2(x,y) / \overline{\varepsilon}(x,y)}{\left(\sum_{k=1}^{n} \varepsilon_k(x,y) I_k(x,y) \right)^2 / (\overline{\varepsilon}(x,y)\Big)^2} \tag{4.5}$$

Here $\overline{\varepsilon}(x,y) = \sum_{k=1}^{n} \varepsilon_k(x,y)$ and $\varepsilon(x,y) = e^{\frac{-(I_k(x,y)-0.5)^2}{2\times 0.2^2}}$. This weighted variance matrix gives the variation in the intensity values of a pixel. Variance in the intensity of a pixel at location (x,y) over all the images in exposure stack is given by $V(x,y)$. Now according to [21], if most of the pixel are static then we can classify a region of the scene as either dynamic or static through this weighed variance measure. However we do not know which LDR image led to the dynamic region of the scene even if a particular region is classified as dynamic. We assume that the pixels, where the variation in the intensities are less then $(0.25)\times\max\big(V(x,y)\big)$ to be static.

Figure 4.2(a) shows the intensity variation between the reference image and a test image. This plot is only for the pixels which are classified as static by the weighted variance measure. There could be some pixels in the test image which are static though they have been classified as dynamic by the algorithm. This is because weighted variance measure works on the variation in the pixel intensities over all the images in the exposure stack. However, if Δt_{ref} and Δt_{test} differ by a greater amount some pixels may also be classified as dynamic. This can be thought of as error due to large exposure time variation. To minimize the error due to large exposure variation we consider a finite width boundary region on both sides of the fourth order polynomial function fitted for the data. Figure 4.2(b) shows the boundary region around the fourth order polynomial function computed.

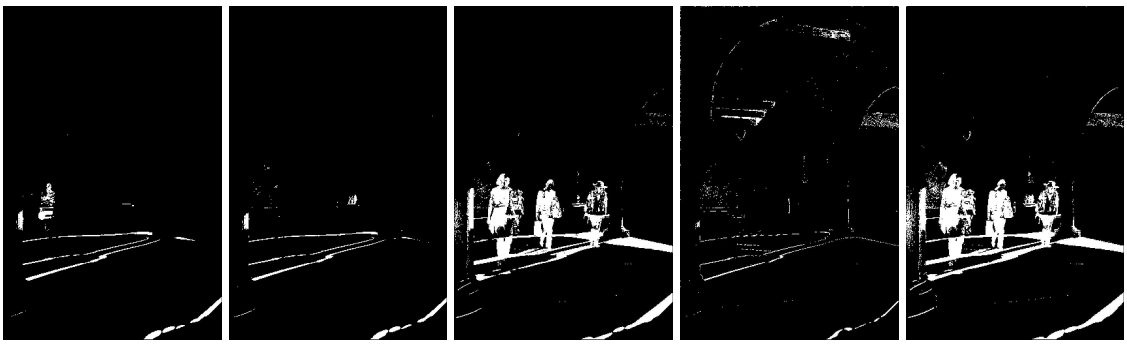
Here image intensities are normalized between 0 and 1. The polynomial function which approximate the comparametric function can be obtained by fitting a polynomial between the intensity values of the reference LDR image and a test LDR image. Figure 4.3(a) shows the dynamic region found by the algorithm in each image in the exposure stack. Figure 4.3(b) shows the dynamic regions in the scene once we combine the dynamic regions found in each LDR image of the exposure stack.



(a) (b)

Figure 4.2: Generating the Comparametric function: (a) plot corresponding to the intensity values of the static pixels found between a reference image and a test image, (b) Plot showing a boundary region around the fourth order polynomial so that error due to high exposure variation is minimized.

Approach suggested by [21] performs superpixel grouping using superpixels (region in the image with minimum texture change). For each superpixel group, if the majority of pixels comes in the dynamic category, the whole superpixel is consider as dynamic region by the algorithm. If the fraction of pixels classified as dynamic are less than a certain predefined threshold, the whole superpixel is consider to be static.



(a) (b)

Figure 4.3: (a) Shows dynamic region found by the algorithm in the input set of LDR images corresponding to the image set of Figure 4.1(a). Here we choose the first image of the exposure stack as the reference image, dynamic regions are defined in other images with respect to this reference image, (b) shows the Combined dynamic regions over all images in the exposure stack. These regions are responsible for introducing ghosting artifacts in the final high contrast LDR image.

However the major drawback of the approach suggested by [21] is that once the algorithm

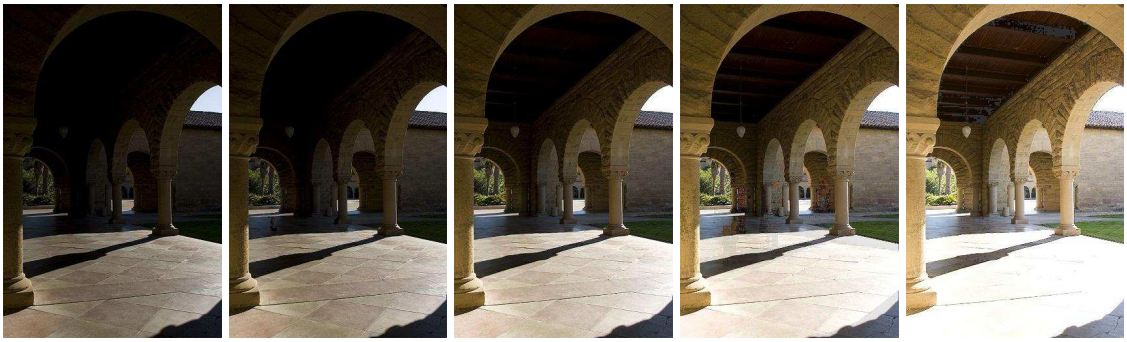
detects a region as dynamic in an image, say I_i, it will discard that region while creating the high contrast LDR image. Consider a small region in image I_i which is detected by the algorithm as dynamic. For instance, let us assume that in all other images this particular region is static. The approach by [21] discards this region while applying exposure fusion. This algorithm only uses images $k = 1$ to n where $k \neq i$ for creating high contrast LDR image for that region. However, one can easily notice that for the dynamic region, we will not get any exposure value due to exposure time Δt_i as we are discarding that region in image I_i.

We can improve this approach by not discarding the region. Instead we can get the intensities for that region from the camparametric function between image I_i and the reference image. Suppose image j is chosen as the reference image. Now the intensities for the dynamic region in I_i can be obtain through $I_i(x, y) = CM_{i,j}(I_j(x, y))$. Figure 4.4(a) shows the static images generated by this approach. Figure 4.4(b) shows the resultant high contrast LDR image obtained by this approach and through [21]. One can easily see the improvement using the proposed approach.

4.2 Aggregate Energy Matrix for Dynamic Scenes

We generate the static images of the dynamic scene corresponding to a reference image using the method discussed in the above section. Now we can merge these images and create a high contrast LDR image through exposure fusion. Figure 4.4(b) shows the result of applying the above mentioned approach on the image set of Figure 4.1(a). One can notice that still there are regions in Figure 4.4(a) which are dynamic and not detected by the algorithm appropriately. This is because those regions are very similar to the background and show a very little intensity variation with respect to the background.

Therefore such false regions will remain in the generated static images of the input exposure sequence. Due to the presence of such regions in the final high contrast LDR image, around the boundaries of the dynamic regions there are some silhouettes. Figure 4.5 shows the silhouette found in the final high contrast LDR image shown in Figure 4.4 (b). Equation (4.6) leads to a binary image D such that if a pixel is found as dynamic by our algorithm then for its pixel location $(x, y), D(x, y) = 1$ else it is zero. Figure 4.3(b) shows the image D for the data



(a)

(b) (c)

Figure 4.4: (a) Exposure sequence created by the Comparametric function after eliminating the dynamic regions. first image of the exposure was taken as the reference image.,(b) Result of exposure fusion on the generated static images. (c) Result generated by the approach describe in [21]. One can notice that contrast of the resultant image (b) is very high compared to individual images as well as image in (c) while using the proposed approach.

set given in Figure 4.1(a).

$$D(x,y) = \begin{cases} 1, & \text{if location}(x,y) \text{ is dynamic in any of input image} \\ 0, & \text{Otherwise} \end{cases} \tag{4.6}$$

Here we are only interested in the boundary (or edges of the dynamic regions). We can get the edges in image D through any of the edge detector algorithms [39]. Figure 4.5 detects the edges for the dynamic scenes. Let us define a new matrix S, showing these silhouettes.

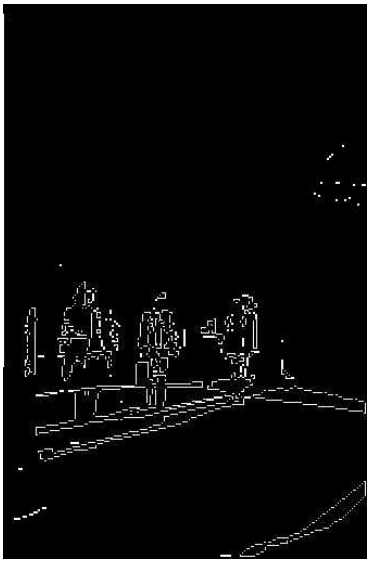


Figure 4.5: Silhouette found in the generated high contrast LDR image of Figure 4.4 (b)

In Chapter 3, we have discussed that aggregate energy matrix having weighted Laplacian as weighting parameter works better than all the other approaches for static scenes. Chapter 5 gives enough evidence about the enhancement we obtain using this approach. In spite of these advantages, we can not apply this weighted Laplacian directly to the dynamic scenes. Because when the output high contrast LDR image is of the same aspect ratio as the input images, these silhouettes are difficult to detect. However when we change the aspect ratio, these silhouettes become dominant and create artifacts in the final high contrast LDR image. To nullify these artifacts, we redefine our previous aggregate energy matrix approach with weighted Laplacian as weighting parameter for the dynamic scenes.

Consider a situation in which we want to reduce the aspect ratio of given set of LDR images of a dynamic HDR scene. We want that in the energy matrix (which decides the seams to be removed first), pixels having the silhouette to be given lesser weight compared to the other pixels. By doing so, we can assure that such pixels will appear first while calculating minimum

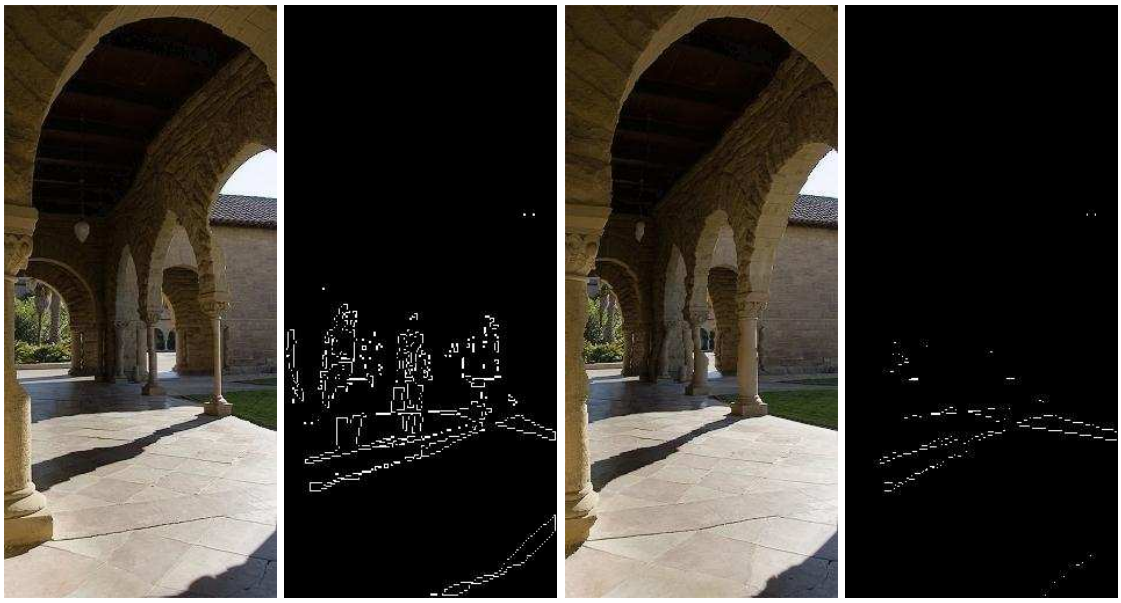
energy seams and will be removed very early in the seam removal process. Therefore once we remove a minimum number of seams, our output high contrast LDR image does not contain any silhouettes. On the other hand when we increase the aspect ratio, if the pixels which are creating silhouettes appear in the minimum energy seams to start with, the algorithm would support the replication of the silhouettes. In this case, the final LDR image tends to have large amount of artifacts introduced by these silhouettes. To avoid that, we need to give very high weight to those pixels having silhouettes so that they do not become part of the seams generated by the algorithm as minimum energy seams to be inserted.

Consider χ as the weighted Laplacian energy matrix defined by equation (3.6). S is a matrix which defines the silhouettes around the dynamic regions as explained earlier. Figure 4.5 shows the image corresponding to the matrix S. Let us define a new measure ψ as an energy matrix for the dynamic scenes as defined in equation (4.7).

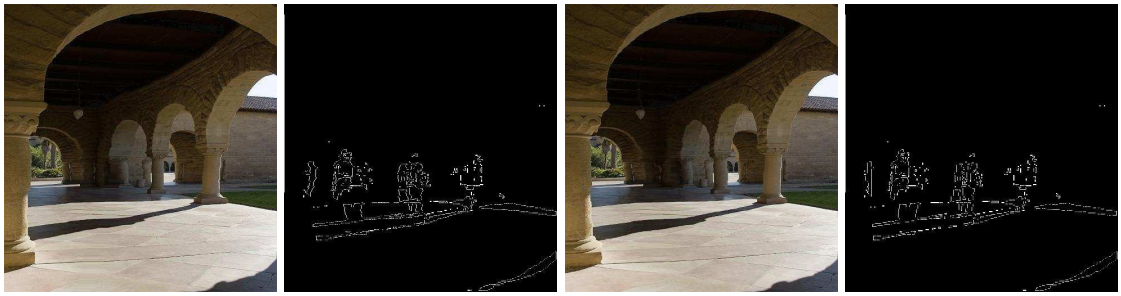
$$\psi(x,y) = \begin{cases} \chi(x,y) + \lambda S(x,y), & \text{while increasing aspect ratio} \\ \chi(x,y) - \lambda S(x,y), & \text{while reducing aspect ratio} \end{cases} \tag{4.7}$$

where λ is a sufficiently large number. This is due to the fact that in the case of reducing the aspect ratio, pixels creating silhouettes will get a very less weight in ψ. On the other hand while increasing the aspect ratio, such pixels will get very high weight in ψ.

Figure 4.6 show the result obtained for dynamic scenes. Figure 4.6 also depicts the results for both enlarging and reducing the aspect ratio. Results are shown for both aggregate energy matrix with weighted Laplacian (equation 3.6) and ψ defined in equation (4.7) used as image energy matrix. We shall discuss more about these results in Chapter 5.



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



(e) (f) (g) (h)

Figure 4.6: (a,b) and (e,f) show results when aggregate energy matrix with Laplacian weighting is used, (c,d) and (g,h) show the result when ψ defined in equation (4.7) is used as image energy function. Top Row: Image downsizing, here we reduce image size by 150 vertical seams. (a,c) and (b,d) show the output high contrast images and the silhouettes in those images respectively. Bottom row : Image enlargement, here we increase image size by 150 vertical seams. (e,g) show output images and (f,h) show the silhouettes in the output images (images are scaled for display purpose).

Chapter 5

Results and Implementation Details

THIS Chapter, presents the results achieved by proposed algorithm using various approaches discussed in previous Chapters. During the retargeting operation an image corresponding to a HDR scene, the main concern is not to lose (or add in case of enlarging) too much energy. In other words, we want resizing in a content aware manner.

Let us first consider the problem for static scenes and later we explain the results for dynamic scenes.

5.1 Static scenes

We took a large set of registered LDR images corresponding to different HDR scenes. The image sets are chosen in such a way that they comprise of different types of natural HDR scenes. In this work, we have shown the results for the image data sets - beach, tubingen river, chameleon, and building.

Table (5.1) depicts analytically that the proposed approach (equation (3.6)) works better than the direct approach. The amount of resizing is shown in terms of input and output aspect ratio for a given image set. The parameters chosen for comparing various results were average energy of a pixel and image entropy.

Consider gray scale output images having intensity values between 0-255 (8 bit per pixel). Entropy of an image defines how well the histogram of the image is distributed. In other words, it defines the uniformity of intensity distribution in the image. More the Entropy is, more the information contained in the image will be. Entropy of an image typically varies between 0 bit

and 8 bit, and defined by equation (5.1).

$$\text{Entropy} = -\sum_{i=0}^{255} p_i log_2 p_i \tag{5.1}$$

$$\text{where}, p_i = \frac{\text{Number of pixels having intensity } i}{\text{Total number of pixels}}$$

For all the image data sets used, Table (5.1) shows that the aggregate energy matrix (with weighting parameter as weighted Laplacian - equation (3.6)) approach for content aware resizing works better than all other approaches. This approach gives enhancement in both energy per pixel and entropy compared to others. Enhancement in entropy specifies that more information is retained by aggregate the energy matrix approach. Similarly, more energy per pixel proves that this approach preserves salient features better.

(a) Exposure sequence

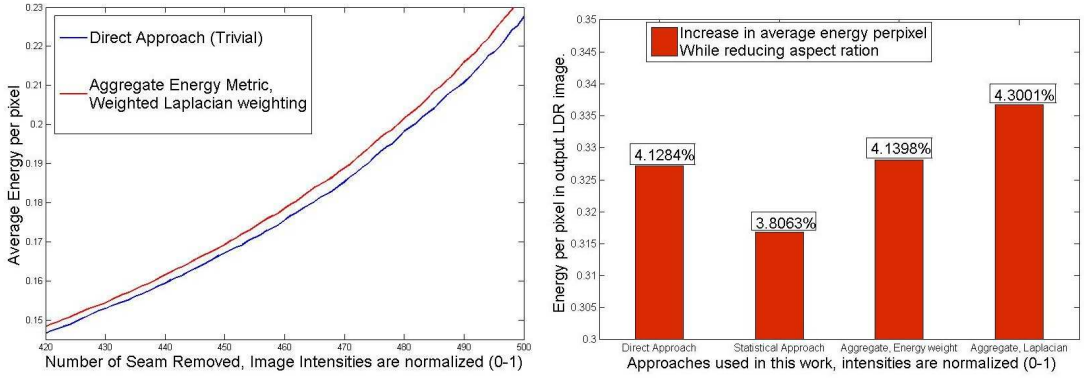
(b) (c) (d)

Figure 5.1: Data Set: Chameleon, Results while increasing the aspect ratio of output high contrast LDR image (70% resizing). Top row: (a) Exposure sequence, Bottom row: results using (b) Direct approach, (c) Statistical approach, (d) Aggregate energy matrix approach. Marked region indicates that the results of aggregate energy matrix using weighted Laplacian has better content preservation compared to direct and statistical approaches. Images Courtesy: Erik Reinhard, University of Bristol.

Figure 5.2(a) shows the change in average energy of a pixel, with removal of minimum energy seams using direct approach and aggregate energy matrix approach. It shows that initially both work similar, but as we move to the higher degree of resizing (in this case downsizing), aggregate energy matrix with weighted Laplacian as a weighting parameter will preserve more energy. Figure 5.2(b) shows the quantitative information about how much energy is increased (in percentage), while resizing through various approaches.

Image Set	Image Data Size (Spatial Resolution)		Direct Approach (Output data)		Statistical Approach (Output data)		Aggregate Energy (Average Energy Weighting) (Output data)		Aggregate Energy (Weighted Laplacian weighting) (Output data)	
	Input	Output	Energy Per pixel	Entropy	Energy Per pixel	Entropy	Energy Per pixel	Entropy	Energy Per pixel	Entropy
Tubengen	560x374	160x374	16.7931	7.5186	16.1600	7.4843	16.7760	7.5297	17.0580	7.5541
Chameleon	480x318	480x58	16.6683	7.2185	8.5193	6.6383	16.6092	7.1372	16.9695	7.2301
Beach	480x360	180x60	23.5856	7.3428	15.8635	7.0173	23.5428	7.3481	23.6575	7.3744
Building	675x450	65x450	18.7093	7.7743	3.3879	7.2286	18.7315	7.7809	18.9671	7.8013

Table 5.1: This result is produced only for down sizing the images. Image Intensities are taken as between 0-255 (8 bit). These results are only for the static scenes.



(a) (b)

Figure 5.2: Change in energy per pixel in the final LDR image corresponding to a static HDR scene. (a) Comparison between the energy retained in final output high contrast LDR image through direct approach and aggregate energy matrix approach. The plot shows that the proposed approach preserves more energy. (b) Percentage gain in the output high contrast LDR image, in all approaches discussed.

Figure 5.1 shows the results while enlarging the final high contrast LDR image by inserting seams. It can be seen that the aggregate energy matrix approach (Figure 5.1(d)) yields the best result. Figure 5.4(a) shows how the artifacts are introduced while enlarging the input images by the direct approach beyond a certain limit. However, in the same case (see Figure 5.4(b)) aggregate energy matrix with Laplacian as weight parameter yields better results. The respective energy distributions are shown alongside.

5.2 Dynamic scenes

In the case of dynamic scenes, the above mentioned aggregate energy matrix with weighted Laplacian as weighting parameter does not give good results. This is because of the limitation in the algorithm describe in Chapter (4). The algorithm creates static images of a set of dynamic images with respect to a given reference image. The problem with this algorithm is that while creating static image there are still silhouettes remaining around the dynamic region boundary found in that image. Figure 4.5 shows the silhouette found in the final LDR image of the image set shown in Figure 4.1 (a).

Figure 4.6 shows the results of the proposed approach for both enlargement and reduction of the aspect ratio of a dynamic scene. Figure 4.6 also shows the results of using the approaches - aggregate energy matrix (weighted Laplacian - equation (3.6)) and ψ defined in equation (4.7).

(a)

(b) (c) (d)

Figure 5.3: Data Set: Tubingen River, Comparison between the energy distribution of final resized high contrast LDR images created through various approaches. Images are reduced horizontally upto 70% of their original resolution. Top row: (a) Input exposure sequence, Bottom row: results using, (b) Direct approach, (c) Statistical approach (minimization of total minimum energy), and (d) Aggregate energy matrix approach with weight assigned according to average pixel energy, (b-d) Left: resized LDR image, Right: Energy distribution of resized LDR images. Marked region shows visually how the proposed approach works better in preserving content information in that region.

Figures 4.6 (a-d) show the results using equation 3.6 and ψ respectively. One can notice that by using ψ in both Figures 4.6((d) and (h)) the silhouettes which remain after resizing are very few compared to the silhouettes that remain in the Figures 4.6((b) and (f)) using aggregate energy matrix approach. The reason behind this has already been explained in Chapter 4.

(a) Retargeting, enlarging (b)

Figure 5.4: Data Set: Beach, Image resizing of the exposure sequence (see figure 3.2), using (a) Direct approach, and (b) Aggregate energy matrix approach with weighted Laplacian as the weighting parameter. Here images are enlarge vertically.

Chapter 6

Discussion

ALL the Chapters discussed so far, presents a novel approach for obtaining the flexibility in both range and spatial resolutions in the digital images. This Chapter discusses the add-ones one can add to the proposed approach. The methods describe here can be apply to the problem of similar nature or can be thought of as an entirely new research paradigm.

6.1 Improved Exposure Fusion

Exposure fusion tries to increase the contrast of an LDR image corresponding to a HDR scene, using multiple LDR images of the same scene [2]. In the proposed approach, we first reduce the size of each LDR image (by our seam removal technique) and then insert the same number of seams in each image, in order to get the same image size back. Finally we apply exposure fusion to obtain the final high contrast LDR image. Number of seams to be removed in the first step depends on the image characteristics. As we know that the minimum energy seams will be removed first, the visualization of the final output image will be better than the exposure fusion approach.

Figure 6.1 shows the comparison between our method and the direct exposure fusion approach. It can be seen that there is a higher increase in average energy per pixel in the proposed approach when compared with the direct approach (see Figure 6.1). However, this approach will not work in all situations. This is due to the fact that first removing certain seams (with least energy), and then inserting the same number of new seams will affect the contrast of the whole image. The resultant image will no longer be the same as the LDR image generated by exposure fusion alone.

(a) (b) (c)

Figure 6.1: Enhancement Over Exposure Fusion: (a) LDR image generated by applying exposure fusion on the input set of the LDR images, (b) LDR image generated through the proposed approach, here we first remove some seams and then again add same number of seams in order to get same aspect ration back. and (c) Comparison between both the techniques, the bar chart shows that the proposed approach preserves more energy.

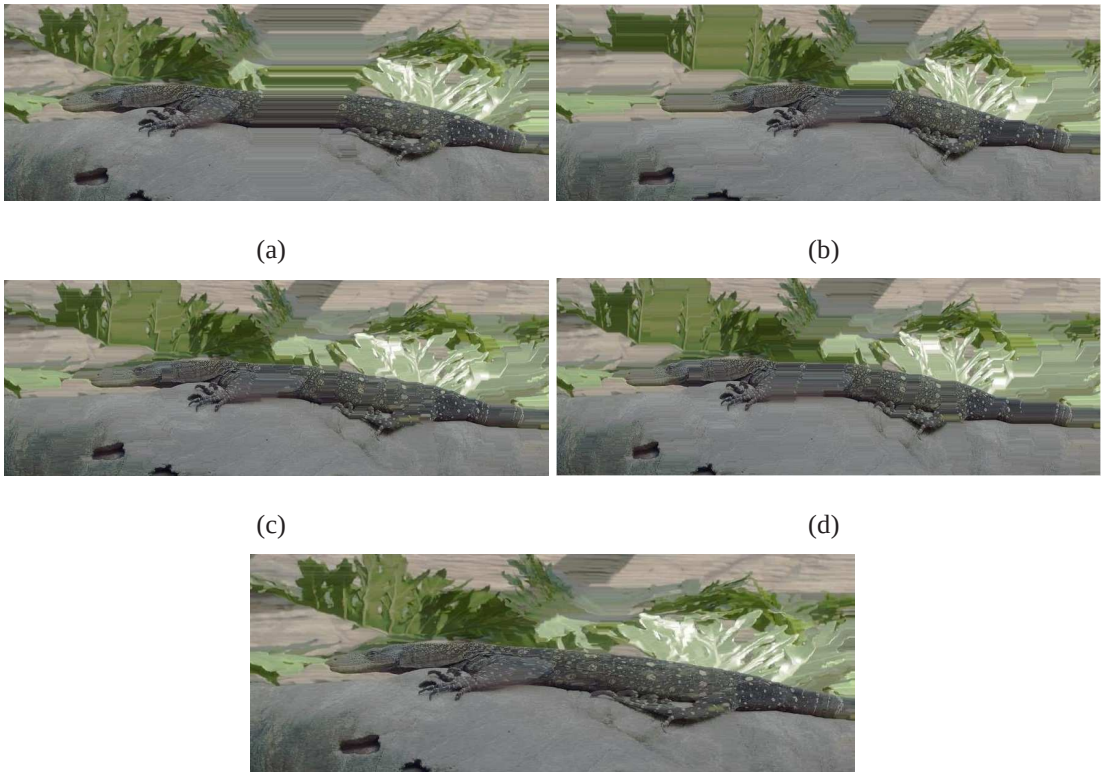
6.2 Efficient way to Enlarge High Contrast LDR Images

While enlarging an image, we first find the seams with minimum energy in each of the input images I_r, say c_r. We now introduce a new seam along the seam c_r, say $c_r^{'}$, as an average of the seam c_r and the seam next to it. It may be noted that the seams are to be inserted in the order of their removal. In the proposed approach, we insert new seams in ascending order of their minimum energy. That is, the first minimum energy seam should be inserted first and so on. If the proposed order is not maintained, the algorithm might choose the same seam for insertion repeatedly.

However after a certain number of seam insertions, it might happen that the same seam starts to repeat itself as shown in the Figure 6.2(a). To overcome this limitation, we modify our algorithm by inserting different number of seams at each stage. We shall illustrate this strategy by considering an exponentially decreasing function to model the number of seams to be inserted at each stage. Consider a situation where we want to insert N seams. Let us insert $\frac{N}{2}$ seams and reconstruct the energy function. Using the resultant energy function, let us insert $\frac{N}{4}$ seams and reconstruct the energy function. In the kth step, insert $\frac{N}{2^k}$ seams and reconstruct the energy function. By this method, the total number of seams inserted are $\frac{N}{2} + \frac{N}{4} + \frac{N}{8} + \ldots = N$. Although the number of stages described above seem infinite, for any $2^k \leq N \leq 2^{k+1}$ we need utmost $k+1$ distinct insertion steps.

It can be seen that inserting N seams in such a way produces a sufficiently better result than the previously stated approaches (see Figure 6.2(e)). This result shows that the use of exponentially decreasing function for choosing the number of seams to be inserted at a given step

yields better results than using other functions - linearly decreasing function (Figure 6.2(b)), quadratically decreasing function (Figure 6.2(c)), and uniform fixed step size function (Figure 6.2(d)).



(a) (b)

(c) (d)

(e)

Figure 6.2: Various resizing methods used to increase an image 75% horizontally. (a) Direct increase without steps, (b) Using a linearly decreasing function for steps, (c) Using a quadratic decreasing function for steps, (d) Using a fixed size at each step, and (e) Using an exponential decreasing function for steps.

6.3 Exploring New Measures for Energy Calculation

In the proposed work, we have used gradient magnitude defined by equation (3.3) in order to model the energy function for detecting seams. Seams with minimum energy have the least resistance (intensity change) in their path. Because of this fact, removing or adding such seams will not affect the salient features in the image. In the original seam carving approach, this energy function is consider better than all other energy measures [5].

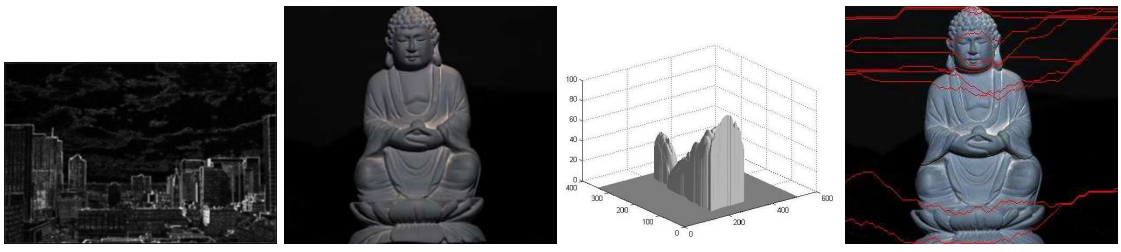
However this particular measure is based only on neighbouring image intensity values. Instead, focused region in the image can also be used as an energy function [40]. In this case,

seams identified as having minimum energy will contain pixels corresponding to the most defocused regions in the image. If we remove such seams first, it implies that we are removing the defocused regions first. Consider $F(x, y)$ to be the degree of focus, such that if $F(x, y) = 1$ then pixel at location (x, y) is at the focus plane. As $F(x, y)$ decreases, location (x, y) becomes more and more defocused. Figure 6.3 (a) shows the focus map of the image set buildings shown in Figure 3.3 using the approach [40].

Another measure for energy function could be the depth map corresponding to the image. One can specify depth boundaries up to which seams need to be retained in the final LDR image. The major challenge here is to find a good depth estimation algorithm for the given scene. However estimation of the depth map of outdoor (natural) scenes is a challenging problem. We created a laboratory set-up for creating a depth map of the scene and used photometric stereo method for creating the depth map [41]. The overhead with this approach is that we need few other images in the exposure stack having different illumination conditions.

If $Z(x, y)$ defines the depth map of the scene in Figure 6.3(b), such that $Z(x, y)$ gives the depth at pixel location (x, y) in the image. Assume that the depth ranges between the limits $lower$ and $upper$. We created a boundary region in the depth map between $lower + \frac{lower+upper}{4}$ and $upper - \frac{lower+upper}{4}$. The pixels whose depths fall in this range are chosen to be having salient features and we require the pixels having these values to be preserved in the final LDR (resized) image.

Figure 6.3(b-c) show an image of Buddha and its corresponding depth map generated by photometric stereo method respectively. Figure 6.3(d) shows the seams having the least energy measures using depth map as the energy function.



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

Figure 6.3: New measures for Image energy function: (a) shows the focus map of image set building of Figure 3.3. If used it as an energy function, the least important seams chosen by the proposed algorithm will be most the defocused seams. (b) A scene containing object (Buddha), (c) its depth map estimated using photometric stereo, and (d) the generated minimum energy seams by the algorithm if the depth map is chosen as the energy matrix.

Chapter 7

Conclusion

We have proposed a novel approach for the content aware resizing of multi-exposure images of static or dynamic HDR scene before fusing them into a high contrast LDR image. The proposed approach efficiently combines the content aware image retargeting and the multi-exposure images to develop a novel application suitable for any digital display device. We showed that the proposed algorithm performs better when compared to the direct approach of fusing the multi-exposure images before content aware resizing. We have shown through experiments that the LDR image results generated using the proposed statistical and aggregate energy matrix approaches to be far better with regard to both the contrast as well as the analytical measures. The optimal selection of seams to be inserted or deleted leads to highly robust retargeting algorithm. The proposed approach is fully automatic with no user intervention.

7.1 Interdisciplinary aspect of the work

In the present work, in order to quantify the results we have used concepts from information theory (Average information or Entropy). The average information shows the amount of randomness present in any physical distribution. We have used this concept for quantifying the distribution on intensity values in an image. On the similar notes Seam carving algorithm can be thought of as an inspiration taken from electricity flow, or heat flow. Consider that salient features of an image (important regions where the highest intensity changes) as the area which produces highest resistance in the path of minimum energy seams. If we have multiple path through which current or heat can be flow, in this case the path which shows high resistance, minimum amount of current (heat) flow through that path. In seam carving our seam with

minimum energy can be thought of as path showing low resistance, there for preserves (do not contain) salient features.

We have generated a mathematical model for implementing the proposed approach. Mathematical modeling is the key work in the thesis. To model the problem for a set of LDR images having different exposure time, both the static as well as dynamic scenes needs complex mathematical tools.

7.2 Applications

As the approach does not involve minimization of any complex cost function, it is computationally inexpensive. Because of this reason the proposed technique can be applied for a large set of computational photography problems.

- The proposed algorithms open up a wide possibility of retargeting and fusion techniques which can be customized for a given display device.
- The developed algorithms can either be included along with the state-of-the-art mobile cameras/digital cameras or,
- The solution can also be provided as applications for post capture image processing softwares.

7.3 Future Pointers

The proposed approaches assume perfectly registered images of a static scene which is a hard constraint to be placed on a real world scene. We hope that the proposed approach can be improved and extended in the case of dynamic scenes which tend to introduce ghosting artifacts. Further, we hope to extend this approach for video image retargeting applications involving HDR scenes. We believe that the novel approach discussed here would lead to more novel ideas in the flexible resolution image retargeting research.

Bibliography

[1] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski, *High dynamic range imaging: acquisition, display, and image-based lighting*. Morgan Kaufmann, 2010.

[2] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion," in *Computer Graphics and Applications, 2007. PG'07. 15th Pacific Conference on*, pp. 382–390, IEEE, 2007.

[3] S. Raman and S. Chaudhuri, "A matte-less, variational approach to automatic scene compositing," in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pp. 1–6, IEEE, 2007.

[4] S. Park, M. Park, and M. Kang, "Super-resolution image reconstruction: a technical overview," *Signal Processing Magazine, IEEE*, vol. 20, no. 3, pp. 21–36, 2003.

[5] S. Avidan and A. Shamir, "Seam carving for content-aware image resizing," in *ACM Transactions on graphics (TOG)*, vol. 26, p. 10, ACM, 2007.

[6] H. Zimmer, A. Bruhn, and J. Weickert, "Freehand hdr imaging of moving scenes with simultaneous resolution enhancement," in *Computer Graphics Forum*, vol. 30, pp. 405–414, Wiley Online Library, 2011.

[7] B. K. Gunturk and M. Gevrekci, "High-resolution image reconstruction from multiple differently exposed images," *Signal Processing Letters, IEEE*, vol. 13, no. 4, pp. 197–200, 2006.

[8] J. Choi, M. K. Park, and M. G. Kang, "High dynamic range image reconstruction with spatial resolution enhancement," *The Computer Journal*, vol. 52, no. 1, pp. 114–125, 2009.

[9] S. Mann and R. W. Picard, "On being undigital with digital cameras: Extending dynamic range by combining differently exposed pictures," in *IS & T Conference*, 1995.

[10] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '97, (New York, NY, USA), pp. 369–378, ACM Press/Addison-Wesley Publishing Co., 1997.

[11] T. Mitsunaga and S. K. Nayar, "Radiometric self calibration," in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, vol. 1, IEEE, 1999.

[12] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcovs, "High dynamic range display systems," *ACM Transactions on Graphics (TOG)*, vol. 23, no. 3, pp. 760–768, 2004.

[13] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Transactions on Graphics (TOG)*, vol. 21, no. 3, pp. 267–276, 2002.

[14] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," *ACM Transactions on Graphics (TOG)*, vol. 21, no. 3, pp. 249–256, 2002.

[15] P. Burt and E. Adelson, "The laplacian pyramid as a compact image code," *Communications, IEEE Transactions on*, vol. 31, no. 4, pp. 532–540, 1983.

[16] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, "Digital photography with flash and no-flash image pairs," in *ACM Transactions on Graphics (TOG)*, vol. 23, pp. 664–672, ACM, 2004.

[17] E. Eisemann and F. Durand, "Flash photography enhancement via intrinsic relighting," in *ACM Transactions on Graphics (TOG)*, vol. 23, pp. 673–678, ACM, 2004.

[18] E. A. Khan, A. Akyuz, and E. Reinhard, "Ghost removal in high dynamic range images," in *Image Processing, 2006 IEEE International Conference on*, pp. 2005–2008, IEEE, 2006.

[19] O. Gallo, N. Gelfandz, W. Chen, M. Tico, and K. Pulli, "Artifact-free high dynamic range imaging," in *Computational Photography (ICCP), 2009 IEEE International Conference on*, pp. 1–7, IEEE, 2009.

[20] F. Pece and J. Kautz, "Bitmap movement detection: Hdr for dynamic scenes," in *Visual Media Production (CVMP), 2010 Conference on*, pp. 1–8, IEEE, 2010.

[21] S. Raman and S. Chaudhuri, "Reconstruction of high contrast images for dynamic scenes," *The Visual Computer*, pp. 1–16, 2011.

[22] W. Zhang and W.-K. Cham, "Gradient-directed multiexposure composition," *Image Processing, IEEE Transactions on*, vol. 21, no. 4, pp. 2318–2323, 2012.

[23] J. Hu, O. Gallo, and K. Pulli, "Exposure stacks of live scenes with hand-held cameras," in *ECCV*, 2012.

[24] P. Sen, N. K. Kalantari, M. Yaesoubi, S. Darabi, D. B. Goldman, and E. Shechtman, "Robust patch-based hdr reconstruction of dynamic scenes," *ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH Asia 2012)*, vol. 31, no. 6, 2012.

[25] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 1, pp. I–511, IEEE, 2001.

[26] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 20, no. 11, pp. 1254–1259, 1998.

[27] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 9, pp. 1167–1183, 2002.

[28] A. Efros and W. Freeman, "Image quilting for texture synthesis and transfer," in *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pp. 341–346, ACM, 2001.

[29] A. Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin, and M. Cohen, "Interactive digital photomontage," *ACM Transactions on Graphics (TOG)*, vol. 23, no. 3, pp. 294–302, 2004.

[30] M. Rubinstein, A. Shamir, and S. Avidan, "Improved seam carving for video retargeting," in *ACM Transactions on Graphics (TOG)*, vol. 27, p. 16, ACM, 2008.

[31] M. Rubinstein, D. Gutierrez, O. Sorkine, and A. Shamir, "A comparative study of image retargeting," in *ACM Transactions on Graphics (TOG)*, vol. 29, p. 160, ACM, 2010.

[32] F. Banterle, A. Artusi, T. Aydin, P. Didyk, E. Eisemann, D. Gutierrez, R. Mantiuk, and K. Myszkowski, "Multidimensional image retargeting," in *SIGGRAPH Asia 2011 Courses*, p. 15, ACM, 2011.

[33] S. G. Narasimhan and S. K. Nayar, "Enhancing resolution along multiple imaging dimensions using assorted pixels," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 4, pp. 518–530, 2005.

[34] A. El Gamal, "High dynamic range image sensors," in *Tutorial at International Solid-State Circuits Conference*, 2002.

[35] M. Granados, B. Ajdin, M. Wand, C. Theobalt, H.-P. Seidel, and H. P. Lensch, "Optimal hdr reconstruction with linear digital cameras," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pp. 215–222, IEEE, 2010.

[36] S. W. Hasinoff, F. Durand, and W. T. Freeman, "Noise-optimal capture for high dynamic range photography," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pp. 553–560, IEEE, 2010.

[37] M. D. Grossberg and S. K. Nayar, "Determining the camera response from images: What is knowable?," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, no. 11, pp. 1455–1467, 2003.

[38] S. Mann, "Comparametric equations with practical applications in quantigraphic image processing," *Image Processing, IEEE Transactions on*, vol. 9, no. 8, pp. 1389–1406, 2000.

[39] J. Canny, "A computational approach to edge detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, no. 6, pp. 679–698, 1986.

[40] S. K. Nayar and Y. Nakagawa, "Shape from focus," *Pattern analysis and machine intelligence, IEEE Transactions on*, vol. 16, no. 8, pp. 824–831, 1994.

[41] R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Optical engineering*, vol. 19, no. 1, pp. 191139–191139, 1980.

Acknowledgments

It gives me immense pleasure in presenting my thesis report. I would like to take this opportunity to express my deepest gratitude to the people, who has contributed their valuable time for helping me to successfully complete my thesis and whose constant guidance and encouragement crown all the efforts with success.

It is my profound privilege to express my deep sense of gratitude to **Dr. Shanmuganathan Raman**, and **Dr. Puneet Sharma**, Asst. Professor, Indian Institute of technology, Jodhpur for their precious guidance, constructive encouragement and whole hearted support throughout the thesis tenure.

Finally, I would like to thank my friends Pooja Singh, Akash Yadav, Sapana Ranwa, Naman Joshi, Preeti Yadav and Anurag, For there kind support. I would also like to thank my college staff and all the people who were directly and indirectly involved in my thesis.

Date: _____ Govind Salvi