# Efficient Image Retargeting For High Dynamic Range Scenes

**Govind Salvi**
**PG201172008**

Indian Institute of Technology, Jodhpur

Supervisor(s):

**Dr. Shanmuganathan Raman**
**Dr. Puneet Sharma**

May 30, 2013

# Outline of the **talk**


1 Introduction
- Problem Statement
- Thesis Objectives

2 Background
- Limitation in Dynamic Range
  - Exposure Fusion
- Limitation in Spatial Resolution
  - Seam Carving

3 Retargeting HDR Scenes
- Static Scenes
  - Proposed Approaches
- Towards the Dynamic Scene
  - Retargeting Dynamic Scene

4 Results

5 Discussion

6 Conclusion

# *Introduction*

*Problem and Objective*

# Introduction

- Real world scenes have a very high dynamic range (HDR).
- The mobile phone cameras and the digital cameras presently available are limited in their capability in both the range and spatial resolution.
- Images while travelling from one devices to other suffers from the problem of aspect ratio change.
- Real time aspect ration change, made it very difficult, for images to preserves their content information upon resizing.

# Problem statement

*Digital images are affected with two serious limitations*

- Limitation In Dynamic Range,
- Limitation In Spatial Resolution.

- Mobile phone cameras, can not capture all the brightness levels of real world,
- Similarly, suffers with constant image resolution,
- Quality, cameras are costly,
- A low cost solution is always desirable.

# Thesis objectives

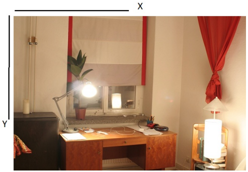
1. Flexible content aware spatial retargeting of an image corresponding to a HDR scene,
2. Depiction of high contrast information of the scene within the user specified spatial resolution,
3. Achieving high quality LDR images without any visible artifacts even in the case of dynamic scenes, and
4. Assumption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

# *Background*

*Understanding the Problem*

# Image formation

## 2D Image Formation




$$I(x, y) = f(E(x, y)\Delta t) \tag{1}$$

- $I(x, y)$, Image Intensity value at pixel location $(x, y)$,
- $E(x, y)$ Image Irradiance at pixel location $(x, y)$, and
- $\Delta t$ Exposure time(Camera shutter speed).

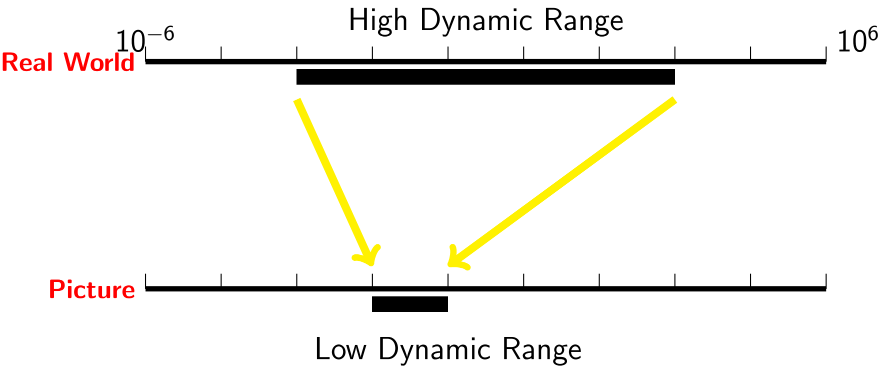
# Limitation In Dynamic range

## Dynamic range

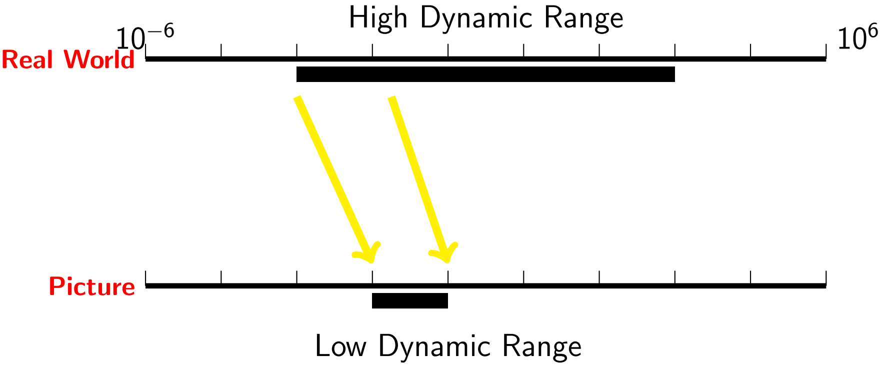
In Computational photography, Dynamic range specifies the Irradiance range of the scene begin photographed. In other words it is the ratio between highest to lowest Irradiance value present in the image.

- Human eye has a very High Dynamic range.
- Human can see things form low star light to high sunlight.
- Mobile phone cameras and other low cost cameras are limited in Dynamic range.

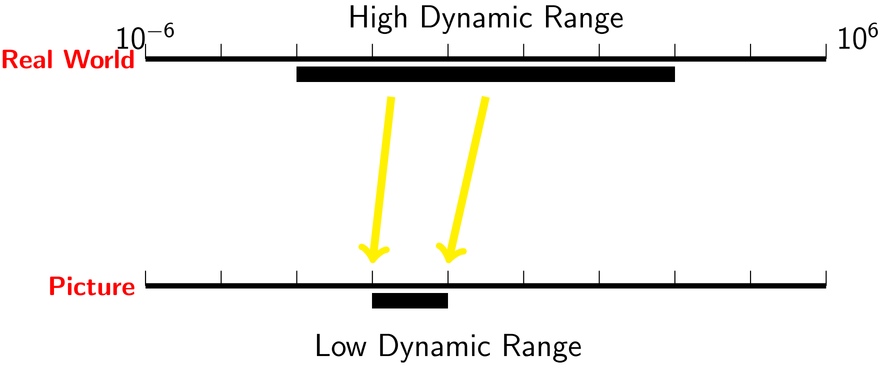
# Multi Exposure Photography



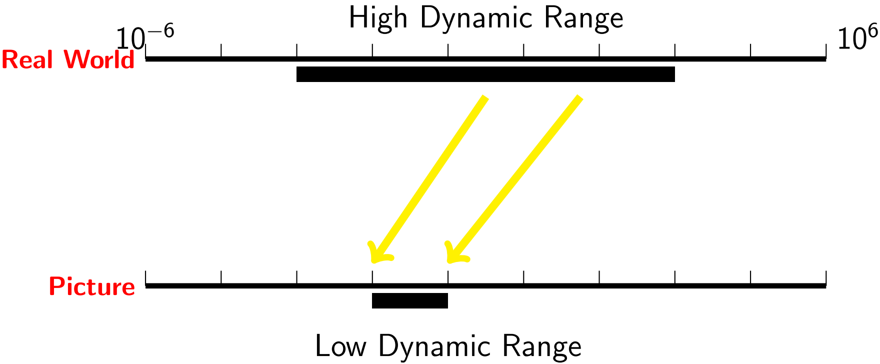
# Multi Exposure Photography



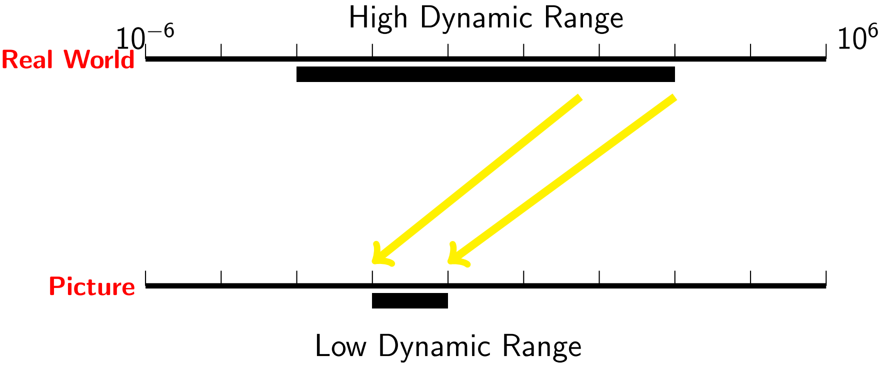

# Multi Exposure Photography




# Multi Exposure Photography




# Multi Exposure Photography




# Exposure Fusion

- Popularly known method, for generating High Contrast LDR image from a given set of LDR images with different exposure time.
- Input images must be perfectly aligned.
- Generation Criteria
  - Contrast(C): it gives a high weight to edges and texture.
  - Saturation(S): standard deviation between R,G and B color channel, at each pixel
  - Well-Exposedness(E): we do not want the intensities which are either underexposed or overexposed.
- The final image is obtained by collapsing the stack using weighted blending.

[1]

[1] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion."

# Exposure Fusion

- *Calculating Weight*

$$W_{ij,k} = (C_{ij,k})^{\omega_C} \times (S_{ij,k})^{\omega_S} \times (E_{ij,k})^{\omega_E} \quad (2)$$

  $\omega$ is weighted assign to that particular parameter that is if $\omega_E$ 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

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

- *Final image*

$$R_{i,j} = \sum_{k=1}^{n} \bar{W_{ij,k}} I_{ij,k} \quad (3)$$

# Exposure Fusion, Example

(a) Exposure Sequence

(b) Exposure Fused Image

2

[2] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion."

# Limitation In spatial Resolution

- Images captured with mobile phone cameras are of constant spatial resolution,
- While moving form one aspect ratio to other, Images should be retargeted,
- Not all the regions of an image are equally important,
- Retargeting should be done, in such a way that object of interest should be minimally effected.

# Seam Carving or Content aware resizing

Figure : Changing the aspect ratio

- Resize Images in content aware manner.
- Use Energy metric to find object of interest.
- Remove(Down sizing) or Add (Enlarging) least important seams, while resizing.

# Seam Carving

*How it Works?*

- We have used Magnitude of gradient as an image energy measure.

$$\chi(x, y) = \left| \frac{\partial I(x, y)}{\partial x} \right| + \left| \frac{\partial I(x, y)}{\partial y} \right| \tag{4}$$

- There could be other energy measures also, like visual saliency. We have defined some new energy measure, will be discussed in section discussion.
- Energy measure, defines the importance of the regions.
- Efficiency of retargeting results highly depend upon this energy matrix.

3

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

| | | | |
|---|---|---|---|
| 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 : (a) Energy matrix (magnitude of gradient), (b) Cumulative matrix, (c) vertical seam on a natural image

Figure : Seam Carving, algorithm

- *Cumulative matrix* $C(x, y) = \chi(x, y) + \min\{C(x-1, y-1), C(x-1, y), C(x-1, y+1)\}$,
- Similarly, minimum energy seams could be found in horizontal direction.

# *Seam Carving*

*Demonstration*

# Seam Carving Demo

## Example video

# *HDR Retargeting*

*Static and Dynamic scenes*

# Static Scenes

Figure : Exposure sequence, for a static HDR scene

- Images are perfectly aligned,
- Images differ in only exposure time (Camera shutter speed),
- No object movement.

# Direct Approach

- A general solution of the given problem is: first generate a high contrast LDR image of the given LDR image sequence (using exposure fusion),
- Apply optimal seam carving on the generated high contrast LDR image for retargeting,
- However, the question to be asked here is: whether this approach is the optimal one or can we find alternate better solutions.

# Motivation towards other solutions

- Performing exposure fusion before resizing deprives from information presents in the original LDR images,
- This operation includes least important seams in the exposure fusion process,
- Hence, the contrast of the final high contrast LDR image depends mainly on the exposure fusion algorithm,
- Instead, minimum energy seams can be found on individual images and removed or added first before exposure fusion.

# Statistical Approach

Figure : 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.

# Statistical Approach

- For a given LDR image $I_r$, let $c_r$ denotes the computed minimum energy seam and $e_r$ denote the corresponding cumulative energy,
- $e_k = \min\{e_r : r = 1, 2, \ldots, n\}$,
- In this case, seam $c_k$ will be chosen as the minimum energy seam and considered for retargeting operation.
- Instead one can also choose the seam having median cumulative energy value,
- The accuracy of the result solely depends upon the scene statistics and geometry.

# Aggregate Energy Metric

- Instead of finding Energy matrices for each LDR image from exposure stack, create a single modified energy matrix for retargeting,
- This modified energy matrix, should be some function of energy matrices of individual LDR images,

$$\chi = h(\chi_1, \chi_2, \chi_3, \ldots, \chi_n) \tag{5}$$

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.

*Assume first, this function* $h$ *to be a linear combination of individual energy matrices*

$$\chi = \sum_{r=1}^{n} \alpha_r \chi_r, \qquad \sum_{r=1}^{n} \alpha_r = 1 \tag{6}$$

- Parameter $\alpha_r$ corresponds to the weight given to the energy matrix of image $I_r$.
- The average energy of a pixel in a given image could be used as a weighting parameter.
- In this case $\alpha_r$ would be the average energy per pixel in image $I_r$ and is defined as $\alpha_r = \sum_{x=1}^{M} \sum_{y=1}^{N} \chi_r(x, y)$.

# Weighted Laplacian

- Laplacian of an image calculates the second order derivative along both the spatial directions, $L_i = \left|\frac{\partial^2 I_i(x,y)}{\partial x^2}\right| + \left|\frac{\partial^2 I_i(x,y)}{\partial y^2}\right|$
- Laplacian operator can also serve as an edge detector kernel,
- Give weight to the energy matrices of individual LDR images through the Laplacian operator.

$$\chi(x,y) = \sum_{i=1}^{n} L_i^*(x,y)\chi_i(x,y) \tag{7}$$

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

Both multiplication and division are performed element wise.

# Retargeting, various proposed approaches

*Image reduced horizontally, 70% of their original resolution.*

(b) Direct approach (c) Statistical approach (d) Aggregate energy matrix

# Retargeting, Enlarging

*Image enlarge vertically, 40% of original resolution*

(a) (b)

Figure : Data Set: Beach, (a) Direct approach, and (b) Aggregate energy matrix approach with weighted Laplacian as the weighting parameter.

# *Dynamic Scenes*

# Towards the Dynamic Scene

## *Image formation for Dynamic scene*

$$I_k(x, y) = f(E_k(x, y)\Delta t_k) \tag{8}$$

- In real situations, capturing a static scene is very difficult,
- Scene become dynamic due to object movement,
- Equation 8 shows image formation of a dynamic scene.

# Dynamic Scenes

*For Dynamic Scenes, retargeting become more complex as*

- Applying Exposure fusion on dynamic scenes results into ghosting effect,
- Before retargeting, we need to get rid of this ghosting effect.

# Retargeting Dynamic Scene

- Consider two images $I_1$ and $I_2$ of the same scene, captured using two different exposure times $\Delta t_1$ and $\Delta t_2$.
- Function $f$ known as Camera response function, is non linear in nature and computationally difficult to estimate accurately.
- Define a new function $CM$, known as Comparametric function

$$I_2(x, y) = CM_{2,1}(I_1(x, y)) \tag{9}$$

$CM_{2,1}$ defines the relation between the intensity values of the images $I_1$ and $I_2$.

# Classification of static pixels

- If a pixel is static in all the images of the exposure stack, variation in its intensity value will be very less,
- However if it is dynamic, the intensity variation will be very high,
- Amount of intensity variation for each pixel, is given by a measure known as weighted variance.

# Classification of static pixels

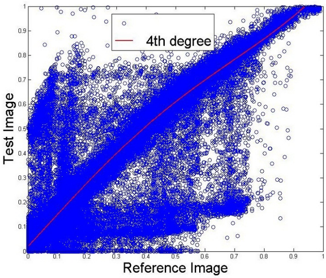
- *Weighted variance*

$$V(x,y) = \frac{\sum_{k=1}^{n} \varepsilon_k(x,y) I_k^2(x,y)/\overline{\varepsilon}(x,y)}{\left(\sum_{k=1}^{n} \varepsilon_k(x,y) I_k(x,y))^2/(\overline{\varepsilon}(x,y)\right)^2} \tag{10}$$

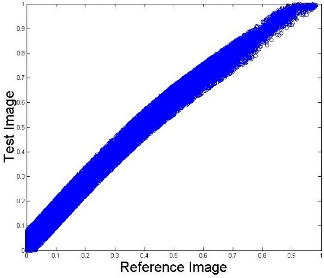
Here $\overline{\varepsilon}(x,y) = \sum_{k=1}^{n} \varepsilon_k(x,y)$ and $\varepsilon(x,y) = e^{\frac{-(I_k(x,y)-0.5)^2}{2\times 0.2^2}}$

- Generate a threshold limit, $(.25) \times \max(V)$.
- Pixel location having, more intensity variation then this limit are considered as dynamic.

# Generating Comparametric Function



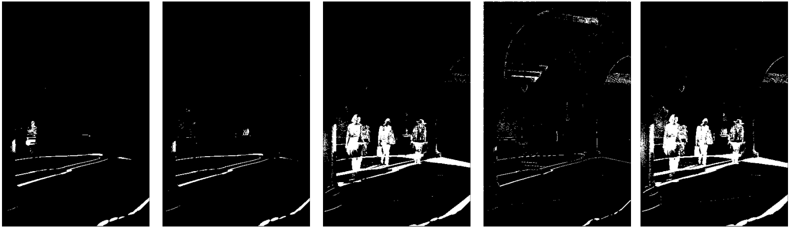

(f)




(g)

Figure : 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.

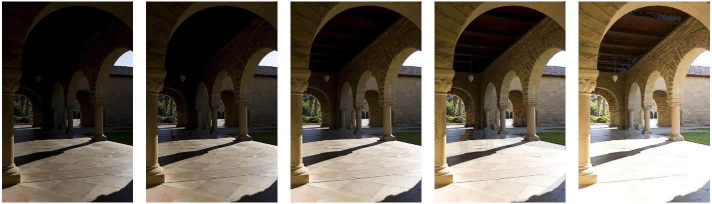
# Detected Dynamic regions



(a) (b)

Figure : First image of the exposure stack is chosen as the reference image, (b) Combined dynamic regions. These regions are responsible for introducing ghosting artifacts in the final high contrast LDR image.

# Generated Static Image Set

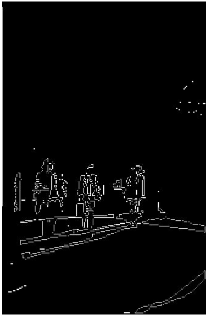


(b) Fused image

# Aggregate Energy Matrix for Dynamic Scenes

- We have generated the static image set corresponding to a dynamic scenes,
- Notice that still there are regions in static image set which are dynamic but not detected by the algorithm,
- This is because these regions are very similar to the background and show a very little intensity variation with respect to the background,
- 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.

# Silhouette



(a)

Figure : Silhouettes found in the generated high contrast LDR image of an image corresponding to a dynamic scene

- Due to these silhouettes, one can not directly apply Aggregate energy matrix (weighting parameter weighted Laplacian) as defined in case of static scenes for retargeting,
- 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.

# Modified Aggregate Energy Matrix, for dynamic scenes

$$\psi(x,y) = \begin{cases} \chi(x,y) + \lambda S(x,y), & \text{while increasing aspect ratio} \\ \chi(x,y) - \lambda S(x,y), & \text{while reducing aspect ratio} \end{cases} \tag{11}$$

where $S$ represents the silhouette matrix.

- where $\lambda$ is a sufficiently large number.
- Because in the case of reducing the aspect ratio, pixels creating silhouettes will get a very less weight in $\psi$.
- On the other hand while increasing the aspect ratio, such pixels will get very high weight in $\psi$ .

# Retargeting result, of dynamic scenes

*Downsizing, vertically, by 150 seams*

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

# Retargeting result, of dynamic scenes

*Enlarging, vertically, by 150 seams*

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

Figure : Retargeting, Dynamic HDR scenes

# *Results*

*Analytically...*

# Results

*Measures*

- If the input images are of dimension $M \times N$, and retargeted dimensions are $(M-p) \times (N-q)$.
- *Energy per pixel*

$$= \sum_{x=1}^{M-p} \sum_{y=1}^{N-q} \chi(x, y)$$

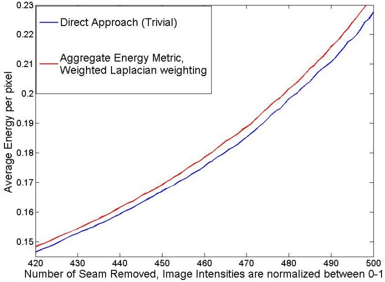
- *Image Entropy*

$$= -\sum_{i=0}^{255} p_i log_2 p_i$$

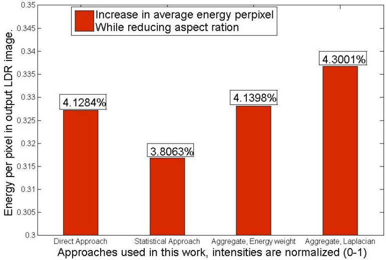
$$\text{where, } p_i = \frac{\text{Number of pixels having intensity } i}{\text{Total number of pixels}}$$

# Results

*Change in Energy per pixel in final high contrast LDR image*




(a) Change in Energy per pixel, corresponding to the removal of seams




(b) Percentage increase in Energy per pixel, through various Proposed Approaches

Figure : Results

# Results

*Analytical Comparison between various Approaches used*

| 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 : 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.

# *Discussion*

*What's More?*

# Enhancement Over Exposure Fusion

*Energy per pixel enhancement*


(a) Exposure fusion (b) Our method (c) Comparison

Figure : Enhance exposure fusion

- First Removal of low energy seams,
- Add new seams,
- Energy per pixel increment
- Limitation: Output image will no longer be the same.

## Better Way to Enlarge Images

(a)

(b)

(c)

(d)

- (a): Add all seams at once, $N$
- (b): Linear decreasing manner, $n + (n-1) + \ldots + 1 \leq N$
- (c): Quadratically decreasing, $n^2 + (n-1)^2 + \ldots + 1 \leq N$
- (d): Add a fix number once and loop through, $nk = N$

# Better Way to Enlarge Images

(e) Exponential

Figure : Better Way to Enlarge Images

- Add seams in a exponential decreasing manner.
- If one wants to add $N$ seams. in first stage insert $\frac{N}{2^k}$ seams and reconstruct the energy function.
- The total number of seams inserted are $\frac{N}{2} + \frac{N}{4} + \frac{N}{8} + \ldots = N$.
- for $2^k \leq N \leq 2^{k+1}$, need utmost $k+1$ distinct insertion steps.

# New measures for image energy

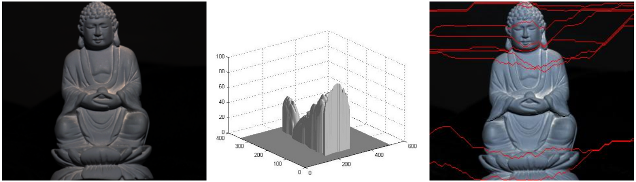
*Images are more than pixel values*

Figure : Focus map of an image

- Energy metric discussed so far are intensity based.
- Define energy measure $F$, indicates focused regions,
- if $F(x, y) = 1$, pixel $(x, y)$ is at focus plan,
- As $F$ decreases, pixels became defocused,
- Removing or adding defocused pixels, make sure that object of interest minimally affected during retargeting.

# New measures for image energy

*Depth map*




(a) Test Image (b) Depth map (c) seam found

- $Z(x, y)$ Scene Depth at location $(x, y)$,
- Using $Z$ we can define object of interest, on depth criteria.

[5]

[5] R. J. Woodham, "Photometric method for determining surface orientation from multiple images," 1980

# Conclusion And Future Pointers

- We have proposed novel approach for content aware resizing of HDR scene,
- We showed that the proposed algorithm performs better when compared to the direct approach,
- The proposed approach is fully automatic with no user intervention,
- We hope to extend this approach for video image retargeting applications involving HDR scenes,
- We believe that novel approach discussed here would lead to more novel ideas in image retargeting research.

# Publications

1. IEEE journal "IEEE Transactions on Image Processing" as a Regular Paper. TIP-10611-2013, entitled "Efficient Image Retargeting for High Dynamic Range Scenes". The work is currently in review process.

arxiv.org, Cornell University Library (online library for scientific papers), 20 May 2013.

# Key References


- T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion."
- S. Avidan and A. Shamir, "Seam carving for content-aware image resizing," in *ACM Transactions on graphics (TOG)*, 2007
- P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," 1997
- A. El Gamal, "High dynamic range image sensors," 2002
- R. J. Woodham, "Photometric method for determining surface orientation from multiple images," 1980

# Key References


- S. Mann and R. W. Picard, "On being undigital with digital cameras: Extending dynamic range by combining differently exposed pictures," 1995
- M. Granados, B. Ajdin, M. Wand, C. Theobalt, H.-P. Seidel, and H. P. Lensch, "Optimal hdr reconstruction with linear digital cameras," 2010
- S. W. Hasinoff, F. Durand, and W. T. Freeman, "Noise-optimal capture for high dynamic range photography," 2010
- S. Raman and S. Chaudhuri, "A matte-less, variational approach to automatic scene compositing," 2007
- ——, "Reconstruction of high contrast images for dynamic scenes," *The Visual Computer*, 2011

# Key References


- T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion," 2007
- S. Park, M. Park, and M. Kang, "Super-resolution image reconstruction: a technical overview," 2003
- S. Baker and T. Kanade, "Limits on super-resolution and how to break them," 2002
- S. K. Nayar and Y. Nakagawa, "Shape from focus," 1994

# *Thank You!*

*Questions?*

*Acknowledgement*