

UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS
TRABALHO DE CONCLUSÃO DE CURSO DE
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

Métodos de geração de imagens HDR

Claudio Mota Oliveira

SÃO CRISTÓVÃO/ SE

2015

UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS
TRABALHO DE CONCLUSÃO DE CURSO DE
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

Claudio Mota Oliveira

Métodos de geração de imagens HDR

Proposta de dissertação parcial apresentada ao Departamento de Computação (DCOMP) da Universidade Federal de Sergipe (UFS) como parte de requisito para obtenção do título de Bacharel em Ciência da Computação.

Orientador: Prof.^a Dr.^a Beatriz Trinchão Andrade

SÃO CRISTÓVÃO/ SE

2015

Resumo

Este documento mostra a pesquisa e implementação de métodos de geração de imagens HDR's a partir de um conjunto de imagens LDR's.

Abstract

This document describes the research and implementation of methods for generating HDR images using LDR images as input.

Lista de Figuras

Lista de Tabelas

2.1	Progressão dos eixos	5
-----	--------------------------------	---

Sumário

1	Introdução	1
2	Conceitos básicos	3
2.1	Faixa dinâmica	3
2.2	Função resposta de uma câmera	4
2.3	Eixo logarítmico	5
2.4	Minimização de Energia	6
2.5	Similaridade Bidirecional	6
2.6	Relaxação de Gauss-Seidel	6
2.7	Spline cúbica	6
3	Revisão bibliográfica	7
4	Métodos HDR	9
4.1	Método Mann e Picard	9
4.1.1	Algoritmo para inferir a função resposta	10
4.1.2	Gerando a imagem HDR	11
4.1.3	Resultados e discussões	11
4.2	Método Robertson	12
4.2.1	Gerando a imagem HDR	12
4.2.2	Algoritmo para inferir a função resposta	14
4.2.3	Resultados e discussões	15
4.3	Método Sen	16
4.3.1	Modelagem do problema	17
4.3.2	Especificidades de implementação	19

4.3.3	Resultados e discussões	19
5	Discussões	21
5.1	Método Mann	21
5.1.1	Vantagens	21
5.1.2	Desvantagens	21
5.2	Método Robertson	22
5.2.1	Vantagens	22
5.2.2	Desvantagens	22
5.3	Método Sen	22
5.3.1	Vantagens	22
5.3.2	Desvantagens	22
6	Considerações finais	23
	Referências	23

Capítulo 1

Introdução

A visão computacional é uma área que possui como objetivo extrair informações significativas a partir de imagens, e assim possibilitar o processamento dessas informações como objetos para fins específicos como o reconhecimento de um ambiente, detecção de movimento entre outros [4]. Um dos principais desafios desta área é a reconstrução de objetos 3D a partir de imagens, desafio este que é de grande utilidade, como ferramenta, para outras diversas áreas do conhecimento como mostrado nos trabalhos de Matsushigue et al. [8], que abordou o uso de reconstrução de tomografias 3D para melhoria de diagnostico de fraturas na região do úmero, e como o trabalho de Beatriz et al. [3], que usou métodos de reconstrução 3D para a preservação digital de obras de arte.

Outro exemplo de aplicação da reconstrução 3D a partir de imagens foi apresentada no método proposto por Argawal et al. [1]. O método proposto aborda a reconstrução 3D sobre a escala de cidades utilizando-se de fotos disponíveis na Internet, que por sua vez, foram capturadas por pessoas aleatórias. Para isto foi usada como objeto de estudo a cidade de Roma que, por se tratar de uma cidade turística, possui milhares fotos disponíveis na Internet, e como resultado foi feita a reconstrução 3D do monumento do Coliseu de Roma.

De maneira geral, grande parte das técnicas de reconstrução 3D funcionam com base no mapeamento de pontos singulares existentes em diferentes imagens, denominados pontos de interesse, que a partir da diferença de posicionamento dos pontos em cada imagem, estima-se a posição dos mesmos no espaço tridimensional, gerando assim uma nuvem de pontos 3D.

Após a realização de um conjunto de pesquisas pela internet em busca de métodos para obtenção de nuvem de pontos 3D a partir de imagens, evidenciou-se que grande parte dos

métodos existentes utiliza-se de imagens LDR (Low Dynamic Range), que possuem uma baixa faixa dinâmica e assim possuindo uma resolução bastante limitada de cores. Sob o ponto de vista de acessibilidade, o uso de imagens LDR é explicado pelo fato das câmeras com alta faixa dinâmica serem muito caras e assim pouco acessíveis.

Trabalhos publicados na literatura, propoem métodos para união de um conjunto imagens LDR em uma imagem HDR (High Dynamic Range), que por sua vez possui melhor representação da cena em relação a uma imagem LDR.

Assim como no trabalho feito por Kontogianni et al. [6], o trabalho descrito neste documento possui como hipótese a ideia que um conjunto de imagens HDR têm o potencial de identificar mais pontos de interesse que um conjunto de imagens LDR, tendo em vista que o primeiro possui mais informação do ambiente que o segundo. Neste contexto, este trabalho visa verificar a viabilidade de implementação de um método para obtenção de nuvem de pontos a partir de imagens HDR esperando que com isso seja possível extrair mais pontos de interesse e reconstruir nuvens de pontos com maior resolução de maneira acessível. Para que este processo possa ser realizado com um baixo custo em relação ao equipamento utilizado se faz necessário o uso das técnicas de união de imagens LDR para geração de imagens HDR.

Sendo assim, o trabalho foi dividido em duas fases:

- Geração de imagem HDR:

Nesta fase métodos de geração de imagens HDR são descritos implementados e comparados para que sejam determinados os métodos que serão utilizados na próxima fase.

- Geração de nuvem de pontos:

Nesta fase será abordada a implementação e verificação de métodos para obtenção de nuvem de pontos 3D extraída de imagens HDR.

Capítulo 2

Conceitos básicos

Os métodos de geração de imagens HDR abordados por este trabalho possuem um conjunto de conceitos que são fundamentais para o entendimento e reprodução das implementações dos mesmos. Neste capítulo abordaremos superficialmente estes conceitos para familiarizar o leitor ao assunto.

2.1 Faixa dinâmica

Cada imagem digital possui uma certa quantidade de bits para representação das cores de cada um de seus pixels. A maior parte das câmeras convencionais do mercado geram imagens com pixels de 8 bits para cada canal de cor, o que resulta numa faixa de representação de 256 valores possíveis. Quando uma imagem possui esta faixa de representação de cores esta é chamada de imagem LDR (Low Dynamic Range) por possuir uma faixa relativamente curta de cores para representar a cena. Quando uma imagem possui uma faixa maior de representação de cores esta é chamada imagens HDR (High Dynamic Range) estas geralmente representam as cores em ponto flutuante o que diminui a deterioração cumulativa ao efetuar operações sobre a imagem sucessivamente, e conseguem representar melhor as cenas capturadas em relação a uma imagem LDR por possuir mais bits para representação das cores.

Supondo uma cena com áreas bem iluminadas, outras com iluminação moderada e outras com baixa iluminação, ao captar uma imagem LDR desta cena muito provavelmente informações das áreas com extremos (pouco/muito iluminadas) serão perdidas pois a imagem não

terá faixa dinâmica suficiente para representar todos os níveis de iluminação da cena. Por outra via ao captar uma imagem HDR desta mesma cena mais detalhes das zonas claras e escuras serão registrados sem perdas significativas.

2.2 Função resposta de uma câmera

Uma câmera digital é composta de uma matriz de elementos sensores de luz, no processo de captura de uma imagem, cada sensor pode ser considerado uma função f que recebe como entradas um valor de irradiação de luz x por um determinado periodo de tempo t , e retorna como saída um valor geralmente inteiro p , chamado de pixel. f é chamada de função resposta da câmera, numa câmera ideal ela pode ser descrita pela seguinte equação:

$$f_{(x,t)} = p \quad (2.1)$$

Porém, por conta de ruídos advindos de diversas fontes essa equação é modificada para a seguinte:

$$f(x, t, R_a) = p + R_b \quad (2.2)$$

Onde

- R_a é a abstração dos ruídos advindos do meio, como ruído termico entre outros.
- R_b é a abstração dos ruídos gerados na conversão de analógico para digital, como corrente de escuro entre outros.

Vale atentar que o valor de iluminação x também é influenciado pela abertura do obturador, mas este trabalho considera que a abertura do obturador irá permanecer constante, pois qualquer mudança da abertura do obturador influenciará no foco da câmera.

Em geral, os algoritmos de geração de imagens HDR se utilizam da inversa da função resposta da câmera para compor uma função g de forma que, dado um pixel p e o tempo de exposição t , seja possível estimar o valor de iluminação x^* que o gerou este pixel.

Tabela 2.1: Progressão dos eixos

Graduação	1º	2º	3º	4º
Valor (linear)	0	1	2	3
Valor (logarítmico)	1	10	100	1000

$$g(p, t) = x^* \quad (2.3)$$

Um dos problemas em relação a esta abordagem é que a confiabilidade dos pixels varia de acordo com o tempo de exposição, além disso, cada pixel p possui um ruído agregado a ele assim como erros de registro entre outros, sendo assim, cada método de geração de imagens HDR trata estes problemas da sua maneira.

Outro problema é que na prática o usuário da câmera geralmente não possui nenhuma informação sobre a função resposta da câmera que está utilizando, sendo assim alguns dos autores dos métodos de geração de imagens HDR, também propõem métodos de inferência da função resposta da câmera a partir das imagens que são passadas como entrada.

2.3 Eixo logarítmico

Um conceito bastante abordado nos métodos de geração de imagem HDR, é o uso do eixo logarítmico. Alguns métodos utilizam o este conceito para ressaltar a confiabilidade que um valor de pixel representa, levando o mesmo a ter um maior ou menor peso na geração da imagem HDR.

O eixo logarítmico é um eixo no qual cada graduação possui uma razão de 10 unidades em relação a graduação anterior. A tabela 2.1 ilustra a variação dos valores nos eixos linear e logarítmico.

As figuras a seguir mostram a mesma função resposta representada em um eixo linear e logarítmico respectivamente.

/a /a /a /a

O cálculo da derivada do ponto de uma função em relação ao eixo logarítmico é dado pela seguinte equação:

$$y'_{\log}(x) = \frac{dy}{dx} \cdot \frac{x}{\log(e)} \quad (2.4)$$

2.4 Minimização de Energia

2.5 Similaridade Bidirecional

2.6 Relaxação de Gauss-Seidel

Consiste num método para resolução de sistemas de equações com múltiplas variáveis simultaneamente de forma iterativa. O método consiste em, a partir de valores especulados iniciais, inferir o valor de uma variável e com este valor encontrado, encontrar o valor da próxima variável, e assim prosseguir no metodo iterativo para todas as variáveis repetitivamente até atingir uma condição de convergência.

2.7 Spline cúbica

Spline cúbica é um método de interpolação que possui como principal característica a geração de uma função contínua, com primeira e segunda derivadas também contínuas. A interpolação consiste em, dado um conjunto de pontos $S = \{(x_i, y_i)\}$, obter uma função $y(x)$ que contenha todos estes pontos.

Capítulo 3

Revisão bibliográfica

Atualmente existem diversos métodos para geração de imagens HDR a partir de imagens LDR na literatura. Neste capítulo é feita uma rápida revisão de alguns dos trabalhos que abordam este assunto, assim como a discriminação dos métodos que foram trabalhados neste documento.

Ao que tudo indica, o primeiro método de geração de imagens HDR a partir de um conjunto de imagens LDR foi proposto por Mann e Picard [7], onde dado um conjunto de imagens estáticas uma em relação a outra, diferindo apenas quanto ao tempo de exposição, é possível estimar o valor de luminosidade que gerou o pixel das imagens com base na operação inversa da função resposta da câmera. Os autores propõem um método para estimar a função resposta da câmera a partir de um conjunto de imagens LDR. Uma vez com a função resposta, o valor de luminosidade é estimado por meio de uma média ponderada com base na derivada da função resposta em relação à exposição em escala logarítmica.

Outro método foi proposto por Mitsunaga e Nayar [9], que trata a questão da exposição tanto pela ótica do tempo, quanto pela variação do diâmetro de abertura do obturador da câmera. Os autores também propõem um método para estimar a função resposta da câmera, que leva em conta o ruído presente nas imagens. Para a geração da imagem HDR, o método assume como peso para os pixels a razão de sinal-ruído presente nas imagens LDR.

Robertson et al. [11], propuseram um método estatístico para inferência tanto da função resposta quanto dos valores de luminosidade, de forma que, os pixels que possuem valores próximos dos extremos da faixa dinâmica possuem menor peso que pixels mais centralizados. Os autores justificam o método pelo pressuposto que os pixels com valores próximos

aos extremos são mais suscetíveis a ruído.

Ali e Mann [2], propuseram uma metodologia para reconstrução de imagens HDR utilizando a função resposta comparamétrica da câmera, a qual os autores definem como sendo uma matriz bidimensional que, dado duas imagens, mapeia qualquer combinação de valores de pixels entre elas, para um pixel que possua informações significativas de ambas imagens. A geração da imagem HDR é feita por meio da combinação sucessiva de pares de imagens. Os autores justificam o método pois, segundo eles, apresenta um resultado mais rápido que os métodos convencionais podendo ser utilizado em aplicações em tempo real como filmagens.

Granados et al. [5], propôs um método de geração que visa encontrar os pesos ótimos para cada pixel na geração de imagens HDR, para isso utiliza-se de uma modelagem da câmera que trata vários tipos de ruídos, tanto espaciais como temporais.

Sen et al. [12] propôs um método iterativo que busca robustez em relação ao movimento da câmera, baseado em técnicas de minimização de energia, que com base em todas as outras imagens infere a luminosidade de uma imagem específica, escolhida como referência no conjunto de imagens LDR. Dessa maneira ao final do método são geradas imagens HDR's em diferentes tempos de exposição, que por sua vez tem aparência da imagem escolhida como referência.

Para o trabalho descrito neste documento, serão abordados e comparados os métodos de geração de imagens HDR dos seguintes autores:

- Mann e Picard: Por se tratar de um método simples que aborda apenas a mudança no tempo de exposição das imagens LDR's para geração de imagens HDR, e ao mesmo tempo suficiente para nosso objetivo final.
- Robertson et al.: Por se tratar de um método mais elaborado que além de abordar da mudança no tempo de exposição, leva em consideração aspectos de erro relativos a ruídos na aquisição das imagens LDR.
- Sen et al.: Por se tratar de um método robusto quanto ao movimento da câmera, característica essa bastante importante para aplicações não profissionais, pois manter a câmera totalmente estática entre a captura das imagens é bastante trabalhoso.

Capítulo 4

Métodos HDR

Neste capítulo serão abordados os métodos de geração de imagens HDR a partir de imagens LDR escolhidos para o trabalho. Cada seção tratará um método, mostrando aspectos teóricos, implementação e resultados obtidos assim como discussões sobre tópicos de interesse.

4.1 Método Mann e Picard

O método proposto por Mann e Picard [7] parte da problemática que para ambientes com áreas bem iluminadas e outras pouco iluminadas, as câmeras convencionais acabam perdendo informações por não possuírem faixa dinâmica suficiente para representação de todos os valores de iluminação ali presentes. A solução de alguns fotógrafos para este problema, é regular manualmente o tempo de exposição de captura da cena de forma que a perda seja a menor possível, ou valorizando apenas áreas de seu interesse. A solução proposta pelos autores parte do pressuposto que ao capturar várias imagens, estáticas uma em relação a outra, variando apenas o tempo de exposição, cada imagem possuirá informações relevantes de determinadas áreas da cena, e ao unir essas imagens obtém-se uma imagem com todas as informações de iluminação deste ambiente.

Para que se possa fazer a união das imagens é necessário saber o valor de iluminação que o pixel representado na imagem mapeia, para isso é necessário fazer o cálculo da inversa da função resposta da câmera, que por sua vez é diferente para cada modelo de câmera no mercado, e muitas vezes é não-linear. Sendo assim os autores propuseram um algoritmo para inferir a função resposta da câmera e assim calcular a inversa da mesma.

4.1.1 Algoritmo para inferir a função resposta

Para o algoritmo são necessárias duas imagens a e b que foram capturadas em diferentes tempos de exposição, assume-se que elas são estáticas uma em relação a outra, considerando k a razão do tempo de exposição de b em relação ao de a ($k > 1$), o objetivo é encontrar a função resposta f da câmera que capturou as imagens seguindo os seguintes passos:

- É escolhido um pixel relativamente escuro na imagem a posicionado em (x_0, y_0) , denotado por $a(x_0, y_0)$. Sabe-se que um valor de irradiação de luz q_0 gerou o valor desse pixel a partir da aplicação da função resposta $f(q_0) = a(x_0, y_0)$.
- Localiza-se o pixel na mesma posição na imagem b , como as imagens são estáticas uma em relação a outra, sabe-se que $b(x_0, y_0)$ possui o mesmo valor de irradiação de luz que $a(x_0, y_0)$, porém com um tempo de exposição de k vezes maior. Sendo assim volta-se a imagem a e procura-se um pixel com o mesmo valor de $b(x_0, y_0)$, denominado $a(x_1, y_1) = f(kq_0)$.
- Localiza o pixel $b(x_1, y_1)$ na imagem b , e procura-se na imagem a um pixel com o valor $b(x_1, y_1)$ denominado $a(x_2, y_2) = f(k^2q_0)$
- Repetindo este passo sucessivamente acha-se a não linearidade de $f(k^n)$

Uma vez com os pontos encontrados elimina-se o termo q_0 , uma vez que não importa a unidade do eixo de irradiação, levando aos pontos mostrados na seguinte figura:

a

a

a

a

a

O método assume então que a função resposta possui a forma:

$$f(q) = \alpha + \beta q^\gamma \quad (4.1)$$

Onde α pode ser considerado zero ao diminuir os valores dos pixels pelo valor obtido em uma imagem com tampa da lente colocada. O β é um valor de escala e é arbitrário. E o γ pode ser inferido através da equação:

$$b(x_i, y_i) = k^\gamma + a(x_i, y_i) \quad (4.2)$$

4.1.2 Gerando a imagem HDR

O método proposto pelos autores para geração de uma imagem HDR se baseia-se na ideia que a derivada da função resposta da câmera em relação ao eixo logarítmico é proporcional a confiabilidade que o pixel apresenta. Sendo assim para cada pixel da imagem, é feita uma média ponderada da confiabilidade que o pixel apresenta em cada tempo de exposição, possuindo maior peso aquele que possuir maior confiabilidade e assim gerando um pixel HDR que foi formado a partir de informações de todas as exposições. Dado um conjunto com N imagens alinhadas uma em relação a outra, variando apenas no tempo de exposição $t_i, i = 1 \dots N$, onde cada imagem possui M pixels representados por $I_{ij}, j = 1 \dots M$, cada um com peso p_{ij} . A equação para inferir o valor de irradiação do pixel x_j é dada por:

$$x_j = \frac{\sum p_{ij} \frac{f^{-1}(I_{ij})}{t_i}}{\sum p_{ij}} \quad (4.3)$$

4.1.3 Resultados e discussões

O método em questão foi implementado na linguagem de programação C++, e foi utilizado um conjunto de imagens para verificação dos resultados. Verificou-se que pelo fato do beta da equação 4.1 ser arbitrário dependendo do valor atribuído ao mesmo, pode-se haver uma incompatibilidade entre as funções respostas dos diferentes canais de cores da imagem, isso é mostrado na figura:

a
a
a
a
a

Para solução deste problema foi usada a mesma solução proposta por Robertson et al. [11], que estabelece que a função resposta deve sempre mapear o pixel com valor

intermediário (128 para pixels com 8 bits) para o valor de uma unidade de luminosidade irradiada. Com essa modificação foi possível obter resultados satisfatórios na união dos canais HDR como mostrados nas figuras

a

a

a

a

a

E as figuras ... mostram as funções resposta dos diferentes canais obtidas utilizando o método proposto.

4.2 Método Robertson

O método proposto por Robertson et al. [11], aborda a geração de imagens HDR de maneira similar ao método proposto por Mann descrito na seção 4.1, pois utiliza a derivada da função resposta em um eixo logaritmico como peso e assume que as imagens são estáticas entre si, variando apenas o tempo de exposição. As principais diferenças entre estes métodos são a postura adotada em relação ao tratamento do ruído, e a forma como a função resposta da câmera é inferida, que neste caso é feita por meio de um método iterativo que supõe uma função inicial padrão, e então, por meio de uma forma de relaxação de Gauss-Seidel, aproxima a função resposta de forma a obter um erro abaixo de um limiar pre-estabelecido.

4.2.1 Gerando a imagem HDR

Ao tentar extrair o valor de irradiação de luz que um pixel mapeia, aplica-se a inversa da função resposta. Porém no momento de registro da imagem vários fatores influenciaram para que o valor registrado fosse ligeiramente diferente do valor ideal, isso se dá por conta da presença do ruído no ambiente, que pode ser provindo de diversas fontes diferentes, como corrente de escuro, conversão de analógico para digital entre outros. Sendo assim a equação da inversa da função resposta da câmera pode ser descrita da seguinte maneira:

$$f_{(p_{ij})}^{-1} = I_{p_{ij}} = xt_i + N_{ij} \quad (4.4)$$

Onde

- N_{ij} encapsula os ruídos tiveram influência na captura do pixel.
- x é o valor de irradiação de luz real que deseja-se encontrar.
- t_i é o tempo de exposição ao qual o sensor da câmera foi submetido para captura do pixel.
- p_{ij} é o valor do pixel da imagem com exposição t_i .

Como o ruído é uma característica singular para cada ambiente/câmera utilizada, modelar o problema de forma a caracterizar cada ruído é uma tarefa bastante complicada. Sendo assim, este método trata o ruído sob o ponto de vista de uma distribuição gaussiana com média zero para simplificar a modelagem devido a diversidade de fontes de ruído.

Com base em técnicas maximização da probabilidade para recuperar o valor de iluminação real com base no valor do píxel, obtem-se a seguinte equação:

$$O(x) = \sum_{i,j} w_{ij} (I_{p_{ij}} - x t_i)^2 \quad (4.5)$$

Que representa o erro quadrático entre o valor inferido e o valor obtido pela inversa. Fazendo o gradiente de $O(x)$ tender a zero os autores do método chegaram a seguinte equação para inferência do valor de irradiação de luz:

$$x_j^* = \frac{\sum_i w_{ij} t_i I_{p_{ij}}}{\sum_i w_{ij} t_i^2} \quad (4.6)$$

Onde

- x_j^* é o valor de irradiação de luz inferido.
- O peso w_{ij} , assim como no método citado na seção 4.1.2, é obtido pela derivada da função resposta em relação um eixo logarítmico.

O conjunto de valores x_j^* , gerados a partir da convolução das formam a imagem HDR final.

4.2.2 Algoritmo para inferir a função resposta

Na maioria dos casos, não se sabe a função resposta de uma câmera e por isso ela deve ser estimada [11]. Para a estimativa da função resposta é necessário a minimização da função obtida de erro quadrático dado pela seguinte equação:

$$O(I, x) = \sum_{i,j} w_{ij} (I_{p_{ij}} - x_{ij} t_i)^2 \quad (4.7)$$

Que é parecida com a equação 4.5, porém com uma variável I de entrada que também será inferida. Como não se sabe o valor de irradiação x_{ij} nem a função resposta para calcular o valor de $I_{p_{ij}}$, utiliza-se o método de Gauss-Seidel para inferir ambas as variáveis simultaneamente. Também é necessário estabelecer uma função para mapear os pesos uma vez que, como não se sabe a função resposta da câmera, não há como obter a sua derivada. Os autores propoem a seguinte função[10] para definir os pesos:

$$e^{-4 \frac{(p_{ij} - 127.5)^2}{127.5^2}} \quad (4.8)$$

O método de inferência da função resposta possui um passo inicial e um conjunto de passos iterativos como segue:

- Passo inicial:

Estabelece-se a função resposta inicial, como sendo uma função linear com o valor $I_{128}^{*(0)} = 1.0$, e então o valor de $x^{*(0)}$ é obtido aplicando a equação 4.6.

- Passos iterativos:

Para cada iteração l são feitos os seguintes passos:

1. Considerando E_m o conjunto de pixels que possuem o valor igual a m , faz-se o relaxamento de $I^{*(l)}$ aplicando a seguinte equação para inferí-lo:

$$I^{*(l)} = \frac{\sum_{(i,j) \in E_m} w_{ij} t_i x_j^{*(l-1)}}{\sum_{(i,j) \in E_m} w_{ij}} \quad (4.9)$$

2. Normaliza-se $I_{p_{ij}}$ de forma que I_{128} seja uma unidade.

3. E então é feito o relaxamento de $x^{*(l)}$ com base na seguinte equação:

$$x_j^{*(l)} = \frac{\sum_i w_{ij} t_i I_{p_{ij}}^{*(l)}}{\sum_i w_{ij} t_i^2} \quad (4.10)$$

Dessa forma, a função resposta e a imagem HDR vão sendo inferidas simultaneamente, e quando as modificações feitas no passo iterativo da equação 4.9 forem menores que um limiar predefinido o passo iterativo é terminado.

Ao final do método, possui-se um conjunto de valores $\{I_{p_{ij}}\}$ que representam informações da função resposta da câmera. Como este conjunto é discreto, há a necessidade de efetuar uma interpolação utilizando a técnica da Spline cúbica 2.7 para obtenção da função resposta final como ilustrado na figura ??.

a

a

a

a

a

4.2.3 Resultados e discussões

O método em questão foi implementado na linguagem de programação C++, e foi utilizado um conjunto de imagens para verificação dos resultados. Os resultados são mostrados nas figuras:

a

a

a

a

a

Notou-se que áreas com exposição muito elevada apresentaram ruído mais perceptível, como mostrado na figura.

a

a

a

a

a

E as figuras ... mostram as funções resposta dos diferentes canais obtidas utilizando o método proposto.

a

a

a

a

a

4.3 Método Sen

O método proposto por Sen et al. [12], aborda a geração de imagens HDR sob uma ótica bastante diferenciada dos outros métodos citados anteriormente, apesar de partir do mesmo pressuposto da união de várias imagens LDR em diferentes tempos de exposição para gerar uma imagem HDR. Um dos princípios básicos deste método é a robustez quanto ao movimento dos elementos das imagens, característica essa que nos outros métodos era assumida como sendo estática.

No processo de captura de uma sequência de várias imagens, assumir que uma imagem e os elementos contidos na mesma estarão estáticos em relação as outras imagens só é possível para ambientes bastante controlados, de forma que os objetos da cena não sejam móveis, e ainda há a necessidade de equipamentos mais sofisticados como tripé da câmera para mantê-la fixa e/ou uso de câmeras controladas por software a distância. Estes casos são bastante específicos e não correspondem a realidade de grande parte dos usuários de câmeras digitais convencionais, muitos destes possuem como única alternativa, a captura das imagens sem tripé, sem software e as cenas capturadas por sua vez possuem uma dinâmica bem diferente de um ambiente estático.

O método em questão se utiliza do princípio de minimização de energia para inferir,

a partir de uma imagem de referência e de varias imagens LDR, como seria o registro da imagem de referência se esta fosse capturada nos tempos de exposição das outras imagens. Por outra via o algoritmo simultaneamente infere a imagem HDR que estas imagens inferidas geram quando combinadas, formando assim um ciclo iterativo.

Ao final, além da imagem HDR, o método gera um conjunto de imagens LDR que conservam os detalhes da imagem de referência mas com diferentes níveis de exposição, ou seja, gera imagens estáticas entre si em diferentes tempos de exposição, o que possibilita o uso dos métodos citados nas seções 4.1 e 4.2 para a geração da imagem HDR.

4.3.1 Modelagem do problema

Dado um conjunto de N imagens LDR (L_1, L_2, \dots, L_N) , onde uma destas é a imagem de referência L_{ref} para que o problema possa ser trabalhado como minimização de energia as seguintes propriedades da imagem HDR H , que será inferida, devem ser satisfeitas:

- Ao mapear um valor de irradiação de H para o tempo de exposição da imagem de referência, o valor mapeado deve ser bem próximo ao valor do pixel da imagem de referência.
- Ao mapear um pixel, propriamente exposto, da imagem de referência para o domínio de irradiação linear, este deve possuir um valor bem próximo do valor de irradiação de H .
- A imagem de referência deve ser utilizada na composição da imagem HDR quando essa estiver propriamente exposta, caso contrário as outras imagens HDR devem assumir participação na inferência do valor de irradiação com base na mútipla similaridade bidirecional.

Sendo assim a equação da energia é dada pela junção destes fatores como segue:

$$E(H) = \sum [\alpha_{ref(p)} \cdot (h(L_{ref})_{(p)} - H_{(p)})^2 + (1 - \alpha_{ref(p)}) \cdot E_{MBDS}(H|L_1, \dots, L_N)] \quad (4.11)$$

Onde

- $E(H)$ é a energia, este valor deve ser diminuído ao máximo para maior confiabilidade da imagem HDR H .
- $\alpha_{ref(p)}$ é uma função de peso que indica o qual bem exposto um pixel da imagem de referência está.
- $h(L)$ mapeia uma imagem L para o domínio de irradiação linear.

Considerando que a equação de energia, receberá como entrada um conjunto de imagens LDR $\{I_k\}, k = 1..N$, onde cada I_k é o mapeamento da imagem HDR H para o tempo de exposição k , os autores do método chegaram a seguinte equação de energia:

$$\begin{aligned}
 E(H, I_1, \dots, I_N) = & \sum_{p \in pixels} [\alpha_{ref(p)} \cdot (h(L_{ref})_{(p)} - H_{(p)})^2 + \\
 & (1 - \alpha_{ref(p)}) \cdot \sum_{k=1}^N MBDS(I_k | g^k(L_1), \dots, g^k(L_N)) + \\
 & (1 - \alpha_{ref(p)}) \cdot \sum_{k=1}^N \Lambda(I_k)_{(p)} \cdot (h(I_k)_{(p)} - H_{(p)})^2]
 \end{aligned} \tag{4.12}$$

Onde

- Λ é uma função de peso que indica a relevância do pixel para a imagem HDR.
- $g^k(L_i)$ mapeia a imagem LDR L com exposição i para uma imagem LDR equivalente com exposição k .

Sendo assim, a equação pode ser solucionada através de um método iterativo que infere simultaneamente os H e os $\{I_k\}$ seguindo os seguintes passos:

- Minimiza-se a equação para I_1, \dots, I_N , usando primeiramente o search-and-vote para resolver o MBDS, e então para cada I_k combina com o mapeamento de H anterior para a exposição k , para que eles possam gradualmente mais informações da imagem de referência.
- Minimiza-se então a equação quanto ao valor H , primeiramente mapea-se a irradiação linear de todos os $\{I_k\}$ e faz a combinação dos mesmos para obter \hat{H}^* como mostra a equação 4.13, após isso o valor de \hat{H}^* é combinado com a imagem de referência

mapeada para o domínio de irradiação linear para obter o novo valor da imagem HDR H como mostrado na equação 4.14.

$$H_{(p)}^* = \frac{\sum_{k=1}^N \Lambda(I_{k(p)}) h(I_{k(p)})}{\sum_{k=1}^N \Lambda(I_{k(p)})} \quad (4.13)$$

$$H_{(p)} = \alpha_{ref(p)} \cdot h(L_{ref})_{(p)} + (1 - \alpha_{ref(p)}) \cdot H_{(p)}^* \quad (4.14)$$

Esses passos são repetidos até atingir um critério de convergência.

4.3.2 Especificidades de implementação

Para implementação do método foi utilizado a própria implementação disponibilizada pelos autores. Para o correto funcionamento deste há a necessidade de efetuar um pré-processamento nas imagens para formatá-las no domínio linear. Para isso é necessário o conhecimento da função resposta da câmera, pois a partir de sua inversa calcula-se o mapeamento de um pixel (geralmente domínio não-linear) para o domínio linear. Uma vez com a imagem linearizada aplica-se uma correção gama de 2.2 para aumentar a eficácia do algoritmo.

Para a primeira iteração do algoritmo, é necessário a atribuição do conjunto de imagens de entrada I_k como sendo $I_k \leftarrow g^k(L_{ref})$.

Para poder definir melhor quais imagens influenciam mais a geração da imagem HDR, os autores utilizaram um conceito de pesos a mais na equação 4.14, de forma que imagens com alta exposição não contribuirão com pixels muito saturados e imagens com baixa exposição não contribuirão com pixels de baixa intensidade.

Outra otimização feita, é a implementação de fases do algoritmo relativas a escalas de forma que a correspondência é feita gradualmente de uma baixa resolução para uma alta resolução.

4.3.3 Resultados e discussões

O método foi testado utilizando um conjunto de imagens estáticas e outras com movimento em diferentes níveis. Os resultados são mostrados nas figuras:

/a /a /a /a /a

Notou-se que ao utilizar imagens com alta resolução o processamento era longo e com consumo de memória excessivo. Para o caso em que foi testado, num computador com memória RAM de 4GB, utilizando imagens de 4000x3000 pixels, não foi possível a execução do método pois a memória era insuficiente para o processo, sendo assim necessário o redimensionamento das imagens para 2000x1500 pixels.

Capítulo 5

Discussões

Com os resultados alcançados, notou-se que todos os métodos atingiram os resultados esperados (geração de imagens HDR), cada um possuindo vantagens e desvantagens em relação aos outros. Nas seções seguintes serão abordadas estas questões referentes a cada método abordado por este documento.

5.1 Método Mann

5.1.1 Vantagens

- Possui a inferência da função resposta mais rápida.
- Possuiu um resultado mais consistente em alguns casos em relação ao método de Robertson.

5.1.2 Desvantagens

- As imagens devem ser estáticas umas em relação as outras.
- A inferência da função resposta leva em conta apenas alguns pixels e não todos.
- A inferência da função resposta leva em conta apenas 2 imagens.
- A falta de especificação da escala na inferência da função resposta leva, em alguns casos, a incoerência na união dos canais de cores.

5.2 Método Robertson

5.2.1 Vantagens

- A inferência da função resposta leva em conta todas as imagens.
- A inferência da função resposta leva em conta todos os pixels da imagem.

5.2.2 Desvantagens

- Na implementação feita algumas imagens HDR apresentaram ruído.
- As imagens devem ser estáticas umas em relação as outras.

5.3 Método Sen

5.3.1 Vantagens

- Robustez quanto ao movimento nas imagens.

5.3.2 Desvantagens

- Necessita de função resposta.
- Método mais lento de obtenção de imagens HDR.
- Não suporta altas resoluções.

Capítulo 6

Considerações finais

Com base nos resultados obtidos, foi considerado que os métodos são complementares e não excludentes. Foi decidido como método mais importante para o próximo passo do projeto o método de geração de imagem HDR de Sen et al. [12], pela sua robustez quanto ao movimento.

Os outros métodos serão utilizados como ferramentas auxiliares para geração da função resposta da câmera uma vez que o método escolhido não possui inferência da função resposta da câmera.

O próximo passo do projeto será a pesquisa e verificação de métodos para geração de nuvem de pontos a partir de imagens HDR.

REFERÊNCIAS

- [1] Sameer Agarwal, Yasutaka Furukawa, Noah Snavely, Ian Simon, Brian Curless, Steven M. Seitz, and Richard Szeliski. Building rome in a day. *Commun. ACM*, 54(10):105–112, October 2011.
- [2] Mir Adnan Ali and Steve Mann. Comparametric image compositing: Computationally efficient high dynamic range imaging. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 913–916. IEEE, 2012.
- [3] Beatriz Trinchão Andrade, Caroline Mazetto Mendes, Jurandir de Oliveira Santos Jr., Olga Regina Pereira Bellon, and Luciano Silva. 3D preserving XVIII century barroque masterpiece: Challenges and results on the digital preservation of Aleijadinho’s sculpture of the Prophet Joel. *Journal of Cultural Heritage*, 13(2):210 – 214, 2012.
- [4] Dana Harry Ballard and Christopher M. Brown. *Computer Vision*. Prentice Hall Professional Technical Reference, 1st edition, 1982.
- [5] M. Granados, B. Ajdin, M. Wand, C. Theobalt, H.-P. Seidel, and H. Lensch. Optimal hdr reconstruction with linear digital cameras. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 215–222, June 2010.
- [6] G. Kontogianni, E. K. Stathopoulou, A. Georgopoulos, and A. Doulamis. Hdr imaging for feature detection on detailed architectural scenes. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-5/W4:325–330, 2015.
- [7] S. Mann and R.W. Picard. Being ‘undigital’ with digital cameras: Extending dynamic range by combining differently exposed pictures. Technical Report 323, M.I.T. Media Lab Perceptual Computing Section, Boston, Massachusetts, 1994.

- [8] Thaís Matsushigue, Valmir Pagliaro Franco, Rafael Pierami, Marcel Jun Sugawara Tamaoki, Nicola Archetti Netto, and Marcelo Hide Matsumoto. A tomografia computadorizada e sua reconstrução 3d aumentam a reprodutibilidade das classificações das fraturas da extremidade proximal do úmero? *Revista Brasileira de Ortopedia*, pages 174–177, 2014.
- [9] T. Mitsunaga and S.K. Nayar. Radiometric Self Calibration. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 374–380, Jun 1999.
- [10] Mark A. Robertson, Sean Borman, and Robert L. Stevenson. Dynamic range improvement through multiple exposures. In *Proceedings of the IEEE International Conference on Image Processing*, volume 3, pages 159–163, Kobe, Japan, October 1999. IEEE.
- [11] Mark A. Robertson, Sean Borman, and Robert L. Stevenson. Estimation-theoretic approach to dynamic range enhancement using multiple exposures. *Journal of Electronic Imaging* 12(2), 219–228 (April 2003)., 12(2):219–228, April 2003.
- [12] Pradeep Sen, Nima Khademi Kalantari, Maziar Yaesoubi, Soheil Darabi, Dan B. Goldman, and Eli Shechtman. Robust patch-based hdr reconstruction of dynamic scenes. *ACM Trans. Graph.*, 31(6):203:1–203:11, November 2012.