

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

**OBTENÇÃO DE NUVEM DE PONTOS A PARTIR DE
IMAGENS HDR GERADAS UTILIZANDO MÉTODO DE
MÚLTIPLAS EXPOSIÇÕES**

CLAUDIO MOTA OLIVEIRA

SÃO CRISTÓVÃO/SE

2016

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

CLAUDIO MOTA OLIVEIRA

**OBTENÇÃO DE NUVEM DE PONTOS A PARTIR DE
IMAGENS HDR GERADAS UTILIZANDO MÉTODO DE
MÚLTIPLAS EXPOSIÇÕES**

Trabalho apresentado ao Departamento de Computação (DCOMP) da Universidade Federal de Sergipe (UFS) como requisito para obtenção do título de Bacharel em Ciência da Computação.

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

SÃO CRISTÓVÃO/ SE

2016

Resumo

Este documento mostra a pesquisa e implementação de métodos de geração de imagens HDR a partir de um conjunto de imagens LDR. A literatura apresenta vários métodos para a geração de imagens com alta faixa dinâmica como os métodos de Mann e Picard que foi um dos primeiros métodos propostos, o método de Robertson et al., que trata do ruído da câmera sob a ótica estatística e o método de Sen et al., que trata o problema com o enfoque em robustez quanto ao movimento. O principal objetivo deste trabalho é definir um método viável para ser utilizado na geração de imagens HDR que servirão de entrada para um posterior trabalho de obtenção de nuvem de pontos a partir de imagens HDR. Os resultados obtidos foram satisfatórios para todos os métodos abordados, sendo então escolhido o método mais robusto quanto ao movimento para os próximos passos do projeto.

Abstract

This document describes the research and implementation of methods for generating HDR images using LDR images as input. There is, in literature, a lot of former works that presents methods of generating high dynamic range images like the works of Mann and Picard who appear to be the first method proposed in history, Robertson et al.method which look the problem at statistical view and the method of Sen et al.who cares about the robustness in relation to moviment between the input LDR images. The main objective of this work is to define a viable method for generating HDR images that will be used as input for a later work of obtaining a 3D point cloud from HDR images. The results obtained were satisfactory for all methods implemented, so it was decided, for the next step, to use the method which is more robust in relation to moviment.

Listas de Figuras

2.1	Funções resposta e suas representações em diferentes eixos.	7
3.1	Note que a parte interna da sala possui uma baixa iluminação e é perdida nas imagens com menor tempo de exposição. Enquanto a parte externa possui alta iluminação e é perdida nas imagens com maior tempo de exposição. . .	13
3.2	Note que partes da cena são perdidas em algumas figuras. Porém estas mesmas partes estão bem representadas em outras.	14
3.3	Note que partes da cena são perdidas em algumas figuras. Porém estas mesmas partes são bem representadas em outras figuras.	15
3.4	Imagen HDR com diferença de escala entre os canais.	18
3.5	Imagen HDR obtida utilizando o método de Mann e Picard [8].	19
3.6	Imagen HDR obtida utilizando o método de Mann e Picard [8].	20
3.7	Imagen HDR obtida utilizando o método de Mann e Picard [8].	21
3.8	Funções resposta da câmera no eixo linear (a) e no eixo logarítmico (b), referentes aos canais vermelho, verde e azul, obtidos com o método de Mann e Picard [8].	22
3.9	Interpolação da função resposta do canal azul da câmera, utilizando o método da Spline cúbica.	26
3.10	Imagen HDR gerada utilizando o método de Robertson et al. [13].	27
3.11	Imagen HDR gerada utilizando o método de Robertson et al. [13].	28
3.12	Imagen HDR gerada utilizando o método de Robertson et al. [13].	29
3.13	Imagen HDR gerada utilizando o método de Robertson et al. [13], que apresentou ruído em partes com muita irradiação de luz.	30
3.14	Funções resposta da câmera relativas aos diferentes canais de cores.	31

3.15	Imagen HDR gerada utilizando o método de Sen et al. [15].	36
3.16	Imagen HDR gerada utilizando o método de Sen et al. [15].	37
3.17	Imagen HDR gerada utilizando o método de Sen et al. [15].	38
4.1	Exemplo de execução do VisualSFM para a geração de nuvem de pontos a partir de imagens LDR.	43
4.2	Exemplo de imagem HDR <i>tonemapped</i>	44
4.3	Filtro de Laplace.	45
4.4	Registro em diferentes tempos de exposição da frente do objeto.	47
4.5	Registro em diferentes tempos de exposição da esquerda do objeto.	47
4.6	Registro em diferentes tempos de exposição da direita do objeto.	48
4.7	Registro em diferentes tempos de exposição da parte de cima do objeto. . . .	48
4.8	Registro em diferentes tempos de exposição do objeto visto por baixo. . . .	49
4.9	Imagenes HDR <i>tonemapped</i> da base de dados.	50
4.10	Nuvens de pontos obtidas utilizando o software VisualSFM.	51

Lista de Tabelas

2.1 Progressão dos eixos	6
4.1 Resultados obtidos com a geração das nuvens de pontos.	51

Sumário

1	Introdução	1
1.1	Objetivo	3
1.2	Metodologia	3
2	Conceitos Básicos	4
2.1	Faixa Dinâmica	4
2.2	Função Resposta de uma Câmera	5
2.3	Eixo Logarítmico	6
2.4	Similaridade Bidirecional	8
2.5	Relaxação de Gauss-Seidel	8
2.6	Spline Cúbica	8
3	Obtenção de Imagens HDR a Partir de Imagens LDR	10
3.1	Revisão Bibliográfica	10
3.2	Métodos HDR Selecionados	12
3.2.1	Base de Imagens e Visualização	12
3.2.2	Mann e Picard	16
3.2.3	Método Robertson	22
3.2.4	Método Sen	31
3.3	Discussões	39
3.3.1	Método Mann	39
3.3.2	Método Robertson	39
3.3.3	Método Sen	40

4	Obtenção de Nuvem de Pontos a Partir de Imagens HDR	41
4.1	Revisão Bibliográfica	41
4.2	Software Utilizado	42
4.3	Introduzindo as Imagens HDR ao Processo	44
4.3.1	Método de Tone Mapping	44
4.4	Proposta para Melhoria na Aquisição da Nuvem de Pontos	45
4.5	Base de Imagens	46
4.6	Resultados e Discussões	49
5	Considerações Finais	52
	Referências	52

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 áreas do conhecimento. Isso é mostrado nos trabalhos de Matsushigue et al. [9], que abordou o uso de reconstrução de tomografias 3D para melhoria de diagnóstico de fraturas na região do úmero, e como o trabalho de Andrade 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. Como resultado do trabalho, 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. Com o uso de conjuntos de pontos de interesse comuns entre as imagens é possível estimar a posição deles em relação a um sistema de coordenadas tridimensional, gerando assim uma nuvem de pontos 3D.

Com base na revisão de literatura feita neste trabalho sobre obtenção de nuvem de pontos

3D a partir de imagens, evidenciou-se que grande parte dos métodos existentes utiliza-se de imagens com baixa faixa dinâmica (em inglês, Low Dynamic Range - LDR), que possuem uma resolução limitada de cores. Por outro lado existem também as imagens com alta faixa dinâmica (em inglês, High Dynamic Range - HDR), que possuem maior resolução de cores.

Sob o ponto de vista de acessibilidade, o uso de imagens LDR é explicado pelo fato de que câmeras com sensores capazes de capturar imagens HDR serem muito caras e assim pouco acessíveis. Porém sob o ponto de vista de qualidade, as imagens HDR possuem maior quantidade de informações do cenário em relação a uma imagem LDR. Um exemplo disso, seria um cenário de uma sala pouco iluminada, que possui uma janela ao fundo a qual mostra o céu de um dia ensolarado. Ao capturar uma imagem deste ambiente com uma câmera LDR, o fotógrafo terá que ajustar as configurações da câmera, para escolher qual parte da cena será melhor registrada (os móveis da sala ou o céu ensolarado). Já para uma câmera HDR todas as partes da imagem seriam propriamente registradas, uma vez que esta possui uma maior faixa de representação de cores.

Trabalhos publicados na literatura, propõem métodos para união de um conjunto de imagens LDR em uma imagem HDR. O uso destas técnicas se mostra uma alternativa economicamente viável para obter imagens de boa qualidade, com alta faixa dinâmica.

Assim como no trabalho feito por Kontogianni et al. [7], 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. Espera-se 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.

1.1 Objetivo

O objetivo deste trabalho é verificar se o uso de imagens HDR, obtidas utilizando conjuntos de imagens LDR, gera nuvens de pontos mais densas em relação às nuvens de pontos geradas a partir de imagens LDR.

1.2 Metodologia

Para que seja possível alcançar o objetivo, este trabalho é dividido em duas etapas:

- Geração de imagem HDR:

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

- Geração de nuvem de pontos:

Nesta etapa são utilizados métodos para geração de nuvem de pontos a partir de imagens, para que assim sejam realizadas as comparações inicialmente propostas.

Capítulo 2

Conceitos Básicos

Este capítulo aborda conceitos fundamentais para a compreensão dos métodos de geração de imagens HDR e obtenção de nuvem de pontos, que serão abordados nos capítulos posteriores.

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 que são representados por 8 bits em cada canal de cor, o que resulta numa faixa de 256 valores possíveis. Quando uma imagem possui esta faixa de representação de cores esta é chamada de imagem com baixa faixa dinâmica (em inglês, Low Dynamic Range - LDR) por possuir uma faixa relativamente curta de cores para representar a cena. Quando uma imagem possui uma faixa maior de representação de cores é denominada imagem com alta faixa dinâmica (em inglês, High Dynamic Range - HDR). As imagens HDR geralmente representam as cores em ponto flutuante o que diminui a deterioração cumulativa ao efetuar operações sobre a imagem sucessivamente. Além disso, 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 de iluminação (pouco/muito iluminadas) serão perdi-

das. Isto acontece devido ao fato da imagem não ter uma faixa de valores suficiente para representar tanto os níveis altos como os níveis baixos de iluminação desta cena. Por outro lado, 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 x e t . x representa o valor de irradiação de luz que incide sobre o sensor, e t representa o período de tempo que esta incidência persiste. Esta função tem como saída um valor geralmente inteiro p , que é o valor do pixel. f é chamada de função resposta da câmera e idealmente ela pode ser descrita pela seguinte equação:

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

Porém, no mundo real há a presença de ruídos, advindos de diversas fontes, que distorcem os sinais tanto de entrada quanto de saída da câmera. Devido a isso, a equação 2.1 pode ser modificada para a seguinte equação que discrimina os ruídos:

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

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 gerou este pixel.

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

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, cada método de geração de imagens HDR trata estes problemas de maneira singular.

Outro problema é que na prática o usuário da câmera geralmente não possui nenhuma informação sobre a função resposta da câmera que está utilizando. Devido a isso alguns 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 este conceito para ressaltar a confiabilidade que um valor de pixel representa, levando-o 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.

Tabela 2.1: Progressão dos eixos

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

A Figura 2.1 mostra uma mesma função resposta representada em um eixo linear (2.1a) e logarítmico, respectivamente (2.1b).

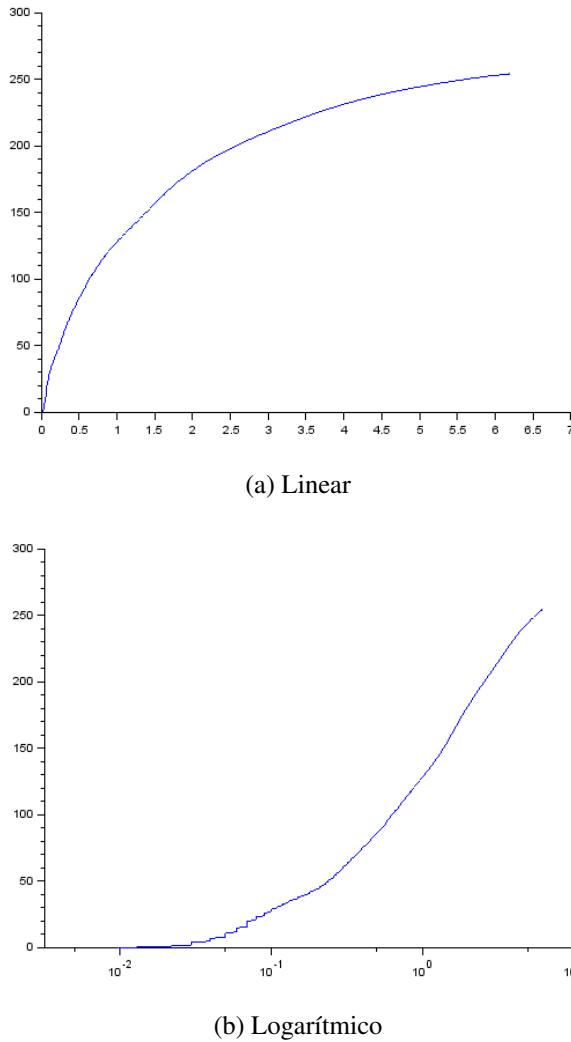


Figura 2.1: Funções resposta e suas representações em diferentes eixos.

A importância do eixo logarítmico para os métodos de geração de imagens HDR se dá por conta de uma suposição assumida por alguns dos autores. Segundo Mann [8] e Robertson [13] a derivada da função resposta da câmera num eixo logarítmico possui relação direta com a confiabilidade que um pixel possui, i.e, quanto maior o valor de confiabilidade menor o ruído associado ao pixel. 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 Similaridade Bidirecional

A Similaridade Bidirecional [16] é um método de maximização das similaridades entre dois sinais S (fonte) e T (destino), que se utiliza de um conjunto de fragmentos de ambos os sinais para definir a similaridade tanto em relação à sua coerência quanto à completude das informações.

Uma imagem T é dita coerente em relação a outra imagem S quando todo fragmento P_i , pertencente a T , está contido em S . Por outro lado, se T contém todos os fragmentos de S , então é dita completa em relação a S .

O método maximiza a similaridade visual por meio da minimização da seguinte função:

$$d(S, T) = \frac{1}{N_S} \sum_{P \subset S} \min_{Q \subset T} D(P, Q) + \frac{1}{N_T} \sum_{Q \subset T} \min_{P \subset S} D(Q, P) \quad (2.5)$$

Onde

- N_S e N_T são os números de fragmentos de S e de T respectivamente.
- $D(,)$ representa a distância, i.e., dissimilaridade entre dois fragmentos.

2.5 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 método iterativo para todas as variáveis de maneira cíclica até atingir uma determinada condição de convergência. Mais informações sobre o método estão disponíveis na Referência [14].

2.6 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. Para isso considera-se que cada intervalo entre dois pontos consecutivos são funções polinomiais de terceiro grau. E a partir de premissas definidas em relação ao primeiro e ultimo ponto, é descoberto os valores das fuções de cada intervalo. Mais informações estão disponíveis na Referência [14].

Capítulo 3

Obtenção de Imagens HDR a Partir de Imagens LDR

Neste capítulo será mostrada a pesquisa bibliográfica e implementação de métodos de geração de imagens HDR a partir de um conjunto de imagens LDR. O principal objetivo deste capítulo é comparar os métodos existentes na literatura, afim de definir um método viável para geração de imagens HDR que servirão de entrada para a próxima etapa do projeto.

3.1 Revisão Bibliográfica

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

Com base no levantamento bibliográfico realizado, o primeiro método de geração de imagens HDR a partir de um conjunto de imagens LDR foi proposto por Mann e Picard [8]. O trabalho deles propõe que dado um conjunto de imagens de uma mesma cena, que diferem 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 primeiramente 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 relativo a cada pixel 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 [10], que trata a questão da exposição dos sensores da câmera sob a ótica do tempo e da 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. [13] 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, dadas 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. [6], 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. [15] propôs um método iterativo que busca robustez em relação ao movimento da câmera, baseado em técnicas de minização de energia. O método proposto escolhe uma imagem de referência no conjunto de imagens LDR de entrada. A partir disso processa todas as outras imagens, buscando inferir detalhes de iluminação da imagem referência fazendo comparações entre os fragmentos das imagens. Isto faz com que sejam geradas versões da imagem referência em diferentes tempos de exposição, que não existiam anteriormente. Possuindo imagens idênticas que diferem apenas no tempo de exposição, é aplicado um método para estimar a iluminação que gerou cada pixel das imagens, obtendo uma imagem HDR.

3.2 Métodos HDR Selecionados

Nesta seção serão abordados os métodos de geração de imagens HDR a partir de imagens LDR escolhidos para este trabalho. Cada subseção tratará um método, mostrando aspectos teóricos, implementação e resultados obtidos, assim como discussões sobre tópicos de interesse. Nas seções que seguem serão abordados e comparados os métodos de geração de imagens HDR dos seguintes autores:

- Mann e Picard: Por se tratar de um método clássico, frequentemente citado em trabalhos da área podendo ser usado como base para comparação. Além de se tratar de um método simples que aborda apenas a mudança no tempo de exposição das imagens LDR para geração de imagens HDR, e ao mesmo tempo se mostra suficiente para o objetivo final deste trabalho.
- Robertson et al.: Por se tratar de um método mais elaborado que, além de abordar a 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 um fator limitante.

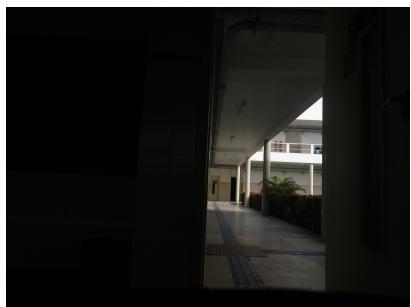
3.2.1 Base de Imagens e Visualização

Para a geração das imagens HDR nos métodos das próximas subseções, será utilizada a base de imagens LDR aqui apresentada. Essa base de dados consiste num conjunto de cenas, cada uma possuindo partes bem iluminadas e partes pouco iluminadas. Cada cena foi registrada por um conjunto de imagens LDR, que diferem apenas em tempo de exposição, cada imagem busca registrar uma faixa de iluminação específica da cena. Com isso visa-se verificar a eficiência de cada método em gerar imagens HDR dessas cenas.

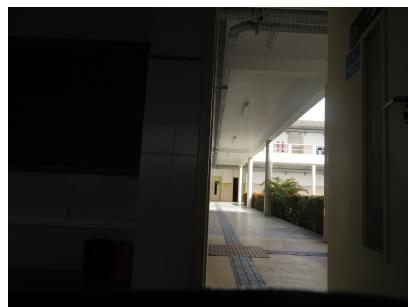
É importante salientar que, apesar de pouco, o movimento entre a captura das imagens LDR de cada cena da base de dados existe. Esse foi um fator limitante neste trabalho, devido aos equipamentos disponíveis para uso. Outro fator limitante devido ao equipamento

utilizado foi o uso de imagens no formato JPG, que possui compressão com perda de informações.

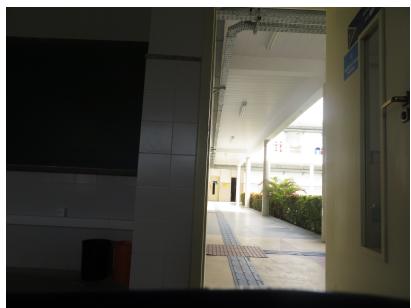
Cena da Didática



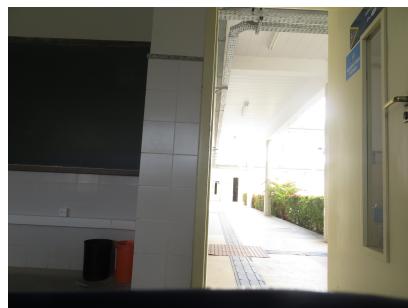
(a) Tempo de exposição de 0.002s.



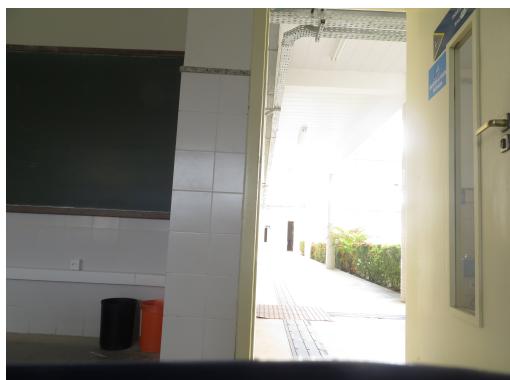
(b) Tempo de exposição de 0.005s.



(c) Tempo de exposição de 0.0167s.



(d) Tempo de exposição de 0.04s.



(e) Tempo de exposição de 0.067s.

Figura 3.1: Note que a parte interna da sala possui uma baixa iluminação e é perdida nas imagens com menor tempo de exposição. Enquanto a parte externa possui alta iluminação e é perdida nas imagens com maior tempo de exposição.

Cena da Sala de Estar

(a) Tempo de exposição de 0.077s.



(b) Tempo de exposição de 0.25s.



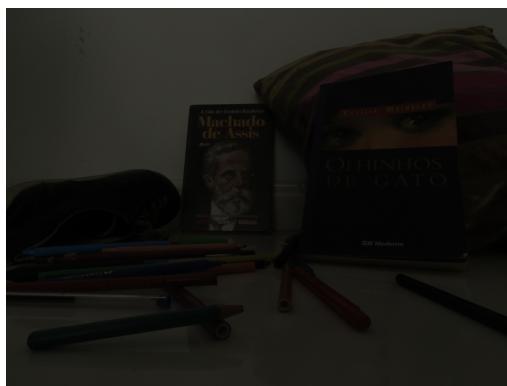
(c) Tempo de exposição de 0.6s.



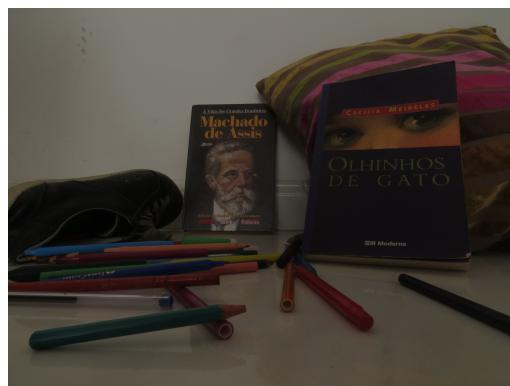
(d) Tempo de exposição de 1.3s.

Figura 3.2: Note que partes da cena são perdidas em algumas figuras. Porém estas mesmas partes estão bem representadas em outras.

Cena dos Livros



(a) Tempo de exposição de 0.02s.



(b) Tempo de exposição de 0.067s.



(c) Tempo de exposição de 0.2s.



(d) Tempo de exposição de 0.6s.



(e) Tempo de exposição de 2.0s.

Figura 3.3: Note que partes da cena são perdidas em algumas figuras. Porém estas mesmas partes são bem representadas em outras figuras.

Software para Visualização de Imagens HDR

Para a visualização das imagens HDR geradas, é necessário o uso de softwares específicos. Neste trabalho é utilizado o Pictureaut, disponível no link da Referência [18]. Este foi escolhido devido a seu fácil manuseio, compatibilidade com diversos tipos de arquivos e por possuir uma ferramenta que realiza tone mapping de imagens HDR.

3.2.2 Mann e Picard

O método proposto por Mann e Picard [8] 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 dos sensores da câmera, de forma que a perda seja a menor possível, ou valorizando apenas áreas de seu interesse. A solução proposta pelos autores restringe o problema à captura de várias imagens de uma mesma cena, variando apenas o tempo de exposição, sem que haja movimentação da câmera entre as capturas. 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 gerou o mapeamento para um dado pixel da imagem. Para isso, é necessário fazer o cálculo da inversa da função resposta da câmera, que é diferente para cada modelo de câmera no mercado, e muitas vezes é não-linear. Os autores propuseram um algoritmo para inferir a função resposta da câmera e assim calcular a inversa da mesma.

Algoritmo para Inferir a Função Resposta

Para o algoritmo são necessárias duas imagens a e b que foram capturadas em diferentes tempos de exposição, e 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 à 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 $a(x_2, y_2) = f(k^2q_0)$.
- Ao aplicar a lógica do passo anterior sucessivamente para os pontos $(x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)$ é feito o mapeamento da não linearidade da função resposta.

Uma vez com os pontos encontrados elimina-se o termo q_0 por meio da mudança de escala, uma vez que não importa a unidade do eixo de irradiação [8]. O que leva nível de irradiação q_0 a ser considerado uma unidade de irradiação.

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

$$f(q) = \alpha + \beta q^\gamma \quad (3.1)$$

Onde α pode ser considerado zero ao subtrair 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:

$$b(x_i, y_i) = k^\gamma + a(x_i, y_i) \quad (3.2)$$

Geração da Imagem HDR

O método proposto pelos autores para geração de uma imagem HDR se baseia na ideia que a derivada da função resposta da câmera em relação ao eixo logarítmico é proporcional à 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. Assim gerando um pixel HDR que foi formado a partir de informações de todas as exposições.

Seja um conjunto com N imagens alinhadas entre si, 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:

$$x_j = \frac{\sum_{i \in N, j \in M} p_{ij} \frac{f^{-1}(I_{ij})}{t_i}}{\sum_{i \in N, j \in M} p_{ij}} \quad (3.3)$$

Resultados e Discussões

O método em questão foi implementado na linguagem de programação C++, e foi utilizado um conjunto de imagens capturadas neste trabalho para a verificação dos resultados. Verificou-se que pelo fato do beta da equação 3.1 ser arbitrário a depender do valor atribuído ao mesmo pode-se haver uma incompatibilidade entre as funções respostas dos diferentes canais de cores da imagem, como mostrado na Figura 3.4.

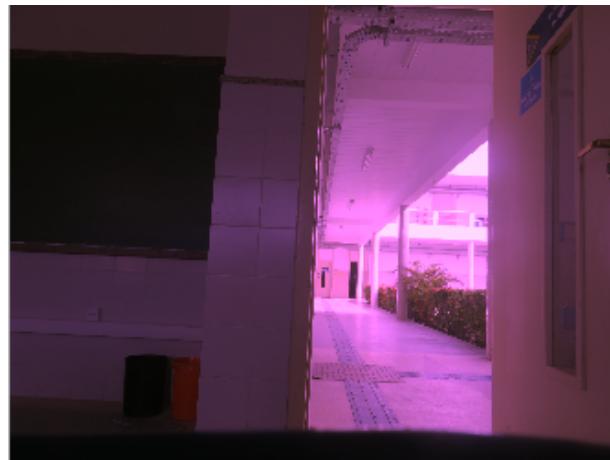


Figura 3.4: Imagem HDR com diferença de escala entre os canais.

Para solução deste problema foi utilizada a mesma solução proposta por Robertson et al. [13], 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, conforme mostrado nas Figuras 3.5, 3.6 e 3.7.



(a) Imagem visualizada com ganho de -9,5EV.



(b) Imagem visualizada com ganho de -4,8EV.

Figura 3.5: Imagem HDR obtida utilizando o método de Mann e Picard [8].



(a) Imagem visualizada com ganho de -3,6EV.



(b) Imagem visualizada com ganho de -0,4EV.

Figura 3.6: Imagem HDR obtida utilizando o método de Mann e Picard [8].



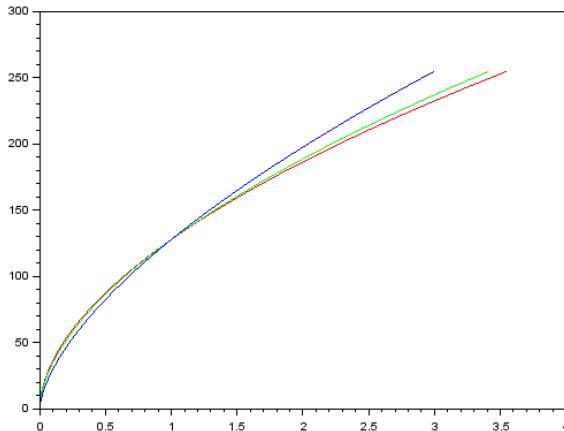
(a) Imagem visualizada com ganho de -2,9EV.



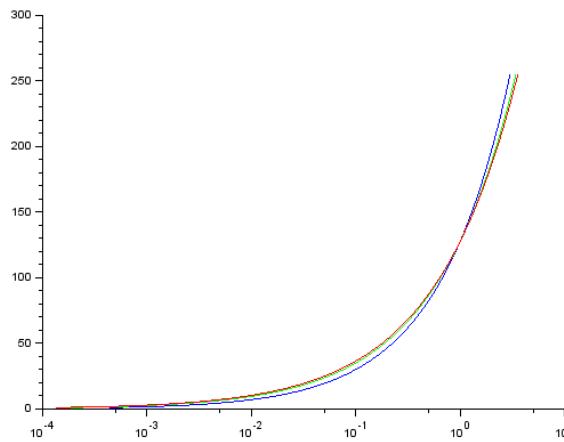
(b) Imagem visualizada com ganho de 0EV.

Figura 3.7: Imagem HDR obtida utilizando o método de Mann e Picard [8].

A Figura 3.8 mostra as funções resposta dos três canais obtidas utilizando o método proposto.



(a) Linear.



(b) Logarítmica.

Figura 3.8: Funções resposta da câmera no eixo linear (a) e no eixo logarítmico (b), referentes aos canais vermelho, verde e azul, obtidos com o método de Mann e Picard [8].

3.2.3 Método Robertson

O método proposto por Robertson et al. [13] aborda a geração de imagens HDR de maneira similar ao método proposto por Mann e Picard (Subseção 3.2.2). Esse método utiliza a derivada da função resposta em um eixo logarítmico como peso e assume que as imagens são estáticas entre si, variando-se apenas o tempo de exposição. As principais diferenças entre os dois métodos são a postura adotada em relação ao tratamento do ruído, e a forma como a função resposta da câmera é inferida. Neste caso a função resposta é inferida por meio de um método iterativo que supõe uma função inicial padrão, e então, com uso da

relaxação de Gauss-Seidel, aproxima a função resposta de forma a obter um erro abaixo de um limiar estabelecido.

Geração da Imagem HDR

Ao tentar extrair o valor de irradiação de luz ao qual um pixel é mapeado, aplica-se a inversa da função resposta. No entanto, no momento de registro da imagem, vários fatores influenciam para que o valor registrado seja ligeiramente diferente do valor ideal. Isso se dá por conta da presença do ruído no ambiente, que pode ser advindo de fontes diferentes, como corrente de escuro, conversão de analógico para digital entre outros. Assim, a equação da inversa da função resposta da câmera pode ser descrita da seguinte maneira:

$$f_{(p_{ij})}^{-1} = I_{p_{ij}} = xt_i + N_{ij} \quad (3.4)$$

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 j 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 complexa. Sendo assim, este método trata o ruído sob o ponto de vista de uma distribuição gaussiana com média zero para simplificar a modelagem.

Utilizando-se de operações algébricas para calcular o erro de um valor de iluminação inferido em relação ao valor real, obtém-se a seguinte equação:

$$O(x) = \sum_{i,j} w_{ij} (I_{p_{ij}} - x_{ij} t_i)^2 \quad (3.5)$$

Essa equação 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:

$$x_j^* = \frac{\sum_i w_{ij} t_i I_{p_{ij}}}{\sum_i w_{ij} t_{ij}^2} \quad (3.6)$$

Onde

- x_j^* é o valor de irradiação de luz inferido.
- O peso w_{ij} é obtido pela derivada da função resposta em relação um eixo logarítmico.

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

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 [13]. Para a estimativa da função resposta é necessária a minimização da função do erro quadrático, dada pela seguinte equação:

$$O(I, x) = \sum_{i,j} w_{ij} (I_{p_{ij}} - x_{ij} t_i)^2 \quad (3.7)$$

Essa equação é parecida com a Equação 3.5, 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 relaxação 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 propõem a seguinte função para definir os pesos [12]:

$$w_{ij} = e^{-4 \frac{(p_{ij} - 127.5)^2}{127.5^2}} \quad (3.8)$$

O método de inferência da função resposta possui um passo inicial e um conjunto de passos iterativos, descritos a seguir:

- Passo inicial:

Estabelece-se a função resposta inicial como sendo uma função linear com o valor $I_{128}^{*(0)} = 1.0$, e o valor de $x^{*(0)}$ é obtido aplicando a equação 3.6.

- 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:

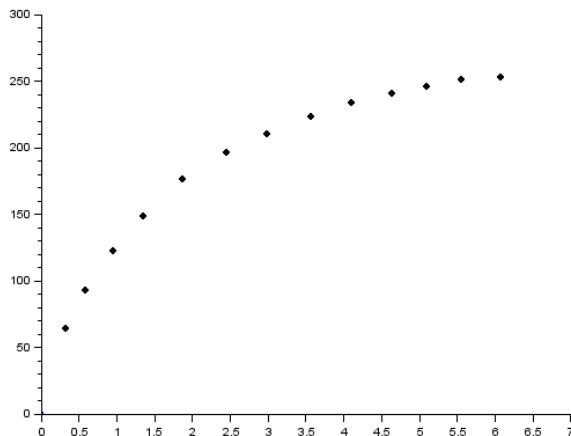
$$I^{*(l)} = \frac{\sum_{(i,j) \in E_m} w_{ij} t_i x_j^{*(l-1)}}{\sum_{(i,j) \in E_m} w_{ij}} \quad (3.9)$$

2. Normaliza-se $I_{p_{ij}}$ de forma que I_{128} seja uma unidade.
3. É feito o relaxamento de $x^{*(l)}$ com base na seguinte equação:

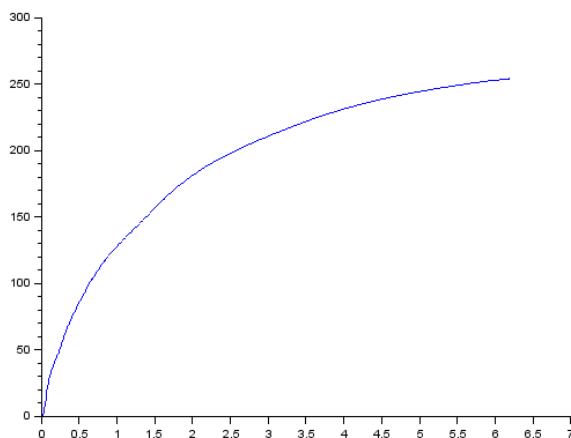
$$x_j^{*(l)} = \frac{\sum_i w_{ij} t_i I_{p_{ij}}^{*(l)}}{\sum_i w_{ij} t_i^2} \quad (3.10)$$

Dessa forma, a função resposta e a imagem HDR vão sendo inferidas simultaneamente, e quando as modificações feitas no passo iterativo da Equação 3.9 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 (vide Seção 2.6) para obtenção da função resposta final como ilustrado na Figura 3.9.



(a) Ilustração dos pontos utilizados na interpolação.



(b) Função resposta obtida com a interpolação.

Figura 3.9: Interpolação da função resposta do canal azul da câmera, utilizando o método da Spline cúbica.

Resultados e Discussões

O método em questão foi implementado na linguagem de programação C++, assim como o método de interpolação por spline cúbica. Os resultados são mostrados nas Figuras 3.10, 3.11 e 3.12.



(a) Imagem visualizada com ganho de -11,0EV.



(b) Imagem visualizada com ganho -4,3EV.

Figura 3.10: Imagem HDR gerada utilizando o método de Robertson et al. [13].



(a) Imagem visualizada com ganho de -4,3EV.



(b) Imagem visualizada com ganho de -2,0EV.



(c) Imagem visualizada com ganho de 0EV.

Figura 3.11: Imagem HDR gerada utilizando o método de Robertson et al. [13].



(a) Imagem visualizada com ganho de -4,3EV.



(b) Imagem visualizada com ganho de -2,7EV.

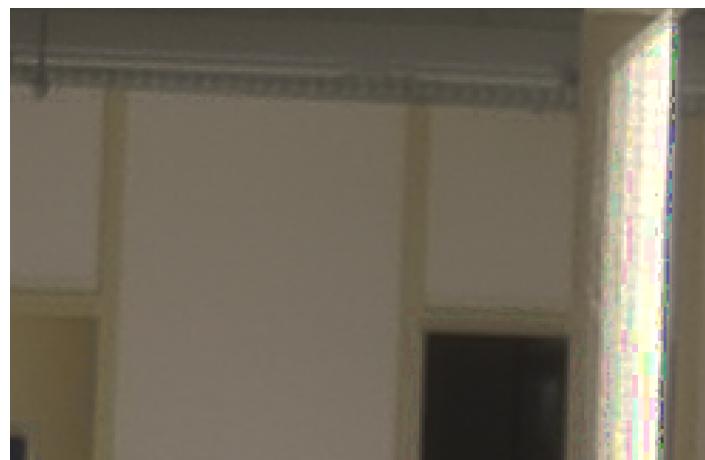


(c) Imagem visualizada com ganho de -0,4EV.

Figura 3.12: Imagem HDR gerada utilizando o método de Robertson et al. [13].



(a) Imagem HDR gerada



(b) A aproximação mostrou a presença de ruído na imagem

Figura 3.13: Imagem HDR gerada utilizando o método de Robertson et al. [13], que apresentou ruído em partes com muita irradiação de luz.

Notou-se que áreas com exposição muito elevada apresentaram ruído mais perceptível, como mostrado na Figura 3.13b. E a Figura 3.14 mostra as funções resposta dos diferentes canais, obtidas utilizando o método proposto.

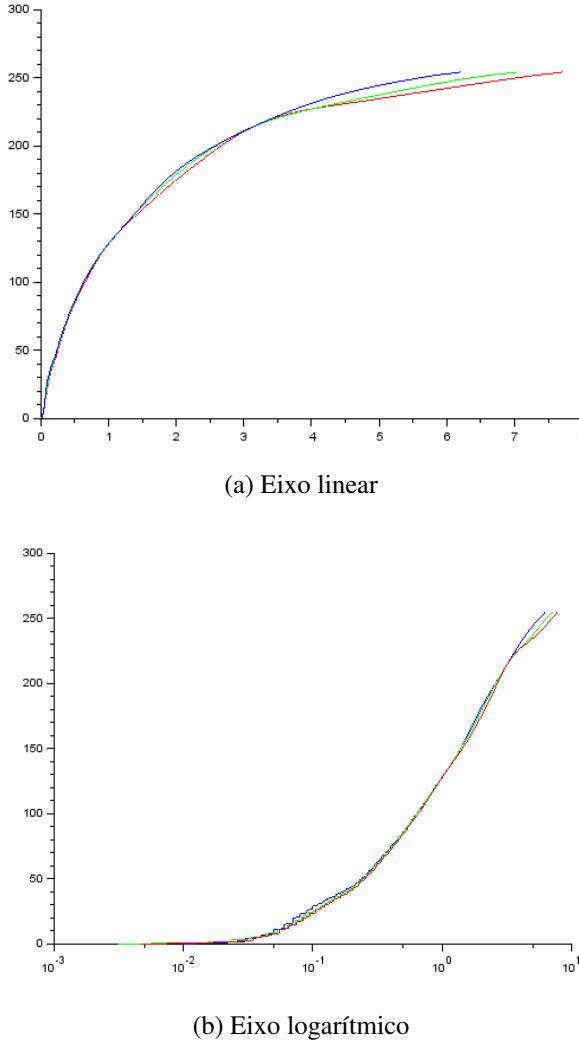


Figura 3.14: Funções resposta da câmera relativas aos diferentes canais de cores.

3.2.4 Método Sen

O método proposto por Sen et al. [15], 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 às outras imagens só é possível

para ambientes bastante controlados. Para isso há a necessidade de que os objetos da cena não sejam móveis, requer o uso de equipamentos mais sofisticados como tripé da câmera para mantê-la fixa e/ou uso de câmeras controladas remotamente. Estes casos são bastante específicos e não correspondem à 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 a cenas capturadas 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 várias imagens LDR, como seria o registro da imagem de referência se esta fosse capturada com os tempos de exposição das outras imagens. Simultaneamente o algoritmo gera a imagem HDR a com base nas imagens de referência, em diferentes exposições, obtidas. Utilizando a imagem HDR obtida o algoritmo verifica e corrige a qualidade das imagens LDR que a gerou. Estas imagens com qualidade melhorada geram uma imagem HDR de melhor qualidade. Este ciclo iterativo continua até atingir um limiar.

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 destas como entrada para os métodos citados nas Seções 3.2.2 e 3.2.3.

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 próximo ao valor do pixel da imagem de referência.
- Ao mapear um pixel, propriamente exposto (mensurado por uma função específica α) da imagem de referência para o domínio de irradiação linear, este deve possuir um valor próximo ao 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 múltipla similaridade bidirecional (MBDS), i.e., versão do conceito similaridade bidirecional para tratar múltiplas entradas.

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

$$E(H) = \sum_{p \in pixels} [\alpha_{ref(p)} \cdot (h(L_{ref})_{(p)} - H_{(p)})^2 + (1 - \alpha_{ref(p)}) \cdot E_{MBDS}(H | L_1, \dots, L_N)] \quad (3.11)$$

Onde

- $E(H)$ é a energia, que deve ser diminuída ao máximo para maior confiabilidade da imagem HDR H .
- $\alpha_{ref(p)}$ é uma função de peso que indica quão 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 à seguinte equação de energia:

$$\begin{aligned} E(H, I_1, \dots, I_N) = & \sum_{p \in pixels} [\alpha_{ref(p)} \cdot (h(L_{ref})_{(p)} - H_{(p)})^2 + \\ & (1 - \alpha_{ref(p)}) \cdot \sum_{k=1}^N MBDS(I_k | g^k(L_1), \dots, g^k(L_N)) + \\ & (1 - \alpha_{ref(p)}) \cdot \sum_{k=1}^N \Lambda(I_k)_{(p)} \cdot (h(I_k)_{(p)} - H_{(p)})^2] \end{aligned} \quad (3.12)$$

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 possuam 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 3.13, 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 3.14.

$$\hat{H}_{(p)} = \frac{\sum_{k=1}^N \Lambda(I_{k(p)}) h(I_k)_{(p)}}{\sum_{k=1}^N \Lambda(I_{k(p)})} \quad (3.13)$$

$$H_{(p)} = \alpha_{ref_{(p)}} \cdot h(L_{ref})_{(p)} + (1 - \alpha_{ref_{(p)}}) \cdot \hat{H}_{(p)} \quad (3.14)$$

Esses passos são repetidos até atingir um critério de convergência.

Especificidades de Implementação

Para implementação do método foi o utilizado a própria implementação disponibilizada pelos autores[15]. Para o correto funcionamento deste há a necessidade de efetuar um preprocessamento nas imagens para colocá-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 necessária a definição 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 extras na Equação 3.14, de forma que imagens com alta exposição não contribuam com pixels muito saturados e imagens com baixa exposição não contribuam 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.

Resultados e Discussões

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.

O método foi testado utilizando o conjunto de imagens mostrado na seção 3.2.1. Os resultados são mostrados nas Figuras 3.15, 3.16 e 3.17. Apesar da ausência de fantasmas nas imagens, algumas delas, conforme mostrado na Figura 3.15b, possuíram artefatos que se assemelham a uma quatização das cores quando o ganho é relativamente alto.

(a) Imagem visualizada com ganho de $-9,8\text{EV}$.(b) Imagem visualizada com ganho de $-1,1\text{EV}$.

Figura 3.15: Imagem HDR gerada utilizando o método de Sen et al. [15].



(a) Imagem visualizada com ganho de -1,3EV.



(b) Imagem visualizada com ganho de 2,5EV.

Figura 3.16: Imagem HDR gerada utilizando o método de Sen et al. [15].

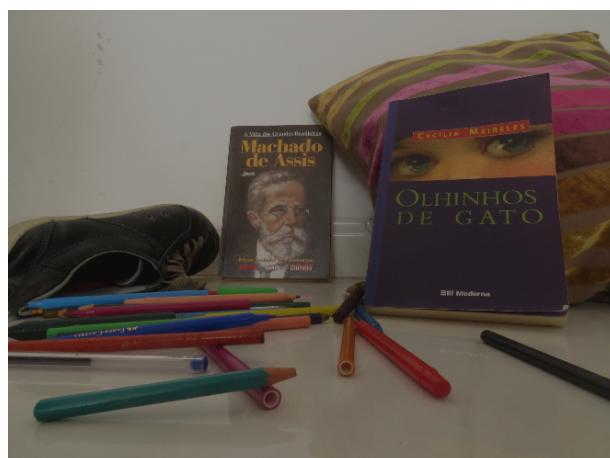
(a) Imagem visualizada com ganho de $-0,6\text{EV}$.(b) Imagem visualizada com ganho de $-3,4\text{EV}$.(c) Imagem visualizada com ganho de $3,6\text{EV}$.

Figura 3.17: Imagem HDR gerada utilizando o método de Sen et al. [15].

3.3 Discussões

Com os resultados alcançados, notou-se que todos os métodos atingiram os resultados esperados (geração de imagens HDR), cada qual com suas vantagens e desvantagens. Nesta seção serão abordadas as questões referentes a cada método de geração de imagens HDR abordado neste documento.

3.3.1 Método Mann

Vantagens

- Possui a inferência da função resposta mais rápida.
- Em alguns casos mostrou resultado mais consistente que o método de Robertson et al.[13].

Desvantagens

- As imagens de entrada não devem apresentar movimento entre suas capturas.
- 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, à incoerência na união dos canais de cores.

3.3.2 Método Robertson

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.

Desvantagens

- Na implementação feita algumas imagens HDR apresentaram ruído.

- As imagens de entrada não devem apresentar movimento entre suas capturas.

3.3.3 Método Sen

Vantagens

- Robustez quanto ao movimento nas imagens.

Desvantagens

- Necessita de função resposta.
- Método mais lento de obtenção de imagens HDR.
- Não suporta imagens em alta resolução.

Tanto o método proposto por Mann e Picard [8], quanto o proposto Robertson et al. [13], supõem que a câmera estará totalmente estática entre a captura das imagens. Para que isso seja verdade, há a necessidade de equipamentos mais sofisticados como tripé e software de captura por computador e de um ambiente muito controlado.

Tendo em vista que usuários comuns de câmeras digitais não costumam possuir equipamento que possibilite a inexistência de movimento entre a captura de imagens. O método mais aplicável para este cenário se mostra o proposto por Sen et al. [15], por se tratar de um método mais robusto quanto ao movimento. Sendo assim, este será o método utilizado na próxima etapa deste trabalho para a obtenção de nuvem de pontos a partir de imagens HDR.

Capítulo 4

Obtenção de Nuvem de Pontos a Partir de Imagens HDR

Neste capítulo serão abordados conceitos relacionados a geração de nuvem a partir de imagens HDR. Para isso serão introduzidos trabalhos e técnicas relacionados ao assunto, assim como softwares e ferramentas utilizadas para geração de nuvens de pontos convencionais. E então será proposto um método para aquisição efetiva da nuvem de pontos a partir de imagens HDR.

4.1 Revisão Bibliográfica

Dentre os principais assuntos que se relacionam ao tema deste capítulo, pode-se destacar a obtenção de pontos de interesse de imagens HDR e a geração de nuvens de pontos. Nesta seção serão abordados trabalhos da literatura que se relacionam com estes temas.

Kontogianni et al. [7] apresentaram um estudo sobre a diferença entre imagens LDR e HDR na obtenção de pontos de interesse. A motivação do trabalho foi importância da preservação da herança cultural e arquitetural que pode ser feita através da reconstrução 3D, e com isso a necessidade de alta precisão e preservação dos detalhes no processo. A hipótese foi que o uso das imagens HDR implicaria num maior número de pontos de interesse, que por sua vez agregariam mais informação para a reconstrução 3D das cenas. Na implementação foram geradas imagens HDR a partir de conjuntos de imagens LDR, seguida pelo *tone mapping* da imagem HDR gerada, para assim passá-la como entrada dos algoritmos de ob-

tenção de pontos de interesse. Os resultados mostraram que o uso das imagens processadas obteve um aumento significativo do número de pontos de interesse encontrados em relação às imagens LDR.

Přibyl et al. [11] também apresentaram um trabalho que verifica a serventia de técnicas HDR para obtenção de pontos de interesse em condições extremas de iluminação, i.e., que possui áreas bem muito iluminadas e áreas pouco iluminadas. Assim como Kontogianni et al. [7], utilizou-se de técnicas de *tone mapping* sobre as imagens HDR para então usá-las como entrada de métodos de obtenção de pontos de interesse. Seus resultados mostraram que o uso das imagens processadas aumentou a taxa de repitibilidade dos pontos de interesse significativamente.

O trabalho mostrado por Wu [17] introduz uma nova técnica de reconstrução 3D a partir de múltiplas imagens. A partir desta foi possível diminuir a complexidade da obtenção de pontos 3D. Este algoritmo é utilizado como parte do software VisualSFM [19] que será utilizado neste trabalho.

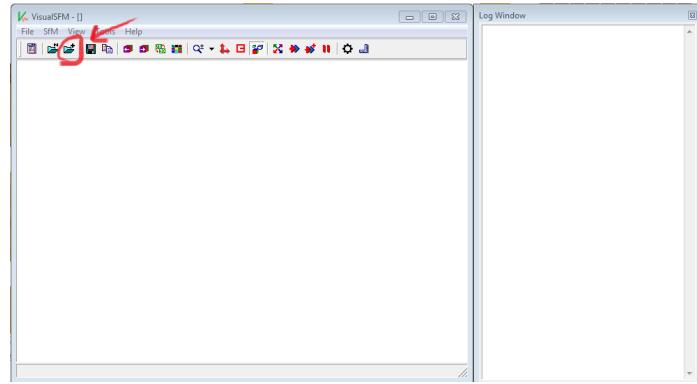
4.2 Software Utilizado

Existe uma gama razoável de softwares que são capazes de obter nuvem de pontos a partir de imagens. Dentre eles pode-se citar o Free-D, Insight3D, PhotoModeler e o VisualSFM. Para este trabalho foi utilizado o VisualSFM [19] por ser um software gratuito, de fácil uso, correspondência automática das imagens e resultados satisfatórios. As desvantagens dos outros softwares citados são as seguintes:

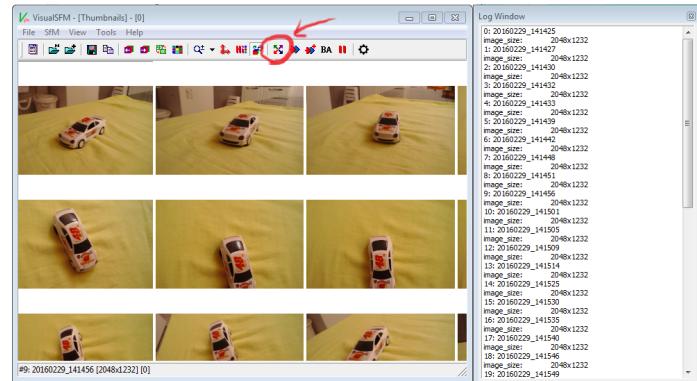
- Free-D: A correspondência das imagens é feita manualmente.
- Insight3d: Ferramenta pouco intuitiva, e de difícil manuseio.
- PhotoModeler: Ferramenta paga.

O VisualSFM recebe como entrada um conjunto de imagens LDR, extrai os pontos de interesse destas imagens e então faz as correspondências entre elas. Uma vez com a correspondência feita, inicializa-se a etapa de reconstrução 3D. Nesta etapa são estimados parâmetros da câmera como: distância focal, e distorção radial, além de estimar, para cada imagem, a posição da câmera.

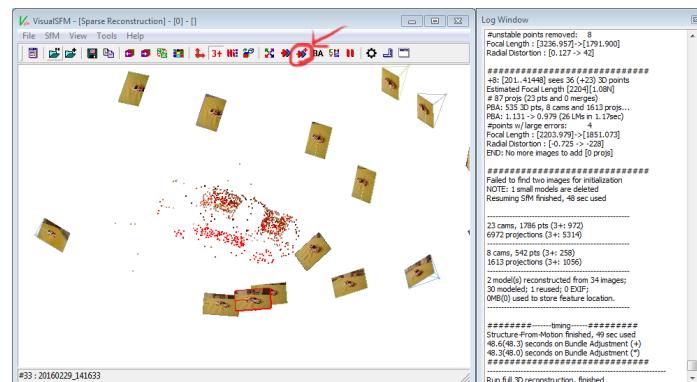
Uma vez com os parâmetros e posições definidas, estima-se as posições dos pontos de interesse que possuem correspondência nas imagens, finalizando a geração da nuvem de pontos. As Figuras 4.1 mostram um exemplo de geração de nuvem de pontos a partir de imagens LDR.



(a) Carrega as imagens LDR.



(b) Encontra a correspondência entre as imagens.



(c) Gera a nuvem de pontos.

Figura 4.1: Exemplo de execução do VisualSfM para a geração de nuvem de pontos a partir de imagens LDR.

4.3 Introduzindo as Imagens HDR ao Processo

Dentre os softwares de geração de nuvem de pontos pesquisados, nenhum possui suporte para leitura e processamento de imagens HDR. Esta problemática exige uma solução alternativa para que seja possível o uso das informações contidas nas imagens HDR para obtenção de nuvem de pontos. Assim como Kontogianni et al. [7] e Přibyl et al. [11], foi decidido pelo uso de imagens HDR processadas com técnicas de tone mapping para geração de uma imagem LDR com mais detalhes e então prosseguir com a obtenção da nuvem de pontos.

4.3.1 Método de Tone Mapping

Como citado na Seção 3.2.1, o Pictureaut possui uma ferramenta mapear tons de uma imagem HDR. Esta funcionalidade é usada para efetuar o tone mapping das imagens HDR com as configurações de filtro bilateral, contraste máximo, e saturação máxima. Essas configurações aumentam a preservação dos detalhes e contraste entre os pixels da imagem processada. A Figura 4.2 exemplifica a configuração do tone mapping de uma imagem.

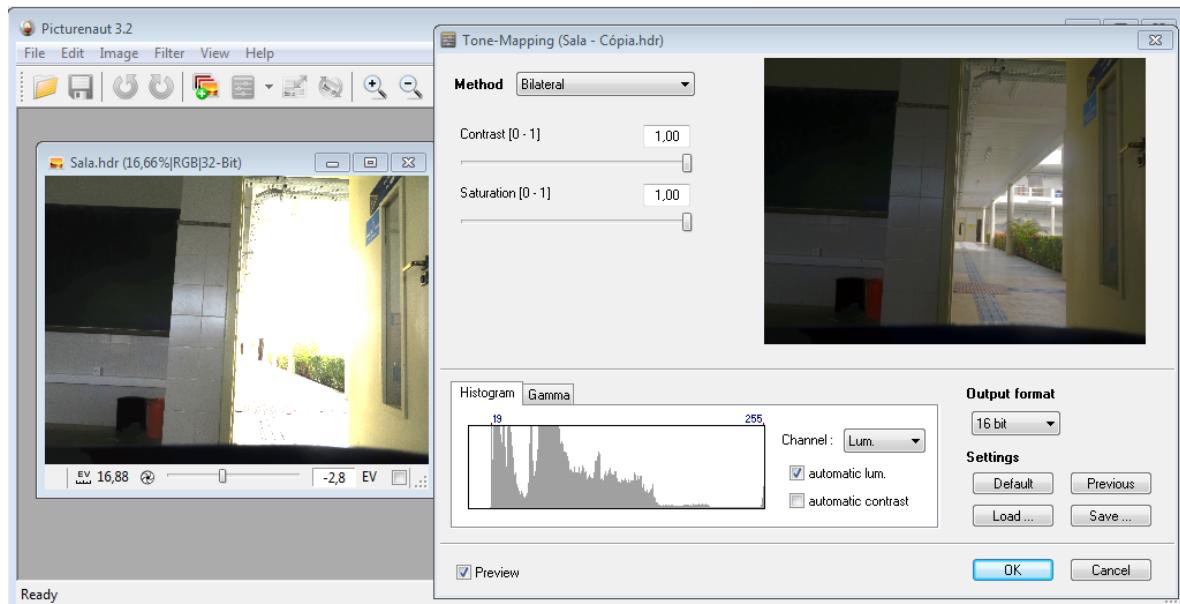


Figura 4.2: Exemplo de imagem HDR tonemapped.

4.4 Proposta para Melhoria na Aquisição da Nuvem de Pontos

Além do uso de imagens HDR tonemapped, outra proposta deste trabalho para melhorar a qualidade e quantidade de elementos da nuvem de pontos gerada, se baseia no processamento de realce de bordas dessas imagens. A hipótese é que com bordas mais acentuadas os algoritmos de identificação de pontos de interesse identificarão mais pontos, com maior qualidade.

Para isso foi proposto o uso do filtro de Laplace, que se trata de um método conhecido em processamento de imagens para realce de bordas. Este método se baseia no conceito da soma das derivadas parciais de segunda ordem [5] para realçar as bordas da imagem. A Equação 4.1 mostra operador de Laplace: [5]

$$\nabla f^2 = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (4.1)$$

Para o processamento digital de imagens essa equação é transformada para a sua forma discreta como mostrado na Equação 4.2. [5]

$$\nabla f^2 = -4f(x, y) + f(x + 1, y) + f(x - 1, y) + f(x, y + 1) + f(x, y - 1) \quad (4.2)$$

Segundo Gonzalez [5], a Equação 4.2 resulta numa máscara que leva em consideração as variações de cores dos pixels vizinhos que formam ângulos múltiplos de 90° em relação ao pixel central. Gonzalez [5] ainda acrescenta as diagonais podem ser adicionadas a máscara resultando na matriz mostrada na Figura 4.3.

-1	-1	-1
-1	8	-1
-1	-1	-1

Figura 4.3: Filtro de Laplace.

Uma vez com as bordas da imagem sendo extraída com o filtro de Laplace. É feito a soma da imagem inicial com as bordas extraídas para que seja obtido o realce das bordas. Neste processo será utilizado um fator α , $0 < \alpha < 1$, para regular a intensidade de realce das bordas. A Equação 4.3 ilustra o processo de realce das bordas.

$$Imagem_{final} = Imagem_{inicial} + \alpha f(Imagen_{inicial}) \quad (4.3)$$

Onde

- f é o filtro de Laplace.

Sendo assim, o processo geral proposto para obtenção da nuvem de pontos a partir de imagens HDR segue as seguintes etapas:

- Geração das imagens HDR a partir de conjuntos de imagens LDR.
- *Tone mapping* das imagens HDR geradas.
- Realce das bordas das imagens *tonemapped*.
- Obtenção da nuvem de pontos a partir das imagens processadas.

4.5 Base de Imagens

Para a execução dos experimentos desta etapa do projeto será utilizada a base de imagens mostrada nas Figuras 4.4, 4.5, 4.6, 4.7 e 4.8. Estas, por sua vez, registram um mesmo objeto em diferentes ângulos num ambiente que possui níveis de iluminação extremos. Para cada ângulo do objeto, foram registradas imagens em diferentes tempos de exposição, para que assim possam ser utilizadas pelos métodos de geração de imagens HDR, e então gerar a nuvem de pontos.

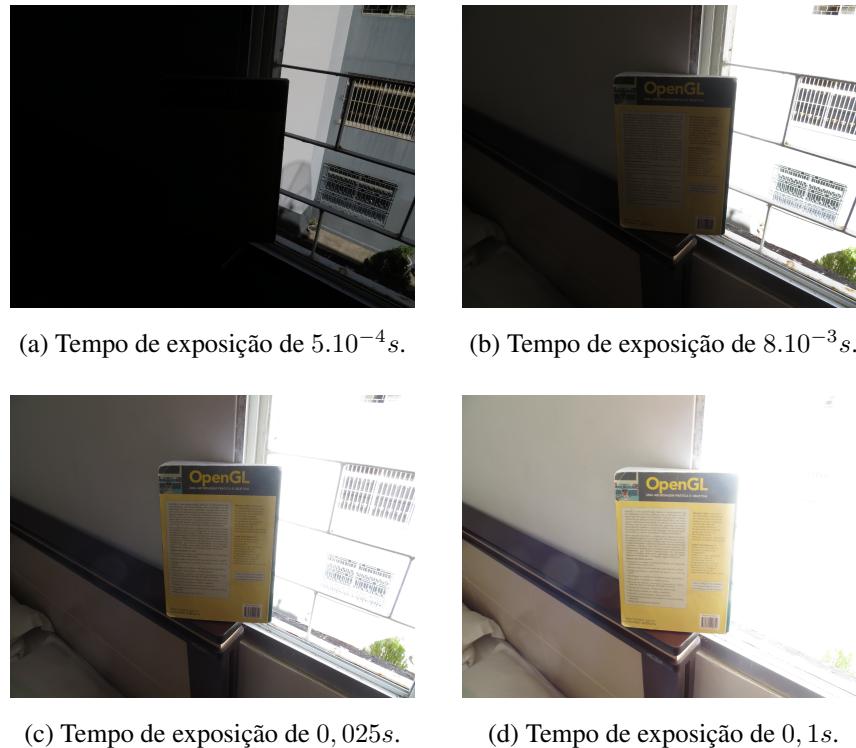


Figura 4.4: Registro em diferentes tempos de exposição da frente do objeto.

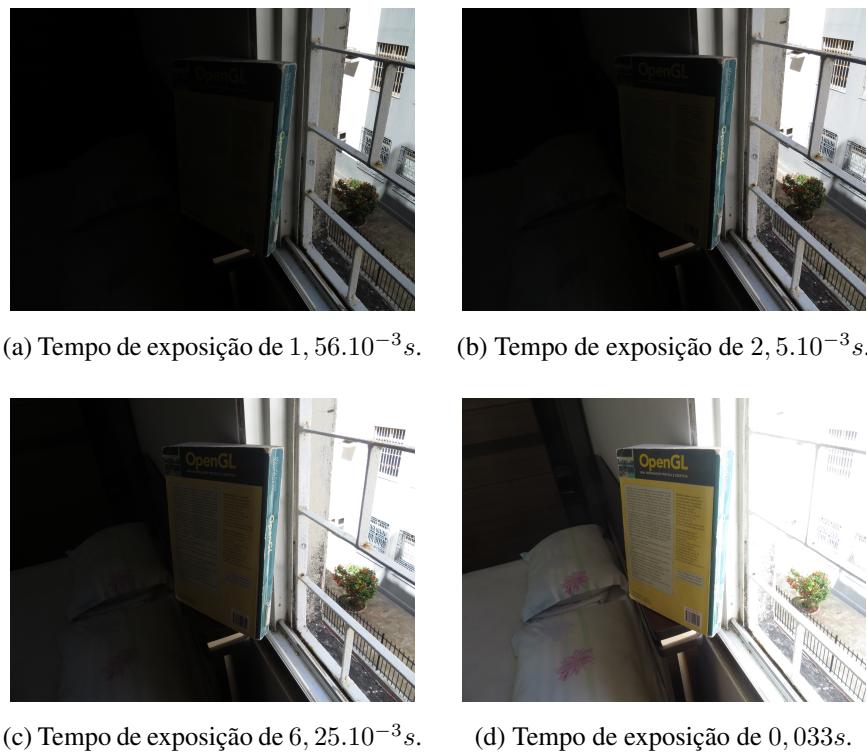


Figura 4.5: Registro em diferentes tempos de exposição da esquerda do objeto.

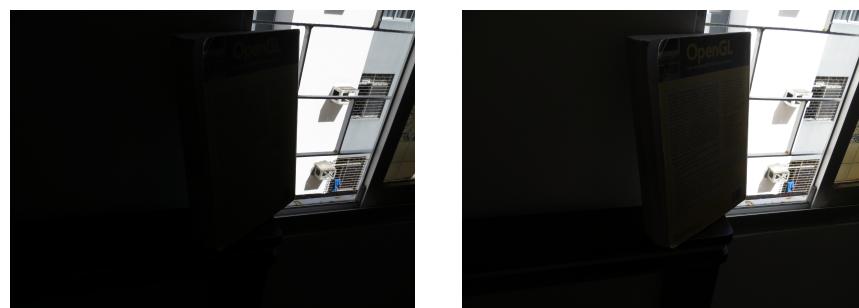
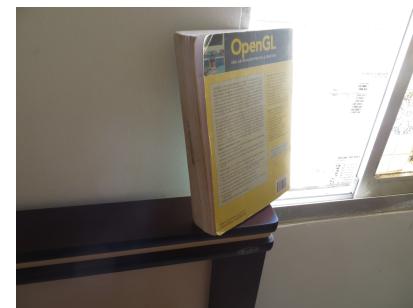
(a) Tempo de exposição de $1,56 \cdot 10^{-3}s$.(b) Tempo de exposição de $5 \cdot 10^{-3}s$.(c) Tempo de exposição de $0,016s$.(d) Tempo de exposição de $0,05s$.

Figura 4.6: Registro em diferentes tempos de exposição da direita do objeto.

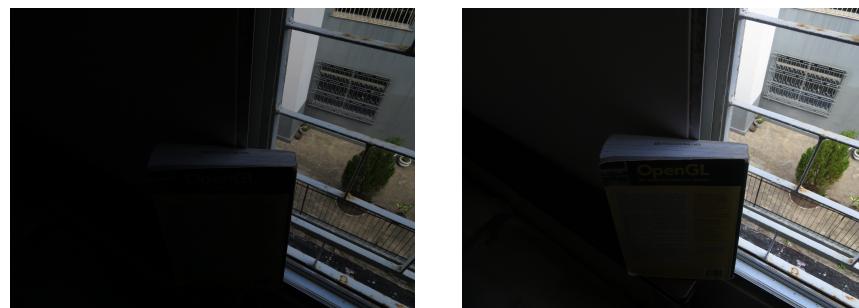
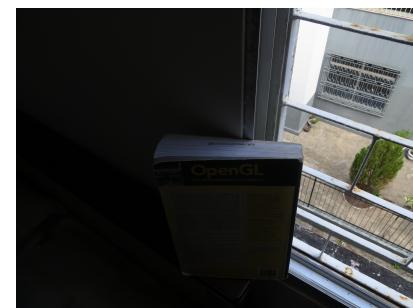
(a) Tempo de exposição de $2,5 \cdot 10^{-3}s$.(b) Tempo de exposição de $0,01s$.(c) Tempo de exposição de $0,04s$.

Figura 4.7: Registro em diferentes tempos de exposição da parte de cima do objeto.

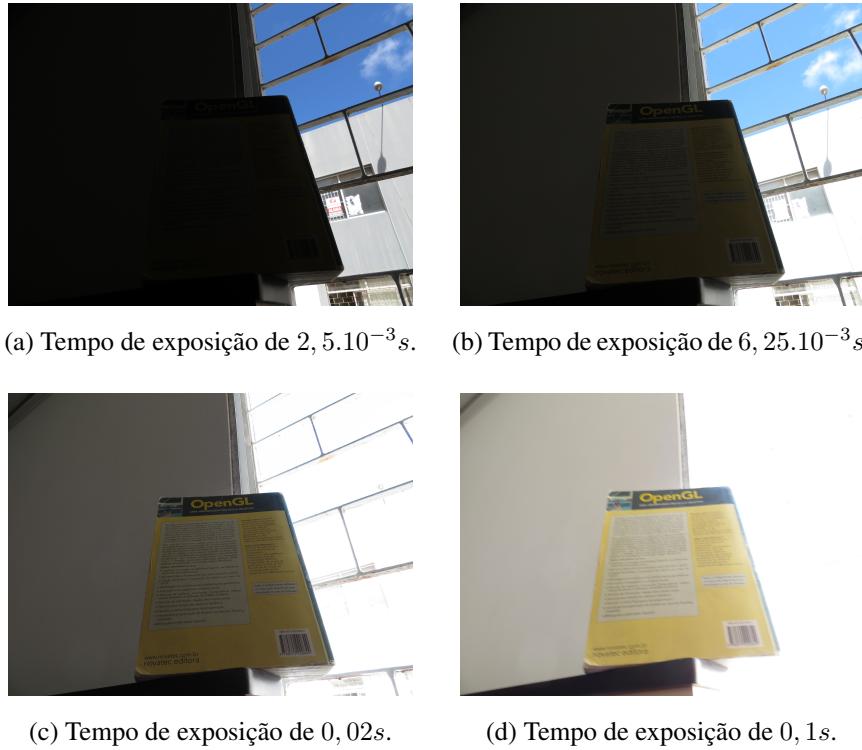


Figura 4.8: Registro em diferentes tempos de exposição do objeto visto por baixo.

Esta base de imagem será utilizada no processo proposto na Seção 4.4. E para fins de comparação as Figuras 4.4b, 4.5c, 4.6c, 4.7c e 4.8c, serão utilizadas no método de obtenção de nuvem de pontos a partir de imagens LDR convencional. Estas foram escolhidas por serem avaliadas visualmente como sendo as mais bem expostas de cada ângulo registrado.

4.6 Resultados e Discussões

Ao aplicar o método de geração de imagens HDR proposto por Sen et al. [15] sobre as imagens da base de dados, notou-se que em algumas delas foram gerados artefatos antes não existentes. A Figura 4.9 mostra a versão *tonemapped* das imagens HDR geradas.

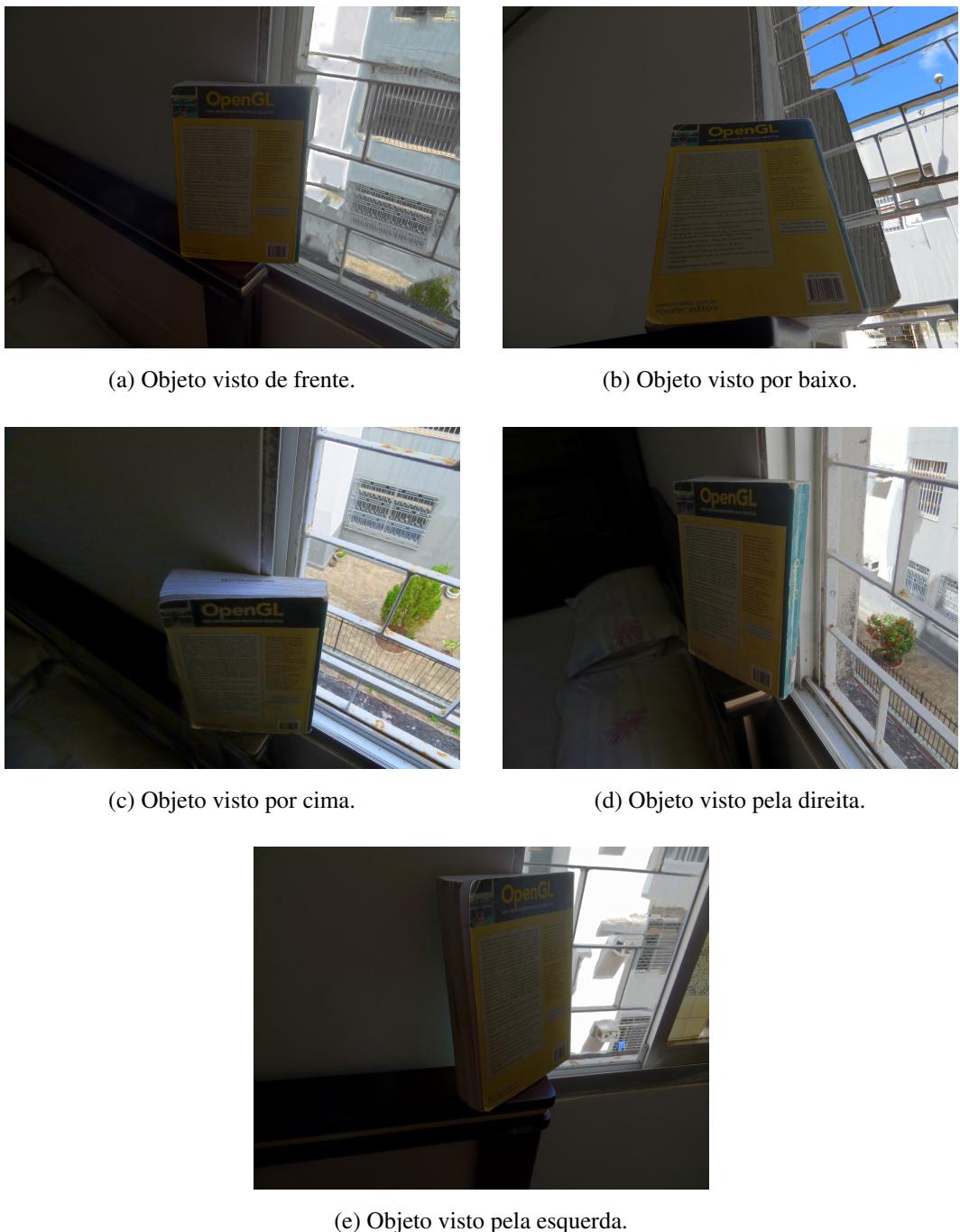


Figura 4.9: Imagens HDR *tonemapped* da base de dados.

É visíveis na Figuras 4.9a, 4.9b, e 4.9e, apresença de distorções e no caso da Figura 4.9b a presença de mais grades que o esperado. Provavelmente estes problemas foram gerados devido ao movimento existente entre as imagens LDR que geraram as respectivas imagens HDR.

Foi feito então a comparação da obtenção de nuvem de pontos utilizando imagens convencionais em relação às imagens *tonemapped* obtidas. Ao gerar a nuvem de pontos utilizando o conjunto de imagens LDR mais bem expostas, citado na Seção 4.5, foi obtido resultado melhor que ao utilizar as figuras HDR ???. A Tabela 4.1 mostra a média de pontos de interesse que foram corretamente correspondidos entre as imagens, assim como o número de pontos 3D obtidos nas nuvens de pontos de cada abordagem. E a Figura 4.10 mostra as nuvens de pontos obtidas.

Tabela 4.1: Resultados obtidos com a geração das nuvens de pontos.

	Média de correspondência dos pontos de interesse por par de imagens	Número de pontos 3D
Imagens LDR convencionais	183,6	1262
Imagens <i>tonemapped</i>	139,3	937



Figura 4.10: Nuvens de pontos obtidas utilizando o software VisualSFM.

Tendo em vista o resultado obtido com as imagens *tonemapped*, optou-se por não dar continuidade ao processo proposto na Seção 4.4. A justificativa disso é que por conta dos artefatos das imagens HDR geradas, ao realçá-las também haverá o realce dessas singularidades, levando a obtenção de pontos de interesse falsos e não úteis.

Capítulo 5

Considerações Finais

Com base nos resultados conclui-se que o uso de imagens HDR obtidas a partir de um conjunto de imagens LDR com diferentes tempos de exposição não se mostra uma alternativa viável para a geração de nuvens de pontos 3D, tendo em vista que as imagens LDR não são capturadas com uso de ferramentas profissionais como tripé.

Os resultados mostraram que as imagens HDR *tonemapped*, que deveriam possuir mais informações do ambiente em relação a uma LDR comum, geraram nuvem de pontos de qualidade inferior à nuvem de pontos obtida utilizando imagens LDR convencionais. O fator que se mostrou limitante para o processo foi a inerente movimentação da câmera entre a captura das imagens LDR, que compuseram as imagens HDR. O fato das imagens não serem alinhadas acabou por gerar artefatos nas imagens HDR o que dificultou o processo de obtenção dos pontos 3D ao invés de melhorá-lo.

Como trabalhos futuros pode-se analisar o processo abordado neste trabalho utilizando-se de ferramentas profissionais que possibilitem a captura de imagens em diferentes tempos de exposição sem que haja movimento entre duas capturas.

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 baroque 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] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing (2nd Ed)*. Prentice Hall, 2002.
- [6] 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.
- [7] 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.

- [8] 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.
- [9] 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 reproduzibilidade das classificações das fraturas da extremidade proximal do úmero? *Revista Brasileira de Ortopedia*, pages 174–177, 2014.
- [10] 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.
- [11] Bronislav Přibyl, Alan Chalmers, and Pavel Zemčík. Feature point detection under extreme lighting conditions. In *Proceedings of the 28th Spring Conference on Computer Graphics*, SCCG '12, pages 143–150, New York, NY, USA, 2012. ACM.
- [12] 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.
- [13] 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.
- [14] Marcia A. Gomes Ruggiero and Vera Lúcia da Rocha Lopes Lopes. *Cálculo Numérico: Aspectos teóricos e computacionais*. Makron Books, 2nd edition, 1996.
- [15] 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.
- [16] D. Simakov, Y. Caspi, E. Shechtman, and M. Irani. Summarizing visual data using bidirectional similarity. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8, June 2008.

- [17] Changchang Wu. Towards linear-time incremental structure from motion. In *Proceedings of the 2013 International Conference on 3D Vision*, 3DV '13, pages 127–134, Washington, DC, USA, 2013. IEEE Computer Society.
- [18] [Online] Pictureonaut. Disponível em: <<http://www.hdrlabs.com/pictureonaut/>>. Acesso em: 14/04/2016.
- [19] [Online] VisualSFM. Disponível em: <<http://ccwu.me/vsfm/>>. Acesso em: 01/05/2016.