



Programa Interdisciplinar de Pós-Graduação em
Computação Aplicada
Mestrado Acadêmico

Marcelo Robson de Azevedo Martins da Silva

Análise Comparativa Entre Suportes para Janelamento na Técnica
Shape From Focus

São Leopoldo, 2017

S586a

Silva, Marcelo Robson de Azevedo Martins da
Análise comparativa entre suportes para janelamento na
técnica Shape From Focus / por Marcelo Robson de Azevedo
Martins da Silva. – 2017.
79 f. : il. ; 30 cm.

Dissertação (Mestrado) — Universidade do Vale do Rio dos
Sinos, Programa de Pós-graduação em Computação Aplicada,
São Leopoldo, RS, 2017.

“Orientadora: Dra. Marta Becker Villamil.”

1. Shape from focus. 2. Focal stack. 3. Agregação. 4. Janela adaptativa.
5. Função de medida de foco. I. Título.

CDU: 004.932

Marcelo Robson de Azevedo Martins da Silva

Análise Comparativa Entre Suportes para Janelamento na Técnica Shape From Focus

Dissertação apresentada à Universidade do Vale do Rio dos Sinos – Unisinos, como requisito parcial para obtenção do título de Mestre em Computação Aplicada.

Aprovado em 27 setembro 2017

BANCA EXAMINADORA

Prof. Dr Luiz Gonzaga da Silveira Júnior – UNISINOS

Prof. Dr^a Adriane Parraga – UERGS

Prof. Dr.^a Marta Becker Villamil (Orientadora)

Visto e permitida a impressão
São Leopoldo,

Prof. Dr. Sandro José Rigo
Coordenador PPG em Computação Aplicada

AGRADECIMENTOS

No final de uma grande caminhada sempre pensamos:

“- Em fim terminei.”

Não podemos de forma alguma esquecer:

“-Terminamos juntos.”

Pois nunca estamos sozinhos.

Agradeço a Deus por estar sempre junto comigo.

Agradeço a minha família por permitir minha dedicação a este trabalho.

Agradeço aos meus professores que me ajudaram e orientaram no esclarecimento dos conhecimentos.

Agradeço aos meus amigos de trabalho que também permitiram minha ausência quando foi necessário.

Agradeço aos meus colegas alunos que trocaram experiências vividas e completaram muitas lacunas de conhecimento

Em fim agradeço a todos que por ventura contribuíram de alguma forma para a conclusão desta caminhada.

RESUMO

Existem muitas técnicas para reconstrução de objetos tridimensionais em computador, algumas são empregadas em ambientes controlados e outras em ambientes que não necessitam de grande precisão. *Shape From Focus* é um método bastante conhecido que utiliza uma pilha de fotografias retiradas com diferentes configurações focais para reconstruir um mapa de profundidade bastante preciso. Este método obtém maior estabilidade na reconstrução de objetos muito pequenos ou microscópios, mas recentemente vem sendo utilizado para reconstrução de ambientes. Com isso, o modelo de reconstrução de mapas de profundidade, *Shape From Focus*, passou a processar maiores quantidades de interferências na pilha de fotografias, como por exemplo, a distorção da lente, o aumento da profundidade de campo, o efeito zoom, entre outros, e também o ruído introduzido pelo ambiente. Este trabalho analisa os efeitos do suporte adaptativo para o janelamento de avaliação do medidor de qualidade de foco do método *Shape From Focus*. Apesar de diferentes trabalhos sobre este tema utilizarem diversas variações do janelamento de avaliação, o suporte adaptativo pode fornecer uma alternativa para encontrar a estabilidade e confiança na obtenção do mapa de profundidade, limitando o erro introduzido por interferências globais.

Palavras chave: *Shape From Focus*, Focal Stack, Agregação, Janela Adaptativa, Função de Medida de Foco.

ABSTRACT

There are many techniques for reconstructing three-dimensional objects in a computer, some are used in controlled environments and others in environments that do not require great precision. *Shape From Focus* is one of the well-known method that uses a stack of cropped photographs with different focal settings to reconstruct a fairly accurate depth map. This method obtains greater stability in the reconstruction of very small objects or microscopes, but has recently been used for reconstruction of environments. As a result, the *Shape From Focus* reconstruction model began to process greater amounts of interference in the photo stack, such as lens distortion, increased depth of field, zoom effect, among others, as well as noise Introduced by the environment. This work analyzes the effects of the adaptive support for the evaluation window of the focus quality meter of the *Shape From Focus* method. Although different works on this theme use several variations of the evaluation window, the adaptive support can provide an alternative to find the stability and confidence in obtaining the depth map, limiting the error introduced by global interferences.

Keywords: *Shape From Focus*, Focal Stack, Aggregation, Adaptive Window, Focus measure function.

SUMARIO

1	INTRODUÇÃO.....	9
1.1	Motivação.....	11
1.2	O problema.....	12
1.3	Objetivos.....	12
1.3.1	Objetivo geral.....	12
1.3.2	Objetivos específicos.....	12
1.4	Contribuição.....	12
1.5	Organização do trabalho	12
2	FUNDAMENTAÇÃO TEÓRICA.....	13
2.1	Construindo um Mapa de profundidade utilizando a focalização.....	13
2.2	Formação da imagem na câmera escura.....	15
2.3	Análise do foco e da profundidade de campo.....	16
2.4	Variações e altas frequências em imagens digitais.....	19
2.5	Os métodos <i>Shape From Focus</i> e <i>Depth From Focus</i>.....	24
2.6	As funções de medida de foco.....	29
2.7	Aglutinando variações no medidor de qualidade de foco.....	33
2.8	Estratégias de reconstrução da profundidade.....	36
2.9	<i>Câmera Calibration</i> para <i>Depth From Focus</i>.....	40
3	TRABALHOS RELACIONADOS.....	45
4	MÉTODO.....	51
4.1	Ferramentas utilizadas	55
4.2	Implementação do algoritmos utilizados.....	55
4.2.1	Método <i>Shape From Focus</i> de Nayar(1989).....	55
4.2.2	Método <i>Adaptive Focus Measure</i> de Aydin e Akgul (2008).....	56
4.2.3	Método <i>Adaptive Depth From Focus</i> de Jing e Yeung(2012).....	56
4.2.4	Método <i>Semi-Global Depth From Focus</i> de Liu e Wu(2015).....	57
4.3	Experimentos e Resultados	57
4.3.1	Descrição dos experimentos.....	57
4.3.2	Resultados.....	58
5	CONCLUSÕES.....	70
	BIBLIOGRAFIA.....	72

LISTA DE FIGURAS

FIGURA 1 -	Formação da imagem na câmera escura	15
FIGURA 2 -	Reflexão do raio de luz parcialmente absorvido pelo objeto	16
FIGURA 3 -	Borramento causado pelos vizinhos do pixel.....	17
FIGURA 4 -	Círculo de confusão em câmeras com lente (2R).....	18
FIGURA 5 -	Imagem desfocada convertida em tons de cinza	21
FIGURA 6 -	Imagem focalizada convertida em tons de cinza	22
FIGURA 7 -	Pilha de foco do mundo real	24
FIGURA 8 -	União de medidores de qualidade de foco em células temporais.....	29
FIGURA 9 -	Medidores de nitidez com sua vizinhança mais próxima no mapa de nitidez.....	30
FIGURA 10-	Tipos de aglutinação de valores em um medidor de qualidade de foco	35
FIGURA 11-	Projeção dos vértices de um plano do quadrilátero em planos focais...	37
FIGURA 12-	Amostras de superfície focalizadas (FIS)	40
FIGURA 13-	Variação de profundidade em cabelos.....	41
FIGURA 14-	Projeção escalar da imagem em câmeras pinhole.....	42
FIGURA 15-	Distorção radial da lente da câmera.....	44
FIGURA 16-	Mapa de profundidade esparso com janelamento não sobreposto.....	45
FIGURA 17-	Mapa de profundidade local com janelamento sobreposto.....	46
FIGURA 18-	Janelamento 3D do medidor de qualidade de foco.....	47
FIGURA 19-	Imagem com mapas de diferentes pesos de profundidade.....	48
FIGURA 20-	Derivadas de duas direções para uma amostra simples.....	49
FIGURA 21-	Problema de oclusão em métodos SFF.....	49
FIGURA 22-	Modelos de janelamento segundo Lee, Shim e Choi	50
FIGURA 23-	Superfícies sintéticas.....	52
FIGURA 24-	Construção de Mapa de profundidade sintético usando Photoshop.....	52
FIGURA 25-	Obtenção da Focal Stack.....	54
FIGURA 26-	Construção da Focal Stack de imagens reais.....	54
FIGURA 27-	Imagem Segmentada usando Mean-Shift-Segmentation.....	64
FIGURA 28-	Mapas de Profundidade Sintéticos.....	64
FIGURA 29-	Reconstrução do Mapa de profundidade da imagem ART.....	66
FIGURA 30-	Reconstrução do Mapa de profundidade da imagem VENUS.....	67
FIGURA 31-	Reconstrução do Mapa de profundidade da imagem REMINDEER.....	68
FIGURA 32	Imagem real toda focalizada com seu Mapa de profundidade.....	69

LISTA DE TABELAS

TABELA 1 - Aplicação do operador Laplaciano 1D na linha 100 da Figura 5.....	22
TABELA 2 - Aplicação do operador Laplaciano 1D na linha 100 da Figura 6.....	23
TABELA 3 - Aplicação do operador Laplaciano 2D na linha 100 da Figura 5.....	23
TABELA 4 - Aplicação do operador Laplaciano 2D na linha 100 da Figura 6.....	23
TABELA 5 - Aplicação do operador Laplaciano Modificado 2D na linha 100 da Figura 5	33
TABELA 6 - Aplicação do operador Laplaciano Modificado 2D na linha 100 da Figura 6	33
TABELA 7 - Amostras com a introdução de ruídos utilizando diversas configurações de ruídos.....	46
TABELA 8 - Erro RMSE para os métodos de reconstrução fina e grosseira.....	59
TABELA 9 - Erro RMSE usando tamanhos variados no janelamento.....	62
TABELA 10- Erro RMSE para superfícies realísticas.....	65

LISTA DE GRÁFICOS

GRÁFICO 1 - Tons de cinza da linha 100 da Figura 5.....	21
GRÁFICO 2 - Tons de cinza da linha 100 da Figura 6.....	23
GRÁFICO 3 - Comparação de medidores de qualidade de foco para autofoco.....	31
GRÁFICO 4 - Interpolação gaussiana para reconstrução fina.....	38
GRÁFICO 5 - Interpolação polinomial para reconstrução fina.....	39
GRÁFICO 6 - Avaliação de medidores de qualidade de foco, para ruído e tamanho da janela.....	47
GRÁFICO 7 Comparação dos métodos de reconstrução Fina e Grosseira.....	60
GRÁFICO 8 Comparação entre os janelamentos e os métodos de reconstrução.....	61
GRÁFICO 9 Comparação do desempenho do tamanho da janela para os medidores.....	63
GRÁFICO 10 Comparação do crescimento da janela para cenas realística.....	65

LISTA DE ABREVIACOES

AFM	Adaptive Focus Measure
ANDF	Anisotropic nonlinear diffusion filtering
CCD	Charge Coupled Device
CPU	Central Process Unit
DFD	Depth From Defocus
DFE	Depth From Focus
DLSR	Digital Single-Lens Reflex
FIS	Focused Image Surface
FM	Focus Measure
GPU	Graphics Process Unit
GRAD	Gradiente
LAP	Laplaciano
MAP	Maximun a posteriori
MRF	Markov Random Field
Pinhole	Cmera sem lente/cmera de orifcio
RAD	Rapid Aplication Development
SFD	Shape From Defocus
SFF	Shape From Focus
SIMD	Single Instruction Multiple Data
SSE	Streaming SIMD Extensions

1 INTRODUÇÃO

Atualmente, os sistemas de mapeamento tridimensional têm obtido grande ênfase para resolução de problemas. Muitos desses sistemas de mapeamento tridimensional utilizam informações do mundo real para reconstruir objetos, como por exemplo: inspeção de produtos industriais para controle de qualidade em linha de produção (RIBNICK, 2012), inspeção de minérios (GAGANOV; IGNATENKO, 2009), inspeção de produtos petroquímicos (NIEDEROST, M.; NIEDEROST, J.; SCUCKA, 2003), digitalização de objetos para modelar cenários de computação gráfica (HASINOFF; KUTULAKOS, 2009). Outra utilidade desses sistemas de mapeamento é a reconstrução de cenas em computação gráfica (AYDIN; AKGUL, 2008).

A reconstrução tridimensional de um objeto é realizada pela representação das medidas geométricas do objeto, no sistema de coordenadas do R^3 (FORSYTH; PONCE, 2002, p.120-122), em um mapa construído sobre uma matriz 3D, onde os componentes de altura, largura e profundidade, são representados pelos eixos de coordenadas ortogonais x , y e z , respectivamente. Esse mapa é construído com valores que são obtidos de amostras de um sinal de imagem digital, fornecido por um dispositivo sensível a radiações do ambiente, e processado por um algoritmo de extração de profundidade.

A obtenção desse sinal digital utiliza dois tipos de sistemas de captura de imagem: sistemas com sensores ativos e sistema com sensores passivos (MACHI; IKEOKA; TAKAYUKI, 2012, p.749). Os sistemas de captura de imagem são caracterizados pela forma com que o sensor de captura é sensibilizado. Um sistema com sensor ativo possui um sensor de imagem e um elemento ativo que produz um sinal luminoso. O sinal luminoso, que foi produzido, será refletido pelo ambiente/objeto, capturado pelo sensor de imagem e analisado por um algoritmo de reconstrução 3D (MACHI; IKEOKA; TAKAYUKI, 2012, p.749). Os sistemas que utilizam sensores passivos, não necessitam do componente que produz o sinal luminoso, pois capturam o sinal já presente no ambiente. Os sistemas com sensores passivos geralmente utilizam as características de um conjunto de imagens para identificar regiões geometricamente semelhantes, onde a projeção do movimento permite encontrar a forma do objeto (FORSYTH; PONCE, 2002, c.4). Outro tipo de reconstrução 3D, que utiliza sensores passivos, são aquelas que obtém mapas de profundidades, onde as coordenadas z são extraídas de um conjunto de imagens retiradas com diferentes configurações da lente do dispositivo de captura de imagem (GROSSMANN, 1987).

Uma imagem digital pode ser definida como uma matriz bidimensional obtida como resposta às amostras discretas de estímulos de intensidade luminosa nos sensores, em

coordenadas espaciais, amostradas de forma, também, discreta. Se os valores das intensidades luminosas da imagem real, forem representados por uma função imagem, contínua, e as coordenadas espaciais pelo par (s, t) , onde s , é a quantidade de linhas e t , a quantidade de colunas, então a imagem digital com coordenadas discretas amostradas e reordenadas em coordenadas matriciais (x, y) pode ser representada pela formula descrita na Equação 1. Cada elemento desta matriz corresponde a um ponto exibível na tela de um monitor de vídeo que é chamado de pixel (GONZALEZ; WOODS, 2010, p.33-34).

$$I(x,y) = f(s,t) \quad (1)$$

Onde: f representa a imagem digital em forma de função bidimensional,

I representa a imagem digital em forma de matriz bidimensional,

x e y são as coordenadas de linha e coluna na matriz bidimensional,

s e t são as coordenadas espaciais na função bidimensional.

Segundo Gonzalez e Woods(2008), a abordagem de matriz bidimensional fornece apenas o entendimento da composição das dimensões da imagem, e não representa as inúmeras possibilidades de representação de cores da imagem (GONZALEZ;WOODS, 2008). Neste trabalho foi utilizado o valor de tons de cinza obtido das intensidades de cores da imagem, calculando a média dos valores do padrão RGB.

Para obtenção do mapa de profundidade duas abordagens baseadas na variação do foco podem ser utilizadas: focalização e desfocalização (MICHII; IKEOKA; TAKAYUKI, 2012, p.749; MAHMOOD; CHOI, 2012, p.2866; JING; YEUNG, 2012, p.1205). No método de desfocalização (*Depth From Defocus* ou *Shape From Defocus*) é utilizada a proporcionalidade do borramento, presente em toda a imagem, para calcular o mapa de profundidade (MOELLER; et al., 2015, p.5369; JING; YEUNG, 2012, p.1205). No método de focalização (*Depth From Focus* ou *Shape From Focus*) é utilizado um algoritmo de extração de altas frequências da imagem para achar a profundidade no conjunto de imagens (NAYAR; NAKAGAWA, 1994, p.825). Em aplicações que exijam precisão o algoritmo *Depth From Focus* (DFF) é mais indicado, pois analisa todas as distâncias, buscando a profundidade no pixel (NAYAR; NAKAGAWA, 1994, p.826; GAGANOV; IGNATENKO, 2009).

Este tipo de mapeamento não pode ser aplicado na reconstrução de cenas dinâmicas, em que existe o movimento, portanto são exemplos de aplicações que não são suportadas por esta técnica: controle de proximidade e controle de veículos (MICHII; IKEOKA; TAKAYUKI, 2012), conversão de vídeos em vídeos 3D ou mapeamento de cenas com várias fotografias em cenas dinâmicas (FORSYTH; PONCE, 2002).

O principal requisito desse tipo de reconstrução é a utilização de cenários estáticos (GAGANOV; IGNATENKO, 2009; JING; YEUNG, 2012, p.1205), sendo a qualidade da reconstrução proporcional à quantidade de imagens e as configurações utilizadas no sensor de captura.

No método *Shape From Focus* (SFF), cada imagem é submetida a uma operação de extração de variações de intensidades, chamada de função de medida de foco ou medidor de qualidade de foco. Este algoritmo utiliza os vizinhos do pixel para calcular a sua taxa de variação, chamada medidor de nitidez. Um medidor de qualidade de foco é obtido pela soma de um ou mais medidores de nitidez. Para cada medidor de qualidade de foco é utilizado um medidor de nitidez com nova vizinhança de região quadrática e que tem como centro o medidor que se deseja obter a variância total. Dentre os principais medidores de nitidez podemos citar o operador Laplaciano e o operador Gradiente. O cálculo do medidor de foco é obtido pela soma dessa região. Segundo Gonzalez e Woods (2008), uma região de imagem é um subconjunto dos pixels da imagem digital, onde cada pixel possui pelo menos um vizinho dentro do mesmo subconjunto, então uma região de um conjunto qualquer é um subconjunto onde os elementos que o compõem possuem pelo menos um vizinho em comum. Quando é utilizada uma função avaliadora de pixels, seu alcance fica restrito segundo um determinado critério, e esse conjunto de sub-amostras que é fornecido para essa função de medida de foco é chamado de suporte. Este trabalho analisa os algoritmos de suporte adaptativo da janela de avaliação do método *Shape From Focus*.

Sendo o medidor de qualidade de foco uma função matemática de soma de variações e o suporte de uma função matemática, o conjunto de valores fornecidos a ela, então qualquer forma de seleção ou modificação desse conjunto de valores, segundo algum critério(diferente da soma), pode ser entendido como adaptação a este critério(adaptativo), mesmo que ocorra a exclusão de valores do medidor.

Os algoritmos de extração de qualidade de foco desse tipo de mapeamento possibilitaram o desenvolvimento e aplicação sobre esse conhecimento em outras áreas de estudo. Segmentação de imagens (JING; YEUNG, 2012, p.1205), fusão da pilha de foco para formar uma imagem de melhor qualidade e a técnica de fotografia HDR são exemplos desse tipo de aplicação que utilizam este tipo de conhecimento.

1.1 Motivação

Os medidores de qualidade de foco usam uma janela de avaliação, somando uma região quadrática de medidores de nitidez(medidores da variação do pixel, medidores de contraste, etc). Este tipo de agregação não leva em consideração as características de cada plano focal,

nem mesmo a característica de cada região do objeto/imagem. Os medidores de qualidade de foco com suporte adaptativo em geral tentam encontrar as características de cada objeto e por isso levem em conta principalmente as características de um ponto específico da cena. Pertuz e Garcia(2013), realizaram um estudo sobre os medidores de qualidade de foco com região quadrática. Este trabalho é um estudo semelhante, porém usando janelamento com suporte adaptativo para o medidor de qualidade de foco.

1.2 O problema

Embora se tenha conhecimento de medidores que adequem-se as características de cada pixel da imagem, não foram encontrados trabalhos que comparem a qualidade desses medidores de qualidade de foco com suporte adaptativo.

1.3 Objetivos

1.3.1 Objetivo geral

Analisar e comparar os suportes para janelamento da técnica *Shape From Focus* .

1.3.2 Objetivos específicos:

- a) Estudar o método *Shape From Focus* e seus sub-algoritmos.;
- b) Estudar o Janelamento com suporte adaptativo.
- c) Avaliar os benefícios de quatro tipos de suporte para o método *Shape from Focus*

1.4 Contribuição

Este trabalho fornece um estudo comparativo dos medidores de qualidade de foco com suporte adaptativo, sob os mesmos parâmetros utilizados para comparação dos medidores de qualidade de foco quadrático, sendo concentrado em quatro métodos.

1.5 Organização do trabalho

Esta dissertação esta organizada da seguinte maneira: O capítulo dois apresenta os principais conceitos utilizados no trabalho, o capítulo três enumera os trabalhos relacionados com o problema de aglutinação de valores ao medidor de qualidade de foco. O capítulo quatro apresenta o método proposto para análise, bem como os resultados encontrados, por fim no capítulo cinco são apresentadas as conclusões.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são apresentados os princípios básicos de focalização usados na construção do mapa de profundidade, bem como o processo de formação da imagem que possibilitam a construção do método SFF, e de seus subalgoritmos.

Também é apresentado um estudo dos subalgoritmos do método SFF, abordando os princípios para realizar o foco parcial em imagem digitais e como isso pode possibilitar construir um mapa de profundidade.

Nas seções 2.1 e 2.4, é analisado o processo de formação da imagem na etapa de aquisição, caracterizando os componentes do mundo real que permitem a obtenção controlada de imagens digitais. Na seção 2.3 são apresentados dois importantes fenômenos, a profundidade de campo e o ponto focal de uma lente, que são bastante empregados nos métodos de extração de mapa de profundidade (SFF, DFF, DFD e SFD), e que ocasionam o borramento em uma imagem. Na seção 2.4 é definido como quantificar matematicamente os efeitos de borramento da imagem que foram descritos na seção 2.3.

Nas seções 2.5 a 2.9, são apresentados os métodos SFF e DFF, bem como seus subalgoritmos. Na seção 2.6 e 2.7, é descrito o medidor de qualidade de foco do método SFF e realizado um estudo do janelamento e dos diversos métodos de aglutinações de variações no medidor de qualidade de foco. Na seção 2.8, são descritos ainda o sub-algoritmo de estratégia de reconstrução do método SFF. Por fim na seção 2.9, é tratado de um processo de calibração de câmera para uma aplicação específica do método SFF, quando empregado lentes com grande profundidade de campo.

2.1 Construindo um Mapa de profundidade utilizando a focalização

A utilização dos princípios básicos de focalização é um dos recursos que podem ser utilizados para realizar a reconstrução tridimensional, através da montagem de um mapa de profundidade. As duas formas conhecidas para extração de profundidade pela manipulação desses princípios de focalização possuem características diferentes e são descritas a seguir.

Uma fotografia de baixa qualidade é adquirida quando é utilizada uma configuração inadequada da lente na máquina fotográfica. Esta configuração inadequada reproduz os mesmos mecanismos utilizados pelo sistema de visão humana e que também são utilizados para reconhecer as distâncias e formas de objetos.

Se uma câmera fotográfica capturar uma foto estando configurada com uma grande abertura da lente, a profundidade de campo limitará a nitidez sobre um ponto de focalização na imagem (GROSSMANN, 1987, p.66). Se a distância focal é traduzida dentro da imagem então essa dica pode ser utilizada para extrair pontos da imagem correspondentes aos planos de

focalização paralelos a sua lente (GROSSMANN, 1987, p.63-67; MOELLER; ET AL., p.5369; SUBBARAO; SURYA, 1994, p.271).

A utilização dos princípios básicos de focalização inicialmente foram usados em algoritmos de autofoco. Horn (1968) utilizou a transformada de Fourier para medir e obter as variações nas intensidades luminosas da imagem, relacionando altas frequências ao conteúdo focalizado. Sua tese de doutorado também sugeriu uma possível utilização desta medida para estimar distâncias.

A transformada de Fourier exigia muitos cálculos computacionais e inviabilizava aplicações práticas (KROTKOV, 1987, p.226). Na década de 70 alguns trabalhos sugeriram sua substituição por outros medidores em aplicações de autofoco. Tenenbaum (1970) propôs o operador Tenengrad aplicando dois detectores de bordas Sobel para obter a magnitude do gradiente em cada ponto da imagem. Kujooory, Mayall e Mendelsohn (1973) propuseram a maximização do valor das intensidades de luminância como medidor de foco. Jarvis (1976) propôs outros medidores de focalização como a soma da diferença dos módulos, a entropia do histograma, variação local do histograma e a variância dos níveis de cinza.

Até a década de 80, os trabalhos acadêmicos apenas analisavam o posicionamento focal da lente da câmera fotográfica, ou estimavam a distância focal da lente. Esses métodos utilizavam um medidor de energia para cada imagem, e posteriormente selecionavam o maior valor absoluto entre eles, levando em consideração a mesma localização para todas as imagens. Neste período, vários artigos acadêmicos como Krotkov (1987), Krotkov e Martin (1986), analisaram exaustivamente as respostas desses medidores de nitidez em muitas aplicações que objetivavam a focalização.

Também no final da década de 80, o artigo de Pentland (1987), propôs utilizar a medida de dispersão de ponto como métrica de estimativa de profundidade de cena, criando a base do método que mais tarde se chamaria DFD. Paralelamente, o artigo de Grossmann (1987), também propôs a extração de profundidade de cena, porém usando os medidores de focalização, com seu algoritmo DFF. Esses dois artigos criaram uma nova perspectiva de reconstrução tridimensional sobre imagens com diferentes configurações focais.

Esses dois tipos de mapeamento de profundidade, considerando-se os princípios básicos de focalização, diferenciam-se principalmente na etapa de obtenção dos dados. Enquanto que o algoritmo DFF utiliza a comparação de medidores de foco retirados de várias imagens, o algoritmo DFD utiliza uma função de espalhamento de foco, para estimar a profundidade de cena do borramento produzido pelo desfocalização, em um pequeno número de imagens (GAGANOV; IGNATENKO, 2009).

Os dois modelos de mapeamento baseado em focalização também se diferenciam de outros que utilizam imagens para reconstrução 3D, por não necessitarem de imagens nítidas e não utilizarem algoritmos de caracterização de pontos chaves na imagem (FORSYTH; PONCE, 2002; SCHECHNER; KIRYATI, 2000).

Embora as bases da reconstrução tridimensional por pilha de foco tenham sido propostas como mapa de profundidade, a primeira aplicação desse tipo foi realizada por Darrell e Wohn (1988), usando o operador Laplaciano e o operador Gaussiano em pilha para medir a nitidez do pixel.

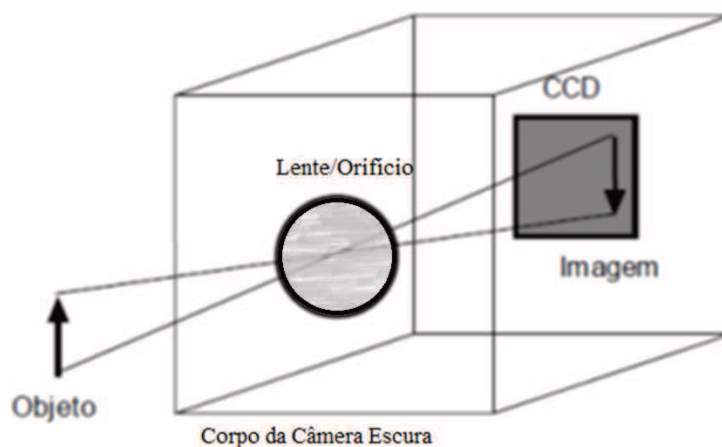
2.2 Formação da imagem na câmera escura

Uma câmara escura é um espaço interior de um compartimento fechado que pode ser, por exemplo, um quarto fechado, uma caverna, uma caixa ou mesmo o interior de uma lata (UFMG,1999).

A formação de uma imagem depende da quantidade de radiância refletida por um objeto (HORN, 1970, p.19-21), pois toda imagem é formada por valores de intensidades amostradas da quantidade de energia da luz (GONZALEZ; WOODS, 2012, pg 29). Se a imagem é formada pela luz visível então sua captação será perceptível em tonalidades de cores e essas intensidades de cor, capturadas pelo sensor, são chamadas de luminâncias.

Para capturar uma imagem é necessário um conjunto de sensores de luminosidade, arranjados em forma de matriz bidimensional (FILHO; NETO, 1999, p.226), em um dispositivo, que geralmente é um Charge Coupled Device (CCD), dentro de uma câmara escura com um orifício, possuindo uma lente ou não, no lado oposto a essa matriz, conforme a Figura 1.

FIGURA 1 - Formação da imagem na câmera escura

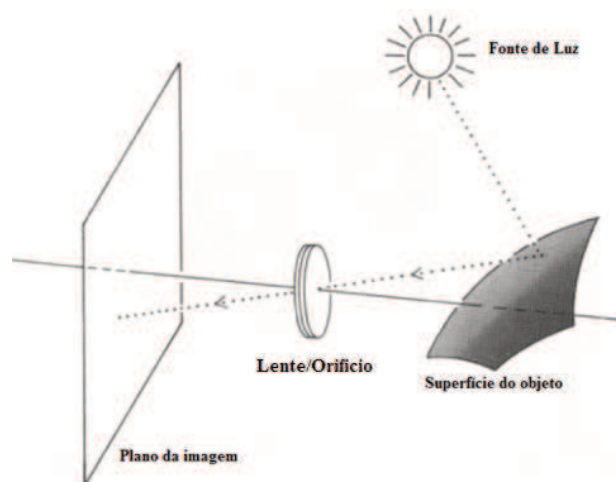


Fonte: Adaptado de FILHO e NETO.(1999, p.21).

A formação da imagem acontece quando os feixes luminosos, refletidos por uma superfície de um objeto, atravessam o orifício da câmera escura e alcançam individualmente os

sensores de luminosidade da matriz de sensores do dispositivo CCD, que se encontra no plano da imagem (NAYAR; NAKAGAWA, 1994, p.824; MAHMOOD; CHOI, 2012, p.2866), conforme a Figura 2. A produção da imagem acontece porque os feixes luminosos seguem uma trajetória retilínea e, também, porque parte da energia da fonte de luz do ambiente não é absorvida pela superfície do objeto, quando o raio de luz é refletido (HORN, 1970, p.2).

FIGURA 2 - Reflexão do raio de luz parcialmente absorvido pelo objeto



Fonte: Adaptado de Horn. (1986, p. 8)

A câmera que não possui uma lente sobre o orifício é chamada de câmera pinhole e a quantidade de borramento da imagem é determinado pelo tamanho do orifício da câmera escura (UFMG, 1999).

As características físicas de uma câmera pinhole tornam difícil a construção aceitável de uma pequena câmera que balanceie a sensibilidade do sensor, o tamanho do orifício da câmera escura e um bom ângulo de visualização do ambiente. Então, a miniaturização dessa câmera é realizada pela adição de uma lente convexa para concentrar os feixes de luz sobre o raio da circunferência da lente.

Todos os sistemas de visão do mundo real que utilizam a luz visível precisam de uma câmera escura e um conjunto de sensores em forma de matriz, para obter imagens.

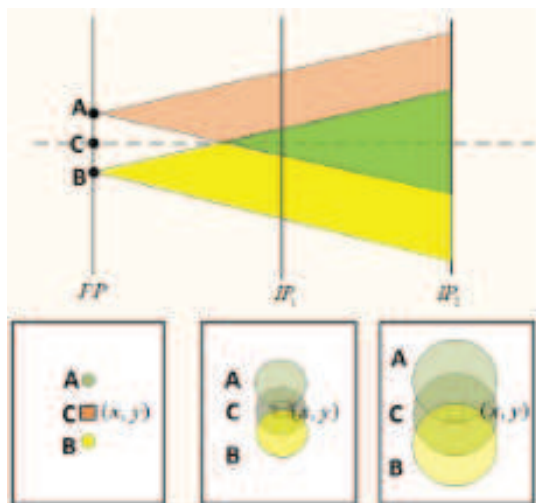
2.3 Análise do Foco e da Profundidade de campo

O processo de fotografar pode sofrer de inúmeras interferências e como o principal componente que permite a captura de uma fotografia é a intensidade luminosa, então todos os agentes que modificam a trajetória dos raios de luz ou diminuem a quantidade de luminosidade na câmera escura, também degradarão a qualidade da foto.

Quando o orifício da câmera escura é limitado por algo, também é limitando a quantidade de luz que será absorvida pelos sensores ópticos. Por outro lado, se aumentar a

quantidade raios de luz, por uma grande abertura, feixes oriundos de diferentes regiões atingirão um mesmo sensor, produzirão uma imagem borrada (NAYAR; NAKAGAWA, 1994, p.827).

FIGURA 3 - Borramento causado pelos vizinhos do pixel



Fonte: Zhang et al. (2014)

A Figura 3 mostra o efeito do borramento ocasionado pela mistura de vizinhanças de um pixel quando a imagem é formada fora do plano focal (FP). Considere uma câmera escura em que três feixes de luz cruzam o orifício de entrada. Em uma determinada posição eles passam por um plano onde formam uma imagem nítida, esse plano é chamado Plano Focal (FP), a medida que se afastam dele e alcançam outros planos de imagem (IP1 e IP2) formam imagens com maior dispersão do borramento. É possível observar que o borramento é produzido tanto antes quanto depois desse plano focal.

Em uma câmera pinhole, quanto maior for o orifício da câmera mais borramento será produzido sobre a imagem, sendo distribuído igualmente em toda a imagem. Entretanto se for acrescentado uma lente convexa ao sistema, os raios sofrerão um desvio de trajetória, se direcionando para proximidades do centro da circunferência da lente. Nesta situação, o raio da circunferência da lente fornecerá a distância focal no qual o objeto se tornará nítido. Na Fórmula abaixo, a equação de Gauss estabelece a distância focal para uma lente fina. (AHMAD; CHOI, 2005, p.566).

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v} \quad \text{Onde: } f - \text{distância focal da lente} \quad (2)$$

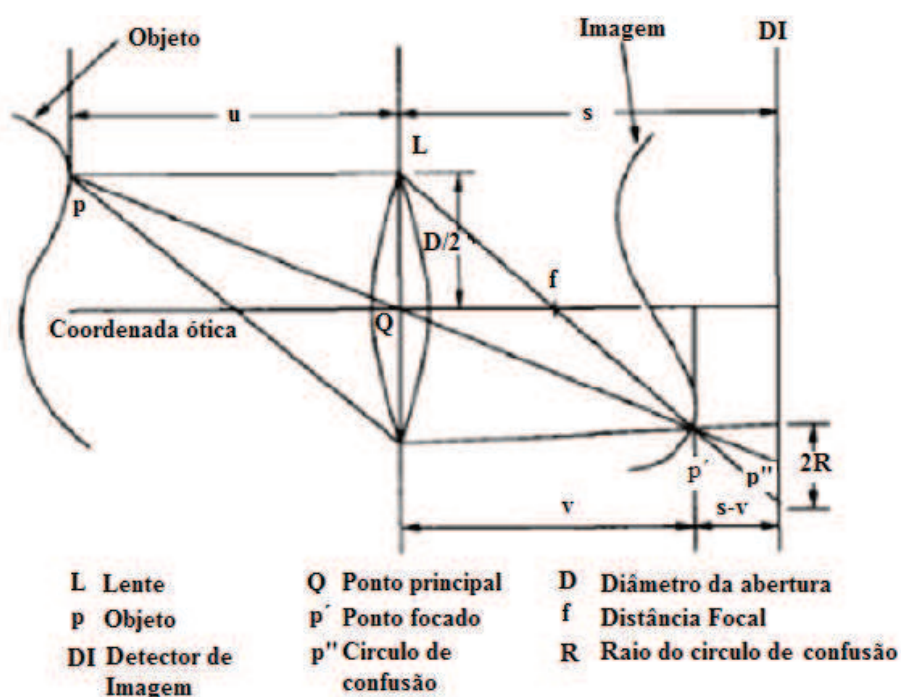
u - distância do objeto
v - distância do plano da imagem

Com a introdução do ponto focal, os diversos raios de luz refletidos por um objeto são concentrados sob um pixel. Os raios que passam pelo centro da lente (ponto Q na Figura 4), são chamados de raios principais, e definem a estrutura básica da formação da imagem. Para que a

imagem esteja nítida, cada raio de luz que reflete de um ponto do objeto do mundo real, deve alcançar somente um pixel da imagem (MAHMUD; CHOI, 2012, p.1057).

Se a matriz de sensores estiver fora da posição determinada pela distância focal, os raios luminosos que irão compor um pixel se cruzarão antes ou depois do sensor que captará o raio principal e se espalharão sobre outros sensores da imagem, interferindo em suas captações e produzindo um borramento naquela região da imagem. Este efeito é conhecido como círculo de confusão (GROSSMANN, 1987, p.64; KROTKOV, 1987, p.234), e pode ser visualizado no plano DI da Figura 4, como um círculo de raio $2R$.

FIGURA 4 - Círculo de confusão em câmeras com lente ($2R$)



Fonte: Adaptado de Subbarão e Choi (1995, p.266).

O alcance do círculo de confusão fornecerá a profundidade de campo que será introduzida na imagem (KROTKOV, 1987, p.235). Quando o círculo de confusão é grande o suficiente para que um raio de luz interfira na captura de outro(s) raio(s) de luz(es) que seja(m) seu(s) vizinho(s) então ocorrerá o desfoque, senão é dito que há um desfoque aceitável (HORN, 1986, p.26).

Em resumo, na câmera pinhole o excesso de feixes luminosos alcança alguns dos sensores da matriz bidimensional em paralelo e, em câmeras com lentes os raios de luz em excesso irão se cruzar e alcançar uma região com vários sensores.

Portanto a profundidade de campo em uma câmera com lente é produzida pela abertura da lente, de maneira que quanto menor a profundidade de campo menos porção do espaço físico ficará nítido na imagem.

2.4 Variações e altas frequências em imagens digitais

O desfoco de uma imagem digital consiste na atenuação das altas frequências espaciais (KROTKOV; MARTIN, 1986, p.1094). A consequência de atenuar altas frequências na imagem causa o aparecimento de borrões. Segundo Krotkov (1987, p.227) o borramento em uma imagem digital é associado às pequenas variações de intensidades de cor. Essas variações de intensidade são parâmetros que permitem construir um mapa que quantifique a nitidez da imagem em taxas de variações.

Embora Horn (1968) tenha utilizado o domínio das frequências para posicionar a lente e focalizar uma cena, a conversão de uma imagem digital para o domínio das frequências mistura informações visuais no domínio das frequências. O aspecto visual do domínio das frequências é conhecido como espectro de Fourier e é calculado somando todos os termos da função imagem multiplicada por um coeficiente. Assim, não é possível medir a qualidade de pequenas regiões com foco e construir um mapa de nitidez com esse tipo de medição. As frequências em imagens digitais são variações de intensidades de cor e as taxas de variação podem ser obtidas pela aplicação das derivadas sobre a imagem digital. A diferença entre a intensidade de cor do pixel corrente, $f(x,y)$, e o consequente, $f(x,y+1)$, fornece a primeira derivada de um pixel da linha da função imagem, contínua, f , e pode ser descrita como a aplicação do operador gradiente sobre a imagem, conforme abaixo.

$$GRAD\ 1D = f(x,y) - f(x,y+1) \quad (3)$$

Onde: f é a função bidimensional que representa as intensidades de cor da imagem.

$x, y, y+1$ são a linha e as colunas dos pixels, respectivamente.

$GRAD\ 1D$ é o operador gradiente para uma dimensão.

Como a função imagem é bidimensional é necessário somar o gradiente da outra coordenada, conforme o seguinte:

$$GRAD\ 2D = f(x,y) - f(x,y+1) + f(x,y) - f(x+1,y+1) \quad (4)$$

$$GRAD\ 2D = 2f(x,y) - f(x,y+1) - f(x+1,y+1)$$

Onde: f é a função bidimensional que representa as intensidades de cor da imagem.

$x, x+1$ são as linhas dos pixels.

$y, y+1$ são as colunas dos pixels.

$GRAD\ 2D$ é o operador gradiente para duas dimensões.

Darrell e Wohn (1988) utilizaram o operador Laplaciano para obter as variações nos pixels. A utilização do operador Laplaciano consiste na aplicação da segunda derivada sobre os tons de cinza da imagem digital, e pode ser obtido pela soma de duas taxas de variações, que

são calculadas pela subtração da intensidade de cinza do pixel corrente, $f(x,y)$, pelo antecedente, $f(x,y-1)$, e outra do pixel corrente, $f(x,y)$, pelo conseqüente, $f(x,y+1)$, conforme o seguinte:

$$LAP\ 1D = f(x,y) - f(x,y+1) + f(x,y) - f(x,y-1) \quad (5)$$

$$LAP\ 1D = 2 \cdot f(x,y) - f(x,y+1) - f(x,y-1)$$

Onde: f é a função bidimensional que representa as intensidades de cor da imagem.

$y-1$, $y+1$ colunas antecedentes e conseqüentes do pixel.

x , y são a linha e coluna do pixel.

$LAP\ 1D$ é o operador Laplaciano para uma dimensão

Para aplicar esse operador em uma imagem somamos o operador da outra coordenada, neste operador, conforme abaixo.

$$LAP\ 2D = 2 \cdot f(x,y) - f(x,y+1) - f(x,y-1) + \quad (6)$$

$$2 \cdot f(x,y) - f(x+1,y) - f(x-1,y)$$

$$LAP\ 2D = 4 \cdot f(x,y) - f(x,y+1) - f(x,y-1) - f(x+1,y) - f(x-1,y)$$

Onde: f é a função bidimensional que representa as intensidades de cor da imagem.

$y-1$, $y+1$ colunas antecedentes e conseqüentes do pixel, respectivamente.

$x-1$, $x+1$ linhas antecedentes e conseqüentes do pixel, respectivamente.

$LAP\ 2D$ é o operador Laplaciano para duas dimensões.

Se a variância, em imagens digitais, é obtida pela comparação dos valores dos pixels adjacentes ao atual e, o borramento por desfoque é causado pela pequena variância entre os pixels adjacentes e o atual, então, pode-se dizer que esse borramento é causado pela interferência da vizinhança do pixel em seu valor, misturando-se ou suavizando a região que eles se localizam. O resultado desse processo é semelhante ao obtido pela aplicação de uma filtragem espacial de baixas frequências. (KROTKOV, 1987, p.226).

Outra forma de medir essa variação é calcular o desvio padrão do pixel segundo uma região da imagem.

$$Var = (f(x,y) - u)^2 \quad (7)$$

Onde: u Média de uma janelada centrada em x,y ;

x, y Coordenadas de um ponto da imagem;

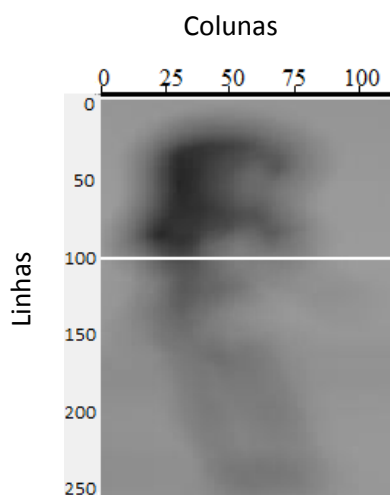
Var Variância e,

f Função que representa as intensidades de cor da imagem.

Uma imagem que não está completamente desfocalizada possuirá maior frequência (portanto maior variância) na região nítida. O desfoque em câmeras com lente acontece em distâncias anteriores ao ponto de foco e posteriores a profundidade de campo. Desta forma Horn (1968) encontrou o ponto focal de uma lente pela mudança de suas configurações.

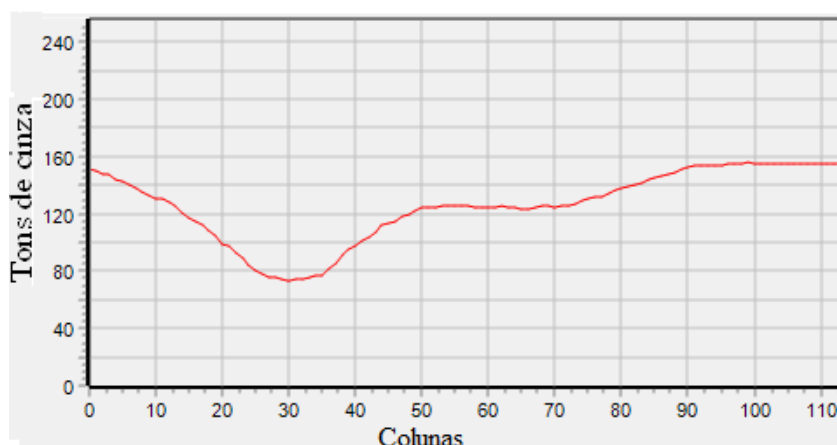
Para demonstrar a aplicação da Equação 5, foi utilizado uma imagem de boneca desfocalizada, conforme a Figura 5. A imagem possui 253 linhas e 115 colunas, e foi convertida em tonalidades de cinza, sendo que cada tonalidade pode ser representada por até 256 valores. A fim de facilitar a identificação do processo foi traçado, sobre a imagem, marcadores espaçados em 50 linhas no eixo vertical, e 25 colunas no eixo horizontal. Uma linha aleatória (linha 100) foi marcada, com a cor branca, para que, depois, seus valores possam ser plotados no Gráfico 1. A Tabela 1 apresenta o processo da aplicação do operador Laplaciano sob os bytes do Gráfico 1, onde é calculado o valor da variação nos pixels da linha, para os primeiros 15 tons de cinza.

FIGURA 5 - Imagem desfocada convertida em tons de cinza.



Fonte: Produzida pelo autor.

GRÁFICO 1 - Tons de cinza da linha 100 da Figura 5



Fonte: Produzida pelo autor.

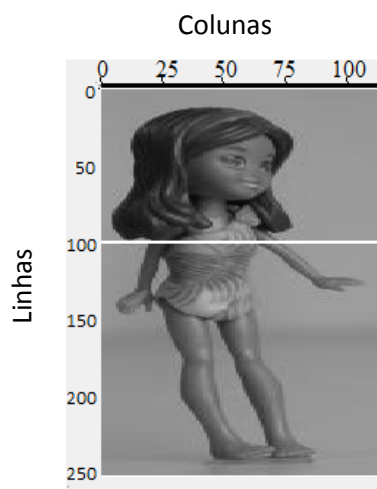
TABELA 1 - Aplicação do operador Laplaciano 1D na linha 100 da Figura 5.

Colunas	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Tons de cinza	151	150	148	147	144	143	140	138	135	133	131	130	128	125	121	117
LAP 1D	N/D	1	-1	2	-2	2	-1	1	-1	0	-1	1	1	1	0	N/D

Fonte: Produzida pelo autor.

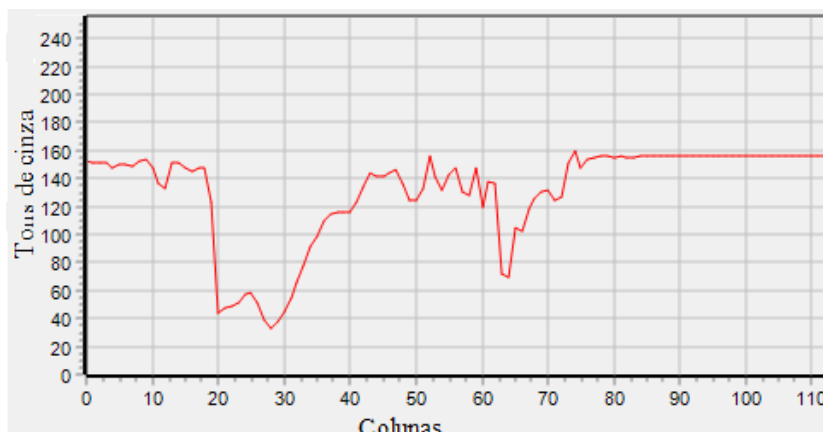
Para qualquer outra imagem (Figura 6), agora focalizada, repetimos o processo de cálculo de variações. A imagem de boneca focalizada (Figura 6) possui 253 linhas e 115 colunas e os valores de cor dessa imagem foram convertidos em tons de cinza. Para visualizar a linha utilizada, foram traçados marcadores acima e no lado esquerdo dela. A cada 50 linhas da Figura 6 (eixo vertical) foi traçado um marcador acima e a cada 25 colunas (eixo horizontal) foi traçado um marcador no lado esquerdo dela. Uma linha aleatória (linha 100) foi escolhida para realizar o cálculo, e pintada com a cor branca facilitar identificá-la. No Gráfico 2 foram plotados as tonalidades de cinza dessa linha e na Tabela 2 são apresentados os resultados do cálculo do operador Laplaciano 1D para os primeiros 15 pixels.

FIGURA 6 - Imagem focalizada convertida em tons de cinza.



Fonte: Produzida pelo autor

GRÁFICO 2 - Tons de cinza da linha 100 da Figura 6



Fonte: Produzida pelo autor.

TABELA 2 - Aplicação do operador Laplaciano 1D na linha 100 da Figura 6.

Colunas	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Tons de cinza	152	151	151	151	147	150	150	149	152	154	147	136	133	151	151	147
LAP 1D	N/D	-1	0	4	-7	3	1	-4	1	9	4	-8	-21	18	4	N/D

Fonte: Produzida pelo autor.

Uma imagem possui coordenadas bidimensionais e as Tabelas 1 e 2 utilizam apenas uma dimensão. O cálculo da variação bidimensional da imagem é realizado aplicando o operador Laplaciano 2D, conforme a Equação 6. As Tabelas 3 e 4 apresentam os mapas de nitidez da Figura 5 e 6, respectivamente, onde, foi aplicação do operador Laplaciano 2D, descrito na Equação 6, para as 15 colunas iniciais da linha 100 desta imagem.

TABELA 3 - Aplicação do operador Laplaciano 2D na linha 100 da Figura 5.

Colunas	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Tons de cinza	Linha 99	151	149	148	147	143	142	140	137	135	132	130	129	127	124	120	117
	Linha 100	151	150	148	147	144	143	140	138	135	133	131	130	128	125	121	117
	Linha 101	152	150	149	148	145	144	141	139	136	134	133	132	129	126	122	119
LAP 2D da linha 100	N/D	2	-2	1	-2	2	-2	1	-2	0	-2	0	1	1	0	N/D	

Fonte: Produzida pelo autor.

TABELA 4 - Aplicação do operador Laplaciano 2D na linha 100 da Figura 6.

Colunas	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Tons de cinza	Linha 99	152	151	151	151	150	152	153	156	156	144	113	83	99	129	149	148
	Linha 100	152	151	151	151	147	150	150	149	152	154	147	136	133	151	151	147
	Linha 101	152	151	151	151	151	153	151	145	143	148	154	157	151	159	145	145
LAP 2D da linha 100	N/D	-1	0	4	-14	-2	-3	-7	6	25	31	24	-5	32	12	N/D	

Fonte: Produzida pelo autor.

Em reconstruções de mapas de profundidade é comum a utilização de variações espaciais ao invés do espectro de frequências. Essas variações fazem parte dos operadores de medida de foco e medem a nitidez da imagem.

2.5 Métodos *Shape From Focus* e *Depth From Focus*

Shape From Focus é um método de reconstrução de profundidade de cena que utiliza um operador de medida de foco (medidor de qualidade de foco) sobre um conjunto de imagens retiradas com diferentes parâmetros da lente (*Focal Stack*), conforme a Figura 7, extraindo regiões com maior variância. Essas regiões correspondem às porções focalizadas.

FIGURA 7 - Pilha de foco do mundo real



Fonte: Adaptado de Moeller et al(2015, p 5369).

A reconstrução de cenários tridimensionais necessita de três coordenadas do sistema de coordenadas do mundo real e uma cor para representar a tonalidade da cor real. Um vóxel é um ponto do cenário tridimensional amostrado do mundo real, que possui coordenadas ortogonais correspondentes a amostras das coordenadas do mundo real, altura, largura e comprimento, bem como a cor amostrada das intensidades luminosas presente na cena. Se as coordenadas do mundo real forem capturadas sob a ótica de fotografia, em que as coordenadas dos pixels da imagem fornecem duas dessas coordenadas, então uma pilha de foco poderá fornecer a terceira coordenada sob a forma de Mapa de Profundidade.

Dentro do cenário tridimensional do mundo real, cada imagem do conjunto da *Focal Stack* possui coordenadas bidimensionais, mas que estão dispostas em um plano de focalização. Cada um desses planos está disposto paralelo e à frente da lente, portanto equivalem a distâncias focais de regiões da imagem.(GOSSMANN, 1987, p.63,66)

Se a profundidade de campo for pequena o suficiente para deixar nítido um pequeno plano focal e o resto da imagem borrada, então a terceira coordenada, bem como o valor da cor do vóxel, será fornecido pelo plano focal.

Como a região focalizada corresponde a região de maior frequência e portanto maior variância, um pico de valor irá se sobressair. Esse pico de variância corresponde ao maior valor da função de medida de foco do conjunto de todas as imagens, para uma determinada localização. Se for utilizada somente a coordenada do plano focal correspondente a esse pico que se sobressaiu, e repetido esse processo para todos os outros pixels, será obtido a profundidade relativa do plano focal. (NAYAR; NAKAGAWA,1994, p.825).

Muitos métodos de reconstrução de profundidade de cena vêm sendo utilizados. Grossmann (1987, p.64) inicialmente propôs o algoritmo *Depth From Focus*, usando um conjunto de imagens obtidas com diferentes pontos focais e pequena profundidade de campo, para retirar suas bordas e comparar os níveis de borrimentos produzidos na imagem. Ele utilizou a seguinte método para descrever seu processo:

- Encontrar um conjunto de primitivas da imagem(bordas, etc);
- Avaliar e medir o borrimento W para cada primitiva;
- Converter W para uma profundidade relativa d (distância do plano Focal);
- Se necessário usar informações de distância para converter d em profundidade absoluta Z .

Darrell e Whon(1988) construíram um mapa de profundidade grosseiro, dividindo melhor o processo de aquisição de primitivas. Eles chamaram seu método de pirâmide de DFF e o descreveram em cinco capítulos, conforme o seguinte:

- Estabelecer um critério de nitidez;
- Obter o mapa de nitidez.;
- Compensar a distorção por zoom da lente;
- Recuperar o mapa de profundidade;
- Parametizar o mapa de profundidade.

Para Darrell e Whon(1988), o janelamento era usado para diminuir os falsos picos que os medidores de nitidez(medidor de qualidade de foco) produzia. Uma compensação por zoom foi adicionada e, estudamos sua aplicação em nossa seção 2.9, como *câmera calibration* para SFF e, a recuperação de profundidade é descrita na seção 2.8, como uma das principais estratégias de reconstrução.

Nayar e Nakagawa(1994), propuseram um novo método para reconstruir o mapa de profundidade chamado de SFF, usando operadores de medida de foco e uma estimativa de profundidade.

- Operador de Medida de Foco
- Estimar Profundidade

Neste trabalho não há o processo de alinhamento de imagens de Darrell e Wohn (1988), após a extração do medida de foco, pois a pequena profundidade de campo, fornecida pela lente de um microscópio, foi apontada por tornar insignificante o efeito zoom da lente. A estimativa de profundidade é encontrada pela interpolação gaussiana entre os maiores valores dos medidores de qualidade de foco no conjunto de imagens para uma dada coordenada. A interpolação gaussiana é apresentada na seção 2.8 como reconstrução fina do mapa de profundidade.

Embora Grossmann (1987) tenha inicialmente estudado os princípios da reconstrução do mapa de profundidade com seu algoritmo DFF, foi o trabalho de Nayar e Nakagawa(1994), SFF, que obteve maior sucesso na reconstrução desse mapa. As duas grandes inovações do algoritmo SFF foram a aplicação de um novo operador de medida de foco chamado Soma dos Laplacianos Modificados em substituição ao operador Soma dos Laplacianos utilizado por Darrell e Whon (1988) em Pirâmide DFF, e a utilização do microscópio, que utiliza lente com pequena profundidade de campo e permite a reconstrução fina do mapa de profundidade para observação de pequenos objetos. O operador Laplaciano Modificado necessitava de menor custo computacional e possuía maior precisão que o operador Tenengrad. Esses dois operadores serão apresentados na seção 2.6.

Para Subbarão e Choi (1995, p.267), o método de Nayar e Nakagawa(1994) não gerava corretamente o mapa de profundidade de um objeto, quando possuía uma superfície inclinada em relação ao plano focal e, propuseram um novo tipo de reconstrução fina com o algoritmo Focused Image Surface-SFF (FIS-SFF). O método FIS-SFF pode ser simplificado conforme o seguinte:

- Gerar uma sequência de imagens em tons de cinza;
- Computar medidores de qualidade de foco;
- Realizar uma estimativa inicial para a profundidade da cena, usando apenas os medidores de qualidade de foco mais confiáveis (que possuem maior variação);
- É realizado um loop para unir partes da estimativa inicial de profundidade;
- Para cada estimativa inicial, uma janela volumétrica é empregada para encontrar uma superfície plana do objeto.

A última parte desse algoritmo é descrita na seção 2.8 como uma das estratégias de reconstrução fina de profundidade dos algoritmos SFF. Segundo Subbarão e Choi (1995, p.267) e Mooler et al.(2015, p.5369) os algoritmos SFF e DFF são equivalentes.

Niederost M. et al. (2003) propuseram aplicar um filtro janelado de média sobre o mapa de profundidade para reduzir as imperfeições que foram geradas pelo medidor de nitidez,

construindo uma ferramenta de reconstrução tridimensional de objetos microscópicos usando a biblioteca Glscene¹.

Kim e Tian (2009) propuseram um método para alinhamento de imagens, antes da aplicação dos processos descritos por Nayar e Nakagawa(1994) e assim reduzir os efeitos de ampliação por zoom, que a lente gera, em seu algoritmo Novo DFF.

Gaganov e Ignatenko(2009) propuseram agregar valor diretamente no mapa de profundidade. Utilizaram para isso uma ferramenta científica, chamada Markov Random Fields (MRF), que possui a propriedade de distribuir uniformemente os valores dos medidores de qualidade de foco. Essa ferramenta usa uma função de minimização de energia que, ora suaviza, ora exalta o valor da vizinhança do pixel e assim constrói o mapa de profundidade final. Esse método foi apresentado como Robusto SFF via MRF e pode ser descrito conforme os seguintes passos:

- Aplicar um medidor de nitidez sobre a imagem;
- Utilizar a estratégia WTA(Winner-takes all), para achar as superfícies focalizadas;
- Regularização por MRF, utilizando a Equação de minimização de energia abaixo:

$$E(d) = \sum_i \left(V_i(d_i) + \lambda \sum_{j \in N(i)} U_{\{i,j\}}(d_i, d_j) \right) \quad (8)$$

Onde: $E(d)$ é a função de energia Markov.

$N(i)$ é o numero de vizinhos conectados

$V_i(d_i)$ é a energia do termo de dado. (medidor de qualidade de foco)

$U_{\{i,j\}}(d_i, d_j)$ é a resposta de suavização ao conjunto dos vizinhos conectados.

No método acima não há agregação antes da estratégia Winner Takes All(WTA) e nem alinhamento de imagens. Este método não utiliza reconstrução fina e a estratégia WTA é descrita como sendo uma estratégia Máxima a posteriori (MAP), que posteriormente será descrita na seção 2.8.

Hasinoff e Kutulakos(2009) propuseram reconstruir o mapa de profundidade com um novo tipo de pilha de imagens utilizando a constância focal. O método Confocal pode ser descrito conforme os seguintes passos:

- Obter imagens com pequenas profundidades de campo e, também grandes profundidade de campo, variando o ponto focal da lente;
- Guarde essas imagens em grupos de posições focais;
- Obter, no grupo, a menor variação conforme a coordenada;

¹ Disponível em www.glscene.org

- Selecionar, somente os menores valores de variação do grupo de imagens;
- Construir o mapa de profundidade, utilizando a localização e as menores variações para regiões.

David et al.(2012), propuseram calcular a imagem focalizada através da Função Integral de Espalhamento de Ponto(IPSF) para comparar sua variação com a variação de cada imagem da *Focal Stack*. Depois usar a localização da imagem para computar a profundidade da região de menor variação. Este método foi apresentado como Integral SFF.

Jing e Yeung(2012) utilizaram as características da segmentação Mean Shift Segmentation para simular uma janela adaptativa. Esse método também realiza o alinhamento de imagens, porém não utiliza a reconstrução fina do mapa de profundidade. Este método pode ser simplificado conforme o seguinte:

- Pré-processamento pelo fluido óptico, para alinhar imagens;
- Estimção pura do mapa de profundidade usando um medidor de nitidez;
- Refinamento de bordas usando Mean-shift-Segmentation;
- Aplicação da ferramenta Markov Random Field, para unir regiões com pouca textura.

O pré-processamento de fluido óptico definido neste algoritmo, é apresentado como compensação de aumento por zoom da lente, na seção 2.9. A estimção de profundidade é o computo do mapa de profundidade para planos focais.

O método de Mahmood(2013), propõe um método que utiliza mais de um medidor de nitidez como hipóteses de profundidade para então aplicar a ferramenta científica variacional, e calcular a energia mínima entre essas hipóteses. Seu método é uma nova abordagem do algoritmo de Gaganov e Ignanteki(2009), sendo apresentado como Variação Total SFF. Esse método pode ser descrito conforme o seguinte:

- Utilizar mais de um medidor de nitidez para cada imagem;
- Montar volumes de medidores para cada imagem.
- Obter hipóteses de profundidades .
- Criar apenas um mapa de profundidade usando a energia variacional, e unindo as hipóteses anteriores, conforme a seguinte fórmula.

$$\begin{pmatrix} d_1 \\ \vdots \\ d_m \end{pmatrix} = \begin{pmatrix} k_1 \\ \vdots \\ k_m \end{pmatrix} d' + \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_m \end{pmatrix} \quad (9)$$

$$d = kd' + \eta$$

Onde: d é a matriz com os vários mapas de profundidade.

m é um número para um dos tipos de medição de foco.

k é uma constante variacional de valorização do mapa (ganho).

η é a constante variacional de valorização do ruído (perda).

d' é o mapa de profundidade final.

d é um dos mapas de profundidade para um método de medição.

Tseng e Wang(2014) propuseram unificar alguns mapas de profundidade para estimar a profundidade utilizando a consistência espacial para SFF, esse método pode ser descrito conforme abaixo:

- Computar medidores de qualidade de foco;
- Aplicação do modelo de consistência espacial;
- Aplicar a busca Máximo a Posteriori (MAP), para encontrar o mapa de profundidade.

Nesse modelo, a consistência espacial é centralizado em células temporais locais, onde um grupo de imagens fornece a consistência espacial, conforme a Figura 8.

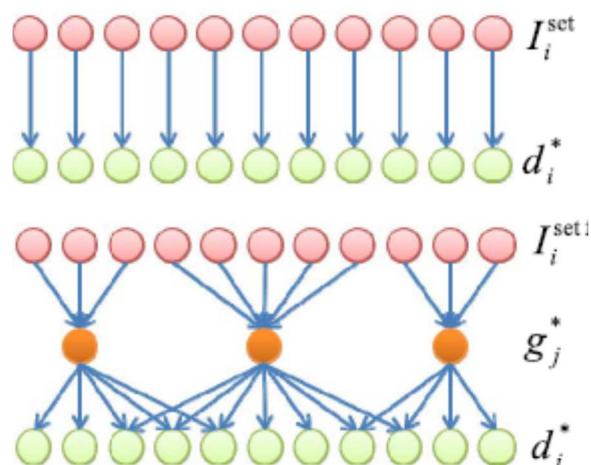
FIGURA 8 - União de medidores de qualidade de foco em células temporais.

Legenda:

I_i^{set} Pixels da imagem digital;

d_i^* Profundidade calculada para o conjunto I_i^{set} , e

g_j^* Células introduzidas pela consistência espacial.



Fonte: Tseng e Wang(2014).

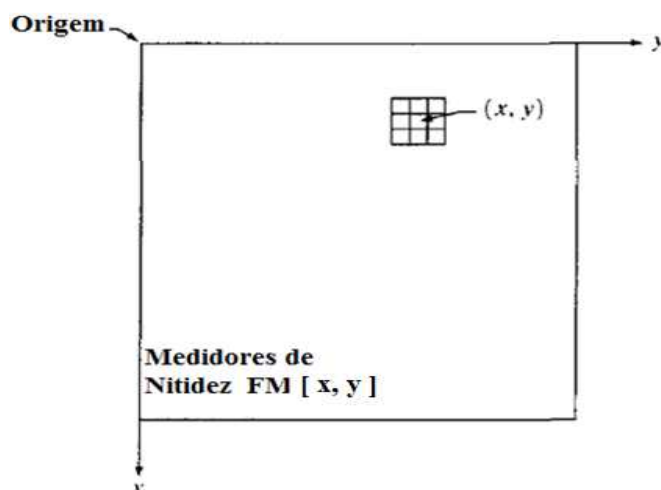
2.6 As funções de medida de qualidade de foco

A primeira parte do algoritmo SFF é a parte mais importante do processo, pois fornece as bases para extrair as altas frequências da imagem. Os medidores de qualidade de foco compõem a parte inicial dos algoritmos SFF e, são compostos pelo medidor de nitidez, estudado como taxa de variação na seção 2.4, e às vezes, uma janela de agregação de valores, obtida com a soma das variâncias de pixels vizinhos de sua região (medidores de nitidez vizinhos). No processo de extração de variação local, um medidor de qualidade de foco é obtido para cada pixel da imagem.

A agregação pode ser obtida somente com a variação dos vizinhos (medidores de nitidez vizinhos) mais próximos ao medidor de nitidez corrente, ou com variação de vários vizinhos deste pixel em uma região (NAYAR; NAKAGAWA, 1994, p.827; SUBBARÃO; CHOI, 1995,

p.267). Essa estrutura é conhecida como janelamento sobreposto no medidor de nitidez e é mostrada na Figura 9, onde uma matriz de medidores de foco FM possui as mesmas dimensões de linhas e colunas que uma matriz imagem.

FIGURA 9 – Medidores de nitidez com sua vizinhança mais próxima no mapa de nitidez



Fonte: Adaptado de Gonzalez e Woods. (2010, p.95)

Se a profundidade de campo produzir uma região focalizada na imagem maior que um pixel então, a variância de somente um pixel não refletirá a profundidade da cena e, a janela sobreposta adaptará a variação do pixel à variação de sua região. Nayar e Nakagawa (1994) utilizaram uma janela de 10 linhas e 10 colunas (10 x 10) no algoritmo SFF, Subbarão e Choi (1995, p.267), usaram uma janela 15 x 15 no algoritmo FIS-SFF. Niederost M., Niederost J. e Scucka (2003), utilizaram uma janela de 9 x 9 com um medidor usando variância de região (Equação 7) nos valores de intensidades de cinza.

Para determinar o tamanho da janela do medidor no algoritmo SFF, muitos autores abordam o problema de forma empírica, realizando experimentos sobre o objeto a ser reconstruído, como os descritos no parágrafo anterior. Malik e Choi (2007) estudaram os efeitos da iluminação sobre o tamanho da janela em medidores de qualidade de foco. Recentemente Srinivas et. al(2015) estudaram o efeito do tamanho da janela na qualidade de reconstrução do mapa, utilizando diversos medidores de nitidez. Mahmood e Choi (2012) usaram um medidor baseado em volume, pela filtragem anisotrópica por difusão não linear (ANDF) em um janelamento 3D.

Além dos medidores de nitidez apresentados na seção 2.4, outros medidores de nitidez foram propostos. Tenenbaum (1970) utilizou o gradiente e propôs seu operador Tenengrad conforme a seguir:

$$\Phi_{x,y} = \sum_{(i,j) \in \Omega(x,y)} (G_x(i,j)^2 + G_y(i,j)^2) \quad (10)$$

Onde: G_x e G_y são os operadores gradiente para a linha e para a coluna, respectivamente

$\Phi_{x,y}$ é o operador Tenengrad para as coordenadas x e y da imagem.

$\Omega(x,y)$ é a janela centrada nos pixels x e y .

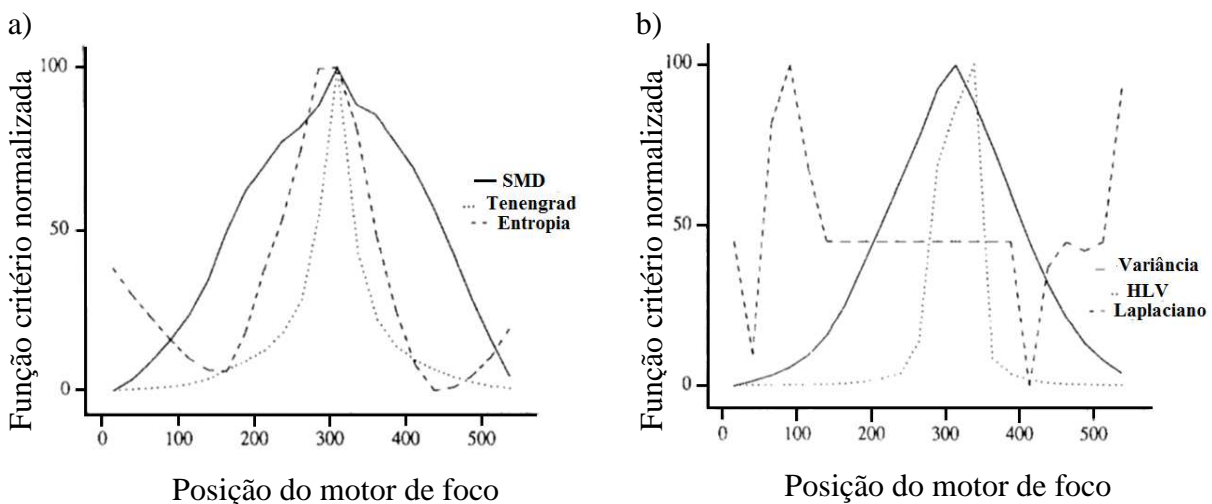
x, y são as coordenadas do medidor Tenengrad.

i, j são as coordenadas utilizadas na janela.

Krotkov(1987) sintetizou e comparou a resposta dos diversos medidores de qualidade de foco em operações de autofoco. Para Krotkov (1987, p.230) o operador Tenengrad demonstrou ser unimodal em aplicações de autofoco, ou seja, variando monotonicamente com a posição do motor de foco em todas as direções, razoavelmente robusto na presença de ruído, mesmo sem qualquer filtragem das amostras de intensidade.

No Gráfico 3 os medidores de qualidade de foco foram normalizados e plotados para apenas uma coordenada(x,y) de um conjunto de imagens. O gráfico foi construído conforme a posição do motor de foco adquiria imagens. Os seguintes medidores de foco foram aplicados para que fosse realizada a comparação de desempenho, a saber: variância dos níveis de cinza(HLV), Laplaciano, Tenengrad, Entropia, e soma de diferença dos módulos (SMD). A função critério é o medidor de qualidade de foco empregado em aplicações de autofoco.

GRÁFICO 3 - Comparação de medidores de qualidade de foco para autofoco.



Fonte: Adaptado de Krotkov(1987).

Nathaniel et al.(2001, p 2795) também analisaram os medidores de qualidade de foco, e obtiveram bom desempenho utilizando a variância e o operador Tenengrad.

Recentemente o trabalho de Pertuz et al.(2013) comparou inúmeros medidores de qualidade de foco, adaptados para a recuperação de mapas de profundidades. Segundo esse trabalho os operadores baseados em Laplaciano demonstraram melhor performance no geral, quando utilizados com amostras de imagens em condições normais (sem distorção, sem alto contraste e sem outras formas de interferência).

O operador Laplaciano Modificado (NAYAR; NAKAGAWA, 1994) é bastante utilizado em artigos acadêmicos que utilizam o algoritmo SFF, sendo calculado pela soma dos módulos de dois operadores Laplacianos: um de linha e outro de coluna. Se um operador Laplaciano é uma segunda derivada, então a soma do módulo das funções derivadas é descrito conforme abaixo.

$$\nabla_M^2 I = \left| \frac{\partial^2 I}{\partial x^2} \right| + \left| \frac{\partial^2 I}{\partial y^2} \right| \quad (11)$$

Onde : $\frac{\partial^2 I}{\partial x^2}$ é a segunda derivada na coordenada x, do conjunto bidimensional I(x,y)

$\frac{\partial^2 I}{\partial y^2}$ é a segunda derivada na coordenada y, do conjunto bidimensional I(x,y).

$\nabla_M^2 I$ é o operador laplaciano modificado.

Na prática, o operador Laplaciano Modificado de duas dimensões, para um sinal imagem, é obtido pela soma dos módulos de dois operadores Laplacianos em duas coordenadas e, portanto o módulo da soma desses operadores pode ser calculado na seguinte equação:

$$\nabla_M^2 I = |2 \cdot f(x,y) - f(x,y+1) - f(x,y-1)| + |2 \cdot f(x,y) - f(x+1,y) - f(x-1,y)| \quad (12)$$

Onde: $\nabla_M^2 I$ é o operador laplaciano modificado.

$x, x+1$ e $x-1$ são a linha do pixel, a linha antecedente e a linha consequente.

$y, y+1$ e $y-1$ são a coluna do pixel, a coluna antecedente e a coluna consequente.

Nas tabelas 5 e 6, o operador Laplaciano Modificado é aplicado sobre os pixels das Figuras 5 e 6 obteremos os medidores de nitidez descritos nas Tabelas 5 e 6, onde o mapa de nitidez da Figura 5 é resultado da aplicação do operador Laplaciano Modificado 2D, descrito na Equação 12, para 15 colunas iniciais, e o mapa de nitidez da Figura 6 é resultado da aplicação do operador Laplaciano Modificado 2D, descrito na Equação 12, para 15 colunas iniciais, imagem.

TABELA 5 - Aplicação do operador Laplaciano Modificado 2D na linha 100 da Figura 5.

Colunas		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Tons de cinza	Linha 99	151	149	148	147	143	142	140	137	135	132	130	129	127	124	120	117
	Linha 100	151	150	148	147	144	143	140	138	135	133	131	130	128	125	121	117
	Linha 101	152	150	149	148	145	144	141	139	136	134	133	132	129	126	122	119
LAP 2D da linha 100		N/D	2	2	3	2	2	2	1	2	0	2	2	1	1	0	N/D

Fonte: Produzida pelo autor.

TABELA 6 - Aplicação do operador Laplaciano Modificado 2D na linha 100 da Figura 6.

Colunas		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Tons de cinza	Linha 99	152	151	151	151	150	152	153	156	156	144	113	83	99	129	149	148
	Linha 100	152	151	151	151	147	150	150	149	152	154	147	136	133	151	151	147
	Linha 101	152	151	151	151	151	153	151	145	143	148	154	157	151	159	145	145
LAP 2D da linha 100		N/D	1	0	4	14	8	5	7	6	25	31	40	37	32	12	N/D

Fonte: Produzida pelo autor.

A grande vantagem do operador Laplaciano Modificado sobre o operador Laplaciano é que a soma de módulos sempre acrescenta valores de variações, pois no operador Laplaciano o sinal poderia gerar subtrações e ocasionalmente zerar o medidor, assim o operador fica mais sensível ao acréscimo.

Os medidores de qualidade de foco possuem grande utilidade em aplicações de SFF ou DFF, autofoco, fusão de imagens de *Focal Stack* (KUTHIRUMMAL et al., 2011, p.67-68), entre outros. Os medidores de autofoco surgiram primeiro e foram posteriormente adaptados para algoritmos de SFF(PERTUZ; PUIG;GARCIA, 2013, p.1415).

2.7 Aglutinando variações no medidor de qualidade de foco.

Aglutinar variações no medidor de qualidade de foco é unir vários medidores de nitidez pela soma dessas variações. A primeira variação é a que se encontra na posição correspondente a linha e coluna corrente, e as demais estão na região vizinha. Essa região é escolhida conforme um determinado critério.

Inicialmente utilizou-se o janelamento sobreposto nos medidores de nitidez como método de aglutinação. Ao unir essas variações, o objetivo era eliminar falsas previsões no mapa de profundidade e reduzir os ruídos provocados pela aquisição da pilha de foco. Posteriormente, verificou-se que apesar dos falsos resultados positivos produzidos por essa janela, ao ser aplicado sobre novas amostras, também incorporava pixels com profundidades diferentes e acabava gerando imperfeições na reconstrução do mapa de profundidade. Se os medidores de qualidade de foco utilizam a soma dos medidores de nitidez de uma região então regiões com grande variações de intensidades de cor influenciarão outras com pouca variação.

Desta forma o tamanho da janela pode introduzir imperfeições em regiões texturizadas ou mesmo com pequena profundidade de campo. Recentemente alguns trabalhos adicionam um validador que atribui pesos para os medidores de nitidez da janela e a torna adaptativa (AYDIN; AKGUL, 2008).

Uma imagem pode possuir muitos objetos. Um objeto ou um local da cena é caracterizado por uma cor em cada uma de suas superfícies. Em geral, mesmo que cada superfície seja irregular, possuirá um conjunto semelhante de intensidades de cor, que produzem uma falsa percepção de única cor. Quando a imagem é formada, mais de uma superfície pode pertencer a um objeto que, estarão sob a mesma distância da câmera, ou seja, com profundidades semelhantes, porém não é possível afirmar que as superfícies pertencem ao mesmo objeto sem conhecer sua composição geométrica.

Para tratar esse problema são validados os vizinhos de um medidor que faz parte de uma janela com algum critério, encontrando semelhanças que permitam aproximar o medidor das características da região em que ele se encontra. A Equação 13, define o janelamento adaptativo AFM segundo Aydin e Akgul(2008), onde pesos de uma função critério ω_{x_0,y_0} , em uma janela $\Omega(x,y)$, são aplicados sobre os medidores de qualidade de foco, $FM(x,y)$, para gerar o novo medidor AFM adaptativo pela multiplicação deles.

$$AFM(x_0, y_0) = \sum_{(x,y) \in \Omega_{x_0,y_0}} \omega_{x_0, y_0}(x, y) FM(x, y) \quad (130)$$

Onde:

ω_{x_0,y_0}	máscara de pesos da janela centrada em x_0 e y_0
$FM(x,y)$	medidor de qualidade de foco original nas coordenadas x e y
$AFM(x_0,y_0)$	medidor adaptativo para as coordenadas x_0 e y_0
x, y	Coordenadas da janela adaptativa
x_0, y_0	Coordenadas do medidor adaptativo

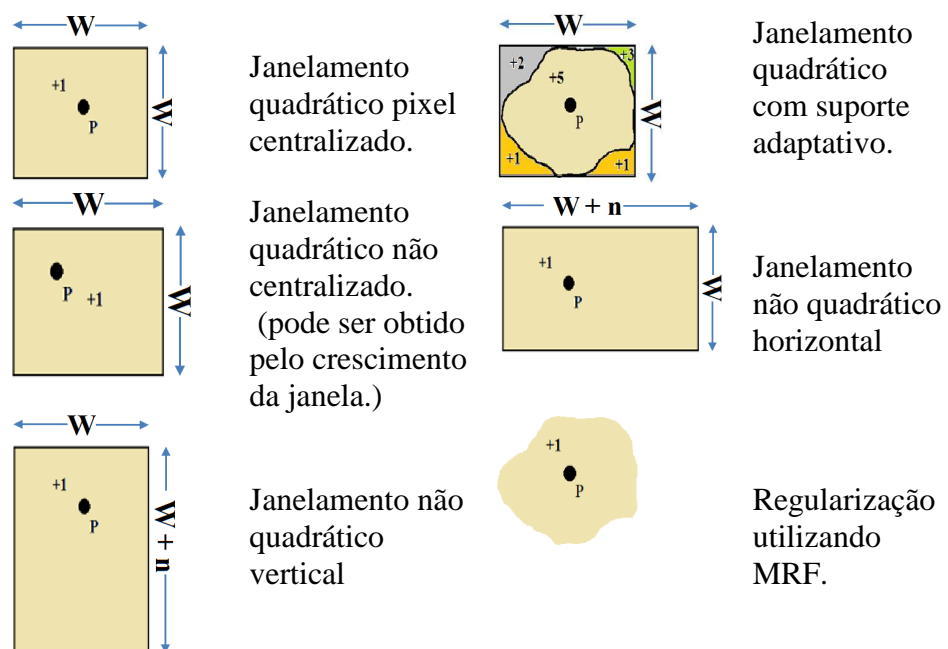
Existem outras formas de agregar valor em um medidor de qualidade de foco, e podem ser descritas conforme o seguinte: Empregar janelas retangulares, onde o maior lado (horizontal ou vertical) é selecionado conforme a direção de maior variação (LEE; SHIM; CHOI, 2013). Empregar o crescimento de janela, definindo um tamanho mínimo e outro máximo, onde a janela de maior variação é selecionada. Apesar do método possuir tamanho mínimo e máximo de janelamento quadrático no medidor de qualidade de foco, pode ocorrer a seleção de uma janela retangular, desde que esteja dentro dos limites escolhidos (LEE, MAHMOOD; CHOI, 2013). Empregar a regularização, onde o mapa de profundidade é inicialmente calculado e depois utilizado um aglutinador (janela, MRF) de valores de profundidade sobre eles para reduzir os efeitos de possíveis imperfeições (GAGANOV; IGNATENKO, 2009).

A regularização é uma forma de agregação de valores realizada após o cálculo dos medidores de qualidade de foco e que usa todas as variações calculadas para a imagem. Este processo consiste em atribuir maior valor a variação na posição corrente, no entanto usa um termo que valida todas as outras variações conforme se distanciam desta posição. O objetivo é suavizar as falsas previsões pelo afastamento das posições de variações.

Enquanto que no janelamento ou janelamento com suporte adaptativo a variação da vizinhança é agregada no medidor de qualidade de foco (MAHMOOD; CHOI, 2012, p.2867), na regularização essa janela é suavizada diretamente no mapa de profundidade (GAGANOV; IGNATENKO, 2009).

A Figura 10 apresenta as diversas formas de agregação ou regularização de valores segundo um dado critério. As imagens representam uma região onde a variação do pixel P é semelhante às regiões coloridas da Figura. A região delimitada por um quadrado representa a janela W centrada no pixel P , e os pesos são utilizados para agregação de valores. Uma região geralmente possui maior peso quando tem características semelhantes de um pixel, exemplo: proximidade do pixel P possuirá maior valor que os mais afastados; variação de intensidade semelhante mais próxima do valor de P , entre outros.

FIGURA 10 - Tipos de aglutinação de valores em um medidor de qualidade de foco.



Legenda:

W Tamanho da janela

P Um ponto da imagem(Pixel)

+1,+2, +3 e +5 pesos

N um valor qualquer

Fonte: Produzida pelo autor.

A regularização pode ser interpretada tanto como uma parte de estratégia de reconstrução de profundidade ou como uma medida de agregação de valores ao medidor de

profundidade, sendo citada nesta seção por compor uma das formas de abordar o problema de aglutinação.

2.8 Estratégias de reconstrução de profundidade

As estratégias de reconstrução são divididas em duas partes: reconstrução grosseira e reconstrução fina.

A reconstrução grosseira utiliza os medidores de qualidade de foco para montar o mapa de profundidade pela localização dos planos focais (AHMAD; CHOI, 2005, p.567), onde o mapa de profundidade de cena pode ser obtido com a extração de regiões focalizadas diferentemente em um conjunto de imagens, ou seja, se a focalização é manifestada pela grande variação das intensidades de cor, conforme estudado nas seções 2.4 e 2.6, então em uma determinada coordenada da *Focal Stack* a imagem com a região em foco possuirá maior variância.

O algoritmo de procura pela maior variância é realizado pela procura do maior valor do medidor de foco, em que todas as imagens fazem parte do universo de pesquisa. O processo de obtenção do maior valor do medidor de foco pode ser descrito da seguinte forma:

- Dada uma pilha de foco com N imagens, L linhas e C colunas, se um pixel localizado nas coordenadas l linha e c coluna, encontra-se com uma variância x , na imagem n , e na imagem conseqüente o pixel localizado nas mesmas coordenadas do anterior, possui maior variância descrito por $x+k$, onde k é um valor qualquer, e essa variação continuar crescendo, na medida em que são avaliadas as próximas imagens do conjunto da pilha, alcançando seu maior valor na posição $n+w$, onde w é um valor que somado a n não alcança o limite de imagens da *Focal Stack*, então o foco para esse pixel estará na imagem que se encontra na posição $n+w$. Para cada pixel um valor de $n+w$ fornecerá a localização da imagem que contem o foco. A localização da imagem que está com foco naquele pixel é o mapa de profundidade dele.

Com essas informações montamos uma estimativa inicial do mapa de profundidade com todas as distâncias focais. Esse mapa inicialmente corresponderá grosseiramente as coordenadas do eixo Z (NAYAR; NAKAGAWA, 1994, p.827).

A maior máxima dentro das variações intensidade de cor dentro do conjunto de imagens como unidade de medida de seleção de profundidade grosseira (*Winer Takes-All*), em probabilidade, é entendido como o *máximo valor a posteriori* (MAP) e é utilizado nos modelos que utilizam emprego de métodos probabilísticos, como em Gaganov e Ignatenko(2009). Segundo Moeller et al.(2015) alguns trabalhos recentes os medidores MAP baseados na

minimização da energia, apresentam bom resultado por regularização. A Equação 14 define a fórmula para obter os valores do mapa de reconstrução de profundidade grosseira.

$$S(x,y) = \text{Arg max}(FM(x,y)) \quad (14)$$

Onde: S é matriz/Mapa de profundidade para cada posição da imagem x,y .

Arg max é a posição de maior valor da função FM .

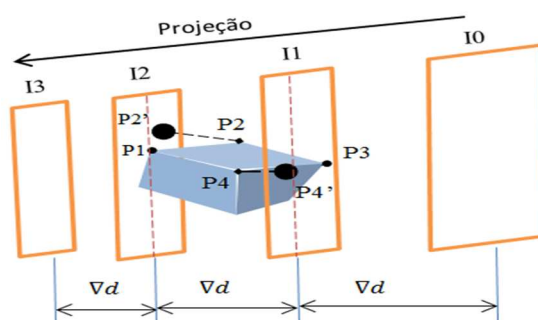
FM Medidor de qualidade de foco

A reconstrução fina, ou interpolação gaussiana, é realizada por estimativa para descobrir os valores de profundidade que não se encontram nos medidores de qualidade de foco da pilha de imagens, utilizando o histórico desses medidores em uma determinada coordenada da *Focal Stack*.

Como o mapa de localizações de posições focais não corresponde exatamente a forma de profundidade dos objetos, um algoritmo de refinamento é aplicado sobre ele, para garantir a formatação correta de profundidade.

Na Figura 11 um quadrilátero está posicionado entre os pontos focais $I1$ e $I2$. Esses planos correspondem às imagens digitais capturadas na *Focal Stack*. Os pontos $P1$ e $P3$ são dois vértices desse quadrilátero e encontram-se nos planos focais $I1$ e $I2$, respectivamente, por isso seus medidores de foco refletem exatamente essa posição no plano focal. Os pontos $P2'$ e $P4'$ são as projeções de outros dois vértices desse quadrilátero (pontos $P2$ e $P4$) nas imagens dos planos focais $I1$ e $I2$. Os medidores dos pontos $P2'$ e $P4'$ são apenas aproximações da posição de focalização dos pontos $P2$ e $P4$ no mundo real.

FIGURA 11- Projeção dos vértices de um plano do quadrilátero em planos focais.



Legenda:

$I0$ a $I3$ Planos focais;

$P1$ a $P4$ Vértices do plano superior de um quadrilátero.

$P2'$ e $P4'$ Projeções dos pontos $P2$ e $P4$ e,

∇d Distância entre os planos focais.

Fonte: Produzida pelo autor.

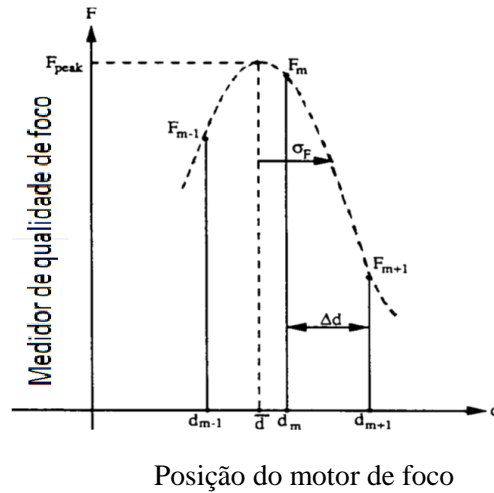
Nayar e Nakagawa (1994) propuseram aplicar uma função gaussiana sobre os medidores de qualidade de foco de uma mesma coordenada do conjunto da *Focal Stack* e utilizar o pico

dessa função (F_{peak}) para representar a localização do pixel na correta posição de profundidade (\bar{d}), conforme o Gráfico 4. Ao aplicar a interpolação gaussiana projetaremos sua posição com base nos medidores de qualidade de foco mais próximos.

GRÁFICO 4 - Interpolação gaussiana.

Legenda:

- F_{peak} Pico da interpolação gaussiana.
- F_m Medidor de qualidade de foco com maior valor.
- F_{m-1} e F_{m+1} Medidores mais próximos de F_m .
- Δd Profundidade de campo ou distância entre os planos focais.
- d_m Ponto focal onde se encontra F_m .
- d_{m-1} e d_{m+1} Pontos focais próximos de d_m .
- \bar{d} Distância onde o pico da curva gaussiana estaria.
- σ_f Variância da curva gaussiana na para reconstrução fina.



Fonte: Adaptado de Nayar e Nakagawa(1994).

O pico dessa função é calculado por um sistema de equações utilizando a Equação 15 para três medidores de qualidade de foco que se encontram próximos ao pico máximo (F_{peak}). F_m é o medidor de qualidade de foco com maior valor, na mesma coordenada da *Focal Stack*, F_{m-1} é o medidor anterior ao medidor de maior valor, e F_{m+1} é o medidor posterior ao medidor de maior valor, portanto os medidores das três imagens mais bem focalizadas em uma mesma coordenada dentro da pilha.

$$F = F_p \exp \left\{ -\frac{1}{2} \left(\frac{d - \bar{d}}{\sigma_F} \right)^2 \right\} \quad (15)$$

Onde: F um medidor de qualidade de foco do plano focal

F_p o medidor de pico da função gaussiana e, que não pertence a um plano focal

d distância do medidor F dentro da *Focal Stack*

\bar{d} distância do medidor de pico F_p da função gaussiana, que não pertence a um plano focal.

σ_F é o desvio padrão da função gaussiana.

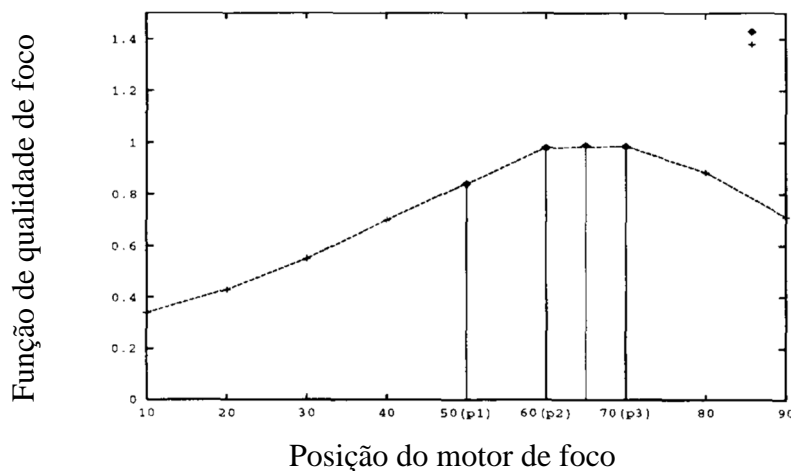
O valor de d e \bar{d} dependem das configurações da *Focal Stack* e por isso Nayar e Nakagawa(1994) propuseram normalizar os valores dos medidores de qualidade de foco, dividindo todos os medidores pelo valor do maior medidor, conforme o seguinte:

$$\begin{aligned} \check{F}_{m-1} &= \frac{F_{m-1}}{F_m} & \check{d}_{m-1} &= -1 & (16) \\ \check{F}_m &= 1 & \check{d}_m &= 0 \end{aligned}$$

$$\check{F}_{m+1} = \frac{F_{m+1}}{F_m} \quad \check{d}_{m+1} = +1$$

Subbarão e Choi (1995) propuseram utilizar a interpolação polinomial para estimar a localização do pixel entre os dois planos focais. No Gráfico 5, as posições focais 50, 60 e 70 são configurações do motor de autofoco, expresso em passos, onde os medidores de qualidade de foco, alcançam os maiores valores, para uma mesma coordenada da *Focal Stack*. A interpolação polinomial é realizada na posição 65, utilizando os medidores dessas três posições.

GRÁFICO 5 - Interpolação polinomial para reconstrução fina.

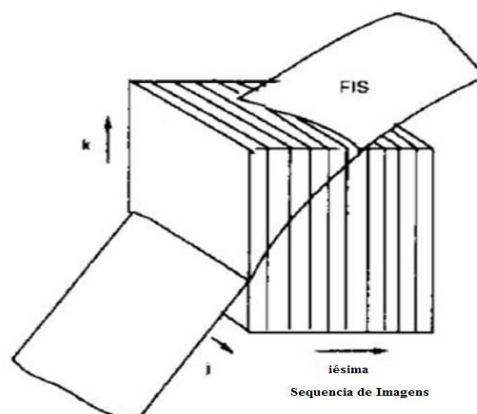


Fonte: Adaptado de Subbarão e Choi (1995),

Subbarão e Choi (1995), também, adicionaram ao algoritmo SFF o conceito de superfícies focalizadas onde uma procura volumétrica por uma superfície plana é realizada sobre um conjunto de imagens. A procura é para encontrar um plano dentro do volume de imagens, onde a inclinação dessa superfície é encontrada procurando a repetição de um padrão de foco, utilizando uma sub-matriz de tamanho $M \times M \times M$ de volume de imagens. A escolha do tamanho desse volume e também do padrão de repetição são encontrados pelo algoritmo de força bruta. Ahmad e Choi (2005) utilizaram a programação dinâmica para otimizar o algoritmo FIS, diminuindo seu custo computacional, e denominaram-no de SFFDP.

Na Figura 12, a superfície (FIS) é particionada focalmente, e cada amostra possui apenas uma parte de suas coordenadas. Através da maximização dos medidores de qualidade de foco no conjunto de amostras, reconstruímos toda a superfície FIS.

FIGURA 12- Amostras de superfície focalizada (FIS)



Fonte: Adaptado de Ahmad e Choi (2005, p.566).

Os algoritmos que utilizam a reconstrução fina fornecem estrutura detalhada de pequenos objetos e são chamados de mapas locais ou densos, e são usados na reconstrução de objetos 3D. Esses algoritmos SFF geralmente são utilizados com a utilização de lentes com pequena profundidade de campo.

Muitos algoritmos SFF não utilizam a reconstrução fina para construir o mapa de profundidade, pois apenas fornecem um mapa de distâncias entre objetos da cena. Os mapas de profundidade desses algoritmos são chamados de mapas globais ou esparsos e em geral utilizam lentes com grande profundidade de campo (NAIR; STEWART, 1992, p.311), sendo comumente chamados de reconstrutores de cenas.

Após a computação do mapa de profundidade grosseiro, muitos autores propõem a regularização como forma de eliminar as imperfeições geradas pelos diversos ruídos introduzidas pelo processo de captura de *Focal Stack* (GAGANOV; IGNATENKO, 2009).

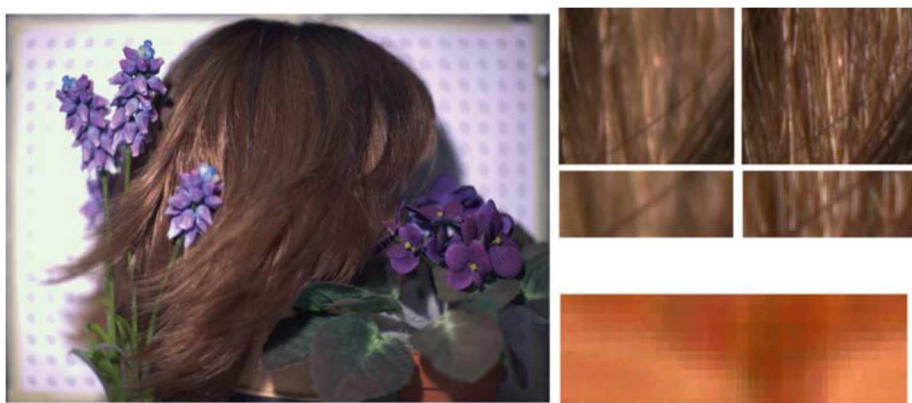
Outra forma de obter o extrair o conteúdo em foco é comparar a variância dos medidores da *Focal Stack* com os medidores de uma imagem totalmente focalizada. Esta comparação procurará a menor variância entre os medidores de uma imagem da *Focal Stack* com os medidores da imagem com focalizada (HASINOFF; KUTULAKUS, 2009) (DAVID; et al, 2012).

2.9 Câmera Calibration para Depth From Focus

Nair e Stewart (1992) sintetizaram e descreveram vários fatores que geram as imprecisões no algoritmo DFF(SFF) afirmando que possivelmente eram causadas pelo processo de captura de pilha de fotos. Essas interferências podem ser enumeradas como o seguinte: o efeito do *zoom*, a distorção pela estrutura curva da lente e o sangramento de borda (ocasionado pelas sombras, pouca luminosidade ou excesso de luminosidade).

Os dois primeiros problemas, geralmente são causados pela movimentação dos pixels na imagem, e o segundo por características advindas do meio de captura, portanto são manipuláveis fora do processo quando utilizado em ambiente controlado. Para tratar o problema de movimentação dos pixels o processo de obtenção de pilha de imagens foi analisado. Durante o processo, a lente é movimentada sobre a coordenada Z, provocando o efeito *zoom*. A curvatura desta lente esférica, também pode gerar aberrações na imagem. Na Figura 13 temos a imagem utilizada no artigo de Hasinoff e Kutulakos (2009), onde a reconstrução do mapa de profundidade de cabelos necessita de grande precisão para extração do formato dos diversos fios de cabelos. O alinhamento de imagens com essas características geralmente possui grande resolução e necessita de alinhamento de pixels para comparar e extrair a profundidade do objeto.

FIGURA 13 - Variação de profundidades em cabelos



Fonte: Hasinoff e Kutulakos (2009, p. 83)

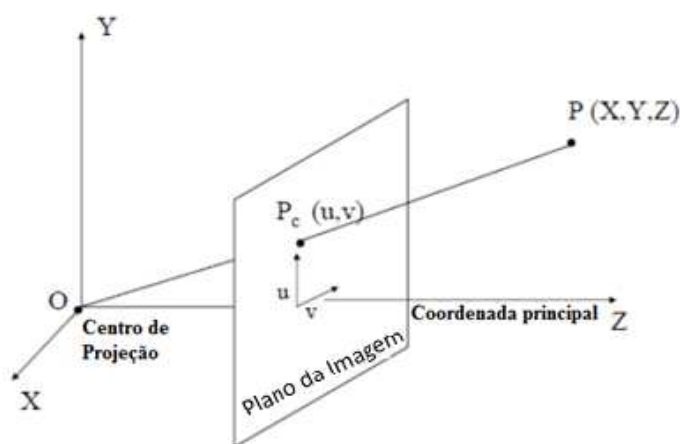
O alinhamento de imagens de uma pilha de foco pode ser realizado por um processo conhecido como *Câmera Calibration*. Esse processo é muito utilizado em processos computacionais onde as imagens são obtidas em diferentes pontos de visão, e por isso necessita de uma adaptação para poder ser utilizado no processo de *Focal Stack*, pois teremos somente um ponto de visão.

Segundo Forsyth e Ponce(2002), a calibração de câmera(*Camera Calibration*) é um processo de obtenção dos parâmetros intrínsecos e extrínsecos de uma câmera. A matriz de parâmetros intrínsecos contem duas matrizes, que são: A matriz de projeção e a matriz de distorção. A matriz de parâmetros extrínsecos, por sua vez, possui outras duas matrizes, que são: A matriz de rotação e a matriz de translação. Esses parâmetros foram descritos computacionalmente por Faugeras e Toscani(1986).

As matrizes de parâmetros extrínsecos são utilizadas para descrever o movimento da câmera quando a mudamos de posição para capturar uma mesma cena sobre diversos pontos de

vista, por exemplo, duas imagens de mesma cena capturadas em duas posições diferentes. Em uma *Focal Stack* a cena é capturada em um mesmo ponto de vista, portanto sem variação da posição da câmera. O efeito de *zoom* é provocado pelo movimento da lente sobre a coordenada *Z* conforme representado na Figura 14, movimentando os pixels da imagem por ampliação (LEE; AHN; MEYYAPPAN, 1997). A curvatura da lente também gera o deslocamento dos pixels da imagem devido à aberração da curvatura angular da lente (TSAI, 1987), mas que se mantem para a mesma localização. Portanto para *Depth From Focus* a calibração da câmera é obtida pelo cálculo das matrizes de distorção e de projeção, sendo computado pela translação.

FIGURA 14 - Projeção escalar da imagem em câmeras pinhole



Fonte: Adaptado de UCI (Acessado em 24Jun2016).

Se $f(u,v)$ corresponde aos valores de intensidade de cor da imagem digital então as coordenadas dessa imagem, u e v , podem ser representadas por um vetor bidimensional, conforme o seguinte: (GONZALEZ; WOODS, 2008, p.59)

$$[u \ v] \quad (17)$$

Uma vez que a operação de *zoom* é obtida pela movimentação do conjunto de coordenadas da imagem, podemos utilizar a transformação afim de Gonzalez e Woods (2008, p. 55) multiplicando a matriz da Equação 17 por um valor escalar, e obter as novas coordenadas dos pixels da imagem, conforme a Equação seguinte:

$$[x \ y] = [u \ v] * w \quad (18)$$

Onde: w é um número real.

u e v são as coordenadas do plano de coordenadas da imagem

x e y são as coordenadas ajustadas para a posição correta.

O processo de compensação de *zoom* da lente foi inicialmente descrito por Darrell e Wohn(1988). Kim e Tian(2009) empregaram o processo de compensação por *zoom* no

algoritmo SFF de Nayar e Nakagawa(1994), em forma de operação escalar. Jing e Yeung (2012) propuseram uma alternativa escalar para empregar a matriz de transformação afim.

Kim e Tian(2009) calcularam o valor da ampliação do tamanho do objeto, correspondente a q , pela divisão da distância da lente para o objeto e a distância da lente para a imagem.

$$q = \frac{v}{u} = \frac{d}{u} \quad (19)$$

Onde : q valor de ampliação da imagem

u, d distância do objeto para a lente

v, D distância da imagem para a lente

Para Jing e Yeung (2012) a diferença entre duas fotos pode ser descrita como uma transformação afim, seguindo a equação abaixo.

$$F_i(x) = F_j(u(x; \theta)), i < j \quad (20)$$

$$\begin{bmatrix} x_j \\ y_j \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

O re-arranjo dessa transformação para o fluido óptico da imagem, gera a Equação 21, e, deve ser resolvida para encontrar os coeficientes da matriz de transformação(a_{11} a a_{23}). Esta matriz é multiplicada sobre cada coordenada da imagem (x, y) representada no lado a direito da Equação 21.

$$\begin{bmatrix} F^x x_i \\ F^x y_i \\ F^x \\ F^y x_i \\ F^y y_i \\ F^y \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{21} \\ a_{22} \\ a_{23} \end{bmatrix} = -F^t + F^x x_i + F^y y_i \quad (21)$$

Se o sistema de projeção for colocado em coordenadas cartesianas do R^3 , cada plano bidimensional da Equação 21, encontra-se em paralelo com outros planos posicionados em diferentes pontos no eixo de coordenadas Z . Se a posição inicial do ponto Z for igual a 1 então a equação será descrita pela Equação 22, onde a movimentação dos pixels ocorre em forma de projeção sobre a coordenada Z .

$$[x \ y \ w] = [u \ v \ 1] * w \quad (22)$$

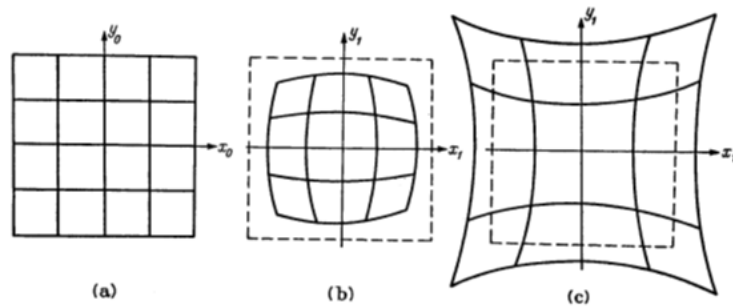
Onde: w é chamado de fator de escala

$$u = \frac{x}{w} \text{ e } v = \frac{y}{w}$$

O sistema acima é conhecido como sistema de coordenadas homogêneas. Podemos imaginar um sistema de coordenadas homogêneas 2D, como o posicionamento do plano $[u \ v]$ na localização w do eixo Z de um sistema 3D qualquer.

A distorção radial da lente foi tratada computacionalmente por Tsai(1987), e posteriormente revisado por Weng, Cohen e Herniou (1992). Esta curva é devido ao formato angular que a lente possui conforme a Figura 15.

FIGURA 15 - Distorção radial da lente da câmera.



Fonte: Born e Wolf (1975, p.217)

O OpenCV utiliza a Equação abaixo para corrigir o posicionamento radial das coordenadas (x,y) da imagem.

$$\begin{aligned} x_{\text{corrigido}} &= x(1 + k_1r^2 + k_2r^4 + k_3r^6) \\ y_{\text{corrigido}} &= y(1 + k_1r^2 + k_2r^4 + k_3r^6) \end{aligned} \quad (23)$$

Onde: $x_{\text{corrigido}}$ é a coordenada x após a correção da distorção radial

$y_{\text{corrigido}}$ é a coordenada y após a correção da distorção radial

k_1 , k_2 e k_3 são coeficientes de distorção radial

r^2 é igual a $x^2 + y^2$.

Para a distorção tangencial a Equação 27 é utilizada. (OPENCV)

$$\begin{aligned} x_{\text{corrigido}} &= x + [2p_1xy + p_2(r^2 + 2x^2)] \\ y_{\text{corrigido}} &= y + [p_1(r^2 + 2y^2) + 2p_2xy] \end{aligned} \quad (24)$$

Onde: $x_{\text{corrigido}}$ é a coordenada x após a correção da distorção tangencial

$y_{\text{corrigido}}$ é a coordenada y após a correção da distorção tangencial

p_1 e p_2 são coeficientes de distorção tangencial

r^2 é igual a $x^2 + y^2$.

Costumeiramente os algoritmos SFF não tratam o problema de aberração esférica da lente, pois o deslocamento de pixels na imagem é mantido na medida em que é obtida a mesma coordenada no mesmo ponto de visão.

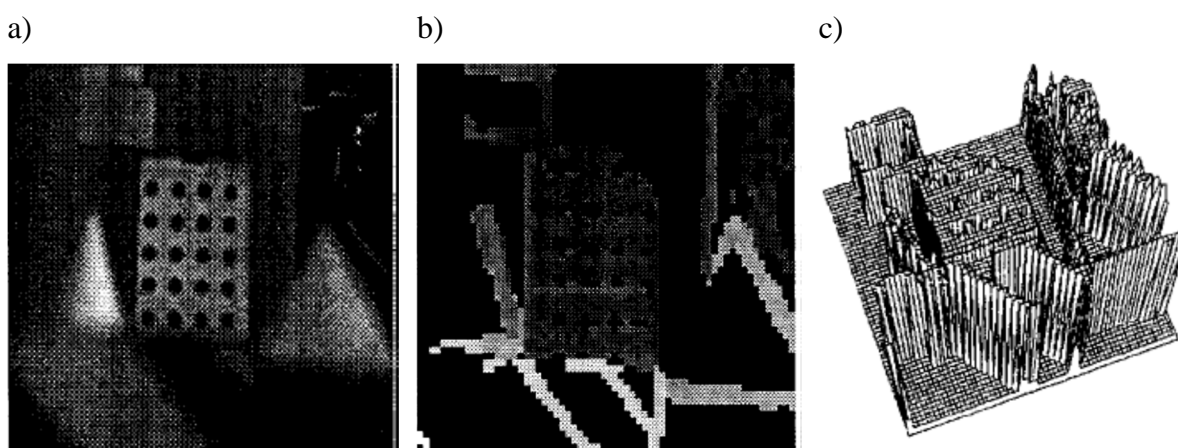
3 TRABALHOS RELACIONADOS

A aglutinação de valores para tornar a medida de profundidade mais confiável foi apresentada na seção 2.7 e tem sido o principal objeto de estudo dos recentes trabalhos de SFF.

Inicialmente, o uso da agregação de valores, utilizando os princípios básicos de focalização, foi realizado em aplicações de autofoco. Nessas aplicações a agregação foi utilizada dentro de uma função critério para encontrar uma região em foco na imagem.

Desde as primeiras aplicações de mapas de profundidades, a agregação de valores vem sendo empregada nos medidores de qualidade de foco dos algoritmos que utilizam os princípios de foco. No início a agregação foi utilizada apenas para compensar os efeitos que o reposicionamento da lente da câmera produzia no deslocamento dos pixels da imagem da *Focal Stack*. Darrell e Wohn(1989) utilizaram a agregação para construir mapas de profundidade esparsos com janelamento não sobreposto, onde cada janela fornecia a profundidade de uma região quadrática. Segundo esse artigo, se a imagem possuísse 256 linhas e 256 colunas e se o tamanho da janela de agregação fosse de 4 linhas e 4 colunas, então o mapa de profundidade possuiria 64 regiões na vertical por 64 regiões na horizontal de profundidades. Na Figura 16, 16.a é uma imagem da Focal Stack, com 256 linhas por 256 colunas. Em 16.b é o mapa de profundidade com 64 linhas por 64 colunas de regiões e em 16.c é exibido o mesmo mapa de 16.b, mas em visualização de perspectiva.

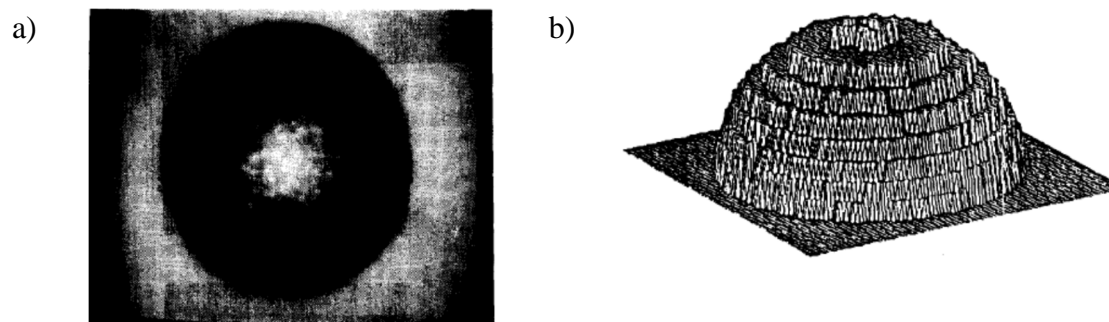
FIGURA 16 - Mapa de profundidade esparsos com janelamento não sobreposto



Fonte: Darrell e Wohn(1989).

Nayar e Nakagawa(1994) utilizaram a agregação para gerar mapas de profundidades locais empregando janelamento sobreposto e desprezar o efeito de zoom gerado pela lente da câmera no conjunto de imagens da pilha de foco. Segundo esse artigo se uma imagem possuísse 256 linhas por 256 colunas de pixels(17.a), o seu mapa de profundidade (17.b) possuiria, também, o mesmo numero de linhas e colunas, conforme a Figura 17.

FIGURA 17 - Mapa de profundidade de local com janelamento sobreposto.



Fonte: Nayar e Nakagawa(1994)

Se uma imagem é obtida através de amostras de intensidades luminosas do mundo real (Seção 2.2), então o movimento da lente para obter uma *Focal Stack*, provavelmente provocará a obtenção de novas amostras diferentes, para formar uma outra imagem, e que será, possivelmente, vizinha às amostras obtidas anteriormente, portanto o tratamento da movimentações de pixels ainda pode ser descrito como uma das utilidades da agregação de valores no medidor de qualidade de foco.

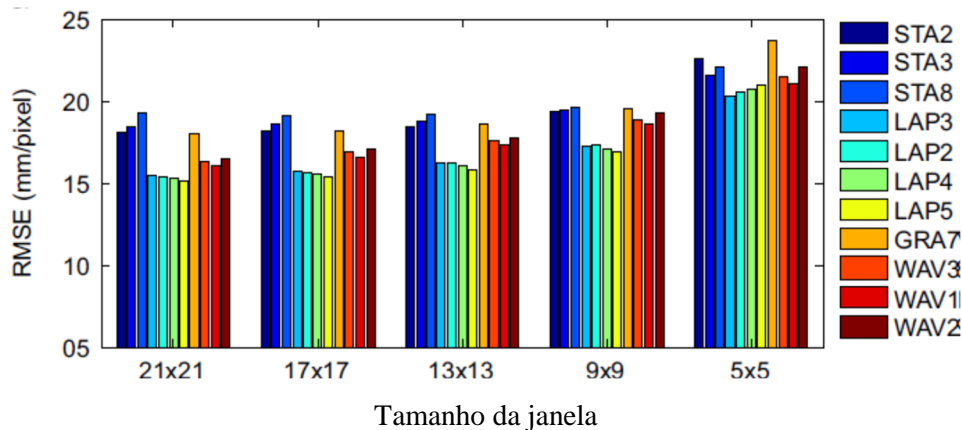
Posteriormente, a agregação de valores foi usada para aumentar a confiança do medidor de qualidade de foco na presença de outros ruídos que provocam a degradação da imagem, como por exemplo: o excesso ou a falta de luminosidade, o ruído gaussiano, imagem com pouco contraste, etc., conforme o Gráfico 6 e a Tabela 7(GAGANOV; IGNATENKO, 2009; PERTUZ, PUIZ; GARCIA, 2013, p.1415).

TABELA 7 - Amostras com a introdução de ruídos utilizando diversas configurações.

Grupo	Janela(Pixel)	Ruído	Contraste(%)	Saturação(%)
Lv1	21 × 21	0.00050	80.0	10
Lv2	17 × 17	0.00176	62.5	20
Lv3	13 × 13	0.00320	45.0	30
Lv4	9 × 9	0.00429	27.5	40
Lv5	5 × 5	0.00555	10.0	50

Fonte: Adaptado de PERTUZ, PUIZ; GARCIA(2013, p.1422).

GRÁFICO 6 - Avaliação de medidores de qualidade de foco, para ruído e tamanho da janela



Legenda:

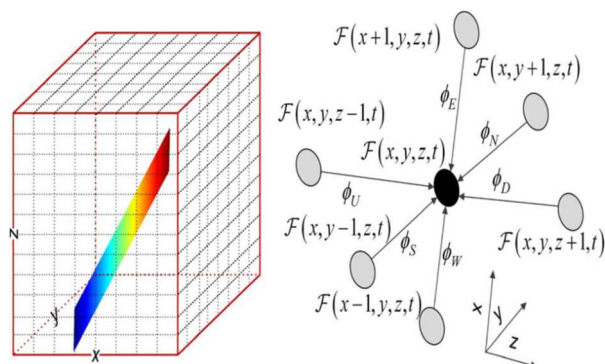
Famílias de medidores de qualidade de foco:

- STA- medidores baseado em medidas estatísticas como por exemplo a variância dos níveis de cinza;
- LAP – medidores baseados em segundas derivadas, tipo laplaciano;
- GRA- medidores baseados em primeiras derivadas, tipo gradiente, e
- WAV – medidores baseados em Wavelets, no domínio tempo/ frequências.

Fonte: Adaptado de Pertuz et al.(2013, p.1423).

Mahmood e Choi(2012) também usaram um agregamento de valores em uma janela 3D para melhorar a resposta desse processo com imagens ruidosas, onde a média da soma de outras janelas em outras imagens no medidor atual, minimizaria a parte do medidor que esteja com ruído, conforme a Figura 18.

FIGURA 18 - Janelamento 3D do medidor de qualidade de foco.

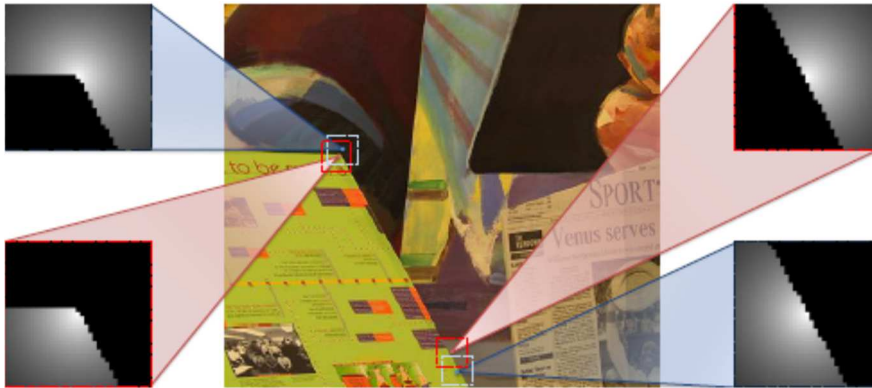


Fonte: Mahmood e Choi (2012, p.2869).

Recentemente a agregação de valores vem sendo utilizada como forma de unir pixels de diferentes profundidades de campo dentro do medidor de qualidade de foco. Nesse contexto Aydim e Akgul(2008) propuseram pensar a agregação de valores do medidor de qualidade de foco com um suporte adaptativo, onde a distância euclidiana e a diferença de cores, entre os pixels fornecesse um peso para os medidores de nitidez. A função de validação não exclui o valor do medidor de foco, somente atribui a ele um peso para ser multiplicado quando for

calculada a agregação de valores. Na Figura 19 as sub-imagens destacadas, pintadas em preto e branco, são exemplos de mapas adaptativos para pixels com pesos para as diversas regiões da imagem, onde quanto mais preto menor valor possui.

FIGURA 19 - Imagem com mapas de diferentes pesos de profundidade



Fonte: Aydim e Akgul (2008).

Em um trabalho mais recente de Aydim e Akgul(2010) propuseram tratar a oclusão no algoritmo SFF, onde a aplicação da Equação de correlação cruzada normalizada(Equação 25), no estágio de obtenção do medidor de foco, separa os pixels que se ligam fora do foco. Na Figura 20, o ponto P do mundo real é refletido no medidor de qualidade de foco q , quando a configuração da lente alcança determinada abertura, contribuindo, também, para o borramento de uma região de foco. Nesse artigo a janela adaptativa também é utilizada com os mesmos parâmetros do trabalho de Aydim e Akgul(2008).

$$NCC_{\Omega}(A, B) = \frac{\sum(A - \bar{A})(B - \bar{B})}{\sum \sqrt{\sum(A - \bar{A})^2 \sum(B - \bar{B})^2}} \quad (25)$$

$$FM_i(x, y) = NCC_{\Omega}(I_f(x, y), I_i(x, y))$$

Onde: NCC é o medidor de correlação cruzada dos conjuntos imagens A e B .

FM_i é o medidor de nitidez/qualidade de foco para a posição x, y
 x e y são a linha e coluna do pixel.

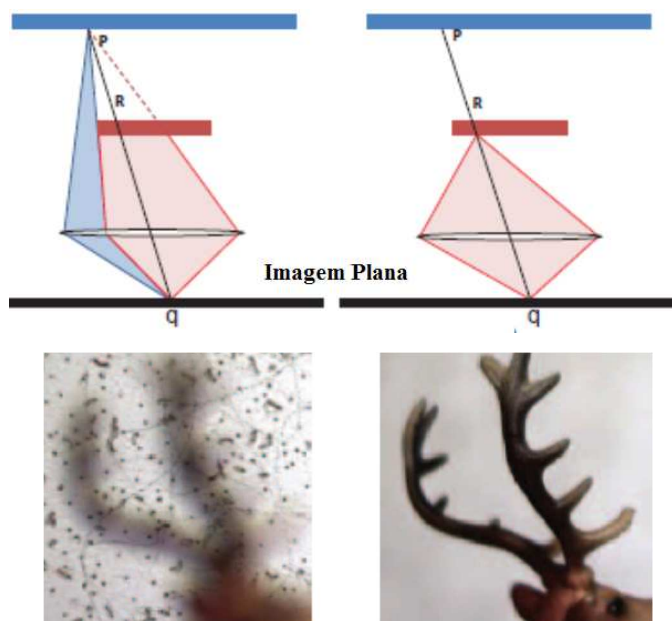
$I_{f(x,y)}$ é a janela da sub-imagem focalizada da cena.

$I_{i(x,y)}$ é a janela da i -ésima sub-imagem do conjunto de foco.

Ω é a janela de avaliação do medidor de qualidade de foco.

A e B são as sub-imagens avaliadas.

FIGURA 20 – Problema de oclusão em algoritmos SFF.

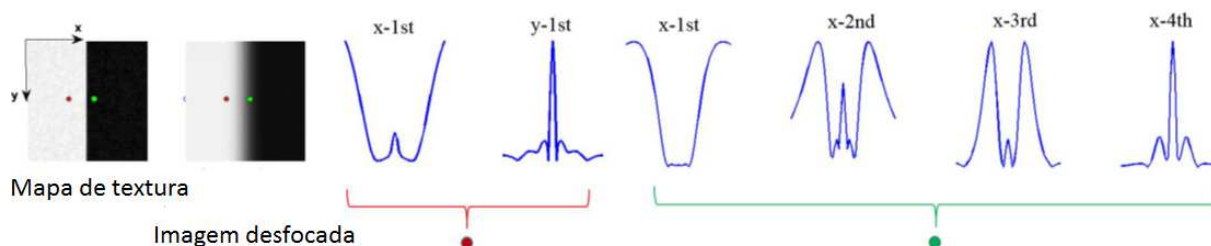


Fonte: Adaptado de Aydim e Akgul(2010).

Jing e Yeung(2012), propuseram utilizar Segmentação para calcular o peso dos elementos da função de medida de foco. Em geral os trabalhos que utilizam uma janela com suporte adaptativo usam grande profundidade de campo e sua metodologia é baseada em análise das características de uma imagem toda em foco.

Frommer et al.(2015) utilizaram a alta ordem de derivadas para atribuir o peso aos medidores de nitidez da janela com suporte adaptativo. Segundo esse artigo a derivada de uma direção é inversa a derivada da outra e por isso elas tendem a se anular(nas direções dos eixos x e y da Figura 20). Utilizando a alta ordem derivativa esse problema poderia ser eliminado.

FIGURA 21 – Derivadas de duas direções para uma amostra simples.

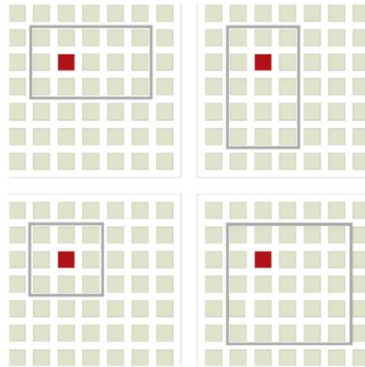


Fonte: Adaptada de Frommer et al.(2013).

Em outros trabalhos a profundidade de cada pixel é avaliada sob a variância da região da imagem, para modificar o tamanho do janelamento. Lee et al.(2013) propuseram utilizar a média das intensidades de cor de uma janela pra calcular a dispersão desses valores, chamando de desvio médio absoluto(MAD), e compará-la a um limiar para identificar a necessidade de crescimento da janela.

Lee et al.(2013) utilizaram o conceito de Semivariograma para determinar o tamanho da uma janela, utilizando diferentes tamanhos tanto na vertical quanto na horizontal.(Figura 22).

FIGURA 22 - Modelos de janelas segundo Lee et al.



Fonte: Lee et al.(2013)

Outra forma de extrair a variação do pixel é transformar o processo de janelamento em pós-processamento, regularizando posteriormente o mapa de coordenadas (LIU; et al., 2010, p.148).

Moeller et al.(2015) abordaram o problema DFF como reconstrução robusta, aplicando os mesmos princípios do framework MRF, ao usar o framework variacional com conceito chamado de *método alternativo de minimização de multiplicadores*(ADMM) que une pixels com semelhanças, para tornar o mapa de profundidade mais confiável.

Liu e Wu.(2015), usaram a segmentação baseada na cor local como um suporte do janelamento e o framework variacional, para reconstruir mapas de profundidades no seu método semi-global DFF.

O suporte adaptativo também pode ser apresentado como uma filtragem anisotrópica bilateral, onde a estrutura local é obtida usando os valores dos pixels (MAHMOOD, M. T. ; et al. , 2011)

4 MÉTODO

O estudo do suporte adaptativo, para medidores de qualidade de foco com janelamento quadrático no método SFF, foi realizado utilizando quatro algoritmos conhecidos de mapeamento de mapas de profundidade por extração de pontos focais. Esses quatro algoritmos foram aplicados sobre dois grupos de amostras de imagens digitais, sendo um grande conjunto de amostras sintéticas e outro conjunto de amostras reais.

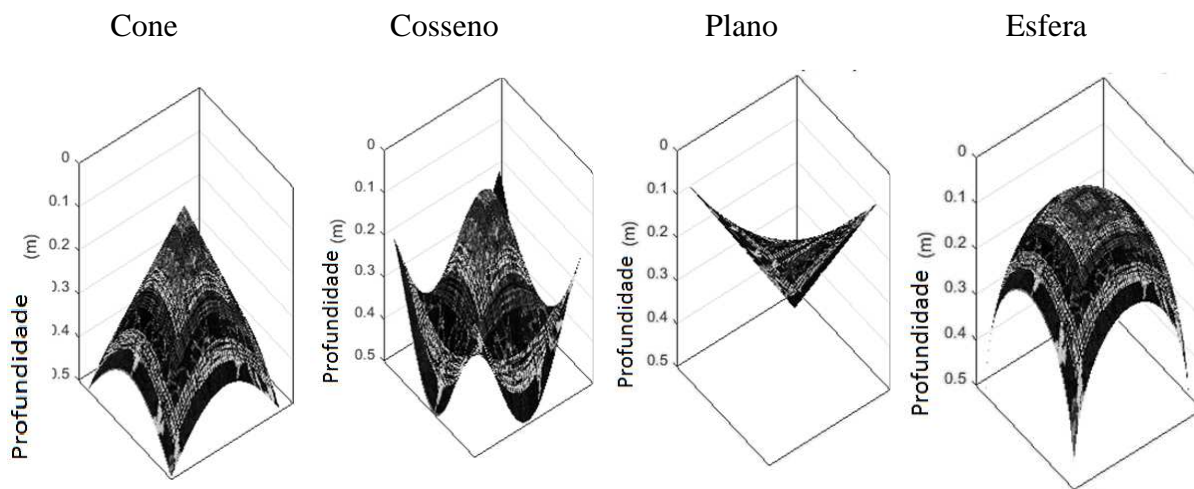
As amostras sintéticas foram produzidas de duas formas: a primeira utilizando o código de programa para MATLAB, implementado por Said Pertuz(2013) em seu trabalho para simulação de desfocalização, e o segundo grupo utilizou o software Photoshop em conjunto com imagens e seus mapas de profundidade disponíveis na base de imagens online Middlebury.

As amostras de imagens reais foram produzidas pelo autor e totalizaram 30 imagens, sendo utilizadas apenas para observação visual dos reais efeitos dos algoritmos avaliados.

A produção dos dois tipos de amostras sintéticas ocorreu devido à seu emprego na maioria dos trabalhos anteriores para avaliar de forma objetiva os resultados dos diversos tipos de mapas de profundidades(objetos simples e cenas realísticas) construídos por esses algoritmos.

O algoritmo de simulação de desfoco, para objetos simples, gera imagens parcialmente focalizadas na simulação da pilha de foco do primeiro grupo de imagens sintéticas. A configuração básica deste algoritmo produz 15 imagens de tamanhos iguais a 360 linhas por 360 colunas, que compõem a *Focal Stack* simulada, um mapa de profundidade verdadeiro (real), de mesmo tamanho das imagens produzidas, e um vetor com 15 distâncias focais(real), sendo os dois últimos normalizados e com precisão de 8 bytes(*double*). O algoritmo pode simular quatro tipos de superfícies, cone, plano, cosseno e esfera, conforme a Figura 23, sendo possível ainda definir o nível de *blur* desejado. Em nossos experimentos iniciais não foi identificado alteração no resultado para alteração do nível de *blur*.

FIGURA 23 – Superfícies sintéticas .

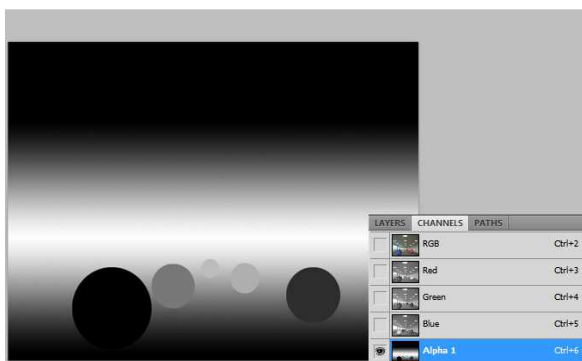


Fonte: Produzida pelo autor.

Para criar o segundo conjunto de imagens sintéticas foi utilizado o método citado nos trabalhos de Gaganov e Ignatenko(2009) e também David et al. (2012), onde a pilha de foco é construída usando o software Adobe Photoshop CS5 e uma a base de dados de mapas de profundidade², conforme a Figura 24.a.

FIGURA 24 - Construção de Mapa de profundidade sintético usando Photoshop.

a)



b)



Fonte: Oprinca, (Abril de 2011, acessado em 10jun2016)

Para construir a pilha de foco no Photoshop foi usada a opção de menu *filter/blur/lens blur* modificando o valor da opção *Blur Focal Distance* da aba suspensa do programa Photoshop(Figura 24.b). Foi utilizado uma lente simulada de 18mm e desprezado os ruídos do meio, bem como as distorções possíveis da lente.

Os mapas de profundidade simulados podem ser construídos no próprio software Photoshop. O processo pode ser resumido da seguinte forma: Inicialmente é carregada uma imagem totalmente focalizada(Figura 24.b), depois são usados pincéis de seleção (*Quick*

² disponível em <http://vision.middlebury.edu>.

Selection Tool) para marcar objetos e montar um mapa de gradientes (*Paint/Gradient*), de forma que os tons mais claros correspondam as posições mais distantes. Esse processo de simulação de mapa de profundidade no software da Adobe, é detalhado no site de tutoriais chamado PSD BOX³. Foi utilizada a base de imagens online Middlebury Stereo. As imagens desta base de dados possuem diferentes tamanhos conforme a identificação nesta base, a saber: ART: 463 linhas por 370 colunas; REMINDEER: 447 linhas por 370 colunas e VENUS:434 linhas por 383 colunas.

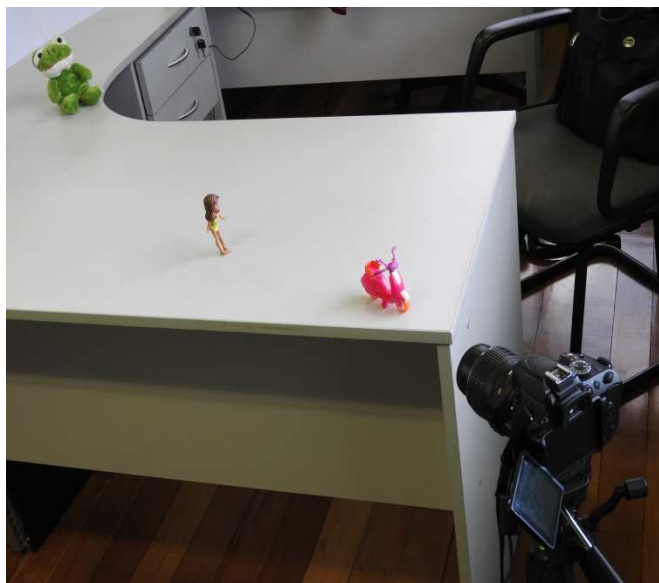
Para melhorar a qualidade das amostras de imagens produzidas segundo o método dois de produção de imagens sintéticas, o mapa de profundidade foi ajustado preenchendo os valores zerados com o(s) seu(s) vizinho(s), conforme o caso. Este ajuste foi necessário para eliminar a ausência de profundidade nessas regiões. Também, para otimizar a construção das amostras, foi mapeado como posição inicial de foco o menor valor de profundidade, e como posição final de foco o maior valor de profundidade, de modo que, posições focais entre o menor valor e o valor zero não gerem amostras e posições entre o maior valor e 255 não gerem amostras desnecessárias.

Diferente de um mapa de profundidade construído com imagens do mundo real, em uma *Focal Stack* simulada é possível eliminar, completamente, a distorção por curvatura da lente e medir matematicamente, de forma objetiva, os resultados do processo de reconstrução do mapa de profundidade. Nos mapas de profundidades construídos a partir de imagens reais, a avaliação é feita somente pela comparação ou avaliação visual.

Para produzir o conjunto de amostras de imagens reais foi utilizada uma câmera Nikon D5200 com lente 18-55 mm 3.5/5.6 e três bonecos dispostos à frente dela, conforme a Figura 25.

³ Disponível em <http://www.psdbox.com/tutorials/photoshop-depth-maps-realistic-depth-of-field>.

FIGURA 25 - Obtenção da Focal Stack



Fonte: Produzida pelo autor.

A obtenção dessas amostras foi realizada através de um software construído pelo autor desse trabalho, em linguagem C, para reposicionar a lente, selecionar a abertura da lente e capturar imagens.

Inicialmente o software realiza a configuração da abertura da lente e posiciona a lente para mais próximo do sensor da máquina (respeitando o limite de construção da lente), em seguida reposiciona o motor da lente, em pontos espaçados empiricamente e à frente, a fim de obter novo ponto focal. A cada reposicionamento da lente, uma imagem é capturada e novamente o processo se repete até completar 30 imagens conforme a Figura 26. As imagens capturadas foram obtidas com tamanho de 640 linhas e 480 colunas no formato Bitmap(BMP).

FIGURA 26 - Construção da Focal Stack de imagens reais



Fonte: produzida pelo autor.

Após experimentos com esta lente, foram escolhidas a configuração de 55mm e abertura 3.5, devido fornecerem maior profundidade de campo. Os espaçamentos focais foram obtidos empiricamente através de testes, pois dependem de cada lente.

4.1 Ferramentas utilizadas

Foi construída uma aplicação em linguagem C/C++ para compor os quatro algoritmos de avaliação dos métodos apresentados e utilizado o software MATLAB como ferramenta de apoio, gerando imagens sintéticas com o algoritmo de simulação proposto por Pertuz(2013). O software MATLAB também foi utilizado para outros suportes de conversão de valores, como por exemplo, gerar um arquivo no formato binário(bin) com os dados de distâncias focais em precisão *double* do algoritmo de Pertuz(2013), a fim de serem lidos pelo programa de avaliação dos métodos que foi construído em linguagem C/C++. Outros software como Photoshop CS6 foram utilizados para gerar a pilha de foco(Focal Stack) e, também, para conversão de formato de arquivos da base Middlebury Stereo para o formato aceito pelas ferramentas utilizadas, como por exemplo converter o formato PNG da base de dados Middlebury, no formato TIF do algoritmo de simulação de desfoco de Pertuz(2013).

O framework de construção de aplicativos para a Máquina fotográfica D5200 da empresa Nikon, também, foi utilizado para construir o software de captura de imagens da pilha de fotografias de imagens reais.

O sistema operacional utilizado para realizar os experimentos foi o Microsoft Windows nas versões 7.0, 8.0 e 8.1 e a execução foi realizada em um notebook com processador Core i3 de terceira geração, 4GB de RAM e clock de 1.5GHZ e um desktop com processador Core i7 de quarta geração, 8GB de RAM, clock de 3.4GHZ e placa de vídeo NVIDIA Geforce GT 640 (384 Shaders).

4.2 Implementação dos algoritmos utilizados.

Embora existam códigos fonte, de alguns modelos estudados, disponíveis na internet, o emprego de diferentes linguagens poderia mascarar resultados, por isso o autor optou por confeccionar um novo código fonte em linguagem C/C++ e aumentar a confiança dos experimentos.

4.2.1 Método *Shape From Focus* de Nayar(89)⁴.

O método de Nayar e Nacagawa(1994) foi apresentado na seção 2.5, 2.6 e 2.8, sendo usado seu modelo como base de comparação com os demais métodos. Dentre os medidores que podem ser aplicados no método, a Soma dos Laplacianos Modificados vem apresentando melhor desempenho na presença de ruídos, conforme o estudo de Pertuz(2013).

⁴ Disponível em http://www1.cs.columbia.edu/CAVE/publications/pdfs/Nayar_ICRA90.pdf.

Inicialmente foi usado os dois modelos de reconstrução, fina e grosseira, para avaliação do mapa de profundidade de imagens sintéticas, e posteriormente selecionado apenas o método de reconstrução que apresentou melhor desempenho.

Objetivando utilizar uma plataforma de comparação, o algoritmo deste método fez parte do programa construído em linguagem C/C++.

4.2.2 Método *Adaptive Focus Measure* de Aydin Akgul(2008).

O suporte adaptativo para janela do medidor de qualidade de foco foi apresentado na seção 2.7, e este, também, serviu de base para a comparação dos demais métodos de seleção. Um recurso usado neste método é o emprego da imagem toda focalizada para calcular o suporte adaptativo. Os demais critérios de reconstrução deste método são semelhantes ao método *Shape From Focus*. O algoritmo deste método também fez parte do programa desenvolvido na linguagem C/C++.

4.2.3 Método *Adaptive Depth From Focus* de Jing e Yeung(2012).

Diferente do método de Aydin e Akgul(2008), este método usa como suporte adaptativo uma segmentação baseada na técnica *Mean-Shift Segmentation*. Como o objetivo é medir o impacto do suporte na reconstrução do método SFF, foi utilizado apenas o modelo de suporte adaptativo descrito no método e desprezado os demais recursos empregados pelo autor na reconstrução do mapa de profundidade. O modelo do artigo não realizou a reconstrução fina, e isso sugere seu melhor desempenho para reconstrução grosseira.

Para realizar a segmentação descrita no artigo de Jing e Yeung(2012) foi utilizado o software fornecido pela The State University of New Jersey- RUTGERS⁵, que foi estudado no trabalho de Comaniciu e Meer(2002). Este software converte o sistema de cor RGB, da imagem, para o formato LUV e então realiza a segmentação *Mean-Shift*. Outra característica deste programa é a necessidade de fornecimento de um tamanho mínimo para a área segmentada.

Para selecionar a região fornecida por esta segmentação foi utilizado um algoritmo de seleção de região, onde os pixels dentro de uma janela, que são iguais ao do centro, são selecionados e os demais descartados. Este algoritmo de seleção de região bem como o algoritmo deste método, fizeram parte do programa construído em linguagem C/C++.

⁵ Disponível em <http://coewww.rutgers.edu/riul/research/robust.html>

4.2.4 Método *Semi-Global Depth From Focus* de Liu e Wu(2015)

Assim como no método de Jing e Yeung(2012), este método utiliza a segmentação, porém baseada em cores. Semelhante a metodologia utilizada no método citado anteriormente não foi implementado os demais recursos utilizados pelo artigo, pois o objetivo deste trabalho é medir o impacto do método de seleção do suporte para a janela do medidor de qualidade de foco. Outra semelhança com o método anterior é que o método não realiza a reconstrução fina.

Este trabalho utiliza a maior variação entre os três canais de cores RGB do pixel para comparar semelhança com o pixel que esta sendo medido, e utiliza um cálculo de limiarização para estabelecer o alcance de similaridade. Por motivo de performance foi desprezado este cálculo e selecionado um limiar empírico de 5 pixels.

Para selecionar a região fornecida por esta segmentação, também foi utilizado um algoritmo de seleção de região igual ao pixel. Este algoritmo fez parte do programa construído em Linguagem C/C++

4.3 Experimentos e Resultados.

4.3.1 Descrição dos experimentos.

Baseado nos trabalhos de de Nayar(89), Aydin Akgul(2008), Jing e Yeung(2012) e Liu e Wu(2015), foi avaliado o desempenho das amostras de imagens sintéticas para os quatro algoritmos produzidos.

Os quatro tipos de superfícies geradas sinteticamente bem como daquelas construídos a partir de um mapa de profundidade da base de imagens *Middlebury Stereo* foram submetidos a esses algoritmos.

Em todos os algoritmos foi utilizado seis tamanhos de janelas bidimensionais quadráticas: 5x5, 9x9, 15x15, 21x21, 31x31 e 41x41 quadráticas e a seleção dos parâmetros para o calculo do suporte adaptativo foi empírica. A comparação não computou no erro a borda do mapa de profundidade, pois o suporte não fornece medidores de nitidez para toda a janela do medidor, de modo a descartar os possíveis valores subjetivos.

Para medir o grau de semelhança entre o mapa gerado e o mapa ideal(real) foi utilizada a raiz quadrada do erro médio quadrático (RMSE), conforme a Equação 26, ajustando o resultado de distâncias para pixels. O cálculo de RMSE foi obtido do trabalho de Pertuz(2013), Mahmood(2013) e David et. al(2012), e o resultado de sua implementação foi comparado com o resultado obtido pelo código MATLAB⁶. O erro RMSE mede o grau de proximidade em

⁶ Disponibilizado em <https://www.mathworks.com/matlabcentral/answers/4064-rmse-root-mean-square-error>

sinais, portanto valores próximos a zero indicam que o mapa de profundidade encontrado é muito parecido com o mapa de profundidade ideal(real).

$$RMSE = \sqrt{\frac{1}{MN} \sum_{(i,j)} (G_T(i,j) - Z(i,j))^2} \quad (26)$$

Legenda:

RMSE raiz quadrada do erro médio quadrático;

MN Número de linhas vezes o numero de colunas da imagem;

$G_T(i, j)$ Mapa de profundidade real(ideal) nas coordenadas i,j;

$Z(i, j)$ Mapa de profundidade Calculado para as coordenadas i,j;.

Embora o resultado final forneça uma imagem, o cálculo do processo é realizado em ponto flutuante de precisão dupla e não pode ser exibidos como imagem na tela do computador, ou seja tanto o mapa de profundidade verdadeiro, fornecido pelo algoritmo de simulação e desfoco de Pertuz(2013), quanto o mapa de profundidade calculado, pelo algoritmo proposto, não pode ser convertido em imagem antes de terminar o cálculo do erro RMSE.

4.3.2 Resultados

Inicialmente foram computados os resultados dos quatro algoritmos para as quatro superfícies simuladas e, gerado uma tabela para cada superfície. Diferente de Aydin e Akgu(2008), o erro RMSE foi computado em pixels, pois a maioria dos trabalhos usam esta unidade. Na Tabela 8 são apresentados os resultados, respectivamente, para as superfícies cone, cosseno, plano e esfera. Nessa tabela os algoritmos foram nomeados conforme o seguinte:

- SFF : *Shape From Focus*;
- ADP: *Adaptive Focus Measure*;
- MEAN : *Adaptive DFF* [tamanho 70] e,
- COLOR: *Semi-Global DFF*.

TABELA 8- Erro RMSE para os métodos de reconstrução fina e grosseira

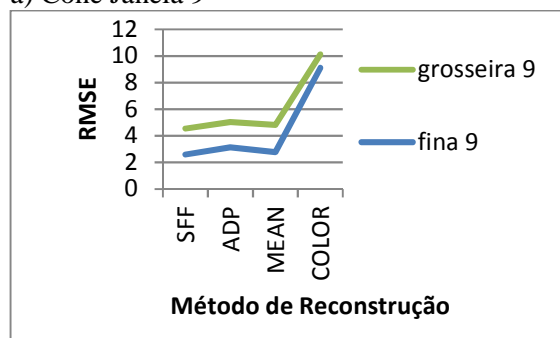
Tipo superfície	Tipo Reconstrução	Janela	Algoritmo			
			SFF	ADP	MEAN	COLOR
Cone	Fina (Interpolação Gaussiana)	9	2,595	3,133	2,784	9,113
		41	2,200	1,286	2,534	20,969
	Grosseira (Máximo a posteriori)	9	4,545	5,032	4,813	10,119
		41	3,124	2,499	3,332	20,951
Cosseno	Fina (Interpolação Gaussiana)	9	2,928	4,254	3,442	12,929
		41	3,537	1,541	3,266	24,376
	Grosseira (Máximo a posteriori)	9	14,306	5,791	4,221	12,159
		41	7,585	2,584	3,936	24,334
Plano	Fina (Interpolação Gaussiana)	9	4,753	8,892	6,136	24,202
		41	1,385	1,388	1,94	29,629
	Grosseira (Máximo a posteriori)	9	5,34	9,237	6,642	24,26
		41	2,506	2,447	2,851	29,45
Esfera	Fina (Interpolação Gaussiana)	9	4,579	8,968	5,779	22,186
		41	2,347	1,833	3,46	44,93
	Grosseira (Máximo a posteriori)	9	4,939	9,113	6,03	22,14
		41	2,968	2,5154	3,822	44,665

Fonte: Produzida pelo autor.

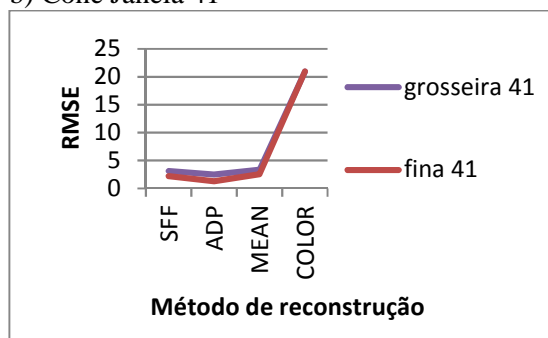
Para analisar estas informações, inicialmente foi plotado oito sub-gráficos, dentro do Gráfico 7, para comparar os modos de reconstrução fina e grosseira. No processo de análise foi identificado que as respostas de cada tipo de reconstrução para os quatro algoritmos podem ser analisadas tanto por um ou outro método de reconstrução, conforme os Sub-gráficos 7.a, 7.b, 7.c, 7.d, 7.e, 7.f, 7.g e 7.h, pois o comportamento similar é reproduzido em forma de linhas paralelas ou quase paralelas nos gráficos.

GRÁFICO 7 – Comparação dos métodos de reconstrução Fina e Grosseira

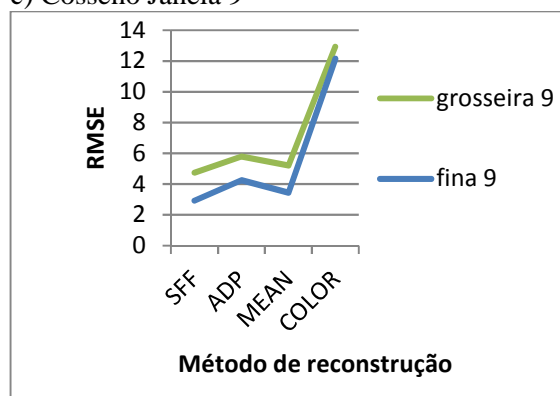
a) Cone Janela 9



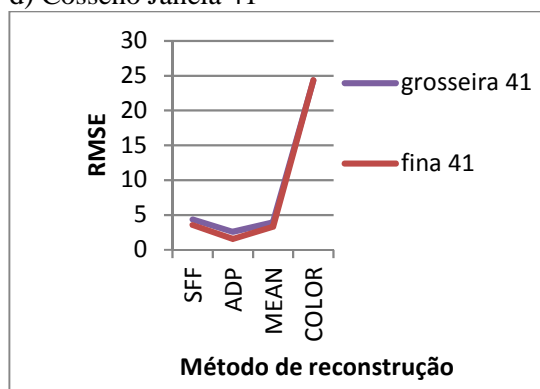
b) Cone Janela 41



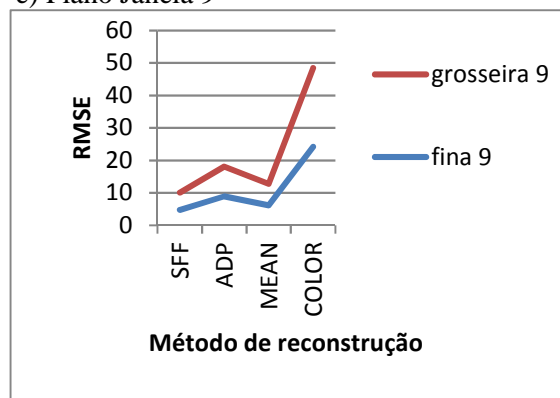
c) Cosseno Janela 9



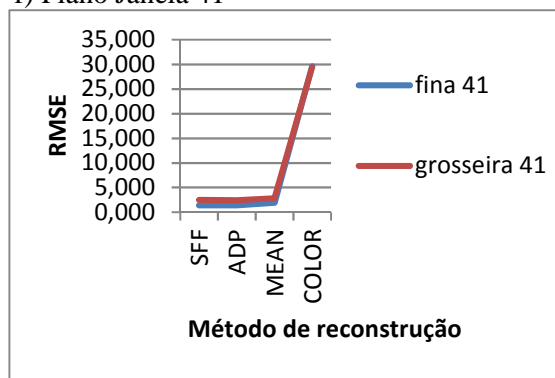
d) Cosseno Janela 41



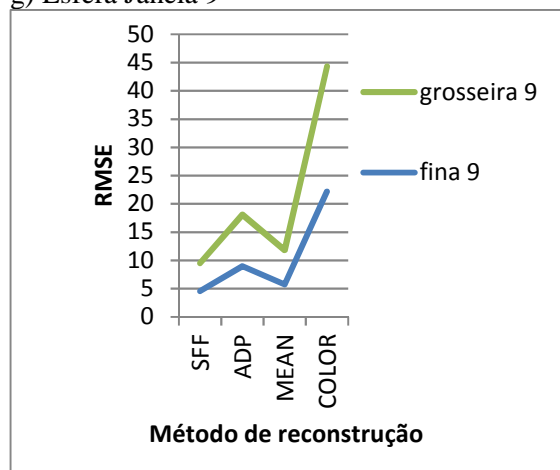
e) Plano Janela 9



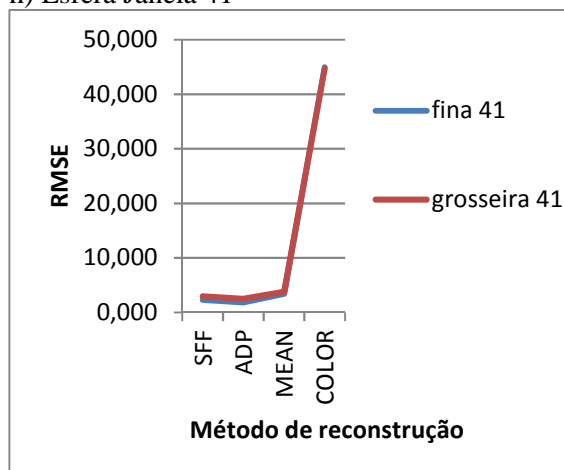
f) Plano Janela 41



g) Esfera Janela 9



h) Esfera Janela 41

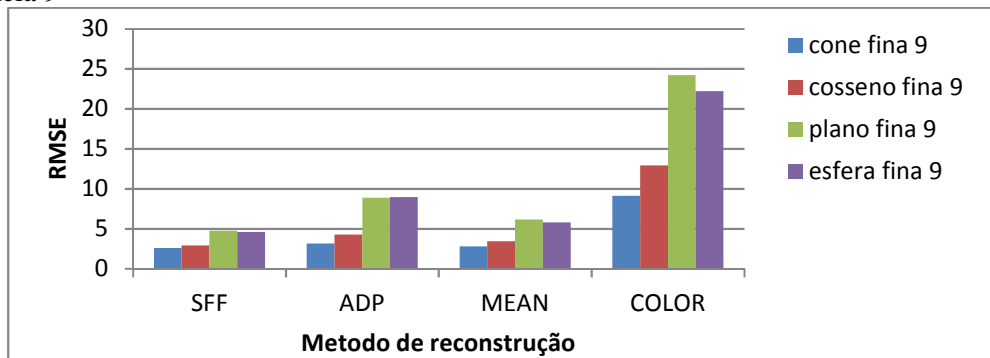


Fonte: Produzida pelo autor

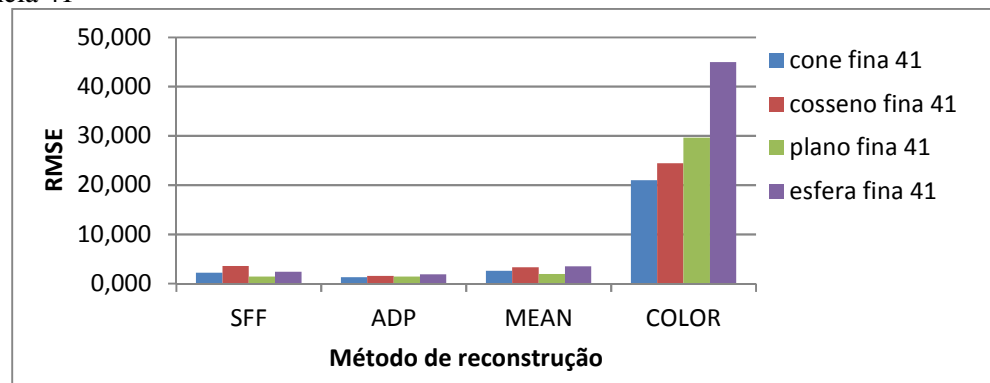
Para comparar o desempenho do tamanho do janelamento, inicialmente, foi plotado quatro sub-gráficos no Gráfico 8. Após análise dos resultados foi identificado o decréscimo do RMSE ao aumentar o tamanho da janela, exceto para o algoritmo COLOR.

GRÁFICO 8- Comparação entre os janelamentos e os métodos de reconstrução

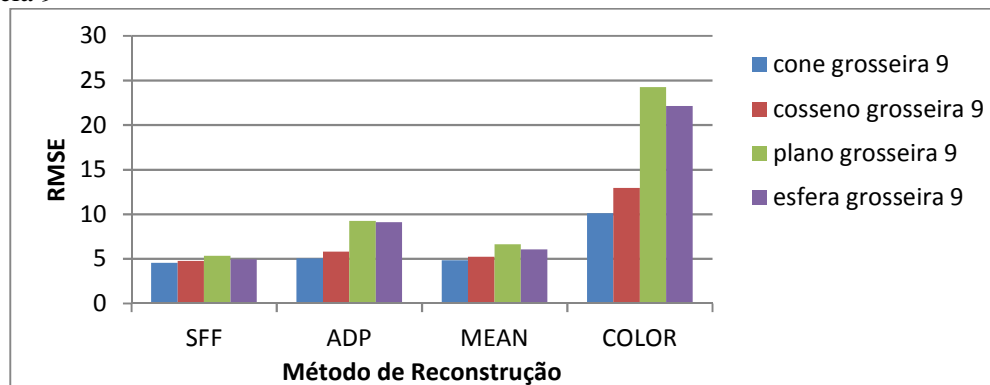
a) Janela 9



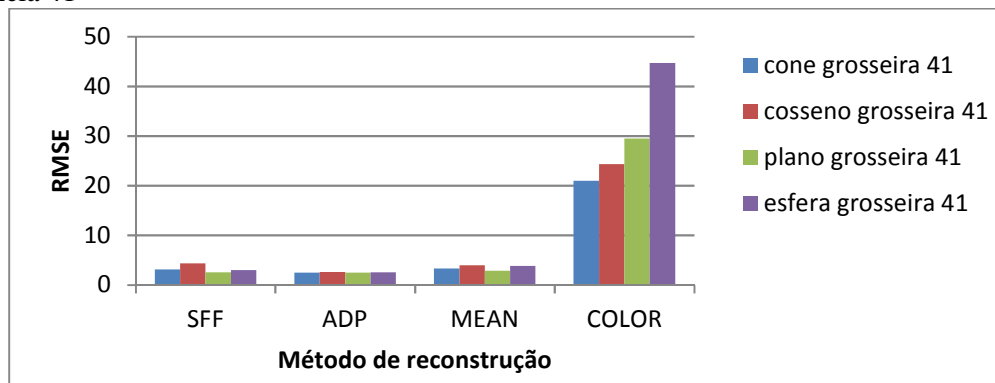
b) Janela 41



c) janela 9



d) janela 41



Fonte: Produzida pelo autor.

Uma segunda amostra mais detalhada foi computada para diversos valores de janelamento, e são apresentados na Tabelas 9. Nesse momento foi escolhida a reconstrução grosseira para o restante das análises.

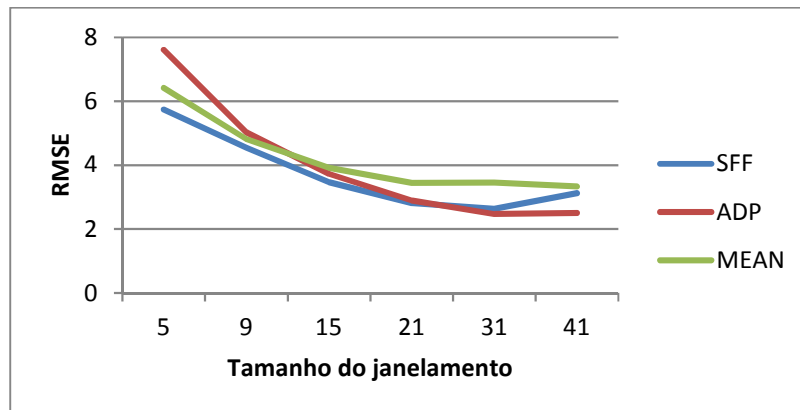
TABELA 9 Erro RMSE usando tamanhos variados no janelamento

superfície	Tipo Reconstrução	Janela	Algoritmo			
			SFF	ADP	MEAN	COLOR
Cone	Grosseira (Máximo a posteriori)	5	5,74	7,609	6,419	9,122
		9	4,545	5,032	4,813	10,119
		15	3,463	3,729	3,919	12,327
		21	2,815	2,896	3,45	14,541
		31	2,633	2,471	3,459	18,639
		41	3,124	2,499	3,332	20,951
Cosseno	Grosseira (Máximo a posteriori)	5	6,306	8,472	7,075	10,808
		9	4,746	5,791	5,221	12,929
		15	3,488	4,015	4,046	15,562
		21	2,821	2,984	3,54	18,652
		31	3,152	2,555	3,568	22,925
		41	4,318	2,584	3,936	24,334
Plano	Grosseira (Máximo a posteriori)	5	9,166	13,035	10,565	18,267
		9	5,34	9,237	6,642	24,262
		15	2,832	4,589	3,588	30,267
		21	2,604	4,589	3,176	32,56
		31	2,48	2,524	2,94	31,802
		41	2,506	2,447	2,851	29,455
Esfera	Grosseira (Máximo a posteriori)	5	7,158	11,188	8,323	13,429
		9	4,939	9,113	6,036	22,14
		15	2,505	6,337	4,538	32,516
		21	2,524	3,849	4,047	37,47
		31	2,702	2,685	3,714	41,197
		41	2,968	2,5154	3,822	44,665

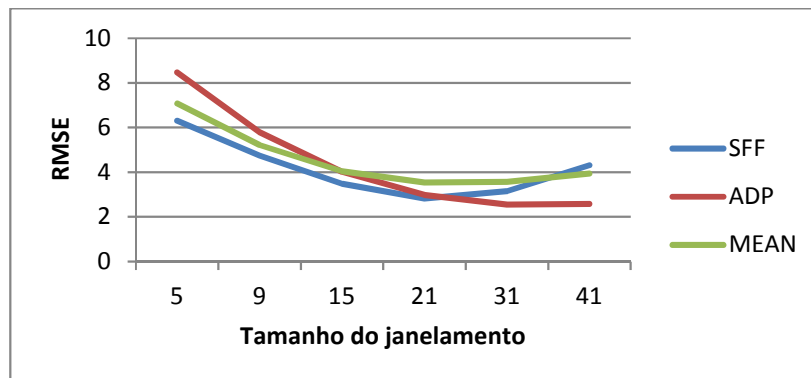
Fonte: Produzida pelo autor.

A análise dos ganhos são apresentadas nos sub-gráficos do Gráfico 9. Foi identificado que o crescimento da janela melhora o desempenho do algoritmo, conforme os gráficos 9.a, 9.b, 9.c e 9.d. Os resultados para o algoritmo COLOR foi omitido tendo em vista seu desempenho ruim e prejudicar a identificação dos valores no gráfico para realizar a análise dos demais algoritmos.

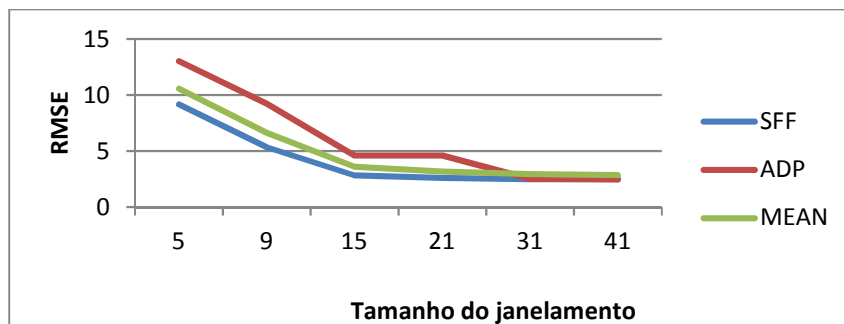
GRAFICO 9 – Comparação do desempenho do tamanho da janela para os medidores
a) Cone



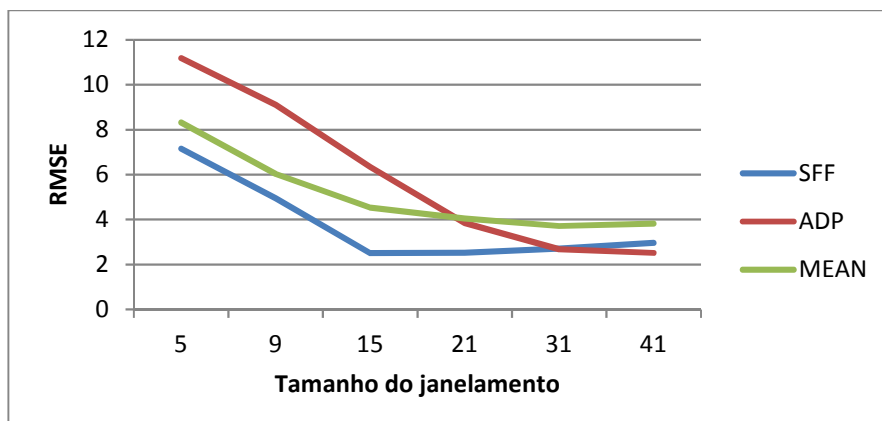
b) cosseno



c) Plano



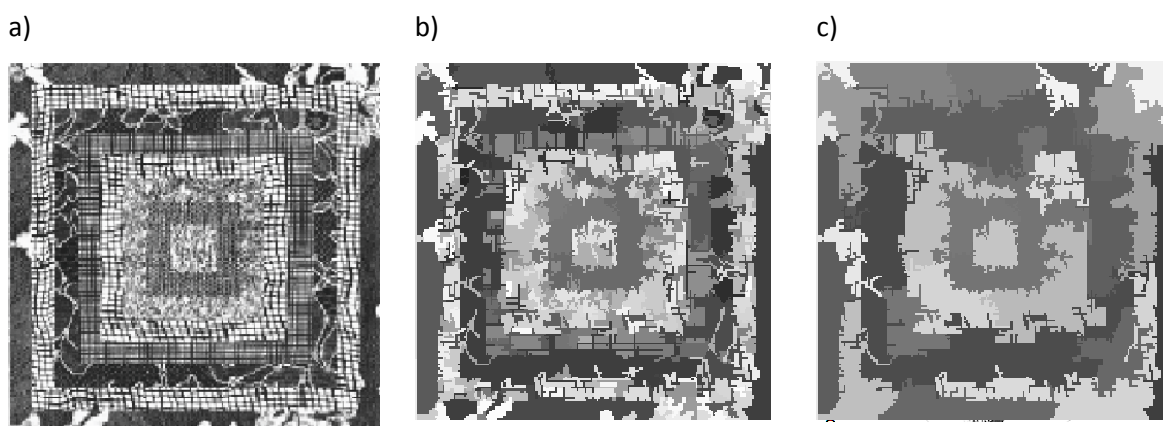
d) Esfera



Fonte: Produzida pelo autor.

Na Figura 27 é apresentada as imagens utilizadas pelo algoritmo de reconstrução do mapa de profundidade utilizando a técnica *Mean-Shift Segmentation*. Inicialmente uma imagem totalmente focalizada é obtida, Figura 27.a, e em seguida é realizada a segmentação *Mean-Shift*, imagens 27.b ou 27.c. O tamanho da área pode ser definida como o tamanho de uma janela, por exemplo, em 27.a é utilizado 70 pixels que correspondem aproximadamente a 70% da área total de uma janela de tamanho 9 linhas por 9 colunas. Em 27.c é utilizado 1681 pixels que correspondem a 100% da área total de uma janela de tamanho 41 linhas por 41 colunas.

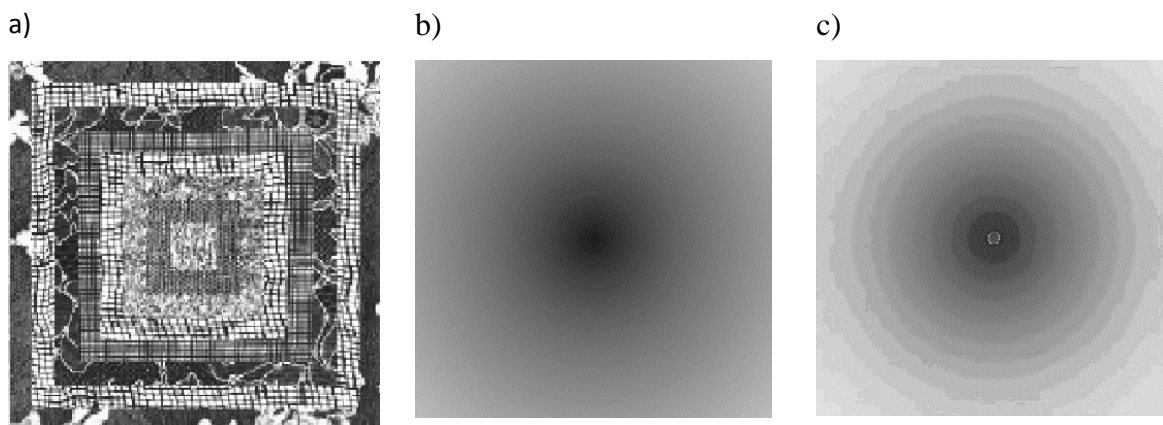
FIGURA 27: Imagem segmentada usando Mean-Shift Segmentation



Fonte: Produzida pelo autor.

A Figura 28 apresenta as figuras que foram utilizadas no processo de reconstrução do mapa de profundidade de um cone utilizando a técnica *Shape From Focus*. A imagem texturizada, Figura 28.a, é utilizada como entrada do algoritmo de simulação do defocus. A imagem 28.b é o mapa de profundidade produzido por esse algoritmo de simulação, e em 28.c, é apresentado o mapa de profundidade que foi reconstruído utilizando a técnica *Shape From Focus*.

FIGURA 28 Mapas de profundidade sintéticos



Fonte: Produzida pelo autor.

Para o grupo de amostras sintéticas número dois, é apresentada a Tabela 9 e o Gráfico 10 computando o processamento sobre imagens realísticas. Diferente das superfícies simuladas, onde existe apenas um objeto, as imagens realísticas são compostas de múltiplos objetos e múltiplas descontinuidades de focalização.

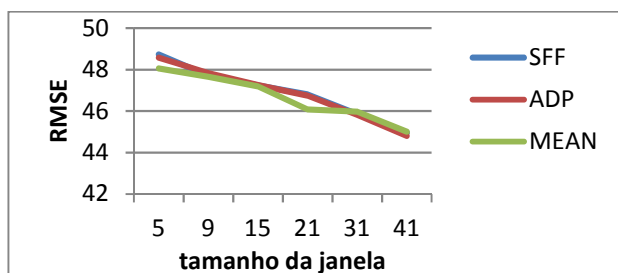
TABELA 10 – Erro RMSE para superfícies realísticas.

Superfície	Tipo Reconstrução	Janela	Algoritmo			
			SFF	ADP	MEAN	COLOR
VENUS	Grosseira (Máximo a posteriori)	5	48,73	48,575	48,056	48,291
		9	47,724	47,833	47,653	48,779
		15	47,231	47,26	47,187	50,239
		21	46,802	46,75	46,084	51,156
		31	45,872	45,825	45,969	52,036
		41	44,965	44,805	45,0067	52,415
ART	Grosseira (Máximo a posteriori)	5	61,299	61,188	65,633	65,189
		9	60,711	60,949	66,45	65,416
		15	60,016	59,617	66,551	65,502
		21	59,557	58,42	66,767	65,432
		31	58,425	57,23	66,542	65,542
		41	57,268	56,747	66,905	65,669
REMINDERR	Grosseira (Máximo a posteriori)	5	58,358	58,422	58,458	58,242
		9	57,885	58,477	58,265	58,309
		15	57,164	57,794	57,785	58,137
		21	56,46	57,176	57,356	57,958
		31	55,417	56,447	57,108	58,109
		41	55,05	55,865	56,855	59,46

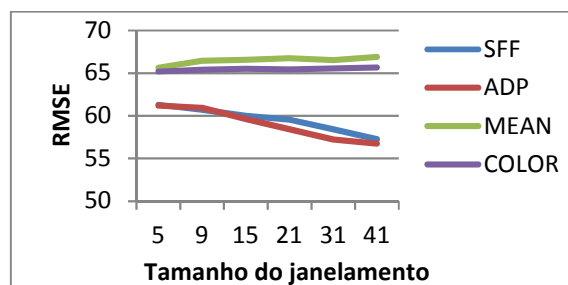
Fonte: Produzida pelo autor.

GRÁFICO 10- Comparação do crescimento da janela para cenas realística

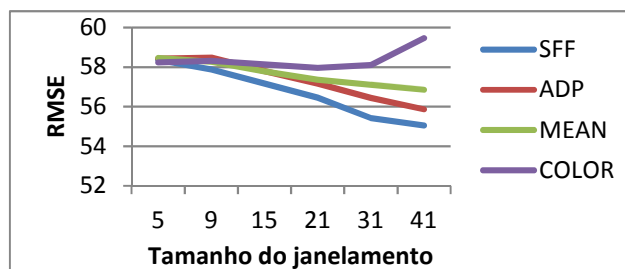
a) Vênus



b) Art



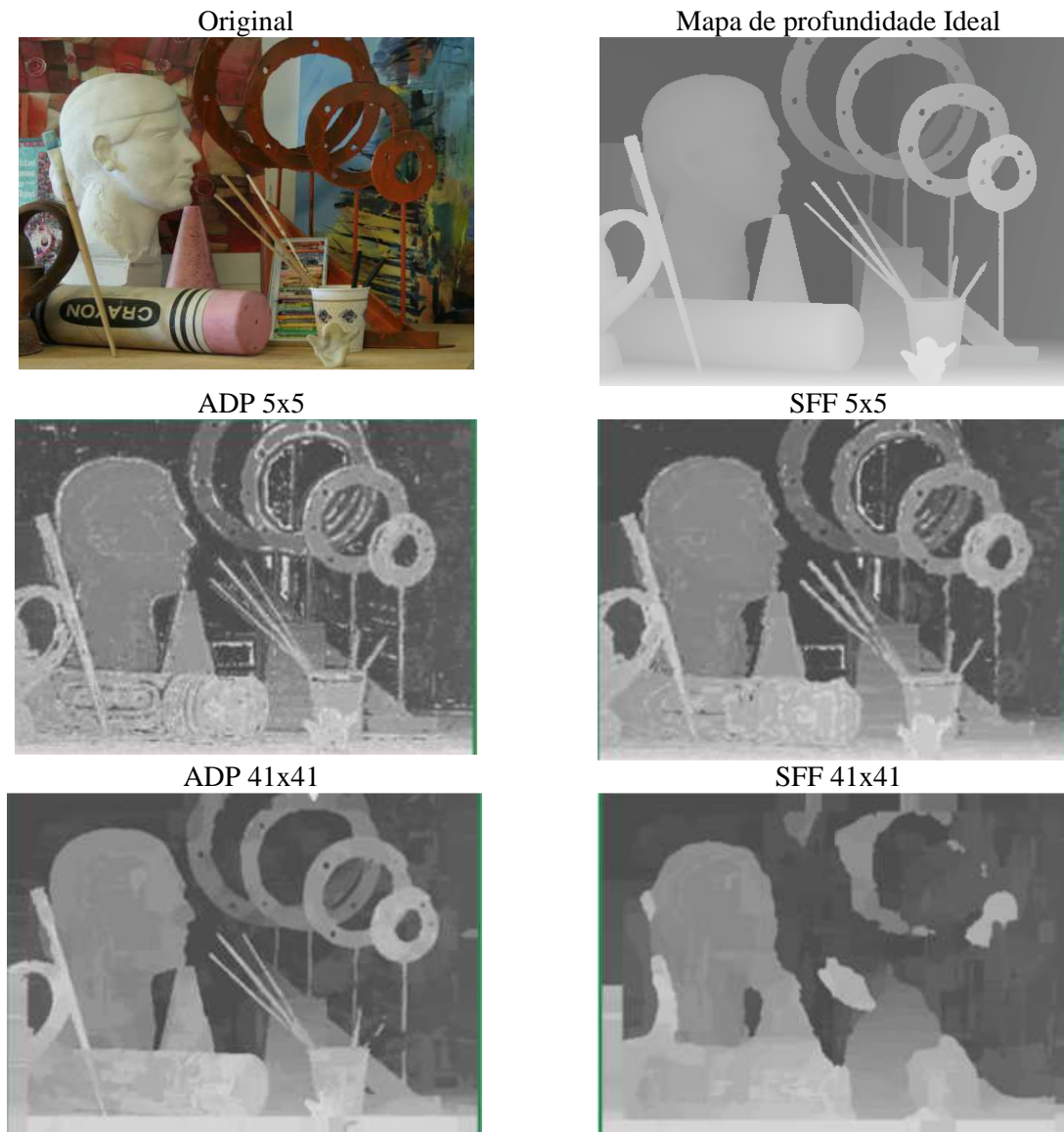
c) Remindeer



Fonte: Produzida pelo autor.

A Figura 29 apresenta o resultado de uma reconstrução de mapa de profundidade com janelamento quadrático de tamanho 41x41 e outro com tamanho 5x5, para os algoritmos SFF e ADP. Na primeira linha é apresentado a imagem original e seu mapa de profundidade verdadeiro(ideal), na segunda linha o resultado comparativo da reconstrução usando o algoritmo SFF e ADP com janelas 5x5 e na terceira linha os mesmos algoritmos usando janelas 41x41. A reconstrução foi utilizada para a base da dados Midlebury ART

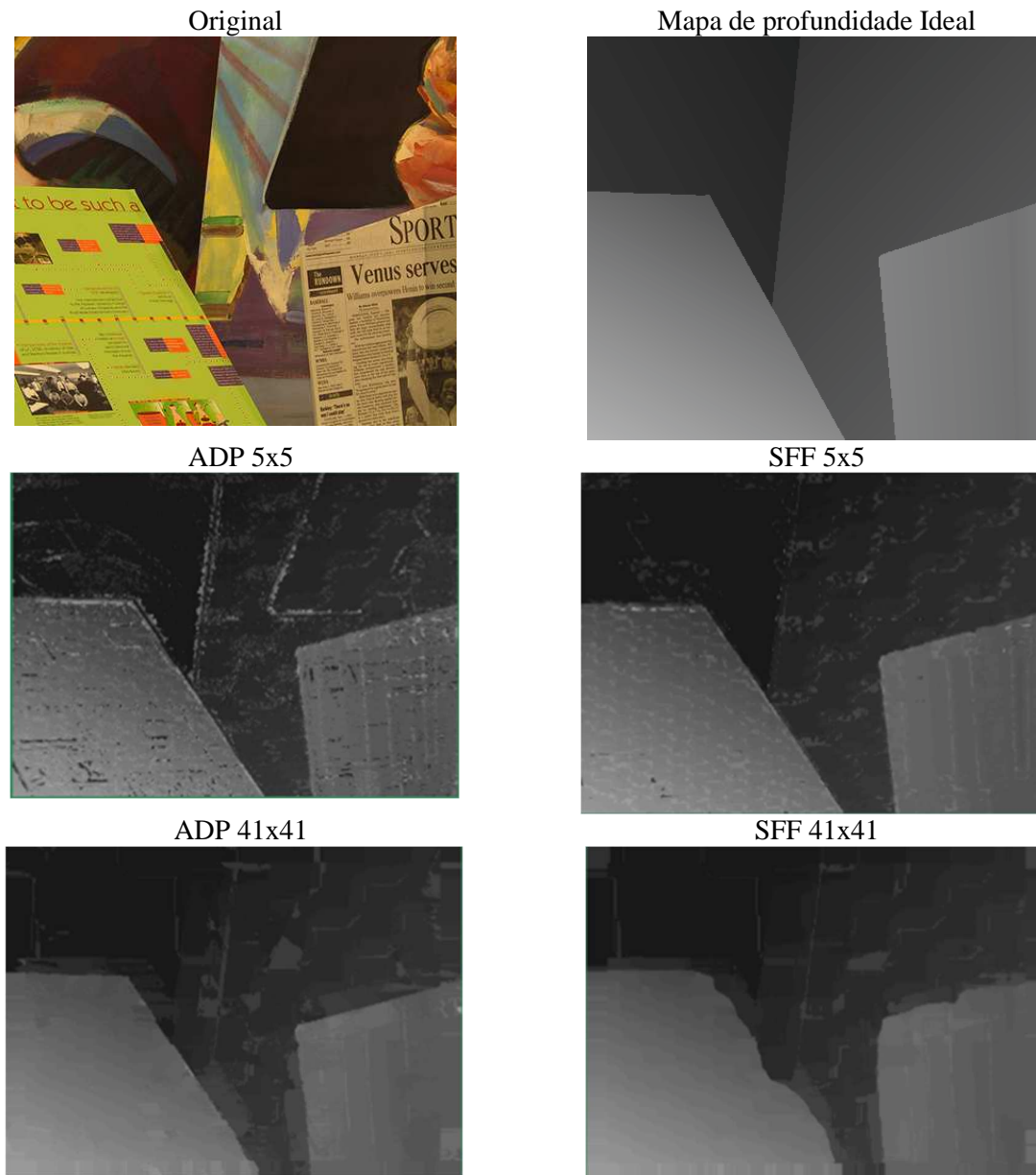
FIGURA 29 – Reconstrução do Mapa de profundidade da imagem ART.



Fonte: Produzida pelo autor

A Figura 30 apresenta o resultado de uma reconstrução de mapa de profundidade com janelamento quadrático tamanho 41x41 e outro com tamanho 5x5, para os algoritmos SFF e ADP. Na primeira linha é apresentado a imagem original e seu mapa de profundidade verdadeiro(ideal), na segunda linha o resultado da reconstrução usando os algoritmos SFF e ADP com janelas 5x5 e na terceira linha o resultado usando janelas 41x41. A reconstrução foi utilizada para a base da dados Midlebury VENUS.

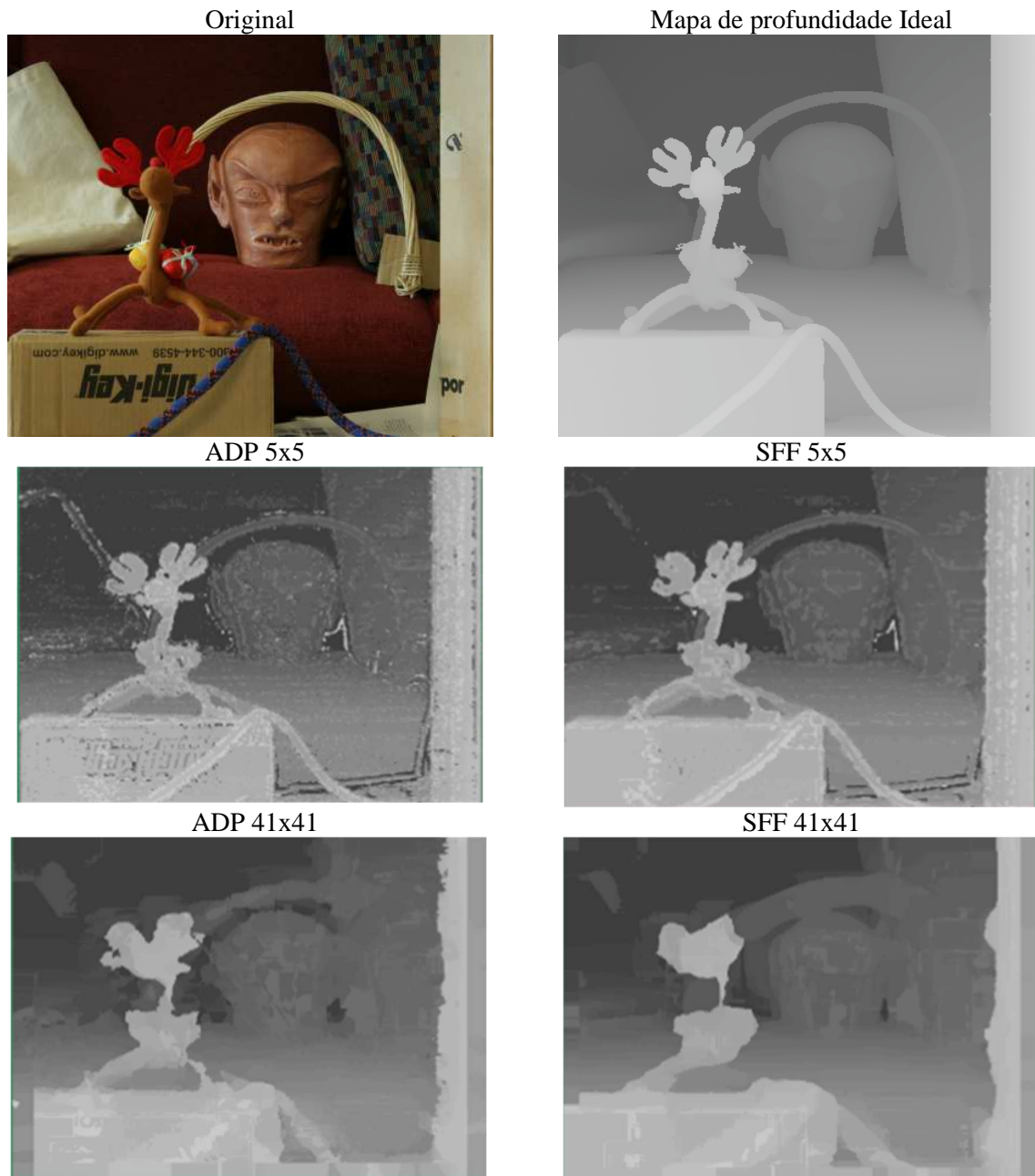
FIGURA 30 – Reconstrução do Mapa de profundidade da imagem VENUS.



Fonte: Produzida pelo autor

A Figura 31 apresenta o resultado de uma reconstrução de mapa de profundidade com janelamento tamanho quadrático 41x41 e outro com tamanho 5x5, para os algoritmos SFF e ADP. Na primeira linha é apresentado a imagem original e seu mapa de profundidade verdadeiro(ideal), na segunda linha o resultado comparativo da reconstrução usando o algoritmo SFF e ADP para janelas 5x5 e na terceira linha o resultado para janelas 41x41. A reconstrução foi utilizada para a base da dados Midlebury REMINDEER.

FIGURA 31 – Reconstrução do Mapa de profundidade da imagem REMINDEER.

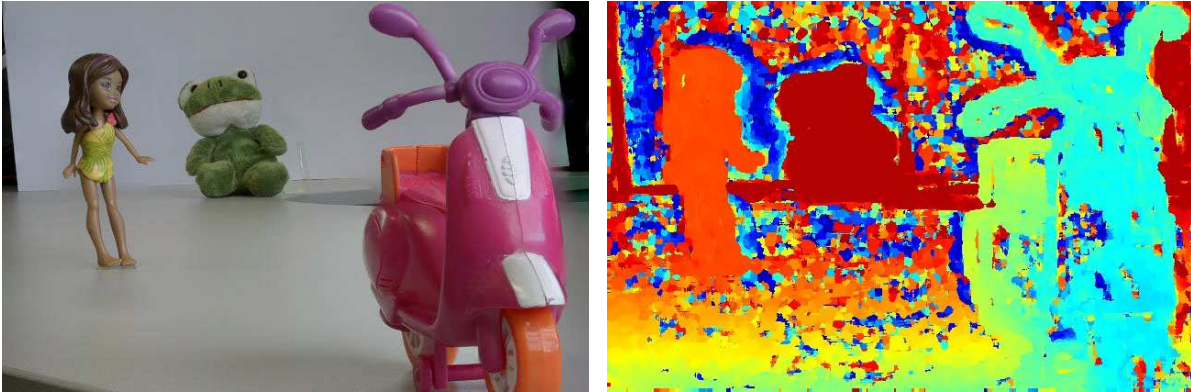


Fonte: Produzida pelo autor

Na Figura 32, é exibido o mapa de profundidade de amostras reais produzidas pelo autor utilizando a técnica *Shape From Focus*, com sua respectiva imagem real. Esse mapa de profundidade foi mapeado segundo a paleta de cor chamada de *colourjet*, a fim de realçar as

tonalidades de profundidade, onde a cor vermelho realça a maior profundidade e a cor azul realça a menor profundidade. Este mapa de profundidade não computou as correções sobre profundidades não confiáveis e por isso elas aparecem em forma de azul escuro e também em forma de chuviscos.

FIGURA 32 Imagem real toda focalizada com seu Mapa de profundidade.



Fonte: produzida pelo autor.

Uma vez que a produção do mapa de imagens reais não possui um mapa de profundidade real que possa ser comparado com o calculado pelo algoritmo, o cálculo do erro RMSE para a pilha de foco de imagens reais não pode ser realizado.

5 CONCLUSÕES

Em pequena profundidade de campo, como na microscopia, a técnica SFF aproveita bem as informações de variações em cada pixel para cada posição focal, porém em grandes profundidades de campo as interferências que se originam de diversos fatores, acabam por inviabilizar apenas um tipo de tratamento da informação de foco. Devido esses fatores, pode-se dizer que a variação em cada pixel é comprometida em grandes profundidades de campo, e tornam inviável compor uma reconstrução fina satisfatória. Em geral a capacidade de reconstrução de detalhes, que é a grande vantagem da técnica SFF sobre as demais de Visão Computacional, acaba se perdendo nessas condições, por outro lado, a pouca captação da rugosidade da superfície acaba atribuindo maior importância tanto à textura (PERTUZ.2013) quanto as cores uniformes (Figuras 27.b e 27.c), e permitindo a captação de detalhes da imagem.

Apesar dos medidores com suporte adaptativo melhorarem visualmente o processo de reconstrução de mapas de profundidade (terceira linha da Figura 29), algumas vezes, os medidores tradicionais parecem superar essas medidores (Gráficos 10.b e 10.c), o que sugere uma aglutinação por aproximação.

Alguns medidores com suporte adaptativo não são tão robustos quanto os tradicionais, mas conforme os trabalhos apresentados, eles tendem a ganhar robustez em grandes janelas de avaliação (Tabela 8 e Gráfico 8). Outro fator que ficou evidenciado é seu desempenho superior em transições de superfícies. O Janelamento muito grande introduziu dois principais, efeitos indesejados em nossos algoritmos: o aumento do tempo de processamento e o maior consumo de memória. Primeiramente, o tempo de processamento pode ser tratado com os modernos processadores multicore, mas o consumo maior de memória pode inviabilizar projetos práticos, principalmente para janelas quadráticas acima de 41 pixels. Portanto, apesar do desempenho inferior nos experimentos, foi identificado a segmentação como critério para tratar este efeito. Embora, aparentemente o algoritmo COLOR tenha apresentado um rendimento muito ruim, pode-se salientar que foi desprezado o processo de cálculo dos parâmetros desse método, o que no desenvolver deste trabalho demonstrou ser de suma importância.

Como trabalhos futuros podemos citar o uso de outros medidores de nitidez (apresentados no Capítulos 2.4 e 2.6) sobre as amostras e a adição de outros algoritmos de suporte adaptativo para *Shape From Focus*. Também, seria importante uma investigação sobre o comportamento dos algoritmos que não tiveram um bom resultado no presente trabalho quando alteramos seus parâmetros.

Por fim o suporte adaptativo não é a solução para todos os problemas de reconstrução de mapas de profundidade usando o método *Shape From Focus*, porém pode ser usado em conjunto com outras técnicas para tornar mais confiável o mapa de profundidade.

BIBLIOGRAFIA

AHMAD, M.B. and CHOI, Tae-Sun. **A heuristic approach for finding best focused shape**. In IEEE Transactions on Circuits and Systems for Video Technology, vol. 15, no. 4, pp. 566-574, April 2005. ISSN : 1051-8215 doi: 10.1109/TCSVT.2005.844450 <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1413274&isnumber=30623>>

AYDIN, Tarkan. and AKGUL, Yusuf Sinan. **A New Adaptive Focus Measure for Shape From Focus**. Proc. British Machine Vision Conference (BMVC) 2008. Disponível em <www.bmva.org/bmvc/2008/papers/13.pdf>

AYDIN, Tarkan. and AKGUL, Yusuf Sinan. **An occlusion insensitive adaptive focus measurement method**. Opt. Express 18, p.14212-14224 (Jun 2010). DOI : 10.1364/OE.18.014212, <<http://www.opticsexpress.org/abstract.cfm?URI=oe-18-13-14212>>

BORN, Max. and WOLF, Emil. **Principles of Optics**. 4 Ed, New York, 1970.

DARRELL, Trevor and WOHN, Kwangyun. **Pyramid based depth from focus**. Computer Vision and Pattern Recognition, 1988. in Proceedings CVPR '88, Computer Society Conference on, Ann Arbor, MI, pp. 504-509, 1988. ISSN : 1063-6919 doi: 10.1109/CVPR.1988.196282 <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=196282&isnumber=5032>>

DAVID, Ferstl; MARTIN, Lenz; MATHIAS, Ruther and HORST, Bischof. **Integral Shape From Focus**. 17th Computer Vision Winter Workshop. in Institute for computer graphic and vision (CVWW). 2012. Disponível em <www.icg.tugraz.at/Members/ferstl/cvww2012>

ENS, John E. **An Investigation of Methods for Determining Depth from Focus**. PhD thesis, University of British Columbia, Canada, July 1990. doi: 10.14288/1.0065642 <<https://open.library.ubc.ca/staticfile/js/build/embed/item.js>>

FAUGERAS , O.D. and TOSCANI,G. **The calibration problem for stereo**. In Proc. Conf. Computer Vision and Pattern Recognition CVPR 1986, pp. 15-20, Miami Beach, 1986.

FILHO, Ogê Marques e NETO, Hugo Vieira. **Processamento Digital de Imagem**. Rio de Janeiro, RJ, 1999. ISBN 8574520098.

FORSYTH, David A. and PONCE, Jean. **Computer Vision: A Modern Approach**. Prentice Hall, 2002. ISBN:0130851981.

FROMMER, Yuval; ARI, Rami Ben and KIRYATI, Nahum. **Shape from Focus with Adaptive Focus Measure and High Order Derivatives**. In Xianghua Xie, Mark W. Jones, and Gary K. L. Tam, editors, Proceedings of the British Machine Vision Conference (BMVC), pages 134.1-134.12. BMVA Press, September 2015. ISBN:1-901725-53-7, DOI:10.5244/C.29.134, <<https://dx.doi.org/10.5244/C.29.134>>

GAGANOV, Victor and IGNATENKO, Alexey. **Robust shape from focus via Markov random fields**. in Proc. GraphiCon, pp. 74–80, 2009. ISBN: 978-5-317-02975-3, disponível em <http://www.graphicon.ru/proceedings/2009/conference/se3/83/83_Paper.pdf>

GONZALEZ, Rafael C and WOODS, Richard E. **Digital Image Processing**. 3rd ed. Prentice-Hall, 2008. Tradução Cristina Yamagami e Leonardo Piamonte, 2010. ISBN : 978-85-7605-401-6.

GROSSMANN, P. **Depth from focus**. Pattern Recognition Letters, vol. 5, pp. 6349, 1987. ISSN 0167-8655, doi:10.1016/0167-8655(87)90026-2 <<http://www.sciencedirect.com/science/article/pii/0167865587900262>>

HASINOFF, Samuel W. and KUTULAKOS, Kiriakos N. **Confocal Stereo**. International Journal of Computer Vision, 81(1), pp. 82-104, 2009. ISSN:1573-1405, doi:10.1007/s11263-008-0164-2 <<http://dx.doi.org/10.1007/s11263-008-0164-2>>

HORN, Berthold Klaus Paul. **Focusing**. Massachusetts Institute Technology, Artificial Intelligence Laboratory. MEMO NO. 160. May 1968.

HORN, Berthold Klaus Paul. **Shape from Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One View**. Ph.D. Thesis, Massachusetts Inst. Of Technology. 1970.

HORN, Berthold Klaus Paul. **Robot Vision**. Cambridge, MA: MIT Press, 1986.

JARVIS , Ray A. **Focus optimisation criteria for computer image processing**.

Microscope, vol. 24, No 2, pp. 163- 180, 1976.

JING, Bing-Zhong and YEUNG, Daniel S. **Recovering Depth From Images Using Adaptive Depth From Focus**. Proceedings 2012 International Conference on Machine Learning and Cybernetics, Xian, 2012, pp. 1205-1211. doi: 10.1109/ICMLC.2012.6359527 <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6359527&isnumber=6359457>>

KIM, Gyung Bum; TIAN, Gui Yun. **A novel depth-from-focus-based measurement system for the reconstruction of surface morphology with depth discontinuity**. 2009. The International Journal of Advanced Manufacturing Technology.2009, pp.1158-1165.ISSN 1433-3015.doi: 10.1007/s00170-008-1438-0 < <http://dx.doi.org/10.1007/s00170-008-1438-0>>

KROTKOV, Eric and MARTIN, Jean-Paul. **Range from focus**. Robotics and Automation. Proceedings. 1986 IEEE International Conference on, 1986. pp. 1093-1098. doi: 10.1109/ROBOT.1986.1087510. < <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1087510&isnumber=23643>>

KROTKOV, Eric. **Focusing**. International Journal of Computer Vision, vol. 1, n. 3, pp. 223-237, 1987. ISSN:1573-1405, doi:10.1007/BF00127822 <<http://dx.doi.org/10.1007/BF00127822>>

KUJOORY, Mohammad Ali; MAYALL, Brian.H. and MENDELSON, Mortimer L. **Focus-assist device for a flying spot microscope**. in IEEE Transactions on Biomedical Engineering, vol. BME-20, pp. 126-132, March 1973. doi: 10.1109/TBME.1973.324174 < <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4120608&isnumber=4120598>>

KUTHIRUMMAL , Sujit; NAGAHARA, Hajime; ZHOU, Changyin and NAYAR, Shree K. **Flexible Depth of Field Photography**. in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 1, pp. 58-71, Jan. 2011. ISSN: 0162-8828 doi: 10.1109/TPAMI.2010.66 <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5432211&isnumber=5639101>>

LEE, Ikhyun; MAHMOOD, Muhammad Tariq and CHOI,Tae-Sun. **Adaptive window selection for 3D shape recovery from image focus**. Optics & Laser Technology, v 45, February 2013, p.21-31. ISSN 0030-3992. DOI:dx.doi.org/10.1016/j.optlastec.2012.08.003. <<http://www.sciencedirect.com/science/article/pii/S0030399212003544>>

LEE, Ik-Hyun; SHIM, Seong-O; CHOI,Tae-Sun. **Improving focus measurement via variable window shape on surface radiance distribution for 3D shape reconstruction**. Optics and Lasers in Engineering. v. 51, Issue 5, May 2013, p.520-526. ISSN 0143-8166. DOI: dx.doi.org/10.1016/j.optlaseng.2012.11.003.< <http://www.sciencedirect.com/science/article/pii/S0143816612003090>>

LEE, S.; AHN , S. C. and MEYYAPPAN, A. **Depth from magnification and blurring**. Robotics and Automation, 1997. Proceedings., 1997 IEEE International Conference on, Albuquerque, NM, 1997, pp. 137-142 vol.1. doi: 10.1109/ROBOT.1997.620028 < <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=620028&isnumber=13464>>

LIU, W and Wu Key, X. **Semi-global depth from focus**. 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), Kuala Lumpur, 2015, pp. 624-629.DOI: 10.1109/ACPR.2015.7486578

MAHMOOD, M. T. ; SHIM, S. and CHOI, T. S. **Depth and image focus enhancement for digital cameras**. 2011 IEEE 15th International Symposium on Consumer Electronics (ISCE), Singapore, 2011, pp. 50-53. DOI: 10.1109/ISCE.2011.5973781 < <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5973781&isnumber=5973167>>

MAHMOOD, Muhammad Tariq. **Shape from focus by total variation**. IVMSP Workshop, 2013 IEEE 11th, Seoul, 2013, pp. 1-4. doi: 10.1109/IVMSPW.2013.6611940 <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6611940&isnumber=6611884>>

MAHMOOD, Muhammad Tariq and CHOI, Tae-Sun. **Nonlinear Approach for Enhancement of Image Focus Volume in Shape From Focus**. in IEEE Transactions on Image Processing, vol. 21, no. 5, pp. 2866-2873, May 2012. ISSN :1057-7149. doi: 10.1109/TIP.2012.2186144 <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6140969&isnumber=6183042>>

MALIK, Aamir Saeed and CHOI, Tae-Sun. **Consideration of illumination effects and optimization of window size for accurate calculation of depth map for 3D shape recovery**, Pattern Recognition, Volume 40, Issue 1, January 2007, Pages 154-170, ISSN 0031-3203, <http://dx.doi.org/10.1016/j.patcog.2006.05.032>. <<http://www.sciencedirect.com/science/article/pii/S0031320306002470>>

MICHI, Hiroyuki; IKEOKA, Hiroshi and TAKAYUKI Hamamoto. **Distance estimation using two different-aperture images obtained by tilted lens optics camera**. Intelligent Signal Processing and Communications Systems (ISPACS), 2012 International Symposium on, New Taipei, 2012, pp. 749-754. ISBN: 978-1-4673-5083-9 doi: 10.1109/ISPACS.2012.6473591 <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6473591&isnumber=6473441>>

MOELLER, Michael; BENNING , Martin; SCHÖNLIEB , Carola and CREMERS, Daniel. **Variational Depth From Focus Reconstruction**. in IEEE Transactions on Image Processing, vol. 24, no. 12, pp. 5369-5378, Dec. 2015. ISSN : 1057-7149 doi:10.1109/TIP. 2015.2479469 <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7271087&isnumber=7271151>>

NAIR, H. N. and STEWART, Charles V. **Robust focus ranging**. Computer Vision and Pattern Recognition, 1992. Proceedings CVPR '92., 1992 IEEE Computer Society Conference on, Champaign, IL, 1992, pp. 309-314. 1992. ISSN : 1063-6919 doi: 10.1109/CVPR.1992.223258 <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=223258&isnumber=5817>>.

NATHANIEL, Ng Kuang Chern; NEOW, Poo Aun and ANG, Marcelo H. **Practical issues in pixel-based autofocus for machine vision.** Robotics and Automation. 2001.

Proceedings 2001 ICRA. IEEE International Conference on, 2001, pp. 2791-2796 vol.3.DOI: 10.1109/ROBOT. 2001.933045 < <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=933045> & isnumber =20184>

NAYAR, Shree K. and NAKAGAWA, Yasuo. **Shape from focus.** in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, no. 8, pp. 824-831, Aug 1994. ISSN : 0162-8828 doi: 10.1109/34.308479 <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=308479&isnumber=7509>>

NIEDEROST, Markus; NIEDEROST, Jana and SCUCKA Jiri. **Automatic 3D Reconstruction and Visualization of Microscopic Objects from a Monoscopic Multifocus Image Sequence.** International Workshop on Visualization and Animation of Reality based 3D Models, Vol. XXXIV-5/W10, 2003. Disponivel em <www.isprs.org/proceedings/XXXIV/5-W10/papers/niederost.pdf>

OPENCV. Camera calibration With OpenCV. **Site official do OpenCv.** Disponível em <http://docs.opencv.org/2.4/doc/tutorials/calib3d/camera_calibration/camera_calibration.html> acessado em 10Jun2016.

OPRINCA, Andrei. **Photoshop Depth Maps – Realistic Depth of Field.** In Effects, Tutorials, PSDBOX. April 19, 2011. Disponível em <<http://www.psdbox.com/tutorials/photoshop-depth-maps-realistic-depth-of-field>> acessado em 10Jun2016.

PENTLAND, Alex Paul 1987. **A new sense for depth of field.** in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-9, no. 4, pp. 523-531, July 1987. ISSN: 0162-8828 doi: 10.1109/TPAMI.1987.4767940 <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4767940&isnumber=4767931>>

PERTUZ, S. **Modeling and applications of the focus cue in conventional Digital cameras.** Phd Thesis. Universitat Rovira I Virgili. 2013. Departament d'Enginyeria Informàtica i Matemàtiques. Tarragona.2013.

PERTUZ,S. , PUIG, D. and GARCIA,M. A. **Analysis of focus measure operators for shape-from-focus**. Pattern Recognit., vol. 46, no. 5, pp. 1415–1432, 2013. ISSN:0031-3203, doi: 10.1016/j.patcog.2012.11.011 <<http://dx.doi.org/10.1016/j.patcog.2012.11.011> >

RIBNICK, Evan. **Non-orthogonal depth from focus for on-the-fly, three-dimensional inspection**. Process SPIE (Journal Paper). p. 89–99, Agu. 2012. doi:10.1117/1.OE.51.8.083608 <<http://opticalengineering.spiedigitallibrary.org/article.aspx?articleid=1352964>>

SCHECHNER, Yoav Y. and KIRYATI, Nahum. **Depth from Defocus vs. Stereo: How Different Really Are They?** International Journal of Computer Vision 39(2), 141–162, 2000. ISSN:1573-1405, doi:10.1023/A:1008175127327, <<http://dx.doi.org/10.1023/A:1008175127327> >

SRINIVAS, CR; SENTHILNATHAN, R ; SUBHASREE, P ; SIVARAMAKRISHNAN, R e SRIVIDYA, R. **Evaluation of Statistical Focus Measures in a Parallax Affected SFF-Inspired Approach**. IJCTA, 20 (10). pp. 1-11. 2015. <<http://eprints.manipal.edu/id/eprint/145094>>

SUBBARÃO , Murali and CHOI , Tae. **Accurate recovery of three-dimensional shape from image focus**. in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 17, no. 3, pp. 266-274, Mar 1995. doi: 10.1109/34.368191 <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=368191&isnumber=8429>>

SUBBARÃO, Murali and SURYA, Gopal. **Depth from defocus: A spatial domain approach**. International Journal of Computer Vision. 1994, v.13, n.3, p.271-294, ISSN:1573-1405, DOI:10.1007/BF02028349, <<http://dx.doi.org/10.1007/BF02028349>>

TENENBAUM , Jay Marty. **Accommodation in computer vision**. Ph.D. Thesis, Stanford Univ., Computer Science Department. MEMO-AIM 134. Stanford, CA, USA. October 1970.

TSAI, Roger Y. **A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the shelf TV cameras and lenses.** IEEE Int. J. Robot. Automat. RA-3 , 323-344.5, USA, 1987. ISBN :0-86720-294-7, disponível em <<http://dl.acm.org/citation.cfm?id=136913.136938>>

TSENG , Chen-Yu. and WANG, Sheng-Jyh. **Shape-From-Focus Depth Reconstruction With a Spatial Consistency Model.** in IEEE Transactions on Circuits and Systems for Video Technology, vol. 24, no. 12, pp. 2063-2076, Dec. 2014. ISSN : 1051-8215 doi:10.1109/TCSVT.2014.2358873 < <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6902761&isnumber=6971052>>

UCI. University of California's First. Donald Bren Schol of Information and Computer Sciences. CS 211A: Visual Computing. **Apostila** www.ics.uci.edu/~majumder/VC/classes/cameracalib.pdf. Acessado em 24Jun2016.

UFMG. Escola de belas artes. Fotografia Pinhole. **Manual Prático de Fotografia Pinhole.** Set 1999. Disponível em <http://eba.ufmg.br/cfalieri/pinhole.html>, acessado em 20 Jul 2016.

UFMG. Escola de belas artes. **Manual Prático de Fotografia Pinhole.** Set 1999. Disponível em <http://eba.ufmg.br/cfalieri/frame.html>, acessado em 22 Jul 2016.

WENG, J.; COHEN, P. and HERNIOU, M. **Camera calibration with distortion models and accuracy evaluation.** IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI 14, 965-980, 1992. ISSN :0162-8828 DOI:10.1109/34.159901 <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=159901&isnumber=4188>>

ZHANG, X.; LIU, Z.; JIANG, M. and CHANG, M. **Fast and accurate auto-focusing algorithm based on the combination of depth from focus and improved depth from defocus.** Opt. Express 22, 31237-31247 (2014).