

Edição tridimensional de fotografias arquitetônicas

Beatriz Silva Villa Alvarez

Orientador: Paulo Cezar Pinto Carvalho

Co-orientador: Marcelo Gattass

Para meu amor Marcos,

que me ensina a ver o mundo de diferentes ângulos,

nossa filha Sofia que está sendo esperada com muito amor,

meus queridos e maravilhosos pais, Fernando e Elisabeth,

que me motivam a estudar sempre,

meus avós Amparo e Florindo, Isaura e Edgar

que começaram minha maravilhosa família

e toda família igualmente amada e querida

... queria mostrar pra vocês que valeu a pena!

Agradecimentos

Aos mestres Paulo Cezar Carvalho e Marcelo Gattas, pela fundamental orientação, compreensão e paciência nos meus momentos de escuridão.

Ao Caco, pela cumplicidade e presteza em me ajudar, já que inicialmente fomos os únicos mestrandos em Computação Gráfica, no ano de 1999.

A Mica e Cuca, pela inestimável amizade, alegria e força nos momentos de desânimo, além das divertidas viradas de noite terminando algum trabalho de PAA.

A Ana Elisa, por sua dedicação ao doutorado que me serviu de exemplo e inspiração, além da grande amizade que surgiu entre nós pelas dificuldades pessoais que compartilhamos.

Aos amigos do Tecgraf, em especial ao Marcio e Rodrigo, que inicialmente muito me ajudaram no meu retorno a programação em C. Também ao Scuri, pelo apoio técnico e a Claudia, pelo estímulo e apoio durante os trabalhos do Tecgraf.

Aos amigos do mestrado, Thyago, Marcel, Beto, Antonio, Paula, Carla e a quase doutora Paty, pelo bom astral das festas e alegria dos lanches no Mr Coffee.

Aos amigos do Visgraf, por onde fiquei nos últimos meses de mestrado, principalmente ao Ari, pelo equilíbrio, calma e generosidade naquelas horas em que eu queria "arrancar os cabelos".

Ao meu pai e minha mãe por não me permitirem fraquejar.

A Tiça, Flavia, Claudia, tia Mimá, tia Teté e Odette, que nunca duvidaram da minha capacidade e sempre me fizeram lembrar dela.

Ao Armando, meu querido sobrinho, por suas visitas à PUC que renovavam as energias das minhas tardes de trabalho.

Ao tio Mayrton pelas brincadeiras no icq.

A fé em Deus, que me fortaleceu nos momentos muito difíceis, especialmente na perda de meu cunhado Claudio.

E finalmente ao Marcos, pelo amor que sinto em meu coração!

A todos os meus queridos amigos, saibam que vocês contribuíram para que eu seguisse um caminho feliz e me tornasse a pessoa que sou hoje.

Esse trabalho não teria existido sem a ajuda e o amor de vocês.

... A você leitor, cuja curiosidade e interesse justificam tantas horas de trabalho!

Resumo

Este trabalho propõe um sistema interativo que possibilita alterar a fotografia de uma cena tridimensional envolvendo elementos arquitetônicos, de modo que o usuário possa avaliar os efeitos estéticos e o impacto que as modificações no ambiente real poderão causar em outras pessoas. O método utilizado se baseia na existência, em fotos arquitetônicas, de três direções principais de interesse, mutuamente ortogonais. A identificação de pontos de fuga destas direções permite calibrar a câmera utilizada nas fotos e inserir novos elementos na cena.

Abstract

This work proposes an interactive system that allows modifying a 3D photograph depicting an architectural scene so that the user may evaluate the aesthetic effects and the impact the changes may cause. The proposed method is based on the existence of three main mutually orthogonal directions of interest in architectural photos. The identification of vanishing points for those directions allows the calibration of the camera used to take the pictures and the insertion of new elements to the scene.

Sumário

Lista de Figuras	iv
Lista de Notações	viii
1 Introdução	1
1.1 Histórico da fotografia	1
1.2 Motivação	2
1.3 Descrição do problema	4
1.4 Trabalhos relacionados	5
1.5 Organização do trabalho	11
2 Calibração de câmera	12
2.1 Modelo de câmera	13
2.1.1 Projeção perspectiva	13
2.2 Método de calibração utilizado	14
2.3 Obtenção da posição da câmera	15

2.4	Equações de projeção	20
3	Edição tridimensional de fotografias	23
3.1	Edição em um plano principal	23
3.2	Correção da perspectiva	25
4	Aplicação e resultados	29
4.1	Aplicação	29
4.1.1	Tecnologia utilizada na implementação	30
4.1.2	Interface	30
4.1.3	Utilização dos recursos	33
4.2	Resultados	39
4.2.1	Descrição dos testes	39
4.2.2	Dados dos testes	40
4.2.3	Cálculo do centro ótico	40
4.2.4	Comparação entre duas lentes fotográficas	44
4.2.5	Testes com inserção de <i>outdoors</i> e logomarcas	51
4.2.6	Testes com correção de distorção perspectiva	55
4.2.7	Testes com recuperação de textura	58
5	Conclusões e trabalhos futuros	61

Lista de Figuras

1.1	(a,b) Fotografias originais, (c) <i>outdoor</i> aplicado na lateral da casa	5
1.2	Função plenóptica	7
2.1	Modelo <i>pin-hole</i> de câmera	13
2.2	(a) Projeção paralela, (b) projeção perspectiva	14
2.3	Interseção entre três segmentos de retas	17
2.4	Localização do centro ótico	17
2.5	(a) Triângulo HF_yF_z , (b) triângulo $F_xF_yF_z$	19
2.6	(a) Imagem toda, (b) imagem parcial	20
2.7	Translação da origem	22
2.8	Sistemas de coordenadas	22
3.1	Etapas do processo de edição de um plano principal	25
3.2	Processo de cálculo da nova câmera	27
3.3	Imagens originais com inclinação vertical e reprojeta- das com remoção de perspectiva	28

4.1	Barra de menus	31
4.2	(a) Menu <i>File</i> , (b) menu <i>Edit</i> , (c) menu <i>Camera</i>	31
4.3	(a) Menu <i>Plane</i> , (b) menu <i>Distortion</i> , (c) menu <i>File (viewport)</i>	33
4.4	Barra de ferramentas	34
4.5	(a) Inicialização das linhas guias, (b) linhas guias posicionadas em retas que projetem as três direções principais	35
4.6	(a) Opção do menu <i>Camera</i> para calibração, (b) centro ótico C desenhado no <i>canvas</i>	35
4.7	(a) Escolha de pontos para a determinação de uma janela de edição, (b) janela projetada	36
4.8	(a) Abrindo um <i>viewport</i> , (b) imagem aberta no novo <i>viewport</i>	37
4.9	(a) <i>Outdoor</i> aberto na <i>viewport</i> , (b) <i>outdoor</i> inserido na fotografia original	37
4.10	(a) Solicitando <i>Warp Image</i> , (b) <i>viewport</i> com a imagem transformada . .	38
4.11	(a) <i>Frustum</i> de visão da lente 24 <i>mm</i> , (b) <i>frustum</i> de visão da lente 60 <i>mm</i>	40
4.12	(a) Centro ótico do Photo 3D, (b) centro ótico do sistema implementado .	41
4.13	(a) Centro ótico do Photo 3D, (b) centro ótico do sistema implementado .	42
4.14	(a) Centro ótico do Photo 3D, (b) centro ótico do sistema implementado .	42
4.15	(a) Centro ótico do Photo 3D, (b) centro ótico do sistema implementado .	42
4.16	(a) Centro ótico do Photo 3D, (b) centro ótico do sistema implementado .	43
4.17	(1) Humaitá 1, (2) Humaitá 2	46

4.18	(3) Humaitá 3, (4) Humaitá 4	47
4.19	(5) NYCity 1, (6) NYCity 2	47
4.20	(7) Barra Tower 1, (8) Barra Tower 2, (9) Barra Tower 3	48
4.21	(10) Rio Design Barra 1, (11) Rio Design Barra 2, (12) Rio Design Barra 3	48
4.22	(13) Rio Design Barra 4, (14) Rio Design Barra 5	49
4.23	(15) Santa Teresa 1, (16) Santa Teresa 2	49
4.24	(17) Santa Teresa 3, (18) Santa Teresa 4, (19) Santa Teresa 5	50
4.25	(20) Santa Teresa 6, (21) Santa Teresa 7	50
4.26	(a) imagem original do <i>outdoor</i> , (b) Barra Tower 1 com <i>outdoor</i> inserido, (c) Barra Tower 2 com <i>outdoor</i> inserido	52
4.27	(a) imagem original do <i>outdoor</i> , (b) Humaitá 3 com <i>outdoor</i> inserido, (c) Humaitá 4 com <i>outdoor</i> inserido	52
4.28	(a) imagem original do <i>outdoor</i> , (b) Santa Teresa 6 com <i>outdoor</i> inserido, (c) Santa Teresa 4 com <i>outdoor</i> inserido	53
4.29	(a) imagem original do <i>outdoor</i> , (b) Humaitá 1 com <i>outdoor</i> inserido sem retoque, (c) Humaitá 1 com <i>outdoor</i> inserido com retoque	53
4.30	(a) imagem original do <i>outdoor</i> , (b) Rio Design Barra 5 com <i>outdoor</i> in- serido sem retoque, (c) Rio Design Barra 5 com <i>outdoor</i> inserido com retoque	54
4.31	(a) 3 imagens originais de <i>outdoor</i> , (b) NYCity 2 com <i>outdoors</i> inseridos sem retoque, (c) NYCity 2 com <i>outdoor</i> inserido com retoque	54
4.32	(a) imagem original do <i>outdoor</i> , (b) Santa Teresa com <i>outdoor</i> inserido	55

4.33	(a) Santa Teresa 1 original, (b) Santa Teresa 1 com remoção de perspectiva	56
4.34	(a) Santa Teresa 8, (b) Santa Teresa 8 com remoção de perspectiva	56
4.35	(a) Rio Design Barra 2 original, (b) Rio Design Barra 2 com remoção de perspectiva	57
4.36	(a) Humaitá 3 original, (b) Humaitá 3 com remoção de perspectiva	57
4.37	(a) Santa Teresa 6 original, (b) imagem da textura corrigida	58
4.38	(a) fotografia original obtida no site [P3D99], (b) imagem da textura corrigida	59
4.39	(a) Humaitá 2 original, (b) <i>outdoor</i> real retirado da imagem Humaitá 2 . . .	59
4.40	(a) Humaitá 4 original, (b) <i>outdoor</i> real retirado da imagem Humaitá 4 . . .	60
4.41	(a) Santa Teresa 8 original, (b) proporções da fachada	60
5.1	(a) Imagem original, (b) imagem transformada, considerando que todos os elementos estão no mesmo plano	62

Lista de Notações

F	ponto de fuga (u_f, v_f) genérico
F_x	ponto de fuga (u_x, v_x) da direção X
F_y	ponto de fuga (u_y, v_y) da direção Y
F_z	ponto de fuga (u_z, v_z) da direção Z
X	direção principal X
Y	direção principal Y
Z	direção principal Z
S_m	sistema de coordenadas do mundo formado pelos eixos
	X, Y e Z
S_i	sistema de coordenadas da imagem formado pelos eixos
	u, v e w
S_c	sistema de coordenadas da câmera formado pelos eixos
	U, V e W
C	centro ótico da câmera (u_c, v_c, w_c) , em coordenadas de
	pixel no sistema S_i
C_p	centro projetivo (u_p, v_p) , em coordenadas de pixel no sistema S_i
$\overrightarrow{CF_x}$	vetor que une o centro ótico C ao ponto de fuga F_x
$\overrightarrow{CF_y}$	vetor que une o centro ótico C ao ponto de fuga F_y
$\overrightarrow{CF_z}$	vetor que une o centro ótico C ao ponto de fuga F_z
$ \overrightarrow{CF_x} $	vetor unitário (X_u, X_v, X_w)

$ \overrightarrow{CF_y} $	vetor unitário (Y_u, Y_v, Y_w)
$ \overrightarrow{CF_z} $	vetor unitário (Z_u, Z_v, Z_w)
H	ortocentro do triângulo
L_1	lado do triângulo que une os pontos de fuga F_x e F_y
L_2	lado do triângulo que une os pontos de fuga F_y e F_z
L_3	lado do triângulo que une os pontos de fuga F_x e F_z
p_1	ponto extremo do quadrilátero $p_1q_1p_2q_2$ a ser editado
q_1	ponto extremo do quadrilátero $p_1q_1p_2q_2$ a ser editado
p_2	ponto extremo do quadrilátero $p_1q_1p_2q_2$ a ser editado
q_2	ponto extremo do quadrilátero $p_1q_1p_2q_2$ a ser editado
P_1	ponto extremo do retângulo $P_1Q_1P_2Q_2$ no espaço
Q_1	ponto extremo do retângulo $P_1Q_1P_2Q_2$ no espaço
P_2	ponto extremo do retângulo $P_1Q_1P_2Q_2$ no espaço
Q_2	ponto extremo do retângulo $P_1Q_1P_2Q_2$ no espaço
T'	transformação projetiva bidimensional
α, β e γ	eixos de coordenadas da nova câmera relacionada a aplicação de correção de perspectiva

Capítulo 1

Introdução

“Desde sua invenção, há cerca de uns 160 anos, a fotografia demonstrou ser um agente de conformação da realidade num processo de montagem e seleção, no qual o mundo se revela semelhante e diferente ao mesmo tempo. Se Nadar, Carjat, LeGray, Adamson e Julia Cameron são exemplos do fotógrafo como artista, atento à captação da interioridade do modelo, muitas vezes próximo de resultados pictóricos, Disdéri representa, ao contrário, o protótipo do fotógrafo industrial, disposto a usar todos os truques a seu alcance para adular e seduzir a clientela.” [Fab91]

1.1 Histórico da fotografia

[Fab91] determina três momentos fundamentais para o aperfeiçoamento dos processos fotográficos – primeiras experiências, colódio úmido, gelatina-bromuro – que levarão, em 1895, à invenção da primeira câmera portátil, carregável e descarregável em plena luz. [Fab91] descreve ainda três etapas da complexa relação entre a fotografia e a sociedade do século XIX: a primeira etapa estende-se de 1839 aos anos 50, quando o interesse se restringe a um pequeno número de abastados, que podem pagar os altos preços cobrados

pelos artistas fotógrafos (Nadar, Carjat, Le Gray, por exemplo). O segundo momento correspondente à descoberta do cartão de visita fotográfico (*carte-de-visite photographique*) por Disdéri, que coloca ao alcance de muitos o que até aquele momento fora privilégio de poucos e confere à fotografia uma verdadeira dimensão industrial, pelo barateamento do produto (1854). Por volta de 1880, tem início a terceira etapa: é o momento da massificação, quando a fotografia se torna um fenômeno prevalentemente comercial, sem deixar de lado sua pretensão de ser considerada arte.

1.2 Motivação

Imagem de muitos significados que se presta, por sua própria natureza, a enfoques diferenciados, a fotografia, assim como as ilustrações, passou a ser usada também em publicidade, arquitetura e artes gráficas em geral. Ela proporciona uma representação precisa e fiel da realidade, retirando da imagem a hipotética da subjetividade. A imagem, além de ser nítida e detalhada, forma-se rapidamente, e o procedimento é simples e acessível a todos, permitindo uma ampla difusão.

Com o advento da informática, as fotografias ganharam o formato digital, e sua utilização contribui para agilizar a dinâmica das criações dos fotógrafos, publicitários, arquitetos e artistas gráficos, resultando num trabalho de arte magnífico e criativo. Ou seja, a tecnologia digital ajudou a ampliar os limites da criatividade, pois as pessoas se sentem mais livres para arriscar suas idéias testando diferentes *caminhos*, uma vez que os resultados podem ser vistos de imediato e não implicam em consumo de material fotográfico.

Com tantas facilidades que a era digital proporcionou a esses profissionais, agora a utilização da fotografia está ainda mais ampla. Contudo, é natural que surjam outras necessidades que estimulem a descoberta de técnicas para a criação de novas ferramentas. Existem no mercado *software* de diversos níveis e recursos, desde os mais simples e amigáveis até os altamente profissionais, que exigem um bom treinamento ou muitas ho-

ras de estudo para sua plena utilização. Estes *software* possibilitam imitar, complementar ou modificar a realidade.

Atualmente, os sistemas de edição de imagens podem ser desenvolvidos em microcomputadores, pois já superou-se o problema de armazenamento dos dados e de eficiência no seu processamento, tornando tais sistemas acessíveis a um maior número de usuários. Nestes sistemas, utiliza-se o processamento de imagens para acrescentar informações que serão sobrepostas à imagem real, gerando melhorias espectrais ou espaciais na mesma. Um exemplo é a manipulação de imagens que envolve mudanças em suas características, desde uma simples alteração da cor de fundo de uma foto ou a remoção de algum indivíduo ou elemento indesejado até operações altamente complexas envolvendo mudanças radicais sobre a tomada original, ajudando a transformá-la em um trabalho artístico. Na verdade, a maior parte das imagens digitais envolve manipulação prévia: no mínimo, é aplicada uma equalização das cores em relação à foto original, devido ao processo de digitalização. Outro exemplo é a composição de imagens que consiste em montar em uma única composição diversas imagens, gráficos, textos e outros elementos. Estas técnicas podem ser consideradas como o ápice da edição de imagens digitais, considerando-se a complexidade dos trabalhos e técnicas envolvidos.

Mesmo com todas as técnicas sofisticadas de edição de imagens já embutidas em produtos profissionais no mercado, mais cedo ou mais tarde surgiria a necessidade de ingressar a fotografia no mundo tridimensional. Muitos dos trabalhos de artistas gráficos consistem na criação de logomarcas, anúncios, placas de sinalização, etc. Estas criações precisam ser inseridas dentro de um contexto real. Assim como os arquitetos, os criadores de logomarcas e embalagens precisam apresentar seus trabalhos com a perspectiva que se assemelhe ao produto final. Este tipo de apresentação ajuda os clientes a entender como os *designs* serão na vida real. Muitos clientes sentem dificuldade em visualizar como um anúncio plano ficará inserido num contexto tridimensional e não conseguem avaliar o impacto que pode causar nas pessoas, correndo o risco, por isso, de rejeitar uma campanha publicitária de sucesso. Antes que seja feito um investimento alto, os criadores

também precisam ter certeza de que suas simulações fotográficas (composições de imagens) pareçam o mais reais possível para que possam avaliar questões importantes da área de *design* como a legibilidade e a visibilidade do material exposto. Outras necessidades comuns aos arquitetos consistem em extrair texturas para serem aplicadas em algum projeto modelado tridimensionalmente e restaurar fotografias de fachadas antigas. Para os artistas gráficos é importante haver inovações no que diz respeito a possíveis novos efeitos que possam ser aplicados às imagens para que elas se tornem diferentes da original. É necessário, portanto, desenvolver ferramentas computacionais capazes de transcender a limitação bidimensional dos sistemas de edição de imagens existentes, inerente à superfície plana de uma fotografia.

1.3 Descrição do problema

Nesta dissertação apresentaremos um sistema simples e interativo, capaz de extrair informações tridimensionais de uma fotografia envolvendo elementos arquitetônicos, permitindo alterá-la – por exemplo, inserindo novos objetos em determinados planos da cena ou corrigindo-a de forma a eliminar deformações. O objetivo é disponibilizar uma ferramenta de composição de imagens que consiga extrair alguma representação geométrica da cena tridimensional, de modo que a sobreposição em planos da fotografia original e a extração de textura com perspectiva corrigida seja feita automaticamente, e outra ferramenta que simule uma lente fotográfica e com isso produza um novo efeito visual.

O sistema visa atender a profissionais das áreas de arquitetura, fotografia, publicidade e artes visuais, quando suas necessidades envolvem a alteração de cenários reais. Por exemplo, arquitetos podem inserir novos objetos tridimensionais em fachadas de prédios e casas para avaliar o resultado de uma reforma ou restauração antes de despendere tempo e dinheiro com a execução real, enquanto publicitários podem ter uma visualização prévia de suas campanhas em mídias de *outdoor* e testar seu impacto visual, além de poder

avaliar questões como visibilidade e legibilidade do material exposto (vide Figura 1.1). Artistas gráficos também podem aprimorar suas pinturas em paredes, atualmente muito comuns nos grandes centros, testando virtualmente o resultado estético e fazendo um estudo de cores e tamanho de seus grafismos, antes de se tornarem reais.

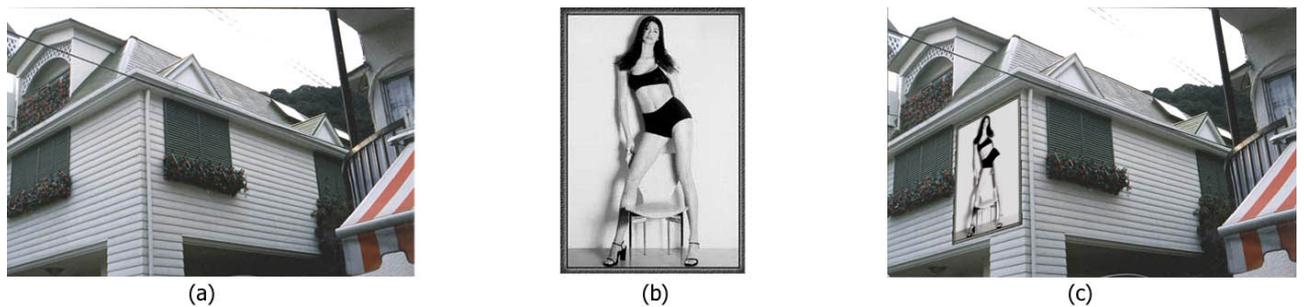


Figura 1.1: (a,b) Fotografias originais, (c) *outdoor* aplicado na lateral da casa

Na composição de imagens fotográficas, é comum o usuário querer inserir novas imagens em determinados planos da fotografia de origem ou obter uma nova vista da cena utilizando uma projeção que preserve proporções paralelas a um determinado plano. Muitas vezes, estas tarefas são executadas *a sentimento*, com o usuário deformando a imagem, por tentativa e erro, até obter o efeito desejado. Este processo, no entanto, é árduo, lento e nem sempre gera bons resultados, mesmo para usuários experientes e habilidosos. Além disso, ele acarreta um desperdício de tempo precioso dentro de prazos geralmente exíguos. Por exemplo, aplicar várias transformações em uma imagem de alta resolução, que resultará na saída, em *plotter*, de um poster de 5 metros, demanda vários minutos, mesmo se executadas em computadores velozes.

1.4 Trabalhos relacionados

Para que haja uma solução que não esteja ligada ao *sentimento* do usuário, é necessário recuperar informações do cenário tridimensional projetado na fotografia. O problema de

obter informações tridimensionais a partir de fotografias tem recebido bastante atenção na literatura, porém quase sempre com objetivos diferentes dos aqui propostos. Um grande número de trabalhos [Che95, McM, Hor97] aborda o problema de visualização utilizando técnicas de *image-based rendering*, cujo objetivo é obter novas vistas de uma cena a partir de um conjunto de imagens. Outros trabalhos abordam a reconstrução tridimensional da cena, utilizando visão estéreo [Fau98] ou adotando modelos matemáticos para os objetos lá presentes [Deba]. A seguir, reveremos brevemente alguns destes trabalhos.

- [Che95] A idéia deste sistema de realidade virtual é utilizar mapeamento de ambientes (em particular, imagens panorâmicas cilíndricas) para compor um ambiente virtual, de forma que possam ser feitas navegações interativas que simulam a movimentação de uma câmera. O mapeamento destes ambientes virtuais é formado por imagens independentes da orientação, permitindo ao observador olhar, arbitrariamente, em torno de qualquer direção. Isto ocorre porque as imagens contêm todas as informações necessárias para se olhar em torno de 360 graus. As imagens panorâmicas podem ser criadas com imagens geradas por computador, com câmeras panorâmicas especializadas ou sobrepondo fotografias tiradas por uma câmera comum para formar a panorâmica. Elas são articuladas de modo que definam uma cena. Além da cena acomodar simultaneamente imagens reais e virtuais, a navegação pode ser *tocada* (reproduzida) em tempo real sem auxílio de aceleração por *hardware*, não exigindo nenhum dispositivo de entrada ou saída especial (ex: *head-mounts*), e é capaz de gerar imagens de alta qualidade independente da complexidade da cena.
- [McM] Este sistema de *image-based rendering* é denominado Modelagem Plenóptica. Seu objetivo é produzir animações realistas, utilizando como entrada uma seqüência de imagens próximas e pouco espaçadas entre si, onde alguma parte de cada imagem na seqüência precisa ser vista na imagem que a sucede e alguma parte da imagem final precisa ser vista na primeira imagem da seqüência. Para capturar as imagens, basta uma câmera de vídeo e um tripé capaz de gerar uma visão pa-

norâmica contínua. Isto implica no problema de muitas imagens de entrada que precisam ser armazenadas na memória do computador. As imagens de referência sofrem uma transformação de *warping* e, combinadas, possibilitam formar as representações da cena a partir de um ponto de vista arbitrário. O sistema baseia-se na função plenóptica, formalizada por Adelson e Bergen, que cria um espaço parametrizado onde idealiza-se um observador que pode ser colocado em qualquer ponto (V_x, V_y, V_z) do espaço, a partir do qual qualquer raio de visualização pode ser selecionado, escolhendo-se um ângulo de fundo e de elevação (Θ, ϕ) , bem como uma banda de comprimento de onda λ e um tempo t , no caso da cena dinâmica. Ou seja, a função plenóptica descreve todas as informações da imagem visível a partir de um ponto de vista especificado, conforme ilustra a Figura 1.2.

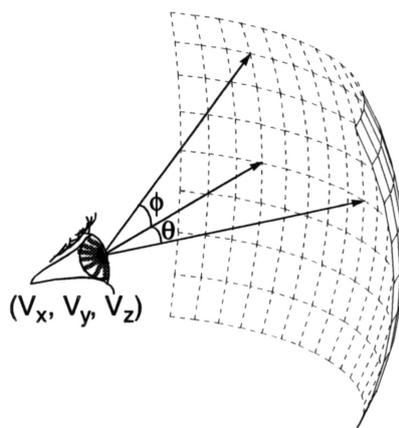


Figura 1.2: Função plenóptica

Os métodos descritos acima recaem no uso de técnicas de visão computacional para determinar automaticamente a estrutura da cena. Dentro do *framework* destes trabalhos, capturar os dados para uma visualização realista requer um número impraticável de fotografias sobrepostas com um pequeno espaçamento entre elas, e derivar a profundidade a partir destas fotografias pode demandar uma quantidade também impraticável de entradas do usuário. Além disso, as fotografias precisam parecer bastante semelhantes para que sejam obtidos resultados confiáveis.

- [Deba] Já este método tem o objetivo de modelar e *renderizar* cenas de arquiteturas existentes a partir de um conjunto esparsa de fotografias estáticas. É classificado como um método híbrido, pois combina técnicas baseadas em geometria e técnicas baseadas em imagens. O modelo geométrico básico é recuperado interativamente utilizando um sistema de modelagem de fotogrametria, que explora as restrições características de uma cena arquitetônica. Mais detalhes geométricos do modelo podem ser recuperados automaticamente através de correspondências estéreo com o usuário, determinando, com segmentos de retas em vez de pontos, bordas na imagem e suas correspondências no modelo. Isto permite recuperar o quanto a cena real se desvia do modelo básico. Desse modo, as novas vistas são criadas colocando uma câmera virtual na posição desejada e as imagens finais podem ser *renderizadas* utilizando-se uma técnica atual de *image-based rendering*, que é um mapeamento de textura dependente do ponto de vista.

Para reconstruir uma cena arquitetônica utilizando o método de modelagem de fotogrametria, foi implementado o sistema *Façade*, no qual determinam-se os parâmetros de um modelo hierárquico com primitivas de poliedros parametrizados. O usuário seleciona um número pequeno de fotografias para começar o modelo, que vai sendo refinado, incluindo mais imagens, até alcançar o nível de detalhe desejado. Componentes do modelo, chamados de blocos, são primitivas geométricas como quadrados, prismas e superfícies de revolução que representam as estruturas não poliédricas (ex: colunas, cúpula, torre de uma mesquita). Um quadrado, por exemplo, é parametrizado por seu comprimento, largura e altura. O usuário não precisa especificar numericamente valores para os parâmetros do quadrado, desde que eles sejam recuperados pelo programa implementado. Fazendo uso do modelo, a abordagem estéreo pode, robustamente, recuperar com precisão a profundidade a partir de imagens com grande *baseline*. Conseqüentemente, é possível modelar grandes ambientes arquitetônicos com quantidades menores de fotografias mais esparsas do que as prévias abordagens de sistemas baseados em imagens, criando assim, com habilidade, ren-

derizações realistas de cenas arquitetônicas a partir de pontos de vista distantes das fotografias originais.

- [Hor97] É um método para criar animações a partir do ponto de vista da câmera de uma fotografia ou pintura 2D de alguma cena, podendo *sobrevoar* tridimensionalmente por esta fotografia ou pintura. Assim como no método que será apresentado nesta dissertação, o método de [Hor97] possui apenas uma imagem disponível e não tem nenhum conhecimento do modelo 3D da cena. Como, teoricamente, recriar animações nestas condições é considerado impossível, este método pode ser visto como um novo tipo de efeito visual, que não tem a pretensão de construir precisamente um modelo 3D rígido da cena para a produção de animações. Para isso, não é possível aplicar as técnicas convencionais de visão computacional no processo de modelagem 3D da cena. Com o auxílio de uma interface gráfica, emprega-se uma malha, na forma de teia de aranha, para estabelecer as condições de perspectiva da imagem e obter o modelo. O processo de modelagem inicia-se com o usuário determinando o ponto de fuga na imagem 2D. Este modelo é como a forma de um poliedro com o ponto de fuga estando na sua base. A estrutura tridimensional é geometricamente formada por uma coleção de polígonos 3D hierárquicos, que constroem e distinguem o fundo da cena dos objetos do primeiro plano.
- [Fau98] Este trabalho propõe algumas soluções para o problema de reconstrução de um modelo tridimensional texturizado (*renderizado*) de um ambiente estático visualizado por uma ou mais câmeras cujos movimentos ou posições relativas sejam desconhecidos e cujos parâmetros intrínsecos, além de também desconhecidos, possam variar de uma fotografia para outra. Com o auxílio de ferramentas de visão computacional, recupera-se a geometria epipolar a partir da determinação de pontos de correspondência entre as imagens, e com isso calcula-se o conjunto completo de matrizes de projeção perspectiva para todas as posições da câmera. A geometria da cena é modelada como um conjunto de poliedros, enquanto as texturas são extraídas automaticamente das imagens e mapeadas nos polígonos. O método mostra

como diversas imagens podem ser combinadas em mosaicos a fim de que artefatos visuais, tais como pedestres ou árvores, sejam removidos automaticamente das texturas. Longe de ser um modelador completamente automático, ele possibilita reduzir significativamente a quantidade de interações manuais geralmente necessárias para obter modelos tridimensionais. A partir de uma série de imagens do ambiente que se deseja modelar, o operador trabalha com um sistema CAD que interage com um servidor de técnicas de visão. O trabalho também mostra que a completa geometria projetiva, afim e euclidiana da cena pode ser capturada com precisão por uma combinação de técnicas que abrange assuntos distintos como detecção de fronteiras, cantos, junções e margens utilizando modelos não paramétricos (baseados em imagens) e paramétricos; busca de características na imagem dentro de uma seqüência de imagens; modelagem geométrica de correspondências nas imagens nos espaços projetivo, afim e euclidiano, com uma distinção clara entre estes espaços e a quantidade de informações que eles requerem; e a estimativa robusta de instanciação algébrica desta geometria.

Diferentemente destas abordagens, o método aqui apresentado não pretende obter uma reconstrução completa da cena. Em compensação, utiliza como dado de entrada uma única fotografia, sobre a qual não se supõe conhecida nenhuma informação métrica prévia, seja sobre a cena (como coordenadas de pontos de referência ou medidas de elementos) ou sobre o posicionamento da câmera.

Este trabalho tem vários pontos em comum com [P3D99], que fornece um ambiente para a extração de informações tridimensionais a partir de uma única foto. A ênfase daquele trabalho, no entanto, está na modelagem parcial de objetos tridimensionais da cena. Nosso objetivo, em contraste, é fornecer ferramentas para que o usuário possa editar de forma geometricamente correta, e não apenas *a sentimento*, planos na fotografia de origem, inserindo neles novos objetos com a perspectiva apropriada. O método também permite que o usuário extraia uma textura da fotografia de entrada com correção automática de

perspectiva e simule uma lente fotográfica capaz de corrigir a distorção angular provocada pela projeção perspectiva - o que nem sempre é possível com uma lente fotográfica real.

1.5 Organização do trabalho

O Capítulo 2 deste trabalho descreve alguns conceitos e técnicas utilizados no método proposto nesta dissertação.

O Capítulo 3 apresenta o trabalho desenvolvido para a edição tridimensional de fotografias arquitetônicas.

O Capítulo 4 descreve a aplicação, recursos disponíveis e interface, além de apresentar exemplos do uso da aplicação e discutir os testes realizados.

Finalmente, o Capítulo 5 conclui o trabalho, fazendo uma revisão e análise do que foi apresentado e sugerindo novos caminhos para trabalhos futuros.

Capítulo 2

Calibração de câmera

Este capítulo apresenta o modelo de câmera adotado e como é realizada a calibração de câmera para os fins deste trabalho.

Uma das condições para a construção de representações geométricas a partir de fotografias é conhecer ou conseguir identificar os parâmetros – divididos em *intrínsecos* e *extrínsecos* – da câmera que capturou a respectiva imagem. Com isso, obtém-se uma transformação que, dado um ponto no espaço, encontra seu ponto correspondente na imagem. Este problema é denominado *calibração de câmera* e um estudo mais detalhado pode ser encontrado em [Fau93, Tom].

Os parâmetros intrínsecos são aqueles que definem a geometria interna da câmera e suas características óticas (*distância focal, centro projetivo, constante de distorção da lente, fator de escala em x*). Já os parâmetros extrínsecos, ou de orientação exterior, representam a *posição* da origem do referencial imagem e sua *orientação* em relação ao referencial objeto (do mundo real). Podemos dizer que os parâmetros extrínsecos produzem a relação espacial entre a câmera e o mundo e correspondem a uma translação, expressa por três números, e três rotações, também expressa por três números cada uma.

Nesta dissertação, estamos preocupados em encontrar, dentre os parâmetros citados, a orientação e a posição virtual da câmera no espaço 3D. O cálculo destes parâmetros será visto em detalhes na seção 2.3.

2.1 Modelo de câmera

É necessário estabelecer o modelo de câmera que será utilizado no método proposto. Neste trabalho, utilizamos o modelo simples de câmera *pin-hole*, que não considera a existência de lentes: a imagem de um objeto 3D é obtida por meio de sua projeção perspectiva em uma superfície plana 2D através do centro ótico da câmera (*pin-hole*).

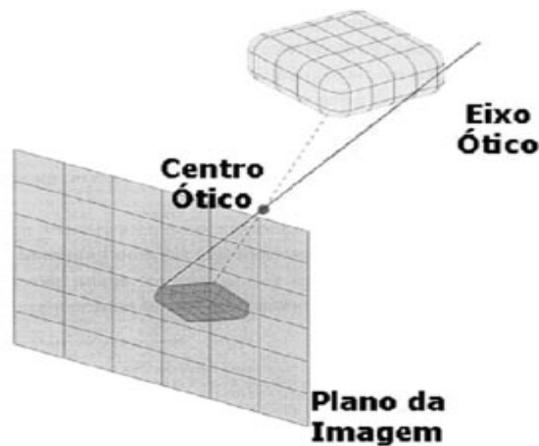


Figura 2.1: Modelo *pin-hole* de câmera

2.1.1 Projeção perspectiva

Para um melhor entendimento do método proposto é necessário que o leitor esteja familiarizado com o conceito de *projeção perspectiva*. A projeção perspectiva e a projeção paralela são classes de projeções geométricas, e a distinção entre elas está na relação do

centro de projeção ao plano de projeção. Enquanto na projeção paralela o centro de projeção está a uma distância infinita e com os raios de projeção paralelos ao plano de projeção, na projeção perspectiva o centro de projeção está explicitamente especificado, ou seja, a distância é finita e os raios de projeção emanam do centro de projeção, passando por cada ponto do objeto e interceptando o plano de projeção [Fei97].

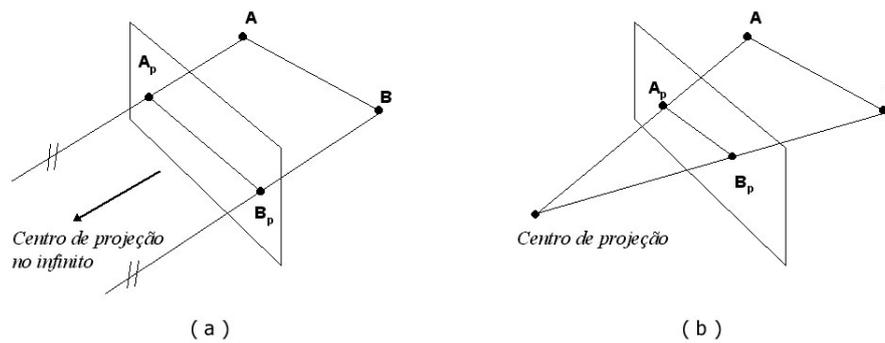


Figura 2.2: (a) Projeção paralela, (b) projeção perspectiva

O efeito visual de uma projeção perspectiva é similar a um sistema fotográfico e ao sistema visual humano. Uma projeção perspectiva preserva os elementos lineares mas não a relação de paralelismo. As projeções perspectivas de um conjunto de linhas paralelas que não estejam paralelas ao plano de projeção convergem para um ponto de fuga, que pode ser pensado como a projeção de um ponto no infinito. A geometria projetiva torna precisa esta noção ([Fau93]). Existe, é claro, uma infinidade de pontos de fuga, um para cada infinidade de direções a quais uma linha pode ser orientada. Como veremos na seção 2.2, estamos interessados em encontrar os pontos de fuga do conjunto de linhas paralelas às três direções principais do objeto arquitetônico presente na fotografia.

2.2 Método de calibração utilizado

Diversos métodos de calibração de câmera são baseados na existência de pontos de referência na imagem cujas coordenadas no mundo real sejam conhecidas [Car98, Tsa].

Outros [Deba] empregam informações métricas conhecidas, tais como medidas de elementos arquitetônicos como janelas ou portas. A semelhança entre estes métodos e o método aqui apresentado deve-se ao fato de que todos permitem escrever um sistema de equações a ser satisfeito pelos parâmetros considerados em cada um deles.

Neste trabalho supomos que nenhuma informação desse tipo está disponível. Isto é importante, por exemplo, em situações envolvendo fotos antigas, para as quais é impossível recuperar tais informações.

O método de calibração usado neste trabalho, que é o mesmo de [P3D99], recorre exclusivamente a informações relativas a direções – mais exatamente, aos pontos de fuga relativos a três direções mutuamente ortogonais. Como mostraremos a seguir, tais pontos de fuga determinam a posição da projeção do centro ótico da câmera e o ângulo de abertura da mesma. No entanto, na ausência de informações métricas complementares, é impossível resgatar a distância focal da câmera. Em consequência, é possível inserir objetos na cena com as proporções corretas, mas sem controle absoluto de seu tamanho. Por exemplo, na Figura 1, o tamanho do *outdoor* pode apenas ser especificado em relação a elementos da cena, como o pé direito da edificação.

A necessidade de identificar pontos de fuga de três direções ortogonais torna este método especialmente aplicável a fotografias arquitetônicas, onde geralmente existem três direções principais facilmente identificáveis. Admitimos também que os elementos a serem inseridos na imagem, ou dos quais se pretenda obter informações a partir da imagem, estejam alinhados com estas direções.

2.3 Obtenção da posição da câmera

Esta seção mostra como obter a posição da câmera em relação à imagem a partir de três pontos de fuga F_x , F_y e F_z relativos a três direções ortogonais X , Y e Z . Em fotos

arquitetônicas, a escolha destas direções é natural, correspondendo às direções principais de altura, largura e profundidade das edificações. Os pontos de fuga são os pontos de interseção de retas paralelas a estas três direções passando pelo centro ótico C com o plano de projeção. Eles determinam, portanto, um triedro tri-retângulo de vértice no centro ótico C . Esta condição permite recuperar a posição do centro ótico C relativamente à imagem.

Na verdade, o usuário não especifica diretamente os pontos de fuga F_x , F_y e F_z : ele indica, para cada uma das direções principais, duas ou mais retas na imagem que sejam projeções de retas da cena que tenham essas direções. O ponto de fuga correspondente é obtido através da interseção dessas retas.

No caso de duas retas, o ponto de interseção é encontrado sem problemas. Mas se três ou mais retas forem consideradas, ocorre que não existirá apenas um ponto de interseção entre todas elas, implicando no surgimento de dois ou mais pontos, em geral um ponto para cada par de retas, como ilustra a Figura 2.3. Para que apenas um ponto de fuga seja determinado, recai-se na resolução de um problema de mínimos quadrados. O ponto desejado é aquele cuja soma dos quadrados das distâncias às retas é mínima. Ou seja, sua determinação recai no seguinte problema de otimização:

$$\min \sum_{i=1}^k (a_i x + b_i y + c_i)^2 \quad (2.1)$$

onde $a_i x + b_i y + c_i = 0$ é a representação de cada reta, com $a_i^2 + b_i^2 = 1$. O problema acima pode ser visto como um sistema de equações lineares da forma $Ax = b$ com mais equações do que incógnitas e pode ser resolvido usando as equações normais $(A^T A)x = A^T b$ ou a decomposição de A em valores singulares [Car99].

Para recuperar a posição do centro ótico $C (u_c, v_c, w_c)$, conhecendo-se os pontos de fuga F_x , F_y e F_z é possível resolver o sistema de equações 2.2, que expressa a ortogonalidade de $\overrightarrow{CF_x}$, $\overrightarrow{CF_y}$ e $\overrightarrow{CF_z}$.

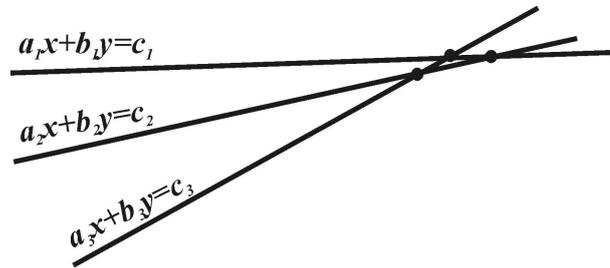


Figura 2.3: Interseção entre três segmentos de retas

$$\begin{aligned}
 (u_c - u_y)(u_c - u_x) + (v_c - v_y)(v_c - v_x) + w_c^2 &= 0 \\
 (u_c - u_y)(u_c - u_z) + (v_c - v_y)(v_c - v_z) + w_c^2 &= 0 \\
 (u_c - u_x)(u_c - u_z) + (v_c - v_x)(v_c - v_z) + w_c^2 &= 0
 \end{aligned} \tag{2.2}$$

Alternativamente, o centro ótico também pode ser determinado utilizando-se o teorema abaixo:

Teorema 2.1 *O centro ótico C projeta-se no ortocentro H do triângulo de vértices $F_x F_y F_z$.*

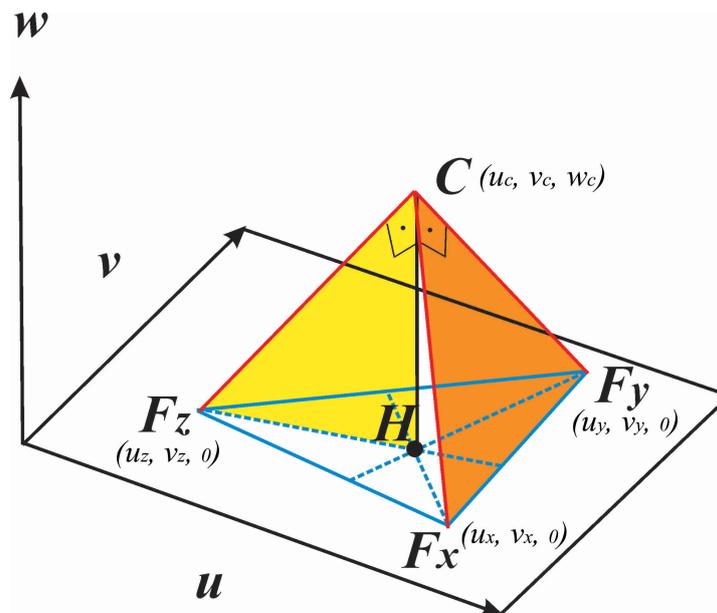


Figura 2.4: Localização do centro ótico

Demonstração

A reta F_zC é perpendicular ao plano do triângulo (C, F_x, F_y) ; logo, ela é ortogonal a qualquer reta contida neste plano. Em particular, F_zC é ortogonal a F_xF_y .

Por outro lado, a reta CH é perpendicular ao plano do triângulo (F_x, F_y, F_z) ; logo, ela é ortogonal a qualquer reta contida neste plano. Em particular, CH é ortogonal a F_xF_y .

Logo, o plano do triângulo (C, F_z, H) é perpendicular à reta F_xF_y , por conter duas retas concorrentes ortogonais a F_xF_y . Assim F_zH é perpendicular a F_xF_y , o que mostra que H está na altura relativa a F_z do triângulo (F_x, F_y, F_z) . De modo análogo, se mostra que H pertence às outras alturas. Portanto, H é o ortocentro do triângulo. ■

Para obter o ortocentro H de $F_xF_yF_z$ podemos exprimi-lo através dos vértices F_x, F_y e F_z na forma:

$$H = r_x F_x + r_y F_y + r_z F_z \quad (2.3)$$

onde r_x, r_y e r_z são suas coordenadas baricêntricas. O valor de r_x é a razão entre as áreas dos triângulos HF_yF_z e $F_xF_yF_z$. Logo:

$$r_x = \frac{HL_2}{F_xL_2} = \frac{L_1 \cos \alpha_y \cdot \cot \alpha_z}{L_1 \text{sen} \alpha_y} = \cot \alpha_y \cdot \cot \alpha_z \quad (2.4)$$

Os valores de r_y e r_z são obtidos de modo análogo, logo:

$$H = (\cot \alpha_y \cdot \cot \alpha_z) F_x + (\cot \alpha_x \cdot \cot \alpha_z) F_y + (\cot \alpha_x \cdot \cot \alpha_y) F_z \quad (2.5)$$

Uma vez obtida a projeção do centro ótico pelo método acima, sua distância w_c pode ser obtida por:

$$\begin{aligned} CF_x^2 &= HF_x^2 + w_c^2 \\ CF_y^2 &= HF_y^2 + w_c^2 \end{aligned} \quad (2.6)$$

Somando as equações 2.6, obtemos:

$$CF_x^2 + CF_y^2 = HF_x^2 + HF_y^2 + 2w_c^2 \quad (2.7)$$

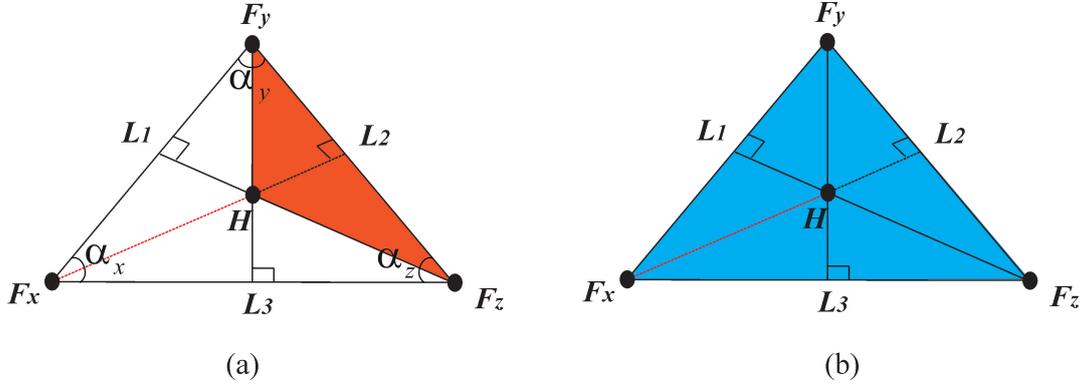


Figura 2.5: (a) Triângulo HF_yF_z , (b) triângulo $F_xF_yF_z$

Mas como o triângulo CF_xF_y é retângulo, temos:

$$CF_x^2 + CF_y^2 = F_xF_y^2 \quad (2.8)$$

Daí, obtemos:

$$F_xF_y^2 = HF_x^2 + HF_y^2 + 2w_c^2 \quad (2.9)$$

e, finalmente:

$$w_c^2 = \frac{F_xF_y^2 - HF_x^2 - HF_y^2}{2} \quad (2.10)$$

A posição (u_c, v_c, w_c) do centro ótico C é expressa no sistema de coordenadas da imagem S_i , que tem origem no seu canto inferior esquerdo, dois eixos u e v alinhados com as bordas da imagem e um terceiro eixo w ortogonal a elas. As coordenadas neste sistema são expressas em *pixels*. Como dito anteriormente, não é possível obter a distância focal real, a menos que se conheçam as dimensões de um *pixel*. É importante ressaltar, também, que a projeção do centro ótico não, necessariamente, coincide com o centro da imagem, vide a Figura 2.6 (b), uma vez que é possível que a imagem utilizada represente apenas uma porção da fotografia original.

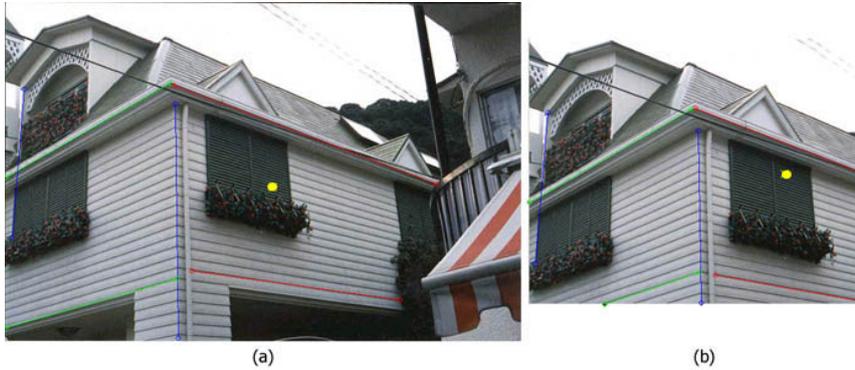


Figura 2.6: (a) Imagem toda, (b) imagem parcial

2.4 Equações de projeção

As equações de projeção da câmera associam a cada ponto do espaço a sua projeção perspectiva na imagem e, segundo [Tom], representam a orientação exterior da imagem. A posição de um ponto P no espaço é descrita através de suas coordenadas no sistema de coordenadas do mundo S_m , que tem origem no centro ótico C da câmera e eixos alinhados com as direções principais X , Y e Z . Vale notar que os eixos deste sistema passam, respectivamente, pelos pontos F_x , F_y e F_z . Como a distância focal real não é conhecida, as coordenadas neste sistema também são expressas em *pixels*. Para expressá-las em medidas convencionais de distância seria necessário, como já foi visto, conhecer as dimensões de um *pixel*.

As equações de projeção, portanto, levam um ponto (X, Y, Z) do mundo em sua projeção (u, v) na imagem. Esta transformação do espaço objeto para o espaço imagem pode ser realizada em dois passos. Primeiramente, para escrever estas equações, é conveniente considerar um sistema de coordenadas intermediário: o sistema da câmera S_c , com origem no centro ótico C e eixos U , V e W paralelos aos eixos do sistema de coordenadas da imagem S_i . A Figura 2.8 ilustra estes sistemas de coordenadas e a transformação espacial. Pode-se dizer que a transformação entre dois sistemas, em geral, envolve uma translação e uma rotação. Como a origem é comum aos sistemas, aqui há somente a rotação. Deste

modo, dado um ponto (X, Y, Z) do mundo, suas coordenadas no sistema da câmara são dadas por:

$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} = \begin{bmatrix} X_u & Y_u & Z_u \\ X_v & Y_v & Z_v \\ X_w & Y_w & Z_w \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (2.11)$$

onde (X_u, X_v, X_w) , (Y_u, Y_v, Y_w) e (Z_u, Z_v, Z_w) são os vetores unitários correspondentes a $C\vec{F}_x$, $C\vec{F}_y$ e $C\vec{F}_z$, respectivamente.

Uma vez obtidas as coordenadas de câmara, sua projeção (u, v) é obtida imediatamente, pelas equações:

$$\begin{aligned} u &= \frac{w_c U}{W} + u_c \\ v &= \frac{w_c V}{W} + v_c \end{aligned} \quad (2.12)$$

Representando (u, v) na forma homogênea (tu, tv, t) , onde $t \neq 0$, essas equações podem ser expressas na forma:

$$\begin{bmatrix} tu \\ tv \\ t \end{bmatrix} = \begin{bmatrix} w_c & 0 & u_c \\ 0 & w_c & v_c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} U \\ V \\ W \end{bmatrix} \quad (2.13)$$

onde u_c e v_c representam a translação da origem no centro ótico C para a origem no canto inferior esquerdo da imagem, e w_c refere-se à distância focal dada em *pixels*.

Assim, a equação de projeção da câmara é expressa por:

$$\begin{bmatrix} tu \\ tv \\ t \end{bmatrix} = \begin{bmatrix} w_c & 0 & u_c \\ 0 & w_c & v_c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_u & Y_u & Z_u \\ X_v & Y_v & Z_v \\ X_w & Y_w & Z_w \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (2.14)$$

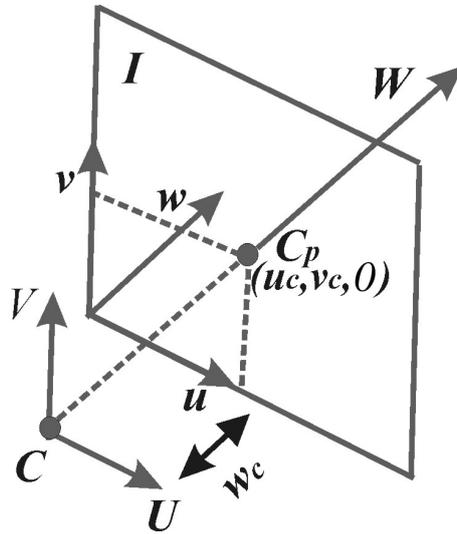


Figura 2.7: Translação da origem

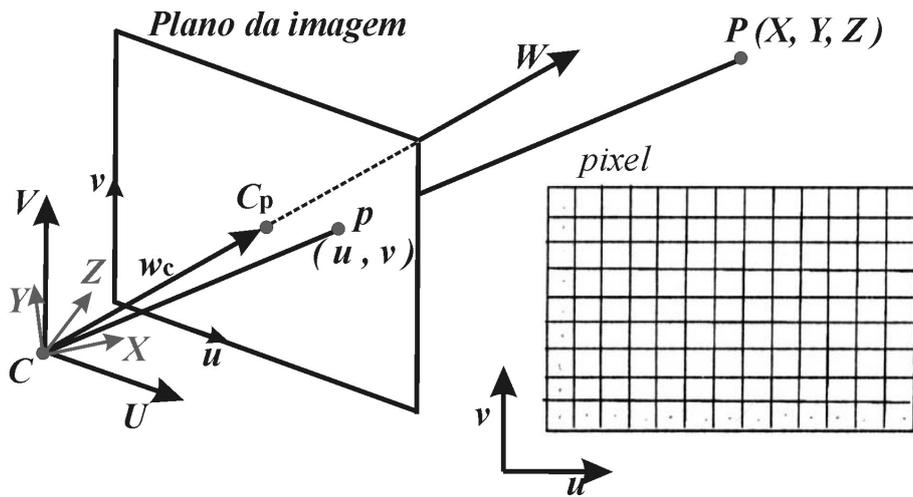


Figura 2.8: Sistemas de coordenadas

Capítulo 3

Edição tridimensional de fotografias

Este capítulo trata tridimensionalmente da edição de planos na fotografia de origem de maneira correta e precisa, que é o objetivo central do trabalho.

Com base nas técnicas descritas no Capítulo 2, desenvolvemos um método que atende a dois objetivos diferentes. Na seção 3.1, veremos em detalhes como é feita a inserção de novas imagens em planos da fotografia, paralelos a duas das direções principais, enquanto na seção 3.2 mostraremos como obter uma nova vista da cena utilizando uma projeção que preserve proporções paralelas a um determinado plano.

3.1 Edição em um plano principal

Uma vez obtidas a posição da câmera e as equações de projeção, é possível editar porções da imagem que correspondam a projeções de figuras (elementos) contidas em planos paralelos a dois dos eixos principais. Na exposição a seguir, consideramos planos paralelos aos eixos X e Y , sendo o tratamento análogo para os demais casos.

Admitimos que a porção da imagem a ser editada é a projeção de um retângulo com lados paralelos aos eixos X e Y , especificada através da escolha de dois pontos diagonalmente opostos p_1 e p_2 , correspondentes às projeções de dois vértices P_1 e P_2 do retângulo. Há uma infinidade de pontos do espaço que se projetam sobre p_1 e sobre p_2 , correspondentes às diversas profundidades Z em que o retângulo pode se situar. Como o nosso objetivo é apenas manter as proporções das dimensões e não recuperar seus valores absolutos, esta profundidade pode ser arbitrada. Fixamos, assim, $Z = 1$ e podemos então recuperar a posição de P_1 e P_2 e, em conseqüência, dos outros dois vértices Q_1 e Q_2 do retângulo e de suas projeções q_1 e q_2 .

Pelas projeções do vetor $\overrightarrow{P_1P_2}$ nas direções X e Y , encontramos os outros dois vértices do retângulo Q_1 e Q_2 , respectivamente, dados por:

$$Q_1 = P_1 + \overrightarrow{P_1P_2} \cdot X, \quad (3.1)$$

onde $\overrightarrow{P_1P_2} \cdot X = (x_2 - x_1, 0, 0)$

$$Q_2 = P_2 + \overrightarrow{P_1P_2} \cdot Y, \quad (3.2)$$

onde $\overrightarrow{P_1P_2} \cdot Y = (0, y_2 - y_1, 0)$

Estabelecemos, deste modo, uma correspondência entre um retângulo do espaço $P_1Q_1P_2Q_2$ e sua projeção – o quadrilátero $p_1q_1p_2q_2$ –, através da transformação projetiva bidimensional T' , que é a projeção perspectiva [Wol90] executada pela câmera restrita ao plano $Z = 1$. Esta correspondência permite recuperar uma imagem não deformada do quadrilátero $p_1q_1p_2q_2$, obtida através de um *warping* dado por T' . Sobre esta imagem não deformada podem ser aplicados novos elementos (letreros ou fotografias, por exemplo, passíveis de serem inseridos de volta na imagem original através do *warping* inverso ao aplicado no passo anterior. Todo esse processo está ilustrado na Figura 3.1.

Em muitos casos, pode ser de interesse executar apenas parte do processo acima – por exemplo, quando se deseja apenas extrair informações da cena, em vez de inserir novos

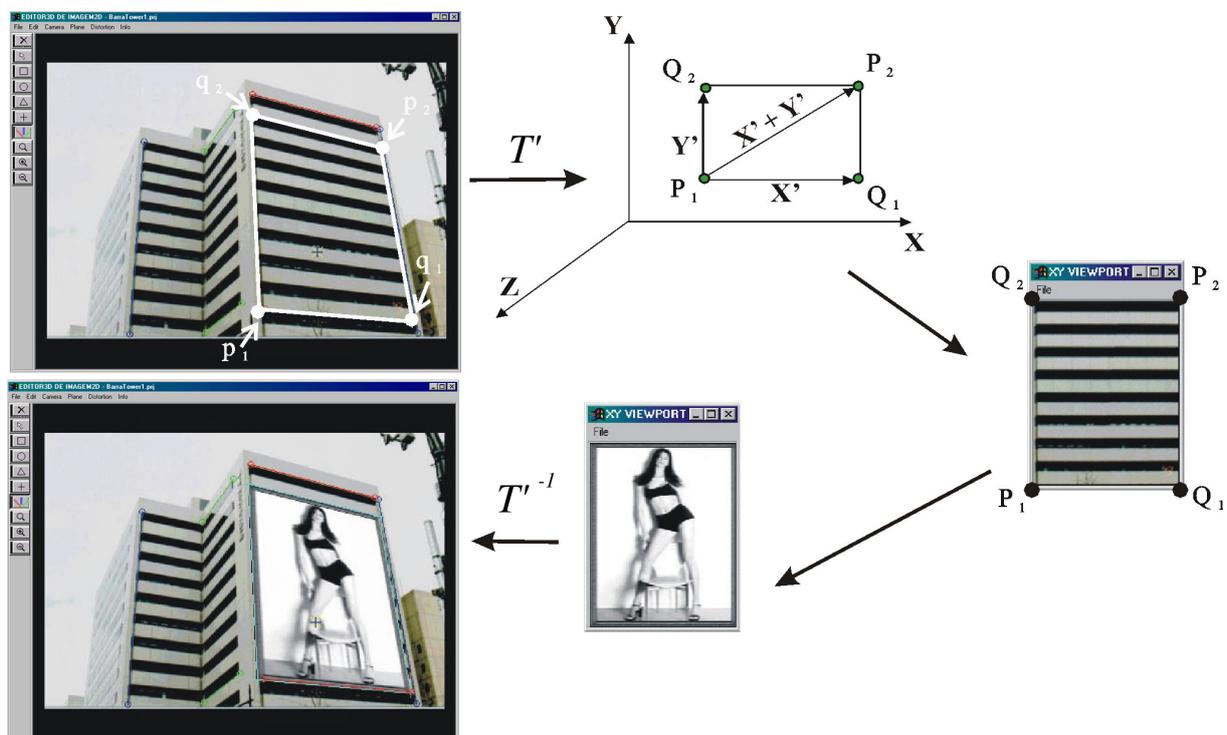


Figura 3.1: Etapas do processo de edição de um plano principal

elementos. A imagem não deformada exhibe corretamente as proporções entre os elementos presentes no plano de edição e preserva os ângulos entre estes elementos, podendo assim, ser empregada para obter relações entre as dimensões dos elementos (por exemplo, portas ou janelas) ou para extrair texturas lá presentes, que podem então ser utilizadas na construção de modelos virtuais da cena.

3.2 Correção da perspectiva

Embora as imagens produzidas através de projeção perspectiva sejam geometricamente corretas, elas nem sempre são a representação ideal do objeto de interesse. Em particular, os ângulos não são preservados, a menos que tenham lados paralelos ao plano de projeção. Por exemplo, quando se produz uma fotografia de um edifício cuja altura é grande em

relação à sua distância ao centro ótico da câmera (ou seja, em situações em que há pouca profundidade de campo), é preciso inclinar a máquina fotográfica, na tentativa de enquadrar a maior parte possível do objeto. Isto faz com que a direção vertical não seja paralela ao plano da foto, o que introduz um ponto de fuga correspondente a esta direção. Ou seja, o observador tem a impressão de que as arestas verticais do prédio se encontram. Embora isto seja perfeitamente natural, os arquitetos muitas vezes preferem imagens em que a direção vertical seja paralela à foto, de modo a evitar este efeito. Uma solução é empregar lentes especiais capazes de produzir fotos com as características desejadas [Nik01]. Nosso sistema oferece uma alternativa a essa solução, permitindo eliminar o ponto de fuga correspondente à direção vertical.

O que fazemos é reprojeter a fotografia original em um novo plano, paralelo à direção vertical Y e situado à mesma distância da câmera do plano de projeção original. Além disso, escolhemos os eixos de coordenadas neste novo plano de modo que Y se projete segundo o eixo vertical na nova foto. Isto equivale a introduzir uma nova câmera, com o mesmo centro ótico, a mesma distância focal e novos eixos de coordenadas α, β e γ . Estes são escolhidos de modo que β tenha a direção de Y e γ tenha a direção o mais próxima possível de W , o que ocorre tomando-se γ como a projeção de W no plano ortogonal a Y , conforme ilustra a Figura 3.2. Assim, os vetores $\vec{\alpha}$, $\vec{\beta}$ e $\vec{\gamma}$ correspondentes a estes novos eixos podem ser obtidos a partir dos vetores unitários \vec{Y} e \vec{W} (respectivamente vertical e ortogonal ao plano de projeção original) por meio das equações abaixo:

$$\begin{aligned}\vec{\beta} &= \vec{Y} \\ \vec{\gamma} &= \vec{W} - (\vec{W} \cdot \vec{Y})\vec{Y} \\ \vec{\alpha} &= \vec{\beta} \times \vec{\gamma}\end{aligned}\tag{3.3}$$

Finalmente, para obter a foto reprojetaada, basta considerar que cada ponto (u, v) da imagem é da forma (u, v, w_c) , transformá-lo para o sistema de coordenadas da nova câmera e aplicar a nova projeção perspectiva. Assim, um ponto (u, v) da foto original é transformado no ponto cujas coordenadas homogêneas (tu', tv', t) são dadas por:

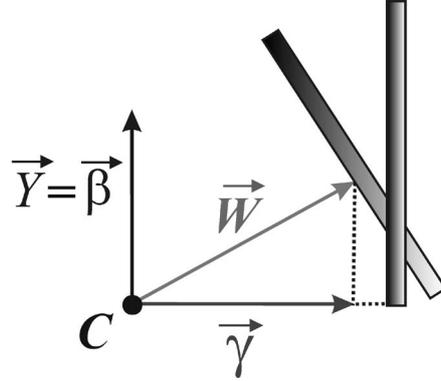


Figura 3.2: Processo de cálculo da nova câmera

$$\begin{bmatrix} tu' \\ tv' \\ t' \end{bmatrix} = \begin{bmatrix} w_c & 0 & 0 \\ 0 & w_c & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_u & \alpha_v & \alpha_w \\ \beta_u & \beta_v & \beta_w \\ \gamma_u & \gamma_v & \gamma_w \end{bmatrix} \begin{bmatrix} u \\ v \\ w_c \end{bmatrix} \quad (3.4)$$

onde $(\alpha_u, \alpha_v, \alpha_w)$, $(\beta_u, \beta_v, \beta_w)$ e $(\gamma_u, \gamma_v, \gamma_w)$ são os vetores unitários correspondentes aos vetores α , β e γ obtidos na equação (3.3).

A equação (3.4) determina uma transformação espacial de *warping* que permite gerar a nova imagem. O mapeamento planar é representado por duas superfícies, onde cada ponto 2D da imagem original se relaciona com o ponto 2D correspondente na imagem transformada.

Na Figura 3.3 podemos visualizar as imagens originais e os resultados das fotografias reprojctadas geradas pelo sistema. Pode-se observar o efeito desejado: as linhas correspondentes à direção vertical nas edificações tornam-se paralelas e verticais na foto reprojctada.



Figura 3.3: Imagens originais com inclinação vertical e reprojctadas com remoção de perspectiva

Capítulo 4

Aplicação e resultados

Neste capítulo é apresentada a aplicação desenvolvida para a edição tridimensional de fotografias arquitetônicas. Em seguida serão mostrados alguns resultados obtidos em testes feitos nessa aplicação e uma análise desses resultados.

4.1 Aplicação

Foi desenvolvida uma aplicação para a edição tridimensional de fotografias arquitetônicas, que disponibiliza para o usuário o método apresentado no capítulo anterior. As ferramentas implementadas possibilitam ao usuário inserir novas imagens em planos paralelos a duas direções principais da fotografia original, extrair texturas com a devida correção de perspectiva, recuperar fachadas para verificar as devidas proporções de seus elementos (por exemplo janelas, portas) e corrigir a perspectiva vertical da imagem original. O uso destas ferramentas e os resultados de testes avaliando os objetivos descritos acima serão expostos ao longo deste capítulo. Iniciaremos descrevendo a tecnologia utilizada na implementação; depois mostraremos a interface da aplicação e seus recursos disponíveis; em seguida, descreveremos uma seqüência referente à utilização das

ferramentas; e finalmente apresentaremos os resultados obtidos com os testes realizados.

4.1.1 Tecnologia utilizada na implementação

A linguagem de programação utilizada na implementação foi C, por apresentar boas características de extensibilidade e desempenho. O compilador utilizado foi o Microsoft Visual C/C++ 6.0 (Windows 2000 e Windows NT).

Para o sistema de janela e interface entre o sistema operacional e o usuário, decidimos pela utilização da biblioteca IUP/C [Lev96] por sua portabilidade, que permite a execução de um programa sem modificações em diferentes sistemas operacionais: Windows 95/98, Windows NT, Windows 2000 e Unix (AIX, IRIX, Linux, SunOS e Solaris). Além disso, a biblioteca oferece uma variedade de funções para a criação de diálogos. A biblioteca gráfica escolhida foi o CanvasDraw (CD) [dTGTa] pela sua simplicidade, por ser independente de plataforma e por estar disponível para diversos compiladores e sistemas operacionais. Para abrir e gravar as imagens de trabalho utilizamos a biblioteca IM [dTGTb], uma biblioteca de acesso a arquivos de imagens *raster*.

4.1.2 Interface

A interface da aplicação é composta por três elementos que compõem o diálogo principal: um *canvas* de trabalho, uma barra de menus e uma barra de ferramentas. A aplicação oferece vários recursos que estão agrupados na barra de menus e na barra de ferramentas. Existe também um diálogo, que chamamos de *viewport*, que será descrito com detalhes mais adiante, junto com sua utilização.

Os recursos da barra de menus estão disponíveis no primeiro nível pelos menus: *File*, *Edit*, *Camera*, *Draw*, *Info* (Figura 4.1).

O menu *File* possui opções como abrir, fechar e salvar uma imagem (*raster*), abrir,



Figura 4.1: Barra de menus



Figura 4.2: (a) Menu *File*, (b) menu *Edit*, (c) menu *Camera*

fechar e salvar um projeto, além de sair da aplicação. As opções de abrir e salvar um projeto agilizam o trabalho do usuário, uma vez que um projeto armazena dados da etapa de calibração da câmera – centro ótico, centro projetivo, posição das linhas guias, direções principais do mundo – e da última porção editada da imagem – pontos do retângulo no espaço e de sua projeção (quadrilátero) (Figura 4.2(a)).

O menu *Edit* tem a opção de limpar todo o *canvas*, excluindo a imagem, o centro ótico, as linhas guias e outras primitivas geométricas da área do *canvas* de trabalho. Também permite limpar o centro ótico ou o centro projetivo caso haja necessidade de manuseiar (mover) as linhas guias e recalculer a calibração de câmera (Figura 4.2(b)).

O menu *Camera* está diretamente relacionado com a etapa de calibração. Com esta opção, o sistema calcula a posição e a orientação da câmera (Figura 4.2(c)).

O menu *Plane* disponibiliza os recursos relacionados à edição em um dos planos da imagem de origem, como foi descrito na seção 3.1. Os itens do menu *Plane* estão agru-

pados em dois processos: *Set Window* e *Show Viewport* (Figura 4.3(a)). O primeiro, *Set Window*, necessita da escolha prévia, pelo usuário, de dois pontos em um dos planos principais da imagem. Então, com a solicitação de um de seus itens de submenu *PlaneXY*, *PlaneYZ* ou *PlaneXZ*, o sistema desenha o quadrilátero projetado no respectivo plano escolhido. Esta etapa permite que o usuário ajuste a posição dos pontos até obter a porção exata da imagem que deseja editar, certificando-se da proporção do quadrilátero onde será inserida a nova foto. O segundo processo, *Show Viewport*, abre uma outra janela de diálogo, que exhibe a porção da imagem contida no quadrilátero desenhado na imagem original. A janela aberta tem as medidas do retângulo no espaço e já exhibe a imagem não deformada, resultante da transformação de *warping*, derivada na seção 2.4. Nesta janela, que chamamos de *viewport*, também existe uma barra com o menu *File*, como ilustra a Figura 4.3(c), onde os itens de submenu são:

- *Open Outdoor*: para abrir a nova imagem *raster* que será inserida sobre a porção da imagem original.
- *Place Outdoor*: insere a nova imagem no quadrilátero desenhado na imagem original, correspondente à projeção do retângulo da janela da *viewport*.
- *Save Outdoor*: permite salvar a nova imagem composta com a imagem de fundo.
- *Save Texture*: salva parte da imagem contida no quadrilátero após a transformação de *warping*, possibilitando a extração de alguma textura da imagem original com a devida correção de perspectiva.
- *Close Images*: fecha tanto a imagem que foi transformada da fotografia original quanto a nova imagem inserida.
- *Exit Viewport*: fecha a janela da *viewport*.

O menu *Distortion* serve para eliminar o ponto de fuga correspondente à direção vertical e com isso corrigir a distorção de perspectiva vertical da imagem, conforme o

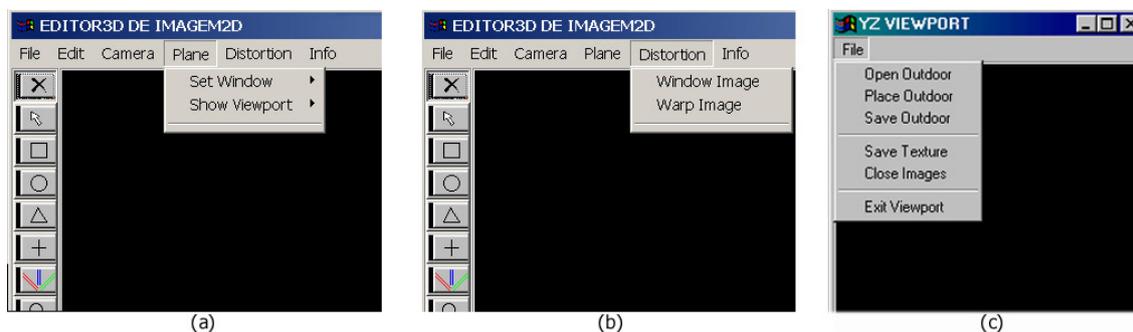


Figura 4.3: (a) Menu *Plane*, (b) menu *Distortion*, (c) menu *File* (*viewport*)

método apresentado na seção 3.2. A transformação espacial de *warping*, derivada na seção 3.2, ocorre na imagem toda. Uma nova imagem é gerada e exibida em outra janela de diálogo, também denominada *viewport* (Figura 4.3 (b)).

Outro elemento da aplicação é a barra de ferramentas (Figura 4.4), que fica a esquerda do diálogo principal, na forma vertical e é composta pelos seguintes botões : *exclusão* – para excluir objetos do canvas de trabalho –, *seleção* – para selecionar e movimentar os objetos no canvas de trabalho –, *desenho de primitivas* (círculo, quadrado e triângulo), *cruz* –para determinar dois pontos iniciais que formarão o retângulo projetado, referente a porção da imagem que se deseja editar– , *linhas guias* – para encontrar os pontos de fuga de cada direção principal – e *zoom*.

4.1.3 Utilização dos recursos

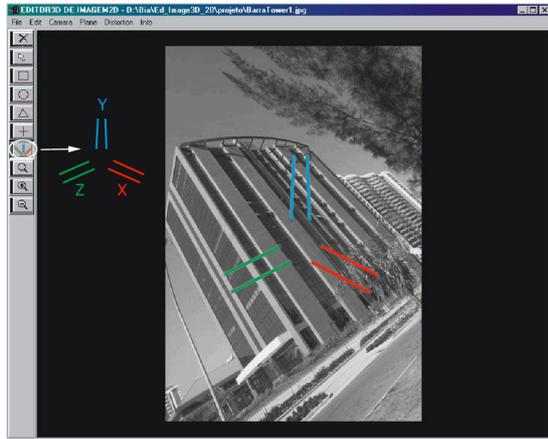
Para extrair informações tridimensionais de uma cena arquitetônica utilizando a aplicação desenvolvida, o usuário começa pela etapa de calibração da câmera, com o intuito de obter sua posição e orientação. Primeiramente, ele deve carregar uma imagem (*Open image*) com extensão .jpg ou .tif, referente à fotografia de alguma edificação de interesse. Como vimos no Capítulo 2, o método de calibração de câmera utilizado implica em encontrar os pontos de fuga relativos a três direções mutuamente ortogonais. Aproveitando-se de pistas na imagem, como margens, silhuetas e cantos retos, o usuário indica, para cada



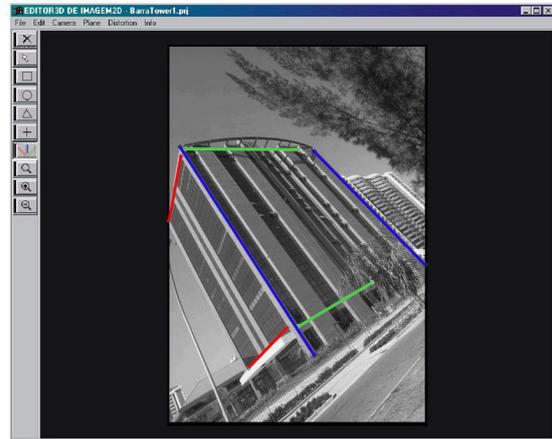
Figura 4.4: Barra de ferramentas

uma das três direções, duas retas na imagem que sejam projeções de retas da cena que tenham a respectiva direção. Para que o usuário indique as retas correspondentes a cada direção, o sistema oferece, na barra de ferramentas, uma interface com três pares de linhas guias para serem posicionadas nas retas da imagem. O par de linhas guias na cor azul corresponde à direção Y , na qual o usuário geralmente determinará a direção mais vertical do objeto. Já os pares de linhas guias nas cores vermelho e verde correspondem às direções X e Z , respectivamente, e estão associadas às direções mais horizontais do objeto, como ilustra a Figura 4.5. Em seguida, o usuário precisa apenas pedir que o sistema calcule a posição e orientação da câmera, escolhendo a opção *Camera/Calibration*, como ilustra a Figura 4.6.

Realizada a etapa de calibração de câmera, o usuário já pode editar porções na fotografia que estejam em planos paralelos às direções principais. Caso ele queira prosseguir em outro momento, o sistema permite salvar um projeto (*File/Save Project*), de extensão .prj, com os dados da calibração de câmera e carregá-lo (*File/Open Project*) posteriormente. É possível inserir imagens de *outdoor* e logomarcas ou apenas extrair parte da imagem original com correção de perspectiva. Para isso, basta indicar dois pontos diago-

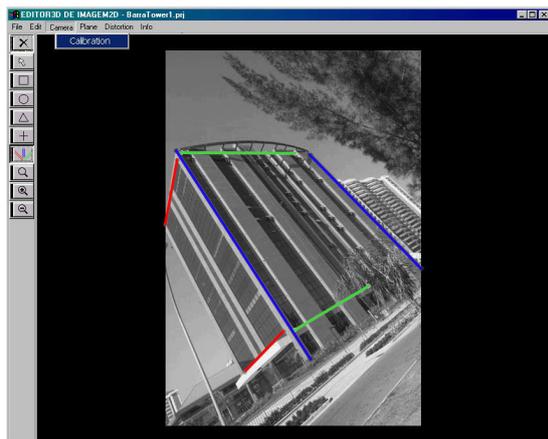


(a)

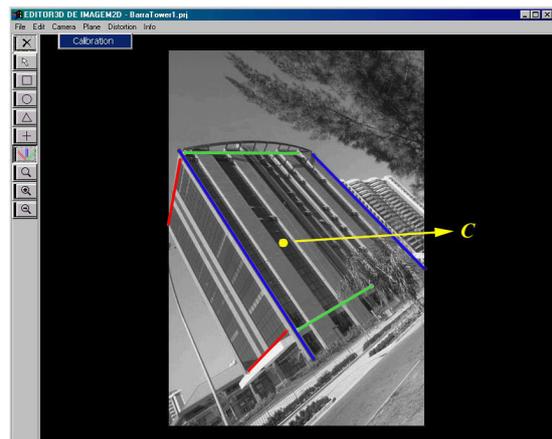


(b)

Figura 4.5: (a) Inicialização das linhas guias, (b) linhas guias posicionadas em retas que projetem as três direções principais



(a)



(b)

Figura 4.6: (a) Opção do menu *Camera* para calibração, (b) centro óptico C desenhado no *canvas*

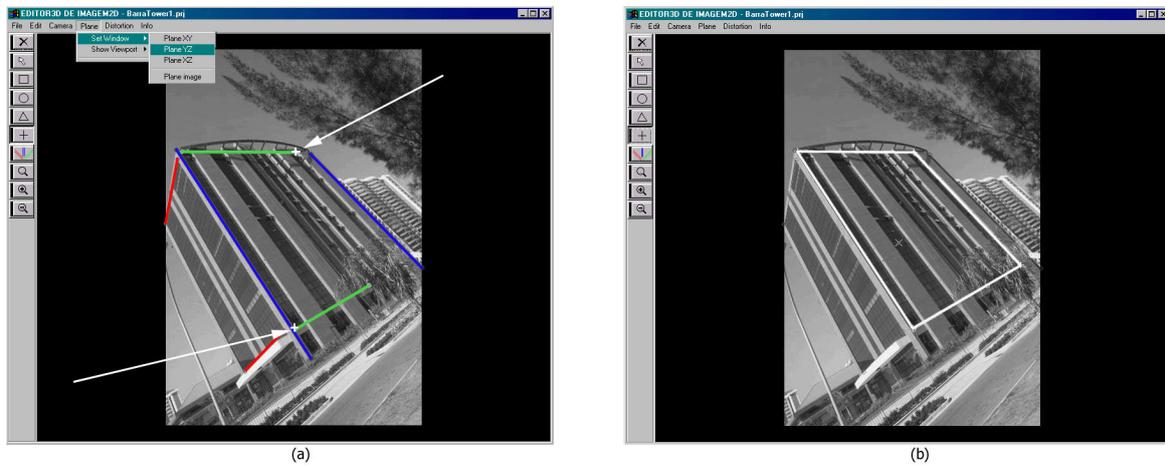


Figura 4.7: (a) Escolha de pontos para a determinação de uma janela de edição, (b) janela projetada

nalmente opostos no plano escolhido, através da *crux* na barra de ferramentas, clicando no primeiro ponto e arrastando até o segundo (Figura 4.7). De acordo com a posição indicada das linhas guias, o plano escolhido será XY , YZ ou XZ . Então, dependendo disso, deve-se escolher a opção *Plane/Set Window/PlaneXY*, *PlaneYZ* ou *PlaneXZ* e visualizar o retângulo projetado correspondente à região da fotografia onde será inserida uma nova imagem ou de onde será extraída uma textura (Figura 4.7). Seguindo a coerência do passo anterior, a *viewport* contendo parte da imagem relativa ao retângulo projetado aparecerá mediante a solicitação da opção de menu *Plane/Show Viewport/ViewXY*, *ViewYZ* ou *ViewXZ*, como ilustra a Figura 4.8. Até esta etapa, o usuário obteve parte da imagem original com a perspectiva corrigida, podendo utilizá-la como textura ou para verificar informações relativas às proporções de elementos da fachada (*File/Save texture*). Para inserir uma nova imagem ou logomarca sobre a fotografia original, basta escolher na barra de menus da janela *viewport* o item *File/Open Outdoor* (vide Figura 4.8 (b)) e *File/Place Outdoor* (vide Figura 4.9).

Se o usuário quiser obter a fotografia original com correção de perspectiva, é necessário seguir os passos *Distortion/Warp Image* (Figura 4.10).

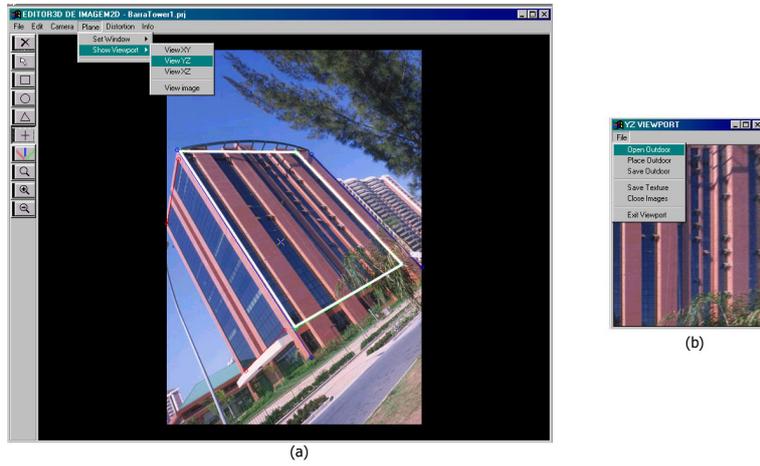


Figura 4.8: (a) Abrindo um *viewport*, (b) imagem aberta no novo *viewport*

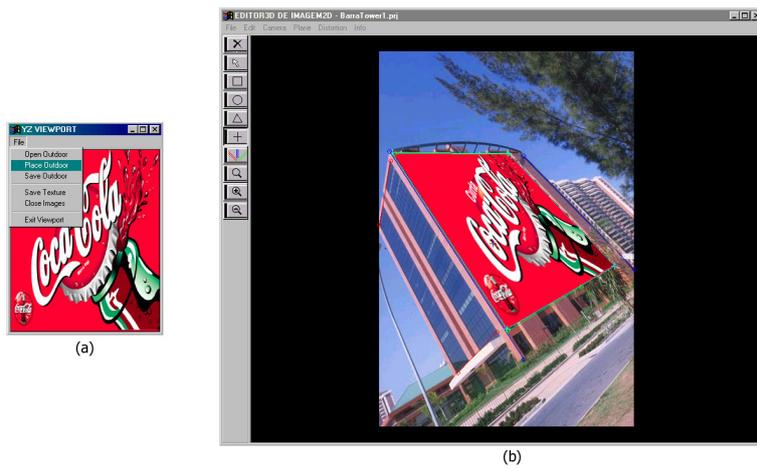
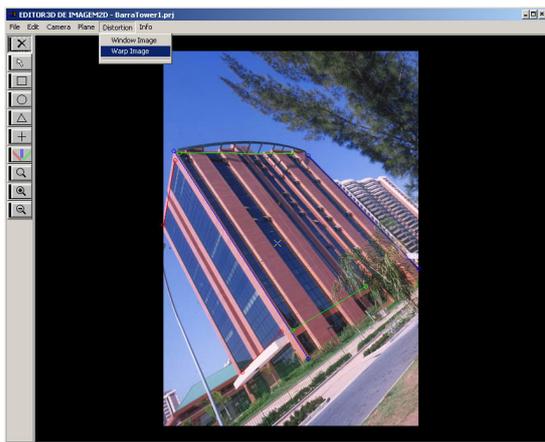
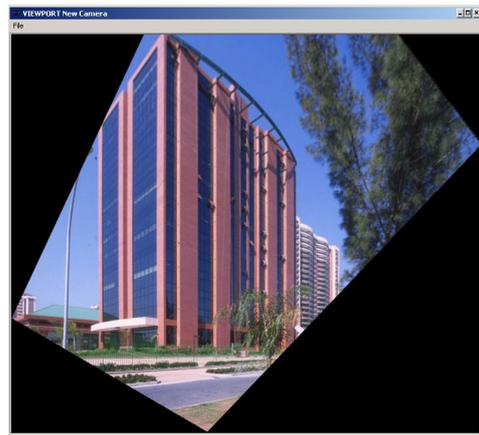


Figura 4.9: (a) *Outdoor* aberto na *viewport*, (b) *outdoor* inserido na fotografia original



(a)



(b)

Figura 4.10: (a) Solicitando *Warp Image*, (b) viewport com a imagem transformada

4.2 Resultados

4.2.1 Descrição dos testes

Os testes foram realizados com a finalidade de avaliar o comportamento do sistema em relação aos objetivos aos quais ele se propôs. Para os testes, foram utilizadas fotografias de diversas edificações, nas quais foi possível identificar pontos de fuga de três direções ortogonais. Embora algumas imagens correspondam à mesma edificação, foram fotografadas de diferentes pontos de vista ou com o mesmo ponto de vista mas com lentes fotográficas diferentes e com variação de ângulos de *pitch* (declividade) e de *roll* (rotação longitudinal).

Inicialmente, avaliamos a etapa de calibração de câmera, comparando os resultados alcançados com o sistema implementado com os obtidos com o sistema [P3D99], devido à semelhança entre eles para o cálculo da posição e orientação de câmera. Em seguida, avaliamos a posição da projeção do centro ótico e do centro da imagem, uma vez que, nas fotografias que foram produzidas por nós, temos o controle de que a imagem projetada no monitor corresponde à imagem projetada no retângulo do sistema ótico da câmera fotográfica. Podemos também verificar a relação entre a lente fotográfica utilizada e o valor encontrado para a coordenada z_c do centro ótico C , que assumimos como o parâmetro de distância focal no sistema implementado. Também verificamos se as novas imagens foram inseridas adequadamente em planos paralelos às direções principais e se a imagem ou textura retirada destes planos foi corretamente transformada a ponto de mostrar uma nova imagem sem distorção de perspectiva, de acordo com o método apresentado na seção 3.1. Por último, testamos as imagens originais e a qualidade das novas imagens geradas com o método da seção 3.2 para correção de perspectiva vertical.

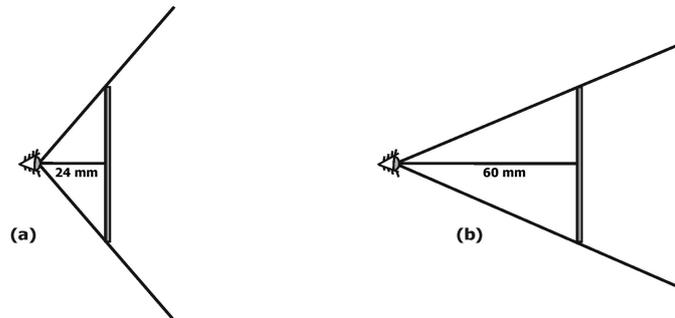


Figura 4.11: (a) *Frustum* de visão da lente 24 mm, (b) *frustum* de visão da lente 60 mm

4.2.2 Dados dos testes

O conjunto de dados testados soma 43 fotografias. Dentre elas, 20 foram obtidas na Internet através do site [P3D99] e as outras 23 foram produzidas utilizando uma câmera profissional (Nikon F4S) e duas lentes fotográficas com distâncias focais iguais a 24 mm e 60 mm. O fato de a distância focal da lente 24 mm ser menor faz com que seu *frustum* de visão seja mais largo (ela é considerada uma lente grande angular), como ilustra a Figura 4.11. A lente de 60 mm é considerada uma lente normal, por produzir resultados semelhantes ao que seria visto a olho nu.

Todos os arquivos referentes às fotografias das edificações possuem uma resolução máxima de 800×600 pixels. As novas fotografias que são inseridas em planos das imagens originais foram campanhas publicitárias reais e são compostas por elementos como produtos, logomarcas, textos e modelos.

4.2.3 Cálculo do centro ótico

Para testar a correção do algoritmo de obtenção do centro ótico, comparamos os resultados encontrados pelo sistema com os produzidos pelo sistema [P3D99], para diversas fotografias de diferentes edificações que foram obtidas no site do próprio [P3D99]. As Figuras 4.12, 4.13, 4.14, 4.15 e 4.16, apresentam alguns dos resultados encontrados. Em



Figura 4.12: (a) Centro ótico do Photo 3D, (b) centro ótico do sistema implementado todas, exceto na Figura 4.16, as projeções do centro ótico do sistema implementado coincidiram com as posições encontradas pelo sistema [P3D99].

Percebe-se que a utilização de linhas como fonte de controle para o processo de calibração de câmera gera bons resultados na medida que sejam bem escolhidas. A configuração bem escolhida destas linhas deve ser estabelecida de modo a evitar certas combinações como, por exemplo, todas as linhas paralelas ou alinhadas. A dimensão das linhas também é um fator importante: elas devem apresentar a maior dimensão possível na imagem e, se possível, serem distribuídas uniformemente.



Figura 4.13: (a) Centro ótico do Photo 3D, (b) centro ótico do sistema implementado



Figura 4.14: (a) Centro ótico do Photo 3D, (b) centro ótico do sistema implementado

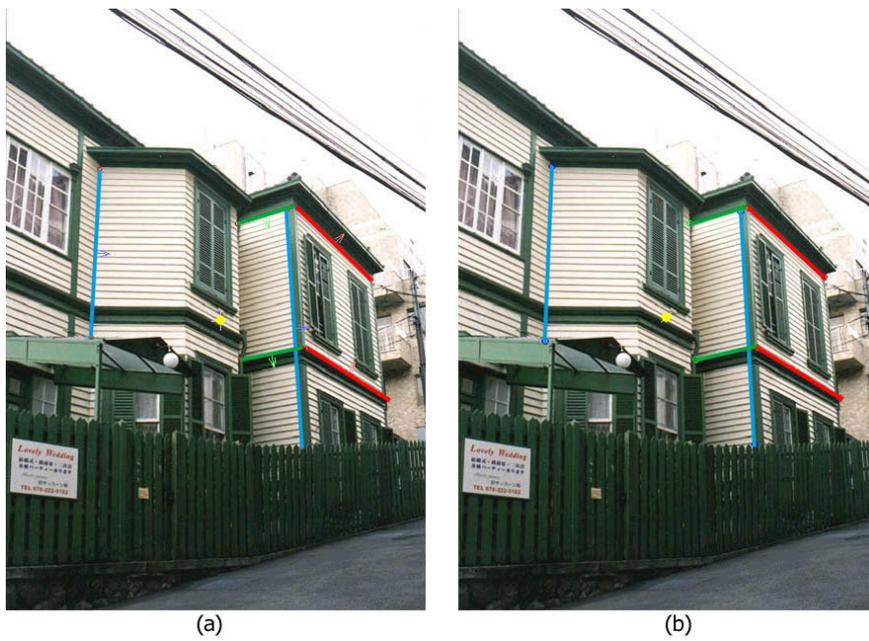


Figura 4.15: (a) Centro ótico do Photo 3D, (b) centro ótico do sistema implementado



Figura 4.16: (a) Centro ótico do Photo 3D, (b) centro ótico do sistema implementado

4.2.4 Comparação entre duas lentes fotográficas

Aqui, temos a oportunidade de comparar a projeção do centro ótico com o centro da imagem, assim como verificar a coerência dos valores obtidos para a coordenada z_c , considerada na implementação do sistema, como parâmetro de distância focal. Para obter a distância focal teórica, usamos o fato de que as dimensões de um negativo ou positivo do filme 35 *mm* são conhecidas (24×36 *mm*). Assim, conhecendo-se a dimensão da imagem em *pixels* e a distância focal da lente em milímetros (*mm*), pode-se obter, por proporcionalidade, o valor esperado em *pixels* para a distância focal. A Tabela 4.1 traz os resultados encontrados para um conjunto de imagens capturadas com lentes 24 *mm* e 60 *mm*.

As Figuras 4.17, 4.18, 4.19, 4.22 e 4.23 foram geradas do mesmo ponto de vista, sendo que as imagens da esquerda (1, 3, 5, 13 e 15) foram fotografadas com a lente 24 *mm* e as da direita (2, 4, 6, 14 e 16), com a lente 60 *mm*. Nota-se que a lente 60 *mm* provoca a sensação de mais proximidade do objeto, pois, devido à menor abertura angular. Já na lente 24 *mm*, uma área maior da cena ocupa o mesmo retângulo do visor da câmera, dando uma sensação de afastamento do objeto e gerando, também, maior distorção angular.

A Figura 4.20 é composta por três fotografias tiradas do mesmo ponto de vista, utilizando a lente 60 *mm* para a imagem da direita (9) e a 24 *mm* para as outras duas (7 e 8). Entretanto, a imagem mais à esquerda (7) teve uma variação de ângulo de *roll*. Os mesmos parâmetros geraram as imagens da Figura 4.21, porém a imagem (11) teve variação de ângulo de *pitch*.

Já as fotografias das Figuras 4.24 e 4.25 foram tiradas de distâncias diferentes do objeto para que tivéssemos a mesma área da edificação de interesse ocupando o retângulo do visor da câmera, tanto com a lente 24 *mm* quanto com a lente 60 *mm*.

Quanto à coincidência do centro projetivo com o centro da imagem, notamos que embora não tenha sido exata, gerou resultados coerentes, próximos ao centro da imagem.

A Tabela 4.1 e as imagens relacionadas a ela podem comprovar esta avaliação. Como foi dito na seção 4.2.3, quanto melhor o posicionamento das linhas guias que identificam as direções principais, melhor o resultado da calibração de câmera. Os valores obtidos para as distâncias focais são todavia consistentes com os resultados teóricos (Tabela 4.1). As diferenças encontradas podem ser explicadas pelo fato de que a distância focal impressa nas lentes das máquinas fotográficas é um parâmetro baseado no modelo de câmera *thick-lens*, que é um modelo mais preciso e envolve distorções radiais causadas pela curvatura das lentes. Embora este parâmetro nos diferentes modelos de câmera – *pin-hole* e *thick-lens* – possua medidas diferentes, ele gera efeitos semelhantes na imagem, e podemos ver pela Tabela 4.1 que são valores coerentes. Notamos, ainda, que os resultados das distâncias focais obtidos para a lente de 60 *mm* são mais próximos aos resultados teóricos. Isto se explica pela maior deformação introduzida pela lente grande angular 24 *mm*.

Tabela 4.1: Medidas referentes ao centro da imagem e à posição do centro ótico, com variação de lente fotográfica

No.	Descrição	Lente	Dim. (<i>pixel</i>)	Pos. teórica C (<i>pixel</i>)	Pos. calculada C (<i>pixel</i>)
1	Humaitá 1	24 <i>mm</i>	399×590	(199.5, 295.0, 393.3)	(208.7, 320.0, 214.3)
2	Humaitá 2	60 <i>mm</i>	396×590	(198.0, 295.0, 990.0)	(193.0, 336.3, 965.1)
3	Humaitá 3	24 <i>mm</i>	406×590	(203.0, 295.0, 393.3)	(193.3, 281.3, 281.42)
4	Humaitá 4	60 <i>mm</i>	399×590	(199.5, 295.0, 983.5)	(185.2, 369.8, 950.9)
5	NYCity 1	24 <i>mm</i>	790×534	(395.0,267.0, 526.6)	(492.8, 330.0, 422.4)
6	NYCity 2	60 <i>mm</i>	790×538	(395.0,269.0, 1316.6)	(493.1, 325.5, 1240.2)
7	BarraTower 1	24 <i>mm</i>	399×590	(199.5, 295.0, 393.3)	(179.7, 302.4, 265.5)
8	BarraTower 2	24 <i>mm</i>	403×590	(201.5, 295.0, 393.3)	(200.3, 305.3, 276.6)
9	BarraTower 3	60 <i>mm</i>	404×590	(202.0, 295.0, 983.3)	(205.5, 301.3, 971.7)
10	RioDesign 1	24 <i>mm</i>	790×540	(395.0, 270.0, 526.6)	(385.5, 258.3, 523.5)
11	RioDesign 2	24 <i>mm</i>	790×539	(395.0, 269.5, 526.6)	(427.1, 254.8, 416.8)
12	RioDesign 3	60 <i>mm</i>	790×551	(395.0, 275.0, 1377.5)	(385.0, 373.5, 1400.0)

continua na próxima página

Tabela 4.1: (continuação)

No.	Descrição	Lente	Dim. (<i>pixel</i>)	Pos. teórica $C(\textit{pixel})$	Pos. calculada $C(\textit{pixel})$
13	RioDesign 4	24mm	790×534	(395.0, 267.0, 526.6)	(400.3, 250.9, 459.3)
14	RioDesign 5	60mm	790×533	(395.0, 266.5, 1332.5)	(328.0, 244.3, 1397.8)
15	StaTeresa 1	24mm	404×590	(202.0, 295.0, 393.3)	(179.6, 346.1, 270.1)
16	StaTeresa 2	60mm	790×537	(395.0, 268.5, 1342.5)	(342.0, 312.5, 1338.8)
17	StaTeresa 3	24mm	790×537	(395.0, 268.5, 526.6)	(451.8, 349.6, 479.1)
18	StaTeresa 4	24mm	790×543	(395.0, 271.5, 526.6)	(394.2, 372.9, 535.4)
19	StaTeresa 5	60mm	790×538	(395.0, 269.0, 1345.0)	(463.0, 399.9, 1821.6)
20	StaTeresa 6	24mm	790×537	(395.0, 268.5, 537.0)	(390.8, 241.3, 442.1)
21	StaTeresa 7	60mm	399×590	(199.5, 295.0, 983.5)	(165.6, 346.7, 915.3)



Figura 4.17: (1) Humaitá 1, (2) Humaitá 2



Figura 4.18: (3) Humaitá 3, (4) Humaitá 4



Figura 4.19: (5) NYCity 1, (6) NYCity 2

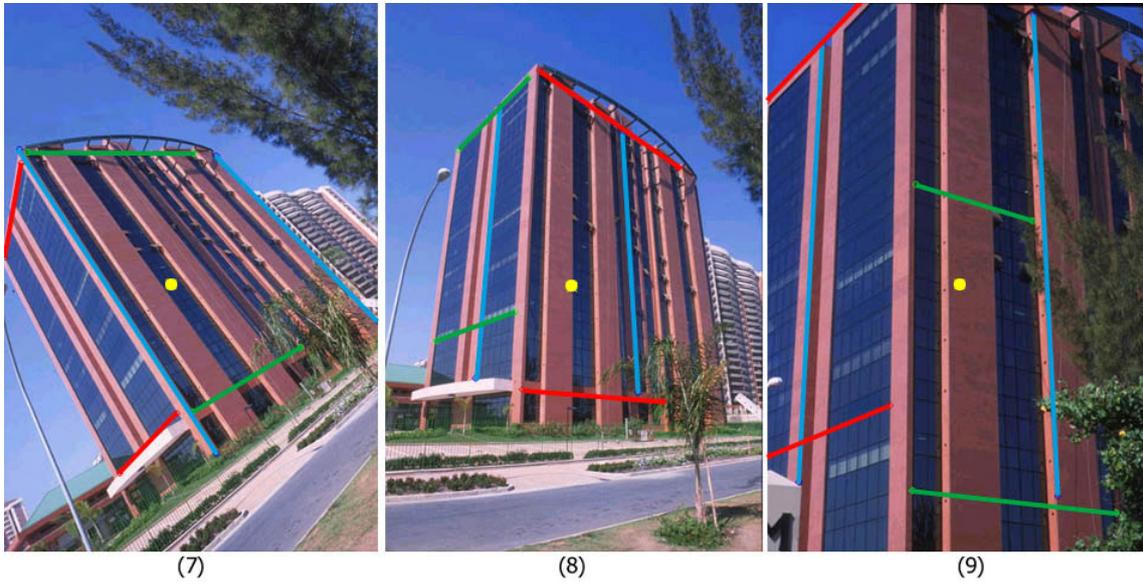


Figura 4.20: (7) Barra Tower 1, (8) Barra Tower 2, (9) Barra Tower 3



Figura 4.21: (10) Rio Design Barra 1, (11) Rio Design Barra 2, (12) Rio Design Barra 3



Figura 4.22: (13) Rio Design Barra 4, (14) Rio Design Barra 5

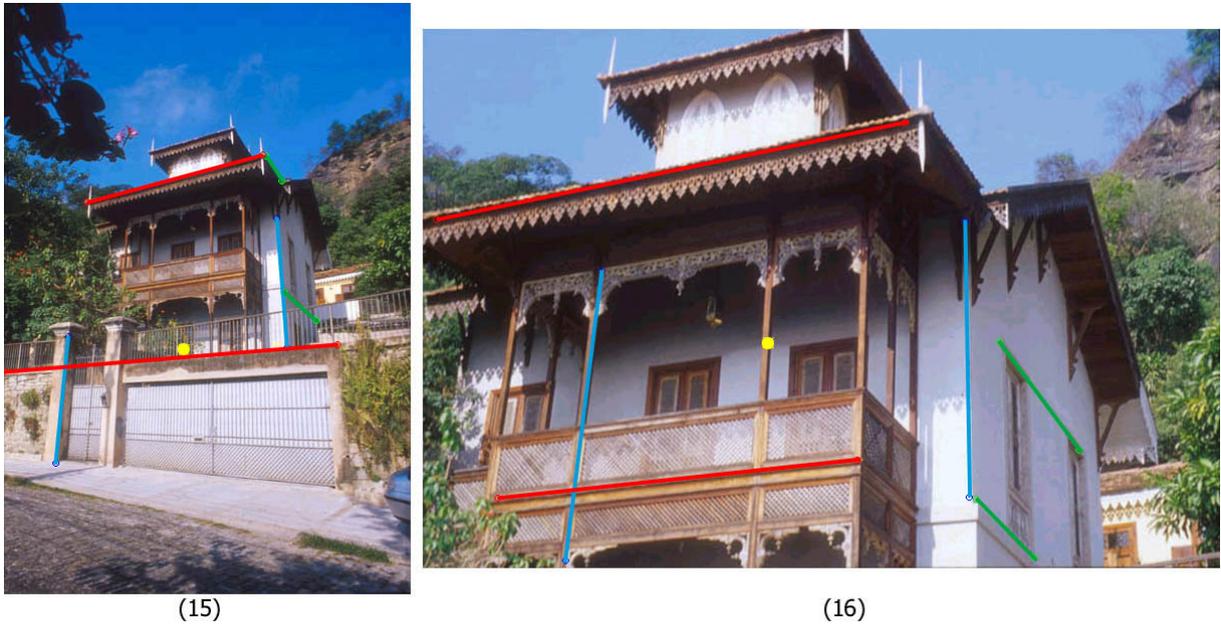


Figura 4.23: (15) Santa Teresa 1, (16) Santa Teresa 2



Figura 4.24: (17) Santa Teresa 3, (18) Santa Teresa 4, (19) Santa Teresa 5



Figura 4.25: (20) Santa Teresa 6, (21) Santa Teresa 7

4.2.5 Testes com inserção de *outdoors* e logomarcas

Para os testes de inserção de objetos em planos da fotografia original, utilizamos as mesmas imagens que foram testadas na seção 4.2.4, mas apenas alguns desses testes serão mostrados. As observações sobre o posicionamento e ângulos de inclinação da câmera fotográfica também valem aqui. Embora as novas imagens inseridas sejam arquivos do tipo *raster*, são referentes a campanhas publicitárias reais e possuem texto, logomarcas e elementos suficientes para a avaliação.

Cada figura ilustra a imagem original de uma campanha publicitária seguida dessa mesma imagem inserida em algum plano de uma edificação. A inserção de imagens de *outdoors* em planos da imagem original foi feita automaticamente e gerou excelentes resultados, porém algumas imagens sofreram uma manipulação posterior em um *software* de edição de imagem, visto que artefatos como postes de luz e árvores precisaram ser novamente inseridos na imagem para aumentar o realismo da cena (Figuras 4.29, 4.30 e 4.31). Isto é um procedimento natural e em nenhum momento tivemos a pretensão de oferecer ferramentas com a finalidade de executar estas tarefas automaticamente. Por outro lado, existem diversos programas que atendem bem ao objetivo de executar, de modo manual, este procedimento.

Na Figura 4.28, nota-se que o muro onde o grafismo foi inserido possui uma ligeira inclinação e não é paralelo a nenhuma das direções principais, produzindo uma imagem pouco real. Outro exemplo que foge a regra é mostrado na Figura 4.32, na qual a edificação de interesse não possui as paredes ortogonais.

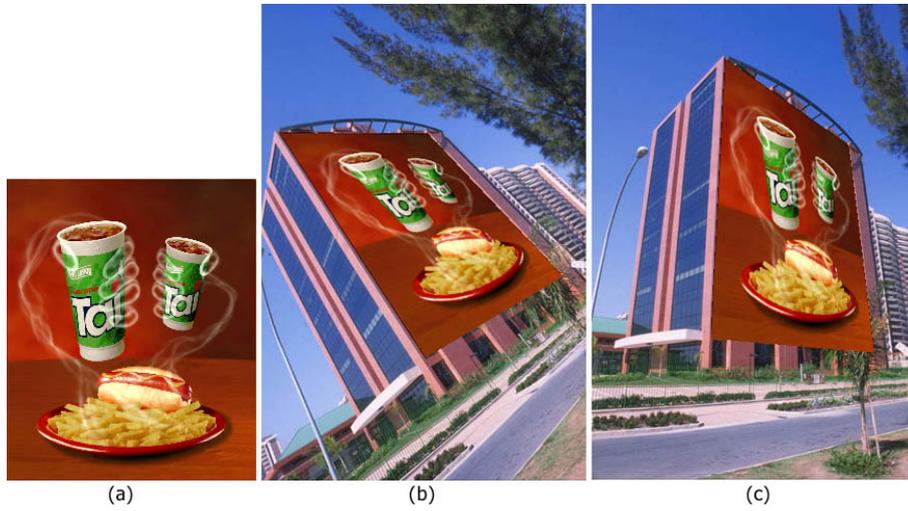


Figura 4.26: (a) imagem original do *outdoor*, (b) Barra Tower 1 com *outdoor* inserido, (c) Barra Tower 2 com *outdoor* inserido



Figura 4.27: (a) imagem original do *outdoor*, (b) Humaitá 3 com *outdoor* inserido, (c) Humaitá 4 com *outdoor* inserido



Figura 4.28: (a) imagem original do *outdoor*, (b) Santa Teresa 6 com *outdoor* inserido, (c) Santa Teresa 4 com *outdoor* inserido

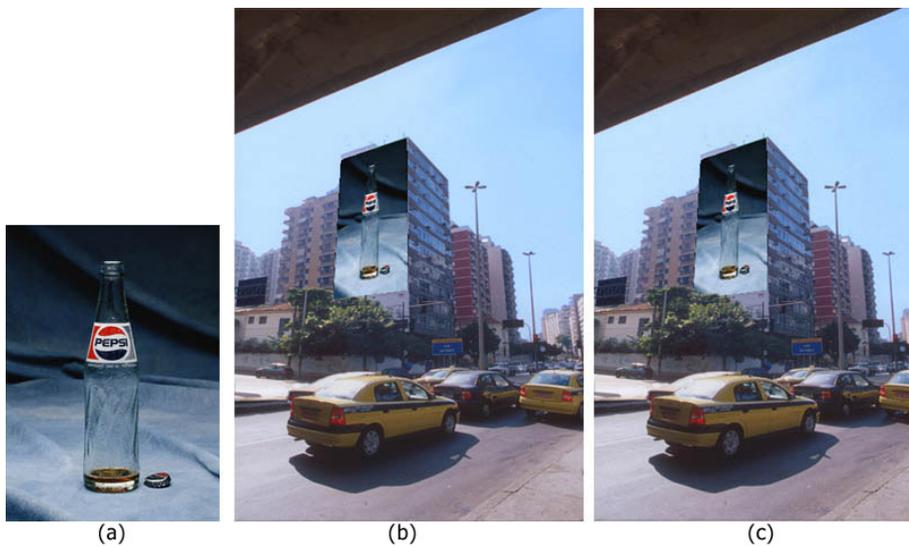


Figura 4.29: (a) imagem original do *outdoor*, (b) Humaitá 1 com *outdoor* inserido sem retoque, (c) Humaitá 1 com *outdoor* inserido com retoque



Figura 4.30: (a) imagem original do *outdoor*, (b) Rio Design Barra 5 com *outdoor* inserido sem retoque, (c) Rio Design Barra 5 com *outdoor* inserido com retoque



Figura 4.31: (a) 3 imagens originais de *outdoor*, (b) NYCity 2 com *outdoors* inseridos sem retoque, (c) NYCity 2 com *outdoor* inserido com retoque



Figura 4.32: (a) imagem original do *outdoor*, (b) Santa Teresa com *outdoor* inserido

4.2.6 Testes com correção de distorção perspectiva

Mostramos as imagens originais seguidas da imagem gerada após a transformação de *warping* que corrige a distorção perspectiva apresentada na seção 3.2. A correção fica bem nítida nas Figuras 4.34 e 4.35, onde os postes ficam visivelmente verticais. Vale ressaltar que em alguns exemplos fica claro que o método não gera um resultado realista pois não corresponde a uma câmera real e pode ser visto como um novo tipo de efeito visual.

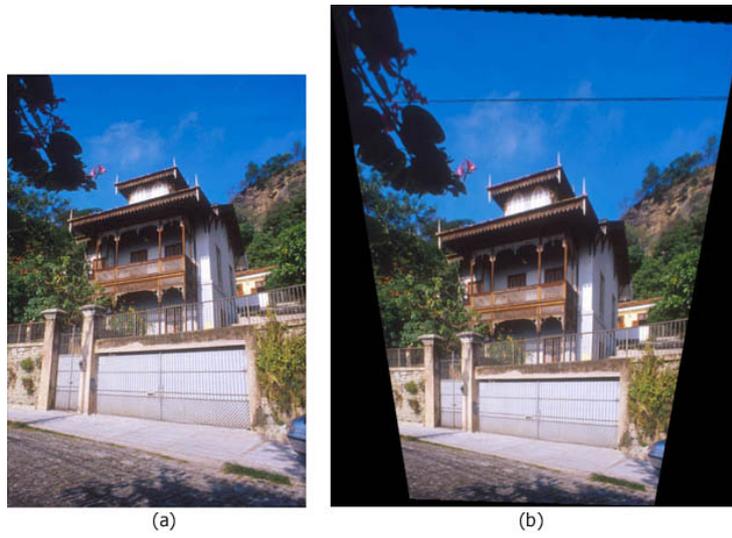


Figura 4.33: (a) Santa Teresa 1 original, (b) Santa Teresa 1 com remoção de perspectiva



Figura 4.34: (a) Santa Teresa 8, (b) Santa Teresa 8 com remoção de perspectiva



Figura 4.35: (a) Rio Design Barra 2 original, (b) Rio Design Barra 2 com remoção de perspectiva

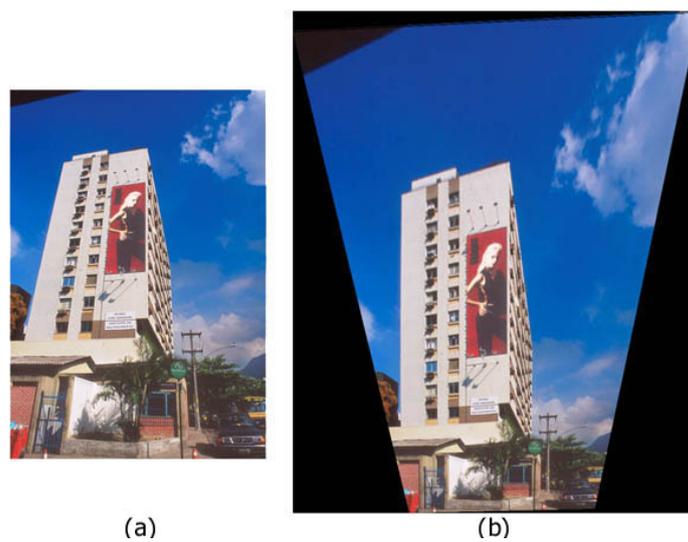


Figura 4.36: (a) Humaitá 3 original, (b) Humaitá 3 com remoção de perspectiva

4.2.7 Testes com recuperação de textura

Além da obtenção de textura com correção de perspectiva, podemos avaliar, em fotografias onde já havia *outdoors* reais, como eles são corretamente retirados da imagem original.

As Figuras 4.37 e 4.38 mostram as texturas com perspectiva corrigida e as áreas na imagem de onde foram extraídas. Em seguida, as Figuras 4.39 e 4.40 são exemplos de fotografias cujas edificações já possuíam *outdoor* reais no local e ilustram como eles são originalmente. Por último, a Figura 4.41 mostra a fachada da casa vista de frente para uma avaliação das proporções de seus elementos.

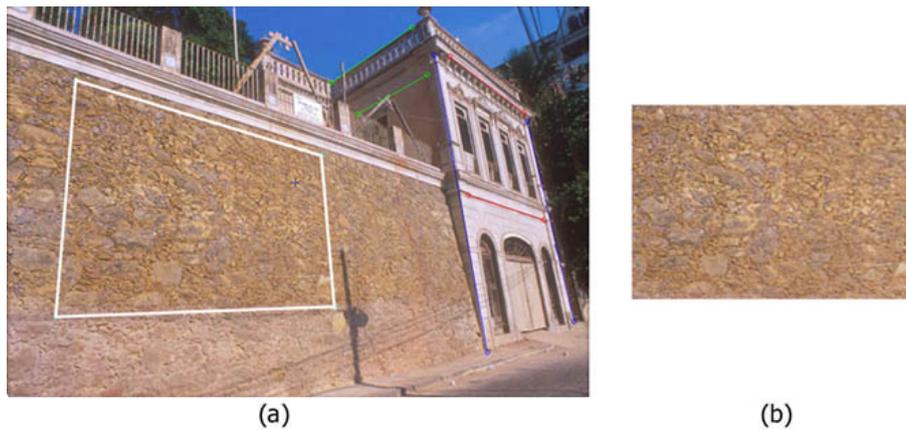


Figura 4.37: (a) Santa Teresa 6 original, (b) imagem da textura corrigida



Figura 4.38: (a) fotografia original obtida no site [P3D99], (b) imagem da textura corrigida



Figura 4.39: (a) Humaitá 2 original, (b) *outdoor* real retirado da imagem Humaitá 2



Figura 4.40: (a) Humaitá 4 original, (b) *outdoor* real retirado da imagem Humaitá 4



Figura 4.41: (a) Santa Teresa 8 original, (b) proporções da fachada

Capítulo 5

Conclusões e trabalhos futuros

Este trabalho apresentou um sistema para a edição tridimensional de fotografias arquitetônicas que serve para ampliar as capacidades dos editores utilizados para a composição de imagens. Ele fornece ferramentas para que o usuário possa editar de forma geometricamente correta, e não apenas a *sentimento*, planos na fotografia de origem, inserindo neles novos objetos com a perspectiva apropriada, extraia uma textura da fotografia de entrada com correção automática de perspectiva e simule uma lente fotográfica capaz de corrigir a distorção angular provocada pela projeção perspectiva – o que nem sempre é possível com uma lente fotográfica real.

O método é de uso simples pelos profissionais a quem se destina, por requerer informações intuitivas e fáceis de serem identificadas nas cenas de interesse e não requer computadores especiais. Ele dispensa informações prévias sobre as cenas, podendo, portanto, ser utilizado em quaisquer fotografias (inclusive históricas) para extrair informações ou inserir novos elementos virtuais. O sistema visa atender a profissionais das áreas de arquitetura, fotografia, publicidade e artes visuais, quando suas necessidades envolvem alteração de cenários reais. Esses profissionais podem modificar a imagem para se expressarem artisticamente ou para avaliarem suas criações.

Uma limitação deste método é que ele assume que a região editada está contida em um único plano paralelo a dois eixos principais. Quando esta hipótese não é satisfeita, a imagem extraída contém elementos que não se apresentam em suas verdadeiras proporções. A Figura 5.1 ilustra este fato. Ao extrair da foto a região correspondente à área assinalada, os elementos da sacada também são extraídos e aparecem deformados (na verdade, tudo se passa como se a sacada estivesse desenhada sobre a parede da casa, em vez de ser um elemento tridimensional). Para atacar este problema, é necessário estender a superfície de edição, passando a considerar superfícies formadas por retângulos paralelos aos eixos principais.

Uma outra extensão natural consiste em considerar planos paralelos a apenas um dos eixos. Este é o caso, por exemplo, de telhados, rampas ou paredes não alinhados com as direções principais.

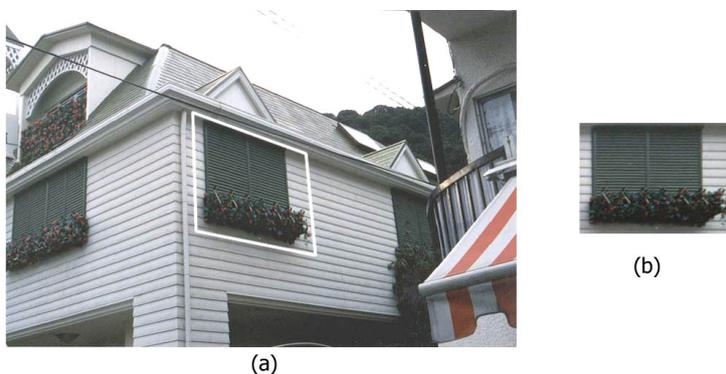


Figura 5.1: (a) Imagem original, (b) imagem transformada, considerando que todos os elementos estão no mesmo plano

O sistema apresentado pode ser utilizado como uma ferramenta de auxílio para estudos onde avalia-se a efetividade de *outdoors*, determinando se cumprem sua missão de veicular apropriadamente informações de interesse social ou publicitário. Outra extensão possível do sistema seria aplicá-lo a um estudo de como a distribuição de *outdoors* pela cidade afetaria a sua harmonia, provocando ou não uma poluição visual urbana.

Também como trabalhos futuros sugerimos que seja introduzido ao sistema o procedimento automático ou semi-automático para obtenção das linhas guias, através do reconhecimento de linhas retas na imagem ([Sze00]) e que essas linhas sejam extraídas com refinamento a nível de sub-pixel. Além disso, sugerimos disponibilizá-lo como um *toolkit* e transformá-lo em um *plug-in* para ser usado em editores de imagem (por exemplo o Adobe Photoshop [PSD]).

Glossário

***Frustum* de visão**

é o volume (sólido) de visualização determinado pela abertura da lente e a profundidade de foco

Ângulo de *pitch*

ângulo de declividade

Ângulo de *roll*

ângulo de rotação longitudinal

Baseline é a linha que une os centros óticos de duas câmeras em estéreo

Centro projetivo

é o ponto onde o eixo ótico do sistema de lentes cruza o plano da imagem.

Constante de distorção ótica

representa a distorção causada pela imperfeição da forma geométrica das lentes da câmera.

Distância focal

é a distância entre o o ponto nodal interior da câmera e o plano da imagem.

Eixo ótico

é a linha que atravessa o centro ótico e intercepta perpendicularmente o plano da imagem.

Fator de escala em x

é oriundo da diferença de elementos sensores na câmera e o número de *pixels* amostrados pelo *hardware* de digitalização da imagem. É o fator de conversão

da unidade-*pixel* na imagem armazenada no *frame-buffer* para a unidade do sensor da câmara, ou seja, uma escala de valores armazenados para valores reais.

Referências Bibliográficas

- [Sze00] Szenberg, F.; Carvalho,P.C. e Gattass,M. Automatic Camera Calibration for Image Sequences of a Football Match. *Interactive Conference on Pattern Recognition*. Rio de Janeiro, RJ, Brasil, Março de 2000.
- [Car98] Szenberg, F.; Carvalho,P.C. e Gattass,M. Image-based Modeling Using a Two-step Camera Calibration Method. In *Proceedings of International Symposium on Computer Graphics, Image Processing and Vision*, 1998, pp 388–395.
- [Car99] Gomes J.; Velho, L.; Carvalho, P. C. e Figueiredo, L.H. *22o. Colóquio Brasileiro de Matemática - Métodos de Otimização Em Computação Gráfica*. Instituto de Matemática Pura e Aplicada - IMPA, Estrada Dona Castorina, 110 - Jardim Botânico cep: 222460-320 - Rio de Janeiro, RJ, Brasil, 1999.
- [Che95] Shenchang Eric Chen. QuickTime VR - An Image-Based Approach to Virtual Environment navigation. In Robert Cook, editor, *Conference Proceedings SIGGRAPH 95*, Annual Conference, pages 29–38, Annual Conference Series, ACM SIGGRAPH, Addison Wesley, Los Angeles, California, EUA, Agosto de 95.
- [Deba] Taylor,C.; Debevec, P. e Malik, J. Modeling and Rendering Architecture from Photographs: A Hybrid Geometry-and Image-Based Approach. In Holly Rushmeier, editor, *Conference Proceedings (SIGGRAPH ' 96)*, pages 11–20, Annual Conference Series, ACM SIGGRAPH, Addison Wesley, Nova Orleans, Louisiana, EUA Agosto de 1996, pp 04-09

- [dTGTa] Grupo de Tecnologia Gráfica TecGraf. Canvas Draw CD. <http://www.tecgraf.puc-rio.br/cd>.
- [dTGTb] Grupo de Tecnologia Gráfica TecGraf. IM. <http://www.tecgraf.puc-rio.br/im>.
- [Fab91] Annateresa Fabris. *FOTOGRAFIA: Usos e Funções No Século XIX*. Edusp - Editora da Universidade de São Paulo, Av. Prof. Luciano Gualberto, Travessa J., no.374/6o.andar - Ed. da Antiga Reitoria - Cidade Universitária - São Paulo, 1991.
- [Fau93] Olivier Faugeras. *Three-Dimensional Computer Vision: A Geometric View-Point*. MIT Press, Cambridge, Massachusetts, EUA, 1993.
- [Fau98] Laveau S.; Csurka, G.; Zeller, C.; Gauclin, C.; Faugeras,O.; Robert,L. e Zogh-lami,I. 3D Reconstruction of Urban Scenes from Image Sequences. *Computer Vision and Image Understanding*, Março de 1998, Vol. 69 No.3(Article No. IV970665.):pp 292–309.
- [Fei97] Van Dam, A.; Feiner, S; Foley, J. e Hughes, J. *Computer Graphics: Principles and Practice Second Edition in*. Addison-Wesley Publishing Company, Inc., EUA, 1997.
- [Hor97] Anjyo,K.; Horry, Y. e Arai,K. Tour into the Picture: Using a Spidery Mesh Interface to make Animation from a Single Image. In *Conference Proceedings SIGGRAPH'97*, 1997, pp 225–232.
- [Lev96] Levy,C; Gattass,M.; Lucena,C.; Figueiredo,L.H. e Cowan,D. IUP : A Portable User Interface Development Tool. *Software: Practice e Experience* 267, pp 737-762 <http://www.tecgraf.puc-rio.br/iup/>, 1996.
- [McM] McMillan, L. e Bishop,G. Plenoptic Modeling: An Image-Based Rendering System. In Robert Cook, editor, *Conference Proceedings (SIGGRAPH '95)*. ACM SIGGRAPH, Addison Wesley, Annual Conference Series, Los Angeles, California, EUA, 06-11 Agosto de 1995, pp 39–46.

- [Nik01] Lense Nikon 28mm f/3.5 PC-Nikkor Perspective Control Technology. <http://www.nikon.com>, 2001.
- [P3D99] Photo 3D. <http://www.photo3D.com>, 1999.
- [PSD] www.adobe.com.
- [Tom] Tommaselli, A. e Tozzi, C.L. Tutorial: Calibração de Câmeras usando Feições Geométricas. In *IV Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens*. SIBGRAPI'91, Centro de Convenções Rebouças, São Paulo, 14-17 Julho de 1991.
- [Tsa] Roger, Y. Tsai. An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Miami Beach, Flórida, EUA, 22-26 Junho de 1986.
- [Wol90] George Wolberg. *Digital Image Warping*. IEEE Computer Society Press, 10662 Los Vaqueros Circle P.O. box 3014, Los Alamitos, CA, EUA, 90720-1264, 1990.