

UNIVERSIDADE ESTADUAL DE SANTA CRUZ
DEPARTAMENTO DE CIÊNCIAS EXATAS E TECNOLÓGICAS
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

Ivan Cezanne Ismerim Bezerra de Meneses Seara

Reconstrução de Objetos 3D a partir de Imagens 2D

Ilhéus– Bahia
2014

Ivan Cezanne Ismerim Bezerra de Meneses Seara

Reconstrução de Objetos 3D a partir de Imagens 2D

Monografia apresentada ao Curso de Bacharelado em Ciência da Computação da UESC, como requisito para a obtenção parcial do grau de BACHAREL em Ciência da Computação.

Orientador: César Alberto Bravo Pariente

Doutor

Ilhéus– Bahia

2014

Ivan Cezanne Ismerim Bezerra de Meneses Seara

Reconstrução de Objetos 3D a partir de Imagens 2D

Monografia apresentada ao Curso de Bacharelado em Ciência da Computação da UESC, como requisito para a obtenção parcial do grau de BACHAREL em Ciência da Computação.

Ilhéus, 27/11/2014.

BANCA EXAMINADORA

César Alberto Bravo Pariente - Doutor
UESC/DCET
(Professor Orientador)

Vânia Cordeiro da Silva - Doutora
UESC/DCET

Marcelo Ossamu Honda - Doutor
UESC/DCET

À família, aos amigos e aos mestres, dedico.

Agradecimentos

Agradeço à UESC, DCET, COLCIC e seus funcionários, magníficos reitores, diretores e coordenadores pelo empenho e sucesso no funcionamento e provimento do curso.

Agradeço ao NBCGIB pelo apoio infra-estrutural durante a execução deste trabalho.

Agradeço, imensamente, aos professores do curso de ciência da computação, contribuidores de conhecimento e experiência da maior valia.

Agradeço ao meu orientador, pelo apoio, esclarecimentos e coordenação; sem os quais este trabalho não se tornaria o que veio a ser.

Agradeço, ainda, aos amigos da TecnoJr e CACIC pelo trabalho, esforço e sucessos compartilhados durante estes anos de graduação.

Por fim, manifesto minha eterna gratidão à minha família - base primordial do meu ser, origem de toda a força, motivação e inspiração, tão fundamentais nesta jornada que agora se encerra.

Resumo

O processo de gravação de imagens fotográficas consiste em projetar regiões visíveis de uma cena em um plano bidimensional, desta forma consegue-se registrar um ambiente tridimensional em um outro bidimensional. Na bibliografia sobre computação gráfica encontram-se vários métodos desta captura, de forma digital. Neste trabalho propõe-se uma técnica capaz de criar uma malha poligonal tridimensional que represente completamente um objeto ora apresentado por uma imagem bidimensional. Utilizando imagens de apelo histórico à computação gráfica e à arte, apresenta-se uma forma de identificação de um sistema de coordenadas de referência já existente e particular a cada objeto representado na imagens, de forma que seja possível o estudo analítico de sua constituição espacial. Uma vez bem referenciados, pode-se obter uma transformação linear ou vetorial que implemente uma correspondência entre pontos 2D de uma imagem com pontos 3D de um sistema de referências. Dada a particularidade das imagens estudadas, é possível inferir padrões de simetria e proporcionalidade nos objetos retratados, permitindo a representação, também, de regiões do objeto oclusas nas imagens.

Palavras-chaves: Processamento Digital de Imagens, Computação Gráfica, Reconstrução 3D

“Há um ditado que ensina ‘o gênio é uma grande paciência’; sem pretender ser gênio, teimei em ser um grande paciente. As invenções são, sobretudo, o resultado de um trabalho teimoso, em que não deve haver lugar para o esmorecimento”.

Alberto Santos Dumont

Sumário

Lista de Tabelas	x
Lista de Figuras	xi
1 Introdução	1
1.1 Justificativa	1
1.2 Problemática	2
1.3 Objetivos	3
2 Estado da Arte	4
2.1 Múltiplas Visões e Geometria Epipolar	4
2.1.1 Cinema 3D através de Estereoscopia	5
2.1.2 Um Abordagem Mais Simples da Estereoscopia	5
2.2 Calibração de Câmera e Geometria Projetiva	6
3 Materiais e Métodos	7
3.1 Recursos Computacionais	7
3.2 Conjunto de Dados	8
3.3 Metodologia Lean	10
3.4 Aferição da corretude	10
4 Fundamentação Teórica	13
4.1 Projeções	13
4.1.1 Projeção Paralela	13
4.1.2 Projeção Cônica	14

4.2	Câmera Virtual	15
4.3	Calibração de câmera	16
5	A Técnica de Reconstrução	19
6	Os Softwares	23
6.1	Aplicativo de detecção de retas	23
6.2	Aplicativo de visualização 3D	25
6.3	Aplicativo de Reconstrução	25
7	Resultados, Discussões e Trabalhos Futuros	26
	Referências Bibliográficas	27

Lista de Tabelas

Lista de Figuras

2.1	Triangulação utilizada na geometria epipolar	5
3.1	O <i>VW Beetle</i> de Sutherland	9
3.2	O Bule de Leite de Newell	9
3.3	A litografia de Dürer (BRUNELLESCHI'S..., 1998)	10
3.4	Cenas para validação da técnica de reconstrução	11
4.1	Projeção paralela em (POTTMANN et al., 2007)	14
4.2	Projeção cônica em (POTTMANN et al., 2007)	15
5.1	Os pontos P_i pertencem à reta de projeção de P , mas suas coordenadas Z são diferentes da coordenada do ponto P	20
5.2	É perceptível que os pontos demarcados com um 'X' têm a mesma coorde- nada Z no espaço	21
6.1	O programa de regressão de rasterização com região de interesse e coefici- entes ressaltados	24
6.2	Reta analítica correspondente à linha da região demarcada na figura 6.1 . .	24

1 Introdução

1.1 Justificativa

O processo fotográfico, surgido no século XIX, determinou um marco na forma de registrar a realidade humana, sendo capaz de capturar a luz proveniente do ambiente e convertê-la em regiões semelhantes, em cores e formas, em uma superfície manuseável chamada de fotografia ou imagem fotográfica. Tal processo tem sido motivo de estudo de diversas áreas do conhecimento e hoje já se é possível executar tal processo, outrora baseado em películas fotossensíveis, de forma digital e computacional.

Uma das áreas de estudo das imagens artificiais é o processamento digital de imagens, campo da ciência da computação que, desde a década de 1960, apresenta discussões relevantes acerca da manipulação de imagens em um meio computacional, vide (ROSENFELD, 1969). Derivado disto, hoje é possível que, através de um computador, se interprete e crie uma imagem. Dentre as demais áreas da computação, o processamento digital de imagens se atém à manipulação de imagens digitais com o intuito de fazê-la apresentar alguma característica baseada em sua configuração anterior, como em (SORKINE-HORNUNG et al., 2012).

Fato é que, o processo de formação de uma imagem digital funciona baseado em um processo de projeção de regiões visíveis de um ambiente tridimensional - universo real - em uma superfície plana bidimensional - imagem em si. Facilmente se percebe que tal técnica consiste em reduzir em uma dimensão o meio de representação de objetos de uma cena. Apesar de avanços tecnológicos neste processo permitirem a reprodução fiel das propriedades da luz, facilitando a noção do observador da imagem acerca dos aspectos físicos tridimensionais(dimensões, distâncias etc.) da cena retratada, os dados que representam tais aspectos foram truncados. Tais dados tridimensionais truncados podem apresentar utilidade para áreas da sociedade que têm nas imagens armazenadas sua fonte de conteúdo. Áreas como conservação e restauro de patrimônio arquitetônico; serviços de GPS e mapas e investigação de cenários reais trabalham constantemente com representações bidimensionais de objetos tridimensionais (edifícios, avenidas urbanas, cômodos

de residências mobiliados etc.) onde o sucesso de suas tarefas depende da compreensão espacial do seu objeto de estudo.

De tal forma, a existência de um processo digital reverso ao processo da projeção, capaz de reconstruir objetos 3D a partir de imagens 2D, representa um avanço significativo na capacidade de se representar a realidade em forma digital.

1.2 Problemática

O *VW Beetle* de Sutherland entrou para a história da computação gráfica por ser o primeiro objeto real digitalizado em três dimensões. Em um processo liderado pelo professor Ivan Edward Sutherland, um grupo de alunos demarcou manualmente, com giz, sobre a carroceria de um veículo Volkswagen Beetle 1967, pontos de interesse para a construção do modelo 3D do carro (MUSEUM, 2014a).

Tendo a experiência sido bem sucedida, um modelo tridimensional foi obtido. No entanto a manutenção dos dados deste modelo é nebulosa, não se sabe ao certo onde obter uma versão correta deste modelo ao passo que já se questiona se os seus dados originais estão preservados para fins de acesso público.

No entanto, a própria Universidade de Utah - sede do experimento - mantém as imagens deste modelo, sob a curadoria do *Computer History Museum*, além de haverem registros individuais com parâmetros úteis a este trabalho (vide capítulo 3). Assim, há a possibilidade de se restaurar os dados tridimensionais deste modelo 3D digital icônico.

Em um caso similar, o Bule de Leite de Newell é, originalmente, parte integrante de um famoso modelo tridimensional chamado *Teaset* de autoria do pesquisador Martin Edward Newell. O modelo *Teaset* pode ser encontrado facilmente em repositórios na *internet*, no entanto, em nenhuma das fontes consta o bule de leite. Sabe-se que tal objeto era parte integrante do conjunto *Teaset* por abordagens e imagens de trabalhos do próprio autor, (NEWELL, 1975).

Por tal circunstância, percebe-se que o Bule de Leite de Newell também se trata de um elemento de apelo histórico do campo de estudo deste trabalho cujo estudo é interessante e não tem seus dados tridimensionais publicados.

Em última instância, há uma litografia datada do século XVII da autoria do

artista Albrecht Dürer que representa uma cena onde os personagens exemplificam a implementação da projeção perspectiva (vide seção 4.1 sobre as projeções). Tal cena, por ser uma documentação pioneira do estudo da perspectiva, apresenta uma oportunidade de trabalhar sob dados antigos e fornecer uma previsão de como ampliar este trabalho de conclusão de curso para a reconstrução 3D de cenários com mais de um objeto.

1.3 Objetivos

Em face da problemática disposta, este trabalho visa, em âmbito específico, obter uma malha poligonal tridimensional que represente o *VW Beetle* de Sutherland, o Bule de Leite de Newell e alguns elementos da cena de Dürer, de forma que seja possível a visualização destes objetos em um software de visualização 3D.

Para a obtenção de sucesso no objetivo específico, os seguintes objetivos gerais foram estabelecidos:

- definir um método de identificação de sistemas de coordenadas de referências nas imagens estudadas;
- inferir informações de padronização deste objetos, tais como simetrias e repetições;
e
- estabelecer uma correspondência entre pontos das imagens com pontos tridimensionais.

2 Estado da Arte

Com o uso de *hardwares* elaborados (câmeras sincronizadas ou câmeras de percepção volumétrica), definição estrita de objeto de reconstrução e distintas vertentes geométricas, a reconstrução 3D já é uma prática difundida, sendo possível encontrar um grande número de artigos acadêmicos sobre o assunto, além de trabalhos de empresas de entretenimento famosas que implementam reconstrução 3D, cada um de uma forma particular.

Neste capítulo é exposto uma breve coleção de exemplos representativos desta difusão da reconstrução a fim de configurar o cenário real onde este trabalho de conclusão de curso se encaixa.

2.1 Múltiplas Visões e Geometria Epipolar

Uma abordagem muito comum para a implementação de um sistema de reconstrução 3D é o uso de diferentes ponto de vista sobre um mesmo objeto. A aquisição de dados de um objeto sob diferentes pontos de vista pode ser feita através de sistemas com múltiplos dispositivos de captura (câmeras) devidamente sincronizados e calibrados ou com imagens já obtidas desde que seja conhecido um certo número de dados tridimensionais e dados de correspondência. Câmeras utilizadas para obtenção de múltiplas visões de um objeto são conhecidas como câmeras estéreo e tal técnica baseada em múltiplas visões, de estereoscopia.

Observando relações de correspondência entre pontos de distintas imagens, o uso da geometria epipolar fornece meios de concluir uma localização tridimensional a partir de dados bidimensionais. De forma sucinta, a geometria epipolar utiliza a técnica de triangulação entre os centros de projeção - que são pontos tridimensionais - de duas câmeras e o ponto de interseção - que também é tridimensional e pode ser chamado de epipolo - entre as retas formadas pelos centros de projeção das câmeras e o ponto da imagem em análise (AGUIAR, 2010, p.11).

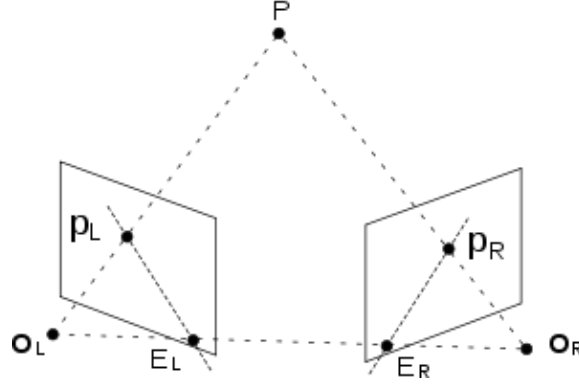


Figura 2.1: Triangulação utilizada na geometria epipolar

2.1.1 Cinema 3D através de Estereoscopia

Em (HEINZLE et al., 2011) vê-se a sugestão de um complexo sistema motorizado de câmeras de vídeo com um ajuste automático de parâmetros de calibração para maior precisão na determinação de características tridimensionais obtidas através do sistema de câmeras estéreo proposto.

Esta abordagem é efetiva para os autores dada a sua finalidade cinematográfica; e já que filmes em 3D são cada vez mais comuns nas salas de cinema em todo mundo, o investimento feito para obter a precisão que este trabalho alcança é facilmente compensado.

Em um contexto mais amplo, que não suporta uma infra-estrutura complexa de uma filmagem dos estúdios de cinema internacionais ou demanda um dinamismo maior, o uso de um sistema motorizado de câmeras em sincronia é uma estratégia pouco eficiente, apesar da sua grande eficácia.

2.1.2 Um Abordagem Mais Simples da Estereoscopia

(NEVES; TSZUKI, 2010) utiliza uma abordagem mais enxuta da reconstrução 3D através de estereoscopia, sob a premissa de que as dimensões do objeto a ser reconstruído são conhecidas. Tal premissa, mesmo válida para o trabalho citado, reduz drasticamente o escopo da técnica sugerida.

Usando um único par de imagens estereoscópicas estáticas e as dimensões do objeto estudado ainda é preciso efetuar alguns processos de normalização a ajuste algébrico para que condições de convergência, como o epipolo, possam ser obtidas. Tais

empecilhos, inclusive, são decorrentes de uma análise do objeto em termos de coordenadas absolutas, o que é irrelevante para a análise de um objeto isolado cuja reconstrução não faz parte de um cenário igualmente importante.

Assim, a reconstrução 3D de (NEVES; TSZUKI, 2010) ainda não se adapta completamente ao objetivo deste trabalho.

2.2 Calibração de Câmera e Geometria Projetiva

Um conceito conhecido na computação gráfica é o de câmera virtual - um modelo matemático computacional que representa o funcionamento de um dispositivo de captura (GOMES; VELHO, 2008a). A técnica de calibração de câmera consiste em estimar os parâmetros de uma câmera virtual capaz de produzir uma imagem. Uma vez que imagens são produzidas através de projeções, com o uso da geometria projetiva é possível haver dedução, sob certas hipóteses, de dados 3D após a análise de uma única imagem 2D.

Em (LEE; FENG; GOOCH, 2008), após um processo de obtenção dos parâmetros de uma câmera, é possível transformar polígonos pertencentes ao plano da imagem em polígonos equivalentes pertencentes a planos tridimensionais distintos do plano da imagem. Infortunadamente, esta referência não fora conhecida tão cedo quanto desejado, impedindo que este trabalho pudesse ter maior embasamento e robustez baseado em tal referência. No entanto, trabalhos parecidos sobre calibração de câmera para reconstrução 3D, que, estes sim, serviram como embasamento inicial, estão presentes em (CARVALHO; SZENBERG; GATTASS, 1998) e (SZENBERG, 2001)

3 Materiais e Métodos

Para a concretização deste trabalho, são necessários materiais tanto de ordem computacional, para processamento e implementação de software, tanto de ordem de dados, principalmente para teste das hipóteses encontradas, verificação de sucesso e embasamento teórico.

Os materiais em forma de dados devem consistir, basicamente, em imagens de objetos isolados e, de preferência, com elementos dispensáveis como plano de fundo e texturas padronizados e sistema ortogonal de coordenadas de referência, a fim de facilitar o processamento dos dados pelo computador e o funcionamento dos algoritmos desenvolvidos; foi conseguido um número suficiente, para o início do trabalho, de imagens que correspondiam a este padrão desejado através de repositórios *on-line*.

A demanda por recursos computacionais de alta eficiência é baixa, portanto, pode-se, utilizar o computador pessoal do autor sem prejuízo de tempo; além disto, serão utilizadas ferramentas computacionais de terceiros, como bibliotecas de linguagens de programação e ambientes de edição de código-fonte, sob licença acadêmica gratuita, portanto, não acarretando em custos financeiros para a execução do trabalho.

Tendo em vista o escopo bem definido e a um conjunto de materiais de aplicação estrito deste trabalho, foi buscado um método de desenvolvimento adequado à esta produção que, por não visar a solução de um problema de grande amplitude no universo da computação gráfica, pode ser dito como enxuto. A escolha considerada apropriada para tal cenário de trabalho é a metodologia Lean, também chamada de metodologia Enxuta.

3.1 Recursos Computacionais

Não havendo grande volume de dados ou alta complexidade no processamento dos dados, será utilizado o computador pessoal textitNotebook Mobile Sim+ 6220 do autor do trabalho, com a seguinte configuração:

- processador Intel®Core®i3 330M (2,13 GHz, 3MB Cache);
- memória RAM Kingston®PC3-8500 (4MB, DDR3);
- chipset gráfico Intel®HM55;
- sistema operacional Windows 7®.

Para fins de construção de programas de computador e otimização do rendimento dos códigos-fontes em relação ao *hardware* disponível, serão usadas as seguintes ferramentas computacionais:

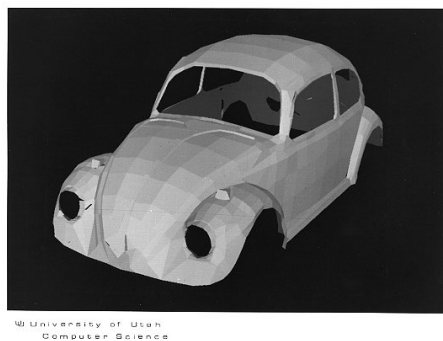
- Ambiente gráfico para programação em linguagem C# Microsoft Visual Studio 2010 v4.5.50938 RTMRel®;
- Biblioteca de métodos algébricos em C# ALGLIB Free Edition v3.8.2
- Software de visualização 3D Meshlab v1.3.3.

3.2 Conjunto de Dados

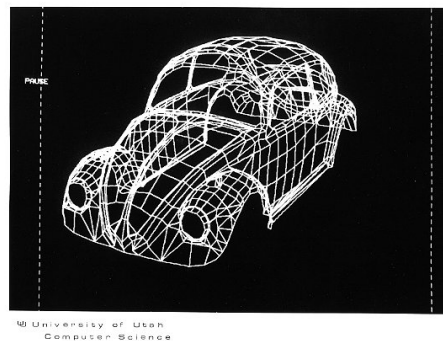
Os dados utilizados como entrada nos algoritmos e programas desenvolvidos neste trabalho consistem em imagens digitais representando objetos apresentados no capítulo 1, o *VW Beetle* de Sutherland (figura 3.1), o Bule de Leite de Newell (figura 3.2) e uma litografia digitalizada de Dürer (figura 3.3). Todas as imagens utilizadas neste trabalho foram obtidas respeitando os direitos autorais e de reprodução.

Quanto ao *VW Beetle* de Sutherland, conseguiu-se obter uma imagem com sistemas ortogonais de coordenadas de referência graficamente evidentes (figura 3.1c), no entanto todas as imagens apresentavam-se em projeção em perspectiva, distorcendo as linhas do sistema de referência. Todas as imagens referentes a este objeto apresentavam regiões desprezíveis, como plano de fundo, facilmente diferenciáveis através de cor. Em formato ‘JPG’, as imagens possuíam resolução de 72 dpi e 24 bits para intensidade de cor.

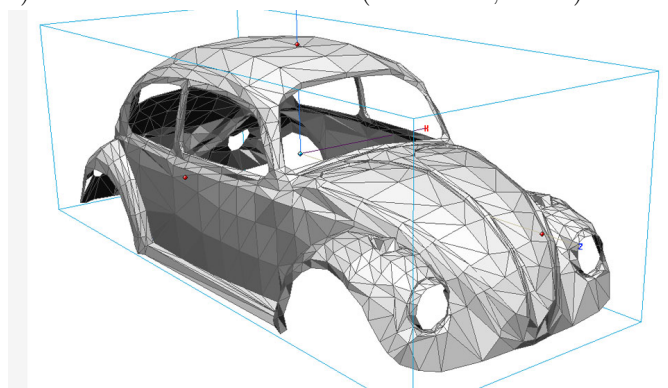
As imagens do Bule de Leite de Newell, por este ser um objeto sem registro oficial dos seus dados, são menos elaboradas, com o sistema de coordenadas de referência mais nebuloso, mas, afortunadamente, apresentando projeção paralela isométrica. Tais



(a) *VW Beetle* de Sutherland renderizado (MUSEUM, 2014b)



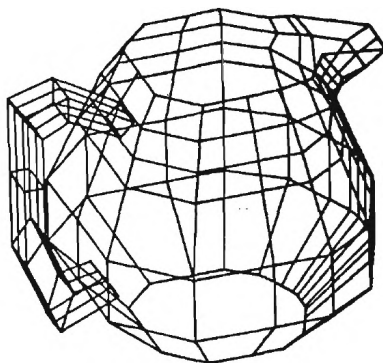
(b) *VW Beetle* de Sutherland em *wireframe* (MUSEUM, 2014c)



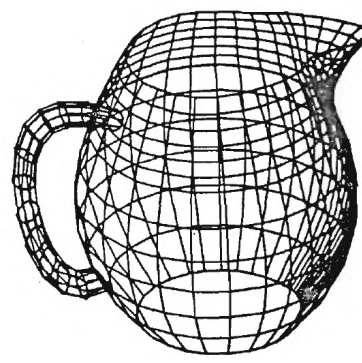
(c) *VW Beetle* de Sutherland com referências (TORCHINSKY, 2013)

Figura 3.1: O *VW Beetle* de Sutherland

imagens são em branco e preto, em formato ‘PNG’ com 32 bits para intensidade de cor e resolução de 95 dpi.



(a) Bule de Leite de Newell convencional (NEWELL, 1975)



(b) Bule de Leite de Newell suavizado (NEWELL, 1975)

Figura 3.2: O Bule de Leite de Newell

A litografia de Dürer, digitalizada e disponibilizada por terceiros, possui um sistema de referências evidente e, como se pode prever, está representada através da projeção em perspectiva. Em formato ‘JPG’ com 96 dpi de resolução e 24 bits de profundidade de cor.

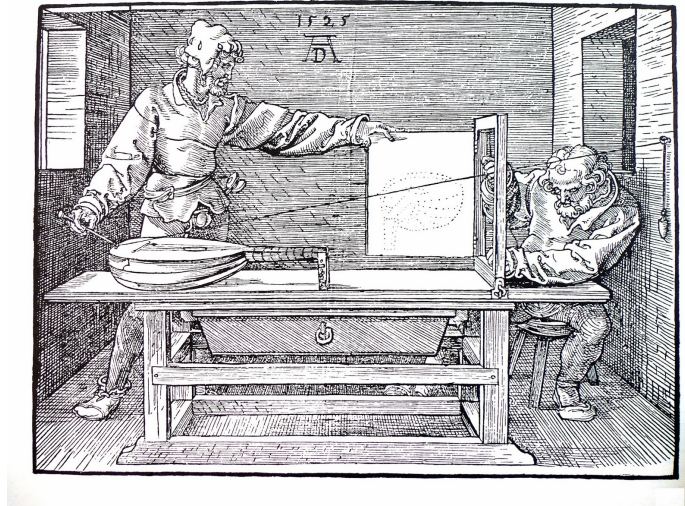


Figura 3.3: A litografia de Dürer (BRUNELLESCHI'S..., 1998)

3.3 Metodologia Lean

A metodologia Lean, surgida no ambiente industrial japonês pós II Guerra Mundial, prega a supressão de fluxos de produção, ou de trabalho, que não contribuem para um melhor resultado. Baseada em sete princípios (PEREIRA; SCHOEFFEL, 2012), a metodologia Lean se fez útil no desenvolvimento deste trabalho através do princípio da "Eliminação de Desperdícios", "Entregar Rápido" e "Visualizar o Todo", que implicaram em um processo eficiente de efetuação das tarefas.

Sendo para fins acadêmicos e não passível de adaptações para públicos-alvo, os resultados deste trabalho se adequaram à filosofia enxuta e focada em qualidade final.

Em adendo ao produto bem definido, o escopo do trabalho, por ser reduzido a um conjunto específico de objetos de estudo, indica que o desenvolvimento, para que fosse bem conduzido, bastava que se mantivesse focado estritamente em seu objetivo, sem desvios ou atualizações.

3.4 Aferição da corretude

Como o objetivo deste trabalho é reconstruir objetos cujas informações tridimensionais são inexistentes, é impossível se fazer uma validação direta dos resultados obtidos, isto é, não há como comparar os modelos tridimensionais obtidos neste trabalho com um modelo paramétrico dos mesmos objetos estudados.

De tal forma, a corretude dos modelos obtidos será obtida de forma implícita,

através da validação da técnica de reconstrução deste trabalho. Utilizando cenas específicas confeccionadas pelo autor (figuras 3.4), nas quais os elementos a serem reconstruídos têm suas informações catalogadas, será estimado um erro geométrico E (fórmula 3.1) entre os pontos de coordenadas conhecidos dos modelos (P) e os pontos de coordenadas obtidos pela técnica desenvolvida neste trabalho (P').

$$E = \sqrt{(x_{P_i} - x_{P'_i})^2 + (y_{P_i} - y_{P'_i})^2 + (z_{P_i} - z_{P'_i})^2} \quad (3.1)$$

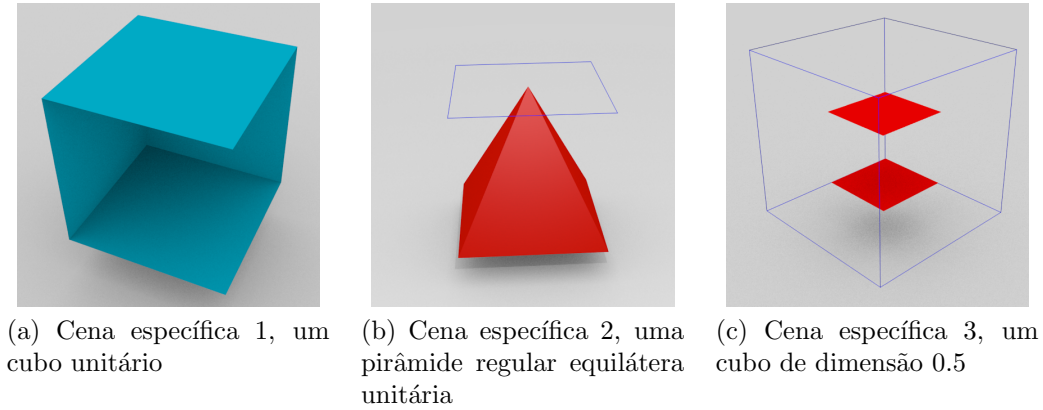


Figura 3.4: Cenas para validação da técnica de reconstrução

Com a cena de validação 1 (figura 3.4a) deseja-se a reconstrução dos vértices de um hexaedro de dimensão unitária; com a cena de validação 2 (figura 3.4b) almejam-se os vértices de uma pirâmide regular equilátera de dimensão unitária - cujas arestas da base têm mesmo tamanho que a altura do sólido; por fim, com a cena de validação 3, buscam-se os vértices de um cubo de dimensão 0.5.

Nas cenas de validação há um plano abaixo do objeto descrito anteriormente, este não será reconstruído, servindo apenas como aparato visual na imagem da cena. Na cena de validação 1 (figura 3.4a), um cubo (hexaedro) teve duas de suas faces removidas, a fim facilitar a demarcação de pontos de referência, originando a figura vista. No caso da figura 3.4b, há, ainda, um quadrado de cor azul paralelo ao plano da base da pirâmide; tal quadrado é um auxílio, inserido pelo autor, para a demarcação dos pontos de referência e também será desconsiderado como objeto em reconstrução. A figura 3.4c, por sua vez, apresenta as arestas (em azul) de um cubo, estas são um auxílio para a demarcação de referências; note que neste cena também foram removidas faces do cubo para facilitar demarcações.

Todas as imagens das cenas de validação têm resolução de 72 dpi com 32 bits

para intensidade de cor.

4 Fundamentação Teórica

Neste trabalho de conclusão de curso, são usados alguns conceitos e técnicas da computação gráfica de forma adaptada ao objetivo proposto. Nas próximas seções são explicados os conceitos utilizados e suas respectivas aplicações no trabalho desenvolvido.

4.1 Projeções

Inicialmente, projeções são representações de objetos sob as restrições de um determinado ambiente. Em geometria analítica usam-se projeções para a descrição de vetores (as componentes do vetor são as projeções deste em cada uma das bases do sistema ordenado); no cotidiano, sombras geradas pelo nosso corpo no chão, por exemplo, é um caso específico de projeção (representação) do nosso corpo (objeto) na superfície plano do chão (ambiente) (POTTMANN et al., 2007).

Em computação gráfica, as projeções servem para representar objetos tridimensionais em dispositivos de visualização bidimensionais. Dependendo da utilização da visualização, diferentes formas de projetar o objeto tridimensional podem ser usadas (POTTMANN et al., 2007); as duas principais são a projeção paralela e a projeção cônica.

Independente do tipo de projeção, com apenas uma imagem não se pode estabelecer uma relação unívoca entre pontos no plano de projeção e pontos no espaço (CARVALHO; VELHO, 2005). No entanto, relações de isometria (seção 4.1.1) ou proporcionalidade (seção 4.1.2) podem incutir restrições aos pontos no espaço de forma a reduzir tal multiplicidade.

4.1.1 Projeção Paralela

A projeção gráfica pode ser feita através de linhas paralelas entre si que ligam pontos do objeto ao plano de projeção (figura 4.1); esta projeção é chamada de projeção paralela (POTTMANN et al., 2007). Neste caso, são mantidas as proporções entre as dimensões do objeto estudado (isometria) (GOMES; VELHO, 2008a).

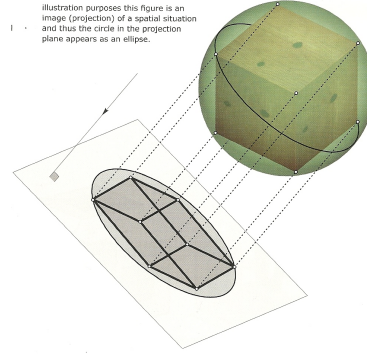


Figura 4.1: Projeção paralela em (POTTMANN et al., 2007)

A isometria é uma característica interessante deste tipo de projeção pois a distância planar entre dois pontos é igual à sua distância no espaço.

A transformação de um ponto 3D $P = (X_P, Y_P, Z_P)$ em um ponto 2D P' sob a projeção paralela é dada pela equação 4.1, em coordenadas baseadas no Sistema de Coordenadas da Câmera (CARVALHO; VELHO, 2005). Como pode-se notar, a projeção de P independe da sua distância - Z_P - ao plano de projeção e isto causa a multiplicidade projetiva.

$$P' = (X_P, Y_P) \quad (4.1)$$

4.1.2 Projeção Cônica

Com o intuito de representar um objeto de forma similar à realidade, é aplicada a projeção cônica ou perspectiva (POTTMANN et al., 2007). Tal projeção provoca uma deformação linear, proporcional à distância Z_P ao plano de projeção, na posição planar do ponto $P = (X_P, Y_P, Z_P)$, que também está expresso em coordenadas do Sistema de Coordenadas da Câmera (equação 4.2) (CARVALHO; VELHO, 2005). Isto se dá pois a projeção P' de P é a interseção da reta que une o ponto 3D e ao ponto P (figura 4.2), sendo e o centro de projeção, que funciona de forma análoga ao olho de um observador humano (POTTMANN et al., 2007).

$$P' = \left(\frac{X_P}{Z_P}, \frac{Y_P}{Z_P} \right) \quad (4.2)$$

A deformação provocada pela projeção cônica impossibilita a existência da

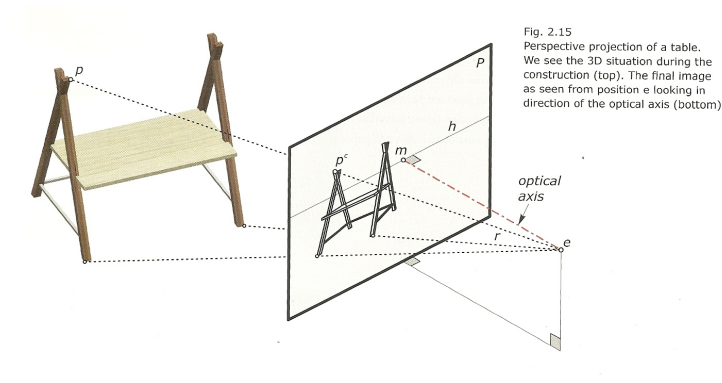


Figura 4.2: Projeção cônica em (POTTMANN et al., 2007)

isometria, no entanto, com o processo de calibração de câmera virtual (seções 4.2 e 4.3) é possível obter a distância entre dois pontos 3D pertencentes a um mesmo plano (CARVALHO; SZENBERG; GATTASS, 1998).

4.2 Câmera Virtual

A operação básica de uma câmera virtual é a projeção (GOMES; VELHO, 2008a). A fim de sintetizar uma imagem de um objeto digital com três dimensões, programas de visualização 3D implementam um modelo matemático que utiliza as técnicas discutidas anteriormente (seção 4.1) para a transformação de 3D para 2D.

Neste trabalho optou-se por desenvolver um modelo de câmera que projeta apenas em perspectiva dada a ausência, inicialmente, de bibliografia sobre modelos de câmeras de projeção paralela. No entanto, na seção 2.2 é exposta uma possibilidade de retificação desta limitação.

Para um modelo altamente fiel à câmera real, devem ser feitas correções acerca de distorções de lentes (AGUIAR, 2010) além de que uma câmera digital real, em termos geométricos, precisa converter as coordenadas de um objeto sob um sistema de referências global para um sistema de coordenadas em pixels passando por um sistema de coordenadas de câmera e de imagem, processo chamado de Transformações de Câmera (CARVALHO; VELHO, 2005). Para este trabalho, não há distorção de lentes nas imagens (uma vez que são artificiais).

Em (CARVALHO; VELHO, 2005) é descrita cada transformação entre cada sistema de coordenadas. No entanto, neste trabalho todas as transformações de câmera foram agrupadas, como em (CARVALHO; SZENBERG; GATTASS, 1998) e (GOMES;

VELHO, 2008a), uma vez que apenas interessa a transformação final e não os parâmetros explícitos da câmera. Assim, o modelo final de câmera C utilizado foi o apresentado na equação 4.3.

$$C = \begin{pmatrix} q_{1,1} & q_{1,2} & q_{1,3} & q_{1,4} \\ q_{2,1} & q_{2,2} & q_{2,3} & q_{2,4} \\ q_{3,1} & q_{3,2} & q_{3,3} & q_{3,4} \end{pmatrix} \quad (4.3)$$

De forma que, aplicando C a um ponto 3D, em coordenadas homogêneas (GOMES; VELHO, 2008b) sob o Sistema de Coordenadas do Mundo (CARVALHO; VELHO, 2005), obter-se-á diretamente o respectivo ponto 2D em termos de pixel na imagem.

4.3 Calibração de câmera

Costumeiramente, o modelo de câmera virtual é construído a partir dos parâmetros (intrínsecos e extrínsecos) (GOMES; VELHO, 2008a) da câmera, pode-se dizer que esta é a forma direta de especificação da câmera.

Neste trabalho, um dos empecilhos era o desconhecimento destes atributos da câmera e sem tal conhecimento não há como construir a transformação 4.3. No entanto, com uma correspondência previamente conhecida entre pontos da imagem e pontos 3D, o modelo da câmera pode ser estimado; este processo é chamado de Calibração de Câmera (costuma-se chamar, também de *Resectioning*, já que calibração pode envolver, ainda, ajuste cromático). (CARVALHO; SZENBERG; GATTASS, 1998), (SZENBERG, 2001) e (LANGER, 2010b) aplicam apenas a calibração geométrica e usam o termo "Calibração".

Como dito na seção 4.2, a aplicação do modelo de câmera em pontos 3D $P = (X, Y, Z)$ leva a pontos 2D $P' = (u, v)$ da imagem; de forma que, segundo (CARVALHO; SZENBERG; GATTASS, 1998), têm-se para cada correspondência de pontos 3D com 2D as equações 4.4 e 4.5.

$$u = \left(\frac{q_{1,1} \cdot X + q_{1,2} \cdot Y + q_{1,3} \cdot Z + q_{1,4}}{q_{3,1} \cdot X + q_{3,2} \cdot Y + q_{3,3} \cdot Z + q_{3,4}} \right) \quad (4.4)$$

$$v = \left(\frac{q_{2,1} \cdot X + q_{2,2} \cdot Y + q_{2,3} \cdot Z + q_{2,4}}{q_{3,1} \cdot X + q_{3,2} \cdot Y + q_{3,3} \cdot Z + q_{3,4}} \right) \quad (4.5)$$

Como a matriz C da câmera (equação 4.3) tem 12 graus de liberdade, são necessários, no mínimo, 6 correspondências, com pontos 3D linearmente independentes, para que se monte um sistema $A^T A.Q = \lambda.Q$ dado por 4.6 possível de ser resolvido (LANGER, 2010b).

$$\begin{pmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -u_1 X_1 & -u_1 Y_1 & -u_1 Z_1 & -u_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -v_1 X_1 & -v_1 Y_1 & -v_1 Z_1 & -v_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X_n & Y_n & Z_n & 1 & 0 & 0 & 0 & 0 & -u_n X_n & -u_n Y_n & -u_n Z_n & -u_n \\ 0 & 0 & 0 & 0 & X_n & Y_n & Z_n & 1 & -v_n X_n & -v_n Y_n & -v_n Z_n & -v_n \end{pmatrix} \cdot \begin{pmatrix} q_{1,1} \\ q_{1,2} \\ q_{1,3} \\ q_{1,4} \\ q_{2,1} \\ q_{2,2} \\ q_{2,3} \\ q_{2,4} \\ q_{3,1} \\ q_{3,2} \\ q_{3,3} \\ q_{3,4} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (4.6)$$

O problema 4.6 é, costumeiramente, impreciso devido a erros na coleta dos valores dos pontos, neste trabalho, isto acontece devido a imprecisão do mouse ao demarcar pixels (vide capítulo 5). Assim, usa-se uma aproximação pelo método dos mínimos quadrados.

Pelo método dos mínimos quadrados, busca-se o menor valor para a matriz C dada uma restrição que impeça se chegar à solução trivial ($C = 0$). Em (CARVALHO; SZENBERG; GATTASS, 1998) buscou-se uma restrição linear. No entanto, por trabalhar em coordenadas homogêneas, C é independente de escala ($\lambda C \cong C$), podendo-se adotar a restrição $|C| = 1$ (LANGER, 2010b).

Por conseguinte, este é um problema de minimização solucionável por multiplicadores de Lagrange, onde se chega à expressão 4.7. Observa-se que a solução Q é um autovetor de $A^T A$ e, para a maior minimização, Q deve ser o autovetor gerado pelo menor autovalor de $A^T A$. (LANGER, 2010a)

$$A^T A.Q = \lambda.Q \quad (4.7)$$

Por fim, a organização dos valores de Q na forma de C completa o processo de calibração da câmera, o que permite, a busca de pontos 3D correspondentes a pontos 2D da imagem (capítulo 5).

5 A Técnica de Reconstrução

Em posse do referencial teórico descrito no capítulo 4, a reconstrução de um objeto se deu através da indicação de pontos de interesse na imagem e da implementação de hipóteses de restrição geométrica, baseadas na transformação de câmera encontrada, a fim de eliminar a multiplicidade inerente às correspondência projetiva (seção 4.1).

Inicialmente, é necessário descobrir qual matriz de transformação projetiva é capaz de gerar tal imagem do objeto. Tal conhecimento permite conhecer a posição 2D, no sistema de coordenadas e sob mesmo ponto de vista da imagem, que corresponde a qualquer ponto 3D do espaço. Como discutido na seção 4.3, 6 correspondências entre pontos 2D e 3D, conhecidas previamente, são o bastante para fornecer uma aproximação da transformação desejada; no entanto, neste trabalho são usadas 8 correspondências, obrigatoriamente entre os pontos 3D:

- $(0,0,0)$;
- $(1,0,0)$;
- $(0,1,0)$;
- $(1,1,0)$;
- $(0,0,1)$;
- $(1,0,1)$;
- $(0,1,1)$;
- $(1,1,1)$.

Tal restrição tem o intuito de simplificar a entrada de dados do usuário no aplicativo de reconstrução (seção 6.3) e garantir a satisfação da condição de independência linear para a boa execução do método de calibração (seção 4.3).

Após o processo de calibração já é possível descobrir se um ponto 3D tem sua projeção coincidente com um ponto de interesse na imagem, podendo, por força-bruta, inferir a posição 3D no espaço deste ponto de interesse, 2D, da imagem.

No entanto, resta, ainda, o problema da multiplicidade de pontos; sabendo-se que isto se dá devido a formação da imagem ser feita através da interseção entre a reta de projeção (reta que liga o ponto 3D retratado ao centro de projeção da câmera) e o plano de projeção (figura 4.2).

Analisando o cenário que contém o objeto, nota-se que a reta de projeção contém vários pontos 3D, mas cada um destes tem um valor exclusivo da coordenada Z (figura 5.1); isto é provado pela geometria analítica, uma vez que a interseção entre uma reta (a reta de projeção, no caso) e um plano (no caso, um plano $Z = k$, com um k qualquer) é um ponto.

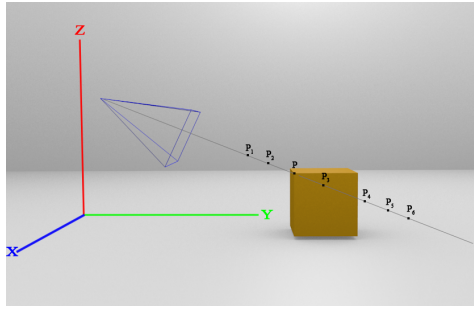


Figura 5.1: Os pontos P_i pertencem à reta de projeção de P , mas sua coordenadas Z são diferentes da coordenada do ponto P

Se, ao indicar o ponto de interesse, fosse fornecido o seu valor da coordenada Z , determinar seu correspondente 3D no espaço seria uma tarefa banal com uma demanda exhaustiva por dados do usuário. Em viés, sem informações além da transformação projetiva e coordenadas 2D do ponto de interesse não há formas de determinar sua coordenada Z (CARVALHO; VELHO, 2005).

No entanto, devido ao apelo visual da imagem, é fácil determinar se um ponto da imagem tem seu correspondente 3D com coordenada Z igual, ou aceitavelmente igual, à coordenada Z do correspondente 3D de outros pontos da imagem (figura 5.2). De forma que, ao fornecer os pontos de interesse para a reconstrução, indicar quais pontos de interesse da imagem têm seu correspondente 3D em uma mesma coordenada Z é simples e provê subterfúgios para a eliminação da multiplicidade projetiva.

Neste trabalho, para cada grupamento é atribuído um identificador numérico inteiro, onde, quanto menor o identificador, menor a coordenada Z do grupamento.

Assim, de posse de um grupamento de pontos da imagem com mesma coordenada Z , iterativamente, são descartados valores para a coordenada Z do grupamento

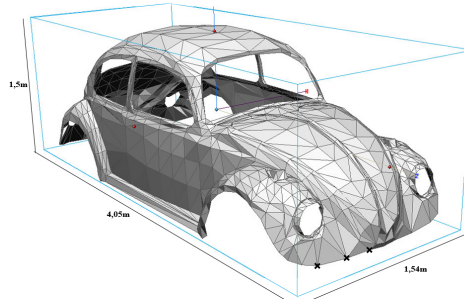


Figura 5.2: É perceptível que os pontos demarcados com um 'X' têm a mesma coordenada Z no espaço

que cujo plano no espaço, pela transformação projetiva encontrada para a imagem, não possui um ponto possível de ser o correspondente 3D de algum dos pontos de interesse do grupamento em questão.

Tal eliminação de multiplicidade por interseção é eficaz dado o seu embasamento geométrico citado anteriormente. No entanto, assim como no estabelecimento das correspondências (seção 4.3), há imprecisão na demarcação dos pontos de interesse na imagem. Então, é compreensível que seja adotada uma tolerância projetiva. A tolerância projetiva diz respeito a quão longe do ponto demarcado como interessante à reconstrução um ponto 3D ainda pode projetar para ser considerado como correspondente a este; tolerâncias grandes levam à imprecisão de reconstrução, ao passo que tolerâncias pequenas levam à dificuldade em achar correspondentes. O valor da tolerância projetiva é definido pelo usuário no momento da reconstrução.

Tal tolerância pode causar a não convergência da eliminação por interseção para um único valor; neste caso, o valor da coordenada Z adotado para o grupamento atual é o valor mais próximo da coordenada Z do grupamento anteriormente analisado.

Ainda em consequência da tolerância projetiva, com a coordenada Z adotada um número finito de pontos com esta coordenada Z projetam próximo o suficiente do ponto de interesse. Assim, como correspondente 3D definitivo do ponto de interesse em questão é adotado o baricentro destes pontos.

Não obstante, é necessário que se determine um espaço finito de busca destes correspondentes 3D, já que há pontos 3D que projetam em um ponto de interesse em infinitas posições do espaço. Para limitar a busca, então, são usados os mesmo pontos 3D aplicados na correspondência para calibração, isto é, os pontos $P = (x, y, z) | 0 \leq x, y, z \leq 1$. Como resultado, os pontos 3D reconstruídos terão suas coordenadas entre 0 e 1.

Dentro do espaço de busca, um quesito ainda importante é a distância ente dois valores consecutivos de uma coordenada (i.e. incremento). Incrementos menores levam à uma maior precisão da reconstrução, enquanto incrementos menores tornam o algoritmo mais ágil. Neste trabalho foi adotado o incremento de 0.005

limitação de busca incremento

6 Os Softwares

Ao decorrer do desenvolvimento deste trabalho, fora o software que implementa a técnica principal deste trabalho, foram sendo necessário que técnicas adjuntas à da estudada aqui também fossem implementadas em softwares a fim de facilitar comprovações de resultados, automatizações de tarefas auxiliares ao projeto; além de agregar determinada riqueza de esforço e de ferramentas a este trabalho de conclusão de curso.

A seguir são apresentados os exemplos mais representativos dos softwares desenvolvidos ao longo do projeto.

6.1 Aplicativo de detecção de retas

Tanto para auxiliar na detecção automática dos segmentos de reta do sistema de coordenadas de referência quanto dos pontos (interseção de segmentos de reta) de interesse à reconstrução. Foi desenvolvido um aplicativo para a obtenção de parâmetros analíticos de retas apresentadas em uma região de uma imagem digital através de linhas.

As linhas inferidas como interessantes, por serem representadas através da rasterização - processo de conversão da representação vetorial para a matricial (AZEVEDO; CONCI, 2003) - precisaram sofrer um processo de regressão linear a fim de se obter uma aproximação dos parâmetros geométricos analíticos de sua formação. A regressão linear executada para a obtenção dos coeficientes angular e linear das retas de interesse se deu através do método dos mínimos quadrados.

O Métodos dos Mínimos Quadrados é uma técnica de regressão linear que busca ajustar corretamente coeficientes de uma curva algébrica de forma que estes minimizem o somatório do quadrado da diferença entre os valores dos dados reais e os valores estimados, minimizando o erro de aproximação.

Com uma interface apropriada, pôde-se carregar uma imagem e, através da demarcação manual de uma região de interesse na mesma, pontos da linha do sistema de referência foram detectados pela sua cor e submetidos ao método dos mínimos quadrados (figura 6.1).

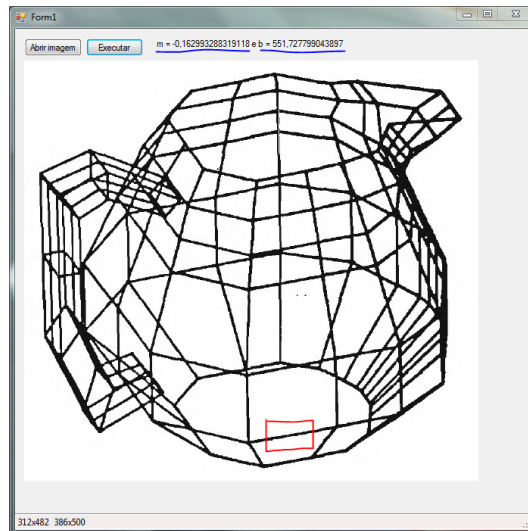


Figura 6.1: O programa de regressão de rasterização com região de interesse e coeficientes ressaltados

Posteriormente os coeficientes encontrados foram submetidos ao programa Microsoft Mathematics® comprovando que os valores encontrados são coerentes (figura 6.2). Note que, como o sentido de orientação dos pontos da imagem (de cima para baixo) é inverso ao do plano cartesiano (de baixo para cima) o aspecto da reta em relação ao aspecto da linha é invertido.

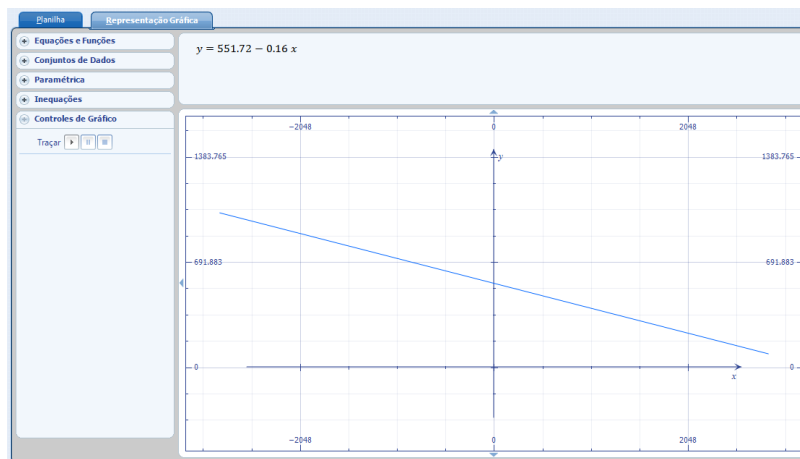


Figura 6.2: Reta analítica correspondente à linha da região demarcada na figura 6.1

Com esta técnica se obteve uma possibilidade de automatização da busca de elementos de interesse nas imagens estudadas.

6.2 Aplicativo de visualização 3D

6.3 Aplicativo de Reconstrução

7 Resultados, Discussões e Trabalhos Futuros

Referências Bibliográficas

- AGUIAR, E. de. Preliminary techniques. In: *Animation and Performance Capture Using Digitized Models*. New York: Springer, 2010. p. 9–18.
- AZEVEDO, E.; CONCI, A. *Computação Gráfica: Teoria e Prática*. Rio de Janeiro: Elsevier, 2003.
- BRUNELLESCHI'S Peepshow & the Origins of Perspective. 1998. Disponível em <https://www.dartmouth.edu/~matc/math5.geometry/unit11/unit11.html> Acesso em: 24 Set 2014.
- CARVALHO, P. C.; VELHO, L. *Fotografia 3D*. Rio de Janeiro, 2005.
- CARVALHO, P. C. P.; SZENBERG, F.; GATTASS, M. Image-based modeling using a two step camera calibration method. In: SIBGRAP'98. *Proceedings of International Symposium on Computer Graphics, Image Processing and Vision*. [S.l.], 1998. p. 388–395.
- GOMES, J.; VELHO, L. Câmera virtual. In: *Fundamentos da Computação Gráfica*. Rio de Janeiro: IMPA, 2008. p. 343–371.
- GOMES, J.; VELHO, L. Geometria. In: *Fundamentos da Computação Gráfica*. Rio de Janeiro: IMPA, 2008. p. 21–52.
- HEINZLE, S. et al. Computational stereo camera system with programmable control loop. In: *ACM SIGGRAPH 2011 Papers*. New York, NY, USA: ACM, 2011. (SIGGRAPH '11), p. 94:1–94:10.
- LANGER, M. *Lecture 18: Least Squares, Singular Value Decomposition*. 2010. Disponível em: <http://www.cim.mcgill.ca/~langer/558/18-leastsquares.pdf> Acesso em 18 Set 2014.
- LANGER, M. *Lecture 19: Camera Calibration*. 2010. Disponível em: <http://www.cim.mcgill.ca/~langer/558/19-cameracalibration.pdf> Acesso em 18 Set 2014.
- LEE, S.; FENG, D.; GOOCH, B. Automatic construction of 3d models from architectural line drawings. In: *Proceedings of the 2008 symposium on Interactive 3D graphics and games*. [S.l.: s.n.], 2008. p. 123–130.
- MUSEUM, C. H. *Mapping Sutherland's Volkswagen*. 2014. Disponível em <http://www.computerhistory.org/revolution/computer-graphics-music-and-art/15/206/560> Acesso em: 14 Ago 2014.
- MUSEUM, C. H. *Rendering of Sutherland's Volkswagen*. 2014. Disponível em <http://www.computerhistory.org/revolution/computer-graphics-music-and-art/15/206/559> Acesso em: 14 Ago 2014.
- MUSEUM, C. H. *Wireframe model of Sutherland's Volkswagen*. 2014. Disponível em <http://www.computerhistory.org/revolution/computer-graphics-music-and-art/15/206/558> Acesso em: 14 Ago 2014.

- NEVES, A. C. das; TSZUKI, M. de S. G. *Reconstrução 3D a Partir de Sequência de Imagens*. Monografia (Graduação) — Departamento de Engenharia Mecatrônica e de Sistemas Mecânicos, Escola Politécnica da Universidade de São Paulo, 2010.
- NEWELL, M. E. The utilization of procedure models in digital image synthesis. Supported by Advanced Research Projects Agency of the Department of Defense. 1975.
- PEREIRA, G. V.; SCHOEFFEL, P. Metodologia lean de desenvolvimento de software: Uma visão geral. Disponível em: http://www.ceavi.udesc.br/arquivos/id_submenu/387/guilherme_metodologia_lean_de_desenvolvimento_de_software_uma_visao_geral.pdf Acesso em: 14 ago 2014. 2012.
- POTTMANN, H. et al. *Projections*. [S.l.]: Bentley Institute Press, 2007. 23-65 p.
- ROSENFELD, A. Picture processing by computer. *ACM Computing Surveys*, v. 1, n. 3, p. 147–176, 1969.
- SORKINE-HORNUNG, A. et al. *Non-linear Disparity Mapping for Stereoscopic 3D*. 2012. Disponível em <http://www.disneyresearch.com/publication/non-linear-disparity-mapping-for-stereoscopic-3d/> Acesso em 27 Ago 2014.
- SZENBERG, F. *Acompanhamento de Cenas Com Calibração Automática de Câmeras*. 144 p. Tese (Doutorado) — Departamento de Informática, Pontifícia Universidade Católica do Rio, 2001.
- TORCHINSKY, J. *The First Real Object Ever 3D Scanned And Rendered Was A VW Beetle*. 2013. Disponível em: <http://floost.com/jalopnik-post-the-first-real-object-ever-3d-scanned-and-rendered-was-a-vw-beetle-2706861> Acesso em: 14 ago 2014.