



UNIVERSIDADE METODISTA DE PIRACICABA
FACULDADE DE CIÊNCIAS EXATAS E DA NATUREZA
MESTRADO EM CIÊNCIA DA COMPUTAÇÃO

**Geração de objetos
virtuais a partir de
imagens**

MARCELO ANTONIO RIBEIRO CAMARGO

ORIENTADOR:

PROF. DR. CLÁUDIO KIRNER

PIRACICABA, SP

2008



UNIVERSIDADE METODISTA DE PIRACICABA
FACULDADE DE CIÊNCIAS EXATAS E DA NATUREZA
MESTRADO EM CIÊNCIA DA COMPUTAÇÃO

**Geração de objetos
virtuais a partir de
imagens**

MARCELO ANTONIO RIBEIRO CAMARGO

ORIENTADOR:

PROF. DR. CLÁUDIO KIRNER

Dissertação apresentada ao Mestrado em Ciência da Computação, da Faculdade de Ciências Exatas e da Natureza, da Universidade Metodista de Piracicaba – UNIMEP, como parte dos requisitos para obtenção do Título de Mestre em Ciência da Computação.

PIRACICABA, SP

2008

Geração de objetos virtuais a partir de imagens

Autor: Marcelo Antonio Ribeiro Camargo

Orientador:

Prof. Dr. Cláudio Kirner

Dissertação de Mestrado defendida e aprovada 31 de Julho de 2008, à Banca Examinadora constituída dos Professores:

Prof. Dr. Cláudio Kirner (Orientador)
UNIMEP

Prof. Dr. José Luís Zem
UNIMEP

Prof. Dr. Marcio Merino Fernandes
UFSCar

Aos

Meus pais, Roberto e Sueli, pelo incentivo e conforto,

Aos

Meus familiares, pelo apoio e confiança,

À

Minha esposa, Denize, pela ajuda e compreensão.

AGRADECIMENTOS

A todos os professores do mestrado da UNIMEP, pelos ensinamentos, em especial ao professor e amigo Cláudio Kirner, que se mostrou sempre à disposição para a solução de qualquer problema, principalmente nos momentos mais difíceis.

Aos professores Luis Augusto Consularo e Regina Célia Coelho, que iniciaram a orientação e deram brilho ao tema escolhido, norteando as principais decisões para o desenvolvimento deste projeto.

À Secretaria de Pós-Graduação da UNIMEP, pelo rápido e ótimo atendimento.

À UNIMEP, pela estrutura de ensino proporcionada, e pela excelente qualidade do ensino dispensada aos alunos.

Aos colegas, pela convivência, rápida porém intensa, onde pudemos nos ajudar com valiosa troca de informações e aprendizado.

“Nós, homens do conhecimento, não nos conhecemos; de nós mesmos somos desconhecidos”.

Friedrich Nietzsche

Geração de objetos virtuais a partir de imagens

RESUMO

Realidade virtual, realidade aumentada e computação gráfica assumem atualmente uma posição privilegiada no cenário tecnológico, tanto da informática quanto da comunicação. A interatividade com mundos virtuais e outras aplicações relacionadas são cada vez mais freqüentes em nossa rotina. Entretanto, existe uma lacuna entre a utilização e a geração de modelos tridimensionais, capazes de representar objetos do mundo real em espaços criados por computador. Dessa maneira, o presente trabalho propõe soluções que facilitam a importação de informações da realidade para os ambientes virtuais, possibilitando a execução desta tarefa por usuários domésticos, popularizando assim o interesse desta área do conhecimento.

PALAVRAS-CHAVE: Realidade Virtual, Realidade Aumentada, Computação Gráfica, Visão Computacional, Morfologia Matemática, Objetos Tridimensionais.

GENERATION OF VIRTUAL OBJECTS FROM IMAGES

ABSTRACT

Virtual reality, augmented reality and graphic computation, actually assume a privileged position in the technological situation. The interactivity with virtual worlds and other related applications are increasingly common in our diary routine. However, there is a gap between the use and the generation of three-dimensional models, able to represent real-world objects in space created by computers. Thus, this paper proposes solutions which facilitate the import of information from reality to virtual environments, enabling the implementation of this task by home users, so popular interest in this knowledge area.

KEYWORDS: Virtual Reality, Augmented Reality, Graphic Computation, Computer Vision, Morphology Mathematics, Three-dimensional Objects.

SUMÁRIO

Lista de Figuras.....	XIII
1. Introdução	01
2. Revisão Bibliográfica.....	04
2.1. Criação de modelos tridimensionais	05
2.2. Criação de modelos tridimensionais a partir de imagens.....	06
2.3. Silhuetas / Visual Hull (Envoltória visual)	07
2.3.1. Image-based (baseada em imagem).....	07
2.3.2. Polyhedral (poliedros)	07
2.3.3. Volume carving (escultura do volume)	08
2.4. Reconstrução com raio laser	08
2.5. Reconstrução com emulsão em líquidos	10
2.6. Reconstrução com Luz Estruturada.....	11
2.7. Técnicas auxiliares.....	13
2.7.1. Realce do Contraste.....	13
2.7.2. Segmentação e Crescimento de Regiões.....	14
2.7.3. Limiarização	15
2.7.4. Classificação de alvos (Registro de imagens).....	18
2.7.5. Template matching (Corresp. por padrão)	18
2.8. Processamento de imagens na área médica	19
2.9. Representação de superfícies.....	21
2.9.1. Voxel coloring.....	23
2.10. Ferramentas para criação 3D.....	24
2.10.1. SketchUp	25
2.10.2. 3D Studio Max	26

3. Descrição do Trabalho.....	28
3.1. Metodologia	28
3.2. Etapas do projeto	29
3.2.1. Captura de imagem.....	29
3.2.2. Filtros	29
3.2.3. Limiarização	30
3.2.4. Seleção das regiões.....	31
3.2.5. Detecção de bordas (contornos)	31
3.2.6. Geração e armazenamento da cadeia de pontos. 31	
3.2.7. Criação da nuvem de pontos.....	32
4. Aquisição de forma 3D utilizando silhuetas.....	34
4.1. Definição de silhueta.....	34
4.2. Detecção de silhueta.....	35
4.3. Representação de silhueta	36
4.4. Distância focal.....	36
4.5. Projeção volumétrica das silhuetas.....	38
4.6. Intersecção das PVSs.....	39
4.6.1. Em um espaço de voxels	39
4.7. Representações do espaço de voxels	40
4.7.1. Octree	41
5. Objetos 3D em ambiente RV e RA	43
5.1. Ambientes de RV e RA	43
5.2. Objetos 3D para Realidade Virtual.....	45
5.3. Objetos 3D para Realidade Aumentada.....	46
5.4. Simplificadores LOD	47
5.4.1. VizUp Optimizer	47

6. Desenvolvimento	49
6.1. Resultados iniciais	49
6.2. Experimentações	51
6.3. Captura de cenas virtuais	52
6.3.1. Cortona ActiveX	53
6.3.2. Testes de execução	53
6.4. Aplicação para geração de objetos 3D	54
6.4.1. Explicações	55
6.4.2. Interface gráfica.....	56
6.4.3. Datasets	58
7. Conclusões.....	61
7.1. Considerações gerais	61
7.2. Contribuições	61
7.3. Resultados	62
7.4. Trabalhos futuros	62
8. Referências Bibliográficas	63

LISTA DE FIGURAS

Figura 01 - Scanner 3D com raio laser	09
Figura 02 - Equipamentos de escaneamento laser para uso doméstico.....	10
Figura 03 - Protótipo de escaner com emulsão em líquido	11
Figura 04 - Escaneamento de face com luz estruturada.....	12
Figura 05 - Exemplo de segmentação de imagem.....	15
Figura 06 - Exemplo de binarização.....	16
Figura 07 - Exemplo de imagem antes e depois da limiarização	17
Figura 08 – Janelas de busca e pesquisa no proc. de correspondência.....	18
Figura 09 - Reconstrução 3D em medicina	20
Figura 10 - Aplicação de objetos 3D na área médica.....	20
Figura 11 - Categorias decimação (A)Pontos (B)Linhas (C)Triângulos	22
Figura 12 - Coloração por voxels	24
Figura 13 - Google SketchUp.....	25
Figura 14 - Interface padrão do 3D Studio Max	26
Figura 15 - 3D Studio Max na indústria cinematográfica.....	27
Figura 16 - Exemplos de silhuetas	35
Figura 17 - Exemplo de variação de distância focal.....	36
Figura 18 - Geometria da câmera	37
Figura 19 - Projeção volumétrica das silhuetas.....	38
Figura 20 - Representação da matriz de projeção	38
Figura 21 - Intersecção das projeções volumétricas	39
Figura 22 - Inteseccão volumétrica em um espaço de voxels.....	40
Figura 23 - Representações pixel (A), voxel (B) e espaço de voxels (C).	41
Figura 24 - Quadtree (Nível 1) e Octree (Nível 2)	41
Figura 25 - Representação de superfície com octree	42

Figura 26 - Realidade Virtual – Cenário e Objetos 3D	45
Figura 27 - Realidade Aumentada – Objetos adic. ao mundo real.....	46
Figura 28 - Software VizUp – Simplificador de polígonos	48
Figura 29 - Protótipo virtual do dispositivo de captação de imagens	50
Figura 30 - Experimento realizado com filtro Sobel.....	52
Figura 31 - Simulação de captura de imagens.....	53
Figura 32 - Miniaturas das imagens capturadas pelo simulador	54
Figura 33 - Tela principal da ferramenta de criação de modelo 3D.....	56
Figura 34 - Tela abertura de imagem e contorno da silhueta.....	56
Figura 35 - Criação modelo 3D baseado na intersecção de 2 imagens.....	57
Figura 36 - Modelo criado a partir da intersecção de várias imagens	57
Figura 37 - Amostra do dataset da estátua de David	58
Figura 38 - Amostra do dataset de uma escultura em barro	58
Figura 39 - Amostra do dataset de uma cabeça de manequim.....	59
Figura 40 - Amostra do dataset de uma estátua de cavalo	59
Figura 41 - Amostra do dataset de um gato (objeto 3D)	59

1. Introdução

O objetivo deste trabalho é analisar diferentes formas de gerar objetos tridimensionais para utilização em cenários virtuais, extraídos do mundo real a partir de imagens bidimensionais, e dentre as diversas técnicas estudadas, selecionar e implementar aquela que apresenta melhores resultados.

Atualmente, o processo de reconstrução tridimensional de objetos reais é um dos temas mais discutidos na literatura relacionada à Visão Computacional e Fotogrametria Digital. Existem no mercado vários sistemas comerciais destinados à realização da reconstrução 3D, porém, possuem diversas desvantagens e restrições, como segue: custo elevado; necessidade de operadores e/ou utilizadores altamente qualificados e com vasto treinamento; inviabilidade de resposta em tempo real ou em um pequeno período de tempo e dificuldade de reconstrução em superfícies com texturas homogêneas.

Através de um ponto de vista mais abrangente, a necessidade de redução de custos permeia praticamente todas as ações do cotidiano, e o avanço tecnológico vem possibilitando aos computadores pessoais tornarem-se cada vez mais rápidos, fazendo com que a Realidade Virtual e a Realidade Aumentada deixem de ser objetos de estudo somente dos grandes centros de pesquisa, passando a ser utilizada também por usuários comuns.

Assim, o projeto e a criação de ferramentas que possibilitem a transferência de informações e objetos do mundo real para o virtual, isento de grandes investimentos, satisfaz os anseios tanto das áreas profissionais e acadêmicas, quanto amadoras e domésticas.

Com a facilidade de criação de modelos 3D, a Computação Gráfica em conjunto com a Realidade Virtual ampliam suas áreas de abrangência, diversificando consideravelmente as possibilidades de utilização destas tecnologias como suporte para quaisquer outros setores profissionais.

Diferenciando-se dos trabalhos similares realizados anteriormente, esta nova proposta apóia-se na idéia de simplificação dos processos manuais e redução considerável dos custos de implantação, direcionando toda a complexidade aos procedimentos computacionais. Tanto para programadores experientes quanto para

usuários iniciantes, o produto deste projeto deve agir de forma transparente, facilitando o entendimento e utilização.

Neste trabalho, também pretende-se demonstrar, de uma forma geral, os obstáculos e dificuldades encontradas durante o processo de planejamento e desenvolvimento, bem como, as possibilidades de solução para os problemas encontrados. Durante as fases de experimentações, existirá a necessidade de criação de algumas ferramentas com funções isoladas, que, ao trabalharem de forma interligada podem apresentar um resultado muito próximo de uma aplicação final.

Na seção “Revisão Bibliográfica” serão expostos alguns conceitos sobre criação de modelos tridimensionais e uma breve explicação sobre envoltória visual que é uma das bases deste trabalho. Serão apresentados alguns trabalhos e pesquisas relacionadas, com aplicação e objetivos similares, por exemplo com a utilização de raio laser, emulsão em líquido e luz estruturada. Também podem ser encontradas as técnicas mais aplicadas e as áreas de atuação com tendência para o uso desta tecnologia em potencial. Durante as fases de pesquisa, alguns tópicos adicionais tornaram-se foco de estudo, os quais foram incluídos nesta seção no sentido de ilustrar e embasar o conhecimento teórico do trabalho. O conteúdo está apresentado e organizado de acordo com a similaridade, possibilitando uma visão crítica sobre o tema e a situação atual.

O tópico “Descrição do Trabalho” está composto pelos itens “Metodologia” e “Etapas do projeto”, os quais respectivamente apresentam detalhes sobre a pesquisa, relacionando os modos encontrados para solução dos problemas e dificuldades encontrados; os recursos de hardware e software necessários para a conclusão do projeto, e finalmente, justificativas e explicações detalhadas.

O capítulo “Aquisição de forma 3D utilizando silhuetas” é destinado aos assuntos sobre computação gráfica, visão computacional e morfologia matemática. Serão expostos os temas referentes à aquisição da silhueta e a sua projeção volumétrica, etapas fundamentais para a reconstrução de um objeto 3D num espaço tridimensional. Outros conceitos como voxel e octree também serão explorados, no sentido de detalhar a metodologia utilizada para a representação do objeto reconstruído.

O quinto capítulo fará menção à Realidade Virtual e Realidade Aumentada, propriamente ditas, colocando-as como o centro de aplicação dos resultados deste

trabalho. Apresenta também algumas soluções para facilitar a implementação e aceitação destes objetos, como é o caso dos simplificadores de nível de detalhamento. Em conjunto, tais ferramentas favorecem a interação entre os objetos 3D reconstruídos e as ferramentas de realidade virtual e aumentada já disponíveis no mercado.

O tópico "Desenvolvimento" como o próprio título revela, apresentará, um-a-um, os acontecimentos e eventos ocorridos durante a fase de desenvolvimento do trabalho. Os resultados obtidos, os testes de execução e as explicações sobre as diferentes ferramentas elaboradas desde o início deste projeto, estarão descritas detalhadamente, com imagens e muitos exemplos. Nesta seção, pode-se encontrar os objetivos alcançados quando da conclusão do projeto, bem como, a descrição de todas as etapas de sua execução, desde os levantamentos iniciais até o término das ações propostas.

Como fechamento, o item "Conclusões" apresentará uma visão geral e resumida sobre a aquisição de objetos 3D a partir de imagens, avaliando os resultados sobre as ferramentas desenvolvidas e soluções encontradas para os problemas, traçando caminhos para trabalhos futuros e relacionados com o tema.

2. Revisão Bibliográfica

A criação ou planejamento de um sistema reconstrução de objetos 3D a partir de imagens requer a integração de conhecimentos de várias áreas, particularmente de visão computacional, morfologia matemática e processamento de imagens, além da área de Realidade Virtual e Realidade Aumentada como foco de aplicação dos resultados obtidos. Existem distintas possibilidades para a solução dos problemas nas várias etapas de construção do sistema, tanto dos componentes de hardware quanto de software, que implementam os métodos de processamento, análise e reconstrução das imagens.

Nas seções seguintes serão apresentados alguns tópicos de interesse que não esgotam a literatura sobre o assunto, mas apresentam apenas algumas das possibilidades para processamento. Uma revisão abrangente sobre todos os métodos disponíveis para o processamento das imagens está além do escopo deste trabalho.

Nesse sentido, os tópicos abaixo terão como objetivo a apresentação de técnicas utilizadas neste trabalho, bem como outras similares ou relacionadas ao assunto principal, que é a criação de objetos tridimensionais.

Também serão expostas técnicas similares, que utilizam ao invés da imagem, outros recursos como o raio laser e emulsão em líquido.

O sistema de reconstrução por Luz Estruturada, por exemplo, numa análise bastante superficial, se assemelha as técnicas anteriores no que diz respeito aos objetivos e ao tipo de obtenção do produto final, sendo que os resultados alcançados têm diversos aspectos em comum. Por outro lado, o conceito é bastante diferenciado, visto que todo o processo é focado em questões de perspectiva ótica.

Os próximos itens como “Limiarização”, “Realce de Contraste”, “Segmentação e Crescimento de Regiões”, “Registro de Imagens” e “Template Matching” expõem variados aspectos práticos e teóricos inerentes ao assunto estudo, e que num primeiro momento, demonstraram-se passíveis de utilização durante a fase de pesquisa e levantamento de informações. Não necessariamente serão utilizados diretamente no projeto, mas tal conhecimento reforça a complexidade do trabalho realizado, preparando e criando soluções secundárias para eventuais problemas ou

dificuldades, bem como, reforçando o embasamento técnico e demais fundamentações.

Ainda, no intuito de apresentar sugestões para aplicação comercial do produto final alcançado, realizou-se um levantamento de áreas profissionais com alto potencial de interesse, tendo como base uma perspectiva dos trabalhos relacionados e com foco direcionado. Como resultado, a área médica apresentou razoável destaque, principalmente no que diz respeito à Visão Computacional, Processamento de Imagens e Realidade Virtual. A reconstrução de objetos 3D, neste caso em particular, estaria voltada à representação de órgãos para os mais diferenciados fins, dentre eles, o treinamento virtual e o ensino médico.

Mesmo com a devida comprovação, a área médica é apenas um dos potenciais segmentos de interesse, paralelo à Engenharia, Robótica, Mecatrônica, entre outros.

2.1. Criação de modelos tridimensionais

Um modelo 3D tem a finalidade de representar um objeto real em uma situação virtual, porém, este não necessita obrigatoriamente existir previamente no mundo real. É o caso, por exemplo, de uma peça de engenharia em fase de planejamento, onde antes de sua produção efetiva, existe a fase de modelagem em ambiente tridimensional. Este processo, geralmente, é realizado manualmente ou através de cálculos matemáticos com o apoio de uma ferramenta CAD, onde o desenho é feito com o auxílio de um computador.

Nos casos onde o objeto já existe e necessita ser transportado para um ambiente virtual, seja qual for a finalidade, o processo costuma ser muito semelhante, a não ser pelo fato de existir uma referência para comparação, ou seja, o próprio objeto real. Desta forma, a modelagem ou o desenho assistido por computador torna-se muito complexo e trabalhoso para ser realizado manualmente.

Algumas ferramentas específicas para esta finalidade fazem parte da rotina de engenheiros, desenhistas, animadores, cineastas, produtores de televisão, e inúmeros outros profissionais de diversas áreas onde há a necessidade e possibilidade de uso de modelos tridimensionais. Entre elas, podemos citar o AutoCAD, 3D Studio Max, Maya, Blender, Lightware, Rinoceros, 3Dviz, Vector

Works, Silo 3D, Google SketchUp, e vários outros aplicativos de uso amador e profissional.

Os modelos 3D também podem ser gerados em um código para modelagem em ambientes de realidade virtual, através do padrão VRML (Virtual Reality Modeling Language) que é uma linguagem de marcação escrita em modo texto. Tal formato apresenta atualmente uma versão revisada e avançada, o X3D, que incorpora recursos disponíveis nos mais recentes dispositivos gráficos, o conceito de codificação XML, além de melhorias na estrutura e arquitetura da linguagem anterior.

Mesmo com todos os recursos existentes, a maior parte do trabalho na criação de modelos 3D e a grande responsabilidade sobre a qualidade dos resultados, fica a cargo do modelador, desenhista, programador ou quaisquer outros recursos humanos envolvidos no processo. Todos os softwares, linguagens e hardwares assumem o papel de simples ferramentas de auxílio na produtividade.

2.2. Criação de modelos tridimensionais a partir de imagens

Mesmo com o auxílio de ferramentas e recursos computacionais, um profissional de modelagem 3D se baseia em sua percepção visual para criar modelos que representem objetos reais em cenários virtuais. E uma forma de auxiliar ainda mais neste trabalho é transferir esta capacidade de percepção para a própria ferramenta, fazendo com que o computador visualize o objeto e o redesenhe automaticamente a partir de sua própria visão.

As atuais técnicas de computação visual (ou visão computacional) e os avanços tecnológicos dos softwares e hardwares gráficos tornam esta idéia perfeitamente possível. Tão possível e viável que podemos encontrar várias linhas de pesquisa sobre este assunto, além de aplicações em desenvolvimento, e inclusive alguns softwares comerciais já disponíveis no mercado.

A seguir, serão analisados alguns métodos para criação de modelos tridimensionais a partir de imagens, apresentando as principais vantagens e desvantagens de cada técnica.

2.3. Silhuetas / Visual Hull (Envoltória visual)

Os Métodos de processamento baseados na envoltória visual se dividem em três categorias: image-based (baseada em imagem), polyhedral (poliedros) e volume carving (escultura do volume). Todas invocam o mesmo princípio básico de que a silhueta de um objeto em relação à posição da câmera produz um cone volumétrico no qual o objeto está inserido. Os cones gerados pelos diversos pontos de visualização se cruzam, e o resultado desta interseção é a representação tridimensional do objeto.

2.3.1. Image-based (baseada em imagem)

É a técnica com o processamento mais rápido, entretanto, com um resultado bastante insatisfatório para os objetivos deste trabalho, visto que não se alcança uma representação tridimensional completa. Ao invés disso, faz uso de algumas amostras de imagens para a geração de novos pontos de vista que transmitem a sensação de tridimensionalidade, tanto do objeto quanto do ambiente, sem necessariamente existir modelos ou cenários 3D. É muito utilizada em jogos com visão em primeira pessoa, onde não há necessidade de interação direta do usuário com os objetos do cenário.

2.3.2. Polyhedral (poliedros)

Tal técnica produz a representação da superfície do objeto (facilmente convertida em uma representação volumétrica, se for necessário) pela intersecção geométrica dos diversos cones poligonais. Este procedimento é relativamente rápido, e fornece uma representação sem detalhamento e com pouca qualidade. Variações desta técnica utilizam-se de splines para obter uma superfície com maior suavização de suas curvaturas (Sullivan e Ponce, 1998).

2.3.3. Volume carving (escultura do volume)

Esta é a técnica mais simples de se aplicar, porém, é também a que apresenta o processamento mais lento. Trabalha com a projeção das silhuetas em relação aos voxels (pixels tridimensionais) para cada um dos diferentes pontos de vista. Um voxel é descartado se a sua posição tridimensional encontra-se totalmente fora do espaço volumétrico de todas as silhuetas projetadas. Isso produz uma representação volumétrica totalmente relacionada a um espaço tridimensional, e que mais se aproxima aos objetivos deste trabalho. Esta técnica também é mais tolerante aos erros de calibragem das câmeras.

Na prática, é como se houvesse um bloco inicial que passa por várias fases de lapidação e apara de suas arestas até chegar numa escultura final, que é a representação volumétrica do objeto captado nas imagens.

2.4. Reconstrução com raio laser

A reconstrução de modelos tridimensionais de objetos bem como a reprodução de superfícies, tanto de objetos como de planos terrestres, ocorre com grande frequência através do uso de recursos como o raio laser (**Figura 01**). Devido à sua eficácia e eficiência na aquisição de medidas e distâncias entre os pontos de início e término do feixe de laser, este método passou a ser comumente utilizado em pesquisas acadêmicas e em produtos comerciais, com a finalidade de capturar informações sobre planimetria. Inúmeros equipamentos foram desenvolvidos com esta finalidade, e passaram a ser denominados “Scanners 3D”, mas sempre estiveram restritos a grandes empresas e centros de pesquisa.



Figura 01 - Scanner 3D com raio laser

Mas esta tecnologia não é pioneira na área. A mais antiga é o levantamento com medidas de distâncias e direções usando instrumentos topográficos. Esta técnica ganhou significativa eficiência com o advento dos medidores eletrônicos de distâncias e direções, instrumentos que, integrados, convencionou-se chamar de estações totais. Os medidores a laser foram significativamente melhorados e passaram a realizar a medição de distâncias sem a necessidade de refletores, o que permitiu a determinação remota de coordenadas, sem tocar o objeto. Além disto, a evolução da tecnologia de controle de servo-mecanismos, aplicada a estes equipamentos, assegurou o desenvolvimento da estação total robotizada, que permite a medição programada de pontos, mesmo sem o operador (Tommaselli, 2003).

Já podemos encontrar dispositivos de escaneamento tridimensional direcionados ao público doméstico (Figura 02), como por exemplo o Zscanner 700, onde a facilidade na utilização é a qualidade primordial. Entretanto, tais equipamentos não oferecem um preço razoável para aquisição, e suficientes para despertar o interesse do mercado consumidor.



Figura 02 - Equipamentos de escaneamento laser para uso doméstico

2.5. Reconstrução com emulsão em líquidos

A forma menos convencional e um tanto quanto excêntrica, é a utilização de emulsão em líquido para reconstrução de objetos tridimensionais. Trata-se de um conceito baseado num misto entre as técnicas de modelagem artesanal e o fatiamento com base em ultrassom. Neste procedimento, o objeto real a ser reconstruído é inserido em um recipiente o qual é ocupado lentamente por um líquido de coloração diferente do objeto, suficiente para gerar um contraste visual identificável por equipamentos de aquisição fotográfica. Em diversos testes estudados, o líquido mais comum utilizado é o leite.

A cada intervalo de tempo uma cena é captada, e a seguir, a quantidade de líquido é aumentada, o que por sua vez, envolve uma camada superior do objeto em descanso no centro do recipiente. Nas cenas capturadas, pode-se identificar as partes do objeto que ainda não foram envolvidas pelo líquido, e que neste caso, coincidem com o contorno bidimensional daquela fatia ou camada.

Percebe-se que a reconstrução com emulsão em líquidos apresenta certa semelhança com a técnica de aquisição de contornos ou silhuetas das imagens, o qual será utilizada no desenvolvimento das ferramentas deste projeto.

Um protótipo foi desenvolvido pelo pesquisador ([Figura 03](#)), entretanto demonstra-se extremamente caseiro e amador, com a utilização de peças de brinquedo lego para a construção da estrutura principal.

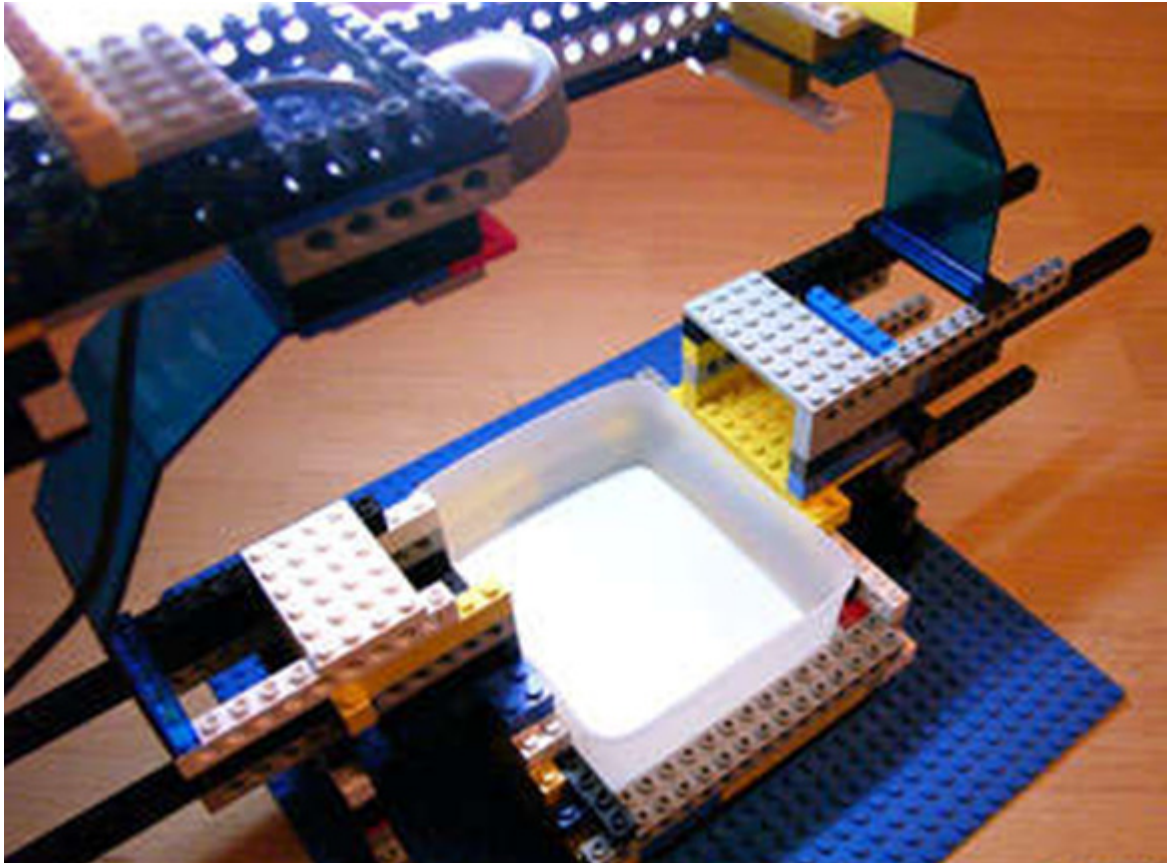


Figura 03 - Protótipo de escaner com emulsão em líquido

2.6. Reconstrução com luz estruturada

Segundo **Newton e Mitchell (1996)**, a técnica de reconstrução por luz estruturada (**Figura 04**) envolve o uso de um projetor e uma câmara, na qual o sistema de projeção é usado para projetar o padrão sobre o objeto na cena e uma câmara captura a imagem dos padrões que são distorcidos pela superfície de projeção. O projetor é equivalente à segunda câmara usada em Fotogrametria e qualquer ponto projetado na cena com direção e identificação conhecidas pode ser reconstruído usando os princípios de intersecção dos raios de luz da Fotogrametria analítica (relação câmara-projetor).

A determinação das relações geométricas existentes entre o sensor e os feixes de raios luminosos do padrão projetado é feita por um processo de calibração. O sistema de luz estruturada captura uma superfície em particular para cada ponto de vista, sendo necessária a combinação de vários pontos de vistas para criar um modelo 3D completo do objeto. Uma das vantagens deste sistema é quando este

usa uma fonte de luz branca e, conseqüentemente, não há a necessidade de cuidados especiais de segurança, ao contrário de sistemas de varredura utilizando laser, cuja luz possui algum nível de radiação.

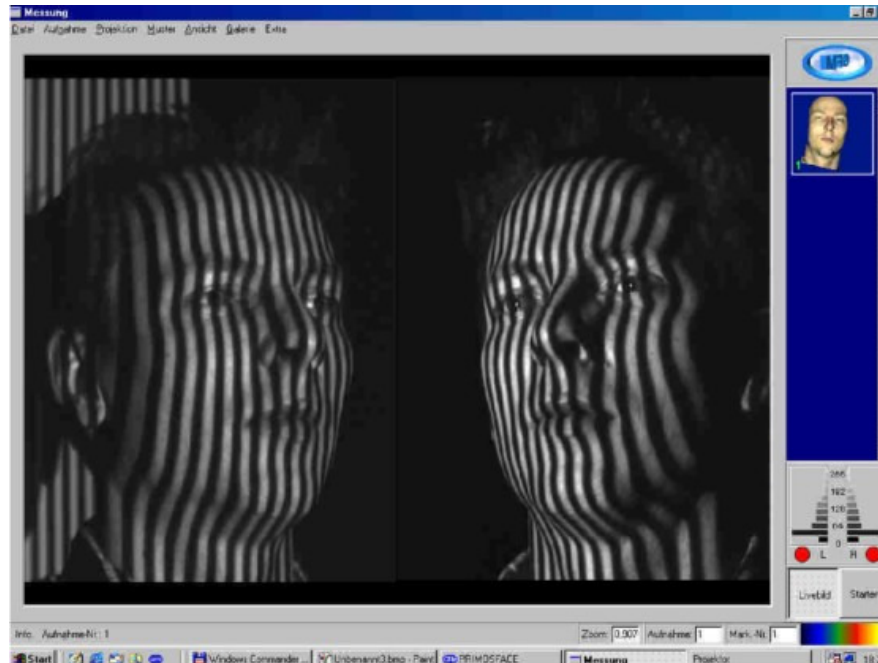


Figura 04 - Escaneamento de face com luz estruturada

Além disso, alguns sistemas de luz estruturada possuem precisões equivalentes às do sistema a laser. Como vantagem adicional deste sistema há a possibilidade de se capturar todos os alvos projetados e a imagem do objeto de uma só vez, sendo desnecessário manter o objeto estático por um longo período de tempo. Também pode haver a necessidade de tomada de múltiplas posições para cobrir inteiramente um objeto, quando ele for maior que o campo de cobertura do sensor.

As vantagens deste sistema em relação a técnicas estereofotogramétricas estão relacionadas com (Battle et al, 1996):

- menor tempo entre a tomada de fotos e a resposta final;
- menor custo dos equipamentos;
- não necessidade de pessoas especializadas para manusear os equipamentos; e,
- menor dificuldade de medir superfícies homogêneas.

Embora os sistemas de reconstrução por luz estruturada minimizem o problema de iluminação do ambiente ao projetar luz estruturada, ainda são necessárias várias operações de realce das imagens, até as etapas de reconhecimento e medição precisa dos alvos projetados. Distintas alternativas de processamento estão disponíveis, como mostra a farta literatura na área, tendo sido exploradas apenas algumas destas técnicas.

2.7. Técnicas auxiliares

2.7.1. Realce do Contraste

O realce de contraste linear consiste no processamento de uma imagem, de modo que o resultado seja mais apropriado para uma aplicação específica do que a imagem original. O contraste nas imagens é afetado pela iluminação não homogênea e abertura incorreta do diafragma da câmara durante a aquisição da imagem. É normalmente usado como uma etapa de pré-processamento para sistemas de reconhecimento de padrões (Gonzalez e Woods, 2000).

Dentre os métodos existentes para solucionar este problema encontram-se: método de expansão do histograma, método de contraste linear percentual e método de transformação local. Segundo Gonzalez e Woods (2000), o método de expansão do histograma, ocupa todo o intervalo de cinza, não provocando nenhuma perda. Esta operação é realizada a partir de uma interpolação linear.

Outra forma de melhoramento do contraste é denominada de contraste linear percentual. Este processo é utilizado quando se deseja especificar os valores de G_{min} e G_{max} para realçar determinado intervalo de cinza. Neste caso, utiliza-se certo percentual de pixels em relação à média do histograma. O problema deste processo é a perda de informação em benefício do realce dos pixels no intervalo que se deseja (Gonzales e Woods, 2000).

Os métodos mostrados anteriormente são globais no sentido de que os pixels são modificados através de uma função de transformação baseada na distribuição dos níveis de cinza sobre uma imagem inteira, não garantindo o realce local desejado (Gonzales e Woods, 2000).

Este método apresenta melhores resultados em relação ao método de expansão do histograma pelo fato de realizar um melhoramento local, porém, possui uma desvantagem para a automação do processo.

2.7.2. Segmentação e Crescimento de Regiões

A segmentação (**Figura 05**) consiste em subdividir uma imagem em regiões homogêneas usando propriedades escolhidas, como o nível de cinza, cor, textura, entre outras (**Sonka et al, 1998**). Os principais objetivos da segmentação são reduzir a quantidade de dados a serem processados na análise de imagens e obter informações importantes a respeito de objetos e feições nelas contidos (**El-Hakim, 1996**).

Existem diferentes formas de classificar os métodos existentes de segmentação. **Sonka et al (1998)** dividem os métodos de segmentação em três categorias: baseados em bordas, regiões ou no conhecimento global sobre a imagem. **Ballard e Brown (1982)** consideram dois tipos de segmentos principais, bordas e regiões, que podem ser separados por diferentes critérios. **Jain (1989)** separa as técnicas de segmentação em seis grupos distintos: limiarização, rotulação de componentes, fronteiras, regiões, correspondência por padrões e textura.

Os métodos de segmentação baseados em regiões procuram dividir a imagem em segmentos que tem características similares (**Jain, 1989**). Um dos métodos mais conhecidos para este tipo de segmentação é o crescimento de regiões. Este método consiste em agrupar pixels vizinhos que representam regiões homogêneas segundo algum critério de similaridade. Um dos critérios que pode ser usado é a comparação da média e do desvio-padrão dos níveis de cinza de duas regiões, verificando-se se estes valores são estatisticamente compatíveis.

Outros métodos de segmentação por regiões podem ser citados, como o de subdivisão e fusão e a segmentação watershed, baseada nos conceitos da morfologia matemática.

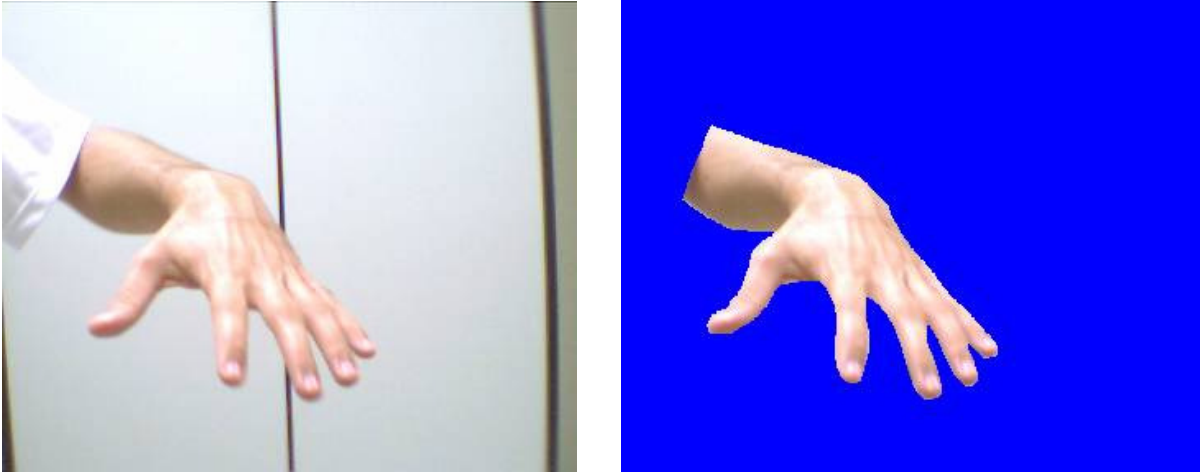


Figura 05 - Exemplo de segmentação de imagem

2.7.3. Limiarização

A operação de limiarização refere-se à definição de um valor ou de uma função, que permita separar elementos de interesse na imagem do fundo. Frequentemente a imagem original é convertida em imagem binária pelo processo de limiarização, atribuindo-se o valor zero aos pixels com tons de cinza menores que o limiar e um aos maiores (**Figura 06**). Este método pode ser considerado como um dos mais simples para a segmentação de imagens.

A limiarização pode ser feita usando técnicas globais ou locais. Nas técnicas globais determina-se um único limiar para toda a imagem, o que traz problemas causados pela refletância, sombras ou iluminação inadequada. Em imagens digitalizadas com scanners, particularmente de documentos, as técnicas globais podem ser apropriadas, porque o histograma da imagem é bimodal e similar em todas as regiões.

Em imagens coletadas com câmaras digitais de cenas 3D, entretanto, ocorre o efeito de escurecimento nas regiões periféricas da imagem. Devido à lei de iluminância, a luz que atinge os cantos da imagem é atenuada por um fator composto pelo ângulo entre o eixo ótico da câmara e o ponto em análise. Com isto a área central da imagem é mais iluminada que a periferia, gerando um efeito análogo ao de vinhete. Com isto, os histogramas das regiões centrais da imagem

apresentam uma média de intensidade maior que os histogramas em regiões mais próximas dos limites da imagem e a adoção de um único limiar não produz bons resultados.

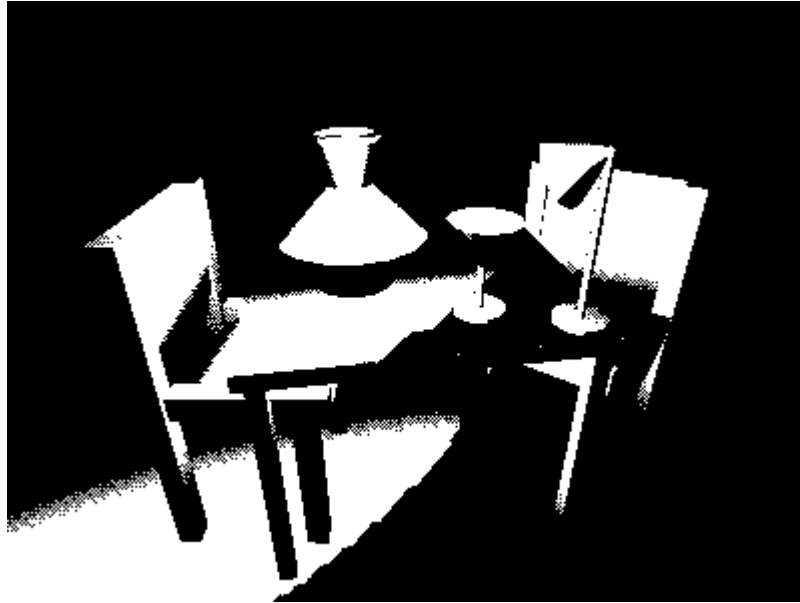


Figura 06 - Exemplo de binarização

Uma melhor alternativa, portanto, são as técnicas locais que consistem em determinar um limiar para cada região estabelecida por uma janela. O tamanho da janela depende das características da imagem original. As técnicas locais são bastante conhecidas e existem vários algoritmos com diferentes respostas dependendo do tipo de imagem (Ballard e Brown, 1982; Sezgin e Sankur, 2004).

A definição de um limiar global pode ser feita interativamente a partir da análise do histograma da imagem ou usando métodos automáticos. Uma análise detalhada das técnicas de limiarização e sua eficiência podem ser encontradas em Sahoo et al (1988) e em Sezgin e Sankur (2004).

Dentre as técnicas existentes, uma das mais conhecidas é a de Otsu (1979), mas várias outras técnicas foram desenvolvidas recentemente, como por exemplo, a de Pun (1980).

A técnica de Otsu (1979) é baseada na análise discriminante e o valor do limiar é obtido supondo-se que os pixels da imagem podem ser classificados em duas classes, que são o objeto e o fundo. Outro método global é o desenvolvido por Pun (1980) que é baseado na teoria da informação e, portanto, parte da premissa de que a geração de informação pode ser modelada como um processo probabilístico.

Uma das técnicas locais, considerada eficiente na avaliação realizada por **Sezgin e Sankur (2004)**, é a de **Niblack (1986)**. Entretanto, dependendo da aplicação analisada o método de Pun foi superior. Uma possibilidade similar é a de usar métodos globais, como os de Otsu e Pun, mas aplicá-los localmente, em pequenas janelas da imagem. **Artero et al (2000)** utilizaram uma técnica de limiarização local da imagem que consiste em determinar um limiar usando os pixels pertencentes a uma janela de 60 x 60 pixels. O valor de limiar assim obtido é utilizado para todos os pixels localizados na região central desta janela (uma região de tamanho 20 x 20 pixels). As dimensões destas janelas foram definidas empiricamente e podem ser alteradas dependendo das características da imagem. O limiar local pode ser definido por critérios como os de Otsu ou Pun, ou qualquer outro.

É importante ressaltar que muitos dos testes realizados por diferentes autores consideram um tipo específico de aplicação, por exemplo, documentos digitalizados, e a eficiência da técnica varia dependendo da aplicação e das condições nas quais as imagens foram tomadas.

Um exemplo clássico de uso de limiarização em imagens, com finalidade de otimizar a identificação de símbolos alfanuméricos é a leitura computadorizada de placas de veículos, para aplicação de multas de trânsito pelos radares fotográficos **(Figura 07)**.

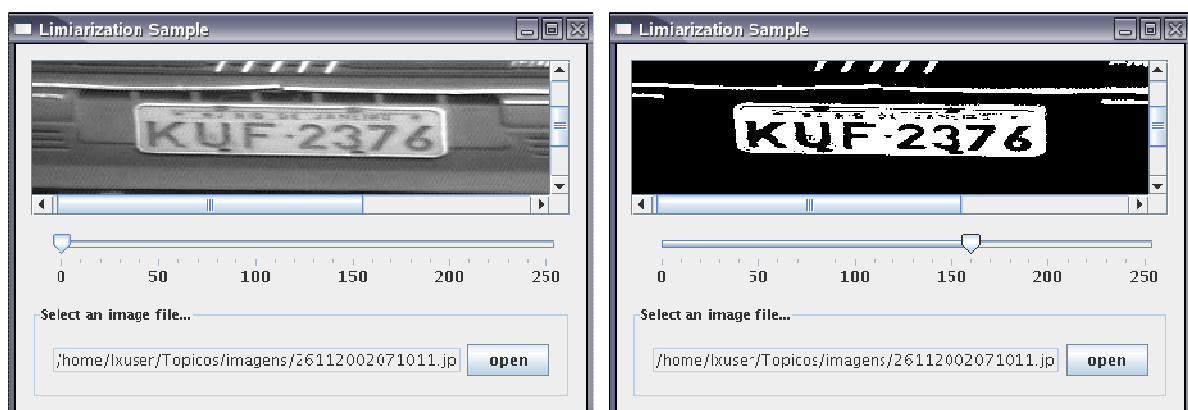


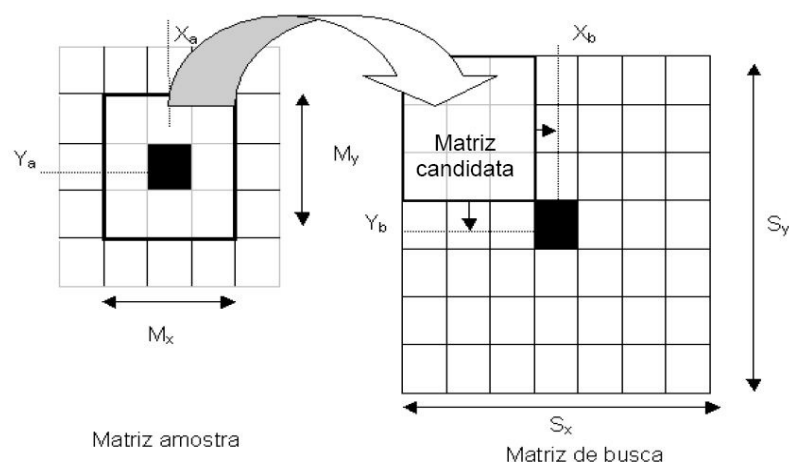
Figura 07 - Exemplo de imagem antes e depois da aplicação de limiarização

2.7.4. Classificação de alvos (Registro de imagens)

A classificação ou identificação de alvos é um processo que consiste em rotular os objetos anteriormente segmentados em diferentes classes definidas a priori. Esta classificação pode ocorrer a partir da similaridade entre cores, níveis de cinza, textura etc. Segundo **Jain et al (2000)**, as melhores abordagens para a classificação de padrões são: correspondência por padrão, classificação estatística, correlação estrutural ou sintática e redes neurais. Geralmente é utilizado o método de correspondência por padrão, com algumas variantes, como a estratégia de reamostragem adaptativa dos padrões.

2.7.5. Template matching (Correspondência por padrão)

Segundo **Schenk (1999)**, o método de correspondência por padrão consiste em comparar a distribuição de níveis de cinza de uma matriz amostra (padrão) com matrizes candidatas pertencentes a uma matriz de busca, usando uma função de correlação adequada (**Figura 08**). Várias são as funções para este tipo de correspondência: erro, erro quadrático, correlação, covariância cruzada e covariância cruzada modificada. A função mais usada na literatura e que proporciona melhores resultados é a covariância cruzada modificada ou coeficiente de correlação.



*Figura 08 – Janelas de busca e pesquisa no processo de correspondência
(Adaptado de **Andrade, 1998**)*

Para a determinação da correspondência a matriz amostra (ou janela de referência) é comparada com cada matriz candidata na janela de referência, comparação esta feita com o cálculo de uma das funções especificadas anteriormente. O pixel que apresentar a melhor resposta é selecionado como ponto de correspondência.

A função covariância cruzada modificada, também chamada de coeficiente de correlação resulta da normalização da função covariância cruzada. O coeficiente de correlação, neste caso, varia de -1 a 1 . O valor 1 para o coeficiente de correlação corresponde à medida de similaridade máxima, o valor 0 indica que não há correlação e o valor -1 indica correlação inversa.

2.8. Processamento de imagens na área médica

Os processos de aquisição de imagens médicas registraram avanços exponenciais nas últimas décadas. O processo de aquisição de imagens tomográficas e a aquisição de imagens funcionais de órgãos através da medicina nuclear são exemplos que atestam esses avanços. Paralelamente a isso, os processos de segmentação de imagens médicas registraram contribuições significativas. Acreditamos, no entanto, existir ainda um largo campo de pesquisa a ser trilhado pelas técnicas de reconhecimento.

No processamento de imagens podemos distinguir entre técnicas de segmentação de imagens e técnicas de interpretação ou reconhecimento. A segmentação decompõe a imagem em regiões, enquanto que o processo de reconhecimento gera uma descrição da imagem, atribuindo marcadores a essas regiões.

Imagens médicas representam estruturas biológicas com uma grande variedade de formas, que não podem ser facilmente descritas em termos de imagens padrões rotacionadas ou escalonadas, sendo difícil a segmentação das mesmas.

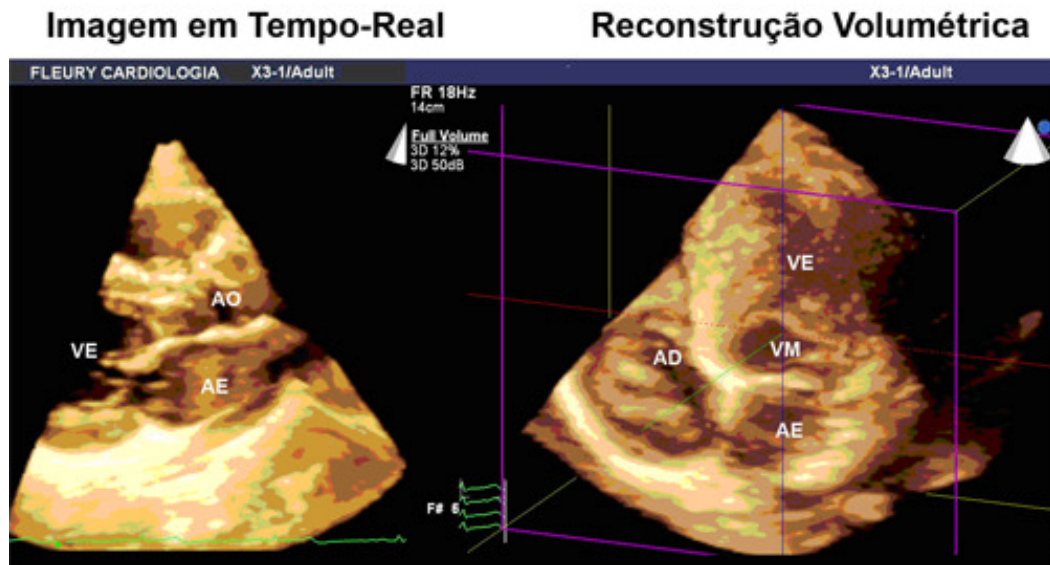


Figura 09 - Reconstrução 3D em medicina

Na literatura, encontramos um vasto campo de aplicação de técnicas de segmentação em imagens médicas: segmentação de estruturas do cérebro, segmentação do baço; segmentação de células; segmentação de nódulos em imagens pulmonares; segmentação do ventrículo esquerdo e segmentação do fígado. Nesses trabalhos são utilizadas uma enorme diversidade de técnicas, que abrangem desde a aplicação de limiar (Badran et al., 1990) até métodos de inteligência artificial, como redes neurais (Costa et al., 1995; Coppini et al., 1995).

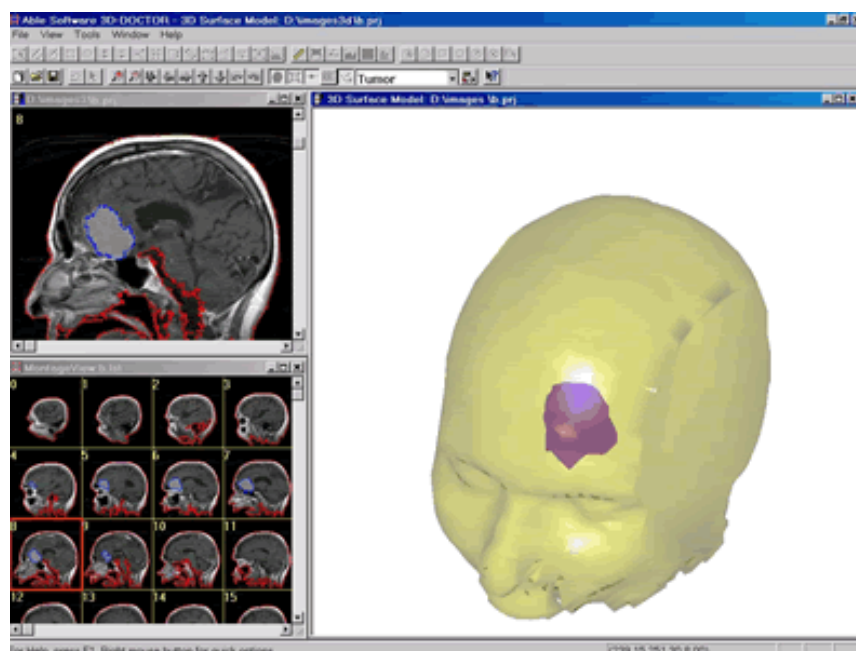


Figura 10 - Aplicação de objetos 3D na área médica

2.9. Representação de superfícies

A idéia principal que descreve a generalização de um modelo de terreno (Chen e Tobler, 1986; Floriani et al., 1996; Voigtmann et al., 1987) é que a representação de uma superfície em um nível arbitrário de detalhe pode ser realizada tanto pela inserção de pontos significativos em um modelo com poucos detalhes ou pela remoção de pontos menos significativos de um modelo com excesso de detalhes. A maioria dos métodos de simplificação de superfícies encontrados na literatura pode ser classificada como métodos de refinamento e decimação (Pedrini, 2000). O método de refinamento inicia com uma aproximação grosseira da superfície e repetidamente adiciona pontos na triangulação até que o modelo satisfaça um critério de aproximação pré-especificado.

O método de decimação inicia com o modelo de triangulação contendo todo o conjunto de pontos, simplificando-o iterativamente até que o critério de aproximação desejado seja atingido. O objetivo principal de um algoritmo de decimação é reduzir o número de triângulos de uma malha, preservando as principais feições com a melhor definição possível. O algoritmo mais utilizado se enquadra dentro deste último grupo.

Grande parte dos algoritmos de decimação pode ser classificada em três categorias, de acordo com as entidades geométricas utilizadas para a remoção. Os principais grupos de métodos são:

a) baseados na redução de vértices da triangulação (Figura 11 A). Alguns exemplos encontrados na literatura são apresentados a seguir.

b) baseados na remoção de linhas (ou faces) da triangulação (Figura 11 B). Neste grupo se enquadra o método de Hoppe et al. (1993), que apresenta um método de simplificação baseado em otimização de malhas, o de Hoppe (1996), que utiliza uma seqüência de eliminações de linhas para construção de malhas progressivas, e o método de Guéziec e Hummel (1995), que é um processo de simplificação que minimiza a medida do volume de erro utilizando programação linear.

c) baseados na remoção de triângulos no resultado da triangulação (Figura 11 C), tal como o método proposto por Scarlatos e Pavlidis (1990);

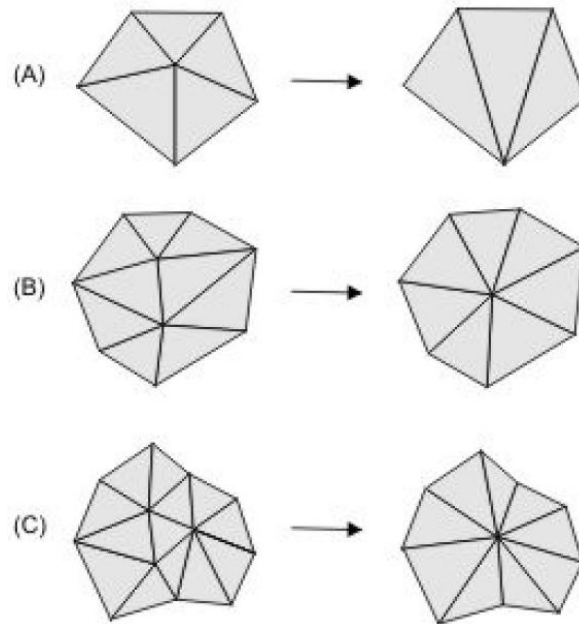


Figura 11

Categorias de decimação – (A) Pontos (B) Linhas (C) Triângulos

Lee (1989) propôs o método heurístico de simplificação de terrenos chamado de “drop heuristic”. O algoritmo faz um tipo de decimação de vértices, removendo um ponto a cada passo. Uma triangulação inicial é criada adicionando uma linha que conecta dois pontos diagonais a cada vizinhança 2×2 . O erro, medido como sendo as diferenças na elevação entre as duas superfícies é calculada para cada ponto remanescente na triangulação e o ponto com o menor erro é excluído.

Um algoritmo geral de decimação é apresentado por Schroeder et al. (1992). O algoritmo executa passagens múltiplas sobre uma triangulação existente, removendo vértices até que um valor especificado de erro é atingido. O erro sobre o vértice é calculado como sendo o valor da distância desde o ponto até o plano médio dos vértices arredores.

A simplificação também pode ser concretizada através da triangulação de um novo conjunto de vértices que substitui a triangulação original, como é proposto por Turk (1992). Nesta abordagem, um procedimento iterativo de repulsão de pontos é empregado para distribuir o novo conjunto de pontos sobre a superfície, concentrando mais pontos na região de maior curvatura. Os pontos originais são

então removido um a um, resultando em uma triangulação que preserva a topologia da superfície original.

Schroeder e Robbach (1994) descreveram um algoritmo de decimação onde a importância de um vértice é avaliada de acordo com a medida de rugosidade do terreno naquele vértice. Um vértice é removido da malha caso não haja mudança significativa na representação geral e a área ao redor do ponto removido é retriangulada. Uma abordagem similar é adotada no presente estudo.

O critério para a seleção de vértices é crucial durante o processo de generalização da triangulação já que isto pode determinar a qualidade e precisão da malha resultante. A métrica mais comumente empregada para avaliação da precisão é conhecida como regra L8, que se baseia na diferença máxima entre os dados originais de elevação e a superfície generalizada. Esta diferença corresponde à medida local de erro. Outra métrica comum é a regra L2, que fornece uma medida de desvio médio entre o modelo original e o aproximado.

2.9.1. Voxel coloring

A técnica de “Voxel coloring” ou “Coloração de voxels” consiste em capturar a cor de uma posição bidimensional de uma imagem, coincidente à posição tridimensional de sua projeção no espaço 3D (**Figura 12**). Em outras palavras, converter a posição de um voxel retornando à posição original de seu pixel na imagem, e assim, utilizar a cor deste ponto para coloração da superfície reconstruída.

Podemos dividi-la em duas propostas distintas, a Coloração de Voxels Generalizada (Generalized Voxel Coloring) e a Coloração de Voxels por Hipóteses Múltiplas.

Ambos algoritmos consistem em percorrer os voxels correspondentes a cada uma das imagens capturadas, projetando-os nas imagens em que estão visíveis e avaliando-os segundo sua consistência. Se um voxel projetado for considerado consistente então ele é colorizado com base nas cores obtidas das imagens em que se encontra visível. Caso contrário, atribui-se uma cor arbitrária ou simplesmente a cor transparente, o que corresponde a removê-lo do modelo. Neste trabalho, optou-se pelo uso da cor azul em casos de não detecção da cor original.

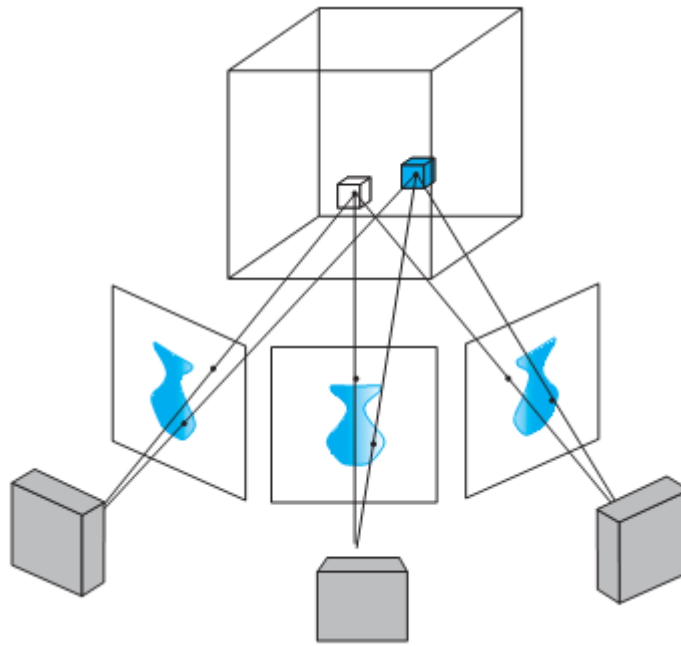


Figura 12 - Coloração por voxels

2.10. Ferramentas para criação 3D

Quando se pensa em ferramentas para criação de objetos ou cenários 3D, logo se destacam os conhecidos aplicativos CAD e os softwares para modelagem. Entretanto, existe uma gama muito mais abrangente, tanto de software quanto de hardware, que atendem aos profissionais desta área. Com os recentes avanços tecnológicos e grande aceitação da computação gráfica e realidade virtual por parte dos usuários domésticos, amadores ou profissionais, tais recursos passam a ocupar um espaço privilegiado em nosso cotidiano. Empresas como imobiliárias, marcenarias e instituições de ensino, que antes ficavam distante desta área de interesse, juntam-se às produtoras de vídeo e cinema, criadoras de jogos e entretenimento, especialistas em arquitetura e engenharia, aumentando consideravelmente o mercado consumidor de recursos de geração e manipulação tridimensional.

Por outro lado, tais ferramentas estão mais focadas na qualidade de sua produção e não na facilidade de sua manipulação, o que ainda restringe o acesso de outros públicos potenciais.

Abordaremos a seguir os softwares de maior destaque atualmente, seja pela forte entrada no mercado e rápida aceitação dos usuários (que é o caso do Google

Sketchup), seja pela tradição e marcante atuação nas áreas profissionais de computação gráfica (entre elas, o 3D Studio Max).

2.10.1. Softwares

O Sketchup (**Figura 13**) é um poderoso software da empresa Google desenvolvido para a criação de modelos em 3D, que possibilita de forma inovadora a criação, visualização e alteração de objetos tridimensionais. permitindo a fácil aplicação desde as fases de concepção do desenho até sua finalização e acabamento. Possui uma interface amigável, intuitiva e vem sendo utilizado para as mais diversas finalidades, como criação de modelos tridimensionais de casas, plataformas, projetos, e outros objetos que podem ser rapidamente compartilhados com outros usuários através da internet, ou simplesmente, posicionados num mundo virtual através de outra ferramenta da mesma empresa, o tão conhecido Google Earth, responsável pelo mapeamento de todo o planeta através de imagens de satélite. Esta integração entre o mundo real e o virtual incentiva ainda mais a produção de modelos e conseqüentemente a divulgação e pulverização no uso destes aplicativos.

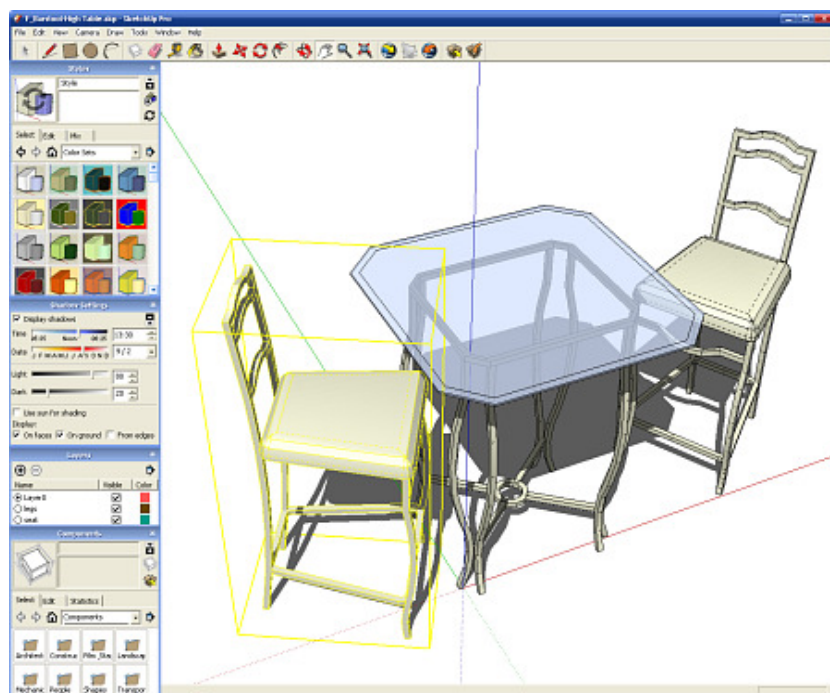


Figura 13 - Google SketchUp

2.10.2. 3D Studio Max

Outra ferramenta de destaque é o 3D Studio Max (Figura 14), da empresa Autodesk, que já está há muitos anos no mercado, e atualmente encontra-se em sua versão 2009. O 3D Max, como é carinhosamente chamado por seus usuários, possui, além dos recursos de modelagem e criação tridimensional, diversas outras funções e facilidades com o objetivo de proporcionar maior sensação de realidade às suas cenas. Iluminação, sombreamento, reflexos e texturização de materiais estão entre as principais preocupações dos criadores deste software.



Figura 14 - Interface padrão do 3D Studio Max

Quanto à questão de animação das cenas e objetos, a ferramenta dispensa alto nível de interatividade, inclusive com as forças da natureza (do mundo real), como gravidade, ventos e dinâmica de fluídos. Destarte, este software é muito utilizado na indústria cinematográfica (Figura 15), justamente por oferecer uma ótima aproximação entre imagens reais e virtuais.



Figura 15 - 3D Studio Max na indústria cinematográfica

Todas essas qualidades carregam consigo um preço muito alto, tanto da própria ferramenta quanto da complexidade de sua utilização. Além disso, torna-se necessária a presença de um equipamento com grande poder de computação, principalmente para as fases de cálculos de interação dos objetos em cena e renderização das imagens.

Segundo o fabricante, o equipamento básico para o funcionamento adequado do sistema, deve ser composto pelos seguintes itens:

- Processador Intel 64bits (EM64T)
- 4Gb de memória RAM
- 2Gb de espaço em disco (livres para “swap”)
- Placa de aceleração de vídeo (compatível com OpenGL e Direct3D)
- Sistema operacional Windows Vista ou Windows XP Professional x64

3. Descrição do Trabalho

O objetivo principal deste projeto é a criação de modelos tridimensionais de objetos a partir de informações extraídas de imagens, aplicando-se técnicas de visão computacional.

Estas imagens podem ser obtidas, inclusive, com o auxílio de dispositivos de captação com baixo custo, como por exemplo, máquinas fotográficas digitais, webcams e celulares com câmera embutida, além de outros equipamentos com recursos ópticos similares.

Os modelos 3D gerados a partir das imagens captadas podem ser utilizados para os mais diferentes propósitos, desde o simples entretenimento até a aplicação profissional, desde a modelagem de peças e componentes industriais até a realização de testes em ambiente simulado com o auxílio de técnicas de Realidade Virtual ou Aumentada.

3.1. Metodologia

A metodologia aplicada neste trabalho está voltada à pesquisa e desenvolvimento, baseada na busca, análise e experimentações de ferramentas e técnicas existentes, além da criação e apresentação de aplicações.

Os trabalhos de pesquisa e análise de resultados serão o foco inicial, tendo como auxílio, o desenvolvimento de ferramentas independentes e isoladas para a realização dos experimentos intermediários. Tais etapas ocorrerão em paralelo, e de forma interdependente.

Todos os testes serão realizados com equipamentos comuns e de baixo custo, como câmeras do tipo webcam e demais dispositivos específicos para esta finalidade.

Ao final de cada fase de testes e experimentos, e de posse dos resultados, será determinado o caminho a ser seguido, impossibilitando assim a determinação prévia dos recursos utilizados.

3.2. Etapas do projeto

Após a determinação do problema inicial e o conhecimento dos pontos-chaves para o sucesso do trabalho, realizou-se uma pesquisa superficial sobre o assunto, destinada à escolha das técnicas e ferramentas utilizadas. Têm-se como fases do projeto os seguintes componentes:

3.2.1. Captura de imagem

Para a realização da captura de imagem, utilizou-se a biblioteca ARToolKit, com o intuito de reaproveitamento das suas funções direcionadas ao reconhecimento de bordas ou contornos. Tal objetivo demonstrou-se inviável por alguns aspectos, entretanto, apresentou grande facilidade na manipulação da imagem capturada.

Decidiu-se, portanto, manter esta biblioteca como base para o desenvolvimento desta fase do projeto.

3.2.2. Filtros

Com a imagem capturada, iniciaram-se os estudos referentes ao processamento das informações obtidas. Diversos filtros foram testados, no intuito de conhecer os possíveis resultados, e selecionar entre eles, a melhor opção, tanto em termos de produto final quanto de velocidade de processamento. Entre os filtros analisados, destacam-se Sobel, Susan, Canny e Threshold.

Ainda, para propiciar uma redução no tempo de processamento, optou-se pela aplicação das técnicas de erosão e dilatação, no sentido de eliminar elementos sem importância na imagem capturada. Consideram-se aqui, elementos sem importância, aqueles que apresentam tamanho demasiadamente reduzido, vistos portanto como ruído na cena.

3.2.3. Limiarização

Neste trabalho, através dos diversos filtros analisados, a limiarização apresentou o melhor custo x benefício, sendo o custo referente ao processamento e o benefício diretamente ligado ao resultado obtido.

Tal projeto prevê a utilização da forma mais simples de limiarização, também conhecida como binarização, pois, a partir de um valor limiar, transforma os dados em níveis de cinza da imagem para preto e branco. Basicamente os pixels com valor menor que o limiar são transformados em pretos e os pixels com valor maior ou igual ao limiar são transformados em brancos.

A limiarização transforma uma imagem em nível de cinza ou colorida em uma imagem binária, reduzindo drasticamente o tamanho da imagem em bytes e assim agilizando os processamentos das fases posteriores.

Se os valores dos pixels dos componentes da imagem e aqueles do fundo são razoavelmente consistentes em seus respectivos valores sobre a imagem inteira, então um simples valor de limiar pode ser encontrado para esta imagem. Esta utilização de um simples limiar para todos os pixels da imagem é chamada de limiarização global. Contudo, para vários documentos, um único valor de limiar global não pode ser usado igualmente por causa das deformidades dentro das regiões de primeiro e segundo planos. Por exemplo, para um documento contendo áreas de fundo brancas bem como áreas de destaque de um fundo colorido diferente, os melhores limiares mudarão por área. Para este tipo de imagem, os valores de limiar diferentes são requeridos para áreas de locais diferentes. Este tipo de uso é chamado de limiarização adaptativa.

Existem diferentes métodos de limiarização. Nenhum método é melhor para todos os tipos de imagens e aplicações. Para problemas mais simples, onde as características da imagem não variam muito dentro dela ou entre diferentes imagens, os métodos mais simples serão suficientes.

3.2.4. Seleção das regiões

Um dos problemas encontrados durante a fase de desenvolvimento, que ocorre em paralelo com as pesquisas sobre as fases posteriores, foi determinar o critério a ser utilizado para a seleção das regiões de importância. Inicialmente, reservaram-se as possibilidades de seleção por semelhança e aproximação de cores, e pelo tamanho e posição da região, este último, determinado pelo maior objetivo detectado no centro da imagem. Demais técnicas e critérios de seleção podem ser utilizados, dependendo da situação, aplicação e ambiente utilizado no processo de captura de imagem.

3.2.5. Detecção de bordas (contornos)

De posse da imagem filtrada, com as informações sem importância de demais ruídos devidamente eliminados e com as regiões de importância respectivamente selecionadas, inicia-se a fase de detecção de bordas ou contornos.

Tendo como matéria-prima uma imagem binarizada, a extração do contorno é realizado lançando-se mão de uma técnica conhecida como “algoritmo do ceguinho”, descrita em [Costa e Cesar \(2001\)](#).

Optou-se pela utilização da conectividade (chain-code) do tipo 8, ou seja, análise dos oitos pontos componentes da vizinhança, e através desta operação, alcançar a representação vetorial do objeto.

3.2.6. Geração e armazenamento da cadeia de pontos

Para o armazenamento dos dados obtidos na fase anterior, o método escolhido foi a criação de uma lista composta diversas filas. Estas filas, por sua vez, compõem-se da cadeia de pontos geradas pelo algoritmo de extração de contorno.

Cada fila apresentará tamanho variado, dependendo do tamanho do objeto ou do segmento selecionado na imagem. A lista, entretanto, com as respectivas filas, terá tamanho conhecido e fixado pela quantidade de imagens analisadas. Quanto maior o nível de precisão necessária ao objeto final, maior será a quantidade de

imagens analisadas, conseqüentemente, o tamanho da lista de cadeias acompanhará tal crescimento.

3.2.7. Criação da nuvem de pontos

Subentende-se que, para a reconstrução completa do objeto, são necessárias outras etapas posteriores, tais como, triangulação (ligação entre os pontos), suavização das faces e finalmente a criação da superfície propriamente dita. Entretanto, existem inúmeras ferramentas e aplicativos comerciais e gratuitos específicos para esta finalidade. De posse da nuvem de pontos, foco primordial e objetivo principal deste trabalho, o usuário dispõe de diversos recursos para a finalização deste procedimento.

Inicialmente, tinha-se a intenção de realizar esta tarefa através da união de todas as cadeias de pontos X,Y referentes às imagens 2D, baseando-se na posição espacial da câmera no momento da captura da imagem, chegando assim, a um conjunto de pontos suficientes para a reconstrução da superfície do objeto 3D. Entretanto, durante o estudo de tal método, percebeu-se algumas restrições, que exigiram o experimento de outras técnicas que serão abordadas a seguir.

Após diversos estudos e análises das técnicas de criação de objetos 3D, seja para obtenção da nuvem de pontos, seja para reconstrução de sua superfície, uma das fases de testes e experimentações foi direcionada à uma técnica não convencional, onde praticamente tornou-se desnecessária a utilização de um modelo matemático.

Primeiramente, um dos modelos experimentados baseia-se no processo inverso ao realizado na etapa de captura. Em outras palavras, tal processo ocorre através da projeção da cadeia de pontos, obtidos pelo contorno de uma imagem, num plano posicionado espacialmente referente à posição da câmera no momento da captura da imagem.

Esta reversão da imagem capturada para a projeção dos pontos de seu contorno busca alcançar maior fidelidade na sua representação, adotando valores

espaciais idênticos em ambas as fases do processo. Deve-se ter em mente, portanto, que no momento da captura, outros dados além da imagem devem ser armazenados, tal como a posição espacial do dispositivo de captura em relação ao centro do objeto, determinada pelo eixo X,Y,Z.

Uma outra informação, não menos importante, porém determinada por um valor constante, é a distância física entre a câmera e o objeto. Utilizando-se como apoio o equipamento de captura e manipulação da câmera desenvolvido especificamente para este projeto, tem-se a distância facilmente calculada ou obtida pelos raios do arcos do dispositivo, tanto horizontal quanto vertical. Atente-se para o fato de que a distância física, neste caso, nem sempre coincide com a distância focal. A possível diferença entre esses valores não interfere no resultado, visto que somente a distância física, já determinada como constante, terá papel fundamental no processo de reconstrução do objeto capturado.

Conhecendo o valor de R (raio ou distância), a matriz X,Y,Z de cada momento de captura, além da cadeia de pontos do contorno do objeto de cada imagem, torna-se possível o reposicionamento dos dados coletados num espaço tridimensional.

4. Aquisição de forma 3D utilizando silhuetas

Após diversos experimentos, análise de resultados das tentativas e erros de algumas das técnicas estudadas e relacionadas acima, encontra-se o método que se demonstra a melhor opção para o cumprimento dos objetivos deste projeto. A aquisição de formas e objetos 3D utilizando-se de silhuetas extraídas de imagens, aborda outros assuntos relacionados, tais como distância focal, projeção volumétrica, envoltória visual, além de termos técnicos como voxel, octree, entre outros, os quais serão descritos detalhadamente a seguir.

4.1. Definição de silhueta

Resumidamente, uma silhueta ([Figura 16](#)) é a visão de algum objeto ou cena, que consiste no esboço de um contorno, desprezando informações em seu interior.

Partindo de uma definição linguística, encontrada em dicionários, pode-se afirmar que uma silhueta consiste em um desenho de perfil que representa as linhas do contorno de um corpo, pessoa ou objeto, ou ainda, de sua sombra projetada.

Relacionada a outros assuntos, técnicos, históricos e específicos da computação gráfica, esta definição alcança outras interpretações.

Históricamente, a origem da palavra silhueta é uma referência a Etienne de Silhouette, ministro das finanças de Louis XV que em 1759 determinou impostos áspers sobre o povo francês, e nesta época se tornou sinônimo de coisas baratas e simples, feita ou construída com poucos recursos. Por este motivo, entende-se uma silhueta como a simplificação de um objeto em uma cena, ou ainda, redução das informações referentes a determinada região de uma imagem, mantendo apenas a sua forma ou delineamento exterior.

Pode também ser entendida como uma forma de arte ou estilo de design, existindo, inclusive, filtros específicos para tratamento de imagens e obtenção de resultados neste formato. Em trabalhos gráficos, encontrar uma silhueta consiste em selecionar uma parcela de uma imagem, caracterizada por uma máscara, de modo a separar, por exemplo, um objeto do fundo da cena.



Figura 16 - Exemplos de silhuetas

4.2. Detecção de silhueta

No processo de detecção de silhueta são considerados três passos fundamentais. Em um primeiro passo, é feita uma abordagem de melhoria da segmentação - aqui são considerados filtros morfológicos para tentar corrigir e diminuir ao máximo os erros de segmentação e ruídos na morfologia do objeto.

No segundo passo, considera-se uma abordagem para a detecção das bordas dos objetos, que está diretamente influenciada pelo passo anterior.

No terceiro passo, é feita a coleta e estruturação dos pontos que fazem parte da silhueta do objeto selecionado.

É fundamental que neste passo estejam previstos os tratamentos de ruídos e erros, ainda presentes após os dois passos anteriores. Numa etapa de refinamento da segmentação, deve-se corrigir ou pelo menos diminuir os erros e ruídos, e para isto é utilizada uma abordagem de processamento de imagens baseada em filtros morfológicos que visam corrigir a imagem do objeto segmentado completando pequenas falhas e buracos, além de eliminar regiões isoladas, compostas por poucos pixels. Os filtros morfológicos utilizados foram especificamente o filtro de Dilatação e o filtro de Erosão.

4.3. Representação de silhueta

A representação de silhuetas deste trabalho, utiliza a técnica de "chain codes" já abordadas em experimentações e capítulos anteriores.

Apenas no aspecto de complementação, a lista de códigos é uma representação que traz algumas vantagens, para a aquisição de silhueta, principalmente no que se refere ao armazenamento dos dados obtidos.

Assim como na técnica baseada em contornos e bordas, o formato de armazenamento e representação chain codes é a solução mais prática, rápida e adequada para o trabalho com silhuetas.

4.4. Distância focal

Em termos técnicos de fotografia, Distância Focal (Focal Length) de uma lente é a distância em milímetros, entre o ponto principal de sua unidade secundária e seu Ponto Focal. Quanto maior o número maior será a magnificação e menor será o campo de visão. Em outras palavras, podemos dizer que quanto maior os milímetros de uma lente menor será o ângulo de visão e maior a distância de foco da imagem. Assim como quanto menor os milímetros da lente maior será o ângulo de visão menor será a distância de foco da lente.

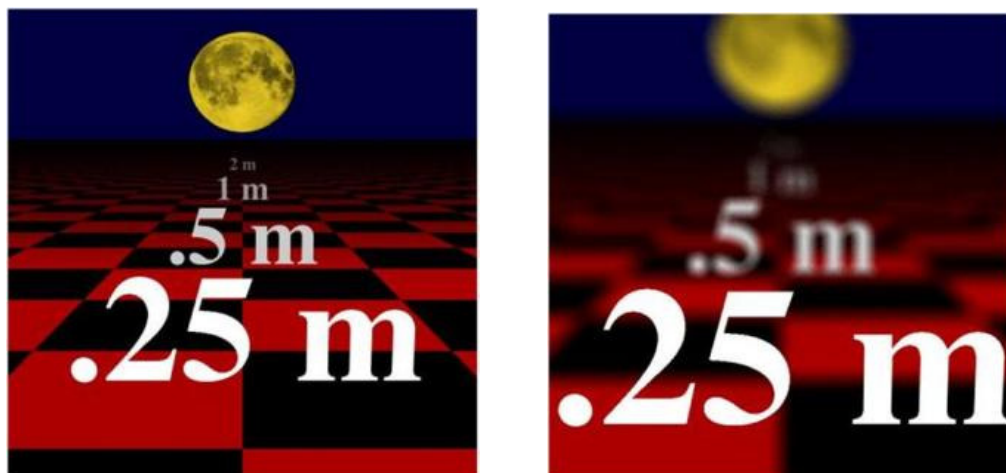


Figura 17 - Exemplo de variação de distância focal

Nas áreas de Computação Gráfica e Visão Computacional, esta definição, mesmo tendo o mesmo significado, apresenta uma aplicação bastante diferente, visto que tal medição tem a finalidade de alimentar variáveis em cálculos e fórmulas matemáticas. De posse de tal valor, é possível realizar inúmeros outros procedimentos, como determinar a profundidade ou posição espacial de um objeto em um espaço tridimensional, além de realizar a calibração de uma câmera, como veremos a seguir.

Em muitas aplicações, é necessário o conhecimento dos parâmetros internos e externos das câmeras com uma certa precisão, para poder usá-los na determinação de posições. A pré-determinação da geometria das câmeras (distância focal, linha de base, centro de perspectiva, fatores de escala) é essencial para identificação da posição (ou distância) de um determinado objeto, a partir da disparidade das suas imagens captadas a partir de diferentes posições e ângulos. A esta pré-determinação dá-se o nome genérico de calibração de câmeras. A precisão pode ser tão acurada quanto se deseje, respeitando os limites das aplicação, do ambiente e do equipamento.

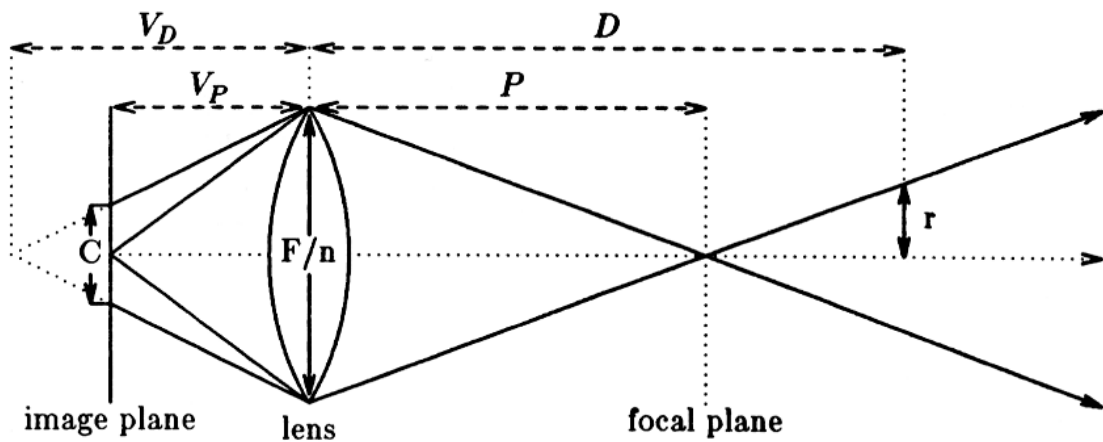


Figura 18 - Geometria da câmera

4.5. Projeção volumétrica das silhuetas

Inicia-se neste ponto o processo de reconstrução do objeto propriamente dito, onde as imagens bidimensionais captadas, filtradas, trabalhadas e com as informações necessárias armazenadas, são novamente reposicionadas em seu espaço tridimensional. A este processo dá-se o nome de Projeção Volumétrica das Silhuetas, ou simplesmente a sigla PVS.

A atividade consiste em, a partir dos dados de contorno da silhueta extraídos das imagens originais, gerar um cone volumétrico com início na posição espacial da câmera, em sentido e direção definidos ao centro do objeto a ser reconstruído (Figura 19).

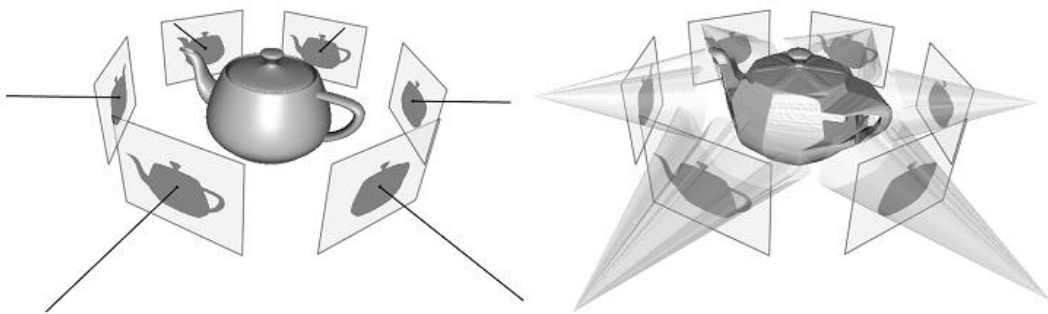


Figura 19 - Projeção volumétrica das silhuetas

Para que esta reposição seja feita de forma adequada, todos os posicionamentos devem estar baseados nas matrizes de projeção das câmeras que são definidas no ato da captura das imagens. Cada imagem está relacionada a uma matriz de projeção (Figura 20), que armazena o ponto de origem da captação.

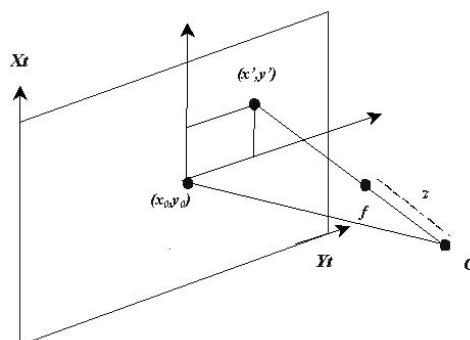


Figura 20 - Representação da matriz de projeção

4.6. Intersecção das PVSs

Após o procedimento de projeção volumétrica das silhuetas, passa-se à fase de detecção das intersecções entre os volumes projetados. Imagina-se que, na prática, um cone volumétrico é responsável pelo “corte” ou “refinamento” de outros cones, e que após todas as interações entre eles, resulte uma massa na região central, composta pelo volume comum a todas as projeções (**Figura 21**).

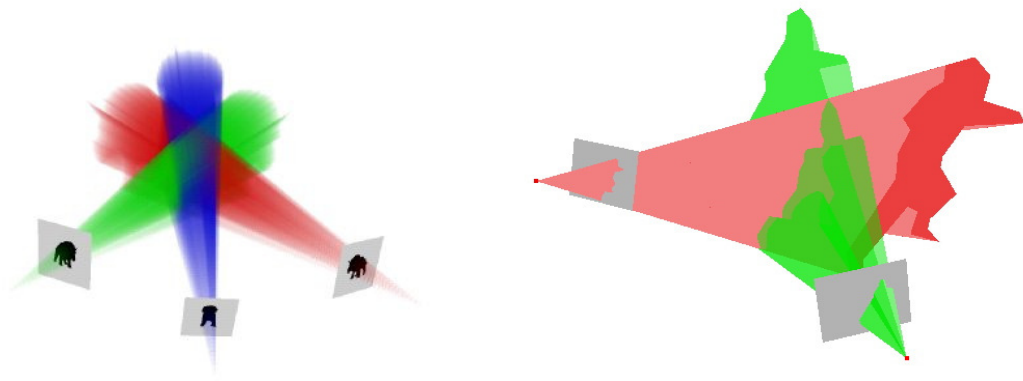


Figura 21 - Intersecção das projeções volumétricas

De certa forma, o processo de reconstrução estaria completado após esta fase, entretanto, existe a necessidade de otimização dos dados obtidos, visto que somente a intersecção volumétrica é insuficiente para a determinação dos vértices e da nuvem de pontos, que posteriormente serão a base da triangulação para reconstrução da superfície.

4.6.1. Em um espaço de voxels

Devido as limitações e restrições apresentadas anteriormente, esta tarefa assume uma fase de preparação e limitação do espaço tridimensional, com o objetivo de conhecimento prévio dos posicionamentos, limites e dimensões.

Após a análise de algumas técnicas e algoritmos existentes, optou-se pela implementação da projeção volumétrica em um espaço de voxels limitado e preparado, com as dimensões previamente conhecidas (**Figura 22**).

Em outras palavras, podemos entender este espaço como um cubo inicial ou uma caixa envolvente, como um bloco inicial de trabalho, que é lapidado a cada projeção do volume de uma silhueta segmentada de uma imagem.

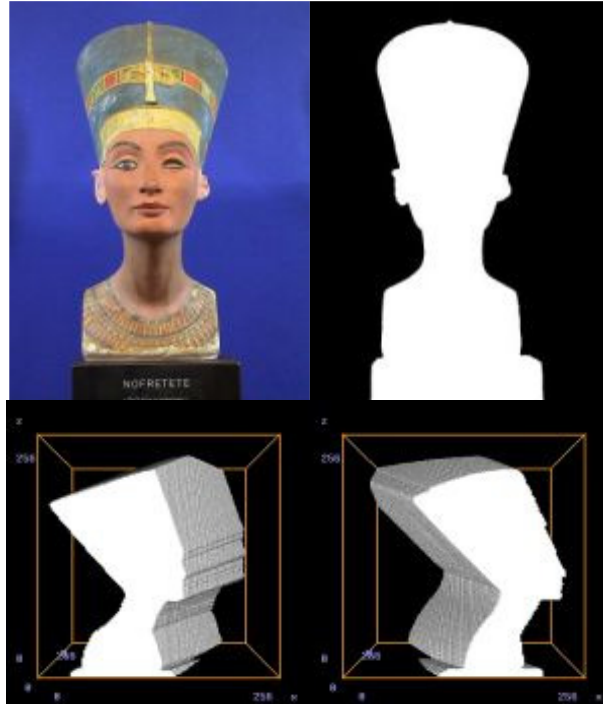


Figura 22 - Inteseccção volumétrica em um espaço de voxels

4.7. Representações do espaço de voxels

Partindo do conceito de pixel (**Figura 23 A**), a unidade mínima da composição de uma imagem digitalizada, entende-se que sua representação pode ser efetuada com o uso de duas informações numéricas referentes à sua posição nos eixos X e Y de um espaço bidimensional. Ampliando tal definição, e levando este conceito a um espaço tridimensional, chegamos ao voxel (**Figura 23 B**), que consiste num pixel com informação adicional de profundidade, no eixo Z.

A um conjunto de voxels dispostos em uma matriz tridimensional, dá-se o nome de “espaço de voxels” (**Figura 23 C**). Existem diversas formas de representação do espaço de voxels, especificamente neste caso, os voxels resultantes do processo de intersecção dos volumes. Os métodos mais conhecidos são a representação de voxels com uso de quadtree e octree, este último

selecionado para implementação neste projeto, por conter um bom balanceamento entre compactação das informações sem grandes perdas de qualidade e definição do objeto representado.

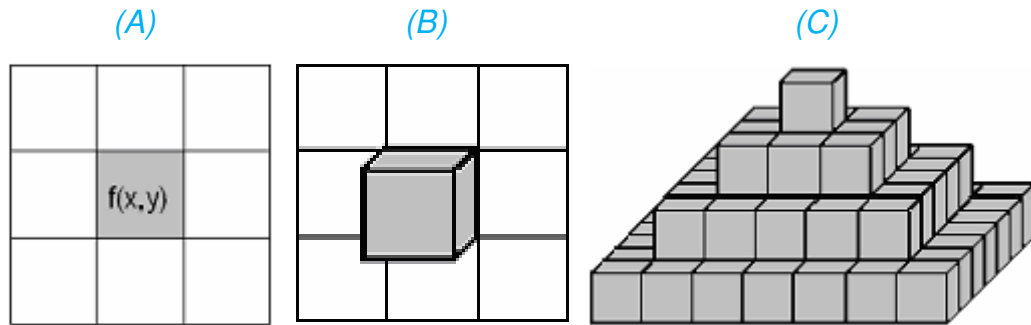


Figura 23 - Representações de pixel (A), voxel (B) e espaço de voxels (C).

4.7.1. Octree

Há numerosas técnicas de estruturação de dados hierárquicos usadas para representar dados no espaço 3D. Uma técnica comumente usada é o quadtree (Figura 24 - Nível 1), e seu desenvolvimento foi motivado em grande parte por uma necessidade de economizar armazenamento agregando dados que têm valores idênticos ou semelhantes. Quando nos referimos a informações bidimensionais, o conceito de quadtree é melhor trabalhado, entretanto, se a intenção é o processo de dados tridimensionais, torna-se necessário o salto para uma variação deste conceito, o denominado octree (Figura 24 - Nível 2).

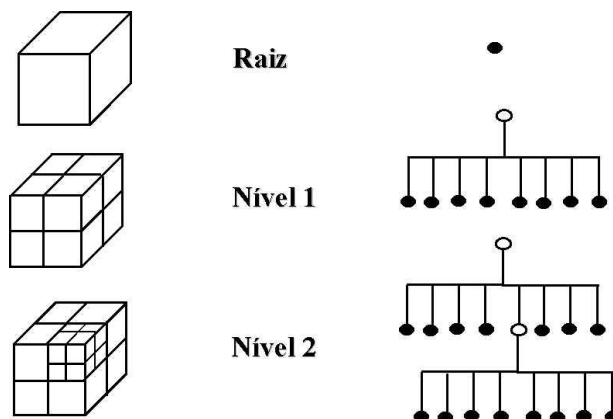


Figura 24 - Quadtree (Nível 1) e Octree (Nível 2)

As octrees são geradas a partir de uma única célula raiz, que através de um algoritmo de refinamento recursivo constrói o restante da árvore, de acordo com a necessidade e do nível de detalhes da representação.

Cada célula ou pequeno cubo resultante da compactação em octree pode ser futuramente convertido em faces de polígonos ou até mesmo em vértices para triangulação, e assim, alcançando diversas formas para representação de superfícies com maior ou menor nível de detalhes e eventual suavização.

Se houver necessidade de maior detalhamento do objeto, sempre há a possibilidade de dividir as células em oito células-filhas, as quais serão submetidas a novo processo de refinamento.

Após todas as etapas de divisão e suavização, somente as células completamente internas ao espaço de voxels previamente definido ficam determinadas como parte integrante do modelo a ser representado. As demais, portanto, são desconsideradas (**Figura 25**).

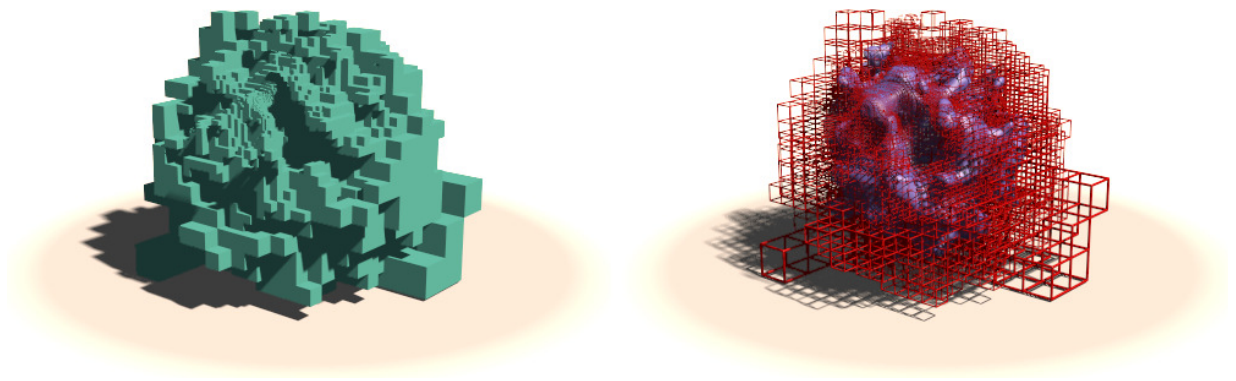


Figura 25 - Representação de superfície com octree

5. Objetos 3D em ambiente RV e RA

Para vislumbrarmos um entendimento geral sobre realidade virtual e aumentada, primeiramente precisamos conceituar alguns fundamentos básicos desta área de conhecimento, abstraindo temporariamente as informações técnicas, voltando-se simplesmente para o modelo de comunicação utilizado pelo ser humano. Os homens são capazes de se comunicar entre si lançando mão de códigos, signos e sinais convencionados, que carregam consigo significado e significância, e que quando decodificados corretamente, traduzem-se em idéias e objetos do mundo real. Tal modelo de comunicação pode assumir diferentes tipos e variações, desde a linguagem verbal, que se manifesta através das formas oral e escrita, até a gestual, artística, inclusive outros tipos de linguagem como a de computador, a matemática e a música.

A grande maioria da comunicação ocorre através da exploração dos sentidos e da percepção humana, seja através do tato, paladar, mas principalmente pela audição e visão. Toda mensagem recebida passa por inúmeros processos de interpretação e compreensão até alcançar a fase do pleno entendimento.

Tanto a realidade virtual e a realidade aumentada aproveitam-se deste modelo de comunicação, explorando todos os sentidos do homem para despertar a percepção de um mundo real através de símbolos visuais e sonoros, gerados por computador, que ocupam um espaço tridimensional e possibilitam a interação entre si, e entre os integrantes deste ambiente. Tal recurso expande os limites impostos pela comunicação visual em duas dimensões, aproximando-se do modelo real de comunicação e interação com o ambiente do mundo real, transformando essa linguagem numa interface natural entre homem e máquina.

5.1. Ambientes de RV e RA

Os ambientes de Realidade Virtual e Realidade Aumentada, como já discutidos, são a representação simbólica do mundo real, com a participação de recursos computacionais. Mas existem algumas diferenças importantes que devem

ser destacadas, principalmente sob o ponto de vista de conceituação de cada modelo.

Realidade Virtual, de uma forma geral, é a criação de um ambiente tridimensional imaginário, com a transferência e representações de objetos do mundo real, possibilitando interação e navegabilidade. Qualquer situação pode ser representada neste mundo virtual, até mesmo com diferentes variações e distorções da realidade. Não existe limites para esta representação, e as leis naturais nem sempre são transferidas juntamente com os cenários e objetos. É o caso dos mundos virtuais onde os membros participantes podem voar, ou os objetos ocupam um mesmo local no espaço.

Segundo **Santin (2008)**, “a realidade virtual (RV) tem se tornado cada vez mais popular na atualidade, devido à queda dos custos de seus equipamentos. Desde a sua criação por volta década de 50, até meados da década de 80, a RV era restrita a algumas instituições altamente equipadas, visto os proibitivos custos dos hardwares utilizados para a sua execução. No final da década de 80, a RV passou a ser viável às empresas e universidades, devido aos avanços na arquitetura dos PCs e ao aumento na demanda de dispositivos de RV, oferecendo, no mercado, sistemas capazes de desenvolver e executar aplicações de realidade virtual a custos mais acessíveis. A popularização da RV ressaltou o uso dessa interface em diversos projetos de pesquisa, o que possibilitou sua propagação a diversas áreas do conhecimento. Devido à sua vasta possibilidade de aplicações, a RV é utilizada no auxílio à venda de produtos, no tratamento de alguns distúrbios psicológicos, no benefício ao aprendizado, seja de funcionários de alguma empresa, alunos das instituições de ensino, ou, até mesmo, de pessoas deficientes, etc.”

Em contra-partida, a realidade aumentada (RA) é uma tecnologia na qual os estímulos sintéticos, ou seja, os elementos gerados por computador, utilizam-se do ambiente real para sua estabilização. O cenário básico, portanto, não necessita de processamento ou geração, visto que o solo, as paredes, o teto e as paisagens, são as mesmas com as quais convivemos na realidade, e apenas alguns objetos adicionais são somados e incluídos à esta situação. De uma forma geral, a RA é uma melhoria do mundo real através de um mundo virtual, variando sua complexidade tanto na quantidade quanto na qualidade dos objetos adicionados, possibilitando a coexistência entre os mundos real e virtual.

Kirner (2006) acrescenta que “realidade aumentada é a inserção de objetos virtuais no ambiente físico, mostrada ao usuário, em tempo real, com o apoio de algum dispositivo tecnológico, usando a interface do ambiente real, adaptada para visualizar e manipular os objetos reais e virtuais”.

5.2. Objetos 3D para Realidade Virtual

A realidade virtual se manifesta através de elementos distintos, tais como os cenários, que podem ser classificados como a parte estática dos mundos virtuais, e pelos objetos 3D que compoem a parte dinâmica e em conjunto com os cenários proporcionam a interatividade com os usuários do espaço tridimensional.

Existe a possibilidade de criação de um mundo virtual, composto apenas por cenários, onde a principal função seria a ambientalização. Neste caso os objetos ou modelos 3D poderiam ser descartados, caso não tivessem uma função específica ou primordial. Mas, mesmo sem a necessidade de objetos numa cena, os próprios cenários são modelos tridimensionais, pois possuem as mesmas características de um objeto 3D e são criados com ferramentas e técnicas semelhantes.

Portanto, de modo geral, na realidade virtual tanto os objetos da cena quanto os cenários podem ser classificados como modelos 3D (**Figura 26**), o que não ocorre com a Realidade Aumentada.



Figura 26 - Realidade Virtual – Cenário e Objetos 3D

5.3. Objetos 3D para Realidade Aumentada

Diferente do que acontece na realidade virtual, na realidade aumentada o cenário não é composto por elementos virtuais ou modelos tridimensionais. Nesta situação o cenário é o próprio mundo real, e apenas os itens adicionados a esta cena são gerados por computador e classificados como objetos ou modelos 3D (Figura 27).

A grande preocupação neste caso está relacionada com a qualidade visual dos objetos tridimensionais, o nível de detalhamento de sua superfície e a perfeição de sua colorização ou textura. Quanto melhor a qualidade dos objetos virtuais, melhor será a sua fusão com o mundo real, e quanto mais imperceptível for esta mistura, melhor será a percepção visual do usuário.



Figura 27 - Realidade Aumentada – Objetos 3D adicionados ao mundo real

5.4. Simplificadores LOD

Para a otimização e aceleração do processamento dos objetos 3D, existe a necessidade de redução das informações sobre a superfície dos modelos, e isto pode ser realizado através da simplificação do nível de detalhes (LOD – Level Of Details). Tal redução deve ser efetuada sem grande perda da qualidade de representação desta superfície.

Para um melhor balanço entre processamento e detalhamento dos objetos, na maioria dos casos onde existe esta preocupação, várias representações de um mesmo objeto são utilizadas, dependendo apenas do distanciamento visual do usuário. Quanto mais longe do objeto, menor a necessidade de detalhamento de sua representação. Ao se aproximar deste modelo, o usuário precisa de mais informações sobre a superfície e a textura.

Para facilitar nesta tarefa, existem alguns softwares específicos para esta finalidade, que conseguem ótimos resultados, reduzindo drasticamente a quantidade de informações sem necessariamente reduzir a qualidade visual de um modelo tridimensional.

5.4.1. VizUp Optimizer

VizUp (**Figura 28**) é uma ferramenta revolucionária, designada a melhorar a velocidade de renderização e qualquer modelo 3D, reduzindo consideravelmente a quantidade de polígonos. O software trabalha com vários tipos de arquivos como CAD, CAM, WRL, entre outros.

Com o auxílio do VizUp é possível alcançar o equilíbrio entre o nível de detalhes necessários e a complexidade de modelos tridimensionais, podendo utilizá-lo para aplicações em realidade virtual, aumentada e outros sistemas de visualização em tempo real. No processo de simplificação, o usuário pode também determinar a porcentagem na redução dos detalhes.

A versão profissional permite reduzir o número de polígonos de um modelo 3D complexo, sem alterar a qualidade e aparência em relação ao original. O resultado final da redução é quase que o mesmo do inicial, sem a necessidade de conhecimentos avançados para a sua utilização, tornando-o uma ferramenta perfeita e indispensável para usuários de realidade virtual.

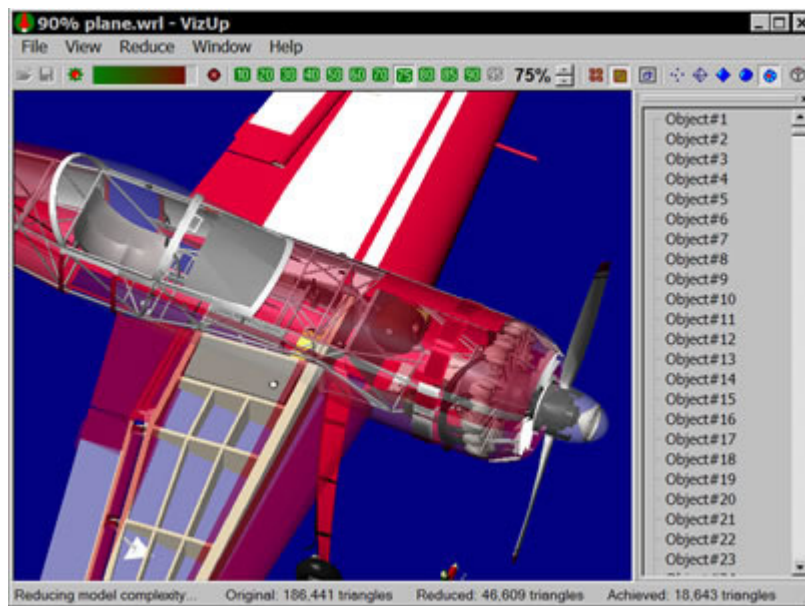


Figura 28 - Software VizUp – Simplificador de polígonos

6. Desenvolvimento

Os maiores obstáculos encontram-se nas fases de tratamento de imagem e criação do modelo tridimensional. Após a captação, todas as imagens devem passar por um processo que consiste dos seguintes passos: (1) reduzir as informações sem importância, (2) separar os objetos relevantes, e (3) identificar as silhuetas. Estas tarefas são realizadas com técnicas de processamento de imagens e visão computacional aplicadas sucessivamente sobre os resultados das etapas anteriores.

Após encontrados todos os contornos de todas as imagens, nos mais diferentes ângulos de visão, aliado ao conhecimento da respectiva posição espacial da câmera no momento da aquisição da imagem, é possível criar uma malha inicial, que consiste no primeiro passo para a reconstrução tridimensional da superfície do objeto desejado.

Inúmeras outras pesquisas buscaram resultados semelhantes, entretanto, utilizaram-se das mais diversas técnicas. Algumas se fundamentaram na criação de novos equipamentos, ou se aproveitaram de outros já existentes, utilizando-os de maneira adaptada. É o caso do escaneamento de objetos 3D com raio laser ou ultrassom (D'Agnoluzzo, 2005). Ainda nesta mesma linha, podemos citar a técnica de projeção de luz estruturada (Reiss, 2003), que consiste na aplicação de projetores e câmeras fotográficas de alta resolução, identificando-se um conjunto de pontos da superfície do objeto a ser reconstruído, conjunto este conhecido como nuvem de pontos, contendo informações da posição em 3D de cada um dos pontos.

6.1. Resultados iniciais

Em princípio, almeja-se disponibilizar o produto final deste projeto sob forma de software ou biblioteca de funções, para uso como auxílio à programação de novas aplicações. Desta forma, tal recurso estará disponível tanto para criação de softwares específicos e avançados quanto para uso amador.

Em ação conjunta com o Núcleo de Gestão da Qualidade e Metrologia (NGQM) da UNIMEP (Campus Santa Bárbara do Oeste), os estudos estão voltados para a utilização de um equipamento, em fase de desenvolvimento de protótipo, que

fará a aquisição das imagens com a câmera em poses e posições precisamente determinadas (Figura 29). Tal equipamento consiste em uma webcam móvel, fixada em um arco que realiza movimentos de translação e rotação, a fim de obter uma superamostragem de posições e poses do objeto para a reconstrução 3D do objeto.

Convém ressaltar que este trabalho está relacionado apenas ao desenvolvimento do software de tratamento das imagens e criação do modelo tridimensional, bem como visualização do modelo criado. Todos os demais itens, como o dispositivo citado anteriormente, são responsabilidades de outras partes integrantes desta mesma pesquisa.

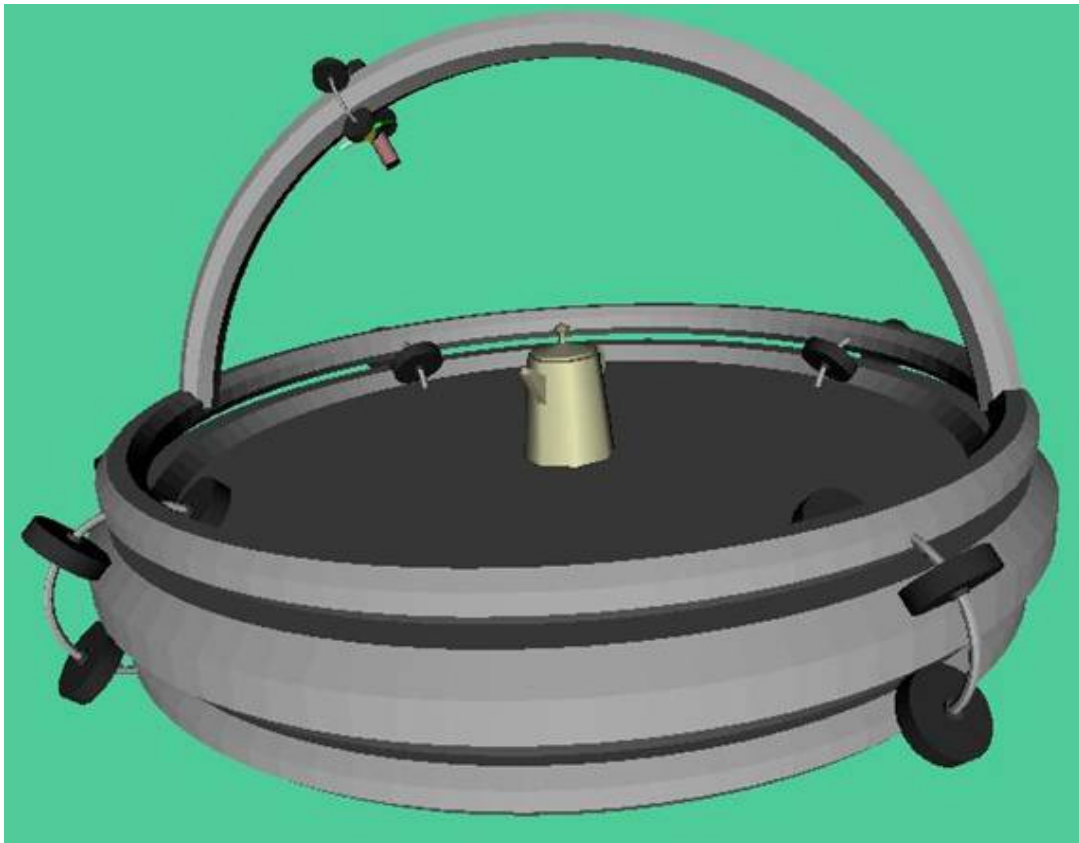


Figura 29 - Protótipo virtual do dispositivo de captação de imagens

Após o processo de aquisição das imagens, de seu processamento digital e conversão dos dados obtidos em posições espaciais torna-se possível a construção de uma malha proveniente dos cruzamentos entre os pontos obtidos. O produto final destes procedimentos é a representação virtual e digital da superfície do objeto escaneado.

Para um melhor aproveitamento do modelo tridimensional gerado, uma das preocupações é a compatibilidade. A análise das ferramentas de manipulação e modelagem de objetos 3D existentes atualmente no mercado, bem como os tipos de arquivos acessíveis em cada uma delas, indica a escolha do formato padrão WaveFront OBJ. Este tipo de arquivo, vastamente utilizado em aplicações CAD, caso não sejam um tipo padrão de alguma das ferramentas, na grande maioria dos casos possui recursos de importação ou conversão para a estrutura e formato adequados.

Uma segunda e não menos importante opção é a geração do modelo tridimensional em formato WRL, visto que grande parte dos casos de uso prevista está relacionada às atividades baseadas em Realidade Virtual e Aumentada. A escolha do tipo de arquivo utilizado no processo de criação amplia ainda mais as possibilidades de aceitação e utilização deste projeto.

6.2. Experimentações

No intuito de otimizar o conteúdo obtido durante o processo de pesquisa, foram realizados em paralelo alguns experimentos, com o objetivo de reconhecer o resultado das diversas técnicas estudadas, bem como, analisar o comportamento de cada método perante sua performance e qualidade do produto final obtido.

Tal metodologia auxiliou na escolha dos caminhos a serem seguidos, possibilitando o embasamento teórico e prático para tais atitudes. Um exemplo claro é a seleção da limiarização como filtro suficiente durante o procedimento de segmentação da imagem.

Outros algoritmos e soluções alternativas demonstraram-se ineficazes ou inapropriadas para este trabalho. A figura a seguir (**Figura 30**), apresenta um momento específico da fase de experimentações, onde o autor do projeto realiza testes com o filtro Sobel, num processamento de imagens em tempo real.

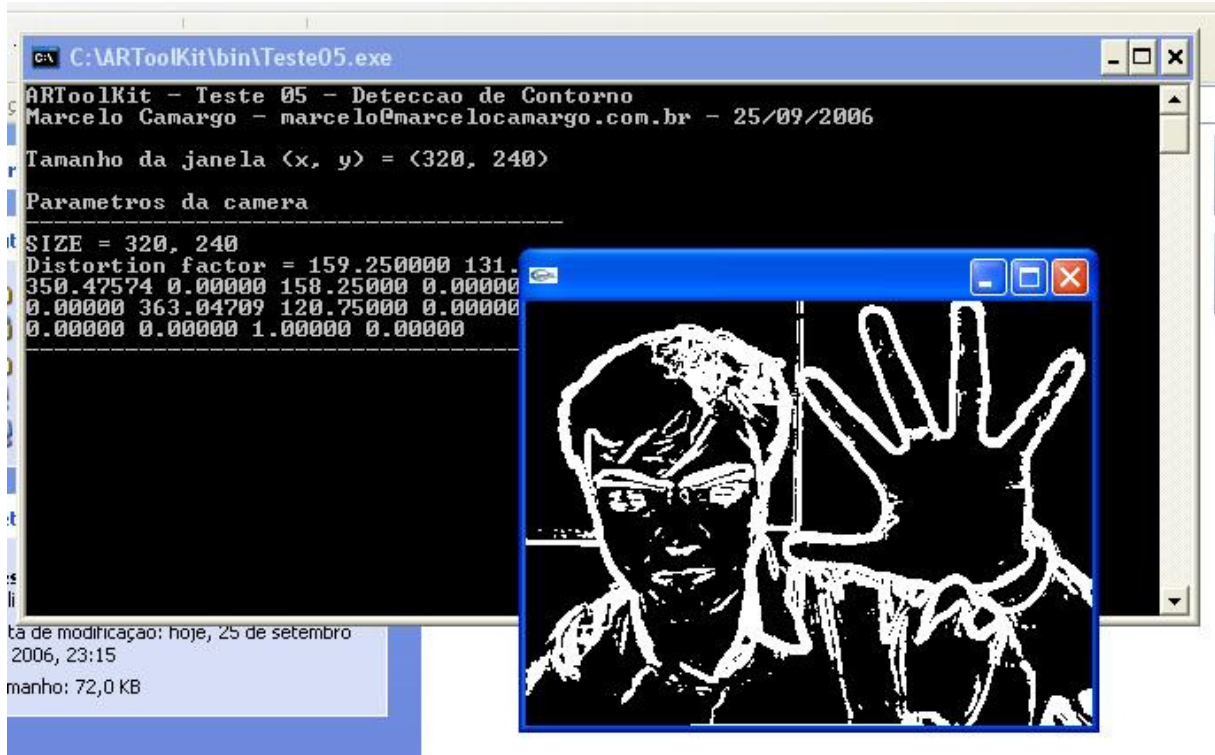


Figura 30 - Experimento realizado com filtro Sobel

6.3. Captura de cenas virtuais

No intuito de aproximar os trabalhos desenvolvidos na fase de experimentos, com a aplicação final do projeto - que neste caso é o funcionamento baseado em imagens obtidas a partir de um equipamento específico – tornou-se necessária a criação de um simulador do elemento mecânico responsável pela captação das imagens.

A versão virtual do dispositivo foi desenvolvida em linguagem VRML, com possibilidade de movimentação do rotor e da grua, além de visualização à partir de uma câmera posicionada no ponto central destas movimentações.

Assim, torna-se possível a realização de experimentos mais aprofundados sem a necessidade de utilização do sistema físico. Eventuais dificuldades podem surgir neste período, favorecendo a criação de soluções antes mesmo do aparecimento dos problemas.

6.3.1. Cortona ActiveX

A primeira tentativa de integração do simulador em VRML com outra linguagem de programação, deu-se através da biblioteca OpenVRML, demonstrando certa dificuldade no processo de comunicação entre os sistemas.

Uma rápida solução foi a utilização do SDK disponibilizado pelo plug-in Cortona, comercializado e distribuído pela empresa ParallelGraphics. Tal produto é encontrado na forma de ActiveX e acompanha a versão gratuita do Cortona.

A grande dificuldade, neste caso, foi a inexistência de documentação ou exemplos de uso, visto que, tais informações encontram-se disponíveis apenas na versão comercial do produto.

6.3.2. Testes de execução

A figura abaixo (**Figura 31**) apresenta um momento do período de testes de execução da ferramenta de simulação de captura de imagens.

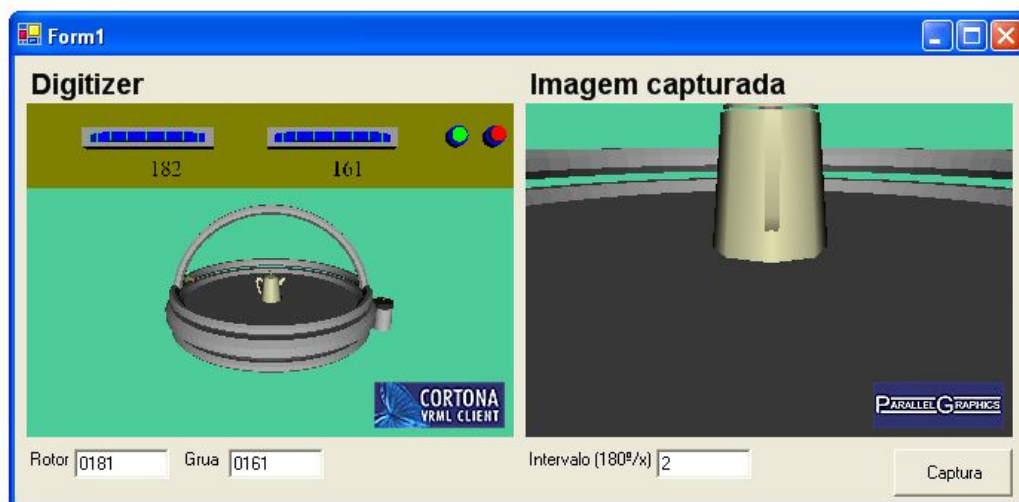


Figura 31 - Simulação de captura de imagens

Pode-se observar, no funcionamento do simulador, à esquerda da tela principal da aplicação, a imagem do dispositivo de captura, e à direita desta mesma tela, a visão da câmera, imagem esta que é gravada em disco para futuro processamento.

Os campos “Rotor” e “Grua” informam a posição atual da câmera no dispositivo, e tais dados são utilizados para nomeação dos arquivos de armazenamento das imagens em disco. Cada imagem capturada é nomeada de acordo com a posição da câmera, e isto possibilitará o reconhecimento espacial dos pontos extraídos de cada imagem.

O campo “Intervalo” é informado pelo usuário e determina a quantidade de graus necessários na movimentação do equipamento para a realização de uma captura de imagem. Em outras palavras, determina o número de imagens a serem capturadas, conseqüentemente a qualidade do objeto final.

A seguir (Figura 32), temos uma relação que exemplifica as imagens capturadas pelo simulador.

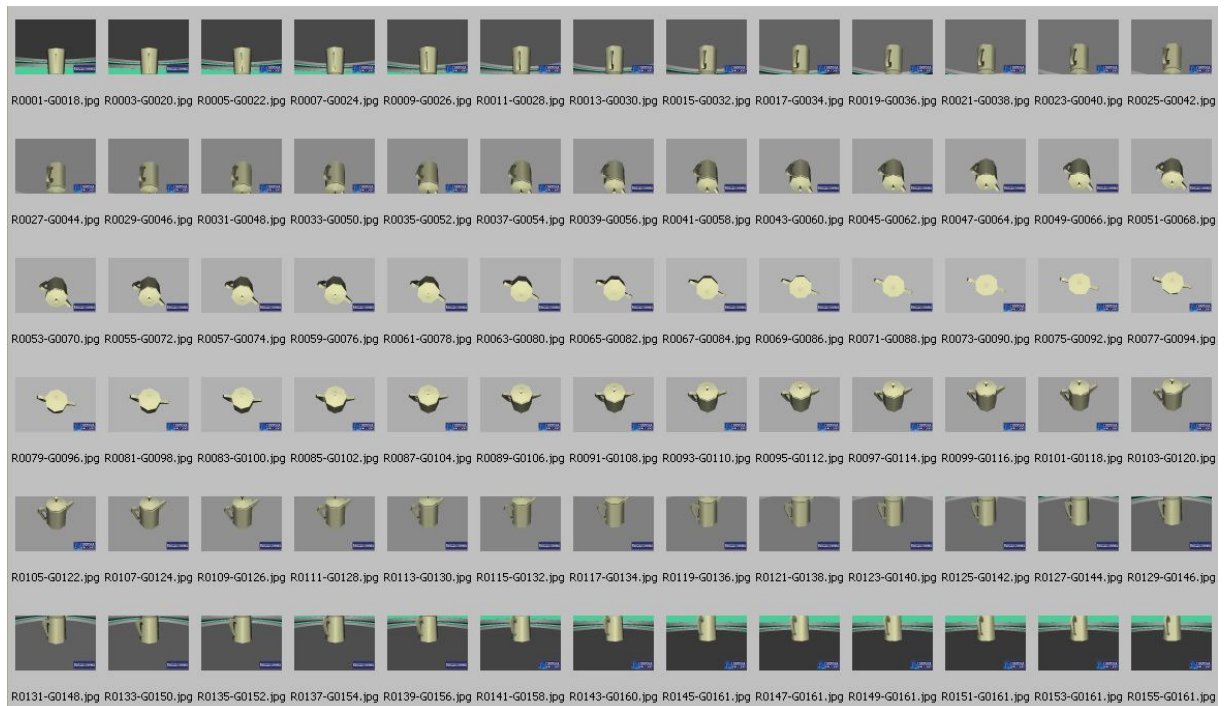


Figura 32 - Miniaturas das imagens capturadas pelo simulador

6.4. Aplicação para geração de objetos 3D

A aplicação para geração de objetos 3D pode ser classificada como a parte mais importante deste projeto, visto que todos os conhecimentos anteriormente obtidos, todas as pesquisas realizadas e todas as análises efetuadas, desembocam na criação desta ferramenta. Denominado “Scanner 3D”, este software contém as

técnicas mais adequadas dentre toda a teoria experimentada, e procura oferecer ao usuário uma fácil utilização, objetividade e rápida aprendizagem.

6.4.1. Explicações

Com a finalidade de alcançar os objetivos principais deste trabalho, o Scanner 3D assume o papel principal e a função de transformar as imagens captadas de um objeto real em um modelo 3D o qual poderá ser utilizado em qualquer ambiente tridimensional ou mundo virtual.

Para o seu funcionamento adequado, são necessárias as seguintes informações:

- várias imagens de um objeto real, captadas em diferentes ângulos e posições espaciais. Cada imagem deve ser representada na forma de arquivo gráfico do tipo JPEG. Todos os arquivos devem conter o mesmo tamanho e resolução. Quanto maior a quantidade de imagens, melhor será a qualidade e o detalhamento do resultado final, entretanto, maior também será o tempo de processamento.

- várias matrizes de projeção, uma para cada imagem captada. A matriz de projeção diz respeito à posição espacial da câmera no momento da captação da imagem, em relação a um ponto central e comum a todas as outras matrizes. Somente os arquivos de imagens com arquivos de matrizes relacionados serão processados. A nomenclatura dos arquivos de matrizes de projeção devem acompanhar o nome da imagem, alterando-se apenas a sua extensão, que neste caso deve seguir o padrão <imagem00.pa>, onde PA significa "Projection Array".

- vários arquivos com informações sobre a silhueta do objeto extraída da imagem, composta pelos pontos X,Y de todo o contorno. A sequência deve ter o ciclo fechado, ou seja, iniciar e terminar no mesmo ponto, com todos os pontos participantes conectados aos seus vizinhos.

Todos estes arquivos, de imagens, matrizes de projeção e pontos de contorno da silhueta, são obtidos com o auxílio das outras aplicações integrantes deste projeto, cada qual específica para uma finalidade. Existem também, softwares de outros autores e empresas, que realizam estes procedimentos.

O Scanner3D, portanto, apenas utiliza tais informações para reconstruir o objeto e transformá-lo em um modelo 3D que pode ser exportado e convertido para diversos formatos.

6.4.2. Interface gráfica

Baseando-se no conceito de simplicidade e objetividade, a interface da ferramenta de criação do modelo 3D foi desenvolvida a fim de conter apenas as funções fundamentais, deixando sempre à mão do usuário os recursos necessários à fácil manipulação e rápida aprendizagem no uso da aplicação.

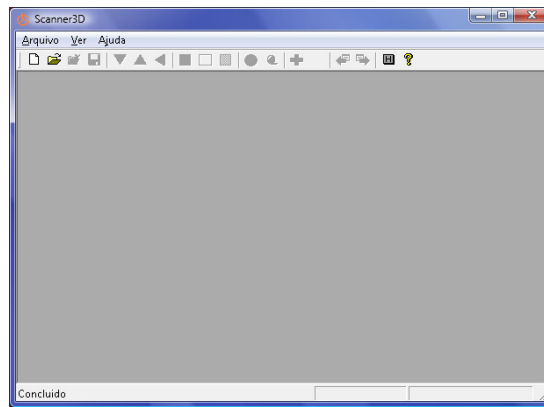


Figura 33 - Tela principal da ferramenta de criação de modelo 3D

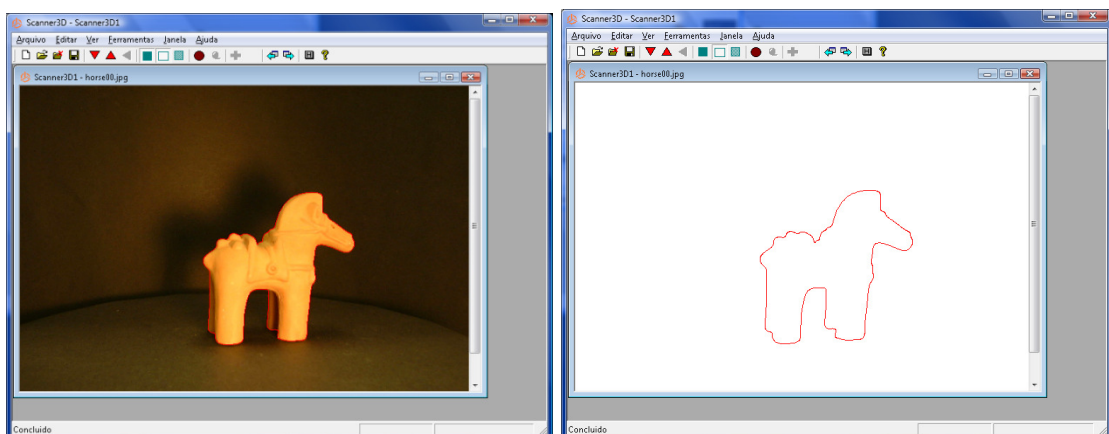


Figura 34 - Telas demonstrando a abertura de imagem e contorno da silhueta

Para agilizar ainda mais os processos de testes e experimentações, a ferramenta disponibiliza a opção de visualização do resultado em tempo real, sem a

necessidade de uso de aplicativos ou ambientes externos para verificar a qualidade dos objetos criados (**Figura 35**).

Desta forma, o usuário pode determinar a quantidade de imagens necessárias para obter um modelo 3D com qualidade satisfatória, ou até mesmo, a número de suavizações e interações dos octrees para alcançar um balanceamento entre qualidade e complexidade dos objetos (**Figura 36**).

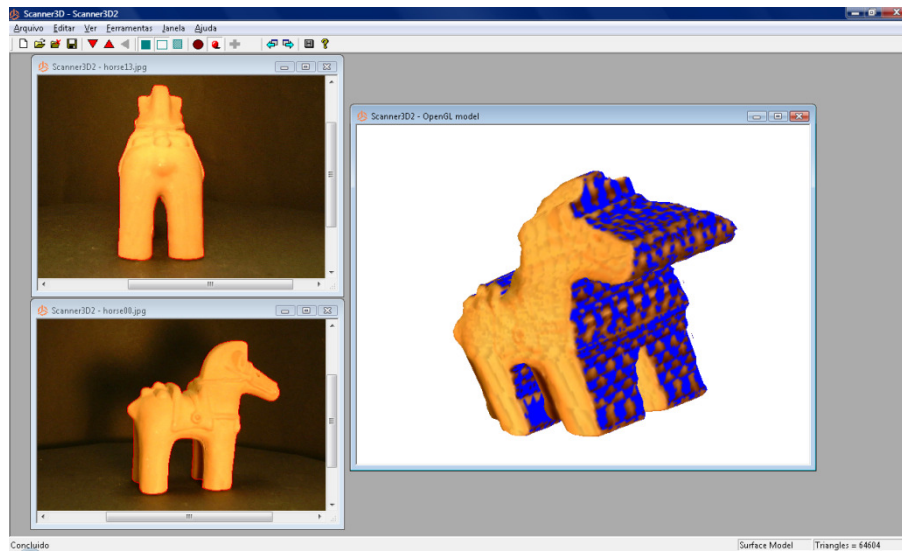


Figura 35 - Criação de modelo 3D baseado na intersecção de apenas 2 imagens

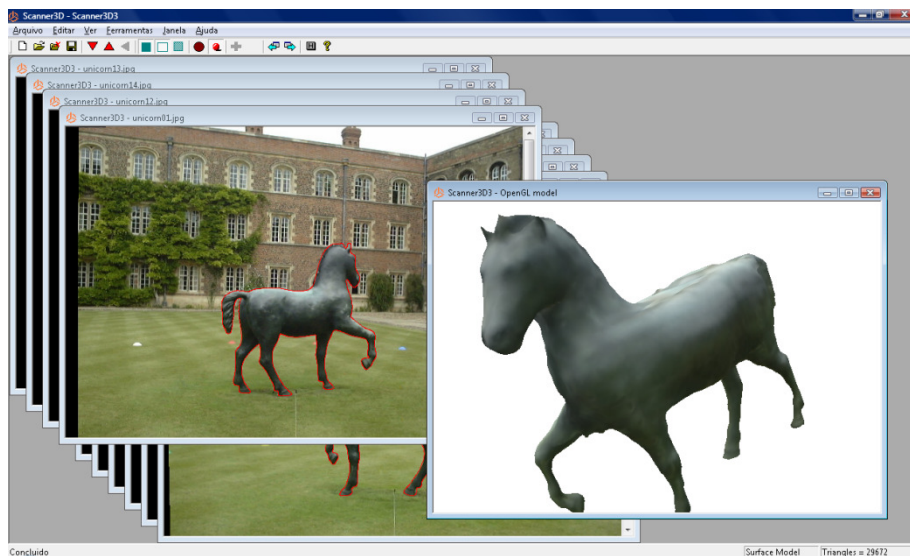


Figura 36 - Modelo 3D criado a partir da intersecção de várias imagens

6.4.3. Datasets

Como o foco principal da pesquisa é a geração dos objetos 3D, optou-se pela utilização de Datasets já preparados, que são conjuntos de arquivos compostos por todos os requisitos expostos anteriormente, ou seja, imagens, matrizes de projeção e contornos das silhuetas.

Estes Datasets foram criados por outros autores e centros de pesquisas em áreas similares, tanto de computação gráfica quanto de realidade virtual. Alguns deles possuíam formatos diferentes e necessitaram de alguns ajustes para total compatibilidade com a ferramenta de geração do objeto 3D.



Figura 37 - Amostra do dataset da estátua de David

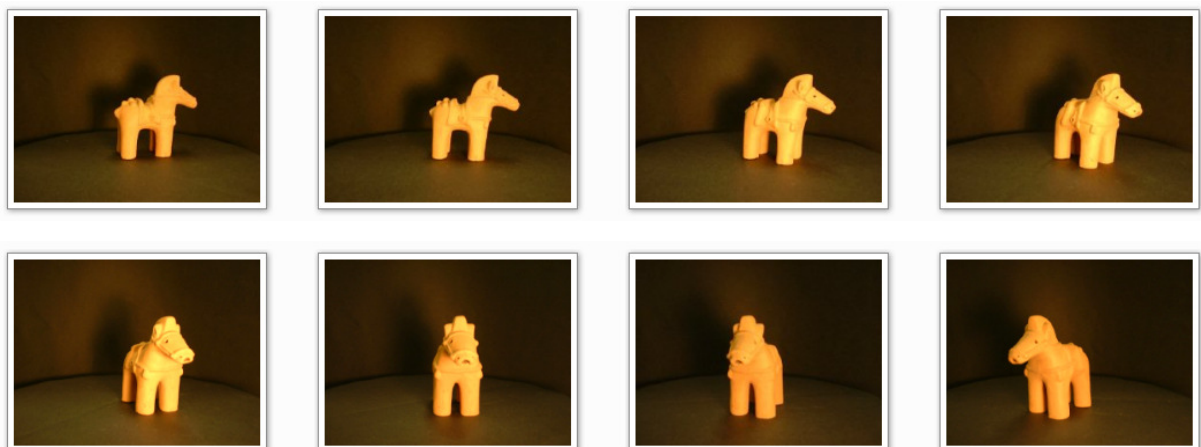


Figura 38 - Amostra do dataset de uma escultura em barro

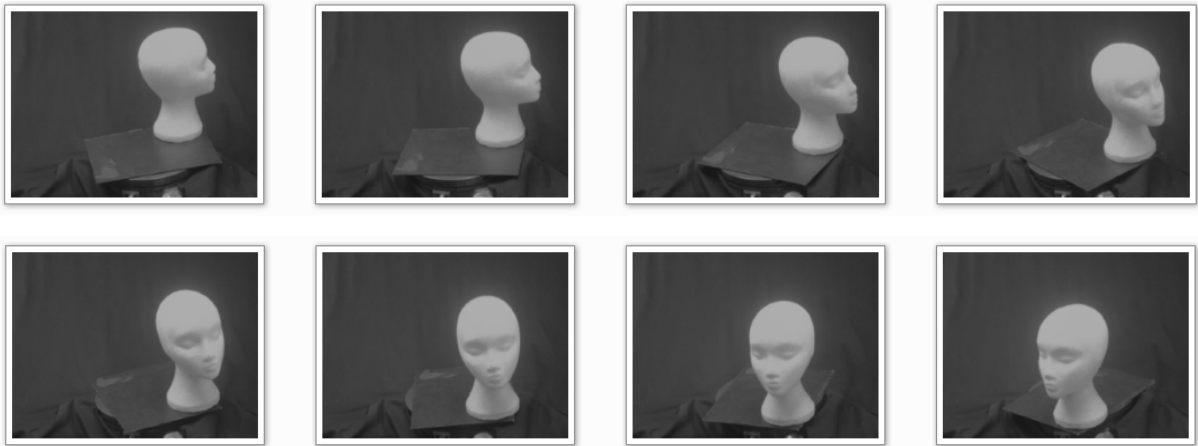


Figura 39 - Amostra do dataset de uma cabeça de manequim

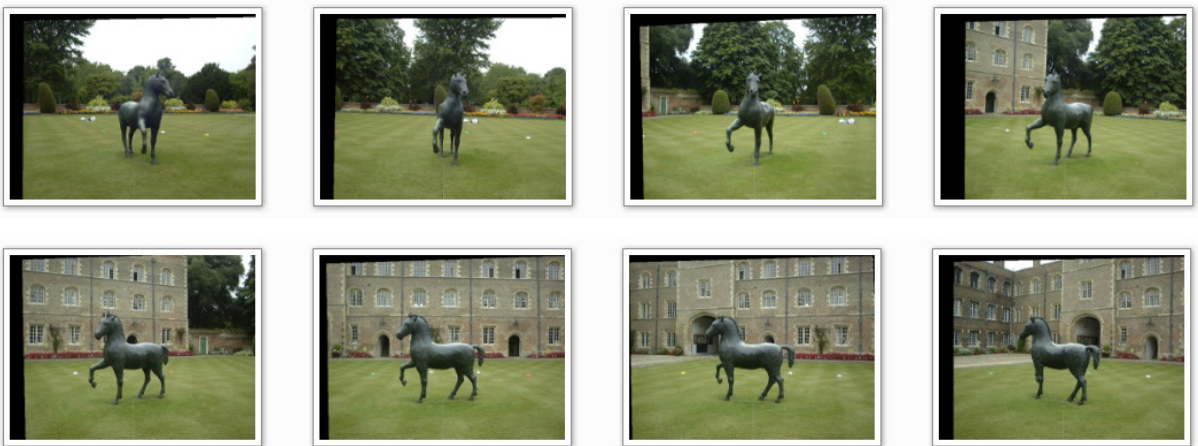


Figura 40 - Amostra do dataset de uma estátua de cavalo

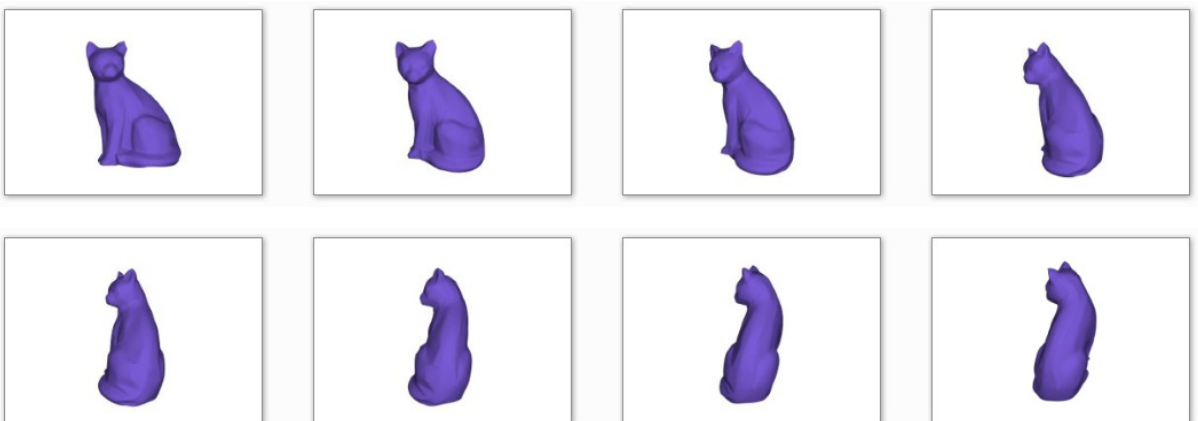


Figura 41 - Amostra do dataset de um gato (objeto 3D)

Um fato que chamou a atenção, foi o nível de complexidade obtido quando do uso do dataset do gato 3D, que mesmo baseado nas mesmas configurações de geração do modelo tridimensional, obteve uma quantidade muito superior de octrees e conseqüentemente, mais triangulações para a representação de sua superfície. Isso se deve ao fato de que as próprias imagens utilizadas como base para a reconstrução possuem melhor qualidade e definição de suas silhuetas.

7. Conclusões

Finalmente, após a conclusão das pesquisas e implementação de algumas ferramentas, principalmente o aplicativo de criação de objetos 3D, determina-se que os objetivos principais e o foco do projeto foram satisfatoriamente alcançados. Em alguns testes realizados, o resultado chegou a ser muito superior ao esperado, comprovando a eficiência das técnicas e métodos selecionados.

7.1. Considerações gerais

Como todo e qualquer outro procedimento computacional que busca resultados em imagens, é importante conhecer a quantidade suficiente de imagens para obter modelos com uma qualidade pré-determinada para o modelo 3D gerado.

Alguns problemas previstos no início do projeto foram solucionados, como a utilização de objetos com superfícies muito irregulares, ou ainda, com áreas de concavidade. Tais detalhes podem ser resolvidos com um número maior de imagens capturadas.

7.2. Contribuições

Acredita-se que, através dos resultados alcançados e comprovações percebidas sobre as técnicas avaliadas, as áreas relacionadas à realidade virtual e aumentada, bem como os setores ligados à computação gráfica se aproximam cada vez mais do usuário doméstico, com a possibilidade de criação de ferramentas de software e hardware capazes de estimular ainda mais o uso destes recursos computacionais.

Este trabalho proporciona também, através da ferramenta de criação de objetos 3D, mesmo em sua primeira versão, auxilia no escaneamento de objetos reais transformando-os em modelos tridimensionais, sem a necessidade de investimentos em complexos softwares ou caros equipamentos atualmente disponíveis no mercado.

Um exemplo disso é a intenção do uso do Scanner 3D para a criação de modelos tridimensionais de figurinos e vestimentas para implantação em um projeto de teatro virtual, onde os atores lançam mão da realidade aumentada para a composição de suas personagens durante a apresentação no palco.

7.3. Resultados

Mesmo com os objetivos principais alcançados, os resultados obtidos ainda assumem uma forma relativamente bruta, necessitando de alguma lapidação e pequenos ajustes para ser entregue como produto acabado ao usuário final. Considera-se portanto adequado para fins de estudo e como comprovação da viabilidade do projeto, inclusive com perspectivas de transformação em ferramenta para distribuição ao uso profissional e acadêmico.

7.4. Trabalhos futuros

Na esteira deste trabalho vislumbra-se a possibilidade de confecção de um equipamento específico e otimizado para auxiliar na captura das imagens e suas respectivas matrizes de projeção da câmera.

Este hardware funcionaria em total sincronia com os softwares deste projeto, criando um ambiente isolado, com condições favoráveis à aceleração dos processos de captura e segmentação das imagens.

8. Referências Bibliográficas

- ANDRADE, José Bittencourt. **Fotogrametria**. Curitiba: SBEE, 1998.
- ARTERO, A. O.; TOMMASELLI, A. M. G. **Limiarização Automática de Imagens Digitais**. Boletim de Ciências Geodésicas. Curitiba: PR, 2000.
- BALLARD, D.H.; BROWN, C.M. **Computer Vision**. New Jersey: Prentice-Hall, 1982.
- BATTLE, J.; MOUADIBB, E.; SALVI, J.; **Recent progress in coded structured light as a technique to solve the correspondence problem: a survey**. *Pattern Recognition*. 1996.
- CESAR JUNIOR, R. M. ; COSTA, L. da F. . **Shape Analysis and Classification: Theory and Practice**. Boca Raton: CRC Press, 2001.
- CHEN Z.T. & TOBLER W.R. **Quadtree representation of digital terrain**, In: **Auto-Carto**. London, England: Proceedings, 1986.
- D'AGNOLUZZO, Vitor Rippel. **Aquisição de modelos 3D por escaneamento a laser**. São Carlos: Universidade de São Paulo, 2005.
- DE FLORIANI L., MARZANO P., PUPPO E. **Multiresolution models for topographic surface description**, *The Visual Computer*. 1996.
- EL-HAKIM, S. F. **Vision-based automated measurement techniques**. In: ATKINSON, K.B. *Close Range Photogrammetry and Machine Vision*. Department of Photogrammetry and Surveying. University College London: Whittles Publishing, 1996.

GONZALEZ, R. C.; WOODS, R. E. **Processamento de Imagens Digitais**, Tradução: Roberto Marcondes César Junior e Luciano da Fontoura Costa. São Paulo: Blücher, 2000.

GUÉZIEC A. & HUMMEL R. **Exploiting triangulated surface extraction using tetrahedral decomposition**. IEEE Transactions on Visualization and Computer Graphics, 1995.

HOPPE H. **Progressive meshes**. In: SIGGRAPH'96 Conference, New Orleans, Louisiana, USA: Proceedings, 1996.

HOPPE H., DEROSE T., DUCHAMP T, MCDONALD J., STUETZLE W. **Mesh optimization**. In: SIGGRAPH'93 Conference, Anaheim, California, USA: Proceedings, 1993.

JAIN, A. K. **Fundamentals of Digital Image Processing**. Prentice Hall, 1989.

JAIN, A. K.; DUIN, R.P.W.; MAO, J. **Statistical Pattern Recognition: A Review**. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000.

KIRNER, C.; TORI, R. **Fundamentos de Realidade Aumentada**. In: Claudio Kirner; Romero Tori; Robson Siscoutto. (Ed.). Fundamentos e Tecnologia de Realidade Virtual e Aumentada. Pré Simpósio SVR 2006, SBC, Belém, 2006.

OTSU, N. **A threshold selection method from gray level histograms**. IEEE Transactions on Systems, Man and Cybernetics, 1979.

PEDRINI H. **An Adaptive Method for Terrain Surface Approximation based on Triangular Meshes**. Rensselaer Polytechnic Institute, Troy, NY, USA: PhD Thesis, 2000.

- PUN, T. **A new method for gray-level picture threshold using the entropy of the histogram.** Signal Processing. 1980.
- REISS, M. L. L.; TOMMASELLI, A. M. G. **Reconstrução 3D por Luz Estruturada: Calibração dos Vetores Diretores dos Feixes de Padrões Projetados.** In: XXI CONGRESSO BRASILEIRO DE CARTOGRAFIA, Belo Horizonte, MG, 2003.
- SAHOO, P. K.; SOLTANI, S.; WONG, A. K. C., **An Survey of Thresholding Techniques, Computer Vision, Graphics and Image Processing,** 1988.
- SANTIN, Rafael. **Sistema de autoria em ambiente colaborativo com realidade aumentada.** Dissertação de Mestrado: UNIMEP, 2008.
- SCARLATOS L. L. & PAVLIDIS T. **Hierarchical triangulation using terrain features.** IEEE Visualization'90, San Francisco, California, USA: Proceedings, 1990.
- SCHENK, T. **Digital Photogrammetry.** TerraScience, 1999.
- SCHROEDER F. & ROBBACH P. **Managing the complexity of digital terrain models.** Computer & Graphics, 1994.
- SCHROEDER W. J., ZARGE J. A., LORENSEN W. E. **Decimation of triangle meshes.** Computer Graphics, 1992.
- SEZGIN, M.; SANKUR, B.; **Survey over image thresholding techniques and quantitative performance evaluation.** Journal of Electronic Imaging, 2004.
- SONKA, M.; VAKLAV, H; BOYLE, R. **Image Processing, Analysis, and Machine Vision.** PWS Publishing, 1998.

S. SULLIVAN AND J. PONCE, **Automatic Model Construction, Pose Estimation, and Object Recognition from Photographs Using Triangular Splines**, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998.

TOMMASELLI, A. M. G. **Um Estudo Sobre as Técnicas de Varredura a Laser e Fotogrametria para Levantamentos 3D a Curta Distância**. GEODÉSIA Online - Revista da Comissão Brasileira de Geodésia. 2003.

TURK G. **Re-tiling polygonal surfaces**. Computer Graphics, 1992.

VOIGTMANN A., BECKER L., HINRICHS K. **A hierarchical model for multiresolution surface reconstruction**. Graphical Models and Image Processing, 1997.