



UNIVERSIDADE FEDERAL DE UBERLÂNDIA  
FACULDADE DE ENGENHARIA ELÉTRICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Daniel Antônio Furtado

**Um Método Computacional Livre de Modelo Esquelético para  
Rastreamento e Reconstrução em Tempo Real de Múltiplos  
Marcadores em Sistemas de Captura de Movimento Ópticos**

UBERLÂNDIA

2013



UNIVERSIDADE FEDERAL DE UBERLÂNDIA  
FACULDADE DE ENGENHARIA ELÉTRICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Daniel Antônio Furtado

**Um Método Computacional Livre de Modelo Esquelético para  
Rastreamento e Reconstrução em Tempo Real de Múltiplos  
Marcadores em Sistemas de Captura de Movimento Ópticos**

Tese apresentada à Coordenação do  
Programa de Pós-Graduação em  
Engenharia Elétrica da Universidade  
Federal de Uberlândia, como requisito  
parcial para obtenção do título de Doutor  
em Engenharia Elétrica.

**Orientador:** Prof. Dr. Adriano A. Pereira

UBERLÂNDIA

2013

Daniel Antônio Furtado

**Um Método Computacional Livre de Modelo Esquelético para  
Rastreamento e Reconstrução em Tempo Real de Múltiplos  
Marcadores em Sistemas de Captura de Movimento Ópticos**

Esta Tese de Doutorado foi julgada adequada para a obtenção do título de Doutor em Engenharia Elétrica e aprovada em sua forma final pelo programa de Pós Graduação em Engenharia Elétrica da Universidade Federal de Uberlândia.

---

Daniel Antônio Furtado

Banca Examinadora:

---

Prof. Dr. Adriano Alves Pereira (orientador)

Universidade Federal de Uberlândia

---

Prof. Dr. Adriano de Oliveira Andrade

Universidade Federal de Uberlândia

---

Prof. Dr. Sílvio Soares dos Santos

Universidade Federal de Uberlândia

---

Prof. Dr. Marcus Fraga Vieira

Universidade Federal de Goiás

---

Prof. Dr. César Ferreira Amorim

Universidade Cidade de São Paulo

UBERLÂNDIA

2013

*Aos meus pais, Rosário e Elza.*

## AGRADECIMENTOS

Gostaria de agradecer, em primeiro lugar, ao professor, orientador e amigo Dr. Adriano Alves Pereira. Muito obrigado por ter me confiado esta maravilhosa pesquisa, que me fez crescer tanto pessoal quanto profissionalmente. Obrigado por ter paciência nos momentos de dificuldade, pela confiança e por todas as revisões, sugestões e ensinamentos.

Agradeço a todos os professores, amigos e colegas do Laboratório de Engenharia Biomédica da Universidade Federal de Uberlândia pelo apoio prestado durante este período de convivência. Obrigado Maria Fernanda, Eder Moura, Guilherme Cavalheiro, Selma Milagre, Jeovane Sousa, Lucas Rocha, Ailton e Raphael Branquinho pelos incansáveis auxílios.

Agradeço ao professor Dr. Sílvio Santos da Faculdade de Educação Física e ao professor Dr. Adriano Andrade pelo apoio oferecido em diversos momentos desta pesquisa e por todas as suas críticas e sugestões durante a avaliação deste trabalho.

Agradeço aos professores Dr. Marcus Fraga e Dr. César Amorim por aceitarem participar como membros da banca examinadora e por todas as suas revisões, críticas e sugestões.

Agradeço ao Edimar Andrade, técnico da oficina mecânica do Instituto de Física da Universidade Federal de Uberlândia, que me auxiliou diversas vezes na construção de dispositivos de calibração e sustentação para as câmeras.

Agradeço à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro.

Por fim, agradeço imensamente aos meus pais, Rosário Furtado Paim e Elza Helena Soares Paim, pelo apoio contínuo e amor incondicional.

*“Your work is going to fill a large part of your life, and the only way to be truly satisfied is to do what you believe is great work. And the only way to do great work is to love what you do.”*

*“Seu trabalho preencherá grande parte de sua vida, e a única forma de estar satisfeito de verdade é fazendo o que você acredita ser um grande trabalho. E a única forma de fazer um grande trabalho é amando o que você faz.”*

*Steve Jobs*

## RESUMO

Nos últimos anos, os sistemas de captura de movimento vêm sendo aplicados em diversas áreas do conhecimento. A animação de personagens virtuais na indústria cinematográfica e a avaliação dos movimentos corporais em áreas da saúde e nos esportes são apenas algumas de suas diversas aplicações. Dos diferentes tipos de sistemas de captura de movimento existentes, os sistemas ópticos baseados em marcadores são reconhecidos como os mais avançados e os que oferecem níveis de precisão e flexibilidade suficientes para suportar o maior número de aplicações. Todavia, esses sistemas estão também entre os mais caros, devido ao alto preço dos elementos de hardware e software. As propostas de baixo custo existentes possuem limitações com relação à precisão, flexibilidade, automatismo e/ou velocidade de processamento, que as tornam inviáveis para uma considerável parcela de aplicações. Neste contexto, esta pesquisa propõe um método computacional completo, considerando as principais deficiências dos trabalhos existentes, para rastrear e reconstruir as trajetórias de uma nuvem de marcadores reflexivos independentes, em tempo real, utilizando múltiplas câmeras especializadas de infravermelho. A técnica proposta envolve um conjunto de algoritmos relativamente simples, os quais integram três etapas principais. Estas etapas realizam o rastreamento e o casamento dos pontos nas imagens e a reconstrução tridimensional das trajetórias dos marcadores. O método foi implementado em software e testado utilizando uma coleção de oito câmeras de infravermelho. Como referência, utilizou-se o software Arena e o sistema *OptiTrack* comercializado pela empresa *NaturalPoint*. Nos experimentos conduzidos, a técnica foi capaz de rastrear automaticamente um conjunto de 38 marcadores em tempo real. Comparada ao software comercial Arena, a técnica revelou uma melhor capacidade de rastreamento automático dos marcadores e ainda reconstruiu trajetórias com menor incidência de ruídos. O método proposto deve incentivar pesquisadores e empresas a desenvolverem sistemas de captura de movimento de alto desempenho e com menor custo.

**Palavras-chave:** sistemas de captura de movimento ópticos; marcadores reflexivos; reconstrução 3D; visão computacional.

## ABSTRACT

In the past years, motion capture has been widely used in many application areas. In movies and games, motion capture is frequently employed to animate virtual characters. In sports, motion capture and analyses focus on optimizing movements of athletes and injury prevention. More applications areas include medicine, military and engineering. Motion capture can be accomplished by several technologies. However, optical marker-based systems are considered the gold standard of the motion capture field. They can offer high precision levels and flexibility to support most applications, but they are also the most expensive systems due to high costs of hardware and software. Although low-cost optical systems have been proposed in the last decade, these systems cannot provide enough precision, flexibility, automatism and/or real-time capability for a number of applications. In this context, the present research aims to develop a complete and high precision approach to track and reconstruct a cloud of independent markers, in real time, using multiple infrared specialized cameras. The proposed method includes a set of relatively simple algorithms which are part of a three-stage procedure. These stages work on the tracking and matching of the image points and spatial reconstruction of the marker trajectories. In order to evaluate the method, a prototype software has been implemented and experiments were performed using a pack of eight infrared cameras. The NaturalPoint's OptiTrack system, which includes the *Arena* software, was used as a reference system. In the experiments, the proposed technique was able to successfully track and reconstruct a set of 38 reflective markers in real time. When compared to the commercial software, the method performed better for automatically marker tracking. In addition, the reconstructed trajectories produced by the prototype software were far less contaminated by noise than the trajectories generated by the *Arena* software. The proposed method should encourage the development of new high performance systems at a more affordable price.

**Key-words:** optical motion capture systems; reflective markers; 3D reconstruction; computer vision.

## LISTA DE SIGLAS E ABREVIATURAS

<i>Mocap</i>	<i>Motion Capture – Captura de Movimento</i>
DLT	<i>Direct Linear Transformation</i>
CCD	<i>Charge-coupled device</i>
SVD	<i>Singular Value Decomposition</i>
RP <sub>2D</sub>	Região de Procura no Espaço Bidimensional

## LISTA DE FIGURAS

FIGURA 2.1 - EXEMPLOS CLÁSSICOS DE APLICAÇÃO DE SISTEMAS DE MOCAP NA INDÚSTRIA CINEMATOGRÁFICA. A IMAGEM (A) ILUSTRA O ATOR TOM HANKS TENDO OS MOVIMENTOS CAPTURADOS PARA O PERSONAGEM DO FILME <i>O EXPRESSO POLAR</i> (IGN, 2004). EM (B), A CAPTURA É FEITA PARA O PERSONAGEM SMEAGOL DA TRILOGIA <i>O SENHOR DOS ANÉIS</i> (IGN, 2003). ....	9
FIGURA 2.2 - IMAGEM DO JOGADOR MAURICE EDU EM ESTÚDIO DE MOCAP (PISSANOS, 2008). O JOGADOR TEVE OS MOVIMENTOS COLETADOS PARA O JOGO DE FUTEBOL <i>FIFA SOCCER 9</i> PRODUZIDO PELA EA SPORTS.....	9
FIGURA 2.3 - SISTEMA DE MOCAP DA QUALISYS SENDO EMPREGADO NA CAPTURA DOS MOVIMENTOS DE UM NADADOR. CÂMERAS SUBMERSAS POSSIBILITAM A COLETA EMBAIXO D'ÁGUA. IMAGEM ADAPTADA DE (QUALISYS, 2013). ....	10
FIGURA 2.4 - (A) INDIVÍDUO USANDO UM CONJUNTO DE MARCADORES PARA CAPTURA APROXIMADA DOS MOVIMENTOS MANDIBULARES; (B) RECONSTRUÇÃO REALIZADA POR UM SISTEMA DE MOCAP APÓS ABERTURA E FECHAMENTO DA MANDÍBULA (FURTADO, PEREIRA ET AL., 2010). A TRAJETÓRIA DO MARCADOR POSTO SOBRE O QUEIXO É DESTACADA. ....	11
FIGURA 2.5 - EXEMPLO DE USO DE MOCAP NO PROJETO DE VEÍCULOS DA FORD. IMAGEM ADAPTADA DE (MOTORTREND, 2011).....	12
FIGURA 2.6 - EXEMPLOS DE APLICAÇÃO DE SISTEMAS DE MOCAP EM TREINAMENTOS MILITARES. (A) TREINAMENTO DE SOLDADOS DOS EUA (SOCIETY, 2011). (B) IMAGEM ADAPTADA DE (HERALD, 2010). ....	13
FIGURA 2.7 - ILUSTRAÇÃO DO SISTEMA DE CAPTURA DE MOVIMENTO ELETROMECÂNICO GYPSY-7 DA EMPRESA ANIMAZOO. IMAGEM ADAPTADA DE (ANIMAZOO, 2013).....	14
FIGURA 2.8 - ILUSTRAÇÃO DO SISTEMA DE CAPTURA DE MOVIMENTO INERCIAL MVN DA EMPRESA XSENS. IMAGEM ADAPTADA DE (XSENS, 2013).....	16
FIGURA 2.9 - EXEMPLOS DE DISPOSITIVOS UTILIZADOS EM SISTEMAS ÓPTICOS PASSIVOS. (A) CÂMERA DE INFRAVERMELHO COMERCIALIZADA PELA EMPRESA VICON; (B) MARCADOR ESFÉRICO DE 1,6 CM COMUMENTE UTILIZADO PARA CAPTURA DE MOVIMENTOS CORPORAIS; (C) CONJUNTO DE MARCADORES SEMIESFÉRICOS DE 3 MM NORMALMENTE UTILIZADOS EM CAPTURAS FACIAIS. ....	18
FIGURA 4.1 - ILUSTRAÇÃO DE UMA CÂMERA DE PINHOLE. A LUZ QUE ENTRA NA CÂMERA PELO ORIFÍCIO PERMITE A FORMAÇÃO DE UMA IMAGEM NO LADO OPPOSTO DA CAIXA.....	33
FIGURA 4.2 - GEOMETRIA DA CÂMERA DE PINHOLE. (A) VISÃO EM PERSPECTIVA. (B) VISÃO NO PLANO YZ. A COORDENADA VERTICAL DO PONTO PROJETADO M NO PLANO DE IMAGEM PODE SER CALCULADA POR SEMELHANÇA DE TRIÂNGULOS, OU SEJA: $F/Z = V'/Y$ , LOGO $V' = F \cdot Y/Z$ . O MESMO PROCESSO PODE SER EMPREGADO PARA O CÁLCULO DA COORDENADA HORIZONTAL DE M NO PLANO DE IMAGEM. FIGURAS BASEADAS EM (IVEKOVIC, FUSIELLO ET AL., 2006).....	34

FIGURA 4.3 – ILUSTRAÇÃO DA MUDANÇA DO SISTEMA DE COORDENADAS DO MUNDO PARA O SISTEMA DE COORDENADAS DA CÂMERA. AS COORDENADAS DE UM PONTO SÃO CONVERTIDAS MEDIANTE APLICAÇÃO DE UMA OPERAÇÃO DE ROTAÇÃO E UMA DE TRANSLAÇÃO.....	37
FIGURA 4.4 – EFEITOS DOS VÁRIOS TIPOS DE DISTORÇÃO NAS IMAGENS: (A) REPRESENTAÇÃO DE IMAGEM IDEAL, SEM DISTORÇÃO; (B) DISTORÇÃO RADIAL PARA O CENTRO (BARREL); (C) DISTORÇÃO RADIAL PARA FORA (PINCUSHION) E (D) DISTORÇÃO TANGENCIAL. NESTE ÚLTIMO CASO, AS LINHAS PONTILHADAS REPRESENTAM PONTOS COM DISTORÇÃO TANGENCIAL. ILUSTRAÇÕES ADAPTADAS DE (WENG, COHEN ET AL., 1992). .....	39
FIGURA 4.5 – (A) OBJETO DE CALIBRAÇÃO FORMADO POR TRÊS EIXOS ORTOGONIAIS. OS NOVE PONTOS DE CONTROLE CORRESPONDEM AOS NOVE MARCADORES REFLEXIVOS, CUJAS POSIÇÕES ESPACIAIS SÃO CONHECIDAS (IMAGEM RETIRADA DE PRIBANIĆ, SURM ET AL. (2007). (B) OBJETO DE CALIBRAÇÃO EM FORMATO DE GRADE SUSTENTANDO DOZE MARCADORES.....	44
FIGURA 4.6 – EXEMPLO DE UM BASTÃO DE CALIBRAÇÃO (CALIBRATION WAND) UTILIZADO PELO FABRICANTE NATURALPOINT PARA CONDUZIR A ETAPA DA CALIBRAÇÃO DINÂMICA DE SEU SISTEMA. IMAGEM ADAPTADA DE (NATURALPOINT, 2013). .....	46
FIGURA 4.7 - GEOMETRIA EPIPOLAR. AS DUAS CÂMERAS SÃO REPRESENTADAS PELOS CENTROS ÓPTICOS $C_e$ E $C_d$ E PELOS PLANOS DE IMAGEM. OS PONTOS $C_e$ E $C_d$ , O PONTO 3D NO ESPAÇO $M$ E SUAS PROJEÇÕES NAS IMAGENS $me$ E $md$ COMPARTILHAM UM PLANO COMUM, CHAMADO PLANO EPIPOLAR. A INTERSEÇÃO DO PLANO EPIPOLAR COM OS PLANOS DE IMAGEM FORMAM AS RETAS EPIPOLARES. A LINHA DE BASE CONECTA OS CENTROS DAS CÂMERAS E PASSAM PELOS PLANOS DE IMAGEM NOS EPIPOLOS ( $ee$ E $ed$ ). .....	47
FIGURA 4.8 – RECONSTRUÇÃO TRIDIMENSIONAL DE UM PONTO $M$ DADOS AS DETECÇÕES $m1$ E $m2$ NOS PLANOS DE IMAGEM DE DUAS CÂMERAS $C1$ E $C2$ , RESPECTIVAMENTE. SITUAÇÃO IDEAL, DESCONSIDERANDO LIMITAÇÕES DA CALIBRAÇÃO, DISTORÇÃO DE LENTES, ETC. ....	50
FIGURA 5.1 – ILUSTRAÇÃO DO FUNCIONAMENTO BÁSICO DA TÉCNICA. PARA UM FRAMESET COLETADO EM UM INSTANTE $T \geq 3$ , UTILIZA-SE SEMPRE OS TRÊS ÚLTIMOS FRAMESETS PROCESSADOS E O FRAMESET CORRENTE PARA OBTENÇÃO DO FRAME 3D CORRENTE ( $3F^r$ ). A LETRA $P$ É UTILIZADA PARA INDICAR QUE O FRAMESET JÁ FOI AVALIADO. ....	55
FIGURA 5.2 – ILUSTRAÇÃO DO PAPEL DESEMPENHADO POR CADA ETAPA DO ALGORITMO PRINCIPAL QUE COMPÕE O MÉTODO DE RECONSTRUÇÃO PROPOSTO. O DIAGRAMA PRESSUPÕE QUE OS FRAMESETS ANTERIORES AO INSTANTE $T$ JÁ TENHAM SIDO PROCESSADOS, POSSIBILITANDO QUE A ETAPA DE RASTREAMENTO 2D INICIE O PROCEDIMENTO DE IDENTIFICAÇÃO DOS MARCADORES NO FRAMESET CORRENTE.....	55
FIGURA 5.3 - RASTREAMENTO 2D - CASO (A). O MARCADOR SENDO RASTREADO POSSUI REGISTRO NOS DOIS ÚLTIMOS FRAMES (COMEÇANDO NO INSTANTE $T-2$ ). SUA POSIÇÃO NO FRAME ATUAL É ESTIMADA COM BASE NAS POSIÇÕES ANTERIORES, QUE SÃO UTILIZADAS PARA A DEFINIÇÃO DE UMA REGIÃO DE BUSCA.	

CASO EXATAMENTE UM MARCADOR SEJA ENCONTRADO DENTRO DESSA REGIÃO, ENTÃO O MESMO É MARCADO COMO IDENTIFICADO UTILIZANDO O RÓTULO DO MARCADOR PRECEDENTE.....	58
FIGURA 5.4 - RASTREAMENTO 2D. CASO EM QUE EXISTE UM CAMINHO DEFINIDO PELO MARCADOR NOS ÚLTIMOS TRÊS FRAMES. O ÂNGULO REFERENTE À "CURVA" DEFINIDA PELO MARCADOR É EMPREGADO PARA ESTIMAR SUA POSIÇÃO NO FRAME CORRENTE.....	59
FIGURA 5.5 - CICLO BÁSICO UTILIZADO DURANTE O DESENVOLVIMENTO DA TÉCNICA E IMPLEMENTAÇÃO DO SOFTWARE CORRESPONDENTE. ....	63
FIGURA 6.1 - ESKEMA DE POSICIONAMENTO DAS CÂMERAS. UM TRIPÉ EM CADA CANTO É UTILIZADO PARA SUSTENTAR DUAS CÂMERAS. NA FIGURA, CADA X REPRESENTA UMA CÂMERA DE INFRAVERMELHO.....	66
FIGURA 6.2 - ESKEMA DO DISPOSITIVO UTILIZADO PARA AVALIAR E COMPARAR A PRECISÃO E EXATIDÃO DO MÉTODO PROPOSTO.....	67
FIGURA 6.3 - NÚMERO DE CÂMERAS VISUALIZANDO O MARCADOR INFERIOR DA PLACA DE PVC AO LONGO DE TRÊS VOLTAS. O GRÁFICO MOSTRA QUE O MARCADOR É VISUALIZADO, A MAIOR PARTE DO TEMPO, POR QUATRO CÂMERAS.....	69
FIGURA 6.4 - (A) DISTÂNCIA MÉDIA ENTRE OS DOIS MARCADORES DO TOPO DO TUBO DE PVC. A DISTÂNCIA CALCULADA UTILIZANDO O PROTÓTIPO FOI INSERIDA NO CONJUNTO DE VALORES OBTIDO POR RICHARDS (1999); (B) DISPERSÃO DOS VALORES DE DISTÂNCIA CALCULADOS QUADRO A QUADRO ENTRE OS DOIS MARCADORES DO TUBO; (C) DISTÂNCIA MÉDIA ENTRE OS DOIS MARCADORES SUPERIORES DA PLACA DE PVC E (D) DISPERSÃO ASSOCIADA À DISTÂNCIA ENTRE OS MARCADORES SUPERIORES DA PLACA. ....	75
FIGURA 6.5 - (A) DIFERENÇA ENTRE O VALOR CALCULADO E VALOR REAL DO ÂNGULO ENTRE OS MARCADORES DA PLACA DE PVC. NO DISPOSITIVO UTILIZADO POR RICHARDS, O MARCADOR DO VÉRTICE FOI COLOCADO FORMANDO UM ÂNGULO DE 95,8º COM OS OUTROS DOIS. JÁ NO DISPOSITIVO CONSTRUÍDO PARA AVALIAR O PROTÓTIPO, O ÂNGULO FORMADO FOI DE 94,2º; (B) DESVIO PADRÃO DAS MEDIDAS DOS ÂNGULOS CALCULADAS QUADRO A QUADRO.....	76
FIGURA 7.1 - COMPOSIÇÃO DA PLATAFORMA GIRATÓRIA DESENVOLVIDA PARA ROTACIONAR O MANEQUIM. ....	81
FIGURA 7.2 - POSICIONAMENTO DOS MARCADORES NO MANEQUIM. PADRÃO DE 38 MARCADORES UTILIZADO PELO SOFTWARE ARENA DA EMPRESA NATURALPOINT. CADA MARCADOR TEM 1,6 CM DE DIÂMETRO..	82
FIGURA 7.3 - CONTAGEM DO TOTAL DE INTERRUPÇÕES NO RASTREAMENTO 3D. CADA X ILUSTRA UM PONTO DE INTERRUPÇÃO NO RASTREAMENTO. NESTE EXEMPLO, HÁ UM TOTAL DE 4 INTERRUPÇÕES. ....	85
FIGURA 7.4 - TRAJETÓRIA RECONSTRUÍDA DE UM MARCADOR COM A PRESENÇA DE RUÍDO DO TIPO SPIKE....	86
FIGURA 7.5 - PROCEDIMENTO GERAL DE CAPTURA. INICIALMENTE OS SISTEMAS FORAM CALIBRADOS E EM SEGUIDA INICIOU-SE AS SESSÕES DE CAPTURA UTILIZANDO OS DOIS SOFTWARES. COLETAS FORAM CONDUZIDAS COM A PLATAFORMA GIRANDO A 10, 20, 30 E 40 RPM.....	89

FIGURA 7.6 - ILUSTRAÇÃO DO PROCESSO UTILIZADO PARA APLICAÇÃO DO MÉTODO COM VARIAÇÃO DO TAMANHO DA REGIÃO DE PROCURA BIDIMENSIONAL, RP <sub>2D</sub> . A FIGURA ILUSTRA O PROCESSO CONSIDERANDO APENAS UM DOS ARQUIVOS DE DADOS (PARA UMA VELOCIDADE DA PLATAFORMA).....	90
FIGURA 7.7 - (A) NÚMERO DE MARCADORES IDENTIFICADOS NO INÍCIO DA SEQUÊNCIA E RASTREADOS ATÉ O FINAL EM FUNÇÃO DO TAMANHO DA REGIÃO DE PROCURA NO RASTREAMENTO 2D. CADA LINHA NO GRÁFICO CORRESPONDE À COLETA REALIZADA EM UMA DADA VELOCIDADE DE ROTAÇÃO (10, 20, 30 OU 40 RPM) DA PLATAFORMA. (B) VARIAÇÃO DO NÚMERO MÁXIMO DE MARCADORES DETECTADOS EM UM FRAME 3D À MEDIDA QUE É AUMENTADO O TAMANHO DA REGIÃO DE BUSCA NO RASTREAMENTO 2D. O NÚMERO MÁXIMO DE MARCADORES DETECTADOS EM UM FRAME 3D ESTÁ RELACIONADO COM A QUANTIDADE DE FALSOS MARCADORES. (C) NÚMERO TOTAL DE INTERRUPÇÕES NO RASTREAMENTO 3D EM FUNÇÃO DO TAMANHO DA REGIÃO DE BUSCA NO RASTREAMENTO 2D.....	91
FIGURA 7.8 - NÚMERO MÉDIO DE RUÍDOS DO TIPO SPIKE DETECTADOS POR TRAJETÓRIA INTEIRAMENTE RECONSTRUÍDA EM FUNÇÃO DO TAMANHO DA REGIÃO DE BUSCA NO RASTREAMENTO 2D.....	92
FIGURA 7.9 - COEFICIENTE GERAL DE RUÍDO EM FUNÇÃO DO TAMANHO DA REGIÃO DE PROCURA 2D, RP <sub>2D</sub> ..	93
FIGURA 7.10 - VALORES NORMALIZADOS DE TODOS OS CRITÉRIOS DE AVALIAÇÃO DOS DADOS 3D EM FUNÇÃO DO TAMANHO DA REGIÃO DE PROCURA 2D. A REGIÃO REPRESENTADA POR UM RETÂNGULO CORRESPONDE À FAIXA DE VALORES DE RP <sub>2D</sub> QUE POSSIBILITOU A GERAÇÃO DOS MELHORES RESULTADOS COM RESPEITO AOS CINCO CRITÉRIOS. ....	95
FIGURA 7.11 - NÚMERO MÁXIMO DE MARCADORES ENCONTRADOS EM UM MESMO FRAME 3D. O PROTÓTIPO NÃO REGISTROU MAIS DE 38 MARCADORES EM NENHUM MOMENTO DOS TESTES. POR OUTRO LADO, O SOFTWARE ARENA REGISTROU MAIS MARCADORES QUE O REAL: 39 OU 40. ....	97
FIGURA 7.12 - NÚMERO MÉDIO DE SPIKES DETECTADOS POR TRAJETÓRIA RECONSTRUÍDA. O NÚMERO FOI CALCULADO PARA CADA UM DOS TESTES, COM A PLATAFORMA GIRANDO A 10, 20, 30 OU 40 RPM. ....	97
FIGURA 7.13 - COORDENADAS X, Y E Z DE UM TRECHO DE APROXIMADAMENTE 5 SEGUNDOS DA TRAJETÓRIA DE UM DOS MARCADORES COLOCADOS NO TRONCO DO MANEQUIM. OS TRÊS GRÁFICOS DA ESQUERDA APRESENTAM AS VARIAÇÕES DAS COORDENADAS OBTIDAS COM O SOFTWARE ARENA. OS TRÊS GRÁFICOS DA DIREITA APRESENTAM OS VALORES EQUIVALENTES OBTIDOS COM O PROTÓTIPO DESENVOLVIDO. NENHUM TIPO DE FILTRO FOI APLICADO.....	98
FIGURA 7.14 - COEFICIENTE GERAL DE RUÍDO CALCULADO PARA OS DADOS 3D GERADOS PELOS DOIS SOFTWARES. PARA TODAS AS COLETAS, O COEFICIENTE DE RUÍDO ASSOCIADO AOS RESULTADOS PRODUZIDOS PELO PROTÓTIPO É PELO MENOS 13 VEZES MENOR DO QUE AQUELE OBTIDO PARA OS DADOS GERADOS PELO ARENA.....	99
FIGURA 8.1 - AMOSTRAS CAPTURADAS DA TELA DO PROTÓTIPO APÓS CAPTURA E RECONSTRUÇÃO DOS MOVIMENTOS DO INDIVÍDUO NA ESTEIRA. UM TOTAL DE 38 MARCADORES FOI UTILIZADO. AS IMAGENS (A), (B) E (C) CORRESPONDEM, RESPECTIVAMENTE, ÀS VISÕES ESPACIAIS DOS MARCADORES ANTEROLATERAL, FRONTAL E PÓSTERO-LATERAL. A IMAGEM (D) EXIBE O TRAÇADO ESPACIAL DO	

MARCADOR POSTO SOBRE O TORNOZELO DIREITO DO PARTICIPANTE, REGISTRADO DURANTE APROXIMADAMENTE 4 SEGUNDOS.....	103
FIGURA 8.2 - CARACTERÍSTICAS DOS DADOS TRIDIMENSIONAIS GERADOS PELO PROTÓTIPO E PELO SOFTWARE COMERCIAL ARENA QUANDO DA APLICAÇÃO DOS MESMOS NOS 20 SEGMENTOS DE DADOS. CADA GRÁFICO DO TIPO BOXPLOT POSSIBILITA UMA AVALIAÇÃO E COMPARAÇÃO DESSAS MEDIDAS COM RELAÇÃO À CONCENTRAÇÃO/DISPERSÃO DOS VALORES OBTIDOS. A LINHA CENTRAL DENTRO DE CADA CAIXA INDICA A MEDIANA.....	106
FIGURA 9.1 – TRAJETÓRIA ESPACIAL DOS TRÊS MARCADORES UTILIZADOS NA BARRA AO LONGO DA EXECUÇÃO DE UM CICLO DE MOVIMENTO NO SUPINO PARALÍMPICO. (A) ATLETA EXERCE CONTROLE PRÓXIMO DO IDEAL SOBRE A BARRA. (B) ATLETA REALIZA UM MOVIMENTO IRREGULAR COM O BRAÇO ESQUERDO..	110
FIGURA 9.2 - CARACTERÍSTICAS DOS MOVIMENTOS DE SUPINO PARALÍMPICO EXTRAÍDAS DAS TRAJETÓRIAS ESPACIAIS RECONSTRUÍDAS PELO PROTÓTIPO. (A) TEMPOS DE DESCIDA E SUBIDA DA BARRA PARA CADA ATLETA. (B) TEMPO DE PERMANÊNCIA DA BARRA PRÓXIMO AO PEITO. (C) MAIOR DIFERENÇA DE ALTURA REGISTRADA ENTRE OS MARCADORES EXTREMOS DA BARRA. (D) ALTURA DOS MARCADORES EXTREMOS AO LONGO DO MOVIMENTO EXECUTADO POR G.....	111
FIGURA 9.3 – EXEMPLO DE USO DO PROTÓTIPO PARA CAPTURAR E RECONSTRUIR OS MOVIMENTOS DE UM ATLETA DO TAEKWONDO. O MOVIMENTO ILUSTRADO CORRESPONDE AO GOLPE <i>TIMYO TCHAGUI</i> . A DETECÇÃO DOS MARCADORES E A RECONSTRUÇÃO DE SUAS TRAJETÓRIAS FORAM REALIZADAS EM TEMPO REAL.....	112

## LISTA DE TABELAS

TABELA 6.1 - DISTÂNCIAS CALCULADAS DOS MARCADORES E ERROS ASSOCIADOS.....	70
TABELA 6.2 - ÂNGULO ENTRE OS MARCADORES DA PLACA DE PVC E ERROS ASSOCIADOS.....	71
TABELA 6.3 - ERRO ASSOCIADO À POSIÇÃO DO MARCADOR FIXO À MEDIDA QUE O MARCADOR DE POSIÇÃO VARIÁVEL É AJUSTADO EM DIREÇÃO AO FIXO. ADICIONALMENTE SÃO APRESENTADAS AS DISTÂNCIAS REAIS E AS DISTÂNCIAS CALCULADAS ENTRE OS CENTROS DOS DOIS MARCADORES.....	72
TABELA 6.4 - SISTEMAS AVALIADOS POR RICHARD UTILIZANDO DISPOSITIVO ILUSTRADO ANTERIORMENTE..	73
TABELA 8.1 – CARACTERÍSTICAS DOS DADOS TRIDIMENSIONAIS GERADOS PELO PROTÓTIPO E PELO SOFTWARE COMERCIAL ARENA QUANDO DA APLICAÇÃO DOS MESMOS NOS 20 SEGMENTOS DE DADOS. PARA CADA CRITÉRIO DE AVALIAÇÃO DEFINIDO NA SEÇÃO 7.3, SÃO APRESENTADOS OS VALORES MÉDIOS OBTIDOS, OS DESVIOS PADRÃO E O RESULTADO DO TESTE DE ANOVA. .....	105

# SUMÁRIO

<b>CAPÍTULO 1 INTRODUÇÃO .....</b>	<b>1</b>
1.1 MOTIVAÇÃO E OBJETIVO GERAL .....	2
1.2 OBJETIVOS SECUNDÁRIOS .....	6
1.3 ORGANIZAÇÃO DO DOCUMENTO.....	6
<b>CAPÍTULO 2 TECNOLOGIAS DE CAPTURA E SUAS APLICAÇÕES.....</b>	<b>7</b>
2.1 CAPTURA DE MOVIMENTO E SUAS APLICAÇÕES.....	8
2.1.1 <i>Indústria Cinematográfica</i> .....	8
2.1.2 <i>Indústria de Jogos</i> .....	8
2.1.3 <i>Esportes</i> .....	10
2.1.4 <i>Saúde e Biomecânica</i> .....	10
2.1.5 <i>Engenharia</i> .....	12
2.1.6 <i>Militar</i> .....	12
2.2 TECNOLOGIAS DE CAPTURA DE MOVIMENTO.....	13
2.2.1 <i>Sistemas Eletromecânicos</i> .....	13
2.2.2 <i>Sistemas Magnéticos</i> .....	14
2.2.3 <i>Sistemas Iniciais</i> .....	15
2.2.4 <i>Sistemas Acústicos</i> .....	16
2.2.5 <i>Sistemas Ópticos</i> .....	17
<b>CAPÍTULO 3 TRABALHOS RELACIONADOS .....</b>	<b>23</b>
3.1 TÉCNICAS BASEADAS EM MODELO.....	25
3.2 TÉCNICAS LIVRES DE MODELO .....	28
3.3 CONSIDERAÇÕES DO CAPÍTULO.....	31
<b>CAPÍTULO 4 FUNDAMENTAÇÃO TEÓRICA.....</b>	<b>32</b>
4.1 MODELO MATEMÁTICO DE UMA CÂMERA .....	33
4.1.1 <i>Modelo de Câmera de Pinhole</i> .....	33
4.1.2 <i>Modelo de Distorção de Brown</i> .....	39
4.2 CALIBRAÇÃO DE CÂMERAS.....	41
4.2.1 <i>Direct Linear Transformation - DLT</i> .....	42
4.2.2 <i>Método de Tsai</i> .....	45
4.2.3 <i>Método de Svoboda</i> .....	45
4.2.4 <i>Métodos Híbridos</i> .....	46
4.3 GEOMETRIA EPIPOLAR .....	47

4.4 RECONSTRUÇÃO TRIDIMENSIONAL .....	48
4.4.1 Reconstrução com base na técnica DLT.....	49
4.4.2 Reconstrução com base na Geometria Epipolar.....	50
<b>CAPÍTULO 5 MATERIAIS E MÉTODOS .....</b>	<b>52</b>
5.1 TERMOS E DEFINIÇÕES .....	52
5.2 VISÃO GERAL DA TÉCNICA .....	54
5.3 DESCREVENDO AS ETAPAS.....	56
5.4 ASPECTOS DE DESENVOLVIMENTO E IMPLEMENTAÇÃO EM SOFTWARE.....	63
<b>CAPÍTULO 6 EXPERIMENTOS UTILIZANDO OBJETO DE TESTE .....</b>	<b>64</b>
6.1 CONDIÇÕES GERAIS DE REALIZAÇÃO DOS EXPERIMENTOS .....	65
6.1.1 Posicionamento das Câmeras.....	65
6.1.2 Calibração .....	66
6.1.3 Parâmetro de Rastreamento 2D .....	66
6.2 DESCRIÇÃO DO DISPOSITIVO.....	66
6.3 PROCEDIMENTO DE CAPTURA E RESULTADOS .....	67
6.3.1 Primeira Etapa - Marcadores Superiores e Marcadores da Placa .....	68
6.3.2 Segunda Etapa - Marcador Fixo e Marcador Variável.....	71
6.3.3 Relacionando os Resultados .....	73
6.4 CONCLUSÃO E CONSIDERAÇÕES DO CAPÍTULO.....	77
<b>CAPÍTULO 7 EXPERIMENTOS UTILIZANDO MANEQUIM .....</b>	<b>79</b>
7.1 DESCRIÇÃO DA PLATAFORMA .....	81
7.2 POSICIONAMENTO DOS MARCADORES.....	81
7.3 CRITÉRIOS DE AVALIAÇÃO DOS DADOS 3D .....	82
7.3.1 Número de Marcadores Rastreados por toda Sequência .....	83
7.3.2 Número Máximo de Marcadores Detectados em um Frame 3D .....	83
7.3.3 Número de Interrupções no Rastreamento 3D .....	84
7.3.4 Número Médio de Spikes por Trajetória .....	85
7.3.5 Coeficiente Geral de Ruído.....	86
7.4 PROCEDIMENTO DE CAPTURA .....	88
7.5 ESTUDO DE DESEMPENHO DO MÉTODO MEDIANTE VARIAÇÃO DO PARÂMETRO $RP_{2D}$ .....	89
7.5.1 Discussão .....	93
7.6 COMPARATIVO: PROTÓTIPO BASEADO NA TÉCNICA E SOFTWARE ARENA .....	95
7.7 CONCLUSÃO E CONSIDERAÇÕES DO CAPÍTULO .....	100
<b>CAPÍTULO 8 EXPERIMENTOS UTILIZANDO ESTEIRA ERGOMÉTRICA.....</b>	<b>101</b>
8.1 PROCEDIMENTO DE CAPTURA .....	102

8.2 AMOSTRA DOS DADOS 3D RECONSTRUÍDOS PELO PROTÓTIPO .....	102
8.3 PROCEDIMENTO DE AVALIAÇÃO E COMPARAÇÃO DOS RESULTADOS .....	103
8.4 RESULTADOS COMPARATIVOS: ARENA X PROTÓTIPO .....	105
8.5 DISCUSSÃO .....	107
8.6 CONCLUSÃO E CONSIDERAÇÕES DO CAPÍTULO.....	108
<b>CAPÍTULO 9 APLICAÇÃO NOS ESPORTES .....</b>	<b>109</b>
9.1 APLICAÇÃO NO HALTEROFILISMO PARALÍMPICO .....	109
9.2 APLICAÇÃO NO TAEKWONDO .....	111
<b>CAPÍTULO 10 CONCLUSÃO E TRABALHOS FUTUROS .....</b>	<b>113</b>
10.1 TRABALHOS FUTUROS.....	116
10.2 PUBLICAÇÕES .....	116

# Capítulo 1

## Introdução

O termo *captura de movimento* ou *mocap* (do inglês *motion capture*) tem sido utilizado nas últimas décadas para fazer referência ao processo de registro tridimensional do movimento de pessoas, animais ou objetos. O conjunto de materiais e os recursos tecnológicos empregados para realizar essa captura e converter as amostras para uma representação digital compõem os chamados *sistemas de captura de movimento* ou *sistemas de mocap* (Gleicher, 1999).

Atualmente, os sistemas de captura de movimento são aplicados em diversas áreas do conhecimento. Na indústria cinematográfica e na indústria de jogos eletrônicos, por exemplo, esses sistemas são amplamente utilizados para animação de personagens virtuais com base em movimentos coletados de seres reais (Oshita, 2006; King e Paulson, 2007; Kitagawa e Windsor, 2008; Gomide, Flam *et al.*, 2009). Em áreas da saúde, os sistemas de *mocap* são utilizados para estudo e avaliação do movimento corporal humano e podem auxiliar especialistas no diagnóstico e tratamento de diversas doenças (Dijkstra, De Bont *et al.*, 1995; Miller, Bookhan *et al.*, 1999; Travers, Buschang *et al.*, 2000; Yeasin e Chaudhuri, 2000; Johnson e Bobick, 2001; Fukui, Tsuruta *et al.*, 2002; Troje, 2002; Naeije, 2003; Saboune e Charpillet, 2005; Gandolla, Ferrante *et al.*, 2011). Outras aplicações incluem a correção e otimização de movimentos de atletas (Zordan e Hodgins, 2002; Moeslund, Hilton *et al.*, 2006), simulações em ambientes de realidade virtual (Hilton, Beresford *et al.*, 1999;

Moeslund, 2000; Zordan e Hodgins, 2002), projeto de interiores de veículos (Lareau, 2011) e treinamentos militares e espaciais (Davies, Mason *et al.*, 2009).

Na fabricação de sistemas de *mocap*, várias tecnologias podem ser empregadas, tais como estruturas eletromecânicas, sensores magnéticos, sensores acústicos, acelerômetros, giroscópios e câmeras de vídeo (Bachmann, 2000; O'brien, Bodenheimer *et al.*, 2000; Guerra-Filho, 2005; Mundermann, Corazza *et al.*, 2007; Vlasic, Adelsberger *et al.*, 2007; Cloete e Scheffer, 2008; Corrales, Candelas *et al.*, 2008; Sato, Nakamura *et al.*, 2011). A tecnologia predominante em um sistema de captura define o seu *tipo* ou *categoria* e traz um conjunto de vantagens e desvantagens que podem torná-lo mais ou menos apropriado para as diferentes aplicações (Kitagawa e Windsor, 2008). De uma forma geral, os sistemas que utilizam câmeras de vídeo, denominados *sistemas ópticos*, se destacam pelos excelentes níveis de precisão e exatidão oferecidos, pelas altas taxas de amostragem e por serem pouco obstrutivos (Oshita, 2006; Thomas, Mcginley *et al.*, 2007). Além disso, esses sistemas podem ser utilizados em um grande número de aplicações (Moeslund e Granum, 2001; Guerra-Filho, 2005; Moeslund, Hilton *et al.*, 2006; King e Paulson, 2007; Barris e Button, 2008).

Um sistema óptico típico emprega câmeras especializadas de infravermelho em conjunto com dezenas de pequenos objetos, denominados *marcadores*, os quais são colocados em um corpo físico como pontos de referência para possibilitar a reconstrução tridimensional de seus movimentos. Embora algumas pesquisas recentes tenham focado no desenvolvimento de sistemas ópticos livres de marcadores (Mundermann, Corazza *et al.*, 2007; Corazza, Mündermann *et al.*, 2010; John, Trucco *et al.*, 2010), as limitações impostas por esses sistemas ainda inviabilizam seu uso para vários propósitos (Field, Pan *et al.*, 2011).

## 1.1 Motivação e Objetivo Geral

Apesar do amplo leque de aplicações, o acesso aos sistemas de captura de movimento ainda é, lamentavelmente, restrito a uma pequena parcela de profissionais, empresas e pesquisadores (Shih-Pin, Yi-Yao *et al.*, 2009). O principal motivo é o seu alto custo. Embora algumas pesquisas tenham buscado construir

sistemas mais baratos utilizando câmeras de vídeo convencionais, o nível de precisão, exatidão, velocidade de captura e processamento exigidos por grande parte das aplicações costumam ser oferecidos apenas pelos sistemas mais caros e sofisticados. Por exemplo, um dos sistemas ópticos mais acessíveis atualmente é o *OptiTrack*, fabricado pela empresa *NaturalPoint*, cujo preço parte de aproximadamente R\$ 45.000,00<sup>1</sup> (Naturalpoint, 2013) para um conjunto com 12 câmeras e ultrapassa, facilmente, os R\$ 200.000,00, quando são exigidas câmeras e softwares mais sofisticados.

No meio acadêmico, os sistemas de captura de movimento têm despertado a atenção de um número cada vez maior de pesquisadores, que têm confirmado o grande potencial da técnica e a diversidade de suas aplicações (Moeslund e Granum, 2001; King e Paulson, 2007). No trabalho de Moeslund *et al.* (2006), por exemplo, os autores fazem um levantamento de mais de 350 artigos científicos relacionados à captura de movimento, publicados entre 2000 e 2006 em importantes veículos de divulgação internacional.

Os trabalhos divulgados nas últimas décadas abordam diversos aspectos vinculados às diferentes tecnologias de *mocap*. Parte desses trabalhos é voltada para o problema de reconstrução do movimento em sistemas ópticos baseados em marcadores. Neste contexto, várias pesquisas têm focado na proposta de técnicas e algoritmos para resolver o problema de identificação, rastreamento e reconstrução tridimensional das trajetórias dos marcadores (Ringer e Lasenby, 2000; Herda, Fua *et al.*, 2001; Shafiq, Tümer *et al.*, 2001; Figueiroa, Leite *et al.*, 2003; Papic, Zanchi *et al.*, 2004; Ringer e Lasenby, 2004; Castro, Medina-Carnicer *et al.*, 2006; Shih-Pin, Yi-Yao *et al.*, 2009; Zou, Zhao *et al.*, 2009; Canton-Ferrer, Casas *et al.*, 2010; Johnson, Xie *et al.*, 2010a).

Muitos autores propuseram métodos baseados em modelos esqueléticos humanos para guiar o processo de reconstrução (Silaghi, Plänkers *et al.*, 1998; Herda, Fua *et al.*, 2000; Ringer e Lasenby, 2000; Herda, Fua *et al.*, 2001; Cerveri, Pedotti *et al.*, 2003; Ringer e Lasenby, 2004). O emprego desses modelos possibilita o aprimoramento do

---

<sup>1</sup> Valor aproximado, incluindo impostos e envio para o Brasil.

processo, pois as restrições do modelo (como os comprimentos dos membros) podem ser utilizadas para resolver ambiguidades e otimizar o rastreamento e o cálculo das coordenadas 3D dos marcadores. Por outro lado, a necessidade de um modelo insere limitações e inconvenientes que podem restringir as aplicações do sistema de captura (Moeslund, Hilton *et al.*, 2006).

Vários outros autores apresentaram técnicas que não se baseiam em modelos esqueléticos (Shafiq, Tümer *et al.*, 2001; Figueroa, Leite *et al.*, 2003; Papic, Zanchi *et al.*, 2004; Castro, Medina-Carnicer *et al.*, 2006; Pinheiro, 2008; Pinheiro, Andrade *et al.*, 2008; Pinto, Buaes *et al.*, 2008). Entretanto, a maior parte desses trabalhos tem foco no processo de rastreamento e reconstrução em modo *off-line*, utilizando imagens obtidas com câmeras de vídeo convencionais. Nesses casos, os métodos propostos geralmente englobam a detecção dos marcadores nas imagens e muitas vezes são utilizados algoritmos de alto custo computacional para realizar o rastreamento e a reconstrução, uma vez que a execução em tempo real não é um pré-requisito. Além disso, muitos desses métodos não são automáticos e podem necessitar da intervenção do usuário durante o processo de rastreamento e reconstrução tridimensional.

Algumas outras propostas buscaram a detecção e a reconstrução das trajetórias dos marcadores em tempo real (Baroni, Ferrigno *et al.*, 1998; Flam, Queiroz *et al.*, 2009; Ping, Yue *et al.*, 2009; Shih-Pin, Yi-Yao *et al.*, 2009). Contudo, limitações com relação à precisão e exatidão dos dados reconstruídos, quanto ao número de câmeras, número de marcadores e quanto à liberdade de execução dos movimentos, restringem o emprego dessas técnicas para um pequeno grupo de aplicações.

A capacidade de operar em tempo real não é essencial para todas as ocasiões, mas mostra-se de grande importância para a maioria delas (Shih-Pin, Yi-Yao *et al.*, 2009). Por exemplo, ela permite que um atleta e seu treinador obtenham parâmetros e realizem avaliações durante a própria sessão de captura, o que pode determinar instantaneamente a necessidade de repetição do movimento, sua correção ou até mesmo o descarte dos dados por eventuais falhas no processo de captura (como pela ocorrência de oclusões). Tais condições também se aplicam para fins

cinematográficos, na criação de jogos e em aplicações diversas envolvendo a saúde e os esportes.

Nesse contexto, considerando a grande demanda por aplicações de *mocap*, o alto custo dos sistemas comerciais ópticos e as deficiências dos trabalhos existentes na literatura aberta, esta pesquisa propõe um procedimento computacional livre de modelo esquelético, de precisão e exatidão a nível comercial, para rastrear e reconstruir automaticamente as trajetórias de uma nuvem de marcadores, em tempo real, utilizando múltiplas câmeras especializadas de infravermelho.

O método proposto foi testado e avaliado mediante o desenvolvimento de um software. Um conjunto de experimentos foi conduzido e parte dos resultados foi analisada tendo como referência um sistema utilizado mundialmente: o *Optitrack* da empresa *NaturalPoint*. Nos testes realizados, a técnica mostrou-se eficiente e trouxe vantagens, em diversos aspectos, sobre o sistema comercial. O método evidenciou potencial para ser utilizado em muitas aplicações, incluindo desde uma simples análise de marcha a até aquelas que demandam a captura de movimentos mais complexos e em tempo real.

A técnica apresentada deve incentivar pesquisadores e empresas a desenvolverem sistemas de captura de movimento de alto desempenho e com menor custo. A contribuição é especialmente importante para o avanço nacional com relação a essa tecnologia, o que pode trazer benefícios reais para o meio acadêmico e áreas como a saúde, os esportes e o entretenimento. Atualmente, a maioria dos sistemas sendo utilizados nos laboratórios de universidades, estúdios e centros de treinamento e reabilitação do país é adquirida do mercado externo e envolve altos custos com importação, hardware e software. Com o domínio das técnicas computacionais envolvidas, há a possibilidade, por exemplo, de fabricar localmente tais sistemas utilizando como material importado apenas o módulo de circuitos principal de cada câmera, que em alguns casos pode ser adquirido separadamente com um custo consideravelmente menor (Naturalpoint, 2013).

## 1.2 Objetivos Secundários

---

Além de propor uma técnica para rastreamento e reconstrução de uma nuvem de marcadores, conforme mencionado anteriormente, esta pesquisa tem como objetivos secundários:

1. Avaliar o desempenho do método com relação à precisão, exatidão e capacidade elementar de rastreamento e reconstrução em tempo real e ainda relacionar os resultados com valores obtidos de sistemas comerciais;
2. Avaliar e comparar a eficiência do método para reconstruir as trajetórias de dezenas de marcadores em um ambiente controlado, onde é possível variar os parâmetros do método e as condições de captura;
3. Avaliar o desempenho do método em aplicações do mundo real, e comparar os resultados com aqueles obtidos de um sistema comercial.

## 1.3 Organização do Documento

---

Este documento está organizado como segue. O Capítulo 2 descreve algumas das principais aplicações dos sistemas de captura de movimento e as principais categorias de sistemas, mostrando suas limitações, vantagens e desvantagens. O Capítulo 3 descreve os principais trabalhos diretamente relacionados a esta pesquisa. O Capítulo 4 fornece a fundamentação teórica que constitui a base da técnica sendo proposta. São apresentados o modelo de câmera de *pinhole*, o modelo de distorção de Brown, a geometria epipolar e aspectos sobre calibração de câmeras e reconstrução tridimensional. O Capítulo 5 descreve o método de reconstrução proposto. No Capítulo 6 são discutidos os resultados dos experimentos iniciais, que foram conduzidos utilizando um pequeno número de marcadores com o objetivo de avaliar a precisão, exatidão e a capacidade básica de rastreamento e reconstrução da técnica. O Capítulo 7 descreve os experimentos realizados utilizando um manequim, um conjunto maior de marcadores e o software comercial Arena da empresa *NaturalPoint*. Os capítulos 8 e 9 apresentam os experimentos conduzidos com o objetivo de avaliar e comparar a eficiência do método em aplicações do mundo real. Por fim, o Capítulo 10 encerra este documento apresentando as considerações finais, as conclusões, os trabalhos futuros e as publicações obtidas.

# Capítulo 2

## Tecnologias de Captura e suas Aplicações

O desenvolvimento das tecnologias utilizadas para captura de movimento se deve a duas fontes principais: a comunidade acadêmica e a indústria. As primeiras aplicações de *mocap* buscaram a mensuração dos padrões de movimentos de seres humanos e animais para possibilitar um melhor entendimento dos mesmos (Thurston, 1999). Os avanços obtidos naquele momento se devem, principalmente, às pesquisas conduzidas no mundo acadêmico. À medida que os recursos computacionais se desenvolveram e se tornaram comuns, começaram a surgir inúmeras novas aplicações para a técnica (Moeslund, Hilton *et al.*, 2006). Desde então, alguns dos principais interessados, com alto poder de investimento, têm sido as grandes indústrias, como a cinematográfica, a indústria de jogos, a médica e a de publicidade.

Juntas, a comunidade acadêmica e essas grandes indústrias atuam continuamente no desenvolvimento e aprimoramento de diversas tecnologias de captura de movimento, as quais podem ser empregadas em um amplo conjunto de aplicações (Moeslund, 2000; King e Paulson, 2007).

Este capítulo fornece uma visão geral dessas tecnologias e de suas aplicações.

## 2.1 Captura de Movimento e suas Aplicações

---

Algumas das aplicações dos sistemas de *mocap* são apresentadas nesta seção em cinco categorias principais: cinematográfica, jogos, esportes, saúde, engenharia e militar.

### 2.1.1 Indústria Cinematográfica

Os sistemas de *mocap* começaram a ser utilizados pela indústria cinematográfica no final da década de 1980 e início dos anos 1990. Os filmes *O Vingador do Futuro* (*Total Recall*, 1990) e *O Passageiro do Futuro* (*The Lawnmower Man*, 1992) são considerados os primeiros longa metragens a utilizar técnicas de *mocap* para registar movimentos corporais (King e Paulson, 2007). Nos últimos anos, grandes estúdios de efeitos visuais como ILM, *Digital Domain* e *Sony Pictures Imageworks* têm empregado amplamente esses sistemas para a criação de inúmeros filmes de sucesso, como *Titanic* (1997), a trilogia *O Senhor dos Anéis* (2001, 2002 e 2003), *O Expresso Polar* (2004), *King Kong* (2005) e *Avatar* (2009). Em geral, os sistemas de captura de movimento são utilizados para a animação de personagens virtuais com base em movimentos coletados de seres reais, o que possibilita que os elementos reconstruídos apareçam com maior realismo, de acordo com a execução dos atores.

Também é comum o uso de *mocap* para animação de elementos em segundo plano, como no caso do filme *Titanic*. Neste caso, a técnica foi intensamente empregada para a animação de figurantes e personagens que "caíam" do navio (Menache, 1999). No recente sucesso *Avatar* (2009), o emprego de técnicas avançadas de *mocap* foi fundamental para que o diretor James Cameron pudesse acompanhar, em tempo real, a atuação de alguns personagens virtuais do filme (Elsaesse, 2011). A Figura 2.1 apresenta imagens clássicas do uso de sistemas de *mocap* nos filmes *O Expresso Polar* (a) e *O Senhor dos Anéis* (b).

### 2.1.2 Indústria de Jogos

O desenvolvimento de jogos eletrônicos constitui atualmente um dos maiores mercados de tecnologias de captura de movimento no mundo (Rizaldy, 2011). Empresas produtoras como a *Valve*, *Naughty Dog* e *Ubisoft* utilizam intensamente sistemas de captura para o desenvolvimento de personagens virtuais em jogos, os

quais se movimentam de acordo com a atuação dos atores. Títulos de sucesso como o *Uncharted Drake's Fortune*, lançado pela *Sony Computer Entertainment*, e o *FIFA Soccer 12*, desenvolvido pela *EA Canada*, são exemplos de produções que demandaram intensa aplicação de *mocap*.

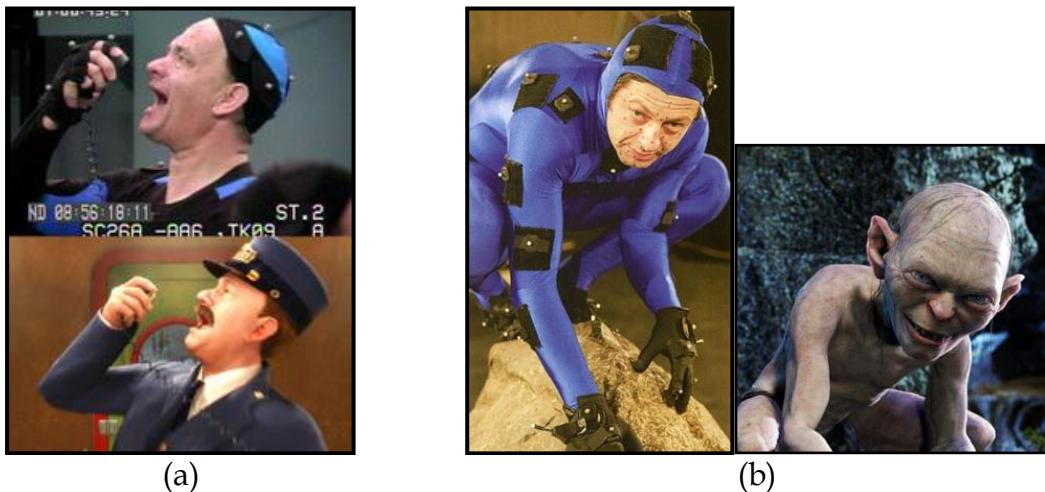


Figura 2.1 - Exemplos clássicos de aplicação de sistemas de *mocap* na indústria cinematográfica. A imagem (a) ilustra o ator Tom Hanks tendo os movimentos capturados para o personagem do filme *O Expresso Polar* (Ign, 2004). Em (b), a captura é feita para o personagem *Smeagol* da trilogia *O Senhor dos Anéis* (Ign, 2003).

Recentemente, os fabricantes Nintendo, Microsoft e Sony desenvolveram para seus respectivos aparelhos, Nintendo Wii, XBOX 360 e Playstation 3, tecnologias de baixo custo para o reconhecimento de alguns movimentos corporais dos jogadores, abrindo espaço para novas possibilidades de imersão e interação em ambientes domésticos.

A Figura 2.2 ilustra o jogador Maurice Edu participando de uma sessão de captura de movimento para o jogo *FIFA Soccer 9*.



Figura 2.2 – Imagem do jogador Maurice Edu em estúdio de *mocap* (Pissanos, 2008). O jogador teve os movimentos coletados para o jogo de futebol *FIFA Soccer 9* produzido pela *EA Sports*.

### 2.1.3 Esportes

Os sistemas de *mocap* têm sido utilizados para avaliação do desempenho, treinamento e correção de movimentos em atletas de diversas modalidades esportivas (Barris e Button, 2008; Dolná, Šimšík *et al.*, 2008; Cockcroft, 2010). Um ginasta, por exemplo, tem a possibilidade de analisar detalhadamente os seus próprios movimentos, de qualquer ângulo de visão, após registrá-los durante uma sessão de *mocap*. Parâmetros cinemáticos como velocidades, acelerações e angulações são facilmente obtidos a partir dos dados tridimensionais calculados. Todas as informações derivadas da reconstrução podem ser utilizadas pelo atleta e pelo treinador no processo de busca por imperfeições, em comparações com outros atletas e até mesmo na busca por limitações físicas, as quais são de grande interesse para ajudar na prevenção de lesões (Barris e Button, 2008).

Os sistemas de *mocap* podem ser utilizados até mesmo para a captura de movimentos em ambientes subaquáticos. A Figura 2.3 ilustra um nadador prestes a ter os movimentos capturados por um sistema óptico da empresa Qualisys. Neste caso, um conjunto de câmeras especiais é mantido sob a água.



Figura 2.3 – Sistema de *mocap* da Qualisys sendo empregado na captura dos movimentos de um nadador. Câmeras submersas possibilitam a coleta embaixo d'água. Imagem adaptada de (Qualisys, 2013).

### 2.1.4 Saúde e Biomecânica

O emprego de *mocap* em áreas da saúde tem se tornado cada vez mais comum. A fisioterapia, a ortopedia, a neurologia e a odontologia são apenas algumas das diversas ciências que têm sido beneficiadas e, ao mesmo tempo, têm contribuído para o desenvolvimento das técnicas (Moeslund, Hilton *et al.*, 2006). Por exemplo, para a

análise clínica da marcha de um indivíduo, as informações temporais e espaciais relativas a seus membros e articulações, obtidas com sistemas de *mocap*, podem ser valiosas no diagnóstico de diversas doenças (Davis, 1988; Yeasin e Chaudhuri, 2000; Johnson e Bobick, 2001; Troje, 2002; Saboune e Charpillet, 2005). Combinando *mocap* e modelagem biomecânica, treinadores, fisioterapeutas, ortopedistas e neurologistas podem atuar juntos na análise do movimento humano para identificarem o estado de saúde do paciente e proporem tratamentos e programas de reabilitação (Davis, 1988; Sutherland, 2002; 2005; McNamara, Bing Feng Johnson *et al.*, 2007; Jeffrey, Ajay *et al.*, 2011). Em testes neurológicos, os sistemas de *mocap* têm auxiliado no diagnóstico antecipado de lesões cerebrais (Gandolla, Ferrante *et al.*, 2011). Isto é feito mediante a detecção de pequenas variações nos movimentos que os diferenciam dos padrões considerados normais. Essas variações nem sempre são evidentes por meio de inspeção visual.

Outra importante aplicação ocorre na odontologia, onde os sistemas de captura têm sido empregados para registrar o traçado tridimensional da mandíbula durante a realização de movimentos como os de abertura e fechamento da boca (Yoon, Zhao *et al.*, 2006; Fang e Kuo, 2008; Pinheiro, Andrade *et al.*, 2008; Santos, Tavares *et al.*, 2008). Uma análise detalhada e quantitativa desses movimentos é de grande valia para auxiliar o especialista no diagnóstico e tratamento de disfunções temporomandibulares (Travers, Buschang *et al.*, 2000; Fukui, Tsuruta *et al.*, 2002; Naeije, 2003; Furtado, Pereira *et al.*, 2010). Isto é ilustrado na Figura 2.4.

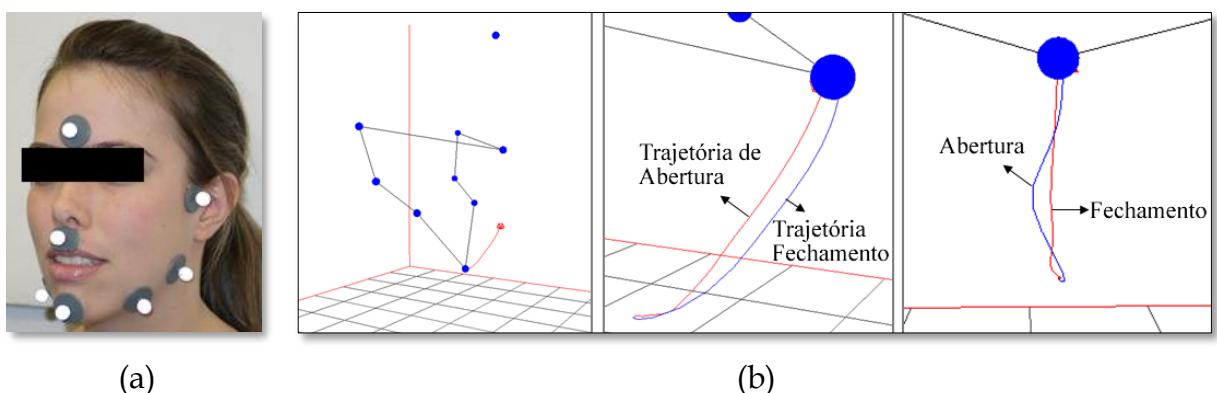


Figura 2.4 - (a) Indivíduo usando um conjunto de marcadores para captura aproximada dos movimentos mandibulares; (b) Reconstrução realizada por um sistema de *mocap* após abertura e fechamento da mandíbula (Furtado, Pereira *et al.*, 2010). A trajetória do marcador posto sobre o queixo é destacada.

### 2.1.5 Engenharia

A indústria automobilística tem empregado os sistemas de *mocap* no projeto de interiores de carros para avaliação e aperfeiçoamento do conforto, da ergonomia e de elementos de segurança para o motorista. Por exemplo, engenheiros e projetistas da Ford utilizam sistemas de *mocap* para entender melhor como as pessoas de diferentes portes se acomodam e interagem com seus veículos (Lareau, 2011). Em uma sessão de captura, um indivíduo usando marcadores executa uma série de movimentos que são normalmente realizados pelos motoristas. Os dados obtidos dos sensores são utilizados para a elaboração de modelos digitais, que são empregados em vários testes conduzidos virtualmente. A Figura 2.5 ilustra um sujeito em um ambiente de captura utilizado pela Ford.



Figura 2.5 – Exemplo de uso de *mocap* no projeto de veículos da Ford. Imagem adaptada de (Motortrend, 2011).

### 2.1.6 Militar

No campo militar, os sistemas de *mocap* têm auxiliado oficiais em sessões de treinamento. A captura do movimento de vários soldados durante uma simulação, por exemplo, permite ao treinador ter uma visão de 360°, em tempo real, de todas as ações conduzidas (Motion, 2011). Caso seja necessário, o treinador pode realizar uma avaliação detalhada do treinamento mediante análise dos dados tridimensionais reconstruídos pelo software de *mocap*. A Figura 2.6 ilustra algumas ações militares sendo capturadas com sistemas ópticos.



Figura 2.6 – Exemplos de aplicação de sistemas de *mocap* em treinamentos militares. (a) Treinamento de soldados dos EUA (Society, 2011). (b) Imagem adaptada de (Herald, 2010).

## 2.2 Tecnologias de Captura de Movimento

Conforme já mencionado, atualmente existem diversas tecnologias e métodos empregados no processo de captura de movimento. Cada tecnologia oferece vantagens e desvantagens, as quais podem torná-la mais ou menos adequada para uma ou outra aplicação de captura (Kitagawa e Windsor, 2008). Esta seção apresenta brevemente as principais categorias de sistemas e os respectivos recursos tecnológicos envolvidos.

### 2.2.1 Sistemas Eletromecânicos

A técnica eletromecânica é conhecida como uma das pioneiras na área de captura de movimento. Os sistemas eletromecânicos se baseiam em uma estrutura rígida, articulada, composta por barras e potenciômetros que atuam como uma espécie de exoesqueleto, o qual é “vestido” pelo usuário. As barras sustentam os potenciômetros nas regiões das articulações e cada potenciômetro tem a função de medir o ângulo da articulação à qual está associado. Os dados coletados são enviados para um módulo de processamento, que calcula o deslocamento relativo dos membros do usuário à medida que ele se movimenta (Corrales, Candelas *et al.*, 2008).

Sistemas eletromecânicos funcionam em tempo real, são relativamente baratos, livres de oclusão (como ocorre em sistemas ópticos) e interferências eletromagnéticas e

possibilitam a captura de movimento em quase todos os locais e ambientes (Vlasic, Adelsberger *et al.*, 2007; Kitagawa e Windsor, 2008). Uma desvantagem desses sistemas está no fato de que o deslocamento global do conjunto é difícil de ser calculado com precisão. Para compensar este problema, alguns sistemas mecânicos incorporam também sensores magnéticos e componentes utilizados em sistemas inerciais, como acelerômetros e giroscópios. Outra desvantagem notável é a obstrução causada pelo exoesqueleto, que pode trazer desconforto quando usado por muito tempo e pode restringir os movimentos do usuário.

Um exemplo de sistema eletromecânico é o Gypsy-7 da inglesa Animazoo. Este sistema tem preço a partir de U\$ 8.000,00 (Animazoo, 2013) e é baseado em 15 sensores de articulação e dois giroscópios. A Figura 2.7 ilustra esse sistema em uso.



Figura 2.7 - Ilustração do sistema de captura de movimento eletromecânico Gypsy-7 da empresa Animazoo. Imagem adaptada de (Animazoo, 2013).

## 2.2.2 Sistemas Magnéticos

Um sistema magnético típico emprega um transmissor eletromagnético estático e uma coleção de sensores magnéticos que é distribuída pelo corpo do usuário. Medindo a percepção espacial do campo magnético em cada sensor, o sistema é capaz de calcular a posição e orientação de todos os sensores com relação ao transmissor (Yoon, Zhao *et al.*, 2006).

Uma vantagem dos sistemas magnéticos sobre os sistemas ópticos está no fato de que os sensores não se tornam oclusos pelo corpo do sujeito ou por objetos não metálicos (Hassan, Jenkyn *et al.*, 2007). Por outro lado, eles podem sofrer de interferência

elétrica e magnética causada por metais e equipamentos eletrônicos do ambiente. Semelhante aos sistemas mecânicos, os fios que conectam os sensores e as baterias também podem restringir a liberdade de movimento do ator (O'brien, Bodenheimer *et al.*, 2000). Outra desvantagem da tecnologia é o limitado volume de captura, que geralmente abrange um raio máximo de 3 metros ao redor do transmissor.

Os sistemas magnéticos são comumente classificados em dois grupos principais: (1) aqueles que utilizam campo eletromagnético gerado por corrente contínua (DC), como o MotionStar da Ascension Technology Corporation e (2) aqueles que geram campo magnético utilizando corrente alternada (AC), como o LIBERTY™ da empresa Polhemus. Os primeiros são bastante sensíveis a alumínio e cobre, enquanto os segundos são mais sensíveis a ferro e aço.

A velocidade de aquisição dos dados (geralmente até 240 amostras por segundo) é suficiente para muitas aplicações, mas ainda é bastante inferior às altas taxas obtidas com os sistemas ópticos (até 2000 amostras por segundo). Sistemas magnéticos geralmente apresentam bons resultados em aplicações de tempo real e costumam ser consideravelmente mais baratos do que os sistemas ópticos e inerciais (Kitagawa e Windsor, 2008).

### **2.2.3 Sistemas Inerciais**

Os sistemas inerciais são baseados em modelos biomecânicos e pequenos sensores, que geralmente são formados por acelerômetros e giroscópios. Os sensores são distribuídos pelo corpo do usuário e conectados por fios a um módulo transmissor (transportado pelo indivíduo), que por sua vez transfere os dados obtidos dos sensores a um computador externo utilizando uma conexão com ou sem fios (veja a Figura 2.8). Alguns sistemas inerciais são capazes de capturar os movimentos de rotação e translação (6 graus de liberdade) dos sensores. Neste caso, o giroscópio é utilizado para o cálculo da orientação do sensor enquanto o acelerômetro fornece medidas a respeito do posicionamento espacial.

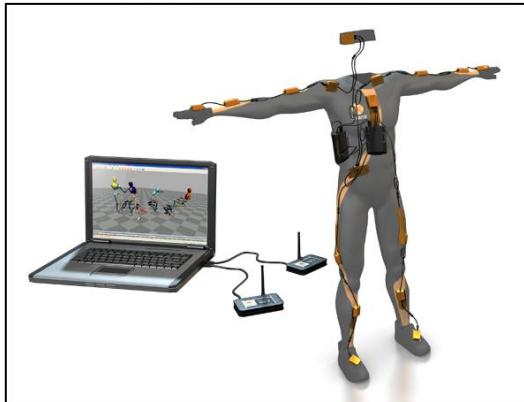


Figura 2.8 - Ilustração do sistema de captura de movimento inercial MVN da empresa Xsens. Imagem adaptada de (Xsens, 2013).

Sistemas inerciais podem ser usados tanto em ambientes internos quanto em ambientes externos, não havendo a necessidade de estúdio especial. Assim como os sistemas magnéticos e os mecânicos, os sistemas inerciais também não sofrem dos problemas de oclusão e perda de marcadores (Bachmann, 2000; Cloete e Scheffer, 2008). Por outro lado, esses sistemas geralmente oferecem baixas taxas de amostragem, trabalham com um número reduzido de sensores (detalhes importantes dos movimentos podem ser perdidos) e são limitados com respeito à exatidão das medidas de translação (Cloete e Scheffer, 2008). O módulo de transmissão e os fios também podem causar obstrução e limitar os movimentos do usuário.

Os sistemas MVN da empresa Xsens e o IGS-190 da Animazoo são exemplos de conjuntos inerciais.

#### 2.2.4 Sistemas Acústicos

Os sistemas de captura de movimento acústicos (ou ultrassônicos) normalmente utilizam três receptores de áudio, que são colocados em locais fixos no ambiente, em conjunto com uma coleção de transmissores espalhados pelo corpo do indivíduo. As distâncias de cada transmissor aos receptores podem ser computadas medindo-se o tempo de transmissão do sinal sonoro, uma vez que é conhecida a velocidade de propagação do som. A posição espacial de cada transmissor é então calculada pela triangulação dessas distâncias (Sato, Nakamura *et al.*, 2011).

Das diversas categorias de sistemas de *mocap*, os sistemas puramente acústicos são os mais raros. Uma desvantagem desses sistemas é a suscetibilidade a ruídos sonoros

do ambiente, que podem interferir na qualidade dos dados reconstruídos. Os sistemas acústicos também são limitados a uma quantidade reduzida de sensores (Vlasic, Adelsberger *et al.*, 2007).

### 2.2.5 Sistemas Ópticos

A tecnologia óptica é reconhecida atualmente como a mais avançada das tecnologias de captura de movimento existentes (Sun, Meng *et al.*, 2010). Um sistema óptico é caracterizado pelo uso de **câmeras de vídeo** que registram **imagens** do indivíduo, animal ou objeto cujos movimentos se deseja capturar. Os sistemas ópticos tradicionais empregam pequenos objetos conhecidos como **marcadores**, os quais são fixados em pontos de interesse do corpo e normalmente aparecem em alto contraste nas imagens capturadas. Tais sistemas têm como meta calcular a trajetória espacial correspondente ao movimento de cada marcador. Recentemente, têm surgido sistemas ópticos que não utilizam marcadores. Esta nova categoria de sistemas tem como objetivo reconstruir os movimentos do corpo pela análise direta dos contornos de suas imagens (Bray, 2006).

Os sistemas que empregam marcadores podem ser classificados em **sistemas passivos** e **sistemas ativos**, dependendo do marcador utilizado ser retrorreflexivo (passivo) ou emissor de luz (ativo), respectivamente. Independente do tipo de marcador, esses sistemas compartilham o mesmo princípio de funcionamento, que consiste na combinação das imagens dos marcadores registradas em várias câmeras para calcular, em um processo de triangulação, suas posições espaciais quadro a quadro (Zhang, 1998a; Hartley e Zisserman, 2004; Guerra-Filho, 2005). Os sistemas passivos, os sistemas ativos e os sistemas que não se baseiam em marcadores são descritos nas próximas seções.

#### 2.2.5.1 Sistemas Baseados em Marcadores Passivos

Um sistema comercial típico baseado em marcadores passivos utiliza uma coleção de câmeras de infravermelho em conjunto com dezenas de marcadores retrorreflexivos que são fixados no corpo do indivíduo. As câmeras desenvolvidas para esses sistemas são equipadas com LEDs (diodos emissores de luz - *light-emitting diodes*) que

emitem luz infravermelha. Parte dessa luz é refletida pelos marcadores e absorvida de volta pela câmera. A iluminação infravermelha em conjunto com o material altamente reflexivo dos marcadores e os filtros de luz integrados às câmeras (que deixam passar apenas o infravermelho) possibilita que os marcadores apareçam em alto contraste nas imagens capturadas. Isto torna possível a detecção dos marcadores nas imagens com menor custo computacional e com maior precisão, o que aumenta a qualidade dos dados tridimensionais reconstruídos a partir dessas imagens (Guerra-Filho, 2005; Ping, Yue *et al.*, 2009).

O tamanho e o formato dos marcadores dependem principalmente do tamanho do volume de captura, da resolução das câmeras e do tipo de movimento a ser capturado (corporal, facial, etc.). Marcadores esféricos com cerca de 1,6 cm ou mais de diâmetro são comumente empregados na captura de movimentos de corpo inteiro, enquanto marcadores semiesféricos com aproximadamente 3 ou 4 mm de diâmetro são utilizados na captura de movimentos faciais. A Figura 2.9 ilustra uma câmera de infravermelho (comercializada pela Vicon), um marcador esférico convencional de 1,6 cm e um conjunto de marcadores semiesféricos de 3 mm.

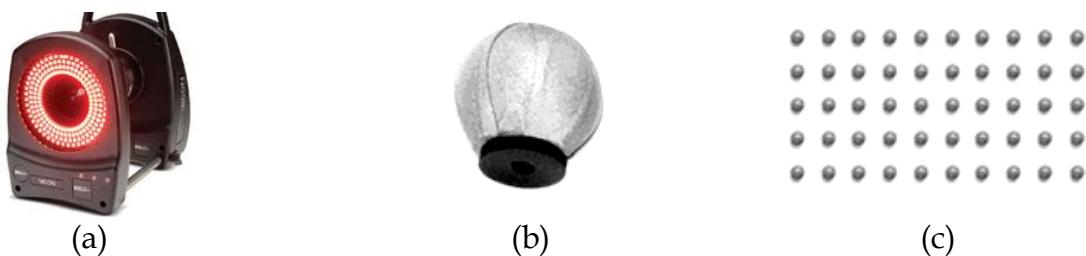


Figura 2.9 – Exemplos de dispositivos utilizados em sistemas ópticos passivos. (a) Câmera de infravermelho comercializada pela empresa Vicon; (b) marcador esférico de 1,6 cm comumente utilizado para captura de movimentos corporais; (c) conjunto de marcadores semiesféricos de 3 mm normalmente utilizados em capturas faciais.

Os marcadores podem ser fixados diretamente sobre o corpo do indivíduo (ou objeto) utilizando fita adesiva, ou podem ser usados em conjunto com roupa especialmente desenvolvida para facilitar a fixação.

Uma das grandes vantagens do uso de câmeras especiais de infravermelho está no fato de que a detecção dos marcadores nas imagens pode ser realizada pelo próprio hardware da câmera (Ping, Yue *et al.*, 2009). Isto possibilita que apenas as coordenadas dos marcadores nas imagens (ao invés da imagem completa) sejam

enviadas para o computador encarregado do processo de triangulação e rastreamento. Este processo pode ser feito em tempo real, para um número consideravelmente alto de câmeras e em taxas de captura que podem chegar a até 2000 quadros por segundo, o que é possível devido à drástica redução do tráfego de dados e pelo processamento distribuído das imagens realizado nas próprias câmeras (Papic, Zanchi *et al.*, 2004; Castro, Medina-Carnicer *et al.*, 2006).

Apesar das câmeras de infravermelho serem utilizadas pela maioria dos sistemas comerciais passivos, há também sistemas de baixo custo que utilizam câmeras de vídeo convencionais. Nesse caso, todos os quadros capturados precisam ser inteiramente transmitidos para um computador, o que nem sempre é possível de ser feito em tempo real, dependendo do número de câmeras, da resolução das imagens, da taxa de captura e dos meios de transmissão utilizados.

Para que a triangulação seja calculada em um sistema óptico é preciso que cada marcador seja visualizado por pelo menos duas câmeras (Shapiro, Zisserman *et al.*, 1995; Zhang, 1998a). Esta exigência pode levar à ocorrência de um dos problemas mais comuns em sistemas ópticos: a **occlusão de marcadores** (Qiong e Boulanger, 2011). Um marcador se torna oculto quando deixa de ser visualizado por pelo menos duas câmeras. Isto pode acontecer, por exemplo, quando o indivíduo realiza um movimento que coloca os próprios membros do corpo cobrindo a visão de um ou mais marcadores.

Enquanto um marcador estiver oculto, sua trajetória não pode ser reconstruída pelo sistema. Em algumas situações, este problema pode ser contornado em uma etapa de pós-processamento dos dados, quando o usuário tem a possibilidade de editar manualmente os resultados tridimensionais (Barca, Rumantir *et al.*, 2008; Kitagawa e Windsor, 2008). Em sistemas ativos, a identidade de cada marcador é consistente ao longo de toda a captura, mesmo que haja lacunas em sua trajetória, mas um sistema passivo nem sempre é capaz de fornecer a mesma identificação a um marcador que reaparece após alguns instantes de oclusão. Como resultado, sistemas passivos podem gerar múltiplas identificações para um único marcador físico.

A quantidade de câmeras que deve ser empregada em um sistema óptico passivo depende de diversos fatores como o tipo de movimento a ser coletado, o tamanho do volume de captura e o número de participantes envolvidos. Em geral, quanto maior o número de câmeras, menores são as chances de haver oclusão de marcadores. Sistemas que utilizam 6, 8, 12, 24 ou 32 câmeras são comuns, embora um número maior também seja possível (Kitagawa e Windsor, 2008).

Em resumo, as principais vantagens dos sistemas ópticos passivos são:

- Possibilita altas velocidades de captura e altos níveis de precisão e exatidão;
- Os movimentos de múltiplos sujeitos podem ser capturados simultaneamente;
- Um alto número de marcadores pode ser usado e a configuração dos mesmos pode ser alterada facilmente;
- O objeto ou indivíduo pode se movimentar livremente dentro do volume de captura, não sofrendo obstrução por fios ou equipamentos que limitam os movimentos;
- O volume de captura pode ser grande o suficiente para vários tipos de aplicações.

As principais desvantagens desses sistemas são:

- Um pós-processamento intenso dos dados pode ser necessário;
- Possibilidade de oclusão dos marcadores, resultando em perda de dados;
- Hardware e software geralmente mais caros do que os de outros sistemas;
- A maioria dos sistemas necessita que a luz do ambiente seja controlada.

Empresas fabricantes de sistemas ópticos passivos incluem a Vicon, a Motion Analysis Corporation, a Qualisys e a NaturalPoint.

#### **2.2.5.2 Sistemas Baseados em Marcadores Ativos**

Diferente dos sistemas passivos, cada marcador empregado em um sistema óptico ativo corresponde a um dispositivo que emite a própria luz, geralmente um LED. Os marcadores são comumente conectados por fios a uma unidade de controle, que é

transportada pelo indivíduo. Alguns sistemas empregam modulação no tempo, ativando um LED (marcador) de cada vez (Adam, O'brien *et al.*, 2005). Outros sistemas acendem todos os LEDs ao mesmo tempo e utilizam modulação em frequência ou amplitude para distingui-los (Kitagawa e Windsor, 2008). Em ambos os casos os marcadores podem ser identificados diretamente, o que é tido como uma das principais vantagens dos sistemas ópticos ativos. A identificação direta simplifica o problema de rastreamento e identificação e elimina algumas falhas naturais dos sistemas passivos, como a possibilidade de troca de identificação dos marcadores (Barca, Rumantir *et al.*, 2008). Por outro lado, a necessidade de porte do módulo de controle e dos fios ou baterias pode comprometer a liberdade de movimento do usuário e limitar a quantidade de sensores suportados. Alguns sistemas ópticos ativos podem funcionar em ambiente externo, sob luz ambiente, enquanto outros ainda são limitados ao uso interno (Maletsky, Sun *et al.*, 2007).

A PhaseSpace Inc. e a Qualisys são exemplos de empresas que comercializam sistemas ópticos ativos.

Independentemente de utilizar marcadores passivos ou ativos, os sistemas ópticos necessitam ser calibrados, antes de iniciar qualquer captura. A calibração desses sistemas tem como objetivo descobrir as características internas das câmeras e o posicionamento e orientação de cada uma com relação a um ponto de referência (ou câmera). A reconstrução tridimensional só pode ser realizada depois que o conjunto de câmeras é calibrado. Algumas técnicas de calibração empregam objetos especiais mantendo marcadores em posições específicas. Esse assunto é abordado em mais detalhes no Capítulo 4.

#### **2.2.5.3 Sistemas Ópticos sem Marcadores**

Embora a tecnologia óptica baseada em marcadores seja a predominante atualmente, diversos trabalhos recentes têm focado no problema de capturar movimentos corporais usando apenas as câmeras de vídeo, sem o emprego de marcadores (Marzani, Calais *et al.*, 2001; Chu, Jenkins *et al.*, 2003; Mikić, Trivedi *et al.*, 2003; Caillette e Howard, 2004; Saboune e Charpillet, 2005; Sundaresan e Chellappa, 2005;

Bray, 2006; Corazza, Mündermann *et al.*, 2006; Kehl e Gool, 2006; Mündermann, Corazza *et al.*, 2006; Rosenhahn, Brox *et al.*, 2006; Mundermann, Corazza *et al.*, 2007; Sigal, Balan *et al.*, 2009; Corazza, Mündermann *et al.*, 2010; Pons-Moll, Leal-Taix *et al.*, 2011). Vários métodos para rastreamento corporal livre de marcadores foram publicados e sistemas como aquele desenvolvido no laboratório de biomecânica de Stanford (Mundermann, Corazza *et al.*, 2007) estão se tornando notáveis. No trabalho de Steele e seus colegas (Steele, Corazza *et al.*, 2008), por exemplo, os autores comparam o desempenho de um sistema tradicional (baseado em marcadores) com o desempenho de um sistema livre de marcadores. Os sistemas foram usados para calcular os centros das articulações de um sujeito durante uma caminhada. Nesse experimento, os valores produzidos pelos dois sistemas foram similares.

A maior parte das técnicas e dos sistemas propostos que não utilizam marcadores é baseada em um modelo corporal humano. A reconstrução dos movimentos é realizada a partir da análise dos contornos das imagens do corpo, registradas por várias câmeras (Mündermann, Corazza *et al.*, 2006). Em geral, os métodos requerem um ambiente bastante controlado para possibilitar a eliminação eficiente das imagens de fundo da cena.

Apesar dos recentes avanços dessas técnicas, os sistemas existentes ainda não são capazes de rastrear movimentos complexos e detalhados, com precisão e velocidade suficientes para possibilitar o processamento em tempo real (Oshita, 2006; Theobalt, Aguiar *et al.*, 2010). Devido a esses fatores, os sistemas que se baseiam em marcadores ainda são preferidos pela maior parte dos usuários de *mocap*. As empresas Organic Motion e Mova são exemplos de fabricantes de sistemas ópticos livres de marcadores.

# Capítulo 3

## Trabalhos Relacionados

Os principais trabalhos diretamente relacionados a esta pesquisa podem ser organizados, sob uma perspectiva geral, em dois grupos. No primeiro grupo estão os trabalhos que se baseiam em algum tipo de *modelo anatômico humano* para conduzir alguma etapa ou tarefa vinculada ao processo de reconstrução do movimento utilizando tecnologia óptica passiva. O segundo grupo compreende os trabalhos que apresentam técnicas ou algoritmos relacionados à captura de movimento, mas que buscam a reconstrução de uma coleção de marcadores que podem se mover independentemente, *sem o suporte de um modelo*.

Os sistemas que se baseiam em um modelo esquelético humano geralmente o fazem para guiar o processo de identificação, rastreamento ou reconstrução dos marcadores. O objetivo essencial é utilizar o conhecimento de como o esqueleto se move para melhor estimar a movimentação dos marcadores. Por exemplo, as restrições cinemáticas impostas pelas limitações angulares das articulações e os comprimentos dos ossos podem ser utilizadas para restringir o espaço de busca e resolver ambiguidades durante o processo de reconstrução da trajetória de cada marcador (Herda, Fua *et al.*, 2001; Pandy, 2001). Como resultado, os dados reconstruídos podem ser mais robustos e podem necessitar de menor esforço do usuário em uma etapa de pós-processamento.

Em contrapartida, o emprego de um modelo esquelético como ferramenta essencial para conduzir a reconstrução geralmente traz várias limitações. Por exemplo, as técnicas que se apoiam em um modelo anatômico humano não podem ser usadas para reconstruir livremente o movimento de outros corpos, como o de objetos e animais. Além disso, muitas dessas técnicas exigem que o usuário execute uma sequência especial de movimentos antes da captura, ou que ele se mantenha por alguns segundos em uma determinada pose (como em forma de T), para que os parâmetros do modelo e o posicionamento relativo dos marcadores sejam adequadamente coletados (Herda, Fua *et al.*, 2000; 2001; Ringer e Lasenby, 2004). Em várias aplicações nos esportes e na saúde, por exemplo, tais exigências são impraticáveis. Um atleta paralímpico ou um indivíduo com debilidade motora severa nos membros superiores pode ser incapaz de executar a posição T, o que poderia impossibilitá-lo de participar de uma sessão de captura. Outra desvantagem é a pouca flexibilidade com relação ao número e posicionamento dos marcadores.

As técnicas que não se baseiam em um modelo não possuem essas limitações, mas também não podem gozar dos benefícios das restrições naturais da anatomia corporal para conduzir a reconstrução. Como resultado, a rotulação automática das trajetórias dos marcadores nem sempre é possível e o processo de rastreamento é mais suscetível a falhas. Isto pode exigir mais esforços do usuário em uma eventual etapa de pós-processamento. Vale destacar que, quando os marcadores são rastreados de maneira independente, as informações relativas ao modelo esquelético subjacente (se houver um) podem ser posteriormente recuperadas, desde que a quantidade e o posicionamento dos marcadores viabilizem o processo (Silaghi, Plänkers *et al.*, 1998; Hiniduma Udugama Gamage e Lasenby, 2001; Adam, O'brien *et al.*, 2005).

Os principais trabalhos disponíveis na literatura que se baseiam em algum tipo de modelo esquelético são apresentados brevemente a seguir. Posteriormente, são apresentados aqueles trabalhos com foco na reconstrução livre de modelo.

### 3.1 Técnicas Baseadas em Modelo

---

Ambrósio e seus colegas (Ambrósio, Abrantes *et al.*, 2001) utilizaram um modelo biomecânico de 16 segmentos para suportar a reconstrução tridimensional dos movimentos de um indivíduo utilizando apenas uma câmera de vídeo. Equações matemáticas foram empregadas para modelar as restrições a respeito do tamanho dos segmentos anatômicos. Tal modelagem substituiu, em conjunto com o algoritmo DLT (*direct linear transformation*) (Abdel-Aziz e Karara, 1971), a necessidade de uma segunda câmera no processo de reconstrução. Os autores aplicaram a técnica para reconstruir apenas movimentos simples e destacaram que o método não pode ser conduzido quando o modelo biomecânico empregado não oferece os graus de liberdade necessários para representação do movimento. Além disso, os baixos valores de precisão e exatidão associados ao método limitam seu uso para um pequeno grupo de aplicações.

Ringer e Lasenby (2000) apresentaram em seu trabalho um modelo matemático para o problema de descobrir a correspondência entre os pontos detectados nas imagens e os marcadores físicos que geraram esses pontos (Sethi, Salari *et al.*, 1988). A resolução desse problema possibilita que as detecções correspondentes ao mesmo marcador físico sejam utilizadas no processo de triangulação de suas coordenadas espaciais. Os autores mostraram que encontrar a solução ideal, quando o número de câmeras é superior a dois, pode ser computacionalmente inviável (devido ao grande número de possibilidades). Em vez disso, Ringer e Lasenby propuseram o uso de heurísticas e de um modelo esquelético cujos parâmetros devem ser conhecidos a priori. Para conduzir o rastreamento dos marcadores nas imagens, os autores empregaram filtros de Kalman estendidos e filtros de partículas (Gordon, Salmond *et al.*, 1993). Assim como o filtro de Kalman, o filtro de partículas tem sido usado para resolver problemas de previsão e controle em sistemas dinâmicos, mas com foco nos problemas não lineares e não gaussianos (Ristic, Arulampalam *et al.*, 2004). O filtro de partículas demanda custo computacional ainda maior que o filtro de Kalman, o que inviabiliza a técnica proposta pelos autores para muitas aplicações de tempo real.

Um dos primeiros trabalhos a propor o emprego de um modelo esquelético humano para melhorar a eficiência do processo de reconstrução de movimentos em sistemas ópticos foi o de Herda *et al.* (2001). A técnica apresentada pelos autores é baseada essencialmente em um modelo esquelético e é capaz de rastrear e reconstruir as trajetórias de uma coleção de marcadores, desde que as informações do modelo corporal do sujeito sejam conhecidas de antemão. Se alguns marcadores isolados se tornam oclusos ao longo de um curto período de tempo, suas posições podem ser estimadas a partir de marcadores da vizinhança que compartilham relações cinemáticas. Como desvantagem, a técnica exige que todos os marcadores sejam detectados e identificados com relação ao modelo esquelético logo no primeiro quadro, momento em que o usuário precisa interagir e identificar manualmente todos os marcadores.

Papic e seus colegas (Papic, Zanchi *et al.*, 2004) propuseram um procedimento para aquisição e processamento dos parâmetros cinemáticos do sistema locomotor humano. Duas câmeras comerciais foram empregadas em conjunto com a técnica DLT (Abdel-Aziz e Karara, 1971) para calibração e reconstrução das coordenadas espaciais dos marcadores. Os autores apresentaram um processo de sincronização das imagens por software utilizando filtros e técnicas de interpolação. O sistema desenvolvido pode ser utilizado para capturar apenas movimentos simples e a técnica é bastante limitada por não considerar o problema de oclusão de marcadores, por não poder ser usada em tempo real e por possuir baixos valores de precisão e exatidão quando comparados aos padrões comerciais (Richards, 1999).

Em um trabalho mais recente, Ringer e Lasenby (2004) apresentaram um procedimento para estimar automaticamente os parâmetros do modelo esquelético sobre o qual os marcadores são colocados, como o comprimento dos membros e a posição espacial das articulações. Os autores propuseram a realização de uma curta sequência de movimentos, antes da captura, que denominaram de *sequência de treinamento*. Esta sequência possibilita a obtenção das trajetórias dos marcadores, as quais são posteriormente empregadas para o cálculo dos parâmetros do modelo. A

sequência de treinamento deve ser executada lentamente e de maneira cuidadosa, para que todos os marcadores permaneçam no campo de visão de pelo menos duas câmeras. Para combinar adequadamente os pontos registrados pelas câmeras e reconstruir as coordenadas 3D dos marcadores, os autores aplicaram uma técnica de aproximação conhecida como Relaxamento de Lagrange (Lemaréchal, 2001). Esta técnica possibilita a obtenção de uma solução aproximada para problemas complexos e se baseia na inserção de relaxamentos nas restrições originais do problema a ser resolvido. O rastreamento 3D dos marcadores durante a sequência de treinamento é feito utilizando estimadores simples, baseados em posições calculadas nos quadros anteriores. Os autores não consideraram o problema de oclusão de marcadores, pois é esperada que a sequência de treinamento seja conduzida com movimentos simples e lentos para que todos os marcadores permaneçam visíveis. Também não mencionaram a capacidade de execução do procedimento em tempo real.

Chai e Hodgins (2005) propuseram um método para animação de modelos humanos tridimensionais utilizando um pequeno número de marcadores e um banco de dados de movimentos pré-gravados. O método é bastante efetivo em reconstruir os movimentos de corpo inteiro usando um reduzido número de marcadores, mas informações a respeito do modelo esquelético precisam ser conhecidas antecipadamente para que o banco de dados de movimentos possa ser escalonado e ajustado ao indivíduo.

Johnson e seus colegas (Johnson, Xie *et al.*, 2010b) apresentaram uma técnica de rastreamento de marcadores no espaço bidimensional para uso em aplicações de análise de movimento no campo da fisioterapia. O método apresentado se baseia na condição de proximidade entre os marcadores de um quadro a outro e nas restrições de distância existentes entre pares de marcadores que são colocados em uma mesma estrutura rígida ou membro do corpo. O trabalho tem foco essencialmente no rastreamento bidimensional dos marcadores, que parte de um quadro de referência pré-identificado. Não são cobertos aspectos de identificação automática, reconstrução e rastreamento tridimensional. Além disso, a necessidade de identificação manual

(ou semiautomática) dos marcadores no primeiro quadro pode tornar a técnica inconveniente para aplicações que demandam maior agilidade no processo de captura e reconstrução.

### 3.2 Técnicas Livres de Modelo

---

Vários sistemas ópticos de baixo custo, empregando câmeras de vídeo convencionais e não baseados em modelo, foram propostos nos últimos anos (Figueroa, Leite *et al.*, 2003; Castro, Medina-Carnicer *et al.*, 2006; Pinheiro, 2008; Pinheiro, Andrade *et al.*, 2008). A maioria desses sistemas processa e sincroniza os quadros de vídeo em software, em modo *off-line*, para detectar e rastrear cada marcador na sequência de imagens. Consequentemente, eles não são focados no rastreamento e reconstrução em tempo real dos marcadores e geralmente disponibilizam uma baixa precisão e um volume de reconstrução insuficiente para muitas aplicações.

Shafiq e seus colegas apresentaram em seu trabalho (Shafiq, Tümer *et al.*, 2001) vários algoritmos e aspectos teóricos empregados no desenvolvimento de um software de *mocap* para análise de marcha. Seis câmeras CCD foram utilizadas e as técnicas propostas incluem algoritmos para detecção dos centros dos marcadores nas imagens e geração das trajetórias tridimensionais. O problema de correspondência das detecções dos marcadores nas imagens de múltiplas câmeras foi tratado com base no conceito de linhas epipolares da geometria epipolar (Zhang, 1998a; Hartley e Zisserman, 2004). Os autores propuseram a realização do casamento dos pontos nas imagens para cada grupo de quadros, o que permite o cálculo de um conjunto de pontos 3D para cada instante. As trajetórias dos marcadores foram reconstruídas considerando os pontos tridimensionais obtidos em instantes consecutivos, que são conectados de acordo com a proximidade espacial e pelo uso de técnica de predição baseada em extração. Uma das principais desvantagens dessa técnica é o alto custo computacional necessário para resolver o problema de casamento dos marcadores quadro a quadro, o que pode inviabilizar seu uso para aplicações que dependem da execução em tempo real. Além disso, nos testes realizados pelos autores, o sistema desenvolvido mostrou dificuldades para reconstruir as trajetórias

3D de marcadores que se cruzam no espaço. Esta característica não chega a ser um problema para a análise de marcha, por exemplo, onde tais cruzamentos geralmente não acontecem, mas pode comprometer os resultados em aplicações que demandam maior liberdade e dinamismo dos movimentos (Kitagawa e Windsor, 2008).

Figueroa e seus colegas (Figueroa, Leite *et al.*, 2003) propuseram um método de rastreamento que pode ser dividido em três etapas: segmentação, casamento (*matching*) e predição. A segmentação visa detectar os marcadores nas imagens capturadas, enquanto as etapas de casamento e predição utilizam filtros de Kalman para prever e rastrear os marcadores de um quadro a outro. O método proposto pelos autores não considera o problema de oclusão e sobreposição dos pontos nas imagens e não é adequado para processamento em tempo real devido ao alto custo computacional associado à detecção dos marcadores nas imagens e ao filtro de Kalman. Nos experimentos conduzidos, a técnica necessitou de dois segundos para processar cada quadro de vídeo de 530 x 100 pixels.

Castro e seus colegas (Castro, Medina-Carnicer *et al.*, 2006) descreveram brevemente o desenvolvimento de um sistema de *mocap* de baixo custo que emprega câmeras de vídeo convencionais e pode ser utilizado em ambientes externos. Os autores apresentaram um objeto de calibração próprio, composto por barras de alumínio que sustentam um total de 16 marcadores. O algoritmo DLT é empregado para calibração e reconstrução tridimensional. A principal desvantagem da técnica está no fato de que o usuário pode ser requisitado para resolver ambiguidades durante o processo de rastreamento e reconstrução.

Pinto e seus colegas (Pinto, Buaes *et al.*, 2008) apresentaram um sistema de baixo custo nomeado *Bratrack*. O sistema proposto utiliza duas câmeras comuns de conexão USB adaptadas para funcionamento com filtros de luz e LEDs de infravermelho. O conjunto foi desenvolvido para uso em aplicações de realidade virtual e realidade aumentada e é capaz de rastrear no máximo 10 objetos contendo de 3 a 5 marcadores reflexivos cada.

Em seu trabalho, Pinheiro (2008) também propôs um sistema de baixo custo para reconstruir os movimentos corporais. O processo de rastreamento e reconstrução tridimensional é guiado manualmente pelo usuário e a técnica DLT empregada para calibração não incorpora um modelo matemático para compensação das distorções das lentes das câmeras (Pribanic, Sturm et al., 2005), o que faz com que os dados reconstruídos possuam baixa exatidão e precisão quando os marcadores se afastam da região definida pelo objeto de calibração (Hartley e Zisserman, 2004).

Shih-Pin e seus colegas (Shih-Pin, Yi-Yao et al., 2009) propuseram o emprego de *webcams* para o desenvolvimento de um sistema de *mocap* de baixo custo e de operação em tempo real. Os autores utilizaram bolas de tênis de mesa coloridas como marcadores e métodos dinâmicos de processamento de imagens para extração de suas posições. Devido à baixa qualidade das imagens, a técnica CONDENSATION (*CONDitional DENSity propagATION*) (Isard e Blake, 1998) foi empregada para aprimorar a qualidade dos dados reconstruídos em um processo de refinamento dos dados atuais com base em aprendizado adquirido de reconstruções passadas. Embora o artigo desperte atenção ao propor um sistema de tempo real utilizando *webcams*, os próprios autores concluem que o método não é rápido o suficiente para processar todos os quadros (30 QPS) das quatro câmeras em tempo real. Baixos valores de exatidão foram notificados, o que pode ser justificado, em parte, pelo uso isolado do algoritmo DLT no processo de calibração e reconstrução. O artigo não cobre aspectos relacionados ao rastreamento dos marcadores.

David Flam e seus colegas (Flam, Queiroz et al., 2009) apresentaram um software de código aberto para sistemas de captura de movimento. O software desenvolvido pelos pesquisadores, denominado *OpenMoCap*, inclui rotinas para detecção dos marcadores nas imagens, reconstrução e rastreamento no espaço bidimensional. Ao contrário de muitos outros trabalhos, os autores não utilizaram o algoritmo DLT para calibração e reconstrução. A técnica empregada por eles, baseada no trabalho de Fraga e Silva (Fraga e Vite-Silva, 2008), não possibilita a reconstrução das coordenadas absolutas dos marcadores no espaço. Os dados resultantes da

triangulação são definidos como fatores de escala e podem ser utilizados apenas para posicionamento relativo. Duas câmeras de infravermelho foram empregadas nos experimentos e capturas em tempo real foram obtidas apenas para movimentos simples e a uma velocidade de 50 quadros por segundo (metade da velocidade máxima suportada pelas câmeras). O método possui outras limitações importantes, como a necessidade de identificação manual dos marcadores nas primeiras imagens e o emprego de um modelo de câmera que não trata as distorções das lentes, o que compromete a precisão e exatidão dos dados 3D (Tsai, 1987).

### **3.3 Considerações do Capítulo**

---

Neste capítulo foram apresentados os principais trabalhos disponíveis na literatura científica que estão diretamente relacionados a esta pesquisa. Os artigos foram organizados em dois grupos: (1) aqueles que se baseiam, em alguma instância, em um modelo esquelético para conduzir a reconstrução dos movimentos; e (2) aqueles trabalhos que propõem métodos e algoritmos ligados ao processo de reconstrução, mas sem o suporte de um modelo pré-definido.

As técnicas baseadas em modelo utilizam restrições e parâmetros acerca da estrutura esquelética para conduzir o rastreamento e a reconstrução tridimensional dos marcadores, mas possuem como desvantagem a necessidade de conhecimento a priori desses elementos.

Por outro lado, as propostas que tratam os marcadores como uma nuvem de pontos independentes não gozam dos benefícios do modelo, mas são mais flexíveis com relação ao tipo de movimento a ser capturado, quanto ao número de marcadores e suas posições e também são mais abrangentes no que diz respeito às possíveis aplicações. Conforme apresentado neste capítulo, a maior parte dos trabalhos que propõem técnicas ou algoritmos para reconstrução de um conjunto de marcadores, sem o suporte de um modelo, o faz para execução em modo *off-line*. As propostas que visaram conduzir a reconstrução em tempo real, ou possuem limitações severas com relação ao número de câmeras, velocidades de captura e/ou número de marcadores, ou possuem baixa precisão e exatidão quando comparadas aos padrões comerciais.

# Capítulo 4

## Fundamentação Teórica

Este capítulo apresenta os aspectos teóricos fundamentais que constituem a base para a compreensão da técnica de reconstrução proposta. São abordados assuntos que abrangem o princípio de funcionamento de uma câmera e a teoria matemática envolvida no processo de reconstrução tridimensional. Esses elementos são expostos para permitir que o leitor compreenda, por exemplo, como acontece o mapeamento (ou projeção perspectiva) de um ponto 3D do espaço para o plano de imagem da câmera. Também são apresentadas propriedades geométricas importantes que relacionam duas câmeras focalizando uma mesma cena espacial, o que é conhecido como *geometria epipolar*.

O capítulo está organizado da seguinte forma. A Seção 4.1 descreve o modelo matemático da câmera de *pinhole* e o modelo de distorção de Brown, que são empregados para modelagem da projeção perspectiva. A Seção 4.2 apresenta aspectos fundamentais sobre o processo de calibração de câmeras. A Seção 4.3 descreve o conceito e as propriedades da geometria epipolar e a Seção 4.4 encerra o capítulo apresentando as principais técnicas empregadas para conduzir a reconstrução tridimensional.

## 4.1 Modelo Matemático de uma Câmera

O modelo matemático de uma câmera estabelece a relação geométrica entre as coordenadas de um ponto no espaço (3D) e sua projeção no plano de imagem da câmera (2D). Nesta seção, são apresentados dois modelos: (a) o *modelo de câmera de pinhole*, que não incorpora nenhum tipo de distorção óptica presente nas câmeras atuais; e (b) o *modelo de câmera de Brown*, que considera distorções radiais e tangenciais.

### 4.1.1 Modelo de Câmera de Pinhole

Uma câmera de *pinhole* é um dispositivo simples, na forma de uma caixa fechada, que possui um pequeno orifício em um de seus lados. A luz do mundo exterior entra na câmera através do orifício, o que possibilita a formação de uma imagem correspondente no lado oposto da caixa (Hartley e Zisserman, 2004) (veja Figura 4.1). A geometria da câmera de *pinhole* é comumente utilizada para composição de um modelo básico de câmera, chamado *modelo de câmera de pinhole*. Este modelo é empregado para representação aproximada de câmeras mais complexas atuais, como aquelas que utilizam sensores CCD (*Charge-coupled device*) (Weng, Cohen *et al.*, 1992).

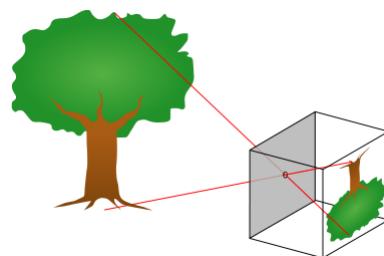


Figura 4.1 - Ilustração de uma câmera de *pinhole*. A luz que entra na câmera pelo orifício permite a formação de uma imagem no lado oposto da caixa.

O modelo de câmera de *pinhole* especifica a relação matemática entre as coordenadas de um ponto no espaço (3D) e sua projeção no plano de imagem da câmera (2D). Vale ressaltar que o modelo não inclui alguns elementos presentes em câmeras atuais, como lentes e suas distorções, sendo ele utilizado como uma aproximação para o mapeamento real (Hartley e Zisserman, 2004; Ivezkovic, Fusiello *et al.*, 2006).

A câmera de *pinhole* é definida pelo seu *centro óptico*  $C$ , que também é conhecido como *centro de projeção da câmera*, e pelo *plano da imagem*  $PI$ . Em uma câmera de *pinhole* real, o centro óptico corresponde à abertura (ou furo) que permite a passagem da luz e o plano da imagem é localizado atrás do centro óptico, como na Figura 4.1. Entretanto, para uma análise teórica conveniente e mais simples, é comum considerar o plano de imagem da câmera como um plano virtual localizado na frente do centro óptico, como mostra a Figura 4.2(a). Esta simplificação não traz prejuízos para a modelagem matemática da câmera e é considerada deste ponto em diante.

A distância do centro óptico ao plano da imagem corresponde à *distância focal*  $f$ . A reta que se origina no centro óptico e é perpendicular ao plano da imagem é denominada *eixo principal* ou *eixo óptico*. O eixo principal intercepta o plano da imagem no *ponto principal*  $P$  e o plano paralelo ao plano da imagem que contém o centro óptico é chamado *plano principal* ou *plano focal* da câmera. Todos esses elementos são apresentados na Figura 4.2(a).

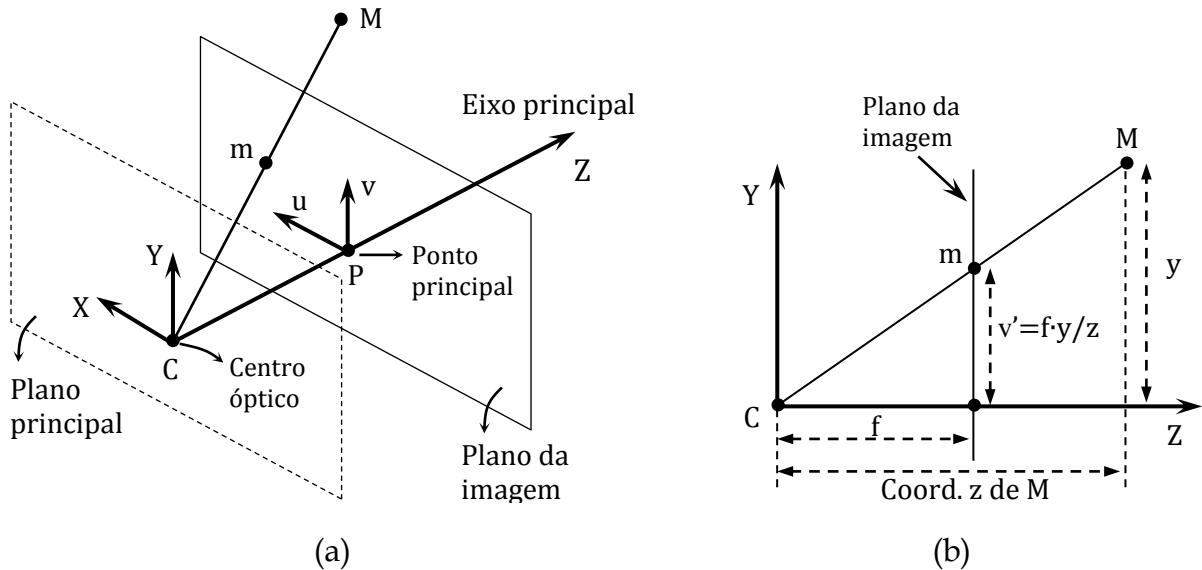


Figura 4.2 - Geometria da câmera de pinhole. (a) Visão em perspectiva. (b) Visão no plano YZ. A coordenada vertical do ponto projetado  $m$  no plano de imagem pode ser calculada por semelhança de triângulos, ou seja:  $f/z = v'/y$ , logo  $v' = f \cdot y / z$ . O mesmo processo pode ser empregado para o cálculo da coordenada horizontal de  $m$  no plano de imagem. Figuras baseadas em (Ivekovic, Fusiello et al., 2006).

A relação entre as coordenadas 3D de um ponto no espaço e as coordenadas de sua projeção no plano da imagem pode ser definida pela *projeção perspectiva*. Um ponto 3D é projetado no plano de imagem da câmera pela reta contendo o ponto e o centro

óptico. Considerando o centro óptico como a origem de um sistema de coordenadas Euclidianas de referência, onde o eixo Z coincide com o eixo principal, é possível visualizar, por semelhança de triângulos, que um ponto  $(x, y, z)$  do mundo é mapeado no ponto  $(fx/z, fy/z, f)$  do plano da imagem (veja a Figura 4.2(b)). Na Figura 4.2(a), o ponto  $M$  do espaço é mapeado no ponto  $m$  pertencente ao plano da imagem. A Figura 4.2(b) ilustra a mesma situação, mas no plano YZ, onde é possível verificar a semelhança de triângulos empregada para calcular a posição do ponto projetado  $m$ .

Ignorando a última coordenada de  $(fx/z, fy/z, f)$ , obtemos um ponto relativo ao plano da imagem. Dessa forma, o mapeamento pode ser escrito como:

$$(4.1) \quad (x, y, z) \mapsto (fx/z, fy/z).$$

A Equação (4.1) especifica o mapeamento linear central das coordenadas do mundo para as coordenadas no plano da imagem (Hartley e Zisserman, 2004).

### **Matriz de Projeção e Matriz de Calibração da Câmera**

Se os pontos são representados por coordenadas homogêneas, então a projeção perspectiva representada pela Equação (4.1) pode ser definida como uma multiplicação de matrizes:

$$(4.2) \quad \begin{pmatrix} fx \\ fy \\ z \end{pmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}.$$

A matriz 3x4 da expressão anterior descreve o mapeamento e é denominada *matriz de projeção da câmera* ( $P$ ). Esta equação também pode ser escrita de forma resumida como

$$(4.3) \quad m = PM$$

onde  $M = (x, y, z, 1)^T$  são as coordenadas homogêneas do ponto 3D e  $m = (fx, fy, z)^T$  são as coordenadas homogêneas do ponto no plano da imagem.

O mapeamento definido na Equação (4.1) considera que a origem do sistema de coordenadas do plano da imagem coincide com o ponto principal. Na prática, entretanto, isto geralmente não acontece e as coordenadas do ponto principal precisam ser incorporadas no mapeamento. Assim, a Equação (4.1) é atualizada para:

$$(x, y, z) \mapsto (fx/z + p_x, fy/z + p_y) \quad (4.4)$$

onde  $(p_x, p_y)$  são as coordenadas do ponto principal. Esta equação também pode ser escrita como uma multiplicação de matrizes utilizando coordenadas homogêneas. A expressão obtida é:

$$\begin{pmatrix} fx + p_x z \\ fy + p_y z \\ z \end{pmatrix} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}, \quad (4.5)$$

que pode ser escrita de maneira concisa como:

$$m = [\beta | 0]M \quad (4.6)$$

onde

$$\beta = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix}. \quad (4.7)$$

A matriz  $\beta$  é denominada *matriz de calibração da câmera*.

### Câmeras CCD

O modelo de câmera de *pinhole* considera as coordenadas no plano da imagem como coordenadas Euclidianas tendo a mesma escala em ambos os eixos. No caso das câmeras que utilizam sensores CCD (*Charge-coupled device*), as coordenadas na imagem são comumente medidas em *pixels* e é possível que o sensor de imagem empregue pixels não quadrados (Hartley e Zisserman, 2004). Logo, o mapeamento da Equação (4.4) pode ser ajustado para expressar a transformação das coordenadas do mundo para as coordenadas no plano da imagem, em pixels:

$$(x, y, z) \mapsto (b_x fx/z + u_0, b_y fy/z + v_0) \quad (4.8)$$

onde  $b_x$  e  $b_y$  correspondem ao número de pixels por unidade de distância nas direções  $x$  e  $y$ , respectivamente. Os termos  $u_0$  e  $v_0$  são as coordenadas do ponto principal em pixels, onde  $u_0 = b_x p_x$  e  $v_0 = b_y p_y$ . A representação do mapeamento

anterior utilizando coordenadas homogêneas e multiplicação de matrizes é dada na Equação (4.9).

$$\begin{pmatrix} b_x f x + u_0 z \\ b_y f y + v_0 z \\ z \end{pmatrix} = \begin{bmatrix} b_x f & 0 & u_0 & 0 \\ 0 & b_y f & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}. \quad (4.9)$$

Para resgatar as coordenadas em pixels  $(u, v)$  do ponto projetado na imagem a partir das coordenadas homogêneas, devemos encontrar o fator  $s$  que quando multiplicado às coordenadas resulta em um ponto da forma  $(u, v, 1)$ . Assim, o mapeamento pode ser representado como na Equação (4.10).

$$s \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} b_x f x + u_0 z \\ b_y f y + v_0 z \\ z \end{pmatrix} = \begin{bmatrix} b_x f & 0 & u_0 & 0 \\ 0 & b_y f & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (4.10)$$

### Sistema de Coordenadas de Referência

As formulações anteriores consideram um único sistema de coordenadas de referência, localizado no centro óptico da câmera e com o eixo  $z$  na direção do eixo óptico. Esse sistema é denominado de *sistema de coordenadas da câmera*. Na prática, todavia, os pontos no espaço são comumente representados em termos de outro sistema de coordenadas, conhecido como *sistema de coordenadas do mundo*. Assim, para mapear um ponto  $M$  do espaço em um ponto  $m$  no plano da imagem, é necessário, primeiramente, converter as coordenadas do ponto  $M$  (dado com relação ao sistema de coordenadas do mundo) para o sistema de coordenadas da câmera.

Esta conversão é ilustrada na Figura 4.3.

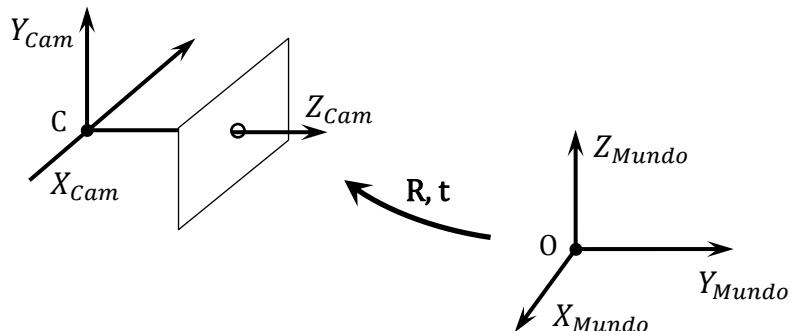


Figura 4.3 – Ilustração da mudança do sistema de coordenadas do mundo para o sistema de coordenadas da câmera. As coordenadas de um ponto são convertidas mediante aplicação de uma operação de rotação e uma de translação.

Como os dois sistemas estão relacionados por uma operação de rotação e uma de translação, a conversão pode ser feita por meio de operações utilizando a matriz adequada de rotação ( $R$ ) e a matriz adequada de translação ( $t$ ):

$$\begin{pmatrix} x_{cam} \\ y_{cam} \\ z_{cam} \end{pmatrix} = R \begin{pmatrix} x_{mundo} \\ y_{mundo} \\ z_{mundo} \end{pmatrix} + t \quad (4.11)$$

onde

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, \quad t = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}.$$

Além de possibilitar a mudança de coordenadas, as matrizes  $R$  e  $t$  podem ser usadas para descrever a posição e orientação da câmera com relação ao sistema de coordenadas do mundo. A Equação (4.11) pode ser reescrita como:

$$\begin{pmatrix} x_{cam} \\ y_{cam} \\ z_{cam} \end{pmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{pmatrix} x_{mundo} \\ y_{mundo} \\ z_{mundo} \\ 1 \end{pmatrix}. \quad (4.12)$$

Utilizando a Equação (4.10) e a Equação (4.12), pode-se definir uma equação geral para descrever o mapeamento de um ponto no espaço, dado no sistema de coordenadas do mundo, para o plano da imagem de uma câmera CCD, em pixels:

$$s \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{bmatrix} b_x f & 0 & u_0 \\ 0 & b_y f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{pmatrix} x_{mundo} \\ y_{mundo} \\ z_{mundo} \\ 1 \end{pmatrix}. \quad (4.13)$$

A equação anterior é comumente escrita de maneira reduzida como:

$$s m = K[R|t]M. \quad (4.14)$$

Repare que a matriz  $K$  apresenta 5 graus de liberdade, correspondentes aos parâmetros  $f$ ,  $b_x$ ,  $b_y$ ,  $u_0$  e  $v_0$ . Como esses parâmetros descrevem características internas da câmera, eles são comumente denominados *parâmetros intrínsecos da câmera*, ou simplesmente *parâmetros internos*. As matrizes  $R$  e  $t$  possuem 3 graus de liberdade cada, que são correspondentes aos três ângulos de rotação em torno dos eixos e aos três valores de translação. Como  $R$  e  $t$  estão relacionadas com o

posicionamento e orientação da câmera, os seis parâmetros que as definem são chamados de *parâmetros extrínsecos da câmera*, ou simplesmente *parâmetros externos*.

Considerando o mapeamento geral expresso pela Equação (4.14), a matriz  $P$  resultante da multiplicação de  $K$  por  $[R|t]$  terá um total de 11 graus de liberdade e é denominada de *matriz de projeção da câmera* (veja a Equação (4.15)).

$$P = K[R|t]. \quad (4.15)$$

A matriz de projeção  $P$  representa o mapeamento completo e pode ser utilizada para projetar um ponto  $M$  do espaço diretamente no plano de imagem da câmera, onde o ponto resultante  $m$  terá as coordenadas em pixels.

#### 4.1.2 Modelo de Distorção de Brown

Imagens geradas por câmeras modernas normalmente apresentam pequenas distorções, as quais não podem ser representadas utilizando a transformação linear do modelo de câmera de *pinhole*. Segundo Weng (Weng, Cohen *et al.*, 1992), essas distorções são causadas por imperfeições no processo de fabricação e montagem das lentes que compõem o sistema óptico da câmera. Dos vários tipos de distorções, existem dois principais: a distorção *radial* e a distorção *tangencial*.

A maior parte da distorção percebida em câmeras comerciais corresponde à distorção radial, que faz com que um ponto na imagem se desloque radialmente para dentro ou para fora (Figura 4.4, imagens (b) e (c)). Já a distorção tangencial é menos predominante e é responsável pelo afastamento dos pontos de seus eixos transversais normais (Figura 4.4(d)).

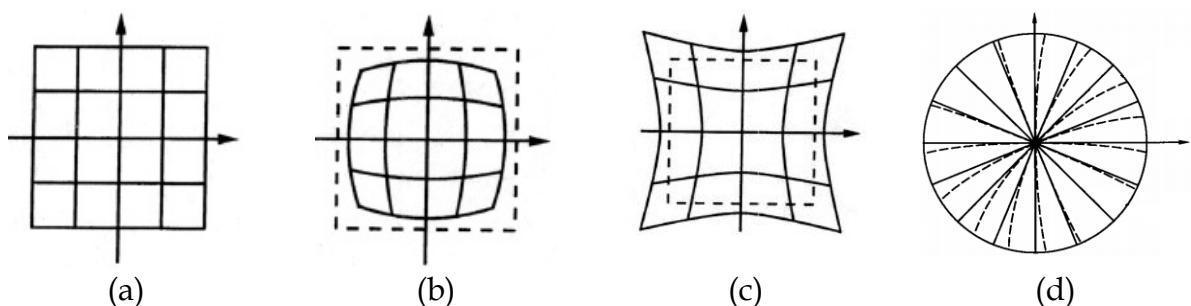


Figura 4.4 – Efeitos dos vários tipos de distorção nas imagens: (a) representação de imagem ideal, sem distorção; (b) distorção radial para o centro (*barrel*); (c) distorção radial para fora (*pincushion*) e (d) distorção tangencial. Neste último caso, as linhas pontilhadas representam pontos com distorção tangencial. Ilustrações adaptadas de (Weng, Cohen *et al.*, 1992).

O modelo de distorção proposto por Brown e Conrady (Conrady, 1919; Brown, 1966) é frequentemente usado para compensar essas distorções. Neste modelo, o par  $(x', y')$  representa as coordenadas normalizadas do ponto no sistema de coordenadas da câmera e  $(m', n')$  corresponde à sua representação após inserção de distorção radial, conforme equação a seguir:

$$\begin{aligned} m' &= (1 + \Delta r)x' \\ n' &= (1 + \Delta r)y' \end{aligned} \quad (4.16)$$

onde

$$\begin{aligned} \Delta r &= K_1 r^2 + K_2 r^4 + K_3 r^6, \\ r &= \sqrt{x'^2 + y'^2}, \\ x' &= x_{cam}/z_{cam}, \\ y' &= y_{cam}/z_{cam}. \end{aligned} \quad (4.17)$$

Os termos  $K_1, K_2$  e  $K_3$  em (4.17) são os coeficientes de distorção radial da câmera, que podem ser obtidos durante a etapa de calibração, como será visto mais adiante.

Já a distorção tangencial é compensada de acordo com as seguintes equações:

$$\begin{aligned} m' &= x' + \Delta t_x \\ n' &= y' + \Delta t_y \end{aligned} \quad (4.18)$$

onde

$$\begin{aligned} \Delta t_x &= 2P_1 x' y' + P_2 (r^2 + 2x'^2), \\ \Delta t_y &= 2P_2 x' y' + P_1 (r^2 + 2y'^2). \end{aligned} \quad (4.19)$$

Os termos  $P_1$  e  $P_2$  em (4.19) são os coeficientes de distorção tangencial da câmera e os termos  $r, x'$  e  $y'$  são obtidos como em (4.17).

Dado um ponto  $(x, y, z)$  já convertido para o sistema de coordenadas da câmera (conforme expressão (4.11)), o mapeamento de suas coordenadas do espaço tridimensional para o plano da imagem, considerando a inserção de distorção radial e tangencial do modelo de Brown, pode ser representado pela seguinte equação:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{bmatrix} b_x f & 0 & u_0 \\ 0 & b_y f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x'' \\ y'' \\ 1 \end{pmatrix}, \quad (4.20)$$

onde

$$(4.21) \quad \begin{aligned} x'' &= x'(1 + \Delta r) + \Delta t_x \\ y'' &= y'(1 + \Delta r) + \Delta t_y. \end{aligned}$$

Em (4.21), os termos  $x'$ ,  $y'$ ,  $\Delta r$ ,  $\Delta t_x$  e  $\Delta t_y$  são obtidos conforme apresentado nas equações (4.17) e (4.19).

## 4.2 Calibração de Câmeras

---

O processo de calibração das câmeras em um sistema de captura óptico tem como objetivo descobrir todos os parâmetros discutidos anteriormente para todas as câmeras que compõem o sistema. A descoberta dos cinco parâmetros internos ( $f$ ,  $b_x$ ,  $b_y$ ,  $u_0$  e  $v_0$ ) permite a construção da matriz  $K$  e a descoberta dos seis parâmetros externos (posição espacial e orientação da câmera com relação a um sistema externo de referência) possibilita a geração das matrizes de rotação  $R$  e translação  $t$ . Conhecendo-se  $K$ ,  $R$  e  $t$ , pode-se calcular a matriz de projeção  $P$  (Equação (4.15)) da câmera, que define o mapeamento linear das coordenadas de um ponto  $(x, y, z)$  do espaço para as coordenadas  $(u, v)$  da imagem, ignorando qualquer tipo de distorção. Quando o modelo de câmera de *pinhole* é estendido para incorporar distorções, os parâmetros de distorção, como  $K_1$ ,  $K_2$ ,  $P_1$  e  $P_2$  do modelo de Brown, também podem ser obtidos durante a calibração.

Nas últimas décadas, vários métodos foram propostos para conduzir o processo de calibração e encontrar esses parâmetros (Abdel-Aziz e Karara, 1971; Tsai, 1986; 1987; Zhang, 1998b; 1999; 2000; Pribanic, Sturm *et al.*, 2005; Svoboda, Martinec *et al.*, 2005; Pribanić, Sturm *et al.*, 2007). Alguns desses métodos são relativamente simples e consistem em estimar diretamente os elementos da matriz  $P$  com base no registro de pontos do espaço com posições previamente conhecidas (Abdel-Aziz e Karara, 1971; Pribanic, Sturm *et al.*, 2005). Outros, entretanto, são mais complexos e podem envolver técnicas de otimização não-linear em processos iterativos para descobrir e refinar os parâmetros (Tsai, 1987; Zhang, 2000; Svoboda, Martinec *et al.*, 2005). Como o procedimento de calibração em si não é o foco desta pesquisa e o conhecimento detalhado dessas técnicas não é necessário para a absorção do método proposto, esta

seção apresenta apenas as ideias principais de alguns dos métodos de calibração mais relevantes. O primeiro deles é o método DLT, o qual é descrito a seguir.

#### 4.2.1 Direct Linear Transformation - DLT

O método DLT original, proposto por Abdel e Karara em (Abdel-Aziz e Karara, 1971), busca a descoberta dos parâmetros internos e externos de uma câmera segundo o modelo puro de *pinhole* e não considera nenhum tipo de distorção. Para uma explicação mais detalhada desse método, considere a matriz  $P$  resultante da multiplicação de  $K$  por  $[R|t]$  na Equação (4.15). Essa matriz pode ser representada da seguinte forma:

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix}. \quad (4.22)$$

Como  $P$  possui apenas 11 graus de liberdade (cinco de  $K$  e seis de  $R|t$ ), pode-se normalizá-la dividindo todos os seus elementos por  $p_{34}$ . A matriz resultante terá a forma:

$$P = \begin{bmatrix} L_1 & L_2 & L_3 & L_4 \\ L_5 & L_6 & L_7 & L_8 \\ L_9 & L_{10} & L_{11} & 1 \end{bmatrix}. \quad (4.23)$$

O método DLT busca descobrir diretamente os 11 coeficientes  $(L_1, \dots, L_{11})$  da matriz de projeção  $P$  utilizando pontos do espaço cujas coordenadas são previamente conhecidas, como é descrito a seguir.

Utilizando a matriz  $P$ , a expressão (4.13) pode ser reescrita como:

$$\begin{pmatrix} su \\ sv \\ s \end{pmatrix} = \begin{bmatrix} L_1 & L_2 & L_3 & L_4 \\ L_5 & L_6 & L_7 & L_8 \\ L_9 & L_{10} & L_{11} & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}. \quad (4.24)$$

O ponto  $(su, sv, s)^T$  resultante do mapeamento (4.24) é representado em coordenadas homogêneas. Para obtenção do par  $(u, v)$  em coordenadas não-homogêneas, é necessário dividir as duas primeiras componentes do ponto pela terceira. Assim, pode-se escrever o mapeamento de (4.24) diretamente como:

$$u = \frac{L_1x + L_2y + L_3z + L_4}{L_9x + L_{10}y + L_{11}z + 1}, \quad v = \frac{L_5x + L_6y + L_7z + L_8}{L_9x + L_{10}y + L_{11}z + 1} \quad (4.25)$$

As equações em (4.25) podem ser rearranjadas da seguinte forma:

$$\begin{aligned} L_1x + L_2y + L_3z + L_4 - uL_9x - uL_{10}y - uL_{11}z &= u \\ L_5x + L_6y + L_7z + L_8 - vL_9x - vL_{10}y - vL_{11}z &= v \end{aligned} \quad (4.26)$$

que podem ser representadas como uma multiplicação de matrizes:

$$\begin{bmatrix} x & y & z & 1 & 0 & 0 & 0 & 0 & -ux & -uy & -uz \\ 0 & 0 & 0 & 0 & x & y & z & 1 & -vx & -vy & -vz \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ L_3 \\ L_4 \\ L_5 \\ L_6 \\ L_7 \\ L_8 \\ L_9 \\ L_{10} \\ L_{11} \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix} \quad (4.27)$$

Observe que, caso os valores de  $x, y, z, u$  e  $v$  sejam conhecidos, a Equação (4.27) pode ser tratada como um sistema linear tendo como variáveis independentes os termos  $L_1, \dots, L_{11}$ . Entretanto, apenas duas equações não são suficientes para a resolução do sistema, que possui onze incógnitas. Contudo, caso sejam conhecidas as coordenadas  $(x, y, z)$  de pelo menos seis pontos no espaço e suas respectivas coordenadas  $(u, v)$  na imagem da câmera, pode-se montar um sistema de doze equações e onze variáveis cuja resolução se torna possível. Esse sistema teria a seguinte forma:

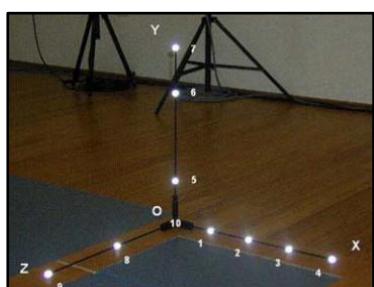
$$\begin{bmatrix} x_1 & y_1 & z_1 & 1 & 0 & 0 & 0 & 0 & -u_1x_1 & -u_1y_1 & -u_1z_1 \\ 0 & 0 & 0 & 0 & x_1 & y_1 & z_1 & 1 & -v_1x_1 & -v_1y_1 & -v_1z_1 \\ x_2 & y_2 & z_2 & 1 & 0 & 0 & 0 & 0 & -u_2x_2 & -u_2y_2 & -u_2z_2 \\ 0 & 0 & 0 & 0 & x_2 & y_2 & z_2 & 1 & -v_2x_2 & -v_2y_2 & -v_2z_2 \\ \vdots & \vdots \\ x_6 & y_6 & z_6 & 1 & 0 & 0 & 0 & 0 & -u_6x_6 & -u_6y_6 & -u_6z_6 \\ 0 & 0 & 0 & 0 & x_6 & y_6 & z_6 & 1 & -v_6x_6 & -v_6y_6 & -v_6z_6 \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ L_3 \\ L_4 \\ L_5 \\ L_6 \\ L_7 \\ L_8 \\ L_9 \\ L_{10} \\ L_{11} \end{bmatrix} = \begin{bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ \vdots \\ u_6 \\ v_6 \end{bmatrix} \quad (4.28)$$

Como o número de equações agora é maior que o número de incógnitas (para um número de pontos  $n \geq 6$ ), o sistema é *sobredeterminado* e não admite uma solução

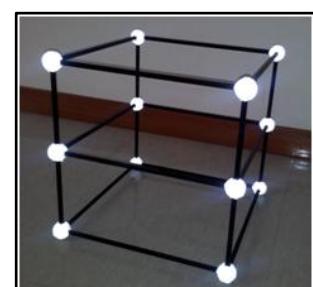
exata. Entretanto, é possível encontrar uma solução que minimize o seu erro residual, utilizando, por exemplo, a técnica SVD (*Singular Value Decomposition*) (Golub e Reinsch, 1970). Uma vez resolvido, os onze coeficientes  $L_1, \dots, L_{11}$  podem ser utilizados para compor a matriz de projeção  $P$  da câmera. Caso seja necessário, os parâmetros internos e externos individuais podem ser obtidos por meio da decomposição da matriz  $P$ .

Os seis pontos espaciais necessários para a resolução do sistema representado na Equação (4.28) são chamados *pontos de controle*. Esses pontos podem ser arranjados de diversas maneiras, mas existem algumas configurações que são degenerativas e não possibilitam que o sistema tenha uma solução válida, como aponta Yihong (Wu e Hu, 2006). Um exemplo de disposição degenerativa é quando todos os pontos são coplanares. Na prática, os pontos de controle são colhidos com o auxílio de marcadores reflexivos, que são colocados em posições previamente conhecidas de um objeto especial, denominado *objeto de calibração estática*. A Figura 4.5 ilustra dois tipos diferentes desse objeto.

O objeto de calibração geralmente define a origem do sistema de coordenadas do mundo e é colocado na região onde o movimento de captura acontecerá. Pelo menos uma imagem do objeto precisa ser registrada de cada câmera e a identificação dos marcadores na imagem pode ser feita manual ou automaticamente.



(a)



(b)

Figura 4.5 – (a) Objeto de calibração formado por três eixos ortogonais. Os nove pontos de controle correspondem aos nove marcadores reflexivos, cujas posições espaciais são conhecidas (imagem retirada de Pribanić, Surm *et al.* (2007)). (b) Objeto de calibração em formato de grade sustentando doze marcadores.

#### 4.2.2 Método de Tsai

A técnica de calibração proposta por Roger Tsai (Tsai, 1986; 1987) considera a distorção radial do modelo de distorção de Brown. A técnica utiliza um objeto de calibração plano, semelhante a um tabuleiro de xadrez, do qual os pontos de controle são extraídos. O objeto deve ser movimentado dentro do campo de visão de uma ou mais câmeras para que alguns de seus parâmetros sejam estimados. O método é dividido em duas etapas. Na primeira, o algoritmo considera parâmetros fornecidos pelo fabricante da câmera e busca uma solução aproximada para o problema de calibração utilizando um conjunto de equações lineares, de forma semelhante àquela empregada no DLT. Na segunda etapa, um processo de refinamento utilizando técnicas de otimização não-linear é empregado para encontrar os parâmetros finais. Em seus resultados, Tsai demonstrou que a incorporação da distorção radial e a etapa de refinamento possibilitaram a geração de resultados com valores de precisão muito superiores àqueles gerados por métodos puramente lineares, como o DLT.

#### 4.2.3 Método de Svoboda

Tomas Svoboda *et al.* (Svoboda, Martinec *et al.*, 2005) propuseram em seu artigo um método conveniente para efetuar a calibração de múltiplas câmeras em ambientes virtuais imersivos. Os autores propõem o uso de apenas uma ponteira laser (ou um LED), que deve ser movida em movimentos ondulares e aleatórios dentro do volume de reconstrução à medida que todas as câmeras (ou pelo menos um subconjunto relevante) registra sua posição. Os pontos coletados quadro a quadro são combinados para representar um objeto de calibração 3D virtual que abrange as principais regiões do volume. Os pontos registrados nas imagens são detectados com alta precisão e validados utilizando as restrições da geometria epipolar<sup>2</sup>. Os parâmetros internos e externos das câmeras são obtidos automaticamente tendo a primeira câmera do conjunto como referência. Opcionalmente, se há um objeto de calibração com pontos de referência cujas posições espaciais são conhecidas, as estruturas Euclidianas obtidas pelo método podem ser alinhadas com o sistema de

---

<sup>2</sup> As propriedades da geometria epipolar são apresentadas na Seção 4.3.

coordenadas do mundo indicado pelo objeto de calibração. O método proposto por Svoboda e seus colegas é conhecido como um método de *auto-calibração*.

#### 4.2.4 Métodos Híbridos

Muitos dos sistemas de captura de movimento comercializados atualmente empregam técnicas de calibração que envolvem uma mistura de diversos conceitos e métodos, incluindo aqueles descritos anteriormente. Uma técnica comumente utilizada é aquela envolvendo duas etapas distintas: a etapa da *calibração estática* e a etapa da *calibração dinâmica*. A primeira etapa consiste em utilizar um algoritmo não-iterativo, como o DLT, para calcular diretamente os parâmetros internos e externos das câmeras utilizando um modelo livre de distorção, como o de *pinhole*. A segunda etapa consiste de um processo de otimização não linear baseado em um modelo que incorpora distorções e que utiliza os parâmetros obtidos na primeira etapa como valores iniciais para a execução de um refinamento. A segunda etapa geralmente envolve a movimentação de um *bastão de calibração* (*calibration wand*, veja Figura 4.6), contendo um ou mais marcadores, dentro do volume de reconstrução. Esse processo pode durar de aproximadamente 30 segundos a 5 minutos, dependendo do tamanho do volume e da precisão desejada. Os pontos registrados pelas câmeras, referentes ao(s) marcador(es) do bastão, são empregados para otimização dos parâmetros visando minimizar o erro de reprojeção dos pontos tridimensionais reconstruídos.

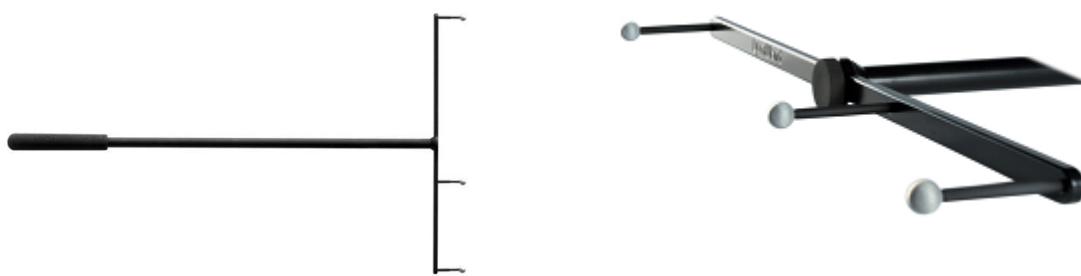


Figura 4.6 - Exemplo de um bastão de calibração (*calibration wand*) utilizado pelo fabricante NaturalPoint para conduzir a etapa da calibração dinâmica de seu sistema. Imagem adaptada de (Naturalpoint, 2013).

Devido à sua praticidade e aos bons níveis de precisão oferecidos (Pribanić, Peharec *et al.*, 2009), esta técnica foi implementada em software e utilizada para calibração do conjunto de câmeras utilizado para teste e avaliação do método proposto.

### 4.3 Geometria Epipolar

A geometria epipolar descreve a relação geométrica entre duas visões da mesma cena 3D, onde cada visão corresponde a uma imagem obtida de um ângulo diferente da cena. A geometria epipolar permite a extração de importantes relações acerca de pontos no espaço e suas projeções nas imagens, o que constitui o princípio de funcionamento das principais técnicas de reconstrução tridimensional (Ivekovic, Fusiello *et al.*, 2006). A geometria é ilustrada na Figura 4.7 e descrita a seguir.

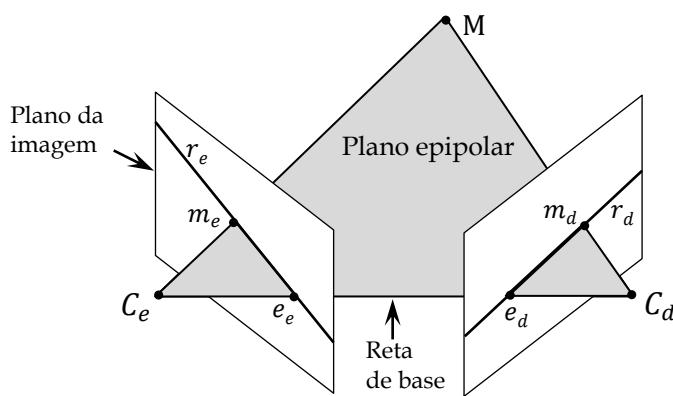


Figura 4.7 - Geometria epipolar. As duas câmeras são representadas pelos centros ópticos  $C_e$  e  $C_d$  e pelos planos de imagem. Os pontos  $C_e$  e  $C_d$ , o ponto 3D no espaço  $M$  e suas projeções nas imagens  $m_e$  e  $m_d$  compartilham um plano comum, chamado plano epipolar. A intersecção do plano epipolar com os planos de imagem formam as retas epipolares. A linha de base conecta os centros das câmeras e passam pelos planos de imagem nos epipolos ( $e_e$  e  $e_d$ ).

A geometria epipolar considera duas câmeras, que são comumente denominadas de *câmera esquerda* e *câmera direita* e representadas pelos seus centros ópticos  $C_e$  e  $C_d$ . Qualquer ponto  $M$  no espaço juntamente com os centros ópticos das câmeras define um plano denominado *plano epipolar*. Os pontos  $m_e$  e  $m_d$  correspondem às projeções de  $M$  nos planos de imagem das câmeras esquerda e direita, respectivamente. Como cada um desses pontos pertence à reta conectando  $M$  ao centro de projeção da respectiva câmera, os dois pontos pertencem também ao plano epipolar. A reta definida pela intersecção do plano epipolar com o plano da imagem é chamada *reta epipolar*. Na Figura 4.7, as duas retas foram denominadas de *reta epipolar esquerda* ( $r_e$ ) e *reta epipolar direita* ( $r_d$ ). A reta conectando os centros de projeção das câmeras ( $C_e$ ,  $C_d$ ) é denominada *reta de base*. Essa reta intercepta cada plano de imagem em um ponto denominado *epipolo*. O epipolo esquerdo ( $e_e$ ) corresponde à projeção do centro

óptico da câmera direita na imagem esquerda e o epipolo direito ( $e_d$ ), o inverso. Todas as retas epipolares na imagem esquerda (que podem ser obtidas de outros pontos no espaço) passam por  $e_e$  e todas as retas epipolares na imagem direita passam por  $e_d$ .

Uma importante propriedade resultante da geometria epipolar é o fato de que os pontos correspondentes nas imagens (isto é, pontos resultantes da projeção de um mesmo ponto 3D) pertencem a retas epipolares correlacionadas (conjugadas). Consequentemente, dado um ponto em uma imagem, é possível buscar pelo ponto correspondente na outra imagem, ao longo da reta epipolar conjugada (ao invés de buscar em toda a imagem). Isto pode ser feito porque um plano epipolar pode ser completamente definido pelos centros das câmeras e por um ponto em uma das imagens. Assim, dado o ponto  $m_e$ , por exemplo, é possível determinar a reta epipolar na imagem direita na qual o ponto  $m_d$  deve estar. Essa reta pode ser facilmente definida projetando-se os pontos  $m_e$  e  $e_e$  na imagem direita.

Adicionalmente, se os pontos  $m_e$  e  $m_d$  são conhecidos, pode-se calcular duas retas que partem dos respectivos centros de projeção das câmeras e seguem em direção ao espaço através dos pontos  $m_e$  e  $m_d$ . Em uma câmera perfeita de *pinhole*, o ponto de interseção das duas retas no espaço corresponde ao ponto tridimensional  $M$ . Este é o princípio básico de um dos métodos de reconstrução 3D, que atua a partir de duas visões de uma mesma cena. Este método é apresentado na próxima seção.

#### 4.4 Reconstrução Tridimensional

---

A reconstrução tridimensional consiste no cálculo das coordenadas espaciais ( $x, y, z$ ) de um ponto a partir de registros desse ponto em duas ou mais imagens. A reconstrução pode ser feita segundo duas formas principais: (1) utilizando as equações e os coeficientes empregados na técnica DLT (Shapiro, 1978); ou (2) utilizando diretamente as restrições resultantes da geometria epipolar para triangular o ponto espacial por meio do cálculo do traçado inverso dos raios de luz (Hartley e Sturm, 1997). As duas formas são discutidas brevemente a seguir.

#### 4.4.1 Reconstrução com base na técnica DLT

A reconstrução tridimensional com base no método DLT utiliza os 11 coeficientes  $L_1, \dots, L_{11}$  para a construção de um sistema linear que permite o cálculo das coordenadas espaciais (Abdel-Aziz e Karara, 1971; Shapiro, 1978).

Considere novamente as expressões da Equação (4.25). Vale lembrar que  $(x, y, z)$  representa um ponto no espaço e o par  $(u, v)$  representa a projeção de  $(x, y, z)$  na imagem. Tais expressões podem ser rearranjadas da seguinte forma:

$$(4.29) \quad \begin{aligned} x(L_1 - uL_9) + y(L_2 - uL_{10}) + z(L_3 - uL_{11}) &= u - L_4, \\ x(L_5 - vL_9) + y(L_6 - vL_{10}) + z(L_7 - vL_{11}) &= v - L_8. \end{aligned}$$

As igualdades em (4.29) podem ser representadas como uma multiplicação de matrizes:

$$(4.30) \quad \begin{bmatrix} (L_1 - uL_9) & (L_2 - uL_{10}) & (L_3 - uL_{11}) \\ (L_5 - vL_9) & (L_6 - vL_{10}) & (L_7 - vL_{11}) \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} u - L_4 \\ v - L_8 \end{bmatrix}.$$

Caso os onze coeficientes da câmera  $L_1, \dots, L_{11}$  e o ponto na imagem  $(u, v)$  sejam conhecidos, a Equação (4.30) pode ser tratada como um sistema linear tendo como variáveis as coordenadas  $x, y$  e  $z$ . Embora o sistema não tenha solução considerando apenas duas equações, é possível solucioná-lo caso o ponto  $(x, y$  e  $z)$  tenha sido detectado também em outra câmera, o que possibilita a obtenção de mais duas equações. A igualdade em (4.31) ilustra o resultado desse procedimento. Nesta equação,  $L_1^1, \dots, L_{11}^1$  são os coeficientes DLT da câmera um e  $(u^1, v^1)$  é o ponto detectado na imagem da câmera um, correspondente à projeção de  $(x, y, z)$ . Similarmente,  $L_1^2, \dots, L_{11}^2$  são os coeficientes da câmera dois e  $(u^2, v^2)$  corresponde ao ponto detectado na câmera dois (que é uma projeção de  $(x, y, z)$ ).

$$(4.31) \quad \begin{bmatrix} (L_1^1 - u^1 L_9^1) & (L_2^1 - u^1 L_{10}^1) & (L_3^1 - u^1 L_{11}^1) \\ (L_5^1 - v^1 L_9^1) & (L_6^1 - v^1 L_{10}^1) & (L_7^1 - v^1 L_{11}^1) \\ (L_1^2 - u^2 L_9^2) & (L_2^2 - u^2 L_{10}^2) & (L_3^2 - u^2 L_{11}^2) \\ (L_5^2 - v^2 L_9^2) & (L_6^2 - v^2 L_{10}^2) & (L_7^2 - v^2 L_{11}^2) \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} u^1 - L_4^1 \\ v^1 - L_8^1 \\ u^2 - L_4^2 \\ v^2 - L_8^2 \end{bmatrix}$$

Uma solução que minimiza o erro residual para o sistema em (4.31) pode ser encontrada utilizando a técnica SVD. Assim, as coordenadas  $x$ ,  $y$  e  $z$  do ponto no espaço são estimadas a partir dos respectivos pontos registrados nas imagens de duas câmeras calibradas. Caso o ponto espacial seja detectado em mais de duas câmeras, um sistema análogo, contendo um número maior de equações, pode ser construído e resolvido.

#### 4.4.2 Reconstrução com base na Geometria Epipolar

Proposto por Hartley e Sturm (Hartley e Sturm, 1997), a ideia básica deste método de reconstrução consiste em realizar o processo inverso ao da projeção. Conforme mencionado na Seção 4.3, dado um ponto  $m_1$  na imagem registrada pela câmera  $C_1$ , pode-se definir uma reta que parte do centro de projeção de  $C_1$  e segue em direção ao espaço através de  $m_1$ , representando o traçado inverso do raio de luz. Em uma câmera ideal de *pinhole*, o ponto espacial que gerou a projeção  $m_1$  deve se encontrar em algum local dessa reta. Como a informação sobre profundidade não está disponível utilizando apenas uma câmera, a posição espacial do ponto  $M$  só pode ser estimada caso  $M$  tenha sido detectado como um ponto  $m_2$  na imagem de uma segunda câmera  $C_2$ . Neste caso, uma segunda reta é definida partindo do centro de  $C_2$  em direção ao espaço através de  $m_2$ . Idealmente, as coordenadas do ponto de cruzamento dessas duas retas equivalem às coordenadas espaciais do ponto  $M$ . Este processo é ilustrado na Figura 4.8.

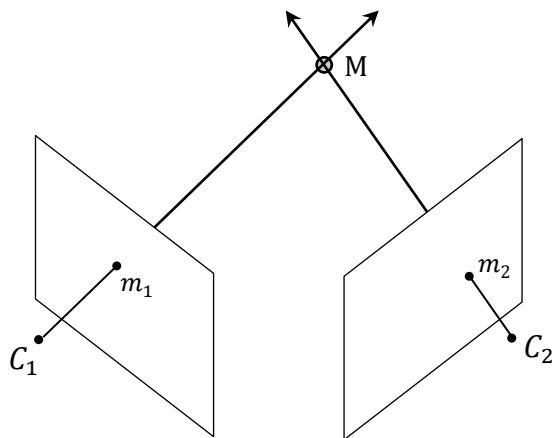


Figura 4.8 – Reconstrução tridimensional de um ponto  $M$  dados as detecções  $m_1$  e  $m_2$  nos planos de imagem de duas câmeras  $C_1$  e  $C_2$ , respectivamente. Situação ideal, desconsiderando limitações da calibração, distorção de lentes, etc.

Entretanto, devido às imperfeições das lentes das câmeras e a limitada precisão do processo de calibração, na prática esses raios dificilmente se cruzam efetivamente. Quando o cruzamento não ocorre, o ponto espacial  $M$  pode ser estimado como sendo o ponto médio  $M'$  da reta que conecta os dois raios formando a menor distância entre eles. Este é o princípio básico empregado pelo método de identificação, rastreamento e reconstrução de um conjunto de marcadores proposto neste trabalho.

# Capítulo 5

## Materiais e Métodos

Este capítulo descreve o método proposto para identificar, rastrear e reconstruir as trajetórias espaciais de múltiplos marcadores sem apoiar-se em qualquer tipo de modelo anatômico humano. Conforme já mencionado, o método é indicado para execução em tempo real, utilizando múltiplas câmeras de infravermelho juntamente com dezenas de marcadores reflexivos. As principais etapas que compõem a técnica são descritas ao longo do capítulo, que está organizado como segue.

A Seção 5.1 descreve diversos termos que são comumente utilizados durante a exposição do método. A Seção 5.2 fornece uma visão geral do método. O tópico compreende o seu funcionamento básico e adianta as ideias gerais a respeito de cada uma das três etapas que compõem o seu algoritmo principal. A Seção 5.3 descreve cada uma dessas etapas e a Seção 5.4 encerra o capítulo apresentando a metodologia básica utilizada para o desenvolvimento do método e sua implementação em software.

### 5.1 Termos e Definições

---

Alguns termos relevantes para a descrição do método são definidos a seguir.

**Marcador Real.** O termo *marcador real* ou *marcador reflexivo* é empregado para fazer referência ao marcador físico, isto é, à esfera coberta com material reflexivo utilizada como ponto de referência em uma sessão de captura.

**Marcador 3D.** O termo *marcador 3D* ou simplesmente *ponto 3D* é utilizado para fazer referência ao marcador “virtual” calculado pela técnica de reconstrução. Na maioria das vezes, um marcador 3D correspondente a um *marcador real* presente na cena, mas também é possível que essa correspondência não exista, caso em que o marcador 3D é chamado de *marcador fantasma* ou *falso marcador*. Um marcador 3D possui um código de identificação (um *Id*), coordenadas espaciais  $x$ ,  $y$  e  $z$  e, eventualmente, um rótulo. A notação  $M^i$  é comumente empregada para representar um marcador 3D com código de identificação  $i$ .

**Marcador 2D.** O termo *marcador 2D* ou simplesmente *ponto 2D* é empregado para fazer referência aos pontos nas imagens registrados pelas câmeras, representados pelas coordenadas  $x$  e  $y$  de seus centros, e que correspondem à projeções em potencial dos marcadores reflexivos (é possível que as câmeras registrem outros pontos luminosos no ambiente, que não se tratam efetivamente de marcadores). Vale lembrar ainda que o método proposto faz uso de câmeras de infravermelho, que já processam as imagens capturadas e devolvem apenas as coordenadas  $x$  e  $y$  correspondentes aos marcadores reflexivos detectados. Utiliza-se frequentemente a notação  $m_i$  para fazer referência a um marcador 2D ( $m_1, m_2$ , etc.). Adicionalmente, o termo *projeção* é algumas vezes utilizado como sinônimo de *marcador 2D*.

**Frame 2D.** Um *frame 2D* ou simplesmente *frame* é definido como um conjunto de  $n$  marcadores 2D obtidos a partir de uma câmera  $k$  no instante  $t$  e é denotado por  $F_k^t = \{m_1, m_2, \dots, m_n\}$ .

**Frameset.** Um *frameset* é definido como um conjunto de  $p$  *frames 2D* obtidos, respectivamente, a partir de  $p$  câmeras de infravermelho no instante  $t$ . Um *frameset* é denotado por  $FS^t = \{F_1^t, F_2^t, \dots, F_p^t\}$ , onde cada  $F_i^t$  corresponde a um *frame 2D* obtido a partir da câmera  $i$  no instante  $t$ , sendo  $p$  o número total de câmeras em uso no sistema. Repare que os *frames* de um *frameset* são sincronizados no tempo, pois são todos obtidos no mesmo instante  $t$ .

**Frame 3D.** Um *frame 3D* é definido como um conjunto de  $n$  pontos 3D calculados para um instante  $t$  a partir de um *frameset*  $FS^t$ . Um *frame 3D* é denotado por  $3F^t =$

$\{M^1, M^2, \dots, M^n\}$ , onde  $M^i$  representa um ponto 3D potencialmente correspondente a um marcador físico utilizado na captura.

## 5.2 Visão Geral da Técnica

---

Tendo definido o termo *frameset*, pode-se dizer que a técnica proposta tem como meta realizar o rastreamento, a identificação e a reconstrução de uma nuvem de marcadores, em tempo real, a partir de um fluxo contínuo de *framesets* sendo coletados de uma coleção de câmeras de infravermelho sincronizadas. Dado cada novo *frameset* recebido do conjunto de câmeras, busca-se usá-lo adequadamente para a produção de um conjunto de pontos tridimensionais (*frame 3D*) correspondente a um conjunto de marcadores reflexivos.

Para o primeiro *frameset* coletado das câmeras,  $FS^0$ , executa-se um procedimento baseado na geometria epipolar para agrupar os pontos 2D de diferentes *frames* que constituem projeções de um mesmo marcador reflexivo. Os pontos 2D “casados” são utilizados para o cálculo dos pontos tridimensionais correspondentes, que irão compor o primeiro *frame 3D*,  $3F^0$ . Neste momento, uma identificação automática (um ID) é fornecida para cada grupo de pontos 2D casados e para o respectivo ponto 3D.

A partir do segundo *frameset*, utiliza-se os pontos 2D já identificados nos *framesets* anteriores para a execução de um rastreamento bidimensional a fim de acelerar a identificação e a reconstrução para o instante corrente. A Figura 5.1 ilustra o funcionamento geral da técnica para um instante  $t \geq 3$ . Conforme esquematizado, o método compreende um algoritmo de rastreamento e reconstrução principal, o qual possui a função de gerar um novo *frame 3D* para cada *frameset* obtido da coleção de câmeras. Este algoritmo pode ser dividido em três etapas principais:

1. Etapa do Rastreamento 2D,
2. Etapa da Reconstrução 3D,
3. Etapa do casamento 2D para novas detecções 3D.

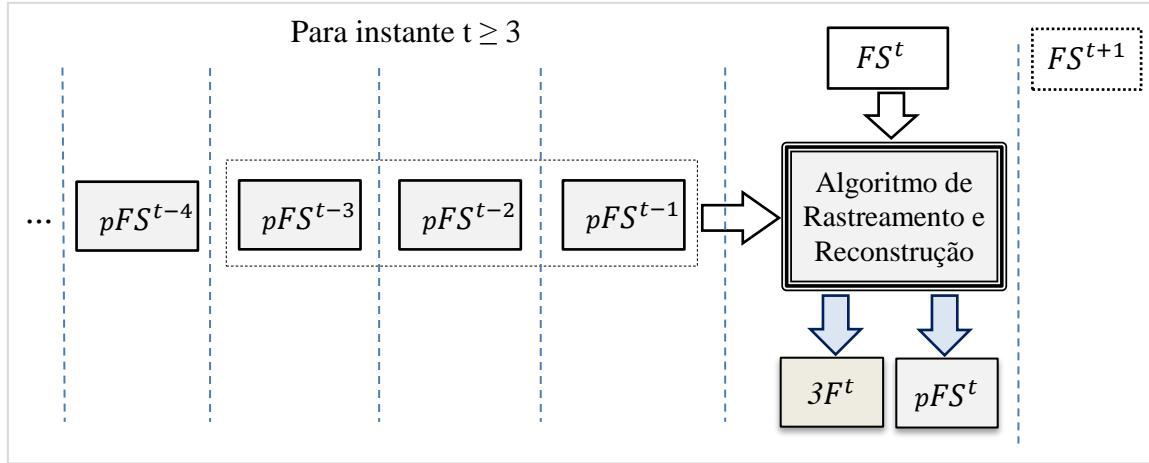


Figura 5.1 – Ilustração do funcionamento básico da técnica. Para um *frameset* coletado em um instante  $t \geq 3$ , utiliza-se sempre os três últimos *framesets* processados e o *frameset* corrente para obtenção do *frame 3D* corrente ( $3F^t$ ). A letra  $p$  é utilizada para indicar que o *frameset* já foi avaliado.

A etapa do *rastreamento 2D* é responsável pela identificação direta dos pontos 2D no *frameset* corrente com base nos pontos 2D já identificados no *frameset* anterior.

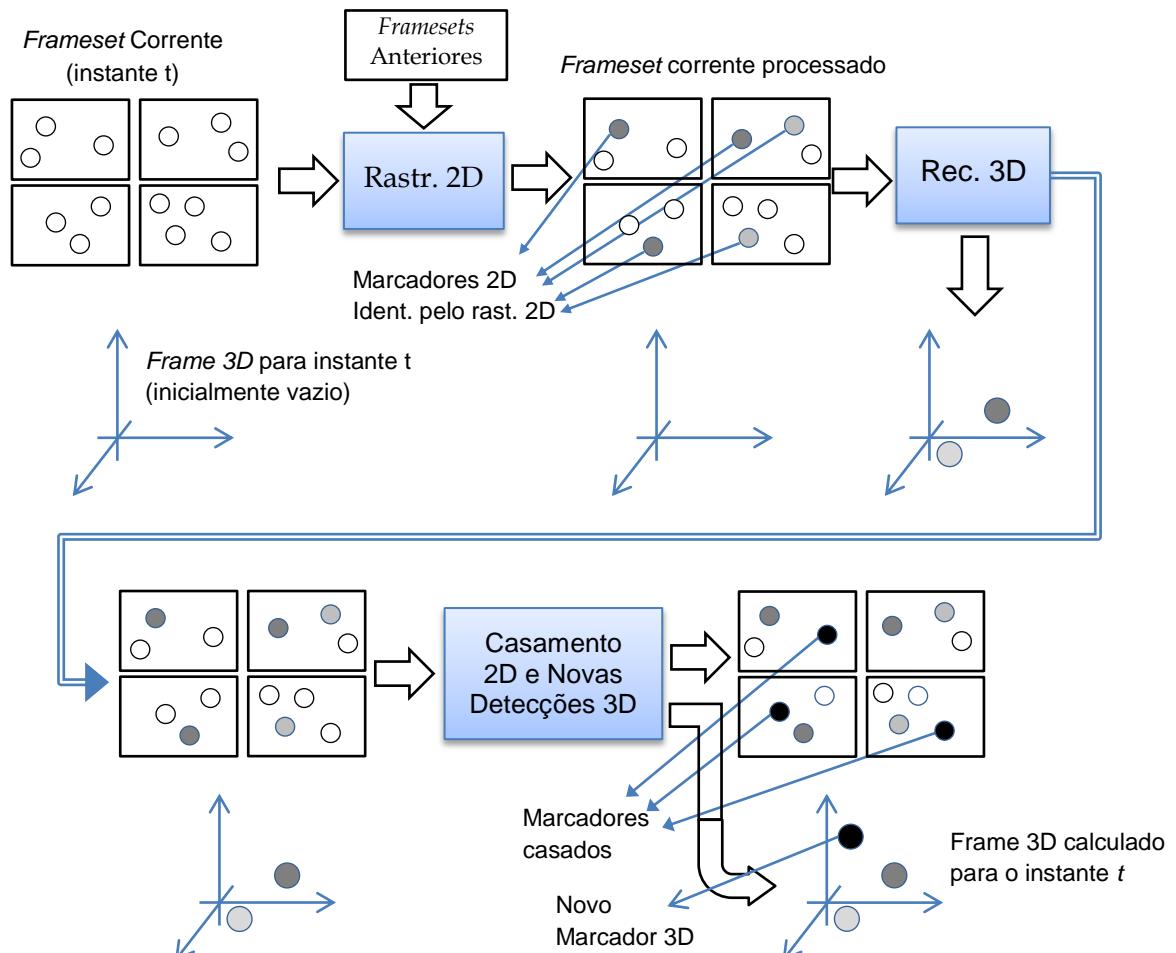


Figura 5.2 – Ilustração do papel desempenhado por cada etapa do algoritmo principal que compõe o método de reconstrução proposto. O diagrama pressupõe que os *framesets* anteriores ao instante  $t$  já tenham sido processados, possibilitando que a etapa de rastreamento 2D inicie o procedimento de identificação dos marcadores no *frameset* corrente.

A etapa da *reconstrução 3D* utiliza todos os marcadores 2D do *frameset* corrente, já identificados pelo rastreamento 2D, para realizar o cálculo das coordenadas x, y e z dos pontos 3D correspondentes.

A etapa do *casamento 2D e novas detecções 3D* exerce seu papel sempre que o *frameset* corrente ainda disponibiliza marcadores sem identificação, os quais podem eventualmente ser combinados para obtenção de novos pontos 3D.

A Figura 5.2 ilustra o papel essencial desempenhado por cada etapa do algoritmo para um instante  $t \geq 3$ .

### 5.3 Descrevendo as Etapas

---

#### *Etapa do Rastreamento 2D*

O objetivo da etapa de rastreamento 2D é rastrear pontos que estão se movendo no espaço bidimensional, que correspondem a projeções de marcadores reais em movimento no espaço 3D. Realizar o rastreamento de marcadores 2D que já foram identificados e rotulados nos *frames* anteriores possibilita a identificação direta desses marcadores nos *frames* correntes, o que contribui para acelerar o processo de reconstrução tridimensional.

Ao desenvolver o algoritmo para efetuar o rastreamento 2D, teve-se como preocupação primordial o seu custo computacional. Devido ao interesse em executá-lo em tempo real, para um número de câmeras e marcadores relativamente alto e com taxas de captura que podem ultrapassar 100 FPS, técnicas mais complexas e de alto custo computacional foram descartadas, como, por exemplo, aquelas que utilizam filtros de Kalman e filtros de partículas (Ringer e Lasenby, 2000; Figueira, Leite *et al.*, 2003). O algoritmo ICP (*Iterative Closest Point*) originalmente proposto em (Besl e Mckay, 1992) e empregado por Johnson em seus experimentos (Johnson, Xie *et al.*, 2010b) também foi descartado devido aos resultados pouco satisfatórios e à possibilidade de não convergência para uma solução. Alguns autores utilizaram métodos de extrapolação empregando técnicas de regressão polinomial (Armstrong, 1985; Shafiq, Tümer *et al.*, 2001). Entretanto, o processo de obtenção dos coeficientes

do modelo polinomial envolve o cálculo de operações não triviais, como a preparação e resolução de sistemas sobredeterminados e operações de decomposição de matrizes (Golub e Reinsch, 1970), o que poderia comprometer a velocidade de execução do método frente às circunstâncias mencionadas anteriormente.

A técnica de rastreamento 2D proposta é similar àquela apresentada por Herda *et al* (2001), onde cada marcador é rastreado de maneira independente com base no fato de que, na maioria dos casos, o deslocamento de um ponto 2D entre dois *frames* consecutivos é mínimo. O objetivo é estimar a posição do marcador no *frame* corrente, dadas as suas posições nos *frames* imediatamente anteriores.

O algoritmo foi proposto para efetuar o rastreamento utilizando, no máximo, os três últimos *frames* processados. Para rastrear um marcador 2D do *frame*  $F_k^{t-1}$  ao *frame*  $F_k^t$ , considera-se três casos, que ocorrem quando:

- a) O marcador 2D foi inicialmente identificado no *frame*  $F_k^{t-2}$  e então rastreado até o *frame*  $F_k^{t-1}$ ;
- b) O marcador 2D foi identificado pela primeira vez no *frame*  $F_k^{t-1}$ ;
- c) O marcador foi rastreado ao longo dos três últimos *frames*:  $F_k^{t-3}$ ,  $F_k^{t-2}$  e  $F_k^{t-1}$ .

No caso (a), o deslocamento do marcador do *frame*  $F_k^{t-2}$  ao *frame*  $F_k^{t-1}$  é utilizado para obtenção de sua direção de movimento e projeção de sua posição em  $F_k^t$ . Isto é feito diretamente utilizando as expressões:

$$(5.2)$$

$$p_x = b_x + (b_x - a_x)$$

$$p_y = b_y + (b_y - a_y)$$

onde  $a$  e  $b$  representam as posições do marcador sendo rastreado nos *frames*  $F_k^{t-2}$  e  $F_k^{t-1}$ , respectivamente; e  $p$  representa a posição estimada do marcador no *frame* corrente.

Uma vez estimada a posição  $p$ , define-se uma *região de procura* circular centrada em  $p$ , onde o raio, denominado  $RP_{2D}$ , é um parâmetro previamente especificado. O

próximo passo é realizar uma busca por marcadores 2D dentro dessa região. Se exatamente um marcador for encontrado dentro da região de procura, então esse marcador deve ser marcado como *identificado* e deve receber o código de identificação oriundo do marcador que o antecede. Para evitar falsa identificação, diferente de Herda *et al* (2001), o algoritmo proposto abandona o procedimento atual de rastreamento quando encontra zero, dois ou mais marcadores dentro da região de busca. Este processo é ilustrado na Figura 5.3.

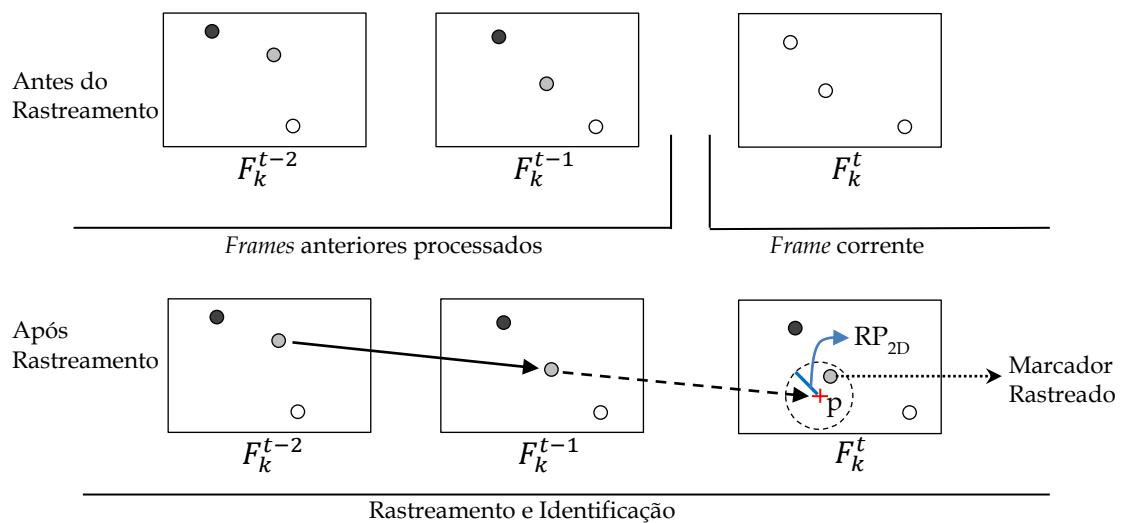


Figura 5.3 - Rastreamento 2D - Caso (a). O marcador sendo rastreado possui registro nos dois últimos frames (começando no instante  $t-2$ ). Sua posição no *frame* atual é estimada com base nas posições anteriores, que são utilizadas para a definição de uma região de busca. Caso exatamente um marcador seja encontrado dentro dessa região, então o mesmo é marcado como identificado utilizando o rótulo do marcador precedente.

Já no caso (b), não é possível definir uma direção, pois apenas a posição anterior do marcador é conhecida. Nesta situação, define-se a região de busca com centro nas coordenadas da última posição do marcador e prossegue-se como no caso (a).

Para o caso (c), a posição do marcador no frame corrente é estimada de maneira semelhante àquela feita em (a), exceto pelo fato de que pode haver um aprimoramento da posição estimada com base na trajetória definida de  $F_k^{t-3}$  a  $F_k^{t-1}$ . Este processo é especialmente importante quando do rastreamento de pontos 2D resultantes de marcadores reais que estão se movendo rápido, traçando uma trajetória curvilínea e a velocidade de captura das câmeras não é alta o suficiente para manter muito pequeno o deslocamento entre pontos 2D consecutivos. Neste

caso, o ângulo alpha formado pelos segmentos conectando os três últimos pontos é utilizado para fazer a correção da posição pré-estimada  $p'$  para a posição  $p$ . Este processo é ilustrado na Figura 5.4.

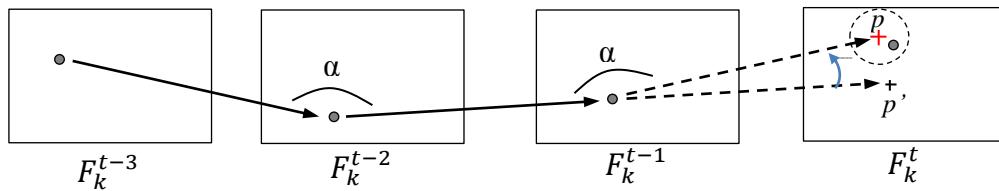


Figura 5.4 - Rastreamento 2D. Caso em que existe um caminho definido pelo marcador nos três frames. O ângulo referente à "curva" definida pelo marcador é empregado para estimar sua posição no *frame* corrente.

Se o ângulo  $\alpha$  é próximo de  $180^\circ$  (de acordo com um limiar pré-definido) ou o deslocamento médio do marcador de  $F_k^{t-3}$  a  $F_k^{t-1}$  é ínfimo (menor que um limiar), então o próprio ponto  $p'$  é escolhido como o centro da região de busca, exatamente como no caso (a). Caso contrário, o ponto  $p'$  é rotacionado  $\beta$  graus em torno da última posição do marcador, onde  $\beta = (180 - \alpha)$ . O sentido da rotação depende das posições do marcador nos três últimos frames, identificadas por **a**, **b** e **c**, respectivamente, que poderão estar em sentido horário ou anti-horário. Para descobrir o sentido dos pontos **a**, **b** e **c**, é executada a varredura de Graham (*Graham scan*) (Graham, 1972), que calcula a cobertura convexa (*convex hull*) de um conjunto finito de pontos no plano. Isto é feito aplicando-se a seguinte equação:

$$A = (x_b - x_a)(y_c - y_a) - (x_c - x_a)(y_b - y_a) \quad (5.1)$$

Caso A seja maior que zero, uma rotação no sentido horário é executada em torno de **b**. Caso A seja menor que zero, uma rotação no sentido anti-horário é considerada.

#### Região de Procura $RP_{2D}$

O valor  $RP_{2D}$  define o tamanho da vizinhança em torno do ponto  $p$  dentro da qual o marcador é procurado. Em geral, quanto maior o valor de  $RP_{2D}$ , maior é a probabilidade de encontrar um marcador dentro da região, mas também são maiores as chances de rastreamento incorreto ou de serem encontrados mais de um marcador. O melhor valor para  $RP_{2D}$  pode depender de vários fatores como o tipo de movimento sendo capturado, a taxa de captura das câmeras e o número de

marcadores em uso. De uma forma geral, valores de  $RP_{2D}$  entre dois e três pixels possibilitaram resultados satisfatórios para diversos experimentos realizados.

### *Etapa da Reconstrução 3D*

A etapa da reconstrução 3D recebe como entrada o *frameset* corrente resultante da etapa de rastreamento 2D. Vale lembrar que esse *frameset* é composto por um grupo de *frames* sincronizados, sendo que cada *frame* pode conter marcadores 2D já identificados (provenientes da etapa anterior), assim como marcadores 2D ainda sem identificação. A etapa da reconstrução 3D tem a responsabilidade de utilizar cada grupo de pontos 2D com a mesma identificação (mesmo Id), os quais correspondem a projeções de um mesmo marcador físico, para reconstruir as coordenadas x, y e z do ponto 3D equivalente. Este procedimento é realizado com base nas propriedades da geometria epipolar (Hartley e Zisserman, 2004).

É importante reforçar que os pontos 3D calculados para o instante  $t$  recebem códigos de identificação sendo carregados pelos respectivos pontos 2D. Como esses códigos são repassados de um *frame* a outro no rastreamento bidimensional, os pontos 3D de *frames* 3D consecutivos também compartilham a respectiva identificação, o que possibilita a extração direta de suas trajetórias no tempo. Dessa forma, ao término de uma sessão de captura que gerou 1000 *framesets*, por exemplo, a técnica terá produzido um total de 1000 *frames* 3D, a partir dos quais é possível obter as trajetórias dos marcadores.

### *Etapa do Casamento 2D e Novas Detecções 3D*

A etapa do rastreamento 2D tem a responsabilidade de rastrear marcadores que já foram previamente identificados. Ela se baseia em marcadores 2D já identificados do *frameset* anterior para identificar e rotular os marcadores 2D nos *frames* atuais, os quais são utilizados na etapa de reconstrução para o cálculo tridimensional. A etapa do *casamento 2D e novas detecções 3D* tem a função de realizar o casamento de marcadores 2D *que ainda não foram identificados*, a fim de encontrar novos pontos 3D para registro em  $3F^t$ . O objetivo inicial é descobrir quais marcadores 2D, de diferentes *frames*, correspondem a projeções de um mesmo marcador real.

A etapa é particularmente importante quando as imagens dos marcadores são registradas pela primeira vez durante uma sessão de captura. Neste momento, a técnica tem a oportunidade de encontrar e agrupar adequadamente as projeções de um mesmo marcador reflexivo e utilizá-las para registrar os pontos 3D no conjunto  $3F^t$  pela primeira vez. Outras situações em que a etapa exerce seu papel são:

- a) Quando um marcador reflexivo que sofreu oclusão temporariamente volta a ser visualizado por três ou mais câmeras do conjunto. Novamente, um novo ponto 3D com um novo ID deve ser registrado em  $3F^t$ ;
- b) Quando a etapa de rastreamento 2D não é capaz de rastrear pelo menos 2 marcadores 2D vinculados a um mesmo marcador reflexivo. Neste caso, uma nova associação desses marcadores pode ser realizada.

Como não há uma informação direta que associe os marcadores reais com suas projeções nas imagens, encontrar as devidas correspondências não é um processo trivial (Ringer e Lasenby, 2004). O problema ainda é dificultado pelo fato de que nem todos os marcadores aparecem no campo de visão de todas as câmeras, ou seja, cada marcador real pode ser detectado por um subconjunto de câmeras diferente.

O método proposto para realizar esse casamento se baseia nas propriedades da geometria epipolar. A ideia geral é calcular a reta epipolar conjugada no *frame*  $F_{k+1}^t$  para cada marcador 2D sem identificação no *frame*  $F_k^t$ . Os pontos próximos à reta calculada são candidatos a formar par com o primeiro marcador. Esses pontos são utilizados para o cálculo de um ponto espacial candidato, que se tornará efetivo caso exista um terceiro ponto 2D em uma terceira câmera mantendo a mesma relação.

O emprego de retas epipolares para reduzir o espaço de busca por marcadores correspondentes foi anteriormente utilizado por Shafiq e seus colegas (Shafiq, Tümer *et al.*, 2001). Entretanto, o método descrito por esses autores envolve a geração de diversas retas epipolares e diversas fases de ordenação de marcadores e câmeras para a escolha da melhor combinação possível, o que o torna computacionalmente inviável para execução em tempo real. Além disso, Shafiq e seus colegas empregaram

uma técnica de reconstrução menos eficiente, baseada no algoritmo DLT (Abdel-Aziz e Karara, 1971; Hartley e Zisserman, 2004).

O algoritmo que a presente pesquisa propõe para efetuar o casamento 2D e o registro de novos pontos 3D é apresentado a seguir.

1. Fazer  $i$  igual a 0;
- 2a. Incrementar  $i$  e escolher o próximo *frame*  $F_i^t$  do *frameset*  $FS^t$ ; se não houver mais *frames*, encerrar o algoritmo;
- 2b. Escolher o próximo marcador não identificado  $m_p$  do *frame*  $F_i^t$  e fazer  $j$  igual a  $i$ ; se não houver mais marcadores, voltar ao passo 2a;
- 3a. Incrementar  $j$  e escolher o próximo *frame*  $F_j^t$ . Utilizar  $m_p$  para gerar a reta epipolar  $E_p$  correspondente em  $F_j^t$ ; se não houver mais *frames*, voltar ao passo 2b;
- 3b. Escolher o próximo marcador não identificado  $m_q$  de  $F_j^t$  que esteja a uma distância inferior a  $W$  de  $E_p$ ; se não houver mais marcadores, voltar ao passo 3a;
- 3c. Gerar os raios correspondentes a  $m_p$  e  $m_q$  e calcular o ponto de cruzamento  $P_1$  (veja Seção 4.4.2). Se a menor distância entre os raios for superior a um limiar previamente definido  $D_{MAX}$ , descartar  $P_1$  e voltar no passo 3b. Caso contrário, faça  $k$  igual a  $j$  e vá para o passo 4a;
- 4a. Incrementar  $k$  e escolher o próximo *frame*  $F_k^t$ . Projetar o ponto  $P_1$  em  $F_k^t$  a fim de obter o ponto projetado  $PR_k$ ; se não houver mais *frames*, voltar ao passo 3a;
- 4b. Escolher o próximo marcador não identificado  $m_r$  de  $F_k^t$  que esteja a uma distância inferior a  $W$  de  $PR_k$ ; se não houver mais marcadores, voltar ao passo 4a;
- 4c. Gerar o raio correspondente a  $m_r$  e calcular sua distância  $d$  do ponto  $P_1$ . Caso  $d$  seja inferior a  $D_{MAX}$ , *registrar um novo marcador 3D em 3F<sup>t</sup>* utilizando as coordenadas de  $P_1$  (atualizadas com o raio de  $m_r$ ) e **marcar  $m_p$ ,  $m_q$  e  $m_r$**  nos respectivos *frames*  $F_i^t$ ,  $F_j^t$  e  $F_k^t$  com o ID fornecido ao novo marcador; faça  $l$  igual a  $k$  e vá para o passo 5a. Caso  $d$  seja maior que  $D_{MAX}$ , vá para o passo 4b;
- 5a. Incrementar  $l$  e escolher o próximo *frame*  $F_l^t$ . Projetar o ponto  $P_1$  em  $F_l^t$  a fim de obter o ponto  $PR_l$ ; se não houver mais *frames*, reiniciar a busca por um novo marcador 3D voltando ao passo 2b.

- 5b. Escolher o próximo marcador não identificado  $m_s$  de  $F_l^t$  que esteja a uma distância inferior a  $W$  de  $PR_l$ ; se não houver mais marcadores, voltar ao passo 5a;
- 5c. Gerar o raio correspondente a  $m_s$  e calcular sua distância  $d$  do ponto  $P_1$ . Caso  $d$  seja inferior a  $D_{MAX}$ , marcar  $m_s$  como identificado utilizando o ID fornecido ao novo marcador 3D no passo 4c e então vá para o passo 5a. Caso contrário, ir para o passo 5b;

## 5.4 Aspectos de Desenvolvimento e Implementação em Software

O método apresentado neste documento foi implementado em software utilizando a plataforma .NET da Microsoft em conjunto com a linguagem de programação C#. Os elementos visuais necessários para apresentação dos dados reconstruídos foram codificados utilizando a biblioteca gráfica DirectX.

A versão atual da técnica foi obtida após realização de diversos ciclos de atividade envolvendo pesquisa, desenvolvimento, aprimoramento, implementações e testes, conforme ilustra a Figura 5.5. Os diversos experimentos utilizando o software desenvolvido foram realizados em um computador com processador Intel® Core™ 2 Quad de 2.66 GHz e 4 GB de memória principal.

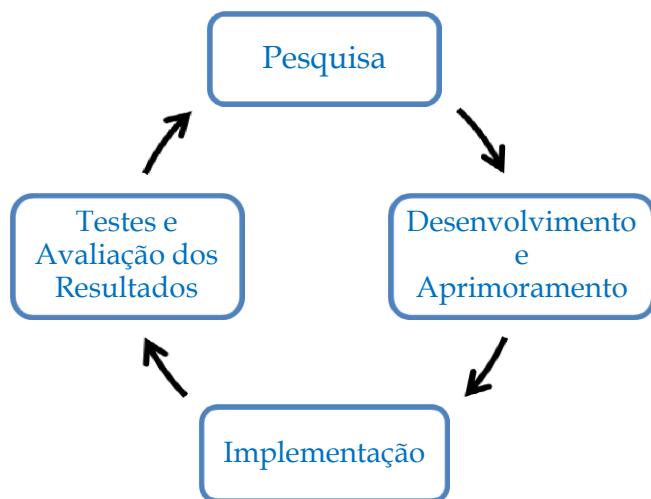


Figura 5.5 – Ciclo básico utilizado durante o desenvolvimento da técnica e implementação do software correspondente.

No restante deste documento, o software de reconstrução experimental desenvolvido com base na técnica proposta é referenciado utilizando o termo *protótipo*. A primeira etapa dos experimentos é apresentada no próximo capítulo.

# Capítulo 6

## Experimentos Utilizando Objeto de Teste

Vários experimentos foram realizados com o objetivo de testar e avaliar a técnica de rastreamento e reconstrução proposta. Os primeiros testes foram executados com a finalidade de avaliar a precisão e exatidão do método, assim como sua capacidade básica de identificação e rastreamento de um conjunto de marcadores.

Para avaliar esses elementos, executou-se um procedimento semelhante àquele realizado por Richards (1999). Em seu trabalho, o autor propõe um dispositivo de teste para avaliar o desempenho de vários sistemas comerciais, incluindo sistemas dos fabricantes Vicon, Qualisys, BTS e Motion Analysis. O dispositivo sugerido por Richards sustenta sete marcadores em posições conhecidas e os movem sistematicamente dentro de um pequeno volume de reconstrução. Neste capítulo, são apresentados os resultados de diversos experimentos realizados utilizando os critérios de avaliação de Richards e um dispositivo similar. O objetivo essencial é verificar o desempenho do método proposto e relacionar os resultados com aqueles disponíveis no artigo a respeito dos sistemas comerciais.

Os principais objetivos específicos são:

- a) Avaliar a precisão e exatidão da técnica proposta para reconstruir e rastrear marcadores quando os mesmos aparecem no campo de visão de todas as câmeras;

- b) Avaliar a precisão e exatidão da técnica proposta considerando uma situação em que apenas um subgrupo dinâmico de câmeras visualiza os marcadores à medida que eles se movem dentro do volume de reconstrução;
- c) Avaliar a capacidade da técnica em identificar e rastrear os marcadores *em tempo real* nas duas situações mencionadas anteriormente;
- d) Avaliar a precisão, exatidão e capacidade de rastreamento em tempo real em situações em que dois marcadores se aproximam efetivamente um do outro;
- e) Avaliar os efeitos da oclusão de marcadores e da sobreposição dos pontos nas imagens sobre a precisão, exatidão e capacidade de rastreamento da técnica proposta.

O capítulo está organizado como segue. A Seção 6.1 apresenta as condições gerais em que os experimentos foram conduzidos, incluindo informações a respeito do conjunto de câmeras, da disposição das mesmas no ambiente, dos aspectos de calibração e dos parâmetros utilizados na reconstrução. A Seção 6.2 descreve o dispositivo de teste. A Seção 6.3 apresenta os procedimentos empregados para realização da captura, assim como os resultados obtidos. Inicialmente, os resultados são apresentados e discutidos isoladamente. Em seguida, os resultados são relacionados com aqueles divulgados por Richards (1999) a respeito dos sistemas comerciais.

## **6.1 Condições Gerais de Realização dos Experimentos**

---

Para conduzir todos os experimentos apresentados nesta tese utilizou-se um conjunto de oito câmeras especializadas de infravermelho, modelo *OptiTrack Flex:V100R2*, comercializadas pela empresa *NaturalPoint*. Cada câmera foi configurada para operar em sua velocidade máxima, a uma taxa de 100 quadros por segundo.

### **6.1.1 Posicionamento das Câmeras**

O conjunto de câmeras foi distribuído em uma configuração retangular (6,5m x 5m) com o auxílio de quatro tripés de sustentação. Cada tripé foi posicionado em um dos cantos do retângulo e utilizado para sustentar duas câmeras, as quais foram fixadas a

alturas de 1,85m e 2,50m, respectivamente, e orientadas para foco na região central do retângulo. A Figura 6.1 ilustra o posicionamento das câmeras e dos tripés. Esta configuração possibilitou um volume de reconstrução de aproximadamente 3m x 2,5m x 2m, em comprimento, largura e altura, respectivamente.

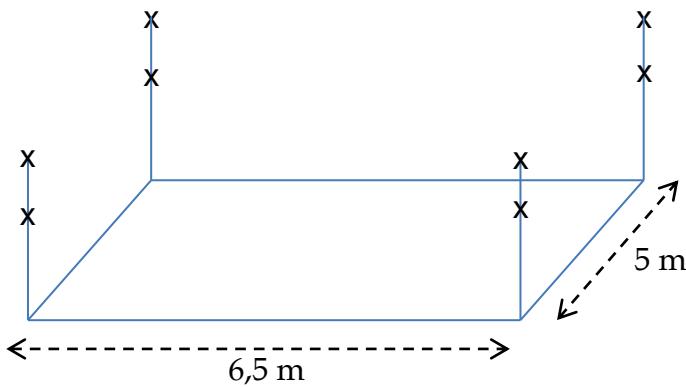


Figura 6.1 - Esquema de posicionamento das câmeras. Um tripé em cada canto é utilizado para sustentar duas câmeras. Na figura, cada X representa uma câmera de infravermelho.

### 6.1.2 Calibração

O conjunto de câmeras foi calibrado imediatamente antes da execução dos experimentos utilizando a técnica descrita na Seção 4.2.4 implementada no protótipo. Todas as coletas utilizando o objeto de teste foram executadas logo após a calibração e compartilharam os mesmos parâmetros internos e externos obtidos.

### 6.1.3 Parâmetro de Rastreamento 2D

Para permitir a identificação e o rastreamento dos marcadores colocados no objeto de teste, o parâmetro de rastreamento bidimensional,  $RP_{2D}$ , foi definido em 2 pixels. Esse valor foi escolhido empiricamente, após observação dos resultados quando da realização de diversos testes. Um estudo detalhado mostrando como a técnica reage à variação desse parâmetro é apresentado no capítulo seguinte.

## 6.2 Descrição do Dispositivo

O dispositivo construído para conduzir os testes (Richards, 1999) é composto essencialmente por um tubo e uma placa de PVC, um motor e uma base de madeira. O tubo de PVC possui comprimento de 80 cm e é sustentado pelo eixo do motor,

conforme mostra a Figura 6.2. O motor é fixado a uma base de madeira e é capaz de girar o tubo horizontalmente a uma velocidade de até 40 RPM.

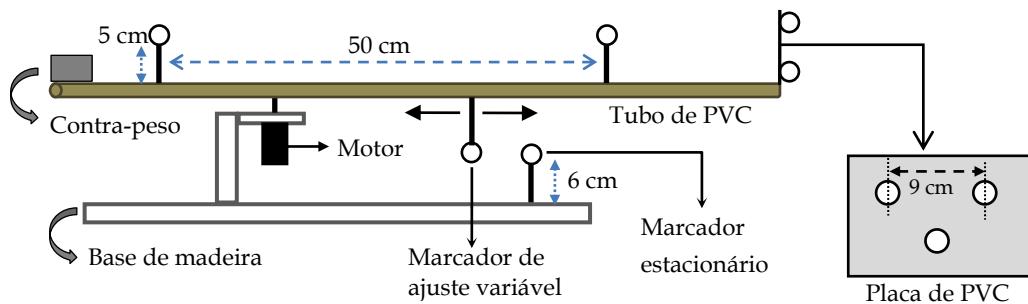


Figura 6.2 - Esquema do dispositivo utilizado para avaliar e comparar a precisão e exatidão do método proposto.

O dispositivo sustenta um total de sete marcadores de 1,6 cm de diâmetro, dos quais seis são suportados e movidos pelo tubo em torno do eixo de rotação do motor. Dois desses marcadores são fixados a 50 cm de distância um do outro e permanecem no campo de visão de todas as câmeras. Três marcadores são fixados em uma placa de PVC formando um padrão triangular e a placa é presa perpendicularmente na extremidade do tubo, com os marcadores voltados para o lado externo. Os dois marcadores superiores da placa são separados por uma distância de 9 cm e o marcador inferior é ajustado na região central formando um ângulo de 94,2° com os outros dois. O objetivo da placa é possibilitar que apenas um subconjunto dinâmico de câmeras visualize esses marcadores à medida que o tubo gira.

O único marcador estacionário do dispositivo é fixado na base de madeira por uma haste de 6 cm e o *marcador de ajuste variável* é sustentado na parte inferior do tubo por uma haste que o deixa nivelado ao marcador estacionário. A haste foi instalada de maneira que sua posição ao longo do tubo pudesse ser ajustada, o que torna possível controlar a distância mínima entre o marcador estacionário e o marcador que orbita o eixo de rotação do motor.

### 6.3 Procedimento de Captura e Resultados

Os experimentos realizados com o dispositivo de teste foram conduzidos em duas etapas. Na primeira etapa, teve-se como objetivo analisar a exatidão, a precisão e a

capacidade básica de rastreamento do método considerando os dois marcadores separados por 50 cm e os três marcadores colocados sobre a placa de PVC. Neste momento, o marcador de ajuste variável foi colocado a uma distância de 5 cm do marcador fixo.

Na segunda etapa, foram realizadas seis coletas onde o marcador de ajuste variável foi colocado a distâncias de 5, 4, 3, 2, 1 e 0 cm, respectivamente, do marcador estacionário.

Em todas as coletas a velocidade do motor foi ajustada para 40 RPM e cada aquisição teve duração de 15 segundos, tempo suficiente para o tubo percorrer 10 voltas ao redor do eixo de rotação. As duas etapas são apresentadas a seguir.

### 6.3.1 Primeira Etapa - Marcadores Superiores e Marcadores da Placa

Na primeira etapa dos experimentos, os seguintes elementos foram avaliados:

- a) A capacidade da técnica em identificar e rastrear os marcadores;
- b) **A distância entre os dois marcadores do topo do tubo.** A distância entre os dois marcadores foi calculada para cada *frameset* a partir das respectivas coordenadas 3D encontradas pelo algoritmo de reconstrução. Utilizando as distâncias calculadas e a distância real conhecida de 50 cm, computou-se a *distância média*, o *erro RMS* ( $E_{RMS}$ ) e o *erro absoluto máximo* ( $E_{MAX}$ ). O ruído associado às medidas foi determinado por meio do cálculo do desvio padrão das medidas de distância obtidas para cada *frameset*. O erro RMS foi calculado com base na equação:

$$E_{RMS} = \sqrt{\frac{1}{n} \sum_{k=0}^n (d_k - D)^2}, \quad (6.1)$$

onde  $n$  é o número de *framesets*,  $d_k$  é a distância obtida para o *frameset*  $k$  e  $D$  é a distância real (50 cm) dos dois marcadores;

- c) **A distância dos dois marcadores superiores da placa de PVC.** Assim como no caso anterior, esta distância também foi calculada para cada *frameset* à medida que a placa se deslocou ao longo de sua órbita. Igualmente, com base nos

múltiplos valores de distância obtidos e na distância real conhecida de 9 cm, pôde-se calcular a distância média, o erro RMS, o erro absoluto máximo e o desvio padrão dessas medidas;

- d) **O ângulo formado pelos três marcadores da placa de PVC.** Além da distância dos marcadores superiores da placa, calculou-se, de maneira análoga, o ângulo médio formado pelos três marcadores e os erros associados.

Conforme já mencionado, a placa de PVC foi proposta para que apenas um subconjunto de câmeras pudesse visualizar os seus marcadores à medida que o tubo rotaciona. Buscando conhecer melhor a variação desse subconjunto, executou-se um procedimento simples, em que todos os marcadores foram retirados do dispositivo, exceto aquele localizado na parte inferior da placa. Com apenas um marcador presente, o dispositivo foi acionado e o software programado para registrar o número de câmeras detectando o único marcador ao longo de três voltas. Como resultado, obteve-se o gráfico da Figura 6.3.

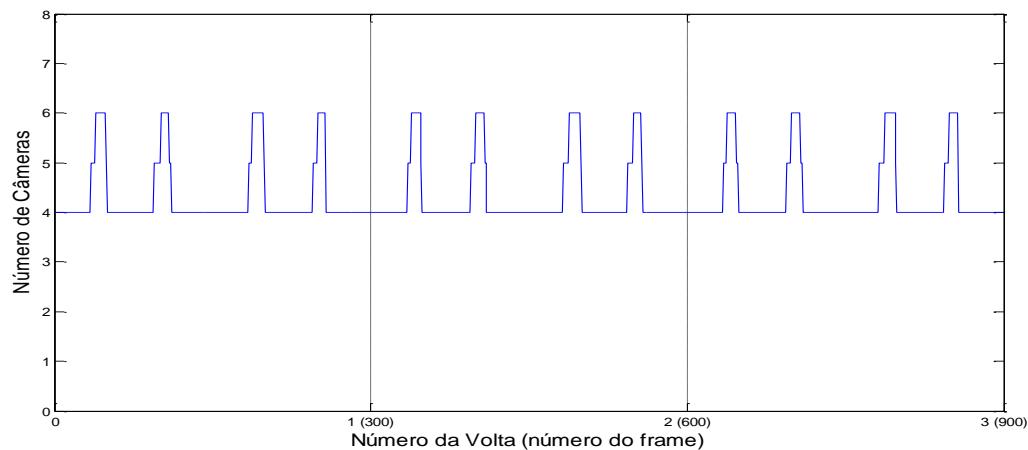


Figura 6.3 - Número de câmeras visualizando o marcador inferior da placa de PVC ao longo de três voltas. O gráfico mostra que o marcador é visualizado, a maior parte do tempo, por quatro câmeras.

Analizando o gráfico, pode-se observar que o marcador se manteve no campo de visão de quatro câmeras a maior parte do tempo. Ao avançar cada quarto de volta, o marcador entra no campo de visão de um novo par de câmeras, mas ainda permanece alguns instantes sendo visualizado pelas duas últimas, o que explica os quatro picos gerados para cada volta do marcador.

## Resultados

Os valores obtidos para a distância média, o desvio padrão, o erro RMS e o erro absoluto máximo para os marcadores do tubo de PVC e os dois marcadores superiores da placa são mostrados na Tabela 6.1.

Tabela 6.1 - Distâncias calculadas dos marcadores e erros associados.

	Dist. Real (cm)	Dist. Média (cm)	Desv. Padrão (cm)	Erro RMS (cm)	Erro Máx. (cm)
Marcadores Superiores do Tubo	50,0	50,083	0,052	0,098	0,317
Marcadores Superiores da Placa	9,0	9,013	0,111	0,112	0,665

A distância média calculada difere da distância real em menos de 1 mm para ambos os casos. O erro RMS relativo à distância dos dois marcadores superiores do tubo também foi inferior a 1 mm e o erro absoluto máximo foi de 3,17 mm. Com relação aos dois marcadores superiores da placa de PVC, o erro RMS foi ligeiramente maior (1,12 mm) e o erro absoluto máximo foi de aproximadamente o dobro daquele obtido para os outros dois marcadores.

Com relação à dispersão dessas medidas, foi encontrado um desvio padrão de 0,52 mm para os dois marcadores do topo e um desvio de 1,1 mm para os marcadores da placa. A maior precisão das medidas associadas aos marcadores do topo já era esperada, uma vez que os dois marcadores permanecem visíveis para todas as câmeras. Por outro lado, os marcadores da placa podem ser parcialmente visualizados durante a entrada e a saída nos campos de visão das câmeras, o que pode causar distorções nas coordenadas dos respectivos pontos nas imagens e consequentemente reduzir a precisão das coordenadas 3D reconstruídas.

A Tabela 6.2 apresenta os valores obtidos para a medida do ângulo definido entre os marcadores da placa de PVC. O valor médio calculado para o ângulo foi de 94,229 graus e o erro RMS das medidas foi de 0,76 grau. O erro absoluto máximo registrado foi de 3,62 graus.

Tabela 6.2 - Ângulo entre os marcadores da placa de PVC e erros associados.

Âng. Real (graus)	Âng. Médio (graus)	Desv. Padrão (graus)	Erro RMS (graus)	Erro Máx. (graus)
94,20	94,229	0,761	0,762	3,625

### 6.3.2 Segunda Etapa - Marcador Fixo e Marcador Variável

Esta segunda etapa dos experimentos foi realizada com o objetivo de verificar o comportamento do método proposto, no que diz respeito à precisão, exatidão e capacidade de rastreamento, à medida que o marcador variável é posto a distâncias de 5, 4, 3, 2, 1 e 0 cm do marcador estacionário. A condição de proximidade dos marcadores no espaço pode levar à ocorrência de projeções muito próximas nas imagens ou até mesmo em estado de sobreposição, o que pode interferir diretamente na capacidade de rastreamento e reconstrução da técnica. Adicionalmente, os dois marcadores inferiores podem ser alvos de oclusões causadas pelo próprio tubo de PVC e os demais elementos do dispositivo, que a cada volta podem cobrir a visão desses marcadores com respeito a uma ou mais câmeras.

Para avaliar as respostas do algoritmo às variações de proximidade dos marcadores, decidiu-se analisar os seguintes elementos:

- A capacidade da técnica em rastrear o marcador fixo e o marcador variável ao longo de cada coleta;
- A variação de posição do marcador fixo ao longo de cada coleta devido à interferência causada pela aproximação do marcador orbitante;
- A distância mínima calculada entre os dois marcadores, para cada coleta.

### Resultados

Com relação à capacidade de identificação e rastreamento, o protótipo foi capaz de identificar e rastrear os dois marcadores ao longo das quatro primeiras coletas (para as distâncias de 5, 4, 3 e 2 cm). Entretanto, o rastreamento correto não foi obtido quando os marcadores foram separados pelas distâncias de 1 e 0 cm. Para cada volta em torno do eixo, ao se aproximar do marcador fixo, o marcador de posição variável teve o rastreamento interrompido (e também o marcador fixo). Além disso, durante

os instantes de cruzamento, um falso marcador foi identificado entre os dois marcadores “reais”. O marcador fantasma permaneceu visível até o afastamento do marcador móvel.

Para avaliar a precisão do sistema com relação ao posicionamento do marcador fixo, foi calculada a posição 3D média desse marcador a partir das posições calculadas quadro a quadro. Em seguida, utilizando sua posição média, foi calculada a variação máxima de sua posição ao longo da coleta, assim como a variação RMS. Esses valores são apresentados nas colunas 2 e 3 da Tabela 6.3.

Tabela 6.3 - Erro associado à posição do marcador fixo à medida que o marcador de posição variável é ajustado em direção ao fixo. Adicionalmente são apresentadas as distâncias reais e as distâncias calculadas entre os centros dos dois marcadores.

Dist. Mín. entre os Marcadores (cm)	Variação RMS Marc. Fixo (cm)	Var. Máx. Marc. Fixo (cm)	Dist. Real Mín. entre Centros dos Marc. (cm)	Dist. Mín. Média Calculada (cm)	Erro (cm)
5	0,028	0,217	6,6	6,713	0,113
4	0,031	0,251	5,6	5,709	0,109
3	0,044	0,387	4,6	4,355	0,245
2	0,081	0,918	3,6	3,225	0,375

Adicionalmente, foi calculada a distância mínima média entre os dois marcadores. Durante os 15 segundos de cada coleta, o marcador orbitante cruzou o marcador fixo um total de 10 vezes. Para cada volta, registrou-se a menor distância entre os dois marcadores e esses valores foram utilizados para o cálculo da distância mínima média. Essas distâncias são mostradas nas colunas 4 e 5 da Tabela 6.3. Repare que o algoritmo calcula a distância entre os *centros* dos marcadores, o que é diferente da distância definida *entre* os marcadores, devido ao seu diâmetro de 1,6 cm.

A variação de posição RMS do marcador estacionário foi inferior a 1 mm para as quatro coletas em que os marcadores foram rastreados com sucesso. Entretanto, observa-se um aumento da variação à medida que a distância diminui. A variação RMS do marcador estacionário quando o marcador orbitante é colocado a 2 cm é cerca de três vezes maior que a variação obtida para a distância de 5 cm. O mesmo comportamento se repete para a maior variação de posição registrada, que aumentou de 2,1 para 9,1 mm. Um dos fatores responsáveis por essa alteração é a ocorrência

mais frequente de sobreposições dos pontos nas imagens, que é mais provável quando os marcadores físicos se aproximam.

A incapacidade de rastreamento dos dois marcadores quando eles foram colocados a distâncias de 1 ou 0 cm, se deve, principalmente, à baixa resolução dos sensores das câmeras. A fim de avaliar este quesito, foram colocados exatamente dois marcadores sobre o chão na região central do volume, separados um do outro pela distância de 1 cm. Ao acionar as câmeras, percebeu-se que 5 das 8 câmeras do conjunto não foram capazes de distinguir os marcadores. Essas câmeras registraram a ocorrência de apenas um ponto na imagem: aquele resultante da sobreposição das projeções dos dois marcadores.

### 6.3.3 Relacionando os Resultados

Nesta seção, os resultados apresentados anteriormente são relacionados com os resultados obtidos por Richards (Richards, 1999), que empregou um dispositivo análogo para avaliar diversos sistemas comerciais.

A Tabela 6.4 apresenta as características básicas dos sistemas avaliados pelo autor. Com exceção do sistema Ariel, todos os demais utilizam um conjunto de 6 câmeras de infravermelho com velocidades de captura que vão de 120 FPS (ElitePlus, da BTS) a 1000 FPS (ProReflex, da Qualisys). O sistema da Ariel Dynamics emprega 5 câmeras de vídeo convencionais que podem registrar até 30 quadros por segundo.

Tabela 6.4 – Sistemas avaliados por Richard utilizando dispositivo ilustrado anteriormente.

Fabricante	Sistema	Câmeras	Velocidade (FPS)
Ariel Dynamics	Ariel System	5	30
Motion Analysis	HiRes System	6	240
Peak Performance	Motus System	6	200
Qualisys	ProReflex System	6	1000
BTS	ElitePlus System	6	120
Vicon	370 System	6	240
Charnwood Dynamic	CODA System	6	200

É importante salientar que os dados apresentados nesta seção não têm como objetivo comparar o protótipo, de maneira exata, com os softwares utilizados nos sistemas comerciais. Uma análise dessa natureza seria possível apenas se todos os parâmetros e condições de captura, assim como os elementos de hardware, pudessem ser os mesmos; o que não é o caso. Os sistemas comerciais utilizam câmeras diferentes, organizadas em arranjos distintos e em posições não conhecidas com exatidão. Embora o dispositivo de teste tenha sido construído com base no artigo de Richards (1999), alguns parâmetros importantes, como as dimensões da placa de PVC, não foram informados pelo autor, o que impossibilitou a criação de um dispositivo com propriedades exatamente iguais ao original. Apesar de todos esses fatores, que podem influenciar os resultados, os dados divulgados pelo autor dizem respeito à precisão, exatidão e capacidade de rastreamento de sistemas de *mocap* de renome mundial e que foram avaliados por meio de um experimento passível de ser reproduzido em essência. Assim, decidiu-se relacionar os resultados, ainda que para uma comparação aproximada entre os diversos softwares comerciais e o protótipo. Essa relação é apresentada a seguir.

No gráfico (a) da Figura 6.4, a distância média entre os dois marcadores do topo, calculada utilizando o protótipo, foi inserida no conjunto de valores divulgado por Richards. Avaliando a proximidade com o valor real de 50 cm, apenas os sistemas da *Vicon* e da *Peak Performance* foram mais exatos que o protótipo desenvolvido.

O gráfico (b) da Figura 6.4 relaciona os resultados com respeito à dispersão das medidas de distância calculadas por cada sistema quadro a quadro (com relação aos dois marcadores separados por 50 cm). Neste contexto, o protótipo se posiciona gerando as medidas mais precisas, seguido pelo sistema da *Motion Analysis*.

O gráfico (c) apresenta as distâncias médias calculadas entre os dois marcadores superiores da placa de PVC. Com o valor de 9,013 cm, o protótipo obteve o segundo melhor resultado, ficando atrás do sistema *ProReflex* da *Qualysis*.

O gráfico (d) exibe a dispersão das distâncias calculadas quadro a quadro para os dois marcadores da placa. Neste caso, o menor valor foi registrado pelo protótipo.

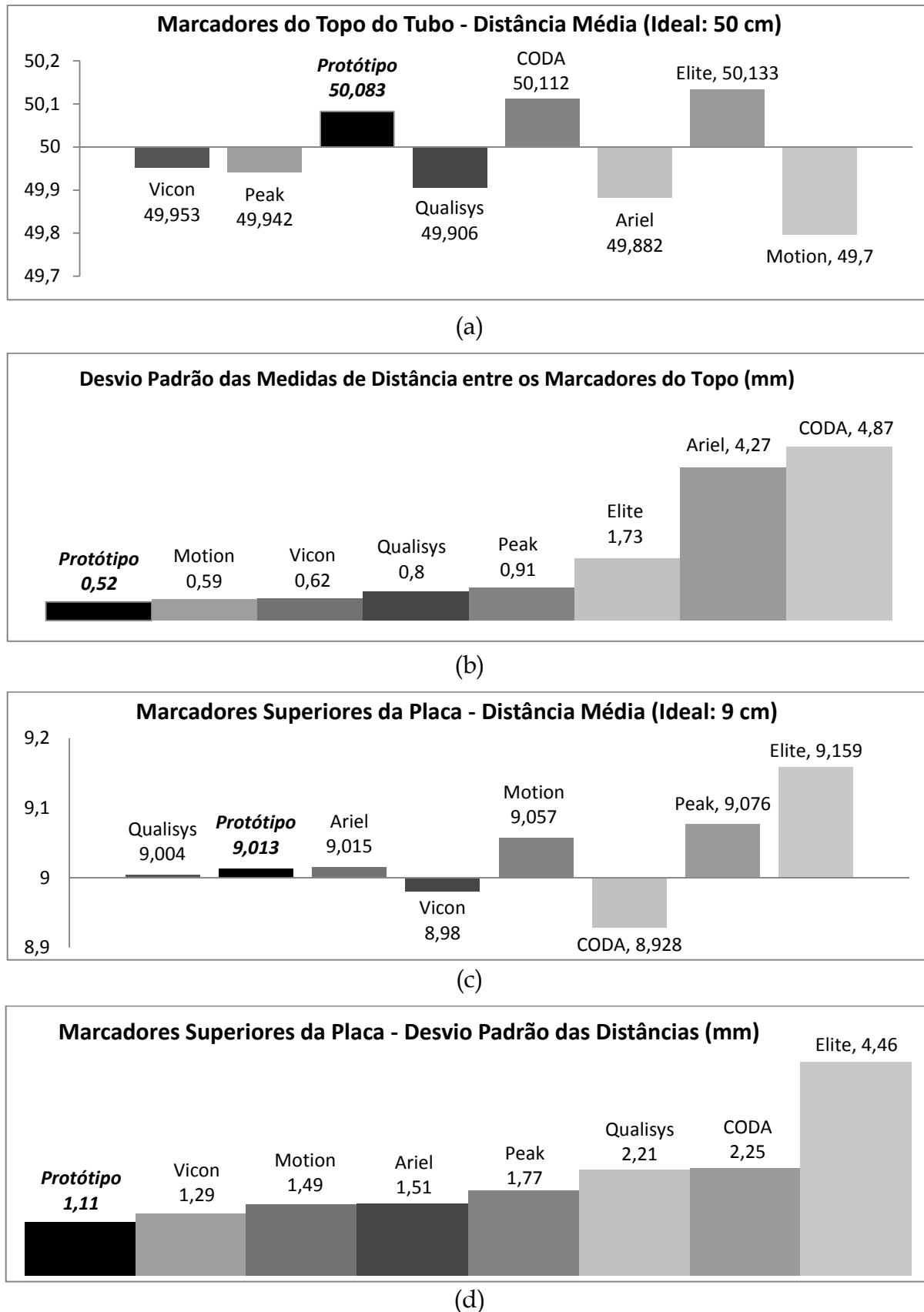


Figura 6.4 – (a) Distância média entre os dois marcadores do topo do tubo de PVC. A distância calculada utilizando o protótipo foi inserida no conjunto de valores obtido por Richards (1999); (b) dispersão dos valores de distância calculados quadro a quadro entre os dois marcadores do tubo; (c) distância média entre os dois marcadores superiores da placa de PVC e (d) dispersão associada à distância entre os marcadores superiores da placa.

Com relação ao ângulo definido pelos três marcadores da placa, os resultados são relacionados de maneira diferente. No objeto construído por Richards, o marcador inferior, tido como o vértice, é posicionado formando um ângulo de 95,8º com os outros dois. Já no dispositivo desenvolvido para avaliar o protótipo, o mesmo foi posto formando um ângulo de 94,2º, diferente do valor original devido à dificuldade em posicionar precisamente tal marcador na placa. Assim, para relacionar os valores, empregou-se a diferença entre o valor real e o valor calculado, conforme apresenta o gráfico (a) da Figura 6.5. O valor do ângulo calculado pelo protótipo corresponde ao segundo mais próximo do valor real.

A dispersão associada às medidas dos ângulos é apresentada no gráfico (b) da mesma figura. O desvio de 0,76 graus obtido pelo protótipo constitui o menor deles.

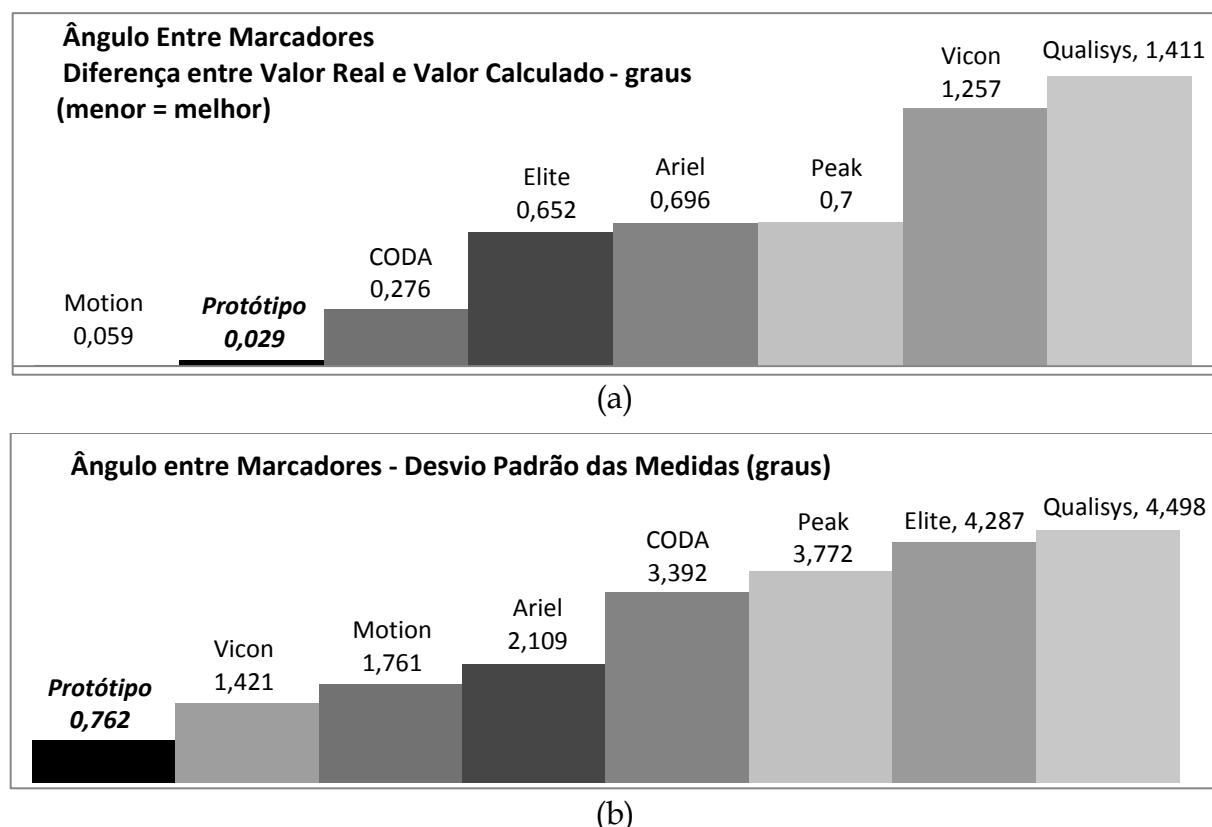


Figura 6.5 - (a) Diferença entre o valor calculado e valor real do ângulo entre os marcadores da placa de PVC. No dispositivo utilizado por Richards, o marcador do vértice foi colocado formando um ângulo de 95,8º com os outros dois. Já no dispositivo construído para avaliar o protótipo, o ângulo formado foi de 94,2º; (b) desvio padrão das medidas dos ângulos calculadas quadro a quadro.

Com relação ao rastreamento do marcador fixo e do marcador de ajuste variável, o artigo informa que todos os sistemas, exceto o Vicon 370 e o Peak Motus, foram capazes de rastreá-los para todas as distâncias consideradas (5, 4, 3, 2, 1 e 0 cm).

Entretanto, foi necessário que os operadores dos respectivos sistemas, em uma etapa de pós-processamento, utilizassem os recursos dos softwares para corrigir/recalcular as trajetórias dos dois marcadores quando a distância mínima considerada foi de 1 ou 0 cm. No caso dos sistemas da Vicon e da Peak Performance, os dois marcadores tiveram as trajetórias trocadas e os recursos de correção e otimização dos respectivos softwares não possibilitaram que os operadores corrigissem o problema.

### ***Discussão***

A quantidade e a velocidade das câmeras correspondem a duas importantes diferenças entre o sistema que utiliza o protótipo e os sistemas comerciais. Enquanto o protótipo usufrui de duas câmeras a mais, os softwares dos sistemas comerciais (exceto o Ariel) gozam de câmeras mais rápidas, o que traz vantagens para o processo de rastreamento (por diminuir a distância entre os pontos em quadros consecutivos). As duas câmeras extras utilizadas pelo protótipo o coloca em posição vantajosa especialmente quando são considerados os critérios de avaliação envolvendo os marcadores da placa de PVC. Naturalmente, há um número maior de câmeras suportando a reconstrução desses marcadores à medida que o tubo rotaciona, o que contribui para uma diminuição do erro associado às medidas. Esse fato colaborou para o posicionamento ótimo do protótipo com relação aos outros sistemas, conforme apresentado no gráfico (d) da Figura 6.4 e no gráfico (b) da Figura 6.5.

## **6.4 Conclusão e Considerações do Capítulo**

---

Os experimentos apresentados neste capítulo foram conduzidos com a finalidade de avaliar a precisão e a exatidão do método proposto, assim como sua capacidade básica de identificação e rastreamento de um conjunto de marcadores. Os testes foram realizados utilizando o software desenvolvido (protótipo) e o objeto de teste proposto por Richards (1999).

Durante os experimentos, o objeto foi posto em movimento circular e o protótipo acionado para reconstruir as trajetórias de seus sete marcadores. A partir das coordenadas reconstruídas, calculou-se as distâncias e a angulação entre os

marcadores. Esses valores foram empregados, juntamente com as medidas reais, para estimar a precisão e a exatidão disponibilizada pelo sistema.

A exatidão RMS encontrada com relação aos marcadores separados por 50 cm foi inferior a 1 mm e a precisão das medidas de distância foi de 0,52 mm. Com respeito aos marcadores superiores da placa de PVC, os quais permaneceram visíveis apenas para um subgrupo de câmeras, a exatidão RMS obtida foi de 1,12 mm e a precisão, 1,11 mm. Já a exatidão RMS associada à posição do marcador estático sofreu variação de 0,28 mm a 0,82 mm.

O protótipo foi capaz de rastrear, automaticamente e em tempo real, os sete marcadores do dispositivo em quatro das seis coletas realizadas. A baixa resolução das câmeras juntamente com as imperfeições intrínsecas da técnica impossibilitou o rastreamento automático do marcador fixo e do marcador variável quando eles foram postos a uma distância igual ou inferior a 1 cm.

Os resultados obtidos sobre precisão, exatidão e capacidade de rastreamento do protótipo foram relacionados com os valores divulgados por Richards a respeito de sete sistemas comerciais. Apesar das diferenças com relação aos elementos de hardware e às condições de captura, uma comparação aproximada dos resultados coloca o protótipo em posição de destaque, especialmente com respeito à precisão, com relação aos vários sistemas comerciais avaliados.

# Capítulo 7

## Experimentos Utilizando Manequim

Os experimentos apresentados no capítulo anterior foram executados com o objetivo de avaliar a capacidade básica de rastreamento do método, sua precisão e exatidão. Considerou-se um ambiente bem controlado e um número reduzido de marcadores, que foram colocados em um objeto de teste especialmente desenvolvido para análise. Este capítulo apresenta os resultados de um experimento que envolveu a reconstrução de dezenas de marcadores, os quais foram colocados em um modelo corporal humano representado por um manequim. A movimentação dos marcadores foi induzida utilizando-se uma base giratória sobre a qual o manequim foi fixado.

A escolha desse experimento foi motivada pelos seguintes fatores:

- A necessidade de avaliar a técnica para capturar movimentos de marcadores organizados em uma estrutura corporal humana;
- A necessidade de avaliar a técnica para rastrear e reconstruir as trajetórias de um número maior de marcadores na presença de oclusões e sobreposições frequentes dos pontos nas imagens;
- O interesse em relacionar diretamente os resultados do protótipo com os resultados de um sistema comercial. Este fator teve especial influência considerando a disponibilidade do sistema *OptiTrack* da empresa NaturalPoint no laboratório onde a pesquisa foi conduzida;

- A necessidade de se trabalhar com movimentos, para fins de avaliação e comparação, que pudessem ser precisamente repetidos e ao mesmo tempo controlados com relação às trajetórias percorridas, velocidades e acelerações dos marcadores.

Após avaliar os fatores listados anteriormente, considerou-se a possibilidade de utilizar um manequim ao invés de um indivíduo, sobre o qual poderiam ser fixados muitos marcadores. Para induzir a ocorrência de sobreposições e oclusões dos marcadores e ao mesmo tempo tornar o experimento controlável, fácil de ser realizado e repetido, decidiu-se pelo desenvolvimento de uma plataforma giratória de velocidade ajustável, sobre a qual o manequim poderia ser posto e ter seus marcadores reconstruídos e rastreados enquanto o conjunto gira em torno de seu eixo.

Os resultados dos experimentos utilizando o manequim e a plataforma giratória são apresentados neste capítulo em duas etapas principais. Primeiramente, os dados 2D coletados durante as sessões de captura são utilizados para análise isolada da técnica de reconstrução. Nesse momento, o desempenho do método proposto é avaliado mediante a variação do parâmetro  $RP_{2D}$  do algoritmo e a comparação dos dados 3D obtidos. Em uma segunda etapa, o desempenho do protótipo é comparado ao desempenho do software comercial Arena por meio da avaliação dos resultados tridimensionais gerados pelos dois programas.

O capítulo está organizado da seguinte forma. A Seção 7.1 descreve brevemente a plataforma giratória desenvolvida. A Seção 7.2 apresenta o padrão escolhido para o posicionamento dos marcadores no manequim. A Seção 7.3 traz em detalhes todos os critérios empregados para avaliar e comparar a qualidade dos dados 3D produzidos. A Seção 7.4 apresenta os procedimentos adotados para a realização das coletas. Na Seção 7.5 é feita uma avaliação do desempenho do método quando da variação do parâmetro de reconstrução  $RP_{2D}$ . Na Seção 7.6 são mostrados os resultados comparativos calculados para o protótipo e para o software comercial Arena e a Seção 7.7 conclui o capítulo com as considerações finais.

## 7.1 Descrição da Plataforma

A plataforma giratória construída para a realização dos experimentos é composta basicamente por uma base de ferro, um rolamento, um eixo de aço, um prato giratório, um motor elétrico e um sistema de polias. O prato giratório é fixado ao eixo, que por sua vez está preso à base de ferro por meio do rolamento. O prato giratório é conectado ao motor pelo sistema de polias e correias, utilizado para aumentar o torque e reduzir a velocidade de rotação.

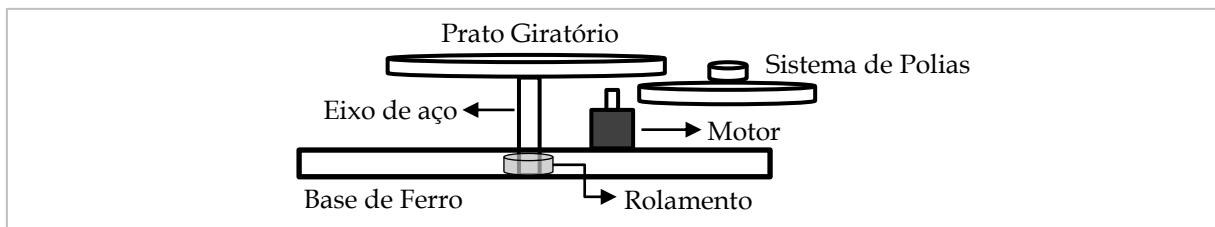


Figura 7.1 – Composição da plataforma giratória desenvolvida para rotacionar o manequim.

A velocidade de rotação do prato é controlada pelo ajuste de velocidade do motor. Para este fim foi desenvolvido um dispositivo eletrônico capaz de aumentar ou diminuir a potência do motor com base no ajuste de um potenciômetro. A velocidade de rotação máxima alcançada pela plataforma sustentando o manequim foi de aproximadamente 43 RPM. O esquema da plataforma é apresentado na Figura 7.1.

## 7.2 Posicionamento dos Marcadores

Atualmente existem vários padrões que definem o posicionamento dos marcadores para a captura de movimentos corporais. A quantidade de marcadores a utilizar e o posicionamento de cada um depende de vários fatores como o tipo de movimento a ser capturado, o propósito da captura (animação, avaliação clínica, simulação, etc.), o tamanho do volume de reconstrução e a quantidade e qualidade das câmeras (Manal, Mcclay *et al.*, 2000; Rosenhahn, Klette *et al.*, 2007; Kitagawa e Windsor, 2008; Kiser, 2010). Em geral, cada fabricante de sistemas de *mocap* adota uma ou mais configurações para o posicionamento dos marcadores. A configuração mais adequada é escolhida de acordo com o propósito da captura. Por exemplo, em um

dos arranjos utilizados pela Vicon, um total de 41 marcadores é distribuído pelos principais segmentos do corpo (Robertson, 2009).

Para conduzir os experimentos apresentados neste capítulo, adotou-se a configuração de 38 marcadores utilizada pela empresa *NaturalPoint* em seu sistema *OptiTrack*. Esta configuração foi escolhida devido ao interesse em relacionar os resultados do protótipo com os resultados do software comercial *Arena* (que integra o sistema *OptiTrack*), sendo esse um dos padrões aceitos pelo software. A Figura 7.2 apresenta o posicionamento dos marcadores no manequim utilizado nos testes.



Figura 7.2 - Posicionamento dos marcadores no manequim. Padrão de 38 marcadores utilizado pelo software *Arena* da empresa *NaturalPoint*. Cada marcador tem 1,6 cm de diâmetro.

### 7.3 Critérios de Avaliação dos Dados 3D

Para avaliar e comparar os dados 3D resultantes da reconstrução, foram definidos os seguintes critérios de análise:

1. Número de marcadores identificados no início da sequência e rastreados automaticamente até o final;
2. Número máximo de marcadores detectados em um *frame* 3D;
3. Número total de interrupções no rastreamento dos marcadores;
4. Quantidade média de ruídos do tipo *spike* detectados por trajetória inteiramente reconstruída;

## 5. Coeficiente geral de ruído.

Esses critérios são descritos a seguir.

### 7.3.1 Número de Marcadores Rastreados por toda Sequência

Encontrar o número de marcadores detectados no início da sequência e rastreados até o final contribui para uma avaliação da capacidade do método em rastrear automaticamente múltiplos marcadores e reconstruir suas trajetórias. Obviamente, se as outras características dos dados forem mantidas, quanto mais marcadores um sistema puder rastrear automaticamente ao longo de toda a captura, menor será a necessidade de intervenção e edição dos dados por parte do usuário.

O rastreamento de um marcador pode ser interrompido basicamente por dois fatores: (1) devido a algum tipo de oclusão, o marcador não pode, de fato, ser detectado por pelo menos duas câmeras do conjunto em algum ponto da captura; ou (2) devido a possíveis ambiguidades e/ou imperfeições intrínsecas da técnica de rastreamento e reconstrução empregada.

No primeiro caso, a interrupção no rastreamento é inevitável e o marcador não terá sua trajetória reconstruída inteiramente, a menos que se utilize alguma técnica adicional para *estimar* suas posições no trecho considerado, o que está fora do escopo deste trabalho. No segundo caso, a parada no rastreamento é um problema inerente à técnica empregada. Entretanto, no caso do protótipo, um eventual ajuste no parâmetro  $RP_{2D}$  do algoritmo pode intensificar ou reduzir as ocorrências de interrupção.

Vale reforçar que, em qualquer situação de captura, quanto maior o número de marcadores automaticamente rastreados do início ao fim da coleta, menor será o esforço realizado pelo usuário em uma eventual etapa de pós-processamento.

### 7.3.2 Número Máximo de Marcadores Detectados em um Frame 3D

Conforme apresentado no Capítulo 5, a técnica de rastreamento e reconstrução calcula um *frame 3D* para cada *frameset* obtido das câmeras. Cada *frame 3D* calculado pode conter um número diferente de marcadores, pois os fatores que influenciam em

suas detecções podem variar ao longo da coleta. O *número máximo de marcadores detectados em um frame 3D* pode ser encontrado após contagem da quantidade de marcadores detectados em cada um dos *frames* 3D gerados ao longo da sessão de captura.

Esse número é uma medida que está relacionada à quantidade de marcadores-fantasmas, ou falsos marcadores, calculados pelo software de captura. Se a quantidade real de marcadores é previamente conhecida, então a detecção de um número de marcadores superior a essa quantidade sugere a existência de pelo menos um falso marcador em algum *frame* 3D da sequência.

A detecção de marcadores-fantasmas é um problema comum em sistemas de captura de movimento ópticos (Johnson, Xie *et al.*, 2010a). Esses marcadores podem ocorrer até mesmo em sistemas comerciais de renome mundial como é o caso daqueles desenvolvidos pelas empresas Vicon e Motion Analysis (Kitagawa e Windsor, 2008). Falsos marcadores são sempre indesejáveis, pois não trazem dados úteis, podem confundir o usuário e, eventualmente, impedirem que outros marcadores reais sejam reconstruídos e/ou rastreados com maior precisão.

### 7.3.3 Número de Interrupções no Rastreamento 3D

Descobrir o número de marcadores que puderam ser rastreados do início ao fim da sequência é de especial importância para a avaliação da capacidade de rastreamento do método porque é uma medida de percepção fácil e direta. Entretanto, este número não traz nenhuma informação a respeito dos demais marcadores. Por exemplo, um marcador pode ter seu rastreamento interrompido no *frameset* 100, mas ser identificado e rastreado novamente a partir do *frameset* 120, enquanto outro marcador pode sofrer várias interrupções no rastreamento seguidas por novas detecções.

O *número total de interrupções no rastreamento 3D* dos marcadores é utilizado como um indicador do grau de fragmentação das trajetórias parcialmente reconstruídas, o que está associado à eficiência da técnica de rastreamento e reconstrução. Eventualmente,

a escolha dos parâmetros do algoritmo de rastreamento, como o valor de  $RP_{2D}$ , pode influenciar no número total de interrupções detectadas.

A Figura 7.3 ilustra como a contagem de interrupções é realizada. Cada X aparecendo na trajetória de um marcador representa uma interrupção em seu rastreamento. Caso este marcador seja detectado novamente (ainda que com outra identificação), é possível que outras interrupções aconteçam, as quais também são contabilizadas (marcador 5). Neste exemplo, há um total de quatro interrupções.

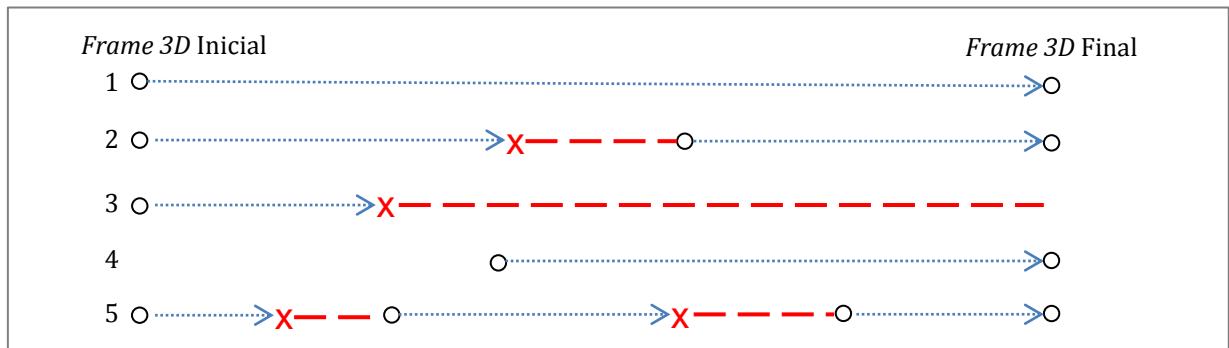


Figura 7.3 - Contagem do total de interrupções no rastreamento 3D. Cada X ilustra um ponto de interrupção no rastreamento. Neste exemplo, há um total de 4 interrupções.

É sempre desejável que o número de interrupções seja mínimo, pois cada parada no rastreamento leva à perda de identidade do marcador (ID) e pode exigir maior esforço do usuário numa etapa de pós-processamento e edição, onde eventualmente seriam aplicadas técnicas de interpolação ou similares para contornar o problema.

### 7.3.4 Número Médio de *Spikes* por Trajetória

Os dados gerados por sistemas de captura de movimento ópticos podem conter diversas formas de ruídos (Giakas e Baltzopoulos, 1997; Barca, Rumantir *et al.*, 2008; Kitagawa e Windsor, 2008; Lou e Chai, 2010). Um tipo de interferência comum normalmente encontrado nesses dados é aquela caracterizada por saídas abruptas do marcador de sua trajetória, geralmente de curta duração, seguida por recuperação do traçado normal. Esse tipo de ruído é comumente denominado de *spike* e é ilustrado na Figura 7.4.

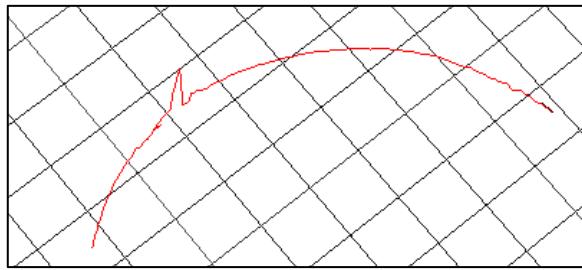


Figura 7.4 – Trajetória reconstruída de um marcador com a presença de ruído do tipo *spike*.

Com o objetivo de avaliar a qualidade dos dados gerados com relação à presença desse tipo de ruído, foi implementado um algoritmo simples capaz de analisar a trajetória reconstruída de um marcador e identificar os *spikes* de maior amplitude. O algoritmo monitora a velocidade corrente do marcador por meio de uma janela deslizante e utiliza o padrão de velocidade detectado para identificar uma eventual mudança brusca de trajetória, que seja de curta duração e incompatível com o traçado corrente do marcador. O algoritmo é empregado na contagem dos *spikes* para cada trajetória inteiramente reconstruída. O número total de *spikes* detectados em todas as trajetórias inteiramente reconstruídas é então utilizado para calcular o *número médio de spikes por trajetória*.

Dados 3D com menor quantidade de *spikes* resultam em trajetórias com traçados mais precisos, suaves e com menor necessidade da aplicação de filtros em etapas de pós-processamento. Vale ressaltar que nenhum tipo de filtro foi aplicado aos dados 3D gerados pelo protótipo.

### 7.3.5 Coeficiente Geral de Ruído

Ruídos do tipo *spike* são especialmente importantes porque eles causam distorções altamente perceptíveis nas trajetórias dos marcadores. Encontrar a quantidade média de *spikes* por trajetória fornece uma dimensão real a respeito da existência desses picos e de sua frequência. Entretanto, dados gerados por sistemas de captura ópticos também podem estar distorcidos por ruídos de menor amplitude (Giakas e Baltzopoulos, 1997). Visando obter um índice que esteja relacionado à quantidade geral de ruído existente nos dados gerados, utilizou-se um filtro digital de *Butterworth* do tipo passa-alta para isolar as componentes do sinal correspondentes a ruídos. O filtro foi aplicado para cada componente (x, y e z) dos pontos espaciais que

constituem cada trajetória reconstruída e em seguida foi calculado a média RMS de cada sinal resultante. Definiu-se como *coeficiente de ruído médio* da componente  $x$  ( $CRM_x$ ) a média aritmética das médias RMS obtidas para todas as trajetórias. O cálculo é ilustrado na Equação (7.1). O termo  $x(Traj_i)$  refere-se aos valores da coordenada  $x$  dos pontos que constituem a  $i$ -ésima trajetória inteiramente reconstruída. O termo *Filt* é usado para indicar a aplicação do filtro passa-alta e RMS para fazer referência ao cálculo da média. O número de trajetórias inteiramente reconstruídas é representado por  $n$ . Os valores  $CRM_y$  e  $CRM_z$  são obtidos de maneira análoga.

$$(7.1) \quad CRM_x = \frac{\sum_{i=1}^n RMS(Filt(x(Traj_i)))}{n}$$

O *coeficiente geral de ruído* (CR) dos dados 3D é definido como sendo a média dos valores  $CRM_x$ ,  $CRM_y$  e  $CRM_z$ , conforme ilustra a Equação (7.2).

$$(7.2) \quad CR = \frac{CRM_x + CRM_y + CRM_z}{3}$$

Para o cálculo do coeficiente geral de ruído dos dados 3D resultantes dos experimentos apresentados neste documento, utilizou-se um filtro *Butterworth* de quatro polos. A frequência de corte (FC) do filtro foi definida de acordo com o experimento. No caso daqueles envolvendo o manequim, utilizou-se uma FC de 2 Hz, uma vez que as componentes úteis do sinal possuem frequência inferior a 1 Hz.

Todos os critérios apresentados nas seções anteriores foram utilizados para avaliar a qualidade dos dados 3D gerados pelo protótipo, o que permitiu uma análise da sensibilidade do algoritmo com relação à mudança do parâmetro de rastreamento RP<sub>2D</sub>, assim como possibilitou uma comparação dos resultados com aqueles obtidos do software comercial Arena. A próxima seção descreve como o procedimento de captura em si foi realizado.

## 7.4 Procedimento de Captura

---

Primeiramente ambos os sistemas foram calibrados. Para o software comercial Arena, o conjunto de câmeras foi calibrado utilizando as funções do software voltadas para maximizar a precisão. Por exemplo, na calibração dinâmica com o bastão, escolheu-se a opção que visa estender o tempo de captura (*Wand data: Slow (More Accurate)*) com o objetivo de maximizar a precisão e a exatidão. O erro médio de reprojeção obtido foi inferior a 0,2 pixels para todas as câmeras do conjunto, marca que é apresentada no programa pelo indicativo *Excellent* (excelente). No caso do protótipo, a calibração foi realizada utilizando técnica análoga, discutida na Seção 4.2.4, e implementada no programa desenvolvido.

Foram realizadas quatro sessões de captura com a plataforma girando a velocidades de 10, 20, 30 e 40 RPM, respectivamente. Em cada sessão, os dados foram capturados durante 30 segundos utilizando o protótipo e por mais 30 segundos utilizando o software Arena.

Inicialmente a plataforma foi posta no centro do volume de reconstrução e o manequim foi fixado sobre a mesma. Em seguida, a plataforma foi acionada e a velocidade de giro ajustada para 10 RPM. Com a plataforma em giro constante, iniciou-se a sessão de captura de 30 segundos utilizando o protótipo desenvolvido. Na sequência (com a plataforma ainda a 10 RPM), executou-se a captura utilizando o software Arena. De maneira análoga, esse procedimento foi repetido mais três vezes, com a plataforma nas velocidades de 20, 30 e 40 RPM, respectivamente. Nenhum tipo de iluminação especial foi aplicado ao ambiente. Um conjunto de lâmpadas fluorescentes comuns manteve o ambiente bem iluminado.

Durante uma sessão de captura, o protótipo desenvolvido é capaz de realizar a identificação, a reconstrução e o rastreamento dos marcadores em tempo real, utilizando a técnica proposta neste trabalho. Os dados 3D resultantes são apresentados na tela do computador durante a sessão. Adicionalmente, os dados 2D

brutos recebidos das câmeras são gravados em arquivo<sup>3</sup> para reconstrução tridimensional futura em modo *off-line*. O salvamento dos dados brutos é especialmente importante porque esses dados são utilizados posteriormente durante repetidas aplicações da técnica, em modo *off-line*, mediante a variação do parâmetro  $RP_{2D}$  do algoritmo e comparação dos resultados.

O procedimento de captura é ilustrado no diagrama da Figura 7.5.

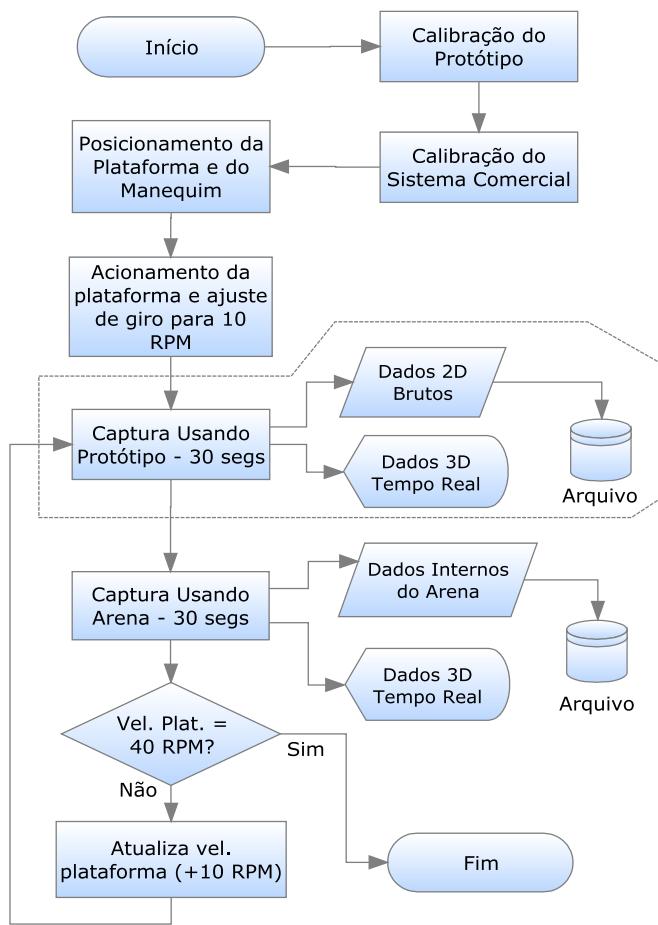


Figura 7.5 - Procedimento geral de captura. Inicialmente os sistemas foram calibrados e em seguida iniciou-se as sessões de captura utilizando os dois softwares. Coletas foram conduzidas com a plataforma girando a 10, 20, 30 e 40 RPM.

## 7.5 Estudo de Desempenho do Método Mediante Variação do Parâmetro $RP_{2D}$

Para melhor avaliar o desempenho do protótipo e ainda testar a sensibilidade do algoritmo com relação ao tamanho da região de procura bidimensional,  $RP_{2D}$ , a técnica de rastreamento e reconstrução tridimensional foi aplicada um total de 50

<sup>3</sup> As coordenadas de todos os pontos nas imagens, para cada câmera, são armazenadas pelo protótipo em arquivo binário, utilizando um formato de dados especialmente desenvolvido para a situação.

vezes para cada arquivo de dados gravado durante a captura, com  $RP_{2D}$  variando de 0,25 a 12,5 pixels (com incrementos de 0,25). O processo é ilustrado na Figura 7.6.

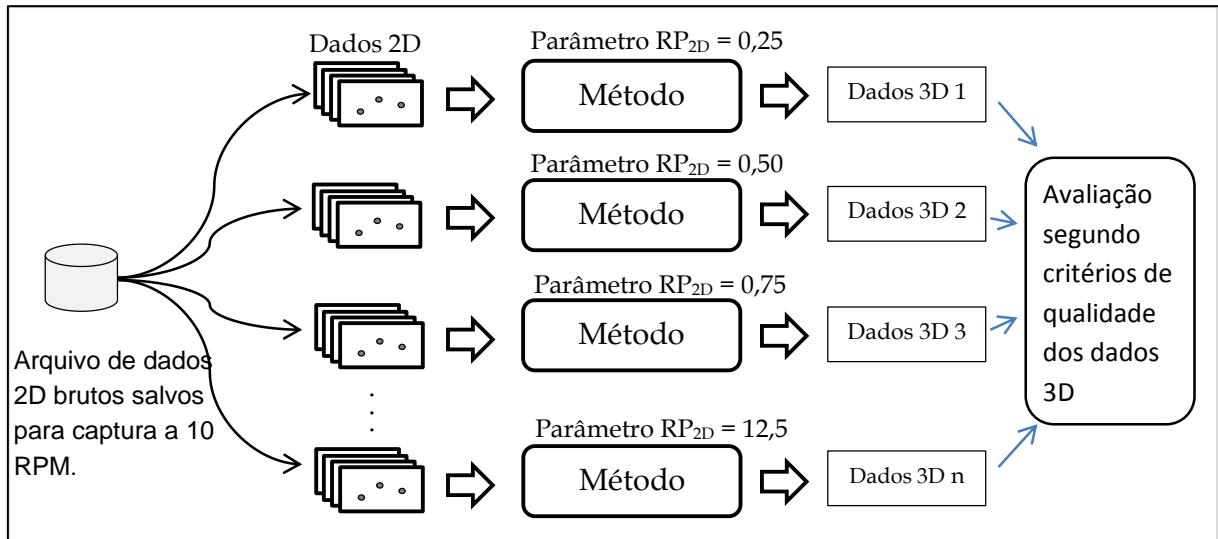


Figura 7.6 - Ilustração do processo utilizado para aplicação do método com variação do tamanho da região de procura bidimensional,  $RP_{2D}$ . A figura ilustra o processo considerando apenas um dos arquivos de dados (para uma velocidade da plataforma).

Repare que cada aplicação do algoritmo pode resultar em uma saída 3D diferente. A seguir, os resultados são avaliados e comparados segundo cada critério de análise dos dados 3D apresentado anteriormente.

### *Número de Marcadores Rastreados por Toda Sequência*

O gráfico (a) da Figura 7.7 ilustra a variação do número de marcadores rastreados por toda a sequência à medida que o parâmetro  $RP_{2D}$  é incrementado. O resultado é fornecido para cada sessão de captura.

Analizando o gráfico percebe-se que, para  $RP_{2D}$  entre aproximadamente 2,25 e 7,25 pixels, o algoritmo foi capaz de detectar os 38 marcadores no início da sequência e rastreá-los com sucesso até o término, para todas as sessões de captura. Quando a região de procura é demasiadamente pequena, a capacidade de rastreamento é debilitada drasticamente. Por outro lado, quando essa região aumenta muito, percebe-se um declínio gradual do número de marcadores totalmente rastreados.

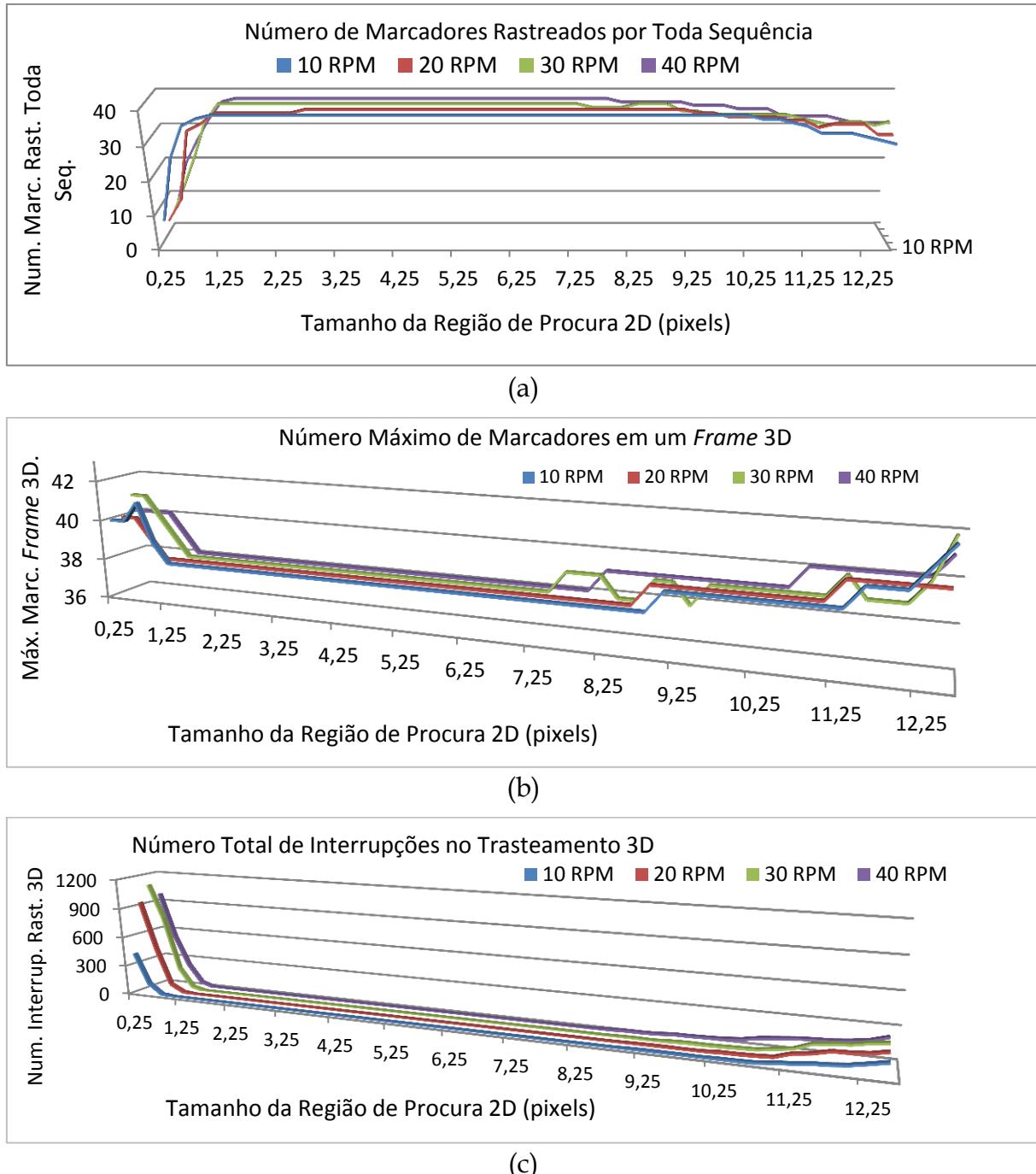


Figura 7.7 – (a) Número de marcadores identificados no início da sequência e rastreados até o final em função do tamanho da região de procura no rastreamento 2D. Cada linha no gráfico corresponde à coleta realizada em uma dada velocidade de rotação (10, 20, 30 ou 40 RPM) da plataforma. (b) Variação do número máximo de marcadores detectados em um frame 3D à medida que é aumentado o tamanho da região de busca no rastreamento 2D. O número máximo de marcadores detectados em um frame 3D está relacionado com a quantidade de falsos marcadores. (c) Número total de interrupções no rastreamento 3D em função do tamanho da região de busca no rastreamento 2D.

### Número Máximo de Marcadores em um Frame 3D

O gráfico (b) da Figura 7.7 ilustra a variação do número máximo de marcadores detectados em um frame 3D em função do tamanho da região de procura 2D. A

existência de marcadores-fantasmas é evidente para  $RP_{2D}$  inferior a 1,25 pixels. Para valores acima de 7,25 pixels, a detecção de falsos marcadores já começa a ocorrer para a captura realizada a 30 RPM. Em todos os casos, para valores de  $RP_{2D}$  acima de 9,25 pixels, é garantida a detecção de pelo menos um falso marcador.

### Número de Interrupções no Rastreamento 3D

O gráfico da (c) da Figura 7.7 exibe o número total de interrupções no rastreamento 3D em função do tamanho da região de procura 2D. Verifica-se um alto número de interrupções para valores muito pequenos desse parâmetro, independente da velocidade do manequim. Para valores de  $RP_{2D}$  entre 1,5 e 8,25, o número de interrupções se mantém próximo de zero e para valores do parâmetro acima de 8,25, o número de interrupções aumenta gradualmente.

### Número Médio de Spikes por Trajetória

O gráfico da Figura 7.8 apresenta o número médio de *spikes* detectados por trajetória inteiramente reconstruída como função do parâmetro  $RP_{2D}$ . Para todas as sessões conduzidas, a faixa de valores de  $RP_{2D}$  que resultou em trajetórias mais limpas, com relação à presença de *spikes*, foi de 1,25 a aproximadamente 2,5 pixels. À medida que  $RP_{2D}$  se torna menor que um pixel, a quantidade de *spikes* aumenta intensamente para as duas coletas mais lentas. Para valores de  $RP_{2D}$  acima de 11 pixels, percebe-se um aumento considerável desse ruído para as velocidades de 20, 30 e 40 RPM. O gráfico deixa claro que o número de *spikes* detectados foi maior, no geral, para a coleta realizada na velocidade mais baixa e menor para a coleta conduzida na velocidade mais alta (40 RPM).

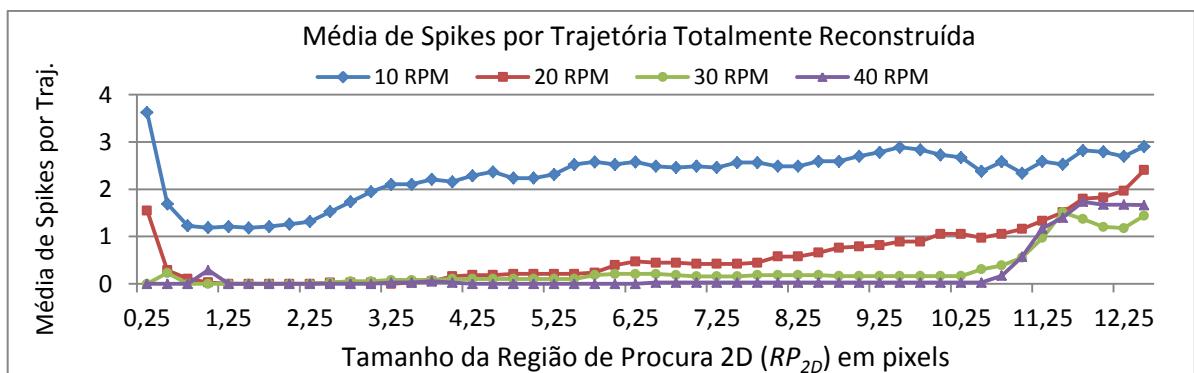


Figura 7.8 - Número médio de ruídos do tipo *spike* detectados por trajetória inteiramente reconstruída em função do tamanho da região de busca no rastreamento 2D.

### Coeficiente Geral de Ruído

O gráfico da Figura 7.9 exibe as variações obtidas para o coeficiente geral de ruído conforme se aumenta o tamanho da região de procura 2D. Os maiores valores do coeficiente foram obtidos para  $RP_{2D}$  inferior a um pixel. À medida que  $RP_{2D}$  aumenta a partir desse limiar, os valores obtidos para o coeficiente aumentam gradativamente.

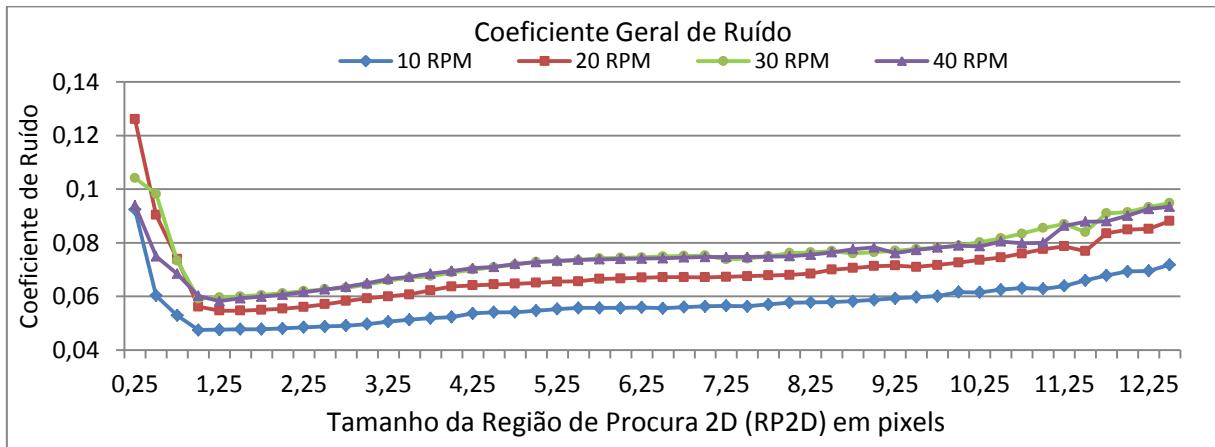


Figura 7.9 - Coeficiente geral de ruído em função do tamanho da região de procura 2D,  $RP_{2D}$ .

#### 7.5.1 Discussão

Conforme apresentado nos gráficos da seção anterior, o tamanho da região de procura no rastreamento 2D exerce influência expressiva nos resultados gerados pela técnica de reconstrução. O gráfico (a) da Figura 7.7 mostrou que o número de marcadores rastreados por toda sequência diminui quando  $RP_{2D}$  se torna inferior a um limiar  $a$  ou superior a um limiar  $b$ . Isto acontece porque quanto menor o valor de  $RP_{2D}$ , menor a probabilidade de o marcador procurado ser localizado dentro da região estimada, o que aumenta as chances de insucesso no rastreamento. Por outro lado, quanto maior o valor de  $RP_{2D}$ , maior é a probabilidade de serem encontrados dois ou mais marcadores dentro da região de busca, o que também pode levar ao insucesso no rastreamento. Assim, existe uma faixa de valores entre dois limites  $a$  e  $b$ , que maximiza o número de marcadores rastreados. As variações do número total de interrupções no rastreamento, exibidas na Figura 7.7(c), podem ser explicadas pelo mesmo motivo.

As chances de detecção de falsos marcadores aumentam à medida que  $RP_{2D}$  ultrapassa o limiar  $b$ , como mostra o gráfico (b) da Figura 7.7. Quando  $RP_{2D}$  aumenta

consideravelmente, há uma maior probabilidade de que seja encontrado um marcador 2D na região de procura que não corresponde ao marcador sendo rastreado. Mesmo assim, o algoritmo de rastreamento pode marcá-lo como uma continuidade do anterior. Esse rastreamento 2D errôneo, dependendo de outros fatores como o posicionamento dos outros marcadores e o número de câmeras que o visualiza, pode fazer com que o marcador 3D correspondente seja deslocado de sua trajetória real e se torne um falso marcador. Também é possível que o rastreamento errôneo leve à ocorrência de *spikes* (reveja a Figura 7.8). Neste caso, o marcador 2D falsamente identificado é desassociado do marcador 3D alguns *frames* depois, o que leva à correção de sua trajetória.

Em geral, o parâmetro  $RP_{2D}$  deve ser suficientemente pequeno para manter baixo o nível de ruído, mas suficientemente grande para não degradar muito a capacidade de rastreamento. O gráfico da Figura 7.10 a seguir foi elaborado para possibilitar uma análise considerando todos os critérios de avaliação em conjunto. Para cada critério, calculou-se a média das quatro sequências de valores correspondentes às quatro velocidades de captura para obtenção de uma sequência média representativa. Em seguida, cada sequência foi normalizada. Em especial, a sequência correspondente à máxima quantidade de marcadores detectados em um *frame* 3D foi adaptada realizando-se a subtração do valor 38, o que resultou em um número diretamente relacionado à quantidade evidente de falsos marcadores. Analisando o gráfico é possível concluir que os melhores resultados obtidos com relação aos cinco critérios foram resultantes de valores de  $RP_{2D}$  no intervalo de 1,25 a 2,25 pixels. Esta faixa é representada por um retângulo translúcido. Note que abaixo de 1,25, percebe-se uma redução imediata da quantidade de marcadores automaticamente rastreados e um aumento dos níveis de ruídos. Acima de 2,25 pixels, por outro lado, torna-se evidente o aumento de *spikes* nas trajetórias.

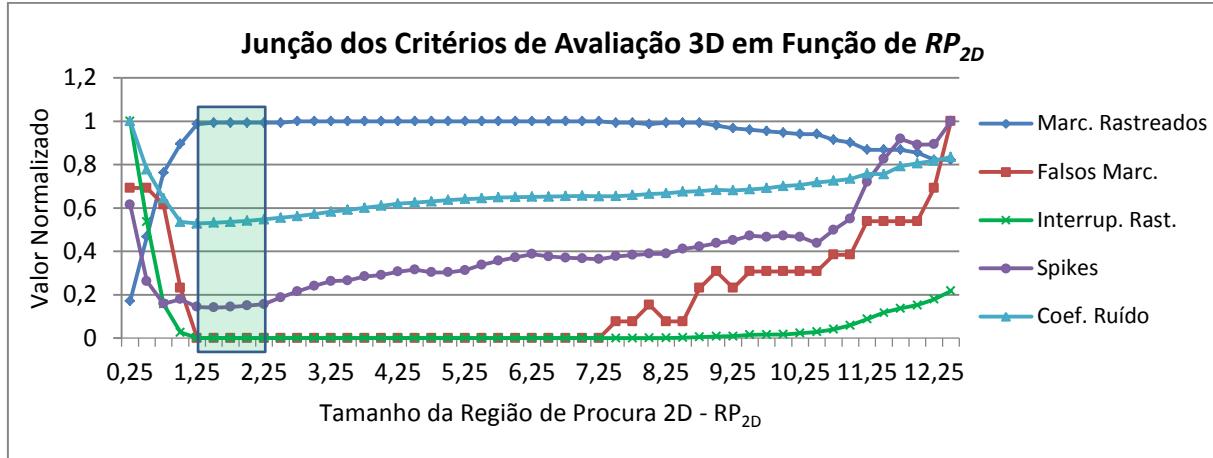


Figura 7.10 - Valores normalizados de todos os critérios de avaliação dos dados 3D em função do tamanho da região de procura 2D. A região representada por um retângulo corresponde à faixa de valores de RP<sub>2D</sub> que possibilitou a geração dos melhores resultados com respeito aos cinco critérios.

## 7.6 Comparativo: Protótipo baseado na Técnica e Software Arena

Para comparar os resultados do protótipo com os resultados do software Arena, o parâmetro RP<sub>2D</sub> foi ajustado para 2,25 pixels, com base nos testes apresentados anteriormente.

### Uso do Software Arena

Empregou-se nos experimentos a versão 1.6 (de 2009) do software Arena, disponível no laboratório onde a pesquisa foi conduzida. Todos os parâmetros relacionados à captura e reconstrução disponibilizados pelo software foram mantidos em seus valores-padrão. Nenhum desses parâmetros possui equivalência direta ao parâmetro RP<sub>2D</sub>.

Para que as trajetórias dos marcadores pudessem ser geradas pelo Arena, foi necessária a calibração de um modelo esquelético (*Skeleton*) no programa visando corresponder às medidas do manequim. Neste momento, o manequim foi posto de braços abertos, na *posição-T* (*T-Pose*).

O software Arena foi configurado para gravar 30 segundos de movimento para cada velocidade da plataforma (10, 20, 30 e 40 RPM), conforme já apresentado na Figura 7.5. Após execução das quatro sessões de captura, cada arquivo registrado pelo software foi aberto para realização de um procedimento denominado no programa

de *trajetorização*. Este procedimento é necessário para que a trajetória de cada marcador seja calculada e é um pré-requisito para que os dados 3D possam ser exportados e usados externamente. No momento da trajetorização, o software solicita o trecho da sequência a considerar e exige que o usuário informe o *frame inicial* e o *frame final*.

Para todas as sessões gravadas, escolheu-se o *frame* inicial como sendo aquele em que o manequim se encontra alinhado pela primeira vez com o eixo Z do sistema de coordenadas de referência (definido no momento da calibração). O *frame* final foi escolhido somando-se 2499 (25 segundos) ao *frame* inicial (um processo análogo também foi feito para o protótipo desenvolvido, visando manter as capturas em sincronia). Dessa forma, apenas 25 dos 30 segundos de cada coleta foram efetivamente processados.

Após executar o cálculo das trajetórias para o arquivo de cada captura, os dados foram exportados para um formato possível de ser interpretado e avaliado externamente: o C3D (*Coordinate 3D*). O C3D é um formato de arquivo binário de domínio público largamente utilizado pela comunidade de biomecânica e animação gráfica para armazenar dados 3D sincronizados.

Os dados tridimensionais gerados pelos dois softwares foram avaliados e comparados, segundo os diversos critérios definidos anteriormente, por um conjunto de funções implementadas no protótipo. Assim, os arquivos exportados pelo Arena contendo as trajetórias reconstruídas foram importados e avaliados pelo protótipo.

### ***Resultados Comparativos***

Com relação à *quantidade de marcadores rastreados do início ao fim*, ambos os softwares foram capazes de rastrear todos os 38 marcadores para todas as sessões de captura (10, 20, 30 e 40 RPM).

Com relação ao número máximo de marcadores encontrados em um *frame* 3D, os resultados são apresentados no gráfico da Figura 7.11. Em nenhum momento durante as quatro sessões de captura o protótipo detectou mais de 38 marcadores. Já

o número máximo detectado pelo Arena foi de 39 ou 40, indicando que pelo menos um (ou, respectivamente, dois) desses marcadores reconstruídos constitui-se um marcador-fantasma.

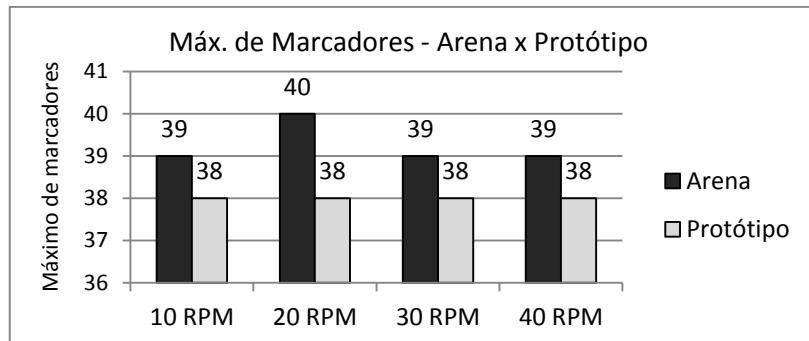


Figura 7.11 - Número máximo de marcadores encontrados em um mesmo *frame* 3D. O protótipo não registrou mais de 38 marcadores em nenhum momento dos testes. Por outro lado, o software Arena registrou mais marcadores que o real: 39 ou 40.

O gráfico da Figura 7.12 ilustra o número médio de *spikes* detectados por trajetória utilizando os dois softwares. Em todos os testes, um número médio menor de *spikes* foi detectado nos dados gerados pelo protótipo.

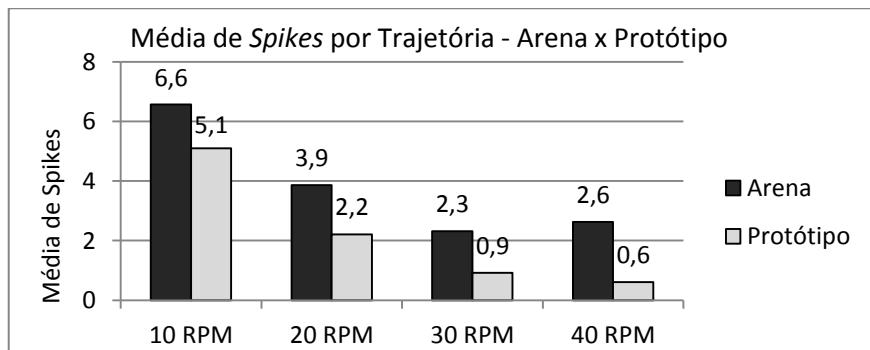


Figura 7.12 - Número médio de *spikes* detectados por trajetória reconstruída. O número foi calculado para cada um dos testes, com a plataforma girando a 10, 20, 30 ou 40 RPM.

Com relação ao critério de avaliação *número de interrupções no rastreamento 3D*, ambos os softwares não registraram interrupções no rastreamento dos 38 marcadores principais. Entretanto, como o software Arena registrou falsos marcadores, algumas interrupções no rastreamento (0, 10, 7 e 9 respectivamente para as velocidades de 10, 20, 30 e 40 RPM) foram registradas para tais marcadores, uma vez que o software naturalmente não manteve esses marcadores ao longo de muitos *frames*. Nos arquivos C3D gerados pelo Arena, tais marcadores foram identificados com o rótulo *unlabeled*.

Os gráficos apresentados na Figura 7.13 ilustram um trecho de aproximadamente 5 segundos da trajetória 3D de um dos marcadores colocados no tronco do manequim. A trajetória é apresentada em termos das coordenadas X, Y e Z do marcador, calculadas usando o software Arena e também usando o protótipo desenvolvido. Comparando os gráficos é possível perceber a maior incidência de ruídos nos dados gerados pelo software Arena.

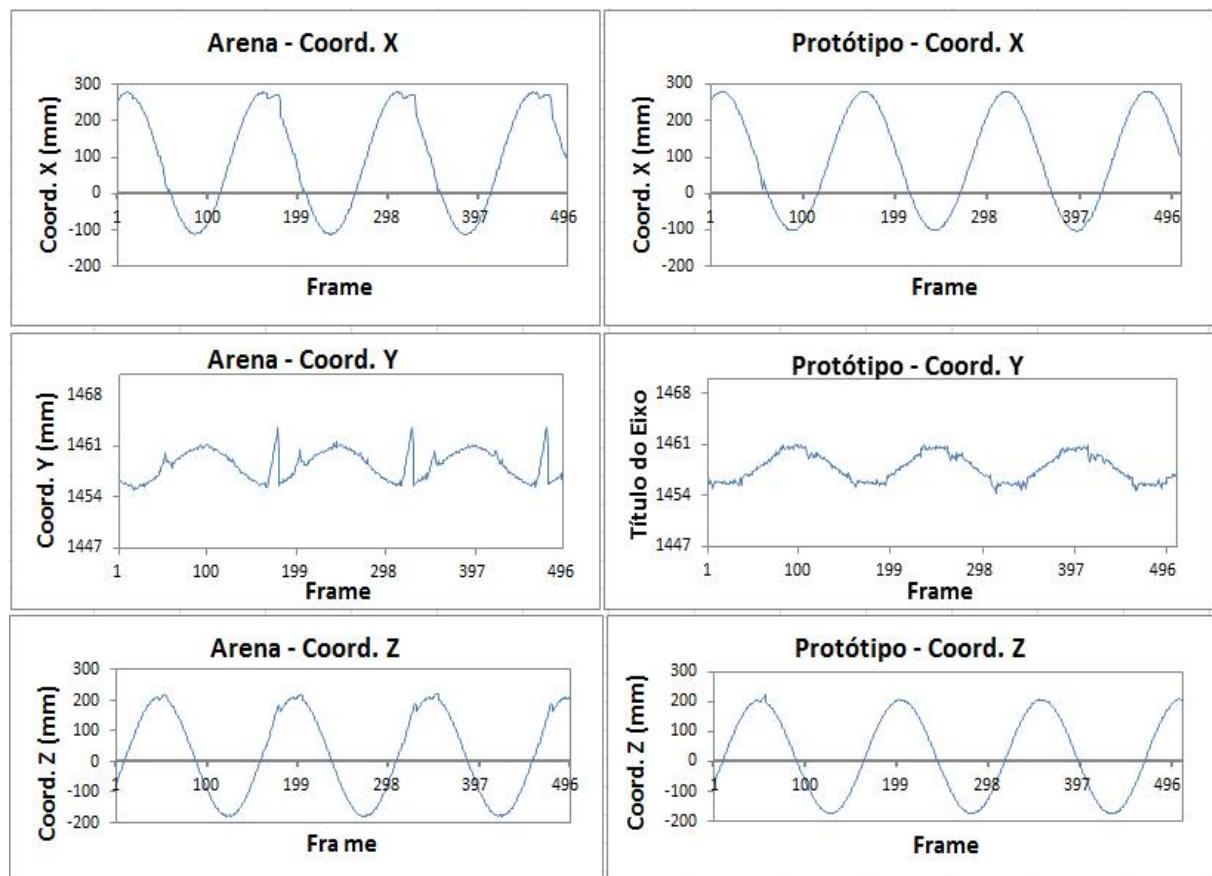


Figura 7.13 - Coordenadas X, Y e Z de um trecho de aproximadamente 5 segundos da trajetória de um dos marcadores colocados no tronco do manequim. Os três gráficos da esquerda apresentam as variações das coordenadas obtidas com o software Arena. Os três gráficos da direita apresentam os valores equivalentes obtidos com o protótipo desenvolvido. Nenhum tipo de filtro foi aplicado.

O gráfico da Figura 7.14 exibe o coeficiente geral de ruído calculado para os dados 3D gerados pelos dois programas. Para todas as coletas, o coeficiente de ruído associado ao protótipo é pelo menos 13 vezes menor do que aquele obtido para o software Arena.

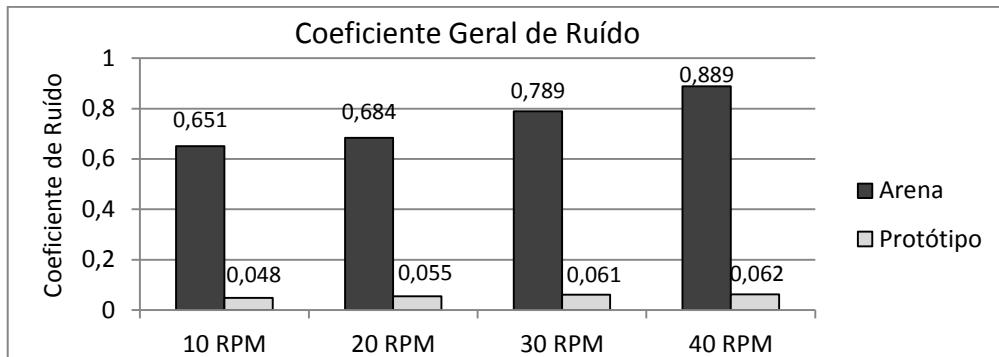


Figura 7.14 - Coeficiente geral de ruído calculado para os dados 3D gerados pelos dois softwares. Para todas as coletas, o coeficiente de ruído associado aos resultados produzidos pelo protótipo é pelo menos 13 vezes menor do que aquele obtido para os dados gerados pelo Arena.

### Discussão

Há dois fatores importantes que devem ser levados em consideração ao comparar os resultados gerados pelos dois softwares. O primeiro deles é o fato de que o software Arena se apoia em um modelo antropométrico, o que viabiliza a criação de vários tipos de restrições que podem ser usadas para otimizar o processo de rastreamento e reconstrução dos marcadores. No caso do protótipo desenvolvido, os marcadores são rastreados e reconstruídos como uma nuvem de pontos independentes e nenhum modelo biomecânico é acoplado ao processo. O segundo fator a ser considerado é que o software Arena necessita executar uma etapa de pós-processamento para o cálculo das trajetórias dos marcadores (etapa da *trajetorização*). Nesse momento, algoritmos de rastreamento mais sofisticados podem ser introduzidos para corrigir e/ou aprimorar as trajetórias, pois todos os dados brutos obtidos das câmeras estão disponíveis e o processamento não precisa acontecer em tempo real. No caso do protótipo, os dados 3D não passaram por nenhum tipo de pós-processamento e são exatamente aqueles calculados pelo algoritmo em tempo real.

Um fato importante de ser observado é que, mesmo em condições teoricamente desfavoráveis, por não usar um modelo e não realizar a etapa de trajetorização, o protótipo foi capaz de rastrear e reconstruir todos os marcadores, gerou trajetórias com menor incidência de ruídos e ainda não detectou nenhum falso marcador, conforme apresentado nos gráficos da Figura 7.11, Figura 7.12 e Figura 7.14. Para a captura realizada com a plataforma girando a 30 RPM, o número médio de *spikes* gerados pelo Arena corresponde a mais que o dobro do número encontrado nos

resultados do protótipo. Para a coleta realizada a 40 RPM, a diferença subiu para quatro vezes e pôde ser percebida até mesmo visualmente durante a reprodução das sequências 3D reconstruídas. Com relação aos ruídos de menor amplitude, a diferença foi ainda maior, conforme ilustrado no gráfico da Figura 7.14.

## 7.7 Conclusão e Considerações do Capítulo

---

Neste capítulo foram apresentados os experimentos conduzidos utilizando um manequim fixado sobre uma plataforma giratória. No manequim foram postos 38 marcadores e capturas foram realizadas com o conjunto girando a 10, 20, 30 e 40 RPM. O experimento teve como principais objetivos:

- ❖ Verificar o desempenho do método quando da reconstrução de dezenas de marcadores postos em uma estrutura anatômica humana e sob influência de oclusões frequentes e sobreposições dos pontos nas imagens;
- ❖ Verificar a sensibilidade da técnica com relação ao tamanho da região de procura do rastreamento bidimensional,  $RP_{2D}$ ;
- ❖ Relacionar diretamente os resultados do método com os resultados de um sistema comercial utilizado mundialmente.

É certo que obter um valor ideal para o parâmetro  $RP_{2D}$  não é um processo trivial, pois a escolha ótima pode depender de diversos fatores, como a qualidade da calibração, a quantidade, resolução e velocidade das câmeras, o número de marcadores e o tipo de movimento a ser capturado. Entretanto, valores próximos de 2,25 pixels possibilitaram resultados satisfatórios não somente para as capturas apresentadas neste capítulo, mas também para vários outros tipos de movimentos testados, incluindo o movimento de corrida em esteira, o qual é apresentado no capítulo seguinte.

O método foi capaz de rastrear completamente todos os 38 marcadores do manequim, assim como o software comercial Arena da empresa NaturalPoint. Todavia, as trajetórias reconstruídas pelo protótipo mostraram-se mais estáveis com relação à presença de ruídos. Adicionalmente, nenhum marcador fantasma foi detectado em nenhum ponto das sequências reconstruídas pelo protótipo, ao contrário daquelas obtidas pelo software Arena.

# Capítulo 8

## Experimentos Utilizando Esteira Ergométrica

Todos os testes apresentados anteriormente foram conduzidos sob circunstâncias controladas, com o objetivo de possibilitar um estudo e avaliação da eficácia da técnica de rastreamento e reconstrução. Usar vários marcadores em um manequim girando simula a ocorrência de sobreposições dos pontos nas imagens e oclusões dos marcadores, situações que acontecem frequentemente em diversas aplicações de captura de movimento. Entretanto, alguns fatores presentes em movimentos corporais reais, como acelerações e desacelerações dos membros do corpo em conjunto com imprevisibilidade, não puderam ser testados. Considerando esses elementos e buscando uma aplicação do mundo real, decidiu-se testar o protótipo na captura de movimentos corporais de um atleta correndo em esteira ergométrica. Para comparar os resultados, utilizou-se novamente o software comercial Arena. Este capítulo apresenta em detalhes como o experimento foi conduzido e todos os seus resultados.

O capítulo está organizado da seguinte forma. A Seção 8.1 apresenta o procedimento de captura de dados. A Seção 8.2 exibe uma amostra dos dados tridimensionais reconstruídos pelo protótipo. A Seção 8.3 descreve o procedimento adotado para avaliação e comparação dos resultados. A Seção 8.4 apresenta os resultados obtidos, a Seção 8.5 discute esses resultados e a Seção 8.6 encerra o capítulo com as considerações finais e as conclusões.

## 8.1 Procedimento de Captura

---

Uma esteira ergométrica elétrica e dobrável foi utilizada para possibilitar a captura do movimento de corrida. Antes de iniciar a coleta, a esteira foi posta no centro da região de captura e seu painel de controle juntamente com as barras de sustentação foram rebaixados até o chão. Esse procedimento foi adotado para que essas partes da esteira não interferissem no registro das imagens.

O padrão de 38 marcadores aceito pelo software Arena foi novamente empregado. O indivíduo vestiu a roupa apropriada e os marcadores foram fixados em suas devidas posições, conforme configuração ilustrada na Figura 7.2.

Uma única sessão de captura foi conduzida, que teve duração de aproximadamente 7 minutos. Durante esse tempo, o sujeito se manteve correndo na esteira em ritmo natural e constante a uma velocidade de 8 km/h. Inicialmente, o protótipo desenvolvido foi utilizado para registrar os três primeiros minutos da atividade. Em seguida, o software Arena foi empregado para gravação de mais três minutos. As câmeras foram calibradas separadamente em ambos os softwares momentos antes do início da coleta e os instantes de aceleração e desaceleração da esteira não foram inclusos nas gravações.

O participante foi instruído a realizar a atividade de corrida de maneira natural, mas procurando manter o mesmo padrão de movimentação dos membros do início ao fim. Em nenhum momento, ao longo dos 6 minutos de captura, o participante realizou qualquer tipo de ação não vinculada ao ato de correr em si, como a inclinação ou giro da cabeça ou a elevação dos braços, que poderia interferir diretamente no registro das imagens.

## 8.2 Amostra dos Dados 3D Reconstruídos pelo Protótipo

---

Para rastrear e reconstruir os movimentos do sujeito, o parâmetro  $RP_{2D}$  do algoritmo foi ajustado para 2,5 pixels. A Figura 8.1 apresenta algumas imagens colhidas do protótipo no momento da visualização 3D dos marcadores reconstruídos. As imagens (a), (b) e (c) correspondem a três visões diferentes do conjunto de

marcadores. A imagem (d) exibe o traçado 3D de aproximadamente quatro segundos da trajetória percorrida pelo marcador colocado sobre o tornozelo direito do participante.

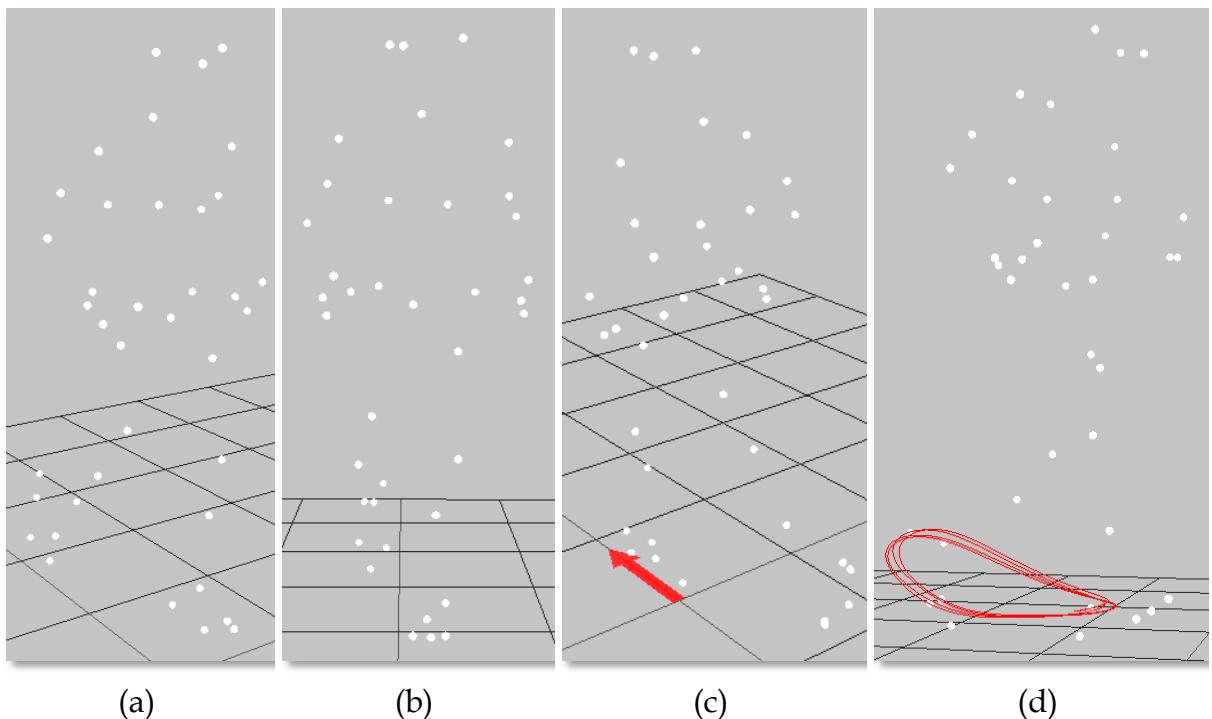


Figura 8.1 – Amostras capturadas da tela do protótipo apóss captura e reconstrução dos movimentos do indivíduo na esteira. Um total de 38 marcadores foi utilizado. As imagens (a), (b) e (c) correspondem, respectivamente, às visões espaciais dos marcadores anterolateral, frontal e póstero-lateral. A imagem (d) exibe o traçado espacial do marcador posto sobre o tornozelo direito do participante, registrado durante aproximadamente 4 segundos.

O protótipo foi capaz de rastrear automaticamente 32 dos 38 marcadores do início ao fim dos 3 minutos de captura. Três dos outros seis marcadores tiveram o rastreamento interrompido momentaneamente devido à oclusão causada pelos próprios membros do corpo e outros três marcadores não foram rastreados por toda a sequência devido a ambiguidades no rastreamento 2D ocasionadas por sobreposição dos pontos nas imagens, o que é tido como uma possibilidade intrínseca da técnica proposta.

### 8.3 Procedimento de Avaliação e Comparaçāo dos Resultados

Conforme já mencionado anteriormente, o participante correu a uma velocidade constante, manteve o mesmo padrão de movimentação dos membros e não executou nenhuma ação desvinculada do ato de correr. Embora esses cuidados tenham sido

tomados para tornar a captura o mais uniforme possível ao longo dos 6 minutos, aplicar ambos os softwares à suas respectivas coletas e relacionar os resultados seria um procedimento vulnerável até mesmo às pequenas variações que poderiam ocorrer no padrão de execução da atividade. Por exemplo, uma elevação adicional de um dos braços do participante, realizada em algum ponto da sessão de captura, poderia ser suficiente para tornar oculto um ou mais marcadores e degradar sensivelmente a qualidade geral dos resultados. Além disso, a disposição dos pontos 2D nos *frames* iniciais da captura pode exercer influência significativa nos primeiros instantes da reconstrução e dificilmente os *frames* iniciais para as duas sessões teriam equivalência, ainda que o descarte de alguns *frames* fosse realizado manualmente.

Visando minimizar esses problemas e tornar a comparação mais robusta, decidiu-se realizar um processo de *janelamento* dos dados para aplicação da técnica. No caso do protótipo, os dados 2D brutos salvos durante a captura foram utilizados para a composição de 20 janelas de 1000 *framesets* cada, o que equivale a 20 trechos de 10 segundos. O ponto de início de cada trecho foi obtido aleatoriamente dentro dos primeiros 17000 *framesets* de um total de 18000 registrados (3 minutos ou 180 segundos, a 100 *frames* por segundo).

Após a geração das janelas de dados 2D, a técnica de rastreamento e reconstrução foi então aplicada para cada um desses trechos e os resultados obtidos foram analisados estatisticamente.

Para o software Arena, o *janelamento* foi realizado de acordo com os recursos disponíveis no programa. Durante a sessão de captura, o software foi empregado para registrar 3 minutos de atividade do sujeito, o que resultou em um único arquivo de dados. Esse arquivo foi posteriormente aberto no próprio programa e o procedimento de *trajetorização*<sup>4</sup> foi realizado um total de 20 vezes visando a obtenção de 20 trechos de dados tridimensionais para fins de avaliação e comparação com o protótipo.

---

<sup>4</sup> Conforme descrito na seção 7.6, o procedimento de *trajetorização* possibilita que as trajetórias dos marcadores sejam calculadas, além de ser uma etapa necessária para a exportação dos dados.

## 8.4 Resultados Comparativos: Arena x Protótipo

Os dados tridimensionais gerados por cada um dos softwares para cada uma das janelas foram avaliados e comparados segundo os cinco critérios apresentados na Seção 7.3. As medidas obtidas para os 20 segmentos de dados foram utilizadas para o cálculo de um valor *médio* representativo. A dispersão dos valores foi estimada pelo cálculo do desvio padrão. Para uma melhor avaliação dos resultados, executou-se também o teste de análise de variância (ANOVA). Os valores obtidos são apresentados na Tabela 8.1.

Tabela 8.1 – Características dos dados tridimensionais gerados pelo protótipo e pelo software comercial Arena quando da aplicação dos mesmos nos 20 segmentos de dados. Para cada critério de avaliação definido na Seção 7.3, são apresentados os valores médios obtidos, os desvios padrão e o resultado do teste de ANOVA.

	Valor Médio		Desvio Padrão		ANOVA
	Arena	Protótipo	Arena	Protótipo	
Num. Marcadores Rastreados por Toda Janela	28,75	33,40	8,98	1,05	p = 0,031
Num. Máximo de Marcadores em um Frame 3D da Janela	42,30	38,20	2,22	0,30	p < 0,001
Num. de Interrupções no Rastreamento	21,00	8,55	13,81	6,57	p < 0,002
Num. de Spikes por Trajetória	0,55	0,10	0,36	0,07	p < 0,001
Coeficiente Geral de Ruído	1,86	0,16	0,27	0,01	p < 0,001

Para uma melhor avaliação e comparação dessas medidas com relação à concentração/dispersão dos valores, gráficos do tipo *boxplot* são apresentados na Figura 8.2. Na média, o software Arena conseguiu rastrear 28,75 dos 38 marcadores ao longo dos segmentos, enquanto o protótipo rastreou 33,4. No pior dos casos, o Arena rastreou apenas 5 marcadores e o protótipo, 31. A variação dos valores pode ser percebida no gráfico (a) da Figura 8.2.

O gráfico (b) da Figura 8.2 ilustra os valores obtidos com relação ao número máximo de marcadores identificados dentro dos segmentos. Na média, o software Arena

identificou um máximo de 42,3 marcadores (4,3 a mais que a quantidade real), enquanto o protótipo identificou 38,2. O gráfico mostra claramente a maior dispersão dos valores relativos ao Arena, além de sugerir a presença de uma quantidade maior de falsos marcadores nos dados gerados por esse software.

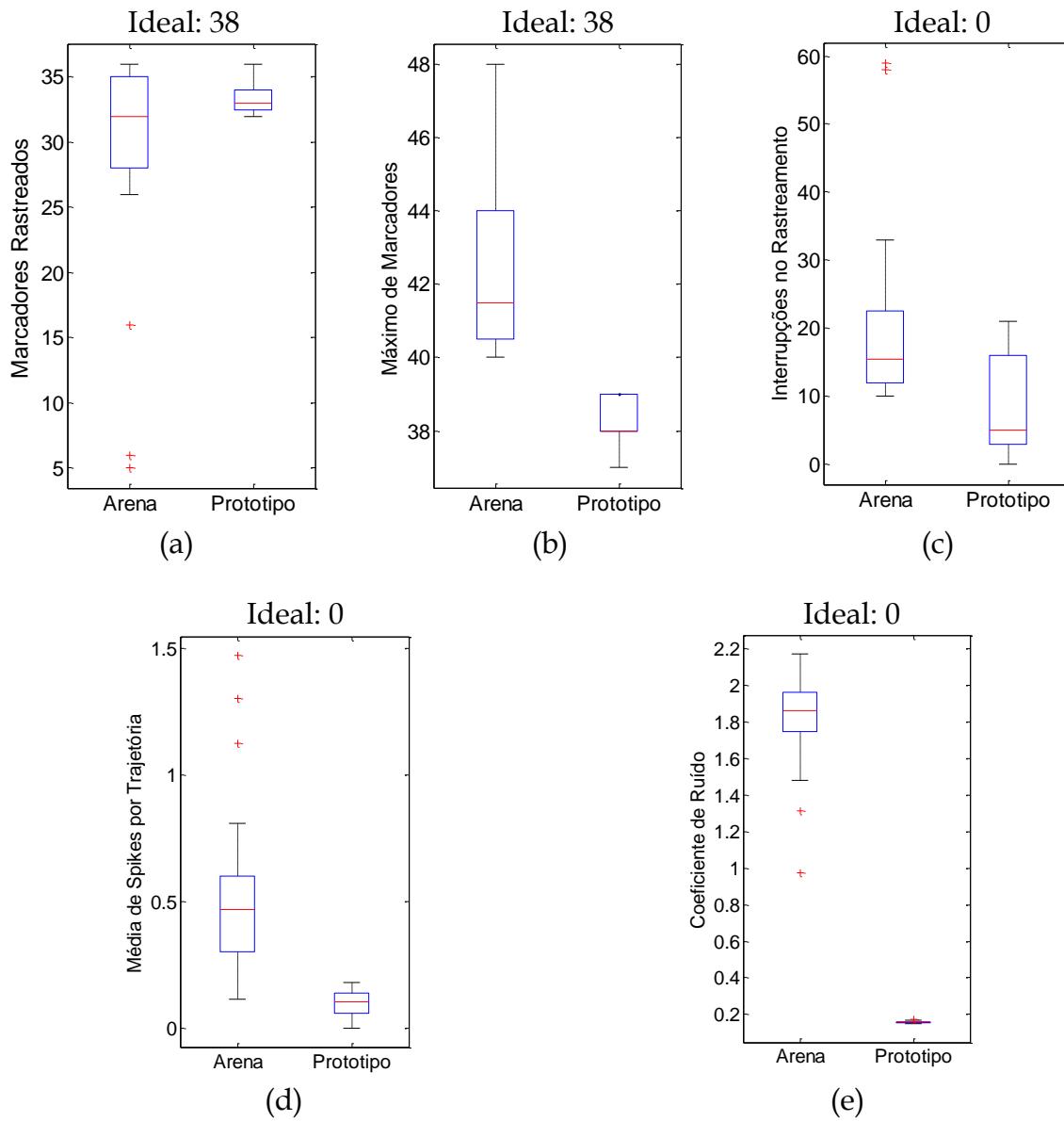


Figura 8.2 - Características dos dados tridimensionais gerados pelo protótipo e pelo software comercial Arena quando da aplicação dos mesmos nos 20 segmentos de dados. Cada gráfico do tipo *boxplot* possibilita uma avaliação e comparação dessas medidas com relação à concentração/dispersão dos valores obtidos. A linha central dentro de cada caixa indica a mediana.

O gráfico (c) da Figura 8.2 exibe os valores obtidos com respeito ao total de interrupções detectadas no rastreamento dos marcadores. Nos dados gerados pelo Arena foram detectadas, em média, 21 interrupções no rastreamento dos marcadores. No caso do protótipo, o valor médio foi de 8,55 interrupções por janela.

Os gráficos (d) e (e) da Figura 8.2 apresentam os resultados obtidos com relação ao número médio de *spikes* detectados por trajetória completa e o coeficiente geral de ruído, respectivamente. A média dos números médios de *spikes* detectados foi de 0,55 para o Arena e 0,10 para o protótipo. O gráfico mostra uma maior dispersão dos valores relativos ao Arena. Com relação ao coeficiente geral de ruído, as médias obtidas foram de 1,86 para o Arena e 0,16 para o protótipo. Neste caso, além de alcançar um coeficiente de ruído médio cerca de 10 vezes menor, os valores obtidos para o protótipo estão claramente concentrados em 0,16, o que não acontece para o software comercial Arena.

## 8.5 Discussão

---

No geral, os gráficos exibidos na seção anterior colocam o protótipo desenvolvido em posição vantajosa, comparado ao software comercial Arena. Com respeito à quantidade média de marcadores rastreados do início ao fim das janelas, por exemplo, a vantagem de quatro marcadores alcançada pelo protótipo (33,4 contra 28,75) é especialmente importante considerando que o software Arena utilizou-se de uma etapa de pós-processamento e de um modelo biomecânico para concluir o cálculo das trajetórias. Como já mencionado, se as outras características forem mantidas, quanto maior o número de marcadores rastreados automaticamente, menor será a necessidade de intervenção e edição dos dados por parte do usuário.

A diferença na quantidade máxima de marcadores detectados, o que está diretamente relacionado à presença de falsos-marcadores, pôde ser observada antes mesmo do processamento dos dados. Ao visualizar as sequências reconstruídas em ambos os softwares, percebe-se um maior número de marcadores-fantasmas nos dados gerados pelo software Arena. Este fato pode ser visualizado pelo gráfico (b) da Figura 8.2.

As diferenças apresentadas nos gráficos (d) e (e) da Figura 8.2 com relação à presença de ruídos também puderam ser percebidas visualmente ao observar os dados tridimensionais em ambos os softwares. A diferença é notável principalmente para os

ruídos de menor amplitude (*não-spikes*), que podem ser identificados pela vibração perceptível de vários marcadores ao longo de suas trajetórias.

## 8.6 Conclusão e Considerações do Capítulo

---

O experimento apresentado neste capítulo foi realizado com o objetivo de avaliar a técnica proposta para reconstruir as trajetórias espaciais de um conjunto de marcadores em uma aplicação real. O protótipo desenvolvido foi empregado para capturar os movimentos de um indivíduo correndo em esteira ergométrica durante 3 minutos. Para melhor avaliar os resultados, utilizou-se como referência o software comercial Arena em conjunto com um procedimento de *janelamento* dos dados.

Considerando os critérios de avaliação definidos no Capítulo 7, o protótipo desenvolvido com base na técnica proposta produziu resultados satisfatórios, conseguindo reconstruir, em tempo real, as trajetórias da maior parte dos marcadores (em média, 33,4 dos 38 marcadores). É importante ressaltar que a atividade alvo de captura oferece oportunidades para a ocorrência de eventos que dificultam o rastreamento e intensificam os ruídos, como oclusões, sobreposições dos pontos nas imagens, acelerações e desacelerações dos marcadores.

Comparado ao software comercial Arena, o protótipo obteve vantagem com relação à capacidade de rastreamento automático dos marcadores, com respeito à presença de falsos-marcadores e também com relação à menor incidência de ruídos nas trajetórias reconstruídas. As diferenças apresentadas tornam-se ainda mais relevantes considerando que o software Arena exige a utilização de um modelo esquelético e ainda depende de uma etapa de pós-processamento para o cálculo das trajetórias dos marcadores.

# Capítulo 9

## Aplicação nos Esportes

O protótipo desenvolvido também foi avaliado na captura de movimentos envolvendo os esportes. Este capítulo apresenta os testes realizados com atletas do halterofilismo paralímpico e com um atleta do taekwondo.

### 9.1 Aplicação no Halterofilismo Paralímpico

---

O protótipo foi empregado na modalidade halterofilismo paralímpico para rastrear o movimento da barra de apoio durante sessões de levantamento de peso (supino reto). O experimento foi conduzido buscando avaliar o desempenho e direcionar o treinamento de oito atletas da seleção brasileira permanente de halterofilismo paralímpico.

A barra olímpica de 2,2 metros de comprimento foi rastreada com o auxílio de três marcadores reflexivos. Dois deles foram fixados nas extremidades e um terceiro marcador foi posto no centro, sobre a barra. Os atletas realizaram um ciclo de movimento cada com o peso ajustado para 90% de sua carga máxima.

A Figura 9.1 ilustra as trajetórias espaciais dos marcadores, reconstruídas pelo protótipo e correspondentes a dois ciclos de movimentos. Os traçados da Figura 9.1(a) mostram que o atleta exerceu um bom controle horizontal sobre a barra, mantendo próximas as trajetórias de descida e subida em ambos os lados. Já os

traçados da Figura 9.1(b) evidenciam o movimento irregular desempenhado por outro atleta, especialmente com o seu braço esquerdo.

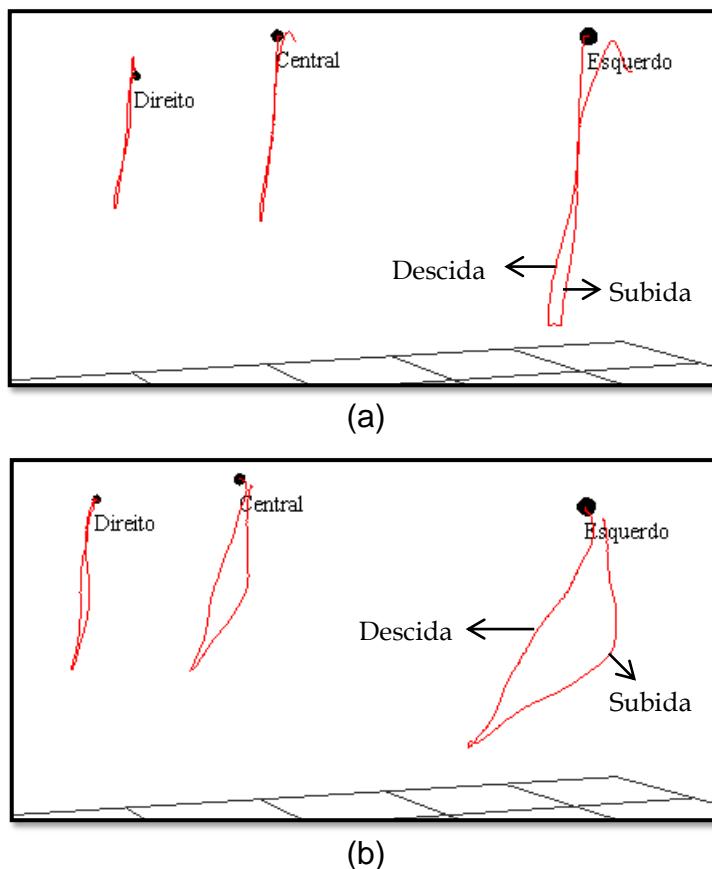


Figura 9.1 – Trajetória espacial dos três marcadores utilizados na barra ao longo da execução de um ciclo de movimento no supino paralímpico. (a) Atleta exerce controle próximo do ideal sobre a barra. (b) Atleta realiza um movimento irregular com o braço esquerdo.

A partir das trajetórias reconstruídas para os oito atletas, foi possível calcular diversos parâmetros a respeito de seus movimentos. O gráfico (a) da Figura 9.2 ilustra o tempo gasto por cada atleta na descida e na subida da barra. Já o gráfico (b) da mesma figura ilustra o tempo em que os atletas permaneceram com a barra próximo ao peito (tempo de parada). O gráfico (c) apresenta a maior diferença de altura registrada para os marcadores extremos da barra. Colunas mais altas indicam um maior desnivelamento da barra em algum ponto do movimento. Por fim, o gráfico (d) da Figura 9.2 exibe a altura dos marcadores *esquerdo* e *direito* em função do tempo, durante a execução de um ciclo de movimento realizado pelo atleta G.

Observe que todos esses parâmetros e trajetórias trazem detalhes acerca dos movimentos que poderiam não ser percebidos, com clareza, por meio de inspeção

visual simples. Um tempo de parada insuficiente, por exemplo, pode causar a desclassificação do atleta em uma competição oficial. Nesse sentido, o sistema de captura de movimento exerce um importante papel no processo de busca pelo aprimoramento contínuo do atleta.

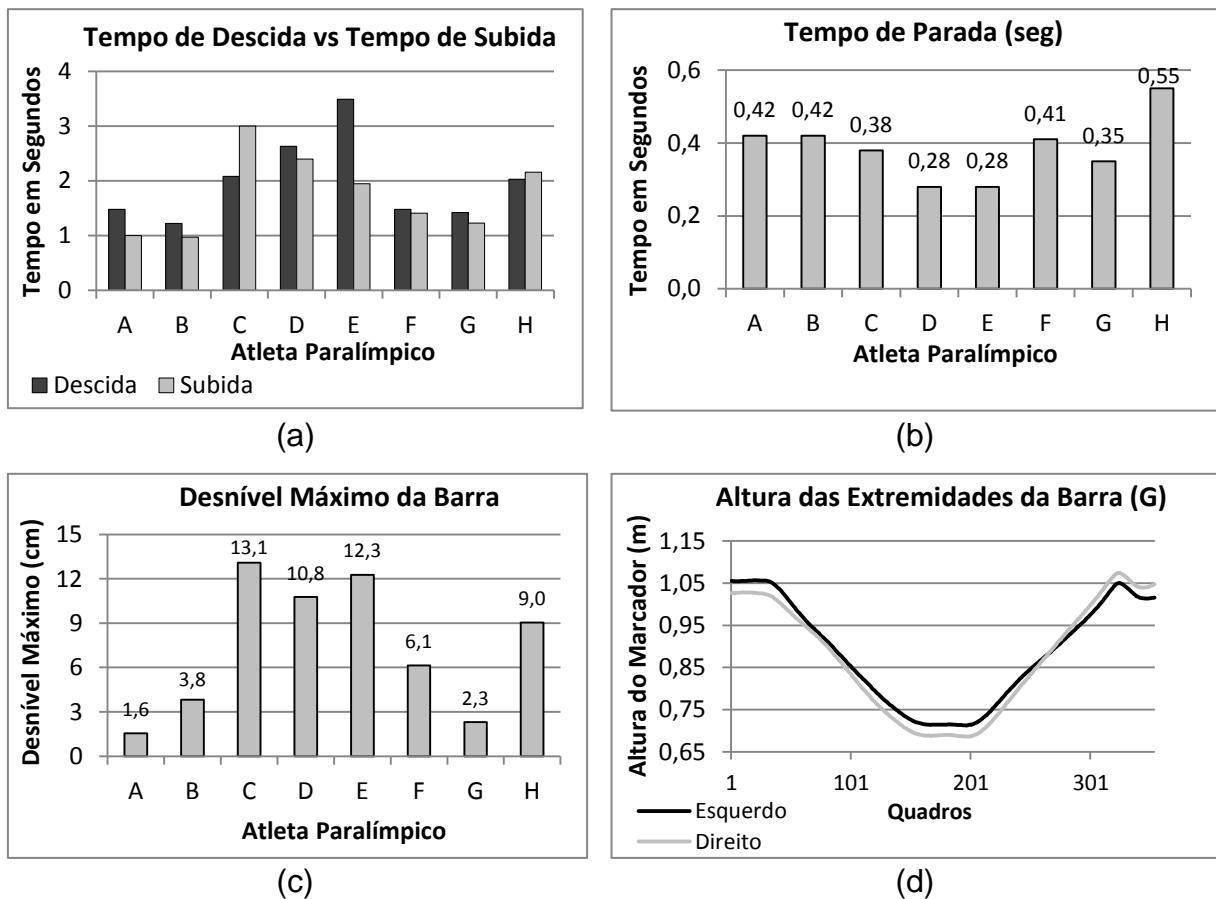


Figura 9.2 - Características dos movimentos de supino paralímpico extraídas das trajetórias espaciais reconstruídas pelo protótipo. (a) Tempos de descida e subida da barra para cada atleta. (b) Tempo de permanência da barra próximo ao peito. (c) Maior diferença de altura registrada entre os marcadores extremos da barra. (d) Altura dos marcadores extremos ao longo do movimento executado por G.

## 9.2 Aplicação no Taekwondo

O protótipo desenvolvido também foi testado na captura dos movimentos de um atleta do taekwondo. O objetivo, novamente, foi auxiliar o atleta e seu treinador nas etapas de avaliação e treinamento durante a busca pelo aperfeiçoamento dos golpes.

No total, 35 marcadores foram utilizados e o protótipo foi capaz de rastrear a maioria deles para a maior parte dos golpes testados, exceto para as situações em que os marcadores se tornaram ocultos.

A Figura 9.3 ilustra os marcadores reconstruídos pelo protótipo durante a execução do movimento *Timyo Tchagui*. A trajetória do marcador posicionado na ponta do pé direito do atleta é apresentada em destaque. Na Figura 9.3(a), o atleta encontra-se pronto para iniciar o movimento. Na Figura 9.3(b), o salto já teve início e a Figura 9.3(c) exibe os marcadores quando o atleta atinge o ponto mais alto do movimento. Por fim, a Figura 9.3(d) ilustra o momento seguinte ao golpe e destaca a trajetória completa definida pelo marcador.

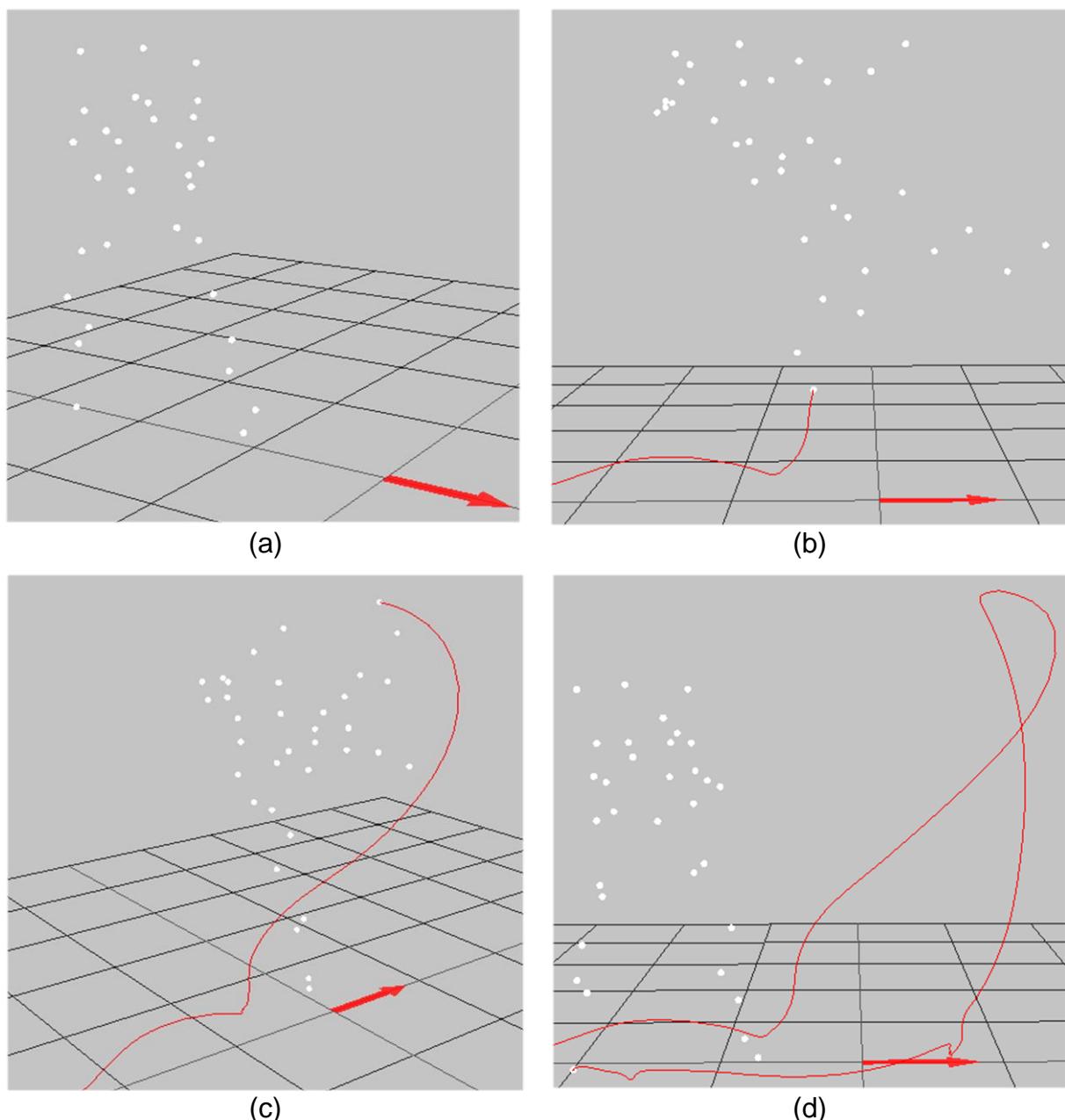


Figura 9.3 – Exemplo de uso do protótipo para capturar e reconstruir os movimentos de um atleta do Taekwondo. O movimento ilustrado corresponde ao golpe *Timyo Tchagui*. A detecção dos marcadores e a reconstrução de suas trajetórias foram realizadas em tempo real.

# Capítulo 10

## Conclusão e Trabalhos Futuros

A principal contribuição deste trabalho consiste na proposta de um método computacional livre de modelo esquelético para reconstruir as trajetórias de uma nuvem de marcadores, em sistemas de captura de movimento ópticos, utilizando múltiplas câmeras especializadas de infravermelho.

Diferente das pesquisas publicadas recentemente, o método proposto agrega um conjunto de características e capacidades que viabilizam seu uso em diversas aplicações, uma vez que:

- é automático e não depende da intervenção do usuário durante a captura;
- oferece níveis de precisão e exatidão comparáveis a sistemas comerciais disponíveis no cenário internacional;
- é capaz de operar em tempo real, considerando dezenas de marcadores, múltiplas câmeras e em velocidades de captura que podem ultrapassar 100 quadros por segundo;
- é flexível com relação ao número e posicionamento dos marcadores e não depende essencialmente de um modelo esquelético subjacente;
- é capaz de rastrear os marcadores em movimentos corporais suficientemente rápidos e complexos para aplicações nos esportes;
- é capaz de operar na presença de oclusões dos marcadores e sobreposições dos pontos nas imagens.

O método proposto foi testado e avaliado mediante o desenvolvimento de um *software*, referenciado pelo termo *protótipo*. Os experimentos empregaram um conjunto de oito câmeras de infravermelho e foram conduzidos em quatro etapas:

1. Utilizando um dispositivo de teste;
2. Utilizando um manequim e uma plataforma giratória;
3. Utilizando uma esteira ergométrica;
4. Realizando aplicações nos esportes: halterofilismo paralímpico e *taekwondo*.

A primeira etapa dos experimentos foi realizada para avaliação da precisão, exatidão e capacidade de rastreamento e reconstrução da técnica. Com valores de precisão e exatidão próximos ou inferior a 1 mm, o protótipo ocupou posições de destaque quando comparado a softwares comerciais submetidos a teste análogo. Com respeito à capacidade de rastreamento, o protótipo foi capaz de rastrear tanto os marcadores que permaneceram visíveis para todo o conjunto de câmeras quanto os marcadores que sofreram de oclusões parciais e foram visualizados apenas por um subconjunto dinâmico de câmeras. A baixa resolução das imagens juntamente com as imperfeições intrínsecas da técnica impossibilitou o rastreamento automático dos marcadores postos a uma distância igual ou inferior a 1 cm.

A segunda e a terceira etapas dos testes envolveram a reconstrução das trajetórias de 38 marcadores. Os experimentos foram escolhidos por oferecerem oportunidades para a ocorrência de sobreposição dos pontos nas imagens, alternâncias de visão, oclusões, acelerações e desacelerações dos marcadores e por viabilizarem uma comparação direta com o software comercial *Arena*. Os resultados comparativos mostraram que:

- a) o protótipo foi capaz de reconstruir automaticamente um número igual ou maior de trajetórias quando comparado ao software *Arena* (número igual quando considerado o teste com o manequim e superior quando considerado o teste com a esteira ergométrica - 33,4 contra 28,75, em média);
- b) o protótipo produziu trajetórias com menor incidência de ruídos do tipo *spike* (o número calculado de *spikes* foi de 22 a 81% menor);

- c) o protótipo produziu trajetórias com menor incidência de ruídos de baixa amplitude, conforme apresentado pelo coeficiente geral de ruído;
- d) o protótipo gerou uma quantidade menor de falsos marcadores, que são indesejáveis e podem interferir nas trajetórias dos marcadores reais.

Vale ressaltar que as diferenças obtidas ao relacionar o protótipo com o software Arena tornam-se ainda mais relevantes considerando que o software da *NaturalPoint* exige a utilização de um modelo esquelético e ainda depende de uma etapa de pós-processamento para finalizar o cálculo das trajetórias dos marcadores. Conforme já mencionado, o emprego obrigatório de um modelo esquelético pode restringir o uso do sistema de captura. Entretanto, o modelo também possibilita a incorporação de otimizações no processo de rastreamento e reconstrução. Adicionalmente, a etapa de geração de trajetórias requerida pelo software Arena consome tempo extra do usuário, mas possibilita que algoritmos de rastreamento mais sofisticados sejam introduzidos para corrigir e/ou aprimorar o cálculo das trajetórias dos marcadores.

Na quarta parte dos testes, o protótipo foi utilizado para rastrear e reconstruir movimentos no halterofilismo paralímpico e no *taekwondo*. Ambos os testes foram conduzidos com rapidez e praticidade e o protótipo foi capaz de calcular as trajetórias dos marcadores em tempo real. No caso do halterofilismo, as trajetórias calculadas e os parâmetros extraídos foram importantes para uma avaliação mais precisa dos movimentos dos atletas e para a proposta de um treinamento direcionado. Benefícios semelhantes foram obtidos quando da aplicação no *taekwondo*.

Os diversos resultados apresentados sugerem que a técnica tem potencial para ser utilizada em um grande leque de aplicações, incluindo desde uma simples análise de marcha a até aplicações que demandam a captura de movimentos mais complexos e com baixo tempo de processamento, como nos esportes, na produção cinematográfica e no desenvolvimento de jogos.

O método apresentado deve incentivar pesquisadores e empresas a desenvolverem sistemas de captura de movimento de alto desempenho e com menor custo. A

contribuição é especialmente importante para o avanço nacional nessa área, pois a grande maioria dos sistemas sendo utilizados nos laboratórios de universidades, estúdios e centros de treinamento e reabilitação do país é adquirida do mercado externo e envolve altos custos com aquisição e manutenção.

## 10.1 Trabalhos Futuros

---

Como trabalho futuro decorrente desta pesquisa, pretende-se ainda:

- ❖ Conduzir experimentos utilizando um número maior de câmeras. Pretende-se realizar um estudo de escalabilidade relacionando o desempenho da técnica e o tempo de execução do algoritmo com o número de câmeras empregado;
- ❖ Conduzir experimentos envolvendo dois ou mais indivíduos em um ambiente de reconstrução significativamente maior;
- ❖ Testar e avaliar o desempenho do método para reconstruir outros tipos de movimentos corporais, como os de um ginasta ou dançarino;
- ❖ Incorporar um modelo esquelético ao método como ferramenta **opcional** e complementar para possibilitar a rotulação automática dos marcadores quando viável.

## 10.2 Publicações

---

As principais publicações obtidas direta ou indiretamente a partir deste trabalho foram:

- FURTADO, D.A.; PEREIRA, A.A.; ANDRADE, A.O.; JUNIOR, D.P. **A specialized motion capture system for real-time analysis of mandibular movements using infrared cameras.** *BioMedical Engineering OnLine*, 2013, 12:17. doi:10.1186/1475-925X-12-17.
- FURTADO, D. A.; PEREIRA, A. A.; BELLOMO JR., D. P.; ALMEIDA, M. F .S. **A Mocap System for 3D Reconstruction of Mandibular Movements in Real Time.** *Anais do XXII Congresso Brasileiro de Engenharia Biomédica*, 2010, Tiradentes.

- ALMEIDA, M. F. S.; CAVALHEIRO, G. L.; ANDRADE, A. O.; PEREIRA, A. A.; FURTADO, D. A. **Investigation of changes in kinetic tremor through analysis of hand-drawing movements - Differences between physiological and essential tremors.** *IV International Joint Conference on Biomedical Engineering Systems and Technologies*, 2011, Roma, Itália. *Annals of International conference on Bio-inspired Systems and Signal Processing*, 2011.
- ALMEIDA, M. F. S.; CAVALHEIRO, G. L.; FURTADO, D. A.; ANDRADE, A. O.; PEREIRA. A. **Investigation of differences between physiological and essential tremors.** *XXII Congresso Brasileiro de Engenharia Biomédica*, 2010, Tiradentes. *Anais do XXII Congresso Brasileiro de Engenharia Biomédica*, 2010.

## Referências Bibliográficas

---

ABDEL-AZIZ, Y. I.; KARARA, H. M. Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry. Proceedings of the Symposium on Close-Range photogrammetry, 1971, Falls Church. American Society of Photogrammetry. p.18.

ADAM, G. K.; O'BRIEN, J. F.; FORSYTH, D. A. **Skeletal Parameter Estimation from Optical Motion Capture Data**. IEEE Conf. on Computer Vision and Pattern Recognition. 2: 782-788 p. 2005.

AMBRÓSIO, J.; ABRANTES, J.; LOPES, G. Spatial reconstruction of human motion by means of a single camera and a biomechanical model. **Human Movement Science**, v. 20, n. 6, p. 829-851, 2001. ISSN 0167-9457.

ANIMAZOO. Homepage oficial da Animazoo. 2013. Disponível em: <<http://www.animazoo.com>>. Acesso em: 16/12/2012.

ARMSTRONG, J. S. Extrapolation Methods. In: (Ed.). **Long-Range Forecasting: From Crystal Ball to Computer**. 2. New York: Wiley-Interscience, 1985.

BACHMANN, E. R. **Inertial And Magnetic Tracking Of Limb Segment Orientation For Inserting Humans Into Synthetic Environments**. 2000. NAVAL POSTGRADUATE SCHOOL, Monterey, California.

BARCA, J. C.; RUMANTIR, G.; LI, R. Noise Filtering of New Motion Capture Markers Using Modified K-Means. In: (Ed.). **Computational Intelligence in Multimedia Processing: Recent Advances**, 2008. p.167-189.

BARONI, G.; FERRIGNO, G.; PEDOTTI, A. Implementation and application of real-time motion analysis based on passive markers. **Medical and Biological Engineering and Computing**, v. 36, n. 6, p. 693-703, 1998. ISSN 0140-0118.

BARRIS, S.; BUTTON, C. A review of vision-based motion analysis in sport. **Sports Medicine**, v. 38, n. 12, p. 18, 2008.

BESL, P. J.; MCKAY, N. D. A Method for Registration of 3-D Shapes. **IEEE Trans. Pattern Anal. Mach. Intell.**, v. 14, n. 2, p. 239-256, 1992. ISSN 0162-8828.

BRAY, J. **Markerless Based Human Motion Capture: A Survey** 2006.

BROWN, D. C. Decentering Distortion of Lenses. **Photogrammetric Engineering**, v. 32, n. 3, p. 19, 1966.

CAILLETTE, F.; HOWARD, T. **Real-Time Markerless Human Body Tracking Using Colored Voxels and 3-D Blobs**. Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality: IEEE Computer Society: 266-267 p. 2004.

CANTON-FERRER, C.; CASAS, J. R.; PARD, M. Marker-based human motion capture in multiview sequences. **EURASIP J. Adv. Signal Process**, v. 2010, p. 1-11, 2010. ISSN 1110-8657.

CASTRO, J. L. G.; MEDINA-CARNICER, R.; GALISTEO, A. M. Design and evaluation of a new three-dimensional motion capture system based on video. **Gait & Posture**, v. 24, n. 1, p. 126-129, 2006. ISSN 0966-6362.

CERVERI, P.; PEDOTTI, A.; FERRIGNO, G. Robust recovery of human motion from video using Kalman filters and virtual humans. **Human Movement Science**, v. 22, n. 3, p. 377-404, 2003. ISSN 0167-9457.

CHU, C.-W.; JENKINS, O. C.; MATARIC, M. J. Towards Model-free Markerless Motion Capture. IEEE Intl. Conf. On Robotics and Automation, 2003.

CLOETE, T.; SCHEFFER, C. Benchmarking of a full-body inertial motion capture system for clinical gait analysis. Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE, 2008. p.4579-4582.

COCKCROFT, J. Evaluating Inertial Motion Capture Systems for Sports Performance Analysis: A Case Study in the Measurement of Road Cycling Kinematics Test & Measurement Conference Proceedings, 2010, Drakensberg, South Africa. p.10.

CONRADY, A. E. Decentered Lens-systems. **Monthly Notices of the Royal Astronomical Society**, v. 79, p. 384-390, 1919.

CORAZZA, S. et al. A Markerless Motion Capture System to Study Musculoskeletal Biomechanics: Visual Hull and Simulated Annealing Approach. **Annals of Biomedical Engineering**, v. 34, n. 6, p. 1019-1029, 2006. ISSN 0090-6964.

CORAZZA, S. et al. Markerless Motion Capture through Visual Hull, Articulated ICP and Subject Specific Model Generation. **International Journal of Computer Vision**, v. 87, n. 1, p. 156-169, 2010. ISSN 0920-5691.

CORRALES, J. A.; CANDELAS, F. A.; TORRES, F. **Hybrid tracking of human operators using IMU/UWB data fusion by a Kalman filter**. Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction. Amsterdam, The Netherlands: ACM: 193-200 p. 2008.

DAVIES, G. R. et al. **EuropaNet Research Infrastructure: Planetary Simulation Facilities**. European Planetary Science Congress. Rome, Italy. 4: 20-24 p. 2009.

DAVIS, R. B. Clinical gait analysis. **Engineering in Medicine and Biology Magazine**, v. 7, n. 3, p. 6, 1988.

DIJKSTRA, P. U. et al. Temporomandibular joint mobility assessment: a comparison between four methods. **J Oral Rehabil**, v. 22, n. 6, p. 439-44, 1995. ISSN 0305-182X.

DOLNÁ, Z. et al. Human motion analysis and its application in sports and sport medicine Proceedings of the 4th Young Biomedical Engineers and Researchers, 2008, Ostrava, Czech Republic. p.4.

ELSAESSE, T. James Cameron's Avatar: access for all. **New Review of Film and Television Studies**, v. 9, n. 3, p. 18, 2011. ISSN 1740-7923.

FANG, J.-J.; KUO, T.-H. Modelling of mandibular movement. **Computers in Biology and Medicine**, v. 38, n. 11-12, p. 1152-1162, 2008. ISSN 0010-4825.

FIELD, M. et al. Human motion capture sensors and analysis in robotics. **Industrial Robot: An International Journal**, v. 38, n. 2, p. 163-171, 2011.

FIGUEROA, P. J.; LEITE, N. J.; BARROS, R. M. L. A flexible software for tracking of markers used in human motion analysis. **Computer Methods and Programs in Biomedicine**, v. 72, n. 2, p. 155-165, 2003. ISSN 0169-2607.

FLAM, D. L. et al. **OpenMoCap: An Open Source Software for Optical Motion Capture**. Brazilian Symposium on Games and Digital Entertainment. Rio de Janeiro, RJ, Brazil: IEEE Computer Society: 151-161 p. 2009.

FRAGA, L. G. D. L.; VITE-SILVA, I. **Direct 3D metric reconstruction from multiple views using differential evolution**. Proceedings of the 2008 conference on Applications of evolutionary computing. Naples, Italy: 341-346 p. 2008.

FUKUI, T. et al. Correlation between facial morphology, mouth opening ability, and condylar movement during opening- closing jaw movements in female adults with normal occlusion. **European Journal of Orthodontic**, v. 24, p. 327-336, 2002.

FURTADO, D. A. et al. A MoCap system for 3D reconstruction of mandibular movements in real-time. XXII Congresso Brasileiro de Engenharia Biomédica, 2010, Tiradentes, MG. p.4.

GANDOLLA, M. et al. fMRI brain mapping during motion capture and FES induced motor tasks: Signal to noise ratio assessment. **Medical Engineering & Physics**, v. 33, n. 8, p. 1027-1032, 2011. ISSN 1350-4533.

GIAKAS, G.; BALTZOPOULOS, V. A comparison of automatic filtering techniques applied to biomechanical walking data. **Journal of Biomechanics**, v. 30, n. 8, p. 847-850, 1997. ISSN 0021-9290.

GLEICHER, M. Animation from observation: Motion capture and motion editing. **SIGGRAPH Comput. Graph.**, v. 33, n. 4, p. 51-54, 1999. ISSN 0097-8930.

GOLUB, G.; REINSCH, C. Singular value decomposition and least squares solutions. **Numerische Mathematik**, v. 14, n. 5, p. 403-420, 1970. ISSN 0029-599X.

GOMIDE, J. V. B. et al. **Captura de Movimento e Animação de Personagens em Jogos**. VIII Brazilian Symposium on Games and Digital Entertainment. Rio de Janeiro, RJ, Brazil 2009.

GORDON, N. J.; SALMOND, D. J.; SMITH, A. F. M. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. **Radar and Signal Processing**, 1993. p.107-113.

GRAHAM, R. L. An Efficient Algorithm for Determining the Convex Hull of a Finite Planar Set. **Information Processing Letters**, v. 1, p. 132-133, 1972.

GUERRA-FILHO, G. B. Optical motion capture: Theory and implementation. **Journal of Theoretical and Applied Informatics (RITA)**, v. 12, n. 2, p. 61-89, 2005.

HARTLEY, R.; ZISSEMAN, A. **Multiple View Geometry in Computer Vision**. New York, NY, USA: Cambridge University Press, 2004. ISBN 0521540518.

HARTLEY, R. I.; STURM, P. Triangulation. **Computer Vision and Image Understanding**, v. 68, n. 2, p. 146-157, 1997. ISSN 1077-3142.

HASSAN, E. A.; JENKYN, T. R.; DUNNING, C. E. Direct comparison of kinematic data collected using an electromagnetic tracking system versus a digital optical system. **Journal of Biomechanics**, v. 40, n. 4, p. 930-935, 2007. ISSN 0021-9290.

HERALD, T. G. Simulação Militar usando MoCap. 2010. Disponível em: <<http://theglobalherald.com/itec-2010-training-and-simulation-for-modern-combat/3811/>>. Acesso em: 29/11/2011.

HERDA, L. et al. Skeleton-Based Motion Capture for Robust Reconstruction of Human Motion. Proceedings of the Computer Animation, 2000, IEEE Computer Society. p.77.

\_\_\_\_\_. Using skeleton-based tracking to increase the reliability of optical motion capture. **Human Movement Science**, v. 20, n. 3, p. 313-341, 2001.

HILTON, A. et al. Virtual People: Capturing Human Models to Populate Virtual Worlds. CA '99: Proceedings of the Computer Animation, 1999, IEEE Computer Society. p.174.

HINIDUMA UDUGAMA GAMAGE, S. S.; LASENBY, J. **Estimation of centres and axes of rotation of articulated bodies in general motion for global skeleton fitting**. University of Cambridge. Cambridge, UK, p.31. 2001

IGN. An Interview with Andy Serkis. 2003. Disponível em: <<http://www.ign.com/articles/2003/01/27/an-interview-with-andy-serkis>>. Acesso em: 20/03/2013.

\_\_\_\_\_. Interview: Tom Hanks. 2004. Disponível em: <<http://www.ign.com/articles/2004/11/10/interview-tom-hanks>>. Acesso em: 20/03/2013.

ISARD, M.; BLAKE, A. CONDENSATION—Conditional Density Propagation for Visual Tracking. **International Journal of Computer Vision**, v. 29, n. 1, p. 5-28, 1998. ISSN 0920-5691.

IVEKOVIC, S.; FUSIELLO, A.; TRUCCO, E. Fundamentals of Multiple-View Geometry. In: (Ed.). **3D Videocommunication**: John Wiley & Sons, Ltd, 2006. p.91-113. ISBN 9780470022733.

JEFFREY, A. R.; AJAY, S.; SCOTT, L. D. Simulation of human movement: applications using OpenSim. **Procedia IUTAM**, v. 2, p. 186-198, 2011. ISSN 2210-9838.

JOHN, V.; TRUCCO, E.; IVEKOVIC, S. Markerless human articulated tracking using hierarchical particle swarm optimisation. **Image and Vision Computing**, v. 28, n. 11, p. 1530-1547, 2010. ISSN 0262-8856.

JOHNSON, A.; BOBICK, A. A Multi-view Method for Gait Recognition Using Static Body Parameters. In: BIGUN, J. e SMERALDI, F. (Ed.). **Audio- and Video-Based Biometric Person Authentication**: Springer Berlin / Heidelberg, v.2091, 2001. p.301-311. (Lecture Notes in Computer Science). ISBN 978-3-540-42216-7.

JOHNSON, G. et al. Efficient Marker Matching Using Pair-Wise Constraints in Physical Therapy. In: BEBIS, G.;BOYLE, R., et al (Ed.). **Advances in Visual Computing**: Springer Berlin / Heidelberg, v.6454, 2010a. p.222-231. (Lecture Notes in Computer Science).

JOHNSON, G. et al. **Efficient marker matching using pair-wise constraints in physical therapy**. Proceedings of the 6th international conference on Advances in visual computing - Volume Part II. Las Vegas, NV, USA: 222-231 p. 2010b.

KEHL, R.; GOOL, L. V. Markerless tracking of complex human motions from multiple views. **Computer Vision and Image Understanding**, v. 104, n. 2, p. 190-209, 2006. ISSN 1077-3142.

KING, B. A.; PAULSON, L. D. Motion Capture Moves into New Realms. **Computer**, v. 40, n. 9, p. 13-16, 2007. ISSN 0018-9162.

KISER, J. **Motion Capture Technology: The Basics**. C-Motion, p.17. 2010

KITAGAWA, M.; WINDSOR, B. **MoCap for Artists: Workflow and Techniques for Motion Capture**. Focal Press 2008. 232 ISBN 978-0-240-81000-3.

LAREAU, J. **Ford uses animation technology to improve worker safety, car quality and purchasing costs**. Automotive News: Crain Communications, Inc. 85: 16 p. 2011.

LEMARÉCHAL, C. Lagrangian Relaxation. In: JÜNGER, M. e NADDEF, D. (Ed.). **Computational Combinatorial Optimization: Optimal or Provably Near-Optimal Solutions**: Springer Berlin / Heidelberg, v.2241, 2001. p.112-156. (Lecture Notes in Computer Science). ISBN 978-3-540-42877-0.

LOU, H.; CHAI, J. Example-Based Human Motion Denoising. **IEEE Transactions on Visualization and Computer Graphics**, v. 16, n. 5, p. 870-879, 2010. ISSN 1077-2626.

MALETSKY, L. P.; SUN, J.; MORTON, N. A. Accuracy of an optical active-marker system to track the relative motion of rigid bodies. **Journal of Biomechanics**, v. 40, n. 3, p. 682-685, 2007. ISSN 0021-9290.

MANAL, K. et al. Comparison of surface mounted markers and attachment methods in estimating tibial rotations during walking: an in vivo study. **Gait & Posture**, v. 11, n. 1, p. 38-45, 2000.

MARZANI, F.; CALAIS, E.; LEGRAND, L. A 3-D marker-free system for the analysis of movement disabilities - an application to the legs. **IEEE Transactions on Information Technology in Biomedicine**, v. 5, n. 1, p. 8, 2001.

MCNAMARA, J. E. et al. **Motion capture of chest and abdominal markers using a flexible multi-camera motion-tracking system for correcting motion-induced**

**artifacts in cardiac SPECT.** Nuclear Science Symposium Conference Record. Honolulu, HI. 6: 4289 - 4293 p. 2007.

**MENACHE, A. Understanding Motion Capture for Computer Animation and Video Games.** San Francisco: Morgan Kaufmann Publishers Inc., 1999. 238 ISBN 0124906303.

**MIKIĆ, I. et al.** Human Body Model Acquisition and Tracking Using Voxel Data. **International Journal of Computer Vision**, v. 53, n. 3, p. 199-223, 2003. ISSN 0920-5691.

**MILLER, V. J. et al.** A mouth opening index for patients with temporomandibular disorders. **Journal of oral rehabilitation**, v. 26, 1999. ISSN 0305-182X

**MOESLUND, T. B.** Interacting with a Virtual World through Motion Capture. In: (Ed.). **Interaction in Virtual Inhabited 3D Worlds**: Springer-Verlag, 2000. cap. 11,

**MOESLUND, T. B.; GRANUM, E.** A Survey of Computer Vision-Based Human Motion Capture. **Computer Vision and Image Understanding**, v. 81, n. 3, p. 231-268, 2001. ISSN 1077-3142.

**MOESLUND, T. B.; HILTON, A.; KRÜGER, V.** A survey of advances in vision-based human motion capture and analysis. **Computer Vision and Image Understanding**, v. 104, n. 2, p. 90-126, 2006. ISSN 1077-3142.

**MOTION, O.** Aplicações militares para sistemas de mocap. 2011. Disponível em: < <http://organicmotion.com/industries/military> >. Acesso em: 29/11/2011.

**MOTORTREND.** Ford Using Motion-Capture Tech to Help Improve Ergonomics. 2011. Disponível em: < <http://wot.motortrend.com> >. Acesso em: 20/03/2013.

**MUNDERMANN, L.; CORAZZA, S.; ANDRIACCHI, T. P.** Accurately measuring human movement using articulated ICP with soft-joint constraints and a repository of articulated models. **Computer Vision and Pattern Recognition**, 2007. CVPR '07. IEEE Conference on, 2007, 17-22 June 2007. p.1-6.

MÜNDERMANN, L.; CORAZZA, S.; ANDRIACCHI, T. P. The evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical app. **Journal of Neuroengineering Rehabilitation**, v. 3, n. 6, 2006.

NAEIJE, M. Measurement of condylar motion: a plea for the use of the condylar kinematic centre. **Journal of oral rehabilitation**, v. 30, n. 3, p. 6, 2003. ISSN 0305-182X.

NATURALPOINT. Homepage oficial da NaturalPoint. 2013. Disponível em: < [www.naturalpoint.com](http://www.naturalpoint.com) >. Acesso em: 20/03/2013.

O'BRIEN, J. F. et al. Automatic joint parameter estimation from magnetic motion capture data. **Graphics Interface**, 2000. p.53-60.

OSHITA, M. Motion-Capture-Based Avatar Control Framework in Third-Person View Virtual Environments. Proc. ACM SIG Computer-Human Interaction Int'l Conf. Advances in Computer Entertainment, 2006, Hollywood, U.S.A.

PANDY, M. G. Computer modeling and simulation of human movement. **Annual Review of Biomedical Engineering**, v. 3, p. 245-73, 2001.

PAPIC, V.; ZANCHI, V.; CECIC, M. Motion analysis system for identification of 3D human locomotion kinematics data and accuracy testing. **Simulation Modelling Practice and Theory**, v. 12, n. 2, p. 159-170, 2004. ISSN 1569-190X.

PING, Z.; YUE, L.; YONGTIAN, W. Adaptive Real-Time Labeling and Recognition of Multiple Infrared Markers Using FPGA. **Image and Graphics**, 2009. ICIG '09. Fifth International Conference on, 2009, 20-23 Sept. 2009. p.983-988.

PINHEIRO, A. P. **Desenvolvimento de um sistema de captura e análise de movimentos baseado em técnicas de visão computacional**. 2008. (Mestrado). Faculdade de Engenharia Elétrica, Universidade Federal de Uberlândia, Uberlândia.

PINHEIRO, A. P. et al. A computational method for recording and analysis of mandibular movements. **Journal of Applied Oral Science**, v. 16, p. 321-327, 2008.

PINTO, F. et al. **BraTrack: a low-cost marker-based optical stereo tracking system**. ACM SIGGRAPH 2008 posters. Los Angeles, California: ACM: 1-1 p. 2008.

PISSANOS, M. FIFA 09: Maurice Edu Motion Capture. 2008. Disponível em: <<http://greeksuperpatch.com/>>. Acesso em: 19/11/2011.

PONS-MOLL, G. et al. **Efficient and robust shape matching for model based human motion capture**. Proceedings of the 33rd international conference on Pattern recognition. Frankfurt, Germany: Springer-Verlag: 416-425 p. 2011.

PRIBANIĆ, T.; PEHAREC, S.; MEDVED, V. A comparison between 2D plate calibration and wand calibration for 3D kinematic systems. **International Journal of Fundamental and Applied Kinesiology**, v. 41, n. 2, p. 147-155, 2009.

PRIBANIC, T.; STURM, P.; CIFREK, M. Camera parameter initialization for 3D kinematic systems. 4th International Symposium on Image and Signal Processing and Analysis, 2005, 15-17 Sept. 2005. p.494-499.

PRIBANIĆ, T.; STURM, P.; CIFREK, M. Calibration of 3D kinematic systems using orthogonality constraints. **Machine Vision and Applications**, v. 18, n. 6, p. 367-381, 2007. ISSN 0932-8092.

QIONG, W.; BOULANGER, P. Real-Time Estimation of Missing Markers for Reconstruction of Human Motion. Virtual Reality (SVR), 2011 XIII Symposium on, 2011, 23-26 May 2011. p.161-168.

QUALISYS. Sistema de *mocap* para nadadores. 2013. Disponível em: <<http://www.qualisys.com/applications/biomechanics/swimming/>>. Acesso em: 20/03/2013.

RICHARDS, J. G. The measurement of human motion: A comparison of commercially available systems. **Human Movement Science**, v. 18, n. 5, p. 589-602, 1999. ISSN 0167-9457.

RINGER, M.; LASENBY, J. Modelling and Tracking Articulated Motion from Multiple Camera Views. In Proc. British Machine Vision Conf, 2000. p.172-182.

\_\_\_\_\_. A procedure for automatically estimating model parameters in optical motion capture. **Image and Vision Computing**, v. 22, n. 10, p. 843-850, 2004.

RISTIC, B.; ARULAMPALAM, S.; GORDON, N. **Beyond the Kalman Filter: Particle Filters for Tracking Applications.** Artech House, 2004.

RIZALDY. Inner Esteem Motion capture Studios. 2011. Disponível em: < <http://www.motioncapturestudios.com/> >. Acesso em: 26/10/2011.

ROBERTSON, D. G. E. **Vicon Workstation Quick Reference Guide.** University of Ottawa. Ottawa. 2009

ROSENHAHN, B. et al. A system for marker-less motion capture. **Künstliche Intelligenz**, v. 1, p. 45-51, 2006.

ROSENHAHN, B.; KLETTE, R.; METAXAS, D. **Human Motion: Understanding, Modelling, Capture, and Animation.** Springer, 2007. 646 ISBN 978-1402066924.

SABOUNE, J.; CHARPILLET, F. Markerless Human Motion Capture for Gait Analysis. **Computing Research Repository - CoRR**, 2005.

SANTOS, I. C. et al. A prototype system for acquisition and analysis of 3D mandibular movement. **International Journal of Mechanics and Materials in Design**, v. 4, n. 2, p. 173-180, 2008. ISSN 1569-1713.

SATO, T. et al. Design and Implementation of a Robust and Real-time Ultrasonic Motion-capture System. International Conference in Indoor Positioning and Indoor Navigation, 2011, Guimarães, Portugal.

SETHI, I. K.; SALARI, V.; VEMURI, S. Feature point matching in image sequences. **Pattern Recognition Letters**, v. 7, n. 2, p. 113-121, 1988. ISSN 0167-8655.

SHAFIQ, M. S.; TÜMER, S. T.; GÜLER, H. C. Marker detection and trajectory generation algorithms for a multicamera based gait analysis system. **Mechatronics**, v. 11, n. 4, p. 409-437, 2001. ISSN 0957-4158.

SHAPIRO, L. S.; ZISSERMAN, A.; BRADY, M. 3D Motion recovery via affine Epipolar geometry. **International Journal of Computer Vision**, v. 16, n. 2, p. 147-182, 1995. ISSN 0920-5691.

SHAPIRO, R. Direct linear transformation method for three-dimensional cinematography. **Res. Quart.**, v. 49, n. 2, p. 197-205, May 1978.

SHIH-PIN, C.; YI-YAO, C.; WU-CHOU, C. The Cost-Effective Method to Develop a Real-Time Motion Capture System. Computer Sciences and Convergence Information Technology, 2009. ICCIT '09. Fourth International Conference on, 2009, 24-26 Nov. 2009. p.494-498.

SIGAL, L.; BALAN, A. O.; BLACK, M. J. HumanEva: Synchronized Video and Motion Capture Dataset and Baseline Algorithm for Evaluation of Articulated Human Motion. **International Journal of Computer Vision**, v. 87, n. 1-2, p. 4-27, 2009.

SILAGHI, M.-C. et al. Local and Global Skeleton Fitting Techniques for Optical Motion Capture. In: MAGNENAT-THALMANN, N. e THALMANN, D. (Ed.). **Modelling and Motion Capture Techniques for Virtual Environments**: Springer Berlin / Heidelberg, v.1537, 1998. p.26-40. (Lecture Notes in Computer Science).

SOCIETY, M. C. Militares utilizando mocap em treinamento. 2011. Disponível em: <<http://www.motioncapturesociety.com/resources/world-records>>. Acesso em: 29/11/2011.

STEELE, K. et al. Markerless versus marker-based motion capture - a comparison of measured joint centres. Annual Meeting (NACOB), 2008, Ann-Arbor.

SUN, S. et al. Adaptive sensor data fusion in motion capture. Information Fusion, 2010, Edinburgh. p.1-8.

SUNDARESAN, A.; CHELLAPPA, R. **Markerless Motion Capture using Multiple Cameras**. Proceedings of the Computer Vision for Interactive and Intelligent Environment: IEEE Computer Society: 15-26 p. 2005.

SUTHERLAND, D. H. The evolution of clinical gait analysis. Part II kinematics. **Gait & Posture**, v. 16, n. 2, p. 159-79, Oct 2002.

\_\_\_\_\_. The evolution of clinical gait analysis part III--kinetics and energy assessment. **Gait & Posture**, v. 21, n. 4, p. 447-61, Jun 2005.

SVOBODA, T.; MARTINEC, D.; PAJDLA, T. A convenient multicamera self-calibration for virtual environments. **Presence: Teleoperators and Virtual Environments**, v. 14, n. 4, p. 407-422, 2005. ISSN 1054-7460.

THEOBALT, C. et al. Performance Capture from Multi-View Video. In: RONFARD, R. e TAUBIN, G. (Ed.). **Image and Geometry Processing for 3-D Cinematography**: Springer Berlin Heidelberg, v.5, 2010. p.127-149. (Geometry and Computing). ISBN 978-3-642-12392-4.

THOMAS, M. D. et al. **Cross-Validation of an Infrared Motion Capture System and an Electromechanical Motion Capture Device**. 2007 Digital Human Modeling Conference. Seattle, WA: SAE International 2007.

THURSTON, A. J. Giovanni Borelli and the Study of Human Movement: An Historical Review. **Australian and New Zealand Journal of Surgery**, v. 69, n. 4, p. 276-288, 1999.

TRAVERS, K. H. et al. Associations between incisor and mandibular condylar movements during maximum mouth opening in humans. **Archives of Oral Biology**, v. 45, n. 4, p. 267-275, 2000. ISSN 0003-9969.

TROJE, N. F. Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. **Journal of Vision**, v. 2, n. 5, p. 17, 2002.

TSAI, R. Y. An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision. Proc. Computer Vision and Pattern Recognition (CVPR), 1986. p.364-374.

\_\_\_\_\_. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. **IEEE Journal of Robotics and Automation**, v. 3, n. 4, p. 323-344, 1987. ISSN 0882-4967.

VLASIC, D. et al. Practical motion capture in everyday surroundings. **ACM Trans. Graph.**, v. 26, n. 3, p. 35, 2007. ISSN 0730-0301.

WENG, J.; COHEN, P.; HERNIOU, M. Camera Calibration with Distortion Models and Accuracy Evaluation. **IEEE Trans. Pattern Anal. Mach. Intell.**, v. 14, n. 10, p. 965-980, 1992. ISSN 0162-8828.

WU, Y.; HU, Z. 7th Detecting Critical Configuration of Six Points. In: NARAYANAN, P.; NAYAR, S., *et al*, Asian Conference on Computer Vision, 2006, Hyderabad, India. Springer Berlin / Heidelberg. p.447-456.

XSENS. Xsens MVN. 2013. Disponível em: <<http://www.xsens.com/en/general/mvn>>. Acesso em: 20/03/2013.

YEASIN, M.; CHAUDHURI, S. Development of an Automated Image Processing System for Kinematic Analysis of Human Gait. **Real-Time Imaging**, v. 6, n. 1, p. 55-67, 2000. ISSN 1077-2014.

YOON, H.-J. et al. Kinematic study of the mandible using an electromagnetic tracking device and custom dental appliance: Introducing a new technique. **Journal of Biomechanics**, v. 39, n. 12, p. 2325-2330, 2006. ISSN 0021-9290.

ZHANG, Z. Determining the Epipolar Geometry and its Uncertainty: A Review. **International Journal of Computer Vision**, v. 27, n. 2, p. 161-195, 1998a.

\_\_\_\_\_. **A flexible new technique for camera calibration**. Redmond. 1998b. (MSR-TR-98-71)

\_\_\_\_\_. Flexible camera calibration by viewing a plane from unknown orientations. Seventh IEEE International Conference on Computer Vision, 1999, Kerkyra, Greece. IEEE. p.666-673.

\_\_\_\_\_. A flexible new technique for camera calibration. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 22, n. 11, p. 1330-1334, 2000.

ZORDAN, V. B.; HODGINS, J. K. **Motion capture-driven simulations that hit and react**. Proceedings of the 2002 ACM SIGGRAPH/Eurographics symposium on Computer animation. San Antonio, Texas: ACM: 89-96 p. 2002.

ZOU, D. et al. **Reconstructing 3D motion trajectories of particles warms by global correspondence selection**. IEEE International Conference on Computer Vision. Shanghai 2009.