



Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação
Departamento de Computação e Automação Industrial



RECUPERAÇÃO DE INFORMAÇÕES TRIDIMENSIONAIS A PARTIR DE MÚLTIPLAS IMAGENS

Autor: Marcelo Marques Gomes

Orientador: Prof. Dr. Clésio Luis Tozzi

Dissertação de Mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos para a obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: **Engenharia de Computação**.

Banca Examinadora

Prof. Dr. Clésio Luis Tozzi (presidente) — DCA/FEEC/UNICAMP
Prof. Dr. José Eduardo Cogo Castanho — FEB/DEE/UNESP
Prof. Dr. José Mario De Martino — DCA/FEEC/UNICAMP

Campinas – SP
2012

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA E ARQUITETURA - BAE - UNICAMP

G585r Gomes, Marcelo Marques
 Recuperação de informações tridimensionais a partir
 de múltiplas imagens / Marcelo Marques Gomes. --
 Campinas, SP: [s.n.], 2012.

 Orientador: Clésio Luis Tozzi.
 Dissertação de Mestrado - Universidade Estadual de
 Campinas, Faculdade de Engenharia Elétrica e de
 Computação.

 1. Automação. 2. Visão por computador. 3. Visão
 binocular. 4. Espaços projetivo. 5. Espaços métrico. I.
 Tozzi, Clésio Luis. II. Universidade Estadual de
 Campinas. Faculdade de Engenharia Elétrica e de
 Computação. III. Título.

Título em Inglês: Recovering of three-dimensional information from multiple images

Palavras-chave em Inglês: Automation, Computer vision, Binocular spaces, Projective
spaces, Metric space

Área de concentração: Engenharia de Computação

Titulação: Mestre em Engenharia Elétrica

Banca examinadora: José Eduardo Cogo Castanho, José Mario De Martino

Data da defesa: 20-01-2012

Programa de Pós Graduação: Engenharia Elétrica

COMISSÃO JULGADORA - TESE DE MESTRADO

Candidato: Marcelo Marques Gomes

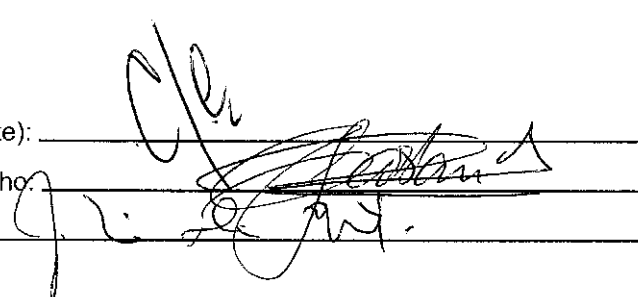
Data da Defesa: 20 de janeiro de 2012

Título da Tese: "Recuperação de Informações Tridimensionais a Partir de Múltiplas Imagens"

Prof. Dr. Clésio Luis Tozzi (Presidente): _____

Prof. Dr. José Eduardo Cogo Castanho: _____

Prof. Dr. José Mario De Martino: _____



RESUMO

Este trabalho aborda a solução para o problema da recuperação da informação tridimensional somente a partir de imagens da cena ou do objeto que se deseja modelar desconhecidas as informações a respeito das câmeras utilizadas. A solução apresentada divide-se em duas partes: na primeira se estabelece a relação entre pares ou trios de câmeras utilizando matriz fundamental ou tensor trifocal, respectivamente e a partir dessas relações obtém-se por meio de triangulação uma reconstrução tridimensional da cena ou do objeto observado em função de uma transformação projetiva arbitrária. Na segunda parte determina-se uma transformação que leva a cena ou o objeto obtidos no espaço projetivo para o espaço métrico. Essa transformação que leva o objeto recuperado no espaço projetivo para o espaço métrico é encontrada por meio de autocalibração utilizando o plano no infinito e a cônica absoluta que apresentam a propriedade de serem invariantes a rotação, translação e escala, em relação a um referencial escolhido arbitrariamente. Dependendo do número de imagens disponíveis é necessário inserir restrições nos parâmetros intrínsecos das câmeras para viabilizar o cálculo da transformação. Essas restrições são inseridas na forma de suposições a respeito dos parâmetros intrínsecos das câmeras, como pontos principais conhecidos ou constantes entre as câmeras, *skew* nulo, relação de aspecto unitária etc. Os resultados da reconstrução no espaço projetivo obtidos com o uso da matriz fundamental e do tensor trifocal foram comparados em relação ao erro de reconstrução utilizando protótipo implementado em Matlab e imagens sintéticas. A solução geral foi avaliada em relação ao erro de reprojeção, reconstrução no espaço métrico com base em imagens sintéticas e imagens reais de objetos conhecidos, utilizando um protótipo desenvolvido na plataforma Embarcadero Delphi.

Palavras-chave: Matriz Fundamental, Tensor Trifocal, Reconstrução Tridimensional, Espaço Projetivo, Espaço Métrico.

ABSTRACT

This work addresses a solution to the problem of recovering three-dimensional information from images of a scene or modeled object based only on images and without any information of the cameras parameters. The presented solution is divided in two parts: in the first part it is established the relationship between a pair or a triple of cameras using the fundamental matrix or trifocal tensor, respectively, and obtained by triangulation a three-dimensional reconstruction of the observed scene or object in function of an arbitrary projective transformation. In the second part it is determined a transformation to covert the obtained scene or object reconstruction from the projective space to the metric space. This transformation is found by auto-calibration using the plane at infinity and the absolute conic which have the property of being invariant to rotation, translation and scale in relation to an arbitrarily chosen reference frame. According to the number of available images, restrictions on the intrinsic parameters may be necessary in order to obtain a valid transformation. These restrictions are inserted in the form of assumptions about the values of the intrinsic parameters of the cameras or relations between them, like known principal points or constant values for the cameras, zero skew, unit aspect ratio and so on. Using synthetic images and a prototype implemented in Matlab, the results of reconstruction in projective space based on the fundamental matrix and trifocal tensor were compared in relation to the error of reconstruction. Based on synthetic images and real images of known objects and a prototype developed in the Embarcadero Delphi platform, the general solution was evaluated in relation to the reprojection error and the error of reconstruction in the metric space.

Keywords: Fundamental Matrix, Trifocal Tensor, Tridimensional Reconstruction, Projective Space, Metric Space.

AGRADECIMENTOS

Ao Prof. Dr. Clésio Luis Tozzi, pela orientação, pelo apoio e pelos conhecimentos transmitidos durante este trabalho, principalmente pela paciência e compreensão na demora da finalização deste trabalho.

À minha esposa Daniela, pela ajuda e pela companhia durante a longa jornada deste trabalho.

LISTA DE FIGURAS

Figura 1.1 – Interface do software Autodesk® ImageModeler (extraído de AUTODESK, 2009a)	12
Figura 1.2 – Observação do espaço por uma câmera	13
Figura 1.3 – Duas imagens de dois pontos do espaço	14
Figura 1.4 – Solução por triangulação.....	15
Figura 1.5 – Duas câmeras observando um quadrilátero (no plano xz)	17
Figura 1.6 – Objeto recuperado por triangulação	18
Figura 2.1 – Detalhes da geometria epipolar.....	22
Figura 2.2 – Tensor trifocal, mostrando os 27 elementos	27
Figura 2.3 - Projeção de uma reta em três imagens.....	27
Figura 2.4 - Geometria com três câmeras (adaptado de HARTLEY e ZISSERMAN, 2000).....	31
Figura 2.5 – Relação de incidência ponto-reta-reta.....	32
Figura 2.6 – Relação de incidência ponto-ponto-reta.....	33
Figura 2.7 – Relação de incidência ponto-ponto-ponto.....	33
Figura 3.1 – Solução por triangulação no espaço projetivo	37
Figura 3.2 – Visão em perspectiva do objeto utilizado como base da geração das imagens sintéticas.....	40
Figura 3.3 – Fluxograma dos testes em ambiente virtual	42
Figura 3.4 – Projeções do objeto em três câmeras paralelas	45
Figura 3.5 – Erro (em pixels) na reprojeção utilizando matriz fundamental com câmeras paralelas.....	46
Figura 3.6 – Erro da reconstrução no espaço euclideo utilizando matriz fundamental	46
Figura 3.7 – Erro (em pixels) na reprojeção utilizando tensor trifocal.....	47
Figura 3.8 – Erro da reconstrução no espaço euclideo utilizando tensor trifocal	47
Figura 3.9 – Projeções do objeto da Figura 3.2, em três imagens	48
Figura 3.10 – Erro (em pixels) na reprojeção utilizando matriz fundamental.....	49
Figura 3.11 – Erro da reconstrução no espaço euclideo utilizando matriz fundamental	49
Figura 3.12 – Erro (em pixels) na reprojeção utilizando tensor trifocal.....	50
Figura 3.13 – Erro da reconstrução no espaço euclideo utilizando tensor trifocal	50
Figura 3.14 – Projeções do objeto para simulação 3	51
Figura 3.15 – Resultado da reconstrução euclidea para simulação 3	51
Figura 3.16 – Erro médio das reprojeções.....	52
Figura 3.17 – Erro médio da reconstrução euclidea	53
Figura 4.1 – Objeto no espaço euclideo.....	54

Figura 4.2 – No espaço métrico, são preservados ângulos e a relação entre os comprimentos das arestas	55
Figura 4.3 – Cubo representado no espaço afim	55
Figura 4.4 – Um ponto representado no espaço projetivo por coordenadas homogêneas.....	56
Figura 4.5 – Um cubo representado no espaço projetivo. Nota-se a intersecção das linhas paralelas (do espaço euclidiano) nos pontos de fuga.....	56
Figura 4.6 – O plano no infinito é definido pelos pontos de fuga do objeto nos espaços afim ou métrico.	60
Figura 5.1 - Visão geral do protótipo	79
Figura 5.2 – Interface principal do protótipo.....	82
Figura 5.3 – Exemplo de zoom para facilitar a tarefa se selecionar pontos na imagem.....	82
Figura 5.4 – Interface que permite relacionar cena e objeto	83
Figura 5.5 – Exemplo de visualização do objeto reconstruído no espaço 3D	83
Figura 5.6 – Exemplo de objeto reprojetoado nas imagens da cena.....	84
Figura 6.1 – Cinco imagens sintéticas utilizadas no teste	86
Figura 6.2 – Erro (em pixels) na reprojeção a partir do espaço métrico	87
Figura 6.3 – Gráfico do erro na reconstrução.....	88
Figura 6.4 – Comparação entre o erro da reprojeção a partir dos espaços projetivo e métrico.....	88
Figura 6.5 – Três imagens reais de uma cena.....	89
Figura 6.6 – Identificação dos pontos utilizados como base para o cálculo do comprimento das arestas ..	90
Figura 6.7 – Resultado da reconstrução visto em ambiente 3D	92
Figura 6.8 – Imagens originais da cena.....	92
Figura 6.9 – Imagens da cena com objeto inserido	93

LISTA DE TABELAS

Tabela 1.1 – Resumo das possibilidades de reconstrução	18
Tabela 2.1 – Número de equações que cada relação de incidência é capaz de gerar (adaptado de HARTLEY e ZISSERMAN, 2000).....	34
Tabela 3.1 - Coordenadas dos pontos do objeto virtual	41
Tabela 3.2 – Parâmetros das câmeras paralelas.....	45
Tabela 3.3 – Erro (em pixels) na reprojeção utilizando matriz fundamental com câmeras paralelas	45
Tabela 3.4 – Erro da reconstrução no espaço euclidiano utilizando matriz fundamental.....	46
Tabela 3.5 – Erro (em pixels) na reprojeção utilizando tensor trifocal	46
Tabela 3.6 – Erro da reconstrução no espaço euclidiano utilizando tensor trifocal.....	47
Tabela 3.7 – Parâmetros de câmeras para a segunda simulação	48
Tabela 3.8 – Erro (em pixels) na reprojeção utilizando matriz fundamental	48
Tabela 3.9 – Erro da reconstrução no espaço euclidiano utilizando matriz fundamental.....	49
Tabela 3.10 – Erro (em pixels) na reprojeção utilizando tensor trifocal	49
Tabela 3.11 – Erro da reconstrução no espaço euclidiano utilizando tensor trifocal.....	50
Tabela 3.12 – Parâmetros de câmeras para a terceira simulação.....	51
Tabela 4.1 – Resumo das transformações, em relação a um objeto do espaço euclidiano	58
Tabela 4.2 - Quadro-resumo da determinação dos parâmetros	73
Tabela 6.1 – Parâmetros das câmeras utilizadas no teste	85
Tabela 6.2 – Erro (em pixels) na reprojeção a partir do espaço métrico	87
Tabela 6.3 – Erro na reconstrução.....	88
Tabela 6.4 – Erro (em pixels) na reprojeção com imagens reais.....	90
Tabela 6.5 – Comparação dos comprimentos das arestas do objeto da Figura 6.6	91
Tabela 6.6 – Ângulos (em graus) entre arestas do objeto real e do objeto reconstruído	91

SUMÁRIO

1. Introdução	11
1.1 Objetivos deste trabalho.....	20
1.2 Notação	20
1.3 Organização deste trabalho	21
2. Geometria entre Câmeras Descalibradas	22
2.1 Matriz fundamental.....	22
2.1.1 Cálculo da matriz fundamental.....	24
2.2 Tensor trifocal.....	26
2.2.1 Cálculo do tensor trifocal	34
3. Reconstrução no Espaço Projetivo	36
3.1 Reconstrução projetiva utilizando matriz fundamental.....	36
3.2 Reconstrução projetiva utilizando tensor trifocal	39
3.3 Reconstrução com imagens sintéticas.....	40
3.3.1 Comparação entre os resultados	52
4. Reconstrução Métrica	54
4.1 Espaço euclideano.....	54
4.2 Espaço métrico.....	55
4.3 Espaço afim.....	55
4.4 Espaço projetivo.....	56
4.5 Transformações entre espaços	57
4.6 Determinação do vetor u	60
4.7 Determinação da matriz Z	62
4.8 Determinação da transformação H	66
4.8.1 Uso de parâmetros conhecidos nas câmeras.....	72
4.8.2 Quantas imagens são necessárias para obter uma reconstrução métrica?.....	78
5. Implementação do Protótipo	79
5.1 Detalhamento do fluxograma.....	80
5.2 Interfaces do protótipo	81
6. Avaliação e Resultados	85
6.1 Imagens sintéticas	85
6.2 Imagens reais	89
7. Conclusão	94
8. Referência Bibliográfica	96

1. INTRODUÇÃO

A proposta deste trabalho é apresentar uma solução para o problema da recuperação da informação tridimensional a partir somente de um conjunto de múltiplas imagens, para as quais são conhecidas somente as correspondências entre pontos ou retas no espaço das imagens e desconhecidas quaisquer informações a respeito das câmeras (parâmetros intrínsecos e extrínsecos).

Há diversas aplicações para esse trabalho, entre elas a modelagem tridimensional de objetos, edifícios (WIEDEMANN et al., 2000 e HEMMLEB, 1999), sítios arqueológicos (POLLEFEYS et al., 2000 e DUCKE et al., 2011), cenas de crimes (GIBSON e HOWARD, 2000), realidade virtual e, mais recentemente, em videogames (IZADI et al., 2011). Muitas outras aplicações estão citadas e tratadas na literatura. Hoje em dia, há uma demanda cada vez maior para modelos 3D em computação gráfica e realidade virtual, sendo que a qualidade visual passa a ser ponto chave.

Para a solução do problema de reconstrução, existem *softwares* comerciais capazes de obter o modelo tridimensional a partir de imagens, cada um utilizando métodos e informações a respeito da cena ou das câmeras diferentes para obter a reconstrução. A seguir, são relacionados alguns desses softwares.

O 3DSOM Pro da CREATIVE DIMENSION SOFTWARE LTD (2011) exige que o objeto a ser modelado seja colocado sobre um suporte que gira o objeto em frente à câmera, de modo que a mesma câmera é utilizada durante o processo. É necessário calibrar a câmera antes de realizar a captura das imagens, por meio de um padrão de calibração fornecido juntamente com o *software*.

O Autodesk® ImageModeler da AUTODESK (2009b) e o PhotoModeler da EOS SYSTEMS INC (2011), apresentam interface com usuário intuitiva, pela qual o usuário relaciona manualmente as projeções de pontos entre as imagens numa fase inicial. A partir dessas relações, esses *softwares* determinam a geometria entre as câmeras para, numa fase posterior, restringir a topologia dos objetos na cena, que é estabelecida manualmente pelo usuário. A Figura 1.1 apresenta a interface do Autodesk® ImageModeler.

O *software* livre insight3d de MACH (2010) relaciona automaticamente pares de pontos entre as imagens (*matching* automático). A partir de poucos pontos relacionados, o *software* calcula a relação entre as câmeras para obter mais relacionamentos entre projeções nas imagens, que permite refinar a relação entre as câmeras. Esse processo permite calcular as posições das câmeras no espaço, os parâmetros intrínsecos

das câmeras e a nuvem de pontos 3D que forma a cena. A partir da nuvem de pontos, o usuário pode definir modelos poligonais com texturas obtidas a partir das próprias imagens.

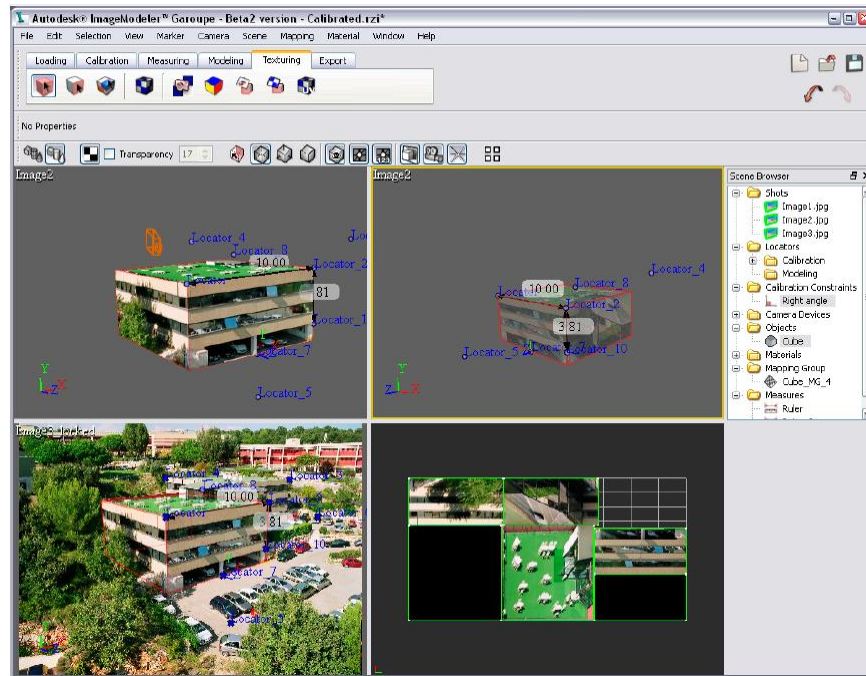


Figura 1.1 – Interface do software Autodesk® ImageModeler (extraído de AUTODESK, 2009a)

A revisão bibliográfica mostra diferentes abordagens na solução de problemas desse tipo. O sistema proposto por Debevec et al. (1996) utiliza um método híbrido, que apresenta uma parte dependente da geometria entre as câmeras e a segunda parte depende do conhecimento prévio do modelo do objeto a ser reconstruído. O resultado desse sistema é um modelo fotorrealista, com aplicação de texturas obtidas a partir das imagens do objeto.

Cippola et al. (1999) propuseram uma abordagem capaz de determinar a orientação das câmeras e os pontos principais a partir dos pontos de fuga das imagens. O trabalho de Yoon et al. (2002) segue na mesma linha, mas propõe a determinação de todos os parâmetros de câmeras a partir dos pontos de fuga. As desvantagens dessa abordagem estão na dificuldade de se determinar os pontos de fuga das imagens e do erro inserido nesse processo.

Os trabalhos de Pollefeys (POLLEFEYS et al. (1997); POLLEFEYS et al. (1998) e POLLEFEYS et al. (2000)), que diferem entre si pelo número de restrições impostas aos parâmetros das câmeras, utilizam uma abordagem estratificada, realizando inicialmente uma reconstrução 3D em um espaço projetivo e, a seguir, encontra uma transformação capaz de elevar o nível dessa reconstrução para o espaço métrico. Em uma terceira etapa, aumenta o número de correspondências entre as imagens e refina o modelo obtido.

Essa abordagem é vantajosa pela flexibilidade atribuída ao processo, já que a reconstrução no espaço projetivo pode ser obtida por qualquer método.

Este trabalho seguirá na linha dos trabalhos de Pollefeys, em razão de não necessitar nenhuma informação a respeito do modelo do objeto (como no trabalho de Debevec et al. (1996)), não realizar operações complexas (como determinar pontos de fuga da imagem) e da liberdade de escolher o método para reconstruir o objeto no espaço projetivo.

A seguir, serão apresentados os conceitos necessários para melhor entendimento do problema, suas especificidades e encaminhamento da solução.

Uma câmera registra uma imagem de um ponto no espaço, conforme a Figura 1.2a.

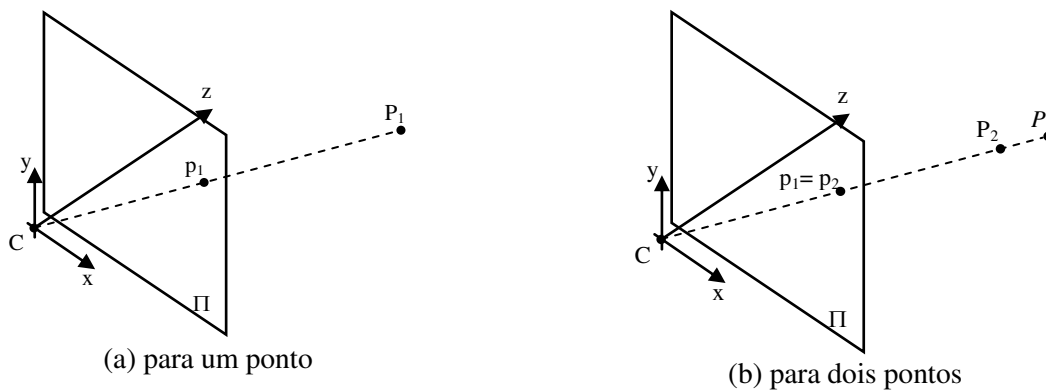


Figura 1.2 – Observação do espaço por uma câmera

A reta tracejada na Figura 1.2a liga o ponto P no espaço ao ponto C , que representa o centro óptico da câmera. Essa reta gera a projeção do ponto no plano da imagem Π , na intersecção da reta tracejada com o plano Π .

Se houver um segundo ponto sobre a reta \overline{CP} , sua projeção no plano da imagem se dará na mesma posição da projeção do ponto p_1 , conforme a Figura 1.2b.

Conforme se nota na Figura 1.2b, os pontos p_1 e p_2 (que são, respectivamente, as projeções dos pontos P_1 e P_2 do espaço) são coincidentes e, portanto, não é possível determinar qual dos dois pontos, P_1 ou P_2 , gerou a projeção. Assim, conhecendo-se somente a projeção $p_1 = p_2$, apenas é possível saber a qual reta esse ponto pertence.

A recuperação da informação tridimensional é possível se houver mais de uma imagem desses pontos. A Figura 1.3 ilustra a situação em que duas câmeras observam esses pontos.

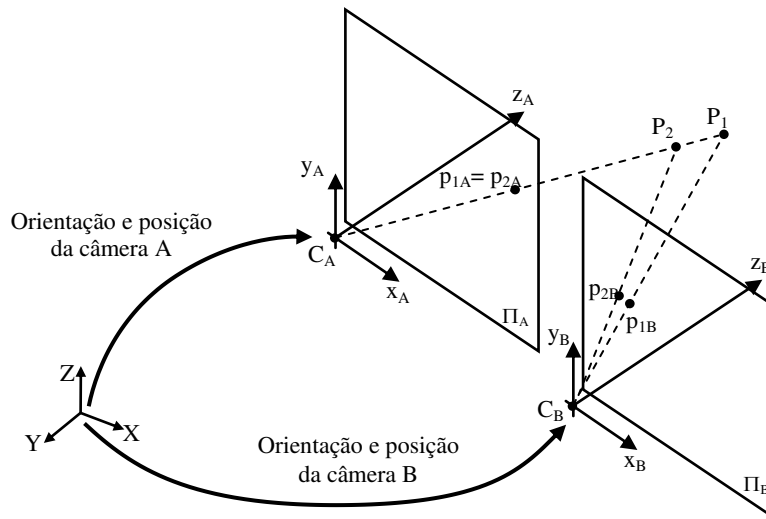


Figura 1.3 – Duas imagens de dois pontos do espaço

Com a informação de uma segunda imagem, é possível resolver a ambiguidade na reprojeção dos pontos P_1 e P_2 , já que as projeções desses pontos na imagem B não são coincidentes. Com isso, torna-se clara a necessidade de se conhecer mais de uma imagem para recuperar a informação tridimensional do objeto.

Se forem conhecidos os parâmetros intrínsecos das câmeras, então são conhecidos os centros ópticos das câmeras e, a partir deles, é possível determinar, por exemplo, as retas $\overline{C_A P_{1A}}$ e $\overline{C_B P_{1B}}$. A intersecção dessas retas se dará no ponto P_1 , o que pode ser obtido conhecendo-se a posição e a orientação dessas câmeras no espaço. Esse procedimento é conhecido como triangulação.

A recuperação da informação tridimensional de uma cena ou de um objeto pode ser resolvida de formas diversas, dependendo das informações que forem conhecidas a respeito das câmeras.

A recuperação pode ser feita de forma simples quando são conhecidos todos os parâmetros extrínsecos das câmeras (que determinam a posição das câmeras no espaço – rotação e translação da câmera em relação ao referencial absoluto ou em relação a uma das câmeras) e os parâmetros intrínsecos (que são os parâmetros internos das câmeras, como distância focal, relação de aspecto, posição do ponto principal etc.). Nessa situação, conhecida na literatura como caso calibrado, a relação entre referenciais é conhecida permitindo a transformação entre referenciais e a solução se dá por triangulação. A Figura 1.4 ilustra essa condição em que são conhecidas as projeções p_A e p_B e todos os parâmetros das câmeras, é possível determinar as retas l^A e l^B no mesmo sistema de referência. O ponto P , no espaço, está na intersecção das duas retas l^A e l^B .

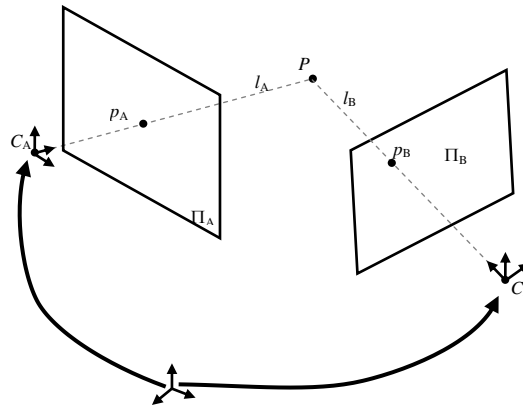


Figura 1.4 – Solução por triangulação

Essa solução exige a calibração prévia das câmeras a fim de obter todos os seus parâmetros intrínsecos e extrínsecos. Um procedimento de calibração de câmera, de modo geral, exige o conhecimento da posição tridimensional de um determinado número de pontos e de suas projeções na imagem. Estão disponíveis na literatura diversas técnicas de calibração, como por exemplo, a técnica proposta por Tsai (1987).

A calibração de câmeras não se aplica na solução que se busca neste trabalho. Para câmeras não calibradas, supõe-se a existência de três situações distintas:

- as câmeras não estão disponíveis para calibração, no caso de, por exemplo, imagens históricas ou fotografias obtidas por terceiros;
- os parâmetros das câmeras são parcialmente conhecidos, quando se conhecem as câmeras utilizadas, mas não se conhecem sua posição e orientação;
- os parâmetros intrínsecos das câmeras não são conhecidos mas são constantes entre as imagens, tal como na situação em que a mesma câmera foi utilizada para gerar as imagens (MAYBANK e FAUGERAS, 1992).

No caso em que, para um conjunto de imagens, é possível a determinação de pontos correspondentes entre elas, a abordagem pela matriz fundamental ou pelo tensor trifocal permite relacionar as câmeras e permite extrair informações a respeito da geometria das câmeras.

A matriz fundamental F (LUONG e FAUGERAS, 1995; ZHANG e KANADE, 1998 e outros) é uma matriz 3×3 que relaciona as projeções dos pontos em coordenadas de imagens, ou seja, realiza uma transformação que leva uma projeção do plano da imagem A para o plano da imagem B.

De maneira análoga, o tensor trifocal (HARTLEY, 1994; SHASHUA e WERMAN, 1995 e outros) permite relacionar projeções em três imagens diferentes. O tratamento pelo tensor trifocal não se restringe ao relacionamento entre projeções de pontos, como no caso da matriz fundamental, mas permite também o relacionamento de projeções de retas.

A matriz fundamental e o tensor trifocal serão detalhados no Capítulo 2 deste trabalho.

Conhecida a matriz fundamental ou o tensor trifocal e dependendo do conjunto de parâmetros conhecidos ou das restrições impostas sobre as câmeras, é possível recuperar a geometria do conjunto de câmeras e, a partir daí, reconstruir a cena ou o objeto no espaço 3D. Dois casos podem ser descritos:

- quando se conhecem os parâmetros intrínsecos das câmeras;
- quando não se conhecem os parâmetros intrínsecos das câmeras.

No caso em que são conhecidos apenas os parâmetros intrínsecos das câmeras, descrito na literatura como caso parcialmente calibrado, não é possível realizar a transferência das retas suportes dos pontos reprojatados no espaço 3D de uma câmera para outra para que seja possível realizar a triangulação, pois os parâmetros extrínsecos não são conhecidos.

A partir da matriz fundamental ou do tensor trifocal, em conjunto com os parâmetros intrínsecos das câmeras, é possível calcular a rotação relativa entre as câmeras e o sentido da translação entre as câmeras, mas não é possível determinar a exata posição relativa entre as câmeras.

A Figura 1.5a ilustra a situação descrita e permite compreender melhor o resultado da reconstrução de um objeto no espaço quando não se conhece o valor absoluto da translação. Nessa figura, duas câmeras paralelas observam o quadrilátero $P_1P_2P_3P_4$ do espaço. Esse quadrilátero projeta os pontos p_{1A} , p_{2A} , p_{3A} e p_{4A} na câmera A e os pontos p_{1B} , p_{2B} , p_{3B} e p_{4B} na câmera B. Na Figura 1.5, para facilitar o entendimento, a análise é feita somente no plano xz.

Mantendo-se as coordenadas das projeções nos referenciais de câmeras e alterando o valor absoluto do vetor de base, obtém-se a reconstrução apresentada na Figura 1.5b que destaca o objeto reconstruído e sua relação com o objeto original. Da figura, nota-se que a variação do comprimento da linha de base na reconstrução resulta no objeto formado pelos pontos P'_1 a P'_4 , que possui a mesma forma do objeto original mas não possui as mesmas dimensões.

Como o objeto reconstruído não guarda a mesma escala do objeto original diz-se que o objeto foi reconstruído no espaço métrico e não no espaço euclidiano, como no caso em que as câmeras são calibradas.

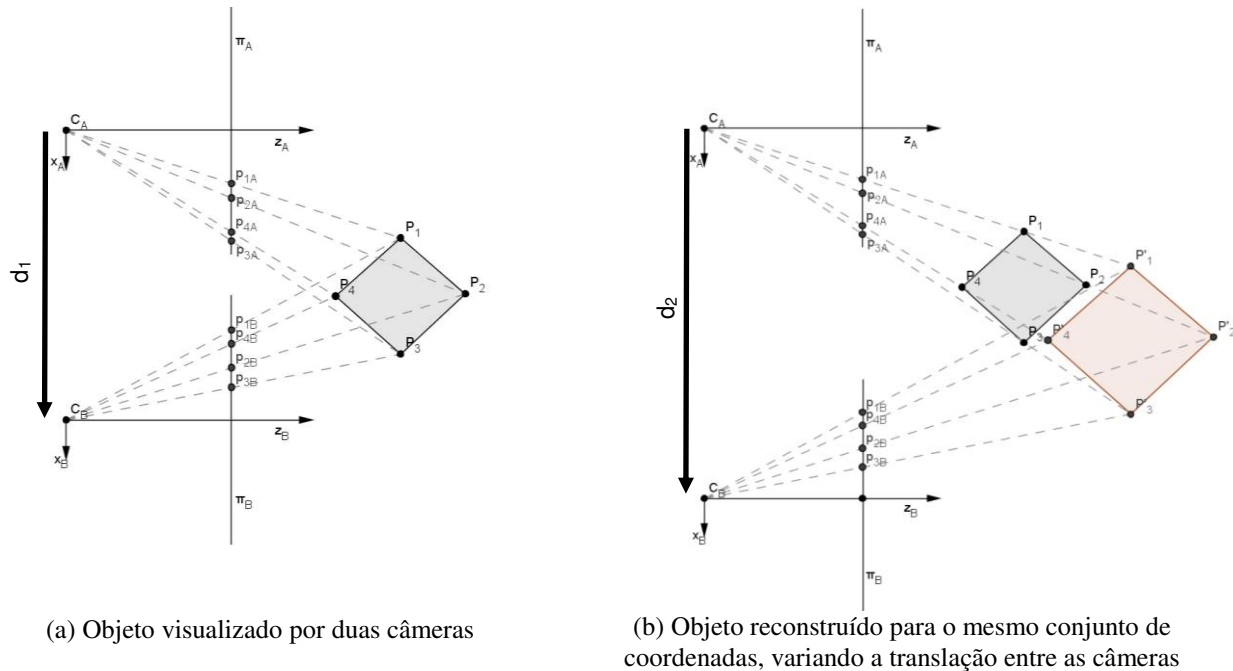


Figura 1.5 – Duas câmeras observando um quadrilátero (no plano xz)

O caso mais geral se apresenta quando também não são conhecidos os parâmetros intrínsecos das câmeras, que é conhecido na literatura como caso descalibrado.

No caso descalibrado, a reconstrução de um objeto por meio da matriz fundamental ou do tensor trifocal exige que sejam arbitrados também os parâmetros intrínsecos, o que equivale à definição de uma transformação projetiva arbitrária.

Para ilustrar o efeito dessa escolha arbitrária, a Figura 1.6a mostra o efeito da reconstrução quando um dos parâmetros internos, como por exemplo a distância focal, é arbitrado.

Nota-se que no objeto recuperado, formado pelos pontos P'_1 a P'_4 , não foram preservados o paralelismo e os ângulos internos do objeto original. Neste caso, diz-se que a reconstrução ocorreu no espaço projetivo.

De modo geral, a reconstrução resulta na arbitração dos parâmetros internos e no valor absoluto do vetor de base, resultando em uma reconstrução conforme a Figura 1.6b.

A Tabela 1.1 resume os três casos de reconstrução, conforme as informações disponíveis a respeito das câmeras, o espaço de reconstrução e as características preservadas do objeto original.

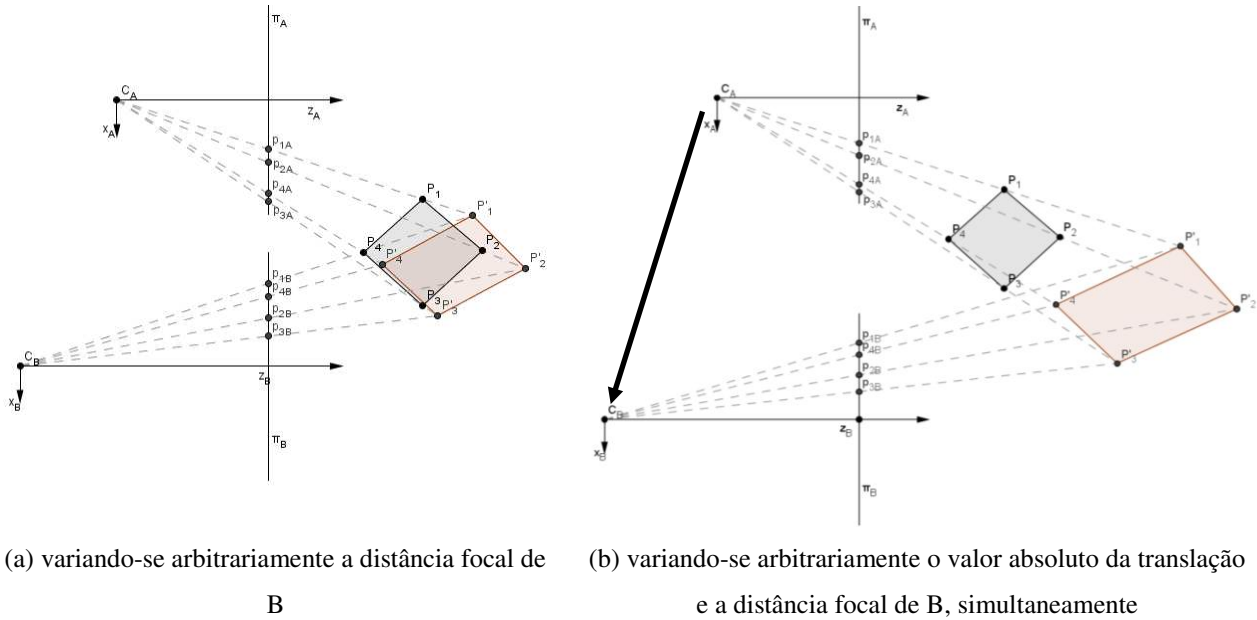
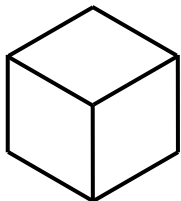
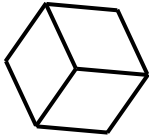



Figura 1.6 – Objeto recuperado por triangulação

Tabela 1.1 – Resumo das possibilidades de reconstrução

Informações das câmeras	Espaço de reconstrução	Características da reconstrução
Todos os parâmetros são conhecidos (caso calibrado)	Euclideano 	Preserva: <ul style="list-style-type: none"> • concorrência entre retas • paralelismo • ângulos entre arestas • distâncias absolutas
Apenas parâmetros intrínsecos são conhecidos (caso parcialmente calibrado)	Métrico 	Preserva: <ul style="list-style-type: none"> • concorrência entre retas • paralelismo • ângulos entre arestas • distâncias relativas
Nenhum parâmetro é conhecido (caso descalibrado)	Projetivo 	Preserva: <ul style="list-style-type: none"> • concorrência entre retas

No caso descalibrado, que resulta em uma reconstrução no espaço projetivo, se estiverem disponíveis número suficiente de imagens, é possível obter uma transformação capaz de levar o objeto do espaço projetivo para o métrico.

Se alguma dimensão do objeto ou da cena for conhecida, será possível então levar o objeto do espaço métrico para o espaço euclidiano.

O objeto recuperado no espaço métrico pode ser obtido por meio da aplicação, ao objeto recuperado no espaço projetivo, de uma transformação do tipo (HARTLEY e ZISSERMAN, 2000):

$$\mathbf{H} = \begin{bmatrix} \mathbf{K}^{-1} & \mathbf{0}_{3 \times 1} \\ \Pi_{\infty}^T & d \end{bmatrix} \quad (1.01)$$

em que \mathbf{K} é uma matriz 3×3 que representa os parâmetros intrínsecos da câmera adotada como referência, $\mathbf{0}_{3 \times 1}$ representa um vetor coluna nulo, $\Pi_{\infty}^T = [a \quad b \quad c]$ e d representam o plano formado pelos pontos de fuga no espaço projetivo.

Na equação (1.01), a transformação \mathbf{H} é definida pelos seis parâmetros de \mathbf{K} , pelos três parâmetros de Π_{∞}^T e pelo parâmetro d , totalizando dez parâmetros a serem determinados.

Como será visto no Capítulo 4, é possível estabelecer uma equação nos parâmetros da matriz \mathbf{H} . Com isso, torna-se necessário conhecer pelo menos 9 pares de imagens para determinar completamente a matriz \mathbf{H} .

O número de pares de imagens pode ser reduzido introduzindo-se suposições ou restrições a respeito dos parâmetros intrínsecos das câmeras. Podem ser utilizadas como restrições para esse problema, por exemplo:

- relações de aspecto conhecidas (ou desconhecidas, porém constantes) entre as câmeras;
- pontos principais conhecidos (ou constantes);
- distâncias focais conhecidas (ou constantes);
- *skew* conhecido.

De acordo com Pollefeys et al. (2000), é necessário impor ao menos uma restrição para que se possa determinar a transformação \mathbf{H} , por exemplo *skew* nulo.

As suposições feitas a respeito das câmeras aumentam o número de equações por par de imagens, no caso de parâmetros supostamente constantes, ou diminuem o número de variáveis a serem determinadas, no caso de parâmetros conhecidos.

A determinação da transformação que permite levar o objeto do espaço projetivo para o métrico, bem como o número de imagens necessárias para diversas situações, será discutida no Capítulo 4.

1.1 *Objetivos deste trabalho*

O presente trabalho tem como objetivos:

- comparação e discussão das abordagens pela matriz fundamental e pelo tensor trifocal na determinação da reconstrução;
- estudo e implementação da transformação da reconstrução do espaço projetivo para o espaço métrico e avaliação de suas implicações;
- avaliação dos resultados obtidos pelas técnicas estudadas, utilizando imagens sintéticas no ambiente MATLAB e obtenção de subsídios para implementação de um protótipo em linguagem de programação;
- implementação do protótipo de um *software* semiautomático capaz de recuperar a informação tridimensional de objetos ou edificações simples, com base em diversas imagens da mesma cena, gerando visualização fotorrealista do objeto recuperado;
- avaliação de resultados da reconstrução para imagens reais.

1.2 *Notação*

Ao longo deste trabalho, será utilizada a notação a seguir para descrever matrizes, vetores, pontos e grandezas escalares:

- matrizes serão representadas por caracteres em negrito, como por exemplo **H** ou **K**. Quando no texto não estiverem claras as dimensões da matriz, a matriz receberá indicativo de suas dimensões em subscrito, como por exemplo $\mathbf{I}_{3 \times 3}$.
- vetores e pontos serão representados com caracteres em itálico. Para pontos ou vetores representados no espaço bidimensional, serão utilizados caracteres em minúsculo, enquanto pontos ou vetores no espaço tridimensional serão representados em maiúsculo, como por exemplo p ou P , respectivamente. Quando necessário, pontos ou vetores receberão indicativo de suas dimensões em subscrito, como por exemplo $o_{3 \times 1}$.
- grandezas escalares não receberão formatação.

1.3 Organização deste trabalho

Os demais capítulos dessa dissertação estão organizados da seguinte maneira:

- o Capítulo 2 apresenta a matriz fundamental e o tensor trifocal como ferramentas para a obtenção da reconstrução no espaço projetivo. São apresentados e discutidos métodos para determinação da matriz fundamental e do tensor trifocal;
- o Capítulo 3 apresenta os métodos para obter a reconstrução no espaço projetivo utilizando a matriz fundamental e o tensor trifocal. Apresenta também uma comparação entre a matriz fundamental e o tensor trifocal na obtenção da reconstrução projetiva realizada no ambiente MATLAB que justifica a escolha do modelo do protótipo;
- o Capítulo 4 apresenta uma breve discussão a respeito da caracterização dos espaços de reconstrução e do método capaz de determinar a transformação projetiva que permite levar o objeto reconstruído do espaço projetivo para o espaço métrico, bem como suas implicações;
- o Capítulo 5 descreve o protótipo e detalha sua implementação;
- o Capítulo 6 apresenta os resultados obtidos pela aplicação do protótipo a imagens reais em duas cenas diferentes. Cada cena foi modelada com um conjunto de suposições a respeito das câmeras e são avaliadas as influências das suposições na reconstrução;
- por fim, no Capítulo 7 são apresentadas as conclusões e considerações finais deste trabalho e sugestões para trabalhos futuros.

2. GEOMETRIA ENTRE CÂMERAS DESCALIBRADAS

Quando não se conhece nenhuma informação a respeito das câmeras ou do objeto observado por duas ou mais câmeras, podem-se relacionar as imagens utilizando a matriz fundamental ou o tensor trifocal.

Neste Capítulo serão apresentados os conceitos associados à matriz fundamental e ao tensor trifocal e métodos para determiná-los.

2.1 *Matriz fundamental*

Um ponto P do espaço tem projeção $p_A = [x_A \ y_A \ w_A]$ na câmera A e $p_B = [x_B \ y_B \ w_B]$ na câmera B, conforme a apresentado na Figura 2.1. Nessa figura, a câmera A possui centro óptico em C_A e plano de projeção dado por Π_A . A câmera B tem centro óptico em C_B e plano de projeção Π_B .

O plano formado pelos centros ópticos das câmeras e pelo ponto P é denominado plano epipolar. Ainda na Figura 2.1, o ponto e_{AB} , que representa a projeção de C_A na imagem B é denominado epipolo de B, enquanto e_{BA} é o epipolo de A.

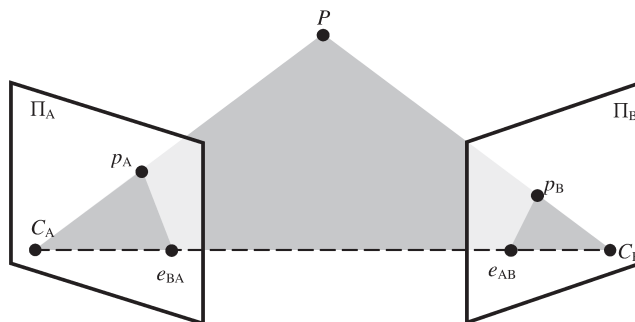


Figura 2.1 – Detalhes da geometria epipolar

A geometria epipolar pode ser modelada pela matriz fundamental, que é uma matriz 3×3 capaz de relacionar as projeções dos pontos em coordenadas de imagens, de modo que (LUONG e FAUGERAS, 1995; ZHANG e KANADE, 1998 e outros):

$$p_B^T \mathbf{F} p_A = 0 \quad (2.01)$$

em que \mathbf{F} é denominada matriz fundamental, p_A representa as coordenadas homogêneas da projeção de um ponto P do espaço 3D na imagem A e p_B as coordenadas homogêneas da projeção de P na imagem B.

A partir da Figura 2.1, pode-se derivar o equacionamento da matriz fundamental a partir de uma relação entre as coordenadas das projeções de um ponto nas duas imagens. Tomando a câmera A como referência (de modo que essa câmera defina o referencial absoluto), pode-se escrever:

$$\begin{cases} p_A = \mathbf{K}_A [\mathbf{I}_{3 \times 3} \mid \mathbf{0}_{3 \times 1}] P \\ p_B = \mathbf{K}_B [\mathbf{R}_{3 \times 3} \mid \mathbf{t}_{3 \times 1}] P \end{cases} \quad (2.02)$$

em que $\mathbf{I}_{3 \times 3}$ representa a matriz identidade, $\mathbf{0}_{3 \times 1}$ é um vetor nulo, \mathbf{R} é a matriz de rotação entre a câmera A e a câmera B, \mathbf{t} é o vetor de translação entre as câmeras, \mathbf{K}_A e \mathbf{K}_B são, respectivamente, as matrizes 3×3 dos parâmetros intrínsecos das câmeras A e B, e P é um ponto no espaço 3D em coordenadas homogêneas.

Reescrevendo as equações (2.02):

$$\begin{cases} P = (\mathbf{K}_A)^{-1} p_A \\ p_B = \mathbf{K}_B \mathbf{R} P + \mathbf{K}_B \mathbf{t} \end{cases}$$

Com o intuito de determinar a relação entre p_A e p_B , substitui-se a primeira equação na segunda:

$$p_B = \mathbf{K}_B \mathbf{R} (\mathbf{K}_A^{-1} p_A) + \mathbf{K}_B \mathbf{t}$$

Trazendo \mathbf{K}_B para o primeiro termo:

$$\mathbf{K}_B^{-1} p_B = \mathbf{R} (\mathbf{K}_A^{-1} p_A) + \mathbf{t}$$

Para eliminar a soma do segundo termo, pode-se pré-multiplicar ambos os termos pela matriz antissimétrica de \mathbf{t} , $[\mathbf{t}]_{\times}$. Como:

$$[\mathbf{t}]_{\times} \mathbf{t} = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} = \begin{bmatrix} -t_3 t_2 + t_2 t_3 \\ t_3 t_1 - t_1 t_3 \\ -t_2 t_1 + t_1 t_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (2.03)$$

então:

$$[t]_{\times}(\mathbf{K}_B^{-1} p_B) = [t]_{\times}(\mathbf{R}(\mathbf{K}_A^{-1} p_A))$$

O primeiro termo pode ser anulado multiplicando-se ambos os termos da equação por $(\mathbf{K}_B^{-1} p_B)^T$:

$$(\mathbf{K}_B^{-1} p_B)^T [t]_{\times}(\mathbf{K}_B^{-1} p_B) = (\mathbf{K}_B^{-1} p_B)^T [t]_{\times}(\mathbf{R}(\mathbf{K}_A^{-1} p_A))$$

Como $(\mathbf{K}_B^{-1} p_B)^T [t]_{\times}(\mathbf{K}_B^{-1} p_B) = 0$, pode-se reescrever:

$$0 = p_B^T (\mathbf{K}_B^{-T} [t]_{\times} \mathbf{R} \mathbf{K}_A^{-1}) p_A$$

Chamando $(\mathbf{K}_B^{-T} [t]_{\times} \mathbf{R} \mathbf{K}_A^{-1})$ de \mathbf{F} , teremos então:

$$p_B^T \mathbf{F} p_A = 0$$

em que \mathbf{F} é a matriz fundamental.

Os epipolos e_{AB} e e_{BA} podem ser determinados a partir da matriz fundamental, por meio das equações (TRUCCO e VERRI, 1998):

$$\begin{aligned} \mathbf{F} e_{BA} &= 0 \\ \mathbf{F}^T e_{AB} &= 0 \end{aligned} \tag{2.04}$$

2.1.1 Cálculo da matriz fundamental

Existem diversas formas de determinar os elementos da matriz fundamental. Todos eles necessitam que as projeções dos pontos nas imagens sejam relacionadas.

Expandindo os termos da equação (2.01):

$$\begin{bmatrix} x_B & y_B & w_B \end{bmatrix} \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} x_A \\ y_A \\ w_A \end{bmatrix} = 0$$

Realizando a multiplicação e isolando os termos em \mathbf{F} :

$$\begin{bmatrix} u_A u_B & u_A v_B & u_A w_B & v_A u_B & v_A v_B & v_A w_B & w_A u_B & w_A v_B & w_A w_B \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix} = 0 \quad (2.05)$$

Para a determinação das 9 variáveis da equação (2.05) é necessário conhecer a relação entre as projeções de pelo menos nove pontos entre as duas imagens para que seja possível escrever um sistema com nove equações:

$$\begin{bmatrix} u_{A_1} u_{B_1} & u_{A_1} v_{B_1} & u_{A_1} w_{B_1} & v_{A_1} u_{B_1} & v_{A_1} v_{B_1} & v_{A_1} w_{B_1} & w_{A_1} u_{B_1} & w_{A_1} v_{B_1} & w_{A_1} w_{B_1} \\ u_{A_2} u_{B_2} & u_{A_2} v_{B_2} & u_{A_2} w_{B_2} & v_{A_2} u_{B_2} & v_{A_2} v_{B_2} & v_{A_2} w_{B_2} & w_{A_2} u_{B_2} & w_{A_2} v_{B_2} & w_{A_2} w_{B_2} \\ u_{A_3} u_{B_3} & u_{A_3} v_{B_3} & u_{A_3} w_{B_3} & v_{A_3} u_{B_3} & v_{A_3} v_{B_3} & v_{A_3} w_{B_3} & w_{A_3} u_{B_3} & w_{A_3} v_{B_3} & w_{A_3} w_{B_3} \\ u_{A_4} u_{B_4} & u_{A_4} v_{B_4} & u_{A_4} w_{B_4} & v_{A_4} u_{B_4} & v_{A_4} v_{B_4} & v_{A_4} w_{B_4} & w_{A_4} u_{B_4} & w_{A_4} v_{B_4} & w_{A_4} w_{B_4} \\ u_{A_5} u_{B_5} & u_{A_5} v_{B_5} & u_{A_5} w_{B_5} & v_{A_5} u_{B_5} & v_{A_5} v_{B_5} & v_{A_5} w_{B_5} & w_{A_5} u_{B_5} & w_{A_5} v_{B_5} & w_{A_5} w_{B_5} \\ u_{A_6} u_{B_6} & u_{A_6} v_{B_6} & u_{A_6} w_{B_6} & v_{A_6} u_{B_6} & v_{A_6} v_{B_6} & v_{A_6} w_{B_6} & w_{A_6} u_{B_6} & w_{A_6} v_{B_6} & w_{A_6} w_{B_6} \\ u_{A_7} u_{B_7} & u_{A_7} v_{B_7} & u_{A_7} w_{B_7} & v_{A_7} u_{B_7} & v_{A_7} v_{B_7} & v_{A_7} w_{B_7} & w_{A_7} u_{B_7} & w_{A_7} v_{B_7} & w_{A_7} w_{B_7} \\ u_{A_8} u_{B_8} & u_{A_8} v_{B_8} & u_{A_8} w_{B_8} & v_{A_8} u_{B_8} & v_{A_8} v_{B_8} & v_{A_8} w_{B_8} & w_{A_8} u_{B_8} & w_{A_8} v_{B_8} & w_{A_8} w_{B_8} \\ u_{A_9} u_{B_9} & u_{A_9} v_{B_9} & u_{A_9} w_{B_9} & v_{A_9} u_{B_9} & v_{A_9} v_{B_9} & v_{A_9} w_{B_9} & w_{A_9} u_{B_9} & w_{A_9} v_{B_9} & w_{A_9} w_{B_9} \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{bmatrix} = 0 \quad (2.06)$$

Como o sistema descrito na equação (2.06) é homogêneo, pode-se determinar sua solução em função de um de seus parâmetros (digamos, f_{33}), reduzindo o número de incógnitas para oito, conforme a equação (2.07).

$$\begin{bmatrix} u_{A_1} u_{B_1} & u_{A_1} v_{B_1} & u_{A_1} w_{B_1} & v_{A_1} u_{B_1} & v_{A_1} v_{B_1} & v_{A_1} w_{B_1} & w_{A_1} u_{B_1} & w_{A_1} v_{B_1} \\ u_{A_2} u_{B_2} & u_{A_2} v_{B_2} & u_{A_2} w_{B_2} & v_{A_2} u_{B_2} & v_{A_2} v_{B_2} & v_{A_2} w_{B_2} & w_{A_2} u_{B_2} & w_{A_2} v_{B_2} \\ u_{A_3} u_{B_3} & u_{A_3} v_{B_3} & u_{A_3} w_{B_3} & v_{A_3} u_{B_3} & v_{A_3} v_{B_3} & v_{A_3} w_{B_3} & w_{A_3} u_{B_3} & w_{A_3} v_{B_3} \\ u_{A_4} u_{B_4} & u_{A_4} v_{B_4} & u_{A_4} w_{B_4} & v_{A_4} u_{B_4} & v_{A_4} v_{B_4} & v_{A_4} w_{B_4} & w_{A_4} u_{B_4} & w_{A_4} v_{B_4} \\ u_{A_5} u_{B_5} & u_{A_5} v_{B_5} & u_{A_5} w_{B_5} & v_{A_5} u_{B_5} & v_{A_5} v_{B_5} & v_{A_5} w_{B_5} & w_{A_5} u_{B_5} & w_{A_5} v_{B_5} \\ u_{A_6} u_{B_6} & u_{A_6} v_{B_6} & u_{A_6} w_{B_6} & v_{A_6} u_{B_6} & v_{A_6} v_{B_6} & v_{A_6} w_{B_6} & w_{A_6} u_{B_6} & w_{A_6} v_{B_6} \\ u_{A_7} u_{B_7} & u_{A_7} v_{B_7} & u_{A_7} w_{B_7} & v_{A_7} u_{B_7} & v_{A_7} v_{B_7} & v_{A_7} w_{B_7} & w_{A_7} u_{B_7} & w_{A_7} v_{B_7} \\ u_{A_8} u_{B_8} & u_{A_8} v_{B_8} & u_{A_8} w_{B_8} & v_{A_8} u_{B_8} & v_{A_8} v_{B_8} & v_{A_8} w_{B_8} & w_{A_8} u_{B_8} & w_{A_8} v_{B_8} \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \end{bmatrix} = \begin{bmatrix} -w_{A_1} v_{B_1} \\ -w_{A_2} v_{B_2} \\ -w_{A_3} v_{B_3} \\ -w_{A_4} v_{B_4} \\ -w_{A_5} v_{B_5} \\ -w_{A_6} v_{B_6} \\ -w_{A_7} v_{B_7} \\ -w_{A_8} v_{B_8} \end{bmatrix} \quad (2.07)$$

O sistema de equações expresso pela equação (2.07) tem solução única, entretanto a precisão da solução depende da precisão das coordenadas das projeções. Uma forma de minimizar o erro é utilizar mais pontos no equacionamento, com o conseqüente aumento de equações, para o mesmo número de incógnitas. Assim, o sistema será super especificado e a solução não será mais exata, mas poderá ser resolvido utilizando métodos de minimização.

O sistema da equação (2.07) pode se tornar mal condicionado (BARROSO et al., 1987) devido à grande variabilidade dos dados, o que causa instabilidade numérica na solução. Técnicas de normalização permitem contornar esse problema e determinar uma solução com erro menor.

De acordo com Hartley (1997), é possível normalizar as coordenadas das projeções a fim de melhorar o condicionamento do sistema de equações. Dois tipos de normalização podem ser utilizados: normalização isotrópica e normalização não-isotrópica:

- a normalização isotrópica translada os pontos da imagem de modo que o centroide da imagem coincida com a origem e em seguida aplica uma escala aos pontos de modo que a distância média entre cada ponto e a origem seja igual a $\sqrt{2}$;
- a normalização não-isotrópica também aplica a translação aos pontos, porém ao invés de normalizar a distância ao centroide, normaliza os momentos principais.

Ainda de acordo com Hartley (1997), a aplicação de técnicas lineares com normalização leva a resultados tão precisos quanto os de métodos não-lineares propostos na literatura.

2.2 *Tensor trifocal*

De maneira análoga à matriz fundamental, o tensor trifocal (HARTLEY, 1994; SHASHUA e WERMAN, 1995 e outros) permite relacionar projeções entre três imagens.

O tensor trifocal pode ser definido, de forma simplificada, como uma matriz $3 \times 3 \times 3$, com total de 27 elementos, conforme ilustrado na Figura 2.2.

Na notação tensorial, em \mathbf{T}_i^{jk} os índices j e k (sobrescritos) são chamados de contravariante, enquanto i (subscrito) é denominado covariante. Neste trabalho, optou-se em utilizar notação matricial para o tensor trifocal, para simplificar o entendimento. Assim, o tensor trifocal neste trabalho será tratado como três matrizes 3×3 $\mathbf{T} = \{\mathbf{T}_1 \quad \mathbf{T}_2 \quad \mathbf{T}_3\}$ conforme a Figura 2.2.

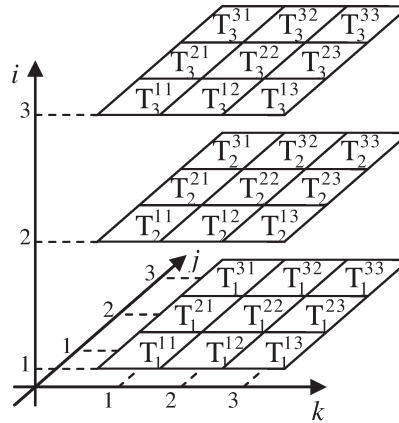


Figura 2.2 – Tensor trifocal, mostrando os 27 elementos

Na Figura 2.3, há um exemplo da projeção de uma reta L em três imagens. Nesse caso, a projeção dessa reta na imagem da câmera A se dá pela intersecção do plano formado pela reta L e o centro óptico da câmera A, com o plano da imagem. A mesma construção é válida para a projeção de L nas imagens de B e C.

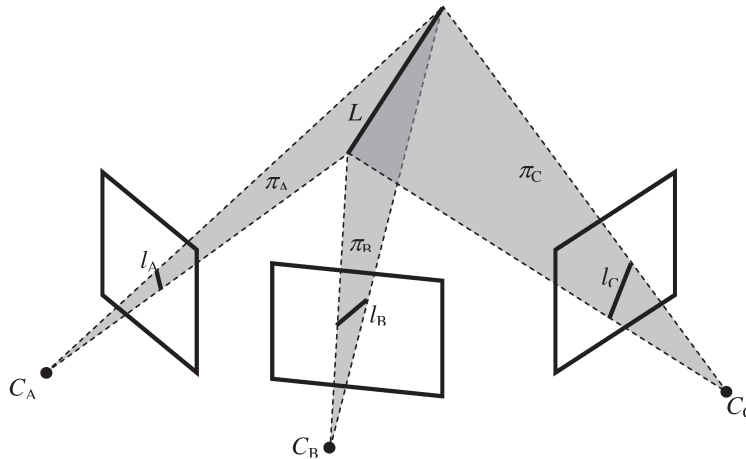


Figura 2.3 - Projeção de uma reta em três imagens

A partir da análise da Figura 2.3 é possível obter um equacionamento para o tensor trifocal, conforme o proposto por Hartley e Zisserman (2000) e por Forsyth e Ponce (2002). Escolhendo a câmera A como referência, as câmeras A, B e C da Figura 2.3 podem ser representadas pelas matrizes \mathbf{M}_A , \mathbf{M}_B e \mathbf{M}_C :

$$\begin{cases} \mathbf{M}_A = [\mathbf{I}_{3 \times 3} \mid \mathbf{0}_{3 \times 1}] \\ \mathbf{M}_B = [b_1 \quad b_2 \quad b_3 \quad b_4] \\ \mathbf{M}_C = [c_1 \quad c_2 \quad c_3 \quad c_4] \end{cases} \quad (2.08)$$

em que b_i e c_i , $i=1\dots 4$, são as colunas (3×1) dos modelos de câmeras. Os planos π_A , π_B e π_C da Figura 2.3 têm seus coeficientes calculados por:

$$\pi_A = \mathbf{M}_A^T l_A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} l_{1_A} \\ l_{2_A} \\ l_{3_A} \end{bmatrix} = \begin{bmatrix} l_{1_A} \\ l_{2_A} \\ l_{3_A} \\ 0 \end{bmatrix}$$

$$\pi_B = \mathbf{M}_B^T l_B = \begin{bmatrix} b_1^T \\ b_2^T \\ b_3^T \\ b_4^T \end{bmatrix} \begin{bmatrix} l_{1_B} \\ l_{2_B} \\ l_{3_B} \end{bmatrix} = \begin{bmatrix} b_1^T l_B \\ b_2^T l_B \\ b_3^T l_B \\ b_4^T l_B \end{bmatrix}$$

$$\pi_C = \mathbf{M}_C^T l_C = \begin{bmatrix} c_1^T \\ c_2^T \\ c_3^T \\ c_4^T \end{bmatrix} \begin{bmatrix} l_{1_C} \\ l_{2_C} \\ l_{3_C} \end{bmatrix} = \begin{bmatrix} c_1^T l_C \\ c_2^T l_C \\ c_3^T l_C \\ c_4^T l_C \end{bmatrix}$$

em que $l_A = [l_{1_A} \ l_{2_A} \ l_{3_A}]^T$ representa os coeficientes da projeção da reta L na imagem A, $l_B = [l_{1_B} \ l_{2_B} \ l_{3_B}]^T$ representa os coeficientes da projeção da reta L na imagem B e $l_C = [l_{1_C} \ l_{2_C} \ l_{3_C}]^T$ representa os coeficientes da projeção da reta L na imagem C, de modo que se um ponto $p = [x \ y \ 1]^T$ fizer parte da projeção l_A , então $l_A^T p = l_{1_A} x + l_{2_A} y + l_{3_A} = 0$. Equacionamentos semelhantes são válidos para as projeções l_B e l_C .

Como esses três planos são gerados a partir da mesma reta no espaço, eles não são independentes. Assim, de acordo com Forsyth e Ponce (2002), a matriz:

$$\mathbf{A} = \begin{bmatrix} \pi_A^T \\ \pi_B^T \\ \pi_C^T \end{bmatrix} = \begin{bmatrix} l_{1_A} & l_{2_A} & l_{3_A} & 0 \\ b_1^T l_B & b_2^T l_B & b_3^T l_B & b_4^T l_B \\ c_1^T l_C & c_2^T l_C & c_3^T l_C & c_4^T l_C \end{bmatrix}$$

tem *rank* igual a 2 e os determinantes das 4 submatrizes 3×3 de A são nulos. Tomando 3 destes determinantes, tem-se:

$$\begin{vmatrix} I_{2_A} & I_{3_A} & 0 \\ b_2^T l_B & b_3^T l_B & b_4^T l_B \\ c_2^T l_C & c_3^T l_C & c_4^T l_C \end{vmatrix} = 0$$

$$\begin{vmatrix} I_{1_A} & I_{3_A} & 0 \\ b_1^T l_B & b_3^T l_B & b_4^T l_B \\ c_1^T l_C & c_3^T l_C & c_4^T l_C \end{vmatrix} = 0$$

$$\begin{vmatrix} I_{1_A} & I_{2_A} & 0 \\ b_1^T l_B & b_2^T l_B & b_4^T l_B \\ c_1^T l_C & c_2^T l_C & c_4^T l_C \end{vmatrix} = 0$$

Resolvendo os determinantes:

$$\begin{aligned} I_{2_A} (b_3^T l_B c_4^T l_C - b_4^T l_B c_3^T l_C) + I_{3_A} (b_4^T l_B c_2^T l_C - b_2^T l_B c_4^T l_C) &= 0 \\ I_{1_A} (b_3^T l_B c_4^T l_C - b_4^T l_B c_3^T l_C) + I_{3_A} (b_4^T l_B c_1^T l_C - b_1^T l_B c_4^T l_C) &= 0 \\ I_{1_A} (b_2^T l_B c_4^T l_C - b_4^T l_B c_2^T l_C) + I_{2_A} (b_4^T l_B c_1^T l_C - b_1^T l_B c_4^T l_C) &= 0 \end{aligned}$$

Rearranjando para:

$$\begin{aligned} I_{2_A} (l_B^T (b_3 c_4^T - b_4 c_3^T) l_C) - I_{3_A} (l_B^T (b_2 c_4^T - b_4 c_2^T) l_C) &= 0 \\ I_{3_A} (l_B^T (b_1 c_4^T - b_4 c_1^T) l_C) - I_{1_A} (l_B^T (b_3 c_4^T - b_4 c_3^T) l_C) &= 0 \\ I_{1_A} (l_B^T (b_2 c_4^T - b_4 c_2^T) l_C) - I_{2_A} (l_B^T (b_1 c_4^T - b_4 c_1^T) l_C) &= 0 \end{aligned} \quad (2.09)$$

Como para dois vetores $\vec{x} = [x_1 \ x_2 \ x_3]^T$ e $\vec{y} = [y_1 \ y_2 \ y_3]^T$, o produto vetorial é dado por:

$$\vec{x} \times \vec{y} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = (x_2 y_3 - x_3 y_2) \vec{i} + (x_3 y_1 - x_1 y_3) \vec{j} + (x_1 y_2 - x_2 y_1) \vec{k} \quad (2.10)$$

pode-se escrever (2.09) como:

$$I_A \times \begin{bmatrix} l_B^T (b_1 c_4^T - b_4 c_1^T) l_C \\ l_B^T (b_2 c_4^T - b_4 c_2^T) l_C \\ l_B^T (b_3 c_4^T - b_4 c_3^T) l_C \end{bmatrix} = 0 \quad (2.11)$$

Chamando:

$$\begin{cases} \mathbf{T}_1 = b_1 c_4^T - b_4 c_1^T \\ \mathbf{T}_2 = b_2 c_4^T - b_4 c_2^T \\ \mathbf{T}_3 = b_3 c_4^T - b_4 c_3^T \end{cases} \quad (2.12)$$

reduz-se (2.11) a:

$$l_A \times \begin{bmatrix} l_B^T \mathbf{T}_1 l_C \\ l_B^T \mathbf{T}_2 l_C \\ l_B^T \mathbf{T}_3 l_C \end{bmatrix} = 0$$

Como o produto vetorial entre dois vetores é nulo somente se eles forem paralelos, então:

$$l_A = \begin{bmatrix} l_B^T \mathbf{T}_1 l_C \\ l_B^T \mathbf{T}_2 l_C \\ l_B^T \mathbf{T}_3 l_C \end{bmatrix} \rightarrow \begin{cases} l_{1_A} = l_B^T [\mathbf{T}_1 l_C] \\ l_{2_A} = l_B^T [\mathbf{T}_2 l_C] \\ l_{3_A} = l_B^T [\mathbf{T}_3 l_C] \end{cases} \quad (2.13)$$

em que as três matrizes \mathbf{T}_1 , \mathbf{T}_2 e \mathbf{T}_3 formam o tensor trifocal.

A equação (2.13) pode ser reescrita como:

$$l_A^T = l_B^T [\mathbf{T}_1 l_C \quad \mathbf{T}_2 l_C \quad \mathbf{T}_3 l_C] \quad (2.14)$$

em que a multiplicação de \mathbf{T}_i por l_C resulta em um vetor 3×1 , de modo que $[\mathbf{T}_1 l_C \quad \mathbf{T}_2 l_C \quad \mathbf{T}_3 l_C]$ resulta em uma matriz 3×3 .

A determinação dos epipolos a partir do tensor trifocal não é direta, como no caso da matriz fundamental e, de acordo com Hartley e Vidal (2004), a solução deve ser obtida pelo cálculo da intersecção das retas epipolares, conforme ilustrado na Figura 2.4.

Na Figura 2.4, o plano A é formado pelos pontos P , C_A e C_B (respectivamente, os centros ópticos das câmeras A e B), de modo que a intersecção desse plano com o plano da imagem A forma a reta epipolar l_{BA} e a intersecção do plano A com o plano da imagem B forma a reta epipolar l_{AB} .

O plano B é formado pelos pontos P , C_B e C_C (centros ópticos das câmeras B e C), de modo que a intersecção desse plano com o plano da imagem B forma a reta epipolar l_{BC} e a intersecção com o plano da imagem C forma a reta epipolar l_{CB} (omitida na figura).

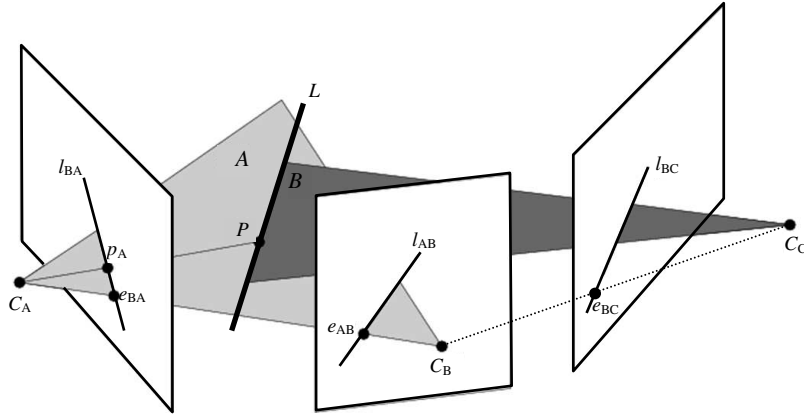


Figura 2.4 - Geometria com três câmeras (adaptado de HARTLEY e ZISSERMAN, 2000)

De acordo com Hartley e Zisserman (2000), a reta l_{AB} satisfaz a equação:

$$l_{AB}^T (x_A \mathbf{T}_1 + y_A \mathbf{T}_2 + w_A \mathbf{T}_3) = 0_{1 \times 3}$$

em que x_A , y_A e w_A são as coordenadas da projeção $p_A = [x_A \quad y_A \quad w_A]^T$, em coordenadas homogêneas, enquanto l_{BC} satisfaz a equação:

$$(x_A \mathbf{T}_1 + y_A \mathbf{T}_2 + w_A \mathbf{T}_3) l_{BC} = 0_{3 \times 1}$$

Como e_{AB} satisfaz a equação $l_{AB}^T e_{AB} = 0$ e fazendo $e_{AB} = [x_{e_{AB}} \quad y_{e_{AB}} \quad 1]^T$ e $l_{AB} = [l_{AB_1} \quad l_{AB_2} \quad l_{AB_3}]^T$:

$$\begin{bmatrix} l_{AB_1} & l_{AB_2} \end{bmatrix}^T \begin{bmatrix} x_{e_{AB}} \\ y_{e_{AB}} \end{bmatrix} = -l_{AB_3} \quad (2.15)$$

Como é necessário determinar a interseção de duas retas epipolares para encontrar o epipolo, será preciso utilizar ao menos dois pontos projetados na imagem A, gerando um sistema de equações cuja solução permite determinar $x_{e_{AB}}$ e $y_{e_{AB}}$ da equação (2.15). Se forem utilizados mais pontos, então uma solução mais robusta será obtida por minimização de erro.

Para os demais epipolos, o equacionamento é semelhante.

Quando se utiliza a matriz fundamental, é possível relacionar a projeção de um ponto em duas imagens. Utilizando o tensor trifocal, é possível relacionar projeções de pontos e/ou retas entre três imagens da mesma cena, assim o leque de possibilidades é maior do que quando se usa a matriz fundamental.

De acordo com Hartley e Zisserman (2000), podem ser definidas as seguintes relações para o tensor trifocal:

- correspondência reta-reta-reta: é representado pela equação (2.14). Nesse tipo de correspondência, o tensor trifocal relaciona a projeção de uma reta em três imagens, como na Figura 2.3, e é modelada pela equação (2.14);
- Correspondência ponto-reta-reta: nessa correspondência, o tensor trifocal permite relacionar a projeção de um ponto P (que faz parte de uma reta L) em uma imagem com a projeção de L nas outras duas imagens, conforme ilustrado na Figura 2.5. Esse tipo de correspondência é modelado pela equação:

$$l_B^T (x_A \mathbf{T}_1 + y_A \mathbf{T}_2 + w_A \mathbf{T}_3) l_C = 0$$

em que x_A , y_A e w_A são as coordenadas da projeção $p_A = [x_A \ y_A \ w_A]^T$, em coordenadas homogêneas, na imagem A . As multiplicações de x_A por \mathbf{T}_1 , y_A por \mathbf{T}_2 e w_A por \mathbf{T}_3 geram e matrizes 3×3 e, por consequência, $(x_A \mathbf{T}_1 + y_A \mathbf{T}_2 + w_A \mathbf{T}_3)$ será uma matriz 3×3 .

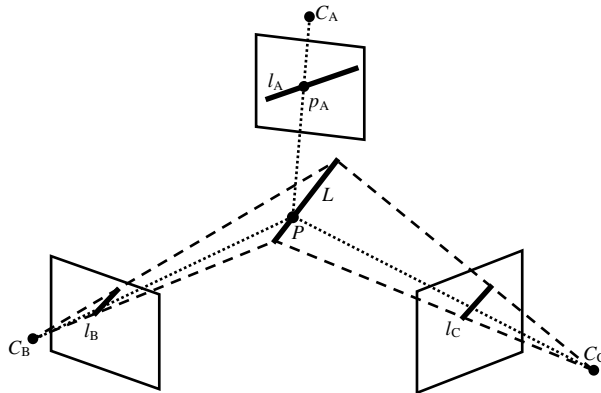


Figura 2.5 – Relação de incidência ponto-reta-reta

- Correspondência ponto-ponto-reta: nessa situação, o tensor trifocal é utilizado para relacionar a projeção de um ponto P (que faz parte de uma reta L) em duas imagens com a projeção de L na terceira imagem, conforme a Figura 2.6.

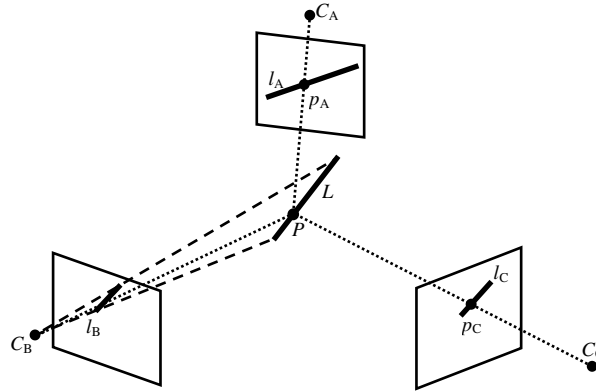


Figura 2.6 – Relação de incidência ponto-ponto-reta

Esse caso é representado pela equação:

$$[p_B]_{\times} (x_A \mathbf{T}_1 + y_A \mathbf{T}_2 + w_A \mathbf{T}_3) l_C = 0_{3 \times 1}$$

em que x_A , y_A e w_A são as coordenadas da projeção $p_A = [x_A \ y_A \ w_A]$, em coordenadas homogêneas, na imagem A e $[p_B]_{\times}$ representa a matriz antissimétrica de p_B .

- Correspondência ponto-ponto-ponto: é um caso semelhante à relação modelada pela matriz fundamental, ou seja, o tensor trifocal permite relacionar as projeções de um ponto P em três imagens, conforme ilustrado na Figura 2.7. Esse caso é representado pela equação:

$$[p_B]_{\times} (x_A \mathbf{T}_1 + y_A \mathbf{T}_2 + w_A \mathbf{T}_3) [p_C]_{\times} = \mathbf{0}_{3 \times 3} \tag{2.16}$$

em que x_A , y_A e w_A são as coordenadas da projeção $p_A = [x_A \ y_A \ w_A]$, em coordenadas homogêneas, na imagem A e $[p_B]_{\times}$ e $[p_C]_{\times}$ representam, respectivamente, as matrizes antissimétricas de p_B e p_C .

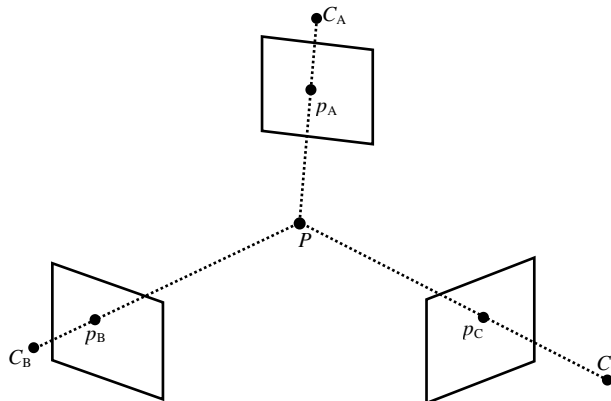


Figura 2.7 – Relação de incidência ponto-ponto-ponto

Essas relações de incidência mostram a maior flexibilidade do tensor trifocal em comparação com a matriz fundamental, que permite relacionar somente a projeção de um ponto em duas imagens. Além disso, o tensor trifocal insere mais restrições à geometria entre as câmeras que a matriz fundamental uma vez que permite relacionar três imagens, ao invés de duas.

2.2.1 Cálculo do tensor trifocal

A partir de uma quantidade suficiente de projeções de pontos ou de retas, relacionadas entre as três imagens, é possível determinar os 27 parâmetros do tensor trifocal a partir das relações de incidência, apresentadas anteriormente.

Cada relação de incidência permite gerar um número diferente de equações nos termos de T , conforme apresentado na Tabela 2.1.

Tabela 2.1 – Número de equações que cada relação de incidência é capaz de gerar (adaptado de HARTLEY e ZISSERMAN, 2000)

Cada relação	gera	pela equação
ponto-ponto-ponto	4 equações	$[p_B]_x (x_A \mathbf{T}_1 + y_A \mathbf{T}_2 + w_A \mathbf{T}_3) [p_C]_x = \mathbf{0}_{3 \times 3}$
reta-reta-reta	2 equações	$l_A^T = l_B^T [\mathbf{T}_1 l_C \quad \mathbf{T}_2 l_C \quad \mathbf{T}_3 l_C]$
ponto-reta-reta	2 equações	$l_B^T (x_A \mathbf{T}_1 + y_A \mathbf{T}_2 + w_A \mathbf{T}_3) l_C = 0$
ponto-ponto-reta	1 equação	$[p_B]_x (x_A \mathbf{T}_1 + y_A \mathbf{T}_2 + w_A \mathbf{T}_3) l_C = 0_{3 \times 1}$

O caso mais facilmente aplicável para a determinação do tensor trifocal é a relação de incidência ponto-ponto-ponto, pois requer o menor número de relações entre projeções (já que cada correspondência gera 4 equações) e esses relacionamentos são mais simples de serem obtidos (que os relacionamentos entre projeções de retas).

Desenvolvendo a equação (2.16):

$$\begin{bmatrix} 0 & -z_B & y_B \\ z_B & 0 & -x_B \\ -y_B & x_B & 0 \end{bmatrix} (x_A \mathbf{T}_1 + y_A \mathbf{T}_2 + z_A \mathbf{T}_3) \begin{bmatrix} 0 & -z_C & y_C \\ z_C & 0 & -x_C \\ -y_C & x_C & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Resolvendo o produto das matrizes à esquerda da equação obtém-se:

$$\begin{aligned}
& x_A z_B z_C T_1^{22} + y_A z_B z_C T_2^{22} + z_A z_B z_C T_3^{22} - x_A y_B z_C T_1^{32} - y_A y_B z_C T_2^{32} - z_A y_B z_C T_3^{32} - x_A z_B y_C T_1^{23} - y_A z_B y_C T_2^{23} - z_A z_B y_C T_3^{23} + x_A y_B y_C T_1^{33} + y_A y_B y_C T_2^{33} + z_A y_B y_C T_3^{33} = 0 \\
& x_A z_B z_C T_1^{21} + y_A z_B z_C T_2^{21} + z_A z_B z_C T_3^{21} - x_A y_B z_C T_1^{31} - y_A y_B z_C T_2^{31} - z_A y_B z_C T_3^{31} - x_A z_B x_C T_1^{23} - y_A z_B x_C T_2^{23} - z_A z_B x_C T_3^{23} + x_A y_B x_C T_1^{33} + y_A y_B x_C T_2^{33} + z_A y_B x_C T_3^{33} = 0 \\
& x_A z_B y_C T_1^{21} + y_A z_B y_C T_2^{21} + z_A z_B y_C T_3^{21} - x_A y_B y_C T_1^{31} + y_A y_B y_C T_2^{31} - z_A y_B y_C T_3^{31} - x_A z_B x_C T_1^{22} - y_A z_B x_C T_2^{22} - z_A z_B x_C T_3^{22} + x_A y_B x_C T_1^{32} + y_A y_B x_C T_2^{32} + z_A y_B x_C T_3^{32} = 0 \\
& x_A z_B z_C T_1^{12} + y_A z_B z_C T_2^{12} + z_A z_B z_C T_3^{12} - x_A x_B z_C T_1^{32} - y_A x_B z_C T_2^{32} - z_A x_B z_C T_3^{32} - x_A z_B y_C T_1^{13} - y_A z_B y_C T_2^{13} - z_A z_B y_C T_3^{13} + x_A x_B y_C T_1^{33} + y_A x_B y_C T_2^{33} + z_A x_B y_C T_3^{33} = 0 \\
& x_A z_B z_C T_1^{11} + y_A z_B z_C T_2^{11} + z_A z_B z_C T_3^{11} - x_A x_B z_C T_1^{31} - y_A x_B z_C T_2^{31} - z_A x_B z_C T_3^{31} - x_A z_B x_C T_1^{13} - y_A z_B x_C T_2^{13} - z_A z_B x_C T_3^{13} + x_A x_B x_C T_1^{33} + y_A x_B x_C T_2^{33} + z_A x_B x_C T_3^{33} = 0 \\
& x_A z_B y_C T_1^{11} + y_A z_B y_C T_2^{11} + z_A z_B y_C T_3^{11} - x_A x_B y_C T_1^{31} - y_A x_B y_C T_2^{31} - z_A x_B y_C T_3^{31} - x_A z_B x_C T_1^{12} - y_A z_B x_C T_2^{12} - z_A z_B x_C T_3^{12} + x_A x_B x_C T_1^{32} + y_A x_B x_C T_2^{32} + z_A x_B x_C T_3^{32} = 0 \\
& x_A y_B z_C T_1^{12} + y_A y_B z_C T_2^{12} + z_A y_B z_C T_3^{12} - x_A x_B z_C T_1^{22} - y_A x_B z_C T_2^{22} - z_A x_B z_C T_3^{22} - x_A y_B y_C T_1^{13} - y_A y_B y_C T_2^{13} - z_A y_B y_C T_3^{13} + x_A x_B y_C T_1^{23} + y_A x_B y_C T_2^{23} + z_A x_B y_C T_3^{23} = 0 \\
& x_A y_B z_C T_1^{11} + y_A y_B z_C T_2^{11} + z_A y_B z_C T_3^{11} - x_A x_B z_C T_1^{21} - y_A x_B z_C T_2^{21} - z_A x_B z_C T_3^{21} - x_A y_B x_C T_1^{13} - y_A y_B x_C T_2^{13} - z_A y_B x_C T_3^{13} + x_A x_B x_C T_1^{23} + y_A x_B x_C T_2^{23} + z_A x_B x_C T_3^{23} = 0 \\
& x_A y_B y_C T_1^{11} + y_A y_B y_C T_2^{11} + z_A y_B y_C T_3^{11} - x_A x_B y_C T_1^{21} - y_A x_B y_C T_2^{21} - z_A x_B y_C T_3^{21} - x_A y_B x_C T_1^{12} - y_A y_B x_C T_2^{12} - z_A y_B x_C T_3^{12} + x_A x_B x_C T_1^{22} + y_A x_B x_C T_2^{22} + z_A x_B x_C T_3^{22} = 0
\end{aligned} \tag{2.17}$$

em que T_i^{jk} representa o elemento j, k da matriz T_i , para $i = 1, 2, 3$.

Cada ponto gera 9 equações em função das incógnitas T_i^{jk} , entretanto, de acordo com Hartley e Zisserman (2000), apenas 4 dessas equações serão linearmente independentes e, portanto, não contribuem para a solução do sistema. No caso mínimo, serão necessárias as relações entre as projeções de 7 pontos entre as imagens para se obter solução única. Se mais correspondências entre pontos forem utilizadas na geração desse sistema, será possível resolvê-lo por minimização de erro, permitindo determinar uma solução mais robusta.

Da mesma forma como no cálculo da matriz fundamental, a normalização das projeções dos pontos é imprescindível para se obter uma solução com pouco erro, dado que o sistema linear que permite o cálculo do tensor trifocal é usualmente mal condicionado (BARROSO et al., 1987).

A normalização isotrópica proposta por Hartley (1997) para o cálculo da matriz fundamental também é aplicável ao tensor trifocal. Essa normalização primeiramente translada as projeções de modo que o centroide da imagem coincida com a origem e em seguida aplica uma escala aos pontos de modo que a distância média entre cada ponto e a origem seja igual a $\sqrt{2}$.

Uma vez determinados os parâmetros do tensor trifocal a partir das projeções normalizadas é necessário que o tensor seja submetido à transformação inversa àquela aplicada aos pontos projetados, de modo que o tensor possa representar a relação entre as projeções dos pontos nas imagens.

3. RECONSTRUÇÃO NO ESPAÇO PROJETIVO

Neste capítulo, será apresentado o equacionamento para determinar modelos de câmeras para pontos no espaço projetivo, a partir da matriz fundamental e do tensor trifocal e o método de triangulação que permite reconstruir o objeto no espaço projetivo.

Serão ainda apresentados e comparados os resultados obtidos pela utilização da matriz fundamental e do tensor trifocal na reconstrução de um objeto no espaço projetivo.

3.1 *Reconstrução projetiva utilizando matriz fundamental*

A partir da matriz fundamental, é possível determinar os modelos de câmeras em função de uma transformação projetiva desconhecida, conforme o equacionamento a seguir.

Como não se conhece o referencial absoluto do objeto, pode-se escolher a origem arbitrariamente. A escolha mais simples consiste em fixar a origem na câmera A, que será a câmera de referência, fazendo com que seu modelo de câmera para pontos no espaço projetivo seja dado por:

$$\mathbf{M}_A = [\mathbf{I}_{3 \times 3} \mid \mathbf{0}_{3 \times 1}]$$

em que $\mathbf{I}_{3 \times 3}$ representa a matriz identidade e $\mathbf{0}_{3 \times 1}$ representa um vetor nulo.

De acordo com Hartley e Zisserman (2000), o modelo da câmera B, considerando a geometria das câmeras expressa pela matriz fundamental F, é dada por:

$$\mathbf{M}_B = \left[[e_{AB}]_{\times} \mathbf{F} + e_{AB} v^T \mid \lambda e_{AB} \right] \quad (3.01)$$

em que $[e_{AB}]_{\times}$ é a matriz antissimétrica do epipolo e_{AB} , v é um vetor qualquer e λ é um escalar não nulo.

Como v é um vetor qualquer, pode ser escolhido livremente. Uma escolha conveniente para v é $v = [0 \ 0 \ 0]^T$. A mesma regra vale para λ e pode-se adotar $\lambda = 1$.

Assim, conhecida a matriz fundamental e o epipolo da imagem B, calculável pela equação (2.04), pode-se determinar as matrizes projetivas das câmeras A e B:

$$\begin{aligned} \mathbf{M}_A &= [\mathbf{I}_{3 \times 3} \mid \mathbf{0}_{3 \times 1}] \\ \mathbf{M}_B &= [[e_{AB}]_x \mathbf{F} \mid e_{AB}] \end{aligned} \quad (3.02)$$

Uma vez determinadas os modelos das câmeras A e B no mesmo referencial, é possível, por triangulação, determinar a posição no espaço dos pontos observados pelas duas câmeras, conforme a Figura 3.1.

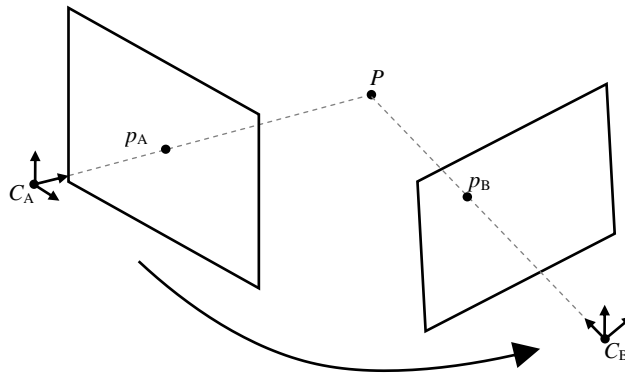


Figura 3.1 – Solução por triangulação no espaço projetivo

O ponto P , de coordenadas $P = [X \ Y \ Z \ W]^T$, projeta o ponto $p_A = [x_A^H \ y_A^H \ w_A^H]^T$ na câmera A por $p_A = \mathbf{M}_A P$:

$$p_A = \begin{bmatrix} x_A^H \\ y_A^H \\ w_A^H \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix}$$

Reescrevendo:

$$\begin{aligned} \frac{x_A^H}{w_A^H} &= x_A = \frac{X}{W} \\ \frac{y_A^H}{w_A^H} &= y_A = \frac{Y}{W} \end{aligned}$$

$$\begin{aligned} W x_A &= X \rightarrow x_A W - X = 0 \\ W y_A &= Y \rightarrow y_A W - Y = 0 \end{aligned}$$

Na forma matricial:

$$\begin{bmatrix} -1 & 0 & 0 & x_A \\ 0 & -1 & 0 & y_A \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix} = 0 \quad (3.03)$$

A projeção de P na câmera é dada por $p_B = \mathbf{M}_B P$:

$$p_B = \begin{bmatrix} x_B^P \\ y_B^P \\ w_B^P \end{bmatrix} = \begin{bmatrix} m_{B_{11}} & m_{B_{12}} & m_{B_{13}} & m_{B_{14}} \\ m_{B_{21}} & m_{B_{22}} & m_{B_{23}} & m_{B_{24}} \\ m_{B_{31}} & m_{B_{32}} & m_{B_{33}} & m_{B_{34}} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix}$$

Reescrevendo:

$$\frac{x_B^H}{w_B^H} = x_B = \frac{Xm_{B_{11}} + Ym_{B_{12}} + Zm_{B_{13}} + Wm_{B_{14}}}{Xm_{B_{31}} + Ym_{B_{32}} + Zm_{B_{33}} + Wm_{B_{34}}}$$

$$\frac{y_B^H}{w_B^H} = y_B = \frac{Xm_{B_{21}} + Ym_{B_{22}} + Zm_{B_{23}} + Wm_{B_{24}}}{Xm_{B_{31}} + Ym_{B_{32}} + Zm_{B_{33}} + Wm_{B_{34}}}$$

$$(Xm_{B_{31}} + Ym_{B_{32}} + Zm_{B_{33}} + Wm_{B_{34}})x_B = Xm_{B_{11}} + Ym_{B_{12}} + Zm_{B_{13}} + Wm_{B_{14}}$$

$$(Xm_{B_{31}} + Ym_{B_{32}} + Zm_{B_{33}} + Wm_{B_{34}})y_B = Xm_{B_{21}} + Ym_{B_{22}} + Zm_{B_{23}} + Wm_{B_{24}}$$

$$x_B Xm_{B_{31}} + x_B Ym_{B_{32}} + x_B Zm_{B_{33}} + x_B Wm_{B_{34}} - Xm_{B_{11}} - Ym_{B_{12}} - Zm_{B_{13}} - Wm_{B_{14}} = 0$$

$$y_B Xm_{B_{31}} + y_B Ym_{B_{32}} + y_B Zm_{B_{33}} + y_B Wm_{B_{34}} - Xm_{B_{21}} - Ym_{B_{22}} - Zm_{B_{23}} - Wm_{B_{24}} = 0$$

Na forma matricial:

$$\begin{bmatrix} x_B m_{B_{31}} - m_{B_{11}} & x_B m_{B_{32}} - m_{B_{12}} & x_B m_{B_{33}} - m_{B_{13}} & x_B m_{B_{34}} - m_{B_{14}} \\ y_B m_{B_{31}} - m_{B_{21}} & y_B m_{B_{32}} - m_{B_{22}} & y_B m_{B_{33}} - m_{B_{23}} & y_B m_{B_{34}} - m_{B_{24}} \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix} = 0 \quad (3.04)$$

Como os sistemas representados nas equações (3.03) e (3.04) apresentam as mesmas variáveis, pode-se escrever:

$$\begin{bmatrix} -1 & 0 & 0 & x_A \\ 0 & -1 & 0 & y_A \\ x_B m_{B_{31}} - m_{B_{11}} & x_B m_{B_{32}} - m_{B_{12}} & x_B m_{B_{33}} - m_{B_{13}} & x_B m_{B_{34}} - m_{B_{14}} \\ y_B m_{B_{31}} - m_{B_{21}} & y_B m_{B_{32}} - m_{B_{22}} & y_B m_{B_{33}} - m_{B_{23}} & y_B m_{B_{34}} - m_{B_{24}} \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix} = 0 \quad (3.05)$$

Fazendo $W=1$, o sistema da equação (3.05) pode ser reescrito como:

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ x_B m_{A_{31}} - m_{B_{11}} & x_B m_{A_{32}} - m_{B_{12}} & x_B m_{B_{33}} - m_{B_{13}} \\ y_B m_{A_{31}} - m_{B_{21}} & y_B m_{A_{32}} - m_{B_{22}} & y_B m_{B_{33}} - m_{B_{23}} \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} -x_A \\ -y_A \\ m_{B_{14}} - x_B m_{B_{34}} \\ m_{B_{24}} - y_B m_{B_{34}} \end{bmatrix} \quad (3.06)$$

Como esse sistema é super especificado e pode não ter solução única, devido a erros na medição das projeções e na determinação dos parâmetros das câmeras projetivas, é necessário resolvê-lo por métodos de minimização de erro.

3.2 Reconstrução projetiva utilizando tensor trifocal

Tal como para a matriz fundamental, de acordo com o equacionamento descrito por Hartley e Zisserman (2000) é possível estabelecer uma relação entre o tensor trifocal e os modelos de câmeras para pontos no espaço projetivo. Da mesma forma que para a matriz fundamental, a câmera A servirá como referência para as demais, que ficarão sujeitas a uma transformação projetiva desconhecida:

$$\begin{aligned} \mathbf{M}_A &= [\mathbf{I}_{3 \times 3} | 0_{3 \times 1}] \\ \mathbf{M}_B &= \left[\sum_{i=1}^3 (\mathbf{T}_i e_{AC}) \mid e_{AB} \right] \\ \mathbf{M}_C &= \left[(e_{AC} e_{AC}^T - \mathbf{I}_{3 \times 3}) \cdot \sum_{i=1}^3 (\mathbf{T}_i^T e_{AB}) \mid e_{AC} \right] \end{aligned} \quad (3.07)$$

em que $\mathbf{I}_{3 \times 3}$ representa a matriz identidade, $0_{3 \times 1}$ representa um vetor nulo, e_{AB} e e_{AC} representam os epipolos nas imagens B e C, respectivamente, decorrentes da câmera A.

A técnica de triangulação aplicada à determinação dos pontos no espaço para a matriz fundamental também pode ser utilizada para o tensor trifocal, de modo que o sistema da equação (3.06) será expandido para 6 equações e será super especificado, podendo ser resolvido por minimização. O sistema da equação (3.06), reescrito para realizar a triangulação para 3 retas, tem a forma:

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ x_B m_{B_{31}} - m_{B_{11}} & x_B m_{B_{32}} - m_{B_{12}} & x_B m_{B_{33}} - m_{B_{13}} \\ y_B m_{B_{31}} - m_{B_{21}} & y_B m_{B_{32}} - m_{B_{22}} & y_B m_{B_{33}} - m_{B_{23}} \\ x_C m_{C_{31}} - m_{C_{11}} & x_C m_{C_{32}} - m_{C_{12}} & x_C m_{B_{33}} - m_{C_{13}} \\ y_C m_{C_{31}} - m_{C_{21}} & y_C m_{C_{32}} - m_{C_{22}} & y_C m_{B_{33}} - m_{C_{23}} \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} -x_A \\ -y_A \\ m_{B_{14}} - x_B m_{B_{34}} \\ m_{B_{24}} - y_B m_{B_{34}} \\ m_{C_{14}} - x_C m_{A_{34}} \\ m_{C_{24}} - y_C m_{A_{34}} \end{bmatrix} \quad (3.08)$$

3.3 Reconstrução com imagens sintéticas

Com o objetivo de testar a reconstrução obtida utilizando a matriz fundamental e o tensor trifocal, foram realizadas reconstruções utilizando imagens sintéticas geradas no Matlab a partir da projeção de um objeto virtual. Esses testes foram conduzidos de modo a tornar possível a comparação dos erros:

- na reprojeção do objeto reconstruído (no espaço projetivo) para as imagens;
- na reconstrução do objeto no espaço métrico, comparando com o objeto original.

Em cada simulação, foram geradas três imagens sintéticas a partir do objeto ilustrado na Figura 3.2.

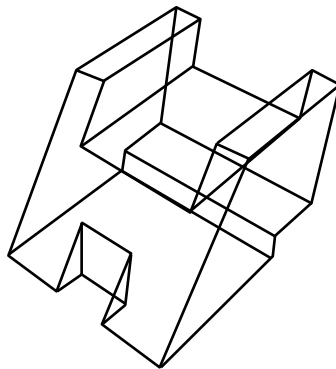


Figura 3.2 – Visão em perspectiva do objeto utilizado como base da geração das imagens sintéticas

Para verificar o efeito dos parâmetros das câmeras virtuais no erro da reprojeção, as câmeras virtuais foram colocadas em posições e com orientações distintas.

Nos testes apresentados a seguir, foram escolhidos parâmetros de câmeras de modo que todo o objeto fosse visível em uma janela de 640×480 pixels. Além disso, tomou-se o cuidado de variar os parâmetros de câmeras (a menos do vetor de *view up*) para garantir que o teste fosse condizente com a proposta deste trabalho.

Para garantir a ausência de oclusões, o objeto foi elaborado em *wireframe*, de modo que 26 pontos estivessem disponíveis em cada imagem. As coordenadas dos pontos do objeto virtual estão apresentadas na Tabela 3.1.

Inicialmente, realizou-se o experimento sem a inserção de erros de medida. Em seguida, com o objetivo de simular o efeito do erro de medida das projeções, adicionou-se um valor aleatório inteiro em cada coordenada das projeções. Variou-se a semente do erro aleatório no intervalo entre 1 e 5, uma vez que,

para sementes maiores que 5, o processo de reconstrução gera muitas soluções com erros muito grandes, inviabilizando o experimento.

Tabela 3.1 - Coordenadas dos pontos do objeto virtual

Ponto	X	Y	Z	Ponto	X	Y	Z
1	-40	29	50	14	28	-8	89
2	-40	29	60	15	-40	-20	50
3	-40	29	100	16	-20	-20	50
4	-40	29	110	17	-20	-30	50
5	10	29	50	18	39	-30	50
6	10	29	60	19	39	-30	71
7	10	29	100	20	28	-30	71
8	10	29	110	21	28	-30	89
9	-40	10	60	22	39	-30	89
10	-40	10	100	23	39	-30	110
11	19	10	60	24	-40	-20	110
12	19	10	100	25	-20	-20	110
13	28	-8	71	26	-20	-30	110

Em cada simulação, determinou-se o objeto no espaço projetivo a partir de duas ou três imagens (conforme a técnica aplicada – pela matriz fundamental ou pelo tensor trifocal) e calcularam-se os erros máximo e médio na reprojeção do objeto recuperado no espaço projetivo em cada imagem.

Com o intuito de verificar a qualidade da reconstrução no espaço projetivo, determinou-se em seguida a transformação do espaço projetivo para o euclidiano, tomando como base o objeto original e o objeto recuperado no espaço projetivo e, com essa transformação, foi possível obter o objeto reconstruído no espaço euclidiano. Por fim, foram calculados os erros máximo e médio entre a reconstrução do objeto no espaço euclidiano com o objeto original.

Cada simulação foi conduzida de acordo com os passos do fluxograma da Figura 3.3.

Os cálculos da matriz fundamental e do tensor trifocal foram realizados com base nos valores normalizados das coordenadas das projeções que foram obtidas da seguinte maneira:

- calcula-se o centroide $(c_x; c_y)$ dos pontos da imagem;
- determina-se a distância média (d) entre os pontos da imagem e o centroide;
- calcula-se a transformação a ser aplicada nos pontos, que terá a forma:

$$\mathbf{N} = \begin{bmatrix} \sqrt{2}/d & 0 & -\sqrt{2}c_x/d \\ 0 & \sqrt{2}/d & -\sqrt{2}c_y/d \\ 0 & 0 & 1 \end{bmatrix}$$

- aplica-se essa transformação aos pontos da imagem, de modo que a distância média dos pontos ao centroide passe a ser de $\sqrt{2}$.

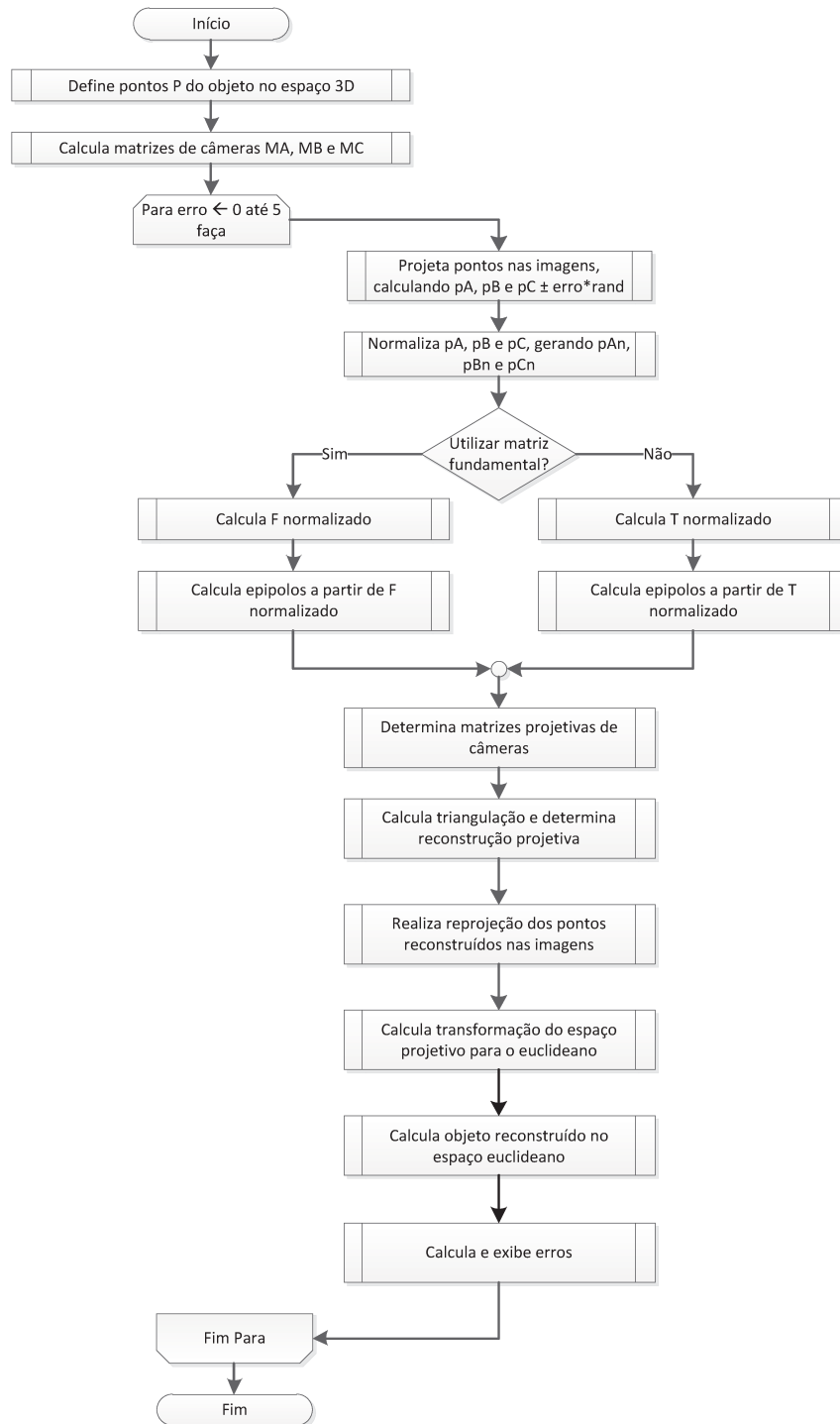


Figura 3.3 – Fluxograma dos testes em ambiente virtual

Para determinar a matriz fundamental, foram utilizadas somente as imagens A e B. A matriz fundamental foi determinada a partir da aplicação do sistema da equação (2.07), expandido para 26 equações, e as matrizes projetivas de câmeras foram determinadas a partir das equações (3.02), calculadas com a matriz

fundamental normalizada. Para retornar ao sistema de coordenadas original, a desnormalização foi efetuada diretamente no modelo da câmera B, conforme a equação (3.09):

$$\mathbf{M}_B = \mathbf{N}_B^{-1} \left[[e_{AB}]_x \mathbf{F} \mathbf{N}_A \mid e_{AB} \right] \quad (3.09)$$

em que \mathbf{F} representa a matriz fundamental, e_{AB} representa o epipolo da imagem B, \mathbf{N}_A e \mathbf{N}_B representam, respectivamente, as matrizes de normalização para as câmeras A e B.

Para o cálculo do tensor trifocal, foram utilizadas as três imagens geradas e o tensor trifocal foi determinado pela solução do sistema da equação (2.17) para 26 pontos e os modelos de câmeras foram determinados a partir das equações (3.07), calculadas com o tensor trifocal normalizado. A desnormalização foi efetuada diretamente nos modelos das câmeras B e C, conforme as equações a seguir:

$$\begin{aligned} \mathbf{M}_B &= \mathbf{N}_B^{-1} \left[\sum_{i=1}^3 (\mathbf{T}_i \cdot e_{AC}) \mathbf{N}_A \mid e_{AB} \right] \\ \mathbf{M}_C &= \mathbf{N}_C^{-1} \left[(e_{AC} \cdot e_{AC}^T - \mathbf{I}_{3 \times 3}) \cdot \sum_{i=1}^3 (\mathbf{T}_i^T \cdot e_{AB}) \mathbf{N}_A \mid e_{AC} \right] \end{aligned} \quad (3.10)$$

em que \mathbf{T}_i , $i=1, \dots, 3$, representa o tensor trifocal, e_{AB} e e_{AC} representam, respectivamente, os epipolos das imagens B e C, $\mathbf{I}_{3 \times 3}$ representa a matriz identidade, \mathbf{N}_A , \mathbf{N}_B e \mathbf{N}_C representam, respectivamente, as matrizes de normalização para as câmeras A, B e C.

Depois de determinados os modelos de câmeras, foram calculados os pontos no espaço projetivo por triangulação, conforme as equações (3.06) ou (3.08), quando calculada pela matriz fundamental ou pelo tensor trifocal, respectivamente.

A partir dessa reconstrução projetiva obtida por triangulação, os pontos do espaço projetivo foram então reprojitados nas imagens geradas pela matriz projetiva de câmera obtida em cada situação.

Para cada ponto reprojitado, foi calculada a distância em relação à projeção original, para um ponto i , conforme a equação:

$$d_i = \sqrt{(x_{R_i} - x_i)^2 + (y_{R_i} - y_i)^2}$$

em que x_i e y_i representam as coordenadas da projeção original, x_{R_i} e y_{R_i} representam as coordenadas das reprojções, com resultado arredondado para o valor inteiro mais próximo, de modo que o erro possa ser avaliado em pixels.

O erro médio foi calculado por:

$$EQM = \frac{\sum_{i=1}^n d_i}{n} \quad (3.11)$$

em que n representa o número de pontos.

O erro médio foi arredondado para o inteiro mais próximo, para que pudesse ser avaliado em pixels.

Com o objetivo de determinar o erro na reconstrução, optou-se por levar o objeto recuperado do espaço projetivo para o espaço euclidiano, de modo que a comparação com o objeto original fosse possível. Como os pontos originais do objeto estavam disponíveis, a determinação da transformação foi realizada por meio da equação:

$$P_E = P_P \mathbf{H}_{P \rightarrow E} \quad (3.12)$$

em que P_E representa um ponto no espaço euclidiano (obtidos a partir do objeto original), P_P representa um ponto no espaço projetivo (obtidos a partir do objeto recuperado no espaço projetivo) e $\mathbf{H}_{P \rightarrow E}$ representa a transformação do espaço projetivo para o espaço euclidiano. Mais informações a respeito dessa transformação serão apresentadas no Capítulo 4 deste trabalho.

Como a transformação da equação (3.12) é obtida em função da posição da câmera de referência, todas as simulações foram realizadas com a mesma rotação nessa câmera, de modo que o objeto recuperado no espaço projetivo possa ser comparado com o objeto original sem a necessidade de se aplicar mais uma transformação (de rotação).

Apesar de terem sido executadas simulações com diversos parâmetros de câmeras, serão apresentados a seguir os resultados de quatro dessas simulações.

Simulação 1 – Câmeras paralelas

Nessa simulação, foram escolhidas três câmeras paralelas com parâmetros intrínsecos constantes, conforme apresentados na Tabela 3.2.

Tabela 3.2 – Parâmetros das câmeras paralelas

Câmera A:	Câmera B:	Câmera C:
$T_x = 0$	$T_x = -80$	$T_x = 50$
$T_y = 0$	$T_y = 0$	$T_y = 0$
$T_z = -1$	$T_z = -1$	$T_z = 0$
$view\ up = [0\ 1\ 0]^T$	$view\ up = [0\ 1\ 0]^T$	$view\ up = [0\ 1\ 0]^T$
$look\ at = [0\ 0\ 1]^T$	$look\ at = [-80\ 0\ 1]^T$	$look\ at = [50\ 0\ 1]^T$
$f = 130$	$f = 130$	$f = 130$
$u = 320$	$u = 320$	$u = 320$
$v = 240$	$v = 240$	$v = 240$

Na Tabela 3.2, a posição da câmera é definida pela translação $T = [T_x\ T_y\ T_z]^T$ e pelos vetores de *view up* (inclinação da câmera) e *look at* (ponto para o qual a câmera está apontada). A distância focal é definida pelo parâmetro f , enquanto a posição do ponto principal é definida pelos parâmetros u e v .

Para essas câmeras, foram obtidas as imagens apresentadas na Figura 3.4.

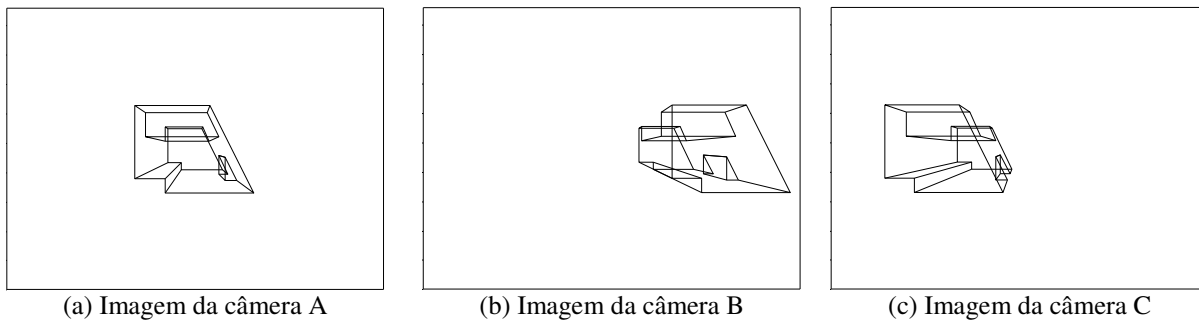


Figura 3.4 – Projeções do objeto em três câmeras paralelas

O erro na reprojeção utilizando a matriz fundamental está apresentado na Tabela 3.3.

Tabela 3.3 – Erro (em pixels) na reprojeção utilizando matriz fundamental com câmeras paralelas

Erro simulado	Imagem A		Imagem B	
	Erro médio	Erro máximo	Erro médio	Erro máximo
0	0	0	0	1
1	0	1	0	2
2	0	2	2	4
3	1	3	3	6
4	1	3	5	8
5	2	4	8	8

Os gráficos da Figura 3.5 apresentam os resultados da Tabela 3.3.

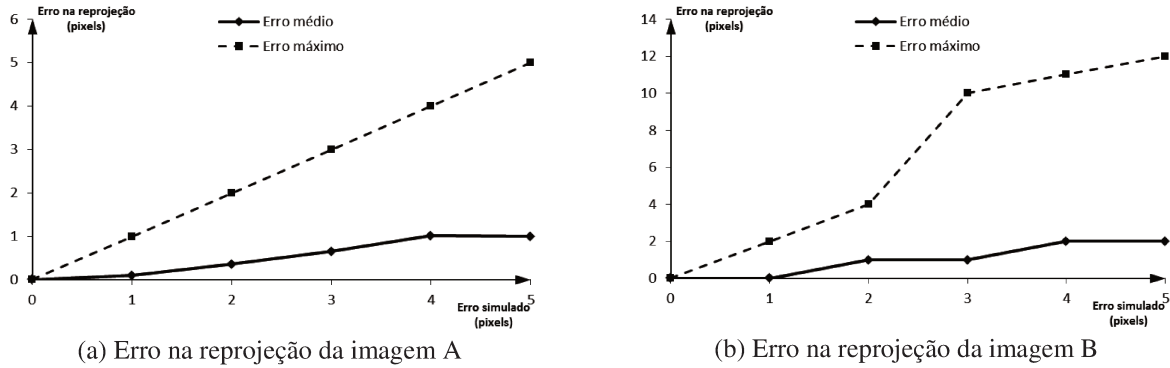


Figura 3.5 – Erro (em pixels) na reprojeção utilizando matriz fundamental com câmeras paralelas

O erro da reconstrução no espaço euclidiano está apresentado na Tabela 3.4. A Figura 3.6 apresenta graficamente os dados da Tabela 3.4.

Tabela 3.4 – Erro da reconstrução no espaço euclidiano utilizando matriz fundamental

Erro simulado	Erro médio	Erro máximo
0	0,637	2,083
1	0,983	4,683
2	1,133	7,332
3	2,767	17,054
4	4,473	26,128
5	4,913	48,642

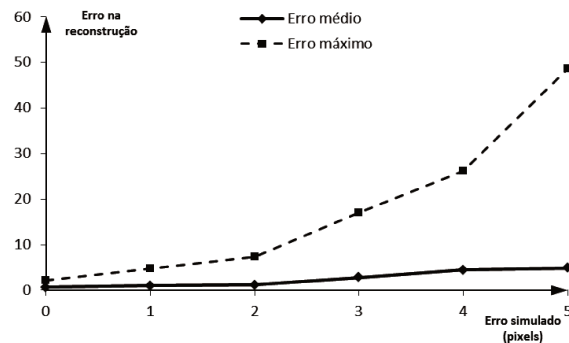


Figura 3.6 – Erro da reconstrução no espaço euclidiano utilizando matriz fundamental

Utilizando o tensor trifocal, foram obtidos os erros na reprojeção conforme apresentado na Tabela 3.5.

Tabela 3.5 – Erro (em pixels) na reprojeção utilizando tensor trifocal

Erro simulado	Imagem A		Imagem B		Imagem C	
	Erro médio	Erro máximo	Erro médio	Erro máximo	Erro médio	Erro máximo
0	0	0	0	1	0	1
1	0	1	1	3	1	2
2	0	1	2	6	1	5
3	1	1	3	11	2	6
4	1	2	4	13	3	13
5	1	5	6	29	4	26

Os gráficos da Figura 3.7 apresentam os resultados da Tabela 3.5.

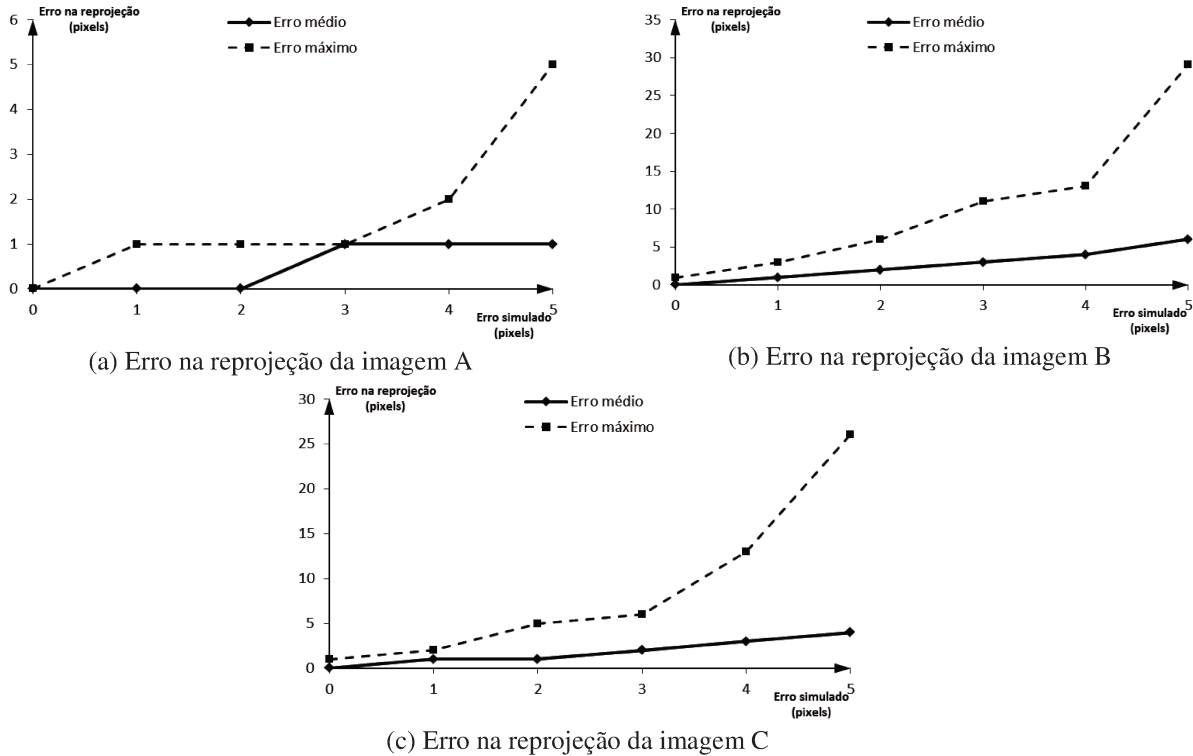


Figura 3.7 – Erro (em pixels) na reprojeção utilizando tensor trifocal

O erro na reconstrução no espaço euclidiano está apresentado na Tabela 3.6. A Figura 3.8 apresenta graficamente os dados da Tabela 3.6.

Tabela 3.6 – Erro da reconstrução no espaço euclidiano utilizando tensor trifocal

Erro simulado	Erro médio	Erro máximo
0	0,301	0,585
1	0,727	2,021
2	1,268	5,123
3	2,055	5,812
4	2,937	9,159
5	3,753	12,548

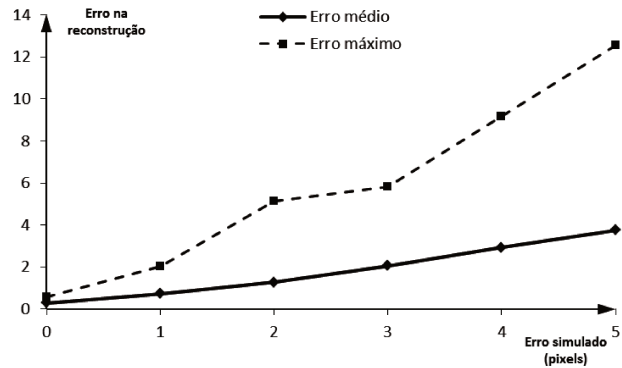


Figura 3.8 – Erro da reconstrução no espaço euclidiano utilizando tensor trifocal

Simulação 2

Foram definidas três câmeras com os parâmetros listados na Tabela 3.7:

Tabela 3.7 – Parâmetros de câmeras para a segunda simulação

Câmera A:	Câmera B:	Câmera C:
$T_x = 0$	$T_x = 400$	$T_x = 200$
$T_y = 0$	$T_y = 0$	$T_y = 200$
$T_z = -1$	$T_z = -100$	$T_z = 200$
$view\ up = [0\ 1\ 0]^T$	$view\ up = [0\ 1\ 0]^T$	$view\ up = [0\ 1\ 0]^T$
$look\ at = [0\ 0\ 1]^T$	$look\ at = [0\ 0\ 1]^T$	$look\ at = [10\ 0\ 10]^T$
$f = 300$	$f = 500$	$f = 750$
$u = 300$	$u = 320$	$u = 320$
$v = 220$	$v = 240$	$v = 250$

As projeções do objeto da Figura 3.2 nas três imagens estão apresentadas na Figura 3.9.

Utilizando a matriz fundamental, o erro obtido na reprojeção está apresentado na Tabela 3.8.

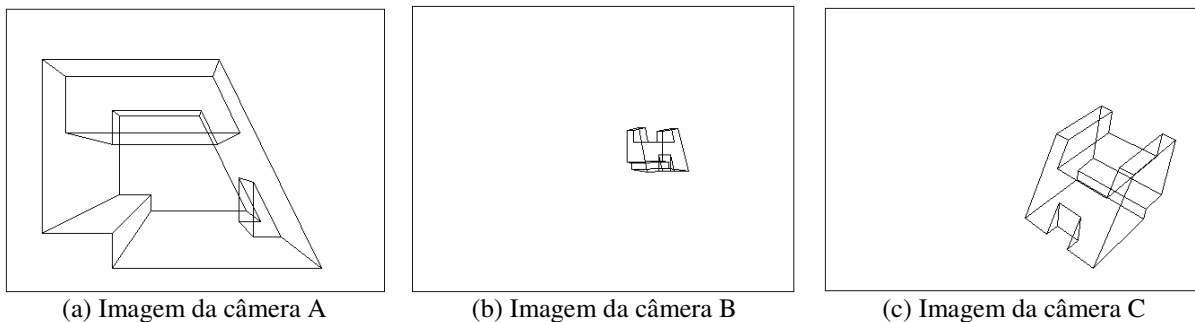


Figura 3.9 – Projeções do objeto da Figura 3.2, em três imagens

Tabela 3.8 – Erro (em pixels) na reprojeção utilizando matriz fundamental

Erro simulado	Imagem A		Imagem B	
	Erro médio	Erro máximo	Erro médio	Erro máximo
0	0	0	0	1
1	0	0	0	1
2	0	0	1	5
3	0	1	1	6
4	0	1	2	11
5	1	1	4	43

Os gráficos da Figura 3.10 apresentam os resultados da Tabela 3.8.

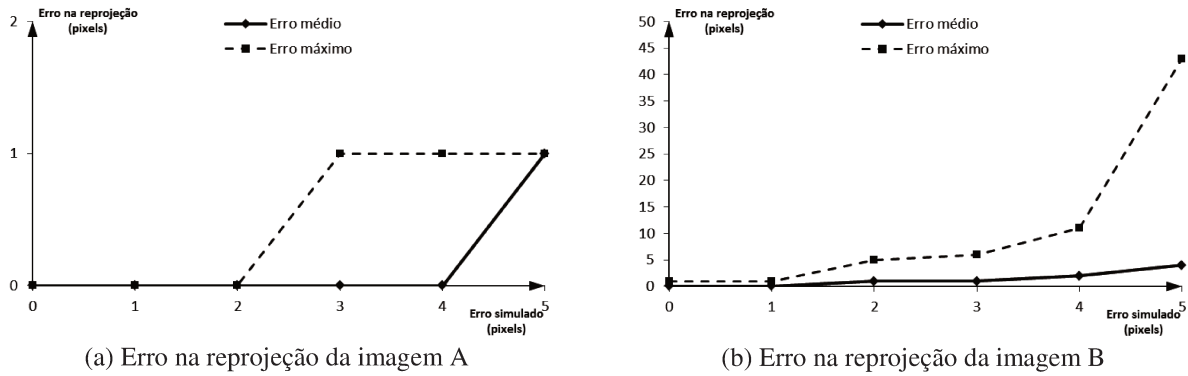


Figura 3.10 – Erro (em pixels) na reprojeção utilizando matriz fundamental

O erro da reconstrução no espaço euclidiano está apresentado na Tabela 3.9. A Figura 3.11 apresenta graficamente os dados da Tabela 3.9.

Tabela 3.9 – Erro da reconstrução no espaço euclidiano utilizando matriz fundamental

Erro simulado	Erro médio	Erro máximo
0	0,397	1,307
1	0,843	3,425
2	1,373	6,895
3	2,207	10,724
4	2,940	12,912
5	4,786	56,317

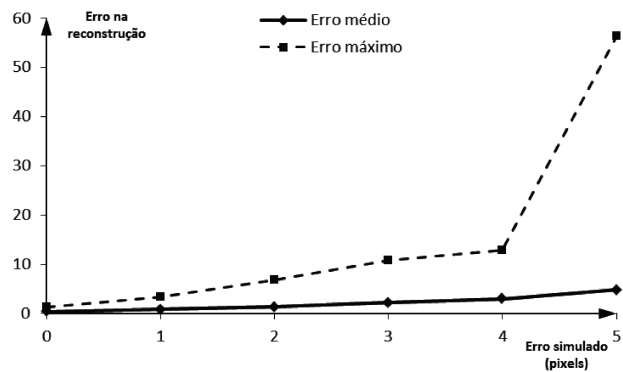


Figura 3.11 – Erro da reconstrução no espaço euclidiano utilizando matriz fundamental

Pelo tensor trifocal, o erro obtido na reprojeção está apresentado na Tabela 3.10. Os gráficos da Figura 3.12 apresentam os resultados da Tabela 3.8.

Tabela 3.10 – Erro (em pixels) na reprojeção utilizando tensor trifocal

Erro simulado	Imagem A		Imagem B		Imagem C	
	Erro médio	Erro máximo	Erro médio	Erro máximo	Erro médio	Erro máximo
0	0	0	0	1	0	1
1	0	0	1	4	2	6
2	0	1	1	20	24	41
3	0	1	2	32	35	77
4	1	2	3	53	41	105
5	1	2	4	44	74	101

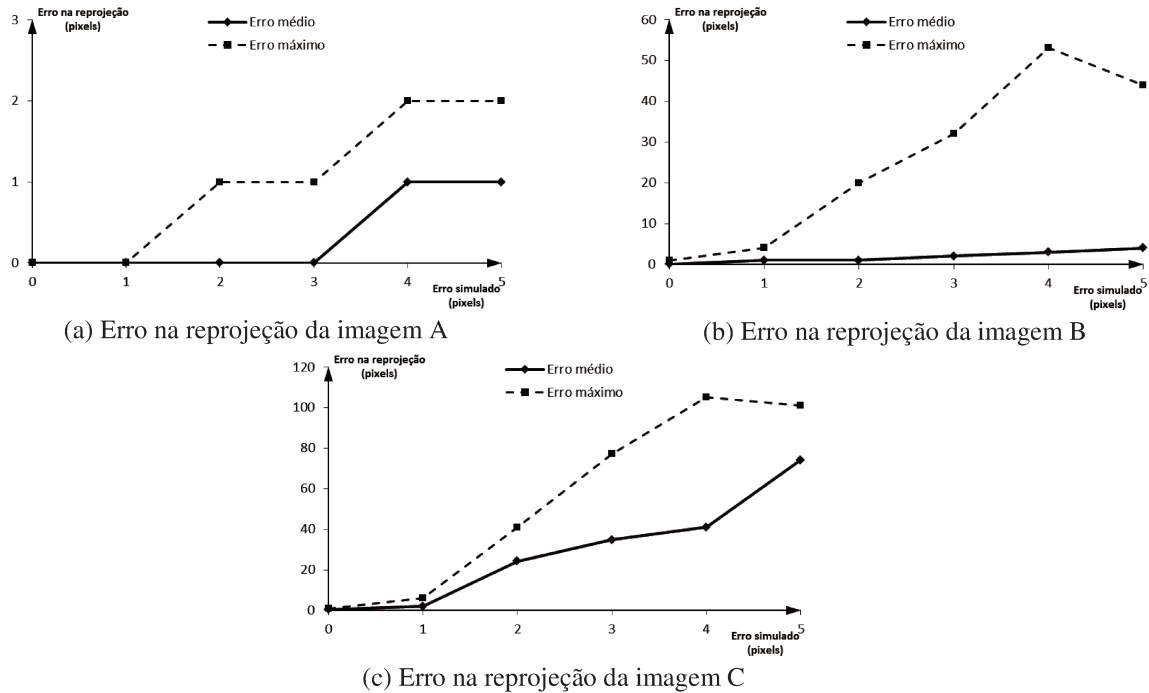


Figura 3.12 – Erro (em pixels) na reprojeção utilizando tensor trifocal

O erro na reconstrução no espaço euclidiano está apresentado na Tabela 3.11 e na Figura 3.13.

Tabela 3.11 – Erro da reconstrução no espaço euclidiano utilizando tensor trifocal

Erro simulado	Erro médio	Erro máximo
0	0,331	0,750
1	0,670	1,930
2	1,008	16,647
3	1,447	28,671
4	1,774	43,612
5	2,126	66,409

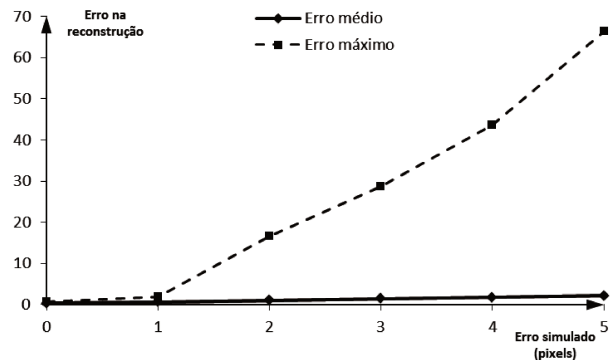


Figura 3.13 – Erro da reconstrução no espaço euclidiano utilizando tensor trifocal

Em comparação com a Simulação 1, percebe-se o aumento geral do erro na reprojeção e na reconstrução no espaço euclidiano, tanto pela matriz fundamental quanto pelo tensor trifocal. A principal razão para esse aumento é o tamanho reduzido do objeto nas imagens da Simulação 2. Tomando-se dois pontos extremos de uma aresta do objeto, nota-se que esses pontos estiverem distantes entre si, o ângulo entre as arestas do objeto real e do objeto reconstruído tenderá a zero. Esse ângulo irá aumentar se os pontos extremos da aresta estiverem próximos, como acontece com as imagens em que os objetos são menores, como nas imagens B e C da Simulação 2.

Simulação 3

Foram definidas três câmeras em posições críticas, conforme a Tabela 3.12. Essas três câmeras diferem entre si somente pela rotação.

Tabela 3.12 – Parâmetros de câmeras para a terceira simulação

Câmera A:	Câmera B:	Câmera C:
$T_x = 0$	$T_x = 0$	$T_x = 0$
$T_y = 0$	$T_y = 0$	$T_y = 0$
$T_z = -1$	$T_z = -1$	$T_z = -1$
$view\ up = [0\ 1\ 0]^T$	$view\ up = [0\ 1\ 0]^T$	$view\ up = [0\ 1\ 0]^T$
$LookAt = [0\ 0\ 1]^T$	$LookAt = [1\ 0\ 1]^T$	$LookAt = [-1\ 0\ 1]^T$
$f = 130$	$f = 130$	$f = 130$
$u = 320$	$u = 320$	$u = 320$
$v = 240$	$v = 240$	$v = 240$

As projeções do objeto da Figura 3.2 nas três imagens estão apresentadas na Figura 3.14.

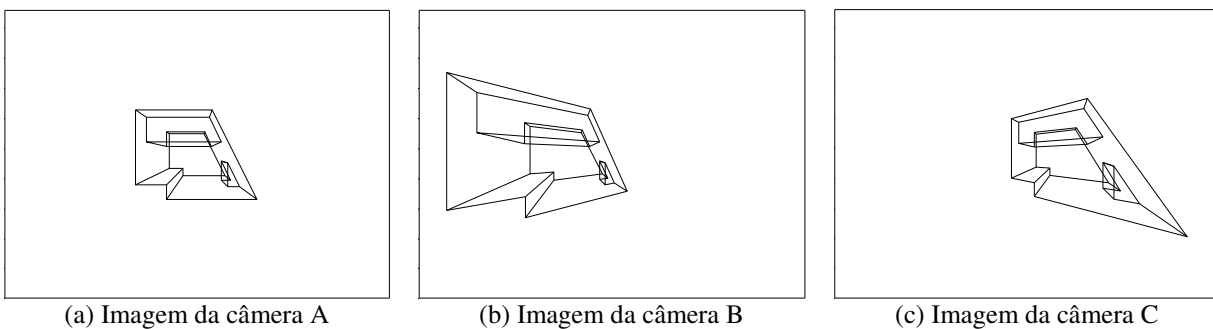


Figura 3.14 – Projeções do objeto para simulação 3

A tentativa de reconstrução no espaço euclidiano resultou nos objetos apresentados na Figura 3.15.

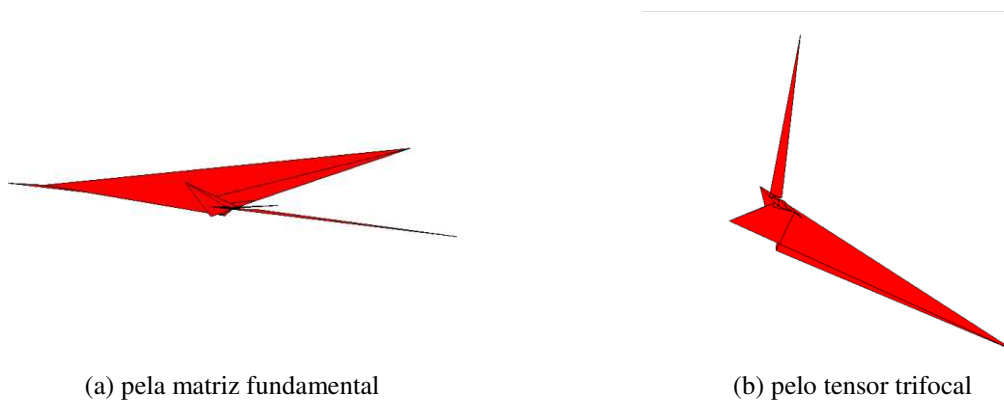


Figura 3.15 – Resultado da reconstrução euclidiana para simulação 3

Na Figura 3.15, nota-se que o objeto reconstruído não guarda semelhança com o objeto original. Como essa é uma posição crítica, apesar da reconstrução no espaço projetivo ser viável, a reconstrução no espaço euclidiano não é possível, mesmo sem a presença de erros de medida. Por essa razão, não serão apresentados os resultados da reprojeção do espaço projetivo.

3.3.1 Comparação entre os resultados

Os gráficos da Figura 3.16 apresentam o erro médio das reprojeções utilizando matriz fundamental e tensor trifocal para as simulações 1 e 2. Nesses gráficos, é possível notar que as abordagens apresentaram resultados com margem de erro semelhantes.

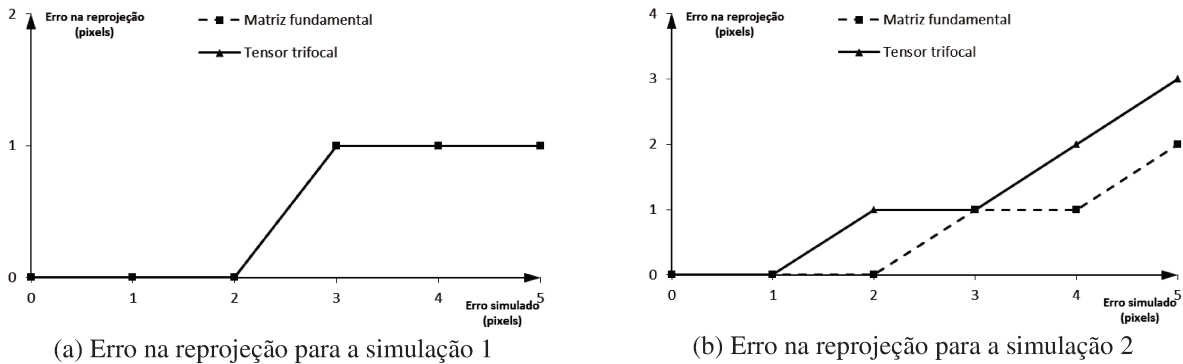
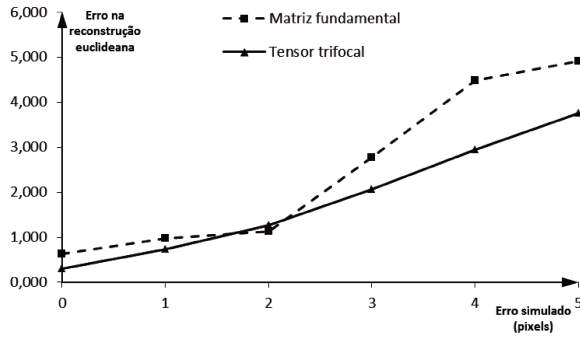


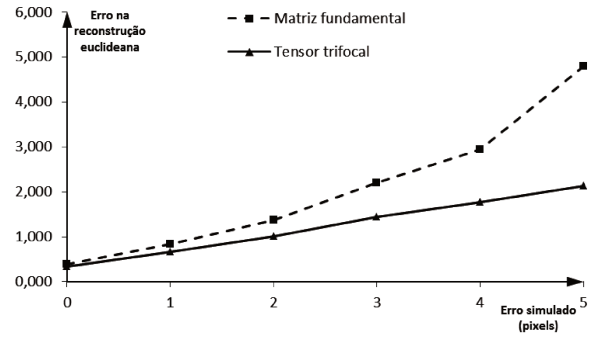
Figura 3.16 – Erro médio das reprojeções

Apesar do tensor trifocal inserir menor erro na reconstrução euclidiana, conforme se nota nos gráficos da Figura 3.17, há situações em que a reconstrução utilizando o tensor trifocal é instável, apresentando solução muito diferente da esperada, gerando reconstruções inaceitáveis. Em todas as simulações, soluções instáveis foram desprezadas e seus resultados não foram computados. Os resultados sugerem que essa instabilidade seja causada pela instabilidade da solução numérica na determinação do tensor trifocal. Além disso, para erros pequenos, não há diferença significativa entre o resultado obtido pelo tensor trifocal e pela matriz fundamental.

Embora o tensor trifocal permita relacionar pontos ou retas entre os trios de imagens, a implementação pela matriz fundamental é mais simples e menos suscetível a erros numéricos.



(a) Erro para a simulação 1



(b) Erro para a simulação 2

Figura 3.17 – Erro médio da reconstrução euclidiana

Com base no que foi exposto, para a segunda parte deste trabalho, que consiste em levar a reconstrução do espaço projetivo para o métrico em que a reconstrução projetiva pode ser realizada tanto pelo tensor trifocal quanto pela matriz fundamental, a reconstrução projetiva utilizando matrizes fundamentais foi escolhida, por apresentar solução mais simples, tanto pelo ponto de vista numérico quanto computacional.

4. RECONSTRUÇÃO MÉTRICA

A aplicação das técnicas associadas à matriz fundamental ou ao tensor trifocal no processo de recuperação tridimensional leva a um resultado no espaço projetivo.

Como essa solução gera um objeto que não guarda a forma do objeto original, é necessário encontrar uma transformação capaz de levar esse objeto recuperado do espaço projetivo para o espaço métrico (onde a forma do objeto recuperado é idêntica à do objeto real, mas a escala não é a mesma) ou para o espaço euclidiano (onde tanto a forma quanto a escala do objeto recuperado são idênticas às do objeto real).

Neste Capítulo serão discutidos os conceitos necessários para se encontrar essa transformação. Para esclarecer o assunto, será discutido a seguir a classificação dos espaços e as transformações capazes de levar um objeto de um espaço a outro.

4.1 Espaço euclidiano

O espaço euclidiano define-se como um espaço vetorial em que o sistema de referência é ortonormal.

As transformações de rotação, translação e “espelhamento”, aplicadas a um objeto no espaço euclidiano, geram um objeto nesse mesmo espaço, de modo que as dimensões do objeto ficam inalteradas, bem como os ângulos relativos entre as arestas (e conseqüentemente o paralelismo entre as arestas).

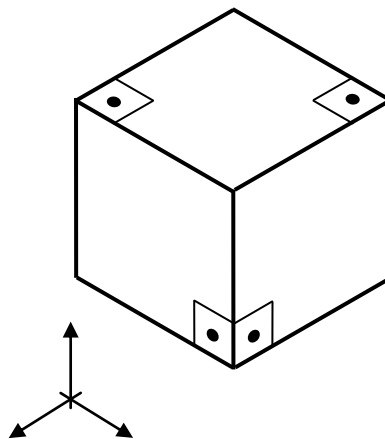


Figura 4.1 – Objeto no espaço euclidiano

4.2 Espaço métrico

O espaço métrico possui as mesmas características do espaço euclidiano, diferindo somente em um fator de escala, já que os valores absolutos dos eixos de referência não são unitários.

Transformações de rotação, translação, “espelhamento” e escala isométrica aplicadas a um objeto do espaço métrico geram um objeto nesse mesmo espaço.

No exemplo da Figura 4.2, os comprimentos das arestas são proporcionais às do objeto da Figura 4.1 e todos os ângulos relativos entre as arestas são mantidos (com a consequente manutenção do paralelismo entre as arestas).



Figura 4.2 – No espaço métrico, são preservados ângulos e a relação entre os comprimentos das arestas

4.3 Espaço afim

Retas que são paralelas no espaço euclidiano mantêm-se paralelas no espaço afim.

Um cubo representado no espaço afim está apresentado na Figura 4.3. Nessa figura, pode-se perceber que as retas que são paralelas no cubo são também paralelas quando é representado nesse espaço.

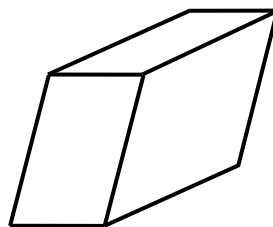


Figura 4.3 – Cubo representado no espaço afim

4.4 Espaço projetivo

O espaço projetivo é caracterizado pela orientação arbitrária dos eixos de referência e pelo uso de coordenadas homogêneas para representar a posição de cada ponto.

De acordo com Topay (2002), um ponto no espaço 3D projetivo é representado por um vetor de 4 dimensões $[X \ Y \ Z \ W]^T$. Nesse vetor, pelo menos uma das coordenadas deve ser diferente de zero e essas coordenadas são chamadas de coordenadas homogêneas.

Um ponto no espaço 2D projetivo pode ser representado como $p = [x \ y \ w]^T$, conforme a Figura 4.4.

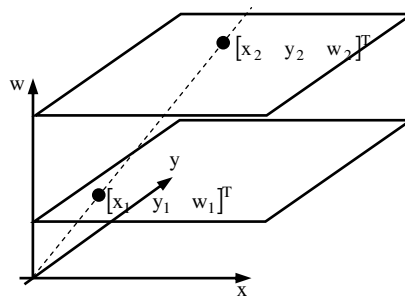


Figura 4.4 – Um ponto representado no espaço projetivo por coordenadas homogêneas

Na Figura 4.4, os pontos $[x_1 \ y_1 \ w_1]^T$ e $[x_2 \ y_2 \ w_2]^T$ serão o mesmo no plano cartesiano se:

$$\frac{x_1}{w_1} = \frac{x_2}{w_2} \quad \text{e} \quad \frac{y_1}{w_1} = \frac{y_2}{w_2}$$

No espaço projetivo, retas que são paralelas no espaço euclidiano interceptam-se nos pontos de fuga (Figura 4.5).

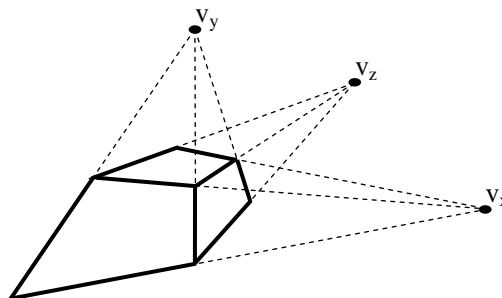


Figura 4.5 – Um cubo representado no espaço projetivo. Nota-se a intersecção das linhas paralelas (do espaço euclidiano) nos pontos de fuga.

4.5 Transformações entre espaços

É possível determinar transformações capazes de levar um objeto de um espaço para outro.

Uma transformação do tipo (HARTLEY e ZISSERMAN, 2000)

$$\mathbf{H}_M = \begin{bmatrix} s\mathbf{R}_{3 \times 3} & \mathbf{t}_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} = \begin{bmatrix} sr_{11} & sr_{12} & sr_{13} & t_x \\ sr_{21} & sr_{22} & sr_{23} & t_y \\ sr_{31} & sr_{32} & sr_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

em que \mathbf{R} representa uma matriz de rotação, s é um fator de escala e \mathbf{t} representa um vetor de translação, permite levar um objeto do espaço euclidiano para o espaço métrico, uma vez que essa transformação aplica rotação, translação e escala isométrica no objeto, preservando a forma geral do objeto.

Uma transformação na forma geral (HARTLEY e ZISSERMAN, 2000):

$$\mathbf{H}_A = \begin{bmatrix} \mathbf{A}_{3 \times 3} & \mathbf{t}_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & t_x \\ a_{21} & a_{22} & a_{23} & t_y \\ a_{31} & a_{32} & a_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

em que \mathbf{A} é uma matriz 3×3 que engloba rotação e escala não-isotrópica, leva o objeto do espaço euclidiano para o espaço afim, cisalhando o objeto nas três direções.

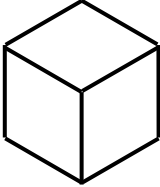
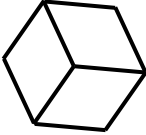
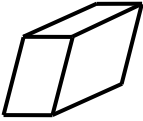
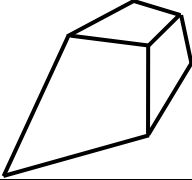
Uma transformação mais geral que permite levar um objeto do espaço euclidiano para o espaço projetivo é dada por:

$$\mathbf{H}_P = \begin{bmatrix} \mathbf{A}_{3 \times 3} & \mathbf{t}_{3 \times 1} \\ \mathbf{v}_{1 \times 3}^T & \delta \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & t_x \\ a_{21} & a_{22} & a_{23} & t_y \\ a_{31} & a_{32} & a_{33} & t_z \\ \alpha & \beta & \gamma & \delta \end{bmatrix} \quad (4.01)$$

em que $[\alpha \ \beta \ \gamma \ \delta]$ representa um vetor qualquer.

A Tabela 4.1 apresenta um resumo das transformações aplicáveis a um objeto do espaço euclidiano e do efeito causado no objeto.

Tabela 4.1 – Resumo das transformações, em relação a um objeto do espaço euclidiano

Espaço	Transformação	Mantém...
Euclidiano 	$\mathbf{H}_E = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$	<ul style="list-style-type: none"> • concorrência entre retas • paralelismo • ângulos entre arestas • escala isométrica
Métrico 	$\mathbf{H}_M = \begin{bmatrix} sr_{11} & sr_{12} & sr_{13} & t_x \\ sr_{21} & sr_{22} & sr_{23} & t_y \\ sr_{31} & sr_{32} & sr_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$	<ul style="list-style-type: none"> • concorrência entre retas • paralelismo • ângulos entre arestas • proporções
Afim 	$\mathbf{H}_A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & t_x \\ a_{21} & a_{22} & a_{23} & t_y \\ a_{31} & a_{32} & a_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$	<ul style="list-style-type: none"> • concorrência entre retas • paralelismo
Projetivo 	$\mathbf{H}_P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{bmatrix}$	<ul style="list-style-type: none"> • concorrência entre retas

De acordo com Hartley e Zisserman (2000), a transformação (4.01) pode ser decomposta em uma cadeia de transformações em que cada matriz representa uma transformação de hierarquia mais alta:

$$\mathbf{H}_{E \rightarrow P} = \mathbf{H}_{E \rightarrow M} \mathbf{H}_{M \rightarrow A} \mathbf{H}_{A \rightarrow P} = \begin{bmatrix} s\mathbf{R}_{3 \times 3} & t_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{Z}_{3 \times 3} & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 1} \\ \mathbf{v}_{1 \times 3}^T & \delta \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{3 \times 3} & t_{3 \times 1} \\ \mathbf{v}_{1 \times 3}^T & \delta \end{bmatrix} \quad (4.02)$$

em que Z é uma matriz triangular superior. Cada uma das matrizes nessa cadeia representa somente a transformação que difere em relação à transformação anterior.

Analisando-se somente as duas últimas matrizes da cadeia da equação (4.02):

$$\mathbf{H}_{M \rightarrow P} = \mathbf{H}_{M \rightarrow A} \mathbf{H}_{A \rightarrow P} = \begin{bmatrix} \mathbf{Z}_{3 \times 3} & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 1} \\ \mathbf{v}_{1 \times 3}^T & \delta \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_{3 \times 3} & \mathbf{0}_{3 \times 1} \\ \mathbf{v}_{1 \times 3}^T & \delta \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} & z_{13} & 0 \\ 0 & z_{22} & z_{23} & 0 \\ 0 & 0 & z_{33} & 0 \\ v_1 & v_2 & v_3 & \delta \end{bmatrix} \quad (4.03)$$

pode-se determinar uma transformação mínima (que não considera rotação e translação) que permite levar um objeto do espaço métrico para o espaço projetivo. Como, de acordo com Hartley e Zisserman (2000), \mathbf{Z} é uma matriz triangular superior:

$$\mathbf{H}_{P \rightarrow M} = \mathbf{H}_{M \rightarrow P}^{-1} = \begin{bmatrix} z_{11} & z_{12} & z_{13} & 0 \\ 0 & z_{22} & z_{23} & 0 \\ 0 & 0 & z_{33} & 0 \\ v_1 & v_2 & v_3 & \delta \end{bmatrix}^{-1}$$

Determinando a transformação inversa de $\mathbf{H}_{P \rightarrow M}$, utilizando a biblioteca de matemática simbólica disponível no *software* Matlab, chega-se a:

$$\mathbf{H}_{P \rightarrow M} = \begin{bmatrix} 1/z_{11} & -z_{12}/(z_{11}z_{22}) & z_{12}z_{23}/(z_{11}z_{22}z_{33}) - z_{13}/(z_{11}z_{33}) & 0 \\ 0 & 1/z_{22} & -z_{23}/(z_{22}z_{33}) & 0 \\ 0 & 0 & 1/z_{33} & 0 \\ -v_1/(\delta z_{11}) & v_1z_{12}/(\delta z_{11}z_{22}) - v_2/(\delta z_{22}) & v_1z_{13}/(\delta z_{11}z_{33}) - v_3/(\delta z_{33}) + v_2z_{23}/(\delta z_{22}z_{33}) - v_1z_{12}z_{23}/(\delta z_{11}z_{22}z_{33}) & 1/\delta \end{bmatrix}$$

Como a inversa da matriz triangular \mathbf{Z} , calculada utilizando a biblioteca de matemática simbólica do Matlab, resulta em:

$$\begin{bmatrix} 1/z_{11} & -z_{12}/(z_{11}z_{22}) & z_{12}z_{23}/(z_{11}z_{22}z_{33}) - z_{13}/(z_{11}z_{33}) \\ 0 & 1/z_{22} & -z_{23}/(z_{22}z_{33}) \\ 0 & 0 & 1/z_{33} \end{bmatrix} = \mathbf{Z}^{-1}$$

e pela troca de variáveis:

$$u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} -v_1/(\delta z_{11}) \\ v_1z_{12}/(\delta z_{11}z_{22}) - v_2/(\delta z_{22}) \\ v_1z_{13}/(\delta z_{11}z_{33}) - v_3/(\delta z_{33}) + v_2z_{23}/(\delta z_{22}z_{33}) - v_1z_{12}z_{23}/(\delta z_{11}z_{22}z_{33}) \end{bmatrix}$$

$d = 1/\delta$

chega-se a:

$$\mathbf{H}_{P \rightarrow M} = \mathbf{H}_{M \rightarrow P}^{-1} = \begin{bmatrix} \mathbf{Z}_{3 \times 3}^{-1} & 0_{3 \times 1} \\ \mathbf{u}_{1 \times 3}^T & d \end{bmatrix} \quad (4.04)$$

A determinação dessa transformação permite elevar o nível de uma reconstrução obtida no espaço projetivo, conforme discutido no Capítulo 3, para o espaço métrico.

4.6 Determinação do vetor u

De acordo com Hartley e Zisserman (2000), o vetor u e o parâmetro d da equação (4.04) podem ser obtidos por meio da determinação do plano no infinito $\Pi_{\infty}^T = ax + by + cz + d$. O plano no infinito é o plano formado pelos três pontos de fuga nos espaços afim ou métrico. Considerando que um objeto como o representado na Figura 4.6 está no espaço projetivo, o plano Π representa o plano formado pelos três pontos de fuga.

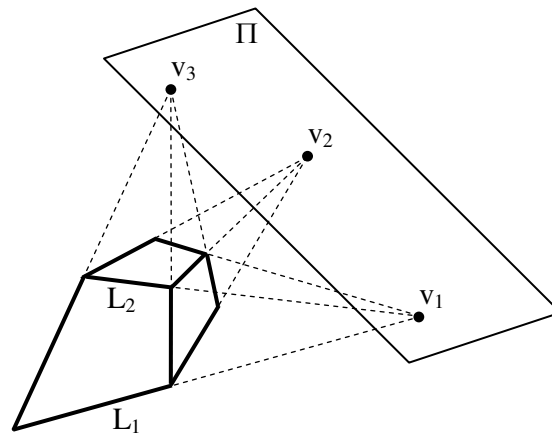


Figura 4.6 – O plano no infinito é definido pelos pontos de fuga do objeto nos espaços afim ou métrico

Para retomar o espaço afim, os pontos de fuga V_1 , V_2 , e V_3 da Figura 4.6 devem ser simultaneamente transladados para o infinito, por meio de uma transformação H_A aplicada a retas e pontos no espaço projetivo.

Dado que, em coordenadas homogêneas, um ponto P_{∞} tem coordenadas $P_{\infty} = [X_{\infty} \ Y_{\infty} \ Z_{\infty} \ 0]^T$ e o plano no infinito tem parâmetros da equação do plano nas coordenadas $\Pi_{\infty} = [0 \ 0 \ 0 \ 1]^T$, pode-se definir uma transformação H na forma de uma matriz 4×4 que leva o ponto de fuga para o infinito. Considerando o ponto de fuga V_1 :

$$\mathbf{H}_A \begin{bmatrix} V_{1x} \\ V_{1y} \\ V_{1z} \\ 1 \end{bmatrix} = \begin{bmatrix} X_{\infty} \\ Y_{\infty} \\ Z_{\infty} \\ 0 \end{bmatrix}$$

Como ponto no infinito deve satisfazer a equação do plano no infinito, então:

$$[0 \ 0 \ 0 \ 1] \mathbf{H}_A \begin{bmatrix} V_{1_x} \\ V_{1_y} \\ V_{1_z} \\ 1 \end{bmatrix} = 0 \quad (4.05)$$

Tomando \mathbf{H}_A como uma matriz com a mesma forma da equação (4.04):

$$\mathbf{H}_A = \begin{bmatrix} h_{11} & h_{12} & h_{13} & 0 \\ 0 & h_{22} & h_{23} & 0 \\ 0 & 0 & h_{33} & 0 \\ a & b & c & d \end{bmatrix} \quad (4.06)$$

e levando (4.06) em (4.05), obtém-se:

$$[0 \ 0 \ 0 \ 1] \begin{bmatrix} h_{11}V_{1_x} + h_{12}V_{1_y} + h_{13}V_{1_z} \\ h_{22}V_{1_y} + h_{23}V_{1_z} \\ h_{33}V_{1_z} \\ aV_{1_x} + bV_{1_y} + cV_{1_z} + d \end{bmatrix} = 0$$

$$aV_{1_x} + bV_{1_y} + cV_{1_z} + d = 0 \quad (4.07)$$

Essa equação é independente de h_{ij} de modo que quaisquer valores podem ser escolhidos para h_{ij} .

Como a transformação H_A deve levar, simultaneamente, todos os pontos de fuga para o infinito, então a equação (4.07) deve ser satisfeita para os pontos de fuga V_y e V_z , resultando em:

$$\begin{cases} aV_{1_x} + bV_{1_y} + cV_{1_z} + d = 0 \\ aV_{2_x} + bV_{2_y} + cV_{2_z} + d = 0 \\ aV_{3_x} + bV_{3_y} + cV_{3_z} + d = 0 \end{cases} \quad (4.08)$$

O sistema da equação (4.08) pode ser resolvido em função do parâmetro d , que pode ser escolhido.

4.7 Determinação da matriz \mathbf{Z}

A transformação \mathbf{H}_A , obtida para h_{ij} arbitrários, restabelece apenas o paralelismo entre as arestas, mas não garante o restabelecimento dos ângulos entre elas e nem garante que as dimensões do objeto estejam de acordo com as dimensões do objeto no espaço euclidiano.

Uma vez determinado Π_∞ , o próximo passo consiste em determinar a matriz \mathbf{Z} da equação (4.04), de modo a restabelecer o ângulo correto entre as arestas do objeto. Como a escolha de h_{ij} é arbitrária, pode-se escolher essa matriz \mathbf{Z} como sendo a matriz de parâmetros intrínsecos da câmera:

$$\mathbf{Z}^{-1} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ 0 & h_{22} & h_{23} \\ 0 & 0 & h_{33} \end{bmatrix} = \begin{bmatrix} f_x & s & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} = \mathbf{K}^{-1}$$

Embora seus parâmetros não sejam conhecidos, essa escolha satisfaz a restrição imposta à equação (4.02), uma vez que a matriz \mathbf{K} é triangular superior e permitirá posteriormente relacionar a transformação \mathbf{H} com os parâmetros intrínsecos da câmera de referência e, a partir daí, reconstruir o objeto no espaço métrico.

Com essa escolha, chega-se a transformação \mathbf{H} , definida em Hartley e Zisserman (2000):

$$\mathbf{H} = \begin{bmatrix} \mathbf{K}^{-1} & 0_{3 \times 1} \\ \Pi_\infty^T & d \end{bmatrix} \quad (1.01)$$

A equação (1.01), apresentada no Capítulo 1, foi aqui repetida para facilidade de leitura.

Como este trabalho tem como objetivo recuperar a informação tridimensional sem depender de informações da câmera ou da cena, deve-se buscar alguma outra entidade geométrica (além do plano no infinito) que auxilie na determinação dessa transformação e, nesse caso, será utilizada a cônica absoluta, que é uma entidade geométrica imaginária definida no plano no infinito. Um ponto $P = [X \ Y \ Z \ W]^T$ na cônica absoluta deve satisfazer as equações:

$$\omega_\infty \begin{cases} X^2 + Y^2 + Z^2 = 0 \\ W = 0 \end{cases} \quad (4.09)$$

Na forma matricial:

$$P^T P = \begin{bmatrix} X & Y & Z & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 0 \end{bmatrix} = 0 \quad (4.10)$$

Como a cônica absoluta é invariante a qualquer transformação métrica (POLLEFEYS et al., 1999), pode ser utilizada para determinar os parâmetros de \mathbf{H} , uma vez que sua projeção nas imagens depende somente dos parâmetros intrínsecos de uma das câmeras, que será definida como câmera de referência. Dessa forma, toda a reconstrução será determinada no sistema de referência dessa câmera.

O referencial da câmera escolhida como referência, conforme abordagens disponíveis na literatura, não é absoluto e a reconstrução ficará em função de rotação, translação e escala arbitrárias, de modo que a reconstrução se dará no espaço métrico.

O equacionamento a seguir relaciona a cônica absoluta com a transformação \mathbf{H} .

Seja a câmera A, que servirá como referência para a reconstrução. Como o objetivo é determinar uma reconstrução no espaço métrico, o objeto recuperado ficará em função de rotação, translação e escala arbitrárias. Dessa forma, pode-se escrever o modelo dessa câmera para pontos no espaço métrico como:

$$\mathbf{M}_{M_A} = \mathbf{K}_A \left[\mathbf{I}_{3 \times 3} \mid \mathbf{0}_{3 \times 1} \right] = \begin{bmatrix} f_{x_A} & s_A & u_{0_A} \\ 0 & f_{y_A} & v_{0_A} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (4.11)$$

em que \mathbf{K}_A representa a matriz dos parâmetros intrínsecos da câmera A e $\mathbf{I}_{3 \times 3}$ representa a matriz identidade. A projeção de um ponto P do espaço métrico nessa câmera terá coordenadas dadas por $p_A = \mathbf{K}_A \left[\mathbf{I}_{3 \times 3} \mid \mathbf{0}_{3 \times 1} \right] P$.

Reprojetando P :

$$P = \left(\mathbf{K}_A \left[\mathbf{I}_{3 \times 3} \mid \mathbf{0}_{3 \times 1} \right] \right)^+ p_A$$

em que $\left(\mathbf{K}_A \left[\mathbf{I}_{3 \times 3} \mid \mathbf{0}_{3 \times 1} \right] \right)^+$ representa a matriz pseudoinversa de $\mathbf{K}_A \left[\mathbf{I}_{3 \times 3} \mid \mathbf{0}_{3 \times 1} \right]$. Como $\mathbf{A}^+ = \left(\mathbf{A}^T \mathbf{A} \right)^{-1} \mathbf{A}^T$, então:

$$\begin{aligned} P &= \left(\left(\mathbf{K}_A \left[\mathbf{I}_{3 \times 3} \mid \mathbf{0}_{3 \times 1} \right] \right)^T \left(\mathbf{K}_A \left[\mathbf{I}_{3 \times 3} \mid \mathbf{0}_{3 \times 1} \right] \right) \right)^{-1} \left(\mathbf{K}_A \left[\mathbf{I}_{3 \times 3} \mid \mathbf{0}_{3 \times 1} \right] \right)^T p_A \\ P &= \left(\left[\mathbf{I}_{3 \times 3} \mid \mathbf{0}_{3 \times 1} \right]^T \mathbf{K}_A^T \mathbf{K}_A \left[\mathbf{I}_{3 \times 3} \mid \mathbf{0}_{3 \times 1} \right] \right)^{-1} \left[\mathbf{I}_{3 \times 3} \mid \mathbf{0}_{3 \times 1} \right]^T \mathbf{K}_A^T p_A \end{aligned}$$

$$P = (\mathbf{K}_A [\mathbf{I}_{3 \times 3} | \mathbf{0}_{3 \times 1}])^{-1} ([\mathbf{I}_{3 \times 3} | \mathbf{0}_{3 \times 1}]^T \mathbf{K}_A^T)^{-1} [\mathbf{I}_{3 \times 3} | \mathbf{0}_{3 \times 1}]^T \mathbf{K}_A^T p_A$$

$$P = [\mathbf{I}_{3 \times 3} | \mathbf{0}_{3 \times 1}]^{-1} \mathbf{K}_A^{-1} \mathbf{K}_A^{-T} [\mathbf{I}_{3 \times 3} | \mathbf{0}_{3 \times 1}]^T [\mathbf{I}_{3 \times 3} | \mathbf{0}_{3 \times 1}]^T \mathbf{K}_A^T p_A$$

Como $[\mathbf{I}_{3 \times 3} | \mathbf{0}_{3 \times 1}]^{-T} [\mathbf{I}_{3 \times 3} | \mathbf{0}_{3 \times 1}]^T = \mathbf{I}_{3 \times 3}$, pode-se reescrever:

$$P = [\mathbf{I}_{3 \times 3} | \mathbf{0}_{3 \times 1}]^{-1} \mathbf{K}_A^{-1} \mathbf{K}_A^{-T} \mathbf{K}_A^T p_A$$

E como $\mathbf{K}_A^{-T} \mathbf{K}_A^T = \mathbf{I}_{3 \times 3}$, pode-se simplificar para:

$$P = [\mathbf{I}_{3 \times 3} | \mathbf{0}_{3 \times 1}]^{-1} \mathbf{K}_A^{-1} p_A \quad (4.12)$$

Considerando o referencial da câmera A, os pontos da projeção da cônica absoluta ω_∞ devem satisfazer a equação (4.10):

$$\begin{aligned} &([\mathbf{I}_{3 \times 3} | \mathbf{0}_{3 \times 1}]^T \mathbf{K}_A^{-1} p_A)^T ([\mathbf{I}_{3 \times 3} | \mathbf{0}_{3 \times 1}]^T \mathbf{K}_A^{-1} p_A) = 0 \\ &(\mathbf{K}_A^{-1} p_A)^T [\mathbf{I}_{3 \times 3} | \mathbf{0}_{3 \times 1}] [\mathbf{I}_{3 \times 3} | \mathbf{0}_{3 \times 1}]^T \mathbf{K}_A^{-1} p_A = 0 \end{aligned}$$

Assim, a projeção da cônica absoluta na câmera A é dada pelos pontos p_A que satisfazem a equação:

$$p_A^T \mathbf{K}_A^{-T} \mathbf{K}_A^{-1} p_A = 0 \quad (4.13)$$

Definindo:

$$\mathbf{K}_A^{-T} \mathbf{K}_A^{-1} = \boldsymbol{\omega}_A \quad (4.14)$$

reescreve-se a equação (4.13):

$$p_A^T \boldsymbol{\omega}_A p_A = 0 \quad (4.15)$$

em que $\boldsymbol{\omega}_A$ é uma matriz 3×3 que relaciona os pontos da projeção da cônica absoluta na imagem A.

Observando a equação (4.15), nota-se que, tomando a câmera de referência como $[\mathbf{I}_{3 \times 3} | \mathbf{0}_{3 \times 1}]$, a projeção dos pontos da cônica absoluta na imagem A só depende dos parâmetros intrínsecos da câmera A. Os pontos da cônica absoluta podem ser transferidos para uma câmera i qualquer por meio de uma transformação do tipo:

$$\mathbf{E} = \begin{bmatrix} \mathbf{R}_{3 \times 3} & t_{3 \times 1} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

em que \mathbf{R} representa a rotação e t a translação entre a câmera i e a câmera de referência.

Como um ponto P da cônica absoluta tem a forma geral $P = [X \ Y \ Z \ 0]^T$, então a aplicação da transformação \mathbf{E} ao ponto P resulta em:

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 0 \end{bmatrix} = \begin{bmatrix} r_{11}X + r_{12}Y + r_{13}Z \\ r_{21}X + r_{22}Y + r_{23}Z \\ r_{31}X + r_{32}Y + r_{33}Z \\ 0 \end{bmatrix} \quad (4.16)$$

que é um ponto no infinito. A equação (4.16) pode ser reescrita como:

$$\begin{bmatrix} r_{11}X + r_{12}Y + r_{13}Z \\ r_{21}X + r_{22}Y + r_{23}Z \\ r_{31}X + r_{32}Y + r_{33}Z \\ 0 \end{bmatrix} = \mathbf{R}P$$

Como todos os pontos da cônica absoluta devem satisfazer a equação (4.10):

$$\begin{aligned} (\mathbf{R}P)^T(\mathbf{R}P) &= 0 \\ P^T \mathbf{R}^T \mathbf{R} P &= 0 \end{aligned}$$

Como a matriz de rotação \mathbf{R} é ortogonal, então $\mathbf{R}^T \mathbf{R} = \mathbf{I}$. Logo:

$$P^T P = 0$$

Esse resultado permite concluir que a representação da cônica absoluta no referencial da câmera i coincide com a representação da cônica absoluta no referencial da câmera A , ou seja, ω_∞ é invariante à rotação e translação entre as câmeras.

Com base nesse resultado, pode-se afirmar que a projeção dos pontos da cônica absoluta na câmera i deve satisfazer a equação:

$$p_i^T \omega_i p_i = 0 \quad (4.17)$$

4.8 Determinação da transformação H

Na câmara i , a projeção de um ponto P_P do espaço projetivo, medido no referencial da câmara A , é dada por:

$$p_i = \mathbf{M}_i P_P \quad (4.18)$$

em que p_i representa a projeção de P_P na câmara i e \mathbf{M}_i representa o modelo da câmara i (em relação ao referencial da câmara A), composta de rotação, translação e parâmetros intrínsecos desconhecidos. A equação (4.18) pode ser escrita como:

$$p_i = (\mathbf{M}_i \mathbf{H}^{-1}) (\mathbf{H} P_P) \quad (4.19)$$

em que \mathbf{H} representa a transformação que permite levar o objeto do espaço projetivo para o métrico, conforme a equação (1.01).

Como um ponto P do espaço métrico pode ser escrito como $P = \mathbf{H} P_P$, pode-se escrever a equação (4.19) como:

$$p_i = (\mathbf{M}_i \mathbf{H}^{-1}) P \quad (4.20)$$

Reprojetando P :

$$P = (\mathbf{M}_i \mathbf{H}^{-1})^+ p_i$$

em que $(\mathbf{M}_i \mathbf{H}^{-1})^+$ representa a matriz pseudoinversa de $\mathbf{M}_i \mathbf{H}^{-1}$. Como $\mathbf{A}^+ = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$, então:

$$\begin{aligned} P &= \left((\mathbf{M}_i \mathbf{H}^{-1})^T (\mathbf{M}_i \mathbf{H}^{-1}) \right)^{-1} (\mathbf{M}_i \mathbf{H}^{-1})^T p_i \\ P &= \left((\mathbf{H}^{-T} \mathbf{M}_i^T) (\mathbf{M}_i \mathbf{H}^{-1}) \right)^{-1} \mathbf{H}^{-T} \mathbf{M}_i^T p_i \\ P &= (\mathbf{M}_i \mathbf{H}^{-1})^{-1} (\mathbf{H}^{-T} \mathbf{M}_i^T)^{-1} \mathbf{H}^{-T} \mathbf{M}_i^T p_i \\ P &= \mathbf{H} \mathbf{M}_i^{-1} \mathbf{M}_i^{-T} \mathbf{H}^T \mathbf{H}^{-T} \mathbf{M}_i^T p_i \end{aligned}$$

Como $\mathbf{H}^T \mathbf{H}^{-T} = \mathbf{I}_{4 \times 4}$, tem-se:

$$P = \mathbf{H} \mathbf{M}_i^{-1} \mathbf{M}_i^{-T} \mathbf{M}_i^T p_i$$

E como $\mathbf{M}_i^{-T} \mathbf{M}_i^T = \mathbf{I}_{4 \times 4}$ chega-se a:

$$P = \mathbf{H} \mathbf{M}_i^{-1} p_i.$$

Considerando que a transformação \mathbf{H} é válida para qualquer ponto, então os pontos da cônica absoluta satisfazem a relação $(\mathbf{H} \mathbf{M}_i^{-1} p_i)^T (\mathbf{H} \mathbf{M}_i^{-1} p_i) = 0$. Rearranjando:

$$p_i^T \mathbf{M}_i^{-T} \mathbf{H}^T \mathbf{H} \mathbf{M}_i^{-1} p_i = 0 \quad (4.21)$$

As projeções p_i dos pontos da cônica absoluta na câmera i devem satisfazer, simultaneamente, as relações (4.13) e (4.21):

$$\begin{cases} p_i^T \mathbf{K}_i^{-T} \mathbf{K}_i^{-1} p_i = 0 \\ p_i^T \mathbf{M}_i^{-T} \mathbf{H}^T \mathbf{H} \mathbf{M}_i^{-1} p_i = 0 \end{cases} \quad (4.22)$$

O sistema da equação (4.22) somente é satisfeito se:

$$\mathbf{K}_i^{-T} \mathbf{K}_i^{-1} \approx \mathbf{M}_i^{-T} \mathbf{H}^T \mathbf{H} \mathbf{M}_i^{-1} \quad (4.23)$$

em que \mathbf{K}_i representa a matriz dos parâmetros intrínsecos da câmera i , \mathbf{M}_i representa o modelo da câmera i , \mathbf{H} representa a transformação que leva do espaço projetivo para o espaço métrico e \approx denota proporcionalidade.

Na bibliografia, é comum representar a equação (4.23) em função de $\boldsymbol{\omega}_i^* = \mathbf{K}_i \mathbf{K}_i^T$, que é a representação *dual* da projeção da cônica absoluta na imagem i . Dessa forma, é necessário reescrever a equação (4.23) como:

$$\mathbf{K}_i \mathbf{K}_i^T \approx \mathbf{M}_i \mathbf{H}^{-1} \mathbf{H}^{-T} \mathbf{M}_i^T \quad (4.24)$$

Explicitando as matrizes de (4.24):

$$\begin{bmatrix} f_{x_i} & s_i & u_{0_i} \\ 0 & f_{y_i} & v_{0_i} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_{x_i} & s_i & u_{0_i} \\ 0 & f_{y_i} & v_{0_i} \\ 0 & 0 & 1 \end{bmatrix}^T \approx \mathbf{M}_i \begin{bmatrix} \mathbf{K}_A^{-1} & 0 \\ H_\infty^T & d \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{K}_A^{-1} & 0 \\ H_\infty^T & d \end{bmatrix}^{-T} \mathbf{M}_i^T \quad (4.25)$$

em que f_{x_i} e f_{y_i} representam, respectivamente, as relações entre a distância focal e as dimensões horizontal e vertical do pixel da câmera i , s_i representa o *skew* da câmera i , u_{0_i} e v_{0_i} representam, respectivamente, a posição horizontal e vertical do ponto principal da câmera i , \mathbf{M}_i representa o modelo da câmera i , \mathbf{K}_A representa a matriz de parâmetros intrínsecos da câmera de referência, Π_∞^T e d representam os coeficientes do plano no infinito. Resolvendo as multiplicações matriciais:

$$\begin{bmatrix} f_{x_i}^2 + s_i^2 u_{0_i}^2 & f_{y_i} s_i + u_{0_i} v_{0_i} & u_{0_i} \\ f_{y_i} s_i + u_{0_i} v_{0_i} & f_{y_i}^2 + v_{0_i}^2 & v_{0_i} \\ u_{0_i} & v_{0_i} & 1 \end{bmatrix} \approx \mathbf{M}_i \begin{bmatrix} f_{y_A}^2 + s_A^2 u_{0_A}^2 & f_{y_A} s_A + u_{0_A} v_{0_A} & u_{0_A} & -\left(\left(f_{x_A}^2 + s_A^2 + u_{0_A}^2\right) \mathbf{a} + \left(f_{y_A} s_A + u_{0_A} v_{0_A}\right) \mathbf{b} + u_{0_A} \mathbf{c}\right) / d \\ f_{y_A} s_A + u_{0_A} v_{0_A} & f_{y_A}^2 + v_{0_A}^2 & v_{0_A} & -\left(\left(f_{y_A} s_A + u_{0_A} v_{0_A}\right) \mathbf{a} + \left(f_{y_A}^2 + v_{0_A}^2\right) \mathbf{b} + v_{0_A} \mathbf{c}\right) / d \\ u_{0_A} & v_{0_A} & 1 & -\left(u_{0_A} \mathbf{a} + v_{0_A} \mathbf{b} + \mathbf{c}\right) / d \\ -\left(\left(f_{x_A}^2 + s_A^2 + u_{0_A}^2\right) \mathbf{a} + \left(f_{y_A} s_A + u_{0_A} v_{0_A}\right) \mathbf{b} + u_{0_A} \mathbf{c}\right) / d & -\left(\left(f_{y_A} s_A + u_{0_A} v_{0_A}\right) \mathbf{a} + \left(f_{y_A}^2 + v_{0_A}^2\right) \mathbf{b} + v_{0_A} \mathbf{c}\right) / d & -\left(u_{0_A} \mathbf{a} + v_{0_A} \mathbf{b} + \mathbf{c}\right) / d & \left(\left(b v_{0_A}\right)^2 + 2 b v_{0_A} \left(\mathbf{c} + a u_{0_A}\right) + \left(b f_{y_A} + a s_A\right)^2 + \left(\mathbf{c} + a u_{0_A}\right)^2 + \left(a f_{x_A}\right)^2 + 1\right) / d^2 \end{bmatrix} \mathbf{M}_i^T \quad (4.26)$$

Os modelos das duas ou três primeiras câmeras podem ser obtidos em relação ao referencial da câmera A por meio da matriz fundamental ou do tensor trifocal, respectivamente, utilizando as equações (3.02) ou (3.07). A partir dessas matrizes, é possível obter a reconstrução do objeto no espaço projetivo por triangulação, conforme procedimento descrito no Capítulo 3 deste trabalho.

Os outros modelos de câmeras \mathbf{M}_i são determinados por calibração implícita, conforme proposto por Gonzalez e Woods (2000), a partir da reconstrução projetiva do objeto e das projeções dos pontos do objeto nas imagens das demais câmeras, pelo equacionamento a seguir.

A projeção de um ponto P , medido na câmera de referência, na imagem i é dada por $p = \mathbf{M}_i P$, em que \mathbf{M}_i é o modelo da câmera i . Essa equação pode ser expandida em:

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Como as coordenadas x e y da projeção p são dadas por $x = u/w$ e $y = v/w$, pode-se escrever:

$$\begin{aligned} xw &= m_{11}X + m_{12}Y + m_{13}Z + m_{14} \\ yw &= m_{21}X + m_{22}Y + m_{23}Z + m_{24} \\ w &= m_{31}X + m_{32}Y + m_{33}Z + m_{34} \end{aligned} \quad (4.27)$$

A substituição da terceira equação de (4.27) nas duas primeiras resulta nas duas equações a seguir, que possuem 12 coeficientes desconhecidos (GONZALEZ e WOODS, 2000):

$$\begin{aligned} m_{11}X + m_{12}Y + m_{13}Z - m_{31}xX - m_{32}xY - m_{33}xZ - m_{34}x + m_{14} &= 0 \\ m_{21}X + m_{22}Y + m_{23}Z - m_{31}yX - m_{32}yY - m_{33}yZ - m_{34}y + m_{24} &= 0 \end{aligned} \quad (4.28)$$

Para ser possível a determinação dos 12 coeficientes m_{ij} das equações (4.28), são necessários 6 pontos conhecidos no espaço e suas respectivas projeções na imagem da câmera i . Podem ser utilizados mais pontos, gerando um sistema de equações super especificado que pode ser resolvido por minimização de erro.

Determinadas as matrizes \mathbf{M}_i , a solução de (4.26) permite determinar os parâmetros da matriz \mathbf{K}_A e a posição do plano no infinito, necessários para determinar a transformação \mathbf{H} .

De acordo com Pollefeys et al (1999), se for conhecido um número suficiente de matrizes projetivas de câmeras e puder ser inserida pelo menos uma restrição aos parâmetros das câmeras, sempre será possível determinar completamente a transformação que leva do espaço projetivo para o métrico.

O caso mais geral consiste em considerar *skew* nulo para todas as câmeras, que permitirá escrever ω_i^* na forma:

$$\omega_i^* = \begin{bmatrix} f_{x_i} & 0 & u_{0_i} \\ 0 & f_{y_i} & v_{0_i} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_{x_i} & 0 & u_{0_i} \\ 0 & f_{y_i} & v_{0_i} \\ 0 & 0 & 1 \end{bmatrix}^{-T} = \begin{bmatrix} f_{x_i}^2 + u_{0_i}^2 & u_{0_i}v_{0_i} & u_{0_i} \\ u_{0_i}v_{0_i} & f_{y_i}^2 + v_{0_i}^2 & v_{0_i} \\ u_{0_i} & v_{0_i} & 1 \end{bmatrix} \quad (4.29)$$

em que f_{x_i} e f_{y_i} representam, respectivamente, as relações entre a distância focal e as dimensões horizontal e vertical do pixel da câmera i , u_{0_i} e v_{0_i} representam, respectivamente, a posição horizontal e vertical do ponto principal da câmera i .

Fazendo-se a troca de variáveis a seguir para simplificar a notação:

$$\mathbf{A} = \begin{bmatrix} f_{x_i}^2 + u_{0_i}^2 & u_{0_i}v_{0_i} & u_{0_i} \\ u_{0_i}v_{0_i} & f_{y_i}^2 + v_{0_i}^2 & v_{0_i} \\ u_{0_i} & v_{0_i} & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & 1 \end{bmatrix} \quad (4.30)$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{K}_A \mathbf{K}_A^T & -\mathbf{K}_A \mathbf{K}_A^T \Pi_\infty^T \\ -\Pi_\infty \mathbf{K}_A \mathbf{K}_A^T & \Pi_\infty \mathbf{K}_A \mathbf{K}_A^T \Pi_\infty^T \end{bmatrix} = \begin{bmatrix} b_1 & b_2 & b_3 & b_4 \\ b_2 & b_5 & b_6 & b_7 \\ b_3 & b_6 & 1 & b_8 \\ b_4 & b_7 & b_8 & b_9 \end{bmatrix} \quad (4.31)$$

$$\mathbf{M}_i = \begin{bmatrix} m_{11_i} & m_{12_i} & m_{13_i} & m_{14_i} \\ m_{21_i} & m_{22_i} & m_{23_i} & m_{24_i} \\ m_{31_i} & m_{32_i} & m_{33_i} & m_{34_i} \end{bmatrix} \quad (4.32)$$

pode-se escrever (4.26) na forma:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & 1 \end{bmatrix} = \begin{bmatrix} m_{11_i} & m_{12_i} & m_{13_i} & m_{14_i} \\ m_{21_i} & m_{22_i} & m_{23_i} & m_{24_i} \\ m_{31_i} & m_{32_i} & m_{33_i} & m_{34_i} \end{bmatrix} \cdot \begin{bmatrix} b_1 & b_2 & b_3 & b_4 \\ b_2 & b_5 & b_6 & b_7 \\ b_3 & b_6 & 1 & b_8 \\ b_4 & b_7 & b_8 & b_9 \end{bmatrix} \cdot \begin{bmatrix} m_{11_i} & m_{21_i} & m_{31_i} \\ m_{12_i} & m_{22_i} & m_{32_i} \\ m_{13_i} & m_{23_i} & m_{33_i} \\ m_{14_i} & m_{24_i} & m_{34_i} \end{bmatrix} \quad (4.33)$$

Por economia de espaço, o desenvolvimento da equação (4.33) continua no Quadro 4.1.

A partir de (4.30), é possível escrever o termo a_{12} na forma do produto de a_{13} por a_{23} :

$$a_{12} = a_{13} a_{23} \quad (4.34)$$

Usando as equações compactas de (4.01), a equação (4.34) pode ser escrita como:

$$\left(\mathbf{M}_i \mathbf{B} \mathbf{M}_{3_i} \right) \left(\mathbf{M}_{2_i} \mathbf{B} \mathbf{M}_{3_i} \right) - \mathbf{M}_i \mathbf{B} \mathbf{M}_{2_i} = 0 \quad (4.35)$$

que gera a seguinte equação não linear nos termos em \mathbf{B} :

$$\begin{aligned} & (m_{11_i} m_{31_i} b_1 + (m_{11_i} m_{32_i} + m_{12_i} m_{31_i}) b_2 + (m_{11_i} m_{33_i} + m_{13_i} m_{31_i}) b_3 + (m_{11_i} m_{34_i} + m_{14_i} m_{31_i}) b_4 + m_{12_i} m_{32_i} b_5 + (m_{12_i} m_{33_i} + m_{13_i} m_{32_i}) b_6 + (m_{12_i} m_{34_i} + m_{14_i} m_{32_i}) b_7 + (m_{13_i} m_{34_i} + m_{14_i} m_{33_i}) b_8 + m_{14_i} m_{34_i} b_9 + m_{13_i} m_{33_i}) \cdot \\ & (m_{21_i} m_{31_i} b_1 + (m_{21_i} m_{32_i} + m_{22_i} m_{31_i}) b_2 + (m_{21_i} m_{33_i} + m_{23_i} m_{31_i}) b_3 + (m_{21_i} m_{34_i} + m_{24_i} m_{31_i}) b_4 + m_{22_i} m_{32_i} b_5 + (m_{22_i} m_{33_i} + m_{23_i} m_{32_i}) b_6 + (m_{22_i} m_{34_i} + m_{24_i} m_{32_i}) b_7 + (m_{23_i} m_{34_i} + m_{24_i} m_{33_i}) b_8 + m_{24_i} m_{34_i} b_9 + m_{23_i} m_{33_i}) - \\ & - m_{11_i} m_{21_i} b_1 - (m_{11_i} m_{22_i} + m_{12_i} m_{21_i}) b_2 - (m_{11_i} m_{23_i} + m_{13_i} m_{21_i}) b_3 - (m_{11_i} m_{24_i} + m_{14_i} m_{21_i}) b_4 - m_{12_i} m_{22_i} b_5 - (m_{12_i} m_{23_i} + m_{13_i} m_{22_i}) b_6 - (m_{12_i} m_{24_i} + m_{14_i} m_{22_i}) b_7 - (m_{13_i} m_{24_i} + m_{14_i} m_{23_i}) b_8 + m_{14_i} m_{24_i} b_9 + m_{13_i} m_{23_i} = 0 \end{aligned} \quad (4.36)$$

A equação (4.36) relaciona a imagem A (imagem de referência) com a imagem i , de modo que a cada relacionamento de A com i gera uma equação com 8 parâmetros em \mathbf{B} (uma vez que o *skew* foi considerado nulo e, portanto, $b_2 = b_3 b_6$), de modo que são necessárias 8 imagens, além da imagem de referência, para se obter a transformação \mathbf{H} .

Efetuada a primeira multiplicação matricial:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & 1 \end{bmatrix} \begin{bmatrix} m_{11}b_1 + m_{12}b_2 + m_{13}b_3 + m_{14}b_4 & m_{11}b_2 + m_{12}b_5 + m_{13}b_6 + m_{14}b_7 & m_{11}b_3 + m_{12}b_6 + m_{13}b_7 + m_{14}b_8 & m_{11}b_4 + m_{12}b_7 + m_{13}b_8 + m_{14}b_9 \\ m_{21}b_1 + m_{22}b_2 + m_{23}b_3 + m_{24}b_4 & m_{21}b_1 + m_{22}b_2 + m_{23}b_3 + m_{24}b_4 & m_{21}b_1 + m_{22}b_2 + m_{23}b_3 + m_{24}b_4 & m_{21}b_1 + m_{22}b_2 + m_{23}b_3 + m_{24}b_4 \\ m_{31}b_1 + m_{32}b_2 + m_{33}b_3 + m_{34}b_4 & m_{31}b_1 + m_{32}b_2 + m_{33}b_3 + m_{34}b_4 & m_{31}b_1 + m_{32}b_2 + m_{33}b_3 + m_{34}b_4 & m_{31}b_1 + m_{32}b_2 + m_{33}b_3 + m_{34}b_4 \end{bmatrix} \cdot \begin{bmatrix} m_{11} & m_{21} & m_{31} \\ m_{12} & m_{22} & m_{32} \\ m_{13} & m_{23} & m_{33} \\ m_{14} & m_{24} & m_{34} \end{bmatrix}$$

Efetuada a segunda multiplicação matricial:

$$\begin{aligned} a_{11} &= m_{11}^2 b_1 + 2m_{11}m_{12}b_2 + 2m_{11}m_{13}b_3 + 2m_{11}m_{14}b_4 + m_{12}^2 b_4 + m_{13}^2 b_4 + 2m_{12}m_{13}b_4 + m_{14}^2 b_4 + m_{13}^2 b_4 \\ a_{12} &= m_{11}m_{21}b_1 + (m_{11}m_{21}b_2 + m_{13}m_{21}b_3 + m_{14}m_{21}b_4 + m_{12}m_{21}b_5 + m_{13}m_{21}b_6 + m_{14}m_{21}b_7 + m_{12}m_{21}b_8 + m_{13}m_{21}b_9 + m_{14}m_{21}b_{10}) \\ a_{13} &= m_{11}m_{31}b_1 + (m_{11}m_{31}b_2 + m_{12}m_{31}b_3 + m_{13}m_{31}b_4 + m_{14}m_{31}b_5 + m_{12}m_{31}b_6 + m_{13}m_{31}b_7 + m_{14}m_{31}b_8 + m_{12}m_{31}b_9 + m_{13}m_{31}b_{10}) \\ a_{22} &= m_{21}^2 b_1 + 2m_{21}m_{22}b_2 + 2m_{21}m_{23}b_3 + 2m_{21}m_{24}b_4 + m_{22}^2 b_4 + m_{23}^2 b_4 + 2m_{22}m_{23}b_4 + m_{24}^2 b_4 + m_{23}^2 b_4 \\ a_{23} &= m_{21}m_{31}b_1 + (m_{21}m_{31}b_2 + m_{22}m_{31}b_3 + m_{23}m_{31}b_4 + m_{24}m_{31}b_5 + m_{22}m_{31}b_6 + m_{23}m_{31}b_7 + m_{24}m_{31}b_8 + m_{22}m_{31}b_9 + m_{23}m_{31}b_{10}) \end{aligned}$$

É possível escrever essas equações de forma mais compacta:

$$a_{11} = M_{11} B M_{11}^T \quad a_{12} = M_{11} B M_{21}^T \quad a_{13} = M_{11} B M_{31}^T \quad a_{22} = M_{21} B M_{21}^T \quad a_{23} = M_{21} B M_{31}^T \quad (4.37)$$

em que o índice i representa a câmera e o índice j representa a linha da matriz M_{ij} .

Quadro 4.1 – Desenvolvimento da equação (4.33)

Uma vez determinadas as incógnitas de \mathbf{B} , determinam-se os parâmetros intrínsecos da câmera de referência e os coeficientes do plano no infinito, comparando-se termo a termo as incógnitas de \mathbf{B} com os termos da equação (4.26):

- as coordenadas do ponto principal são obtidas diretamente, já que $b_3 = u_0$ e $b_6 = v_0$;
- a partir de v_0 e de b_5 , obtém-se f_y , fazendo $f_y = \sqrt{b_5 - v_0^2}$;
- de forma análoga, obtém-se f_x , fazendo $f_x = \sqrt{b_1 - u_0^2}$;
- para determinar a , b e c , é necessário resolver o sistema de equações:

$$-\mathbf{K}_A \mathbf{K}_A^T \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} b_4 \\ b_7 \\ b_8 \end{bmatrix}$$

- por fim, d pode ser obtido pela equação:

$$d = \sqrt{(a^2 f_x^2 + b^2 f_y^2 + (c + v_0 b + a u_0)^2 + 1)} / b_9$$

4.8.1 Uso de parâmetros conhecidos nas câmeras

O sistema gerado pela aplicação da equação (4.35) a 9 imagens pode ser reduzido e resolvido linearmente se forem impostas outras restrições (além de *skew* nulo) nos parâmetros intrínsecos das câmeras.

Essas informações têm o intuito de viabilizar a solução de (4.25), mas reduzindo o número de imagens necessárias para a reconstrução. Essas informações podem ser dadas de várias maneiras:

- distâncias focais conhecidas;
- distâncias focais desconhecidas, porém constante entre as câmeras utilizadas;
- pontos principais conhecidos ou constante entre as imagens;
- relação de aspecto conhecida (f_y/f_x);
- outras combinações dos parâmetros intrínsecos.

Conforme o número de informações disponíveis sobre os parâmetros intrínsecos, o número de imagens necessárias para obter a transformação necessária está apresentado na Tabela 4.2.

Tabela 4.2 - Quadro-resumo da determinação dos parâmetros

Parâmetros conhecidos	Parâmetros constantes entre as câmeras	Imagens necessárias
• <i>skew</i> nulo	–	9
–	• ponto principal em x • ponto principal em y	5
• ponto principal na origem em x • ponto principal na origem em y • <i>skew</i> nulo	–	5
• relação de aspecto unitária • ponto principal na origem em x • ponto principal na origem em y • <i>skew</i> nulo	–	3
• relação de aspecto unitária • ponto principal na origem em x • ponto principal na origem em y • <i>skew</i> nulo	• distâncias focais	2

A seguir, serão apresentados casos nos quais foram aplicadas mais restrições sobre parâmetros intrínsecos das câmeras, com o intuito de se determinar equacionamento apropriado para diversas situações.

Assim, dependendo do número de informações conhecidas a respeito desses parâmetros intrínsecos (distância focal conhecida ou igual para todas as imagens etc.), é possível obter um equacionamento específico para cada situação, a partir da equação (4.25). O equacionamento segue, em parte, do que foi proposto por Pollefeys et al (1999) e por Heyden e Aström (1996).

Caso 1 – Ponto principal na origem

Supondo que o ponto principal esteja na origem do sistema de cada câmera, a matriz **A** pode ser escrita como:

$$\mathbf{A} = \begin{bmatrix} f_{x_i}^2 + s_i^2 & s_i f_{y_i} & 0 \\ s_i f_{y_i} & f_{y_i}^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Com essas restrições, os dois elementos a_{13} e a_{23} da matriz **A** anulam-se, resultando em restrições que podem ser utilizadas na determinação das equações no parâmetros de **B**.

Referindo-se às equações (4.38):

$$a_{12} = \mathbf{M}_i \mathbf{B} \mathbf{M}_{2_i}^T = 0$$

$$a_{13} = \mathbf{M}_i \mathbf{B} \mathbf{M}_{3_i}^T = 0$$

e particularizando-as para $a_{13} = 0$ e $a_{23} = 0$, tem-se duas equações para a imagem i :

$$\left\{ \begin{array}{l} a_{12} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \end{bmatrix} \begin{bmatrix} b_1 & b_2 & 0 & b_4 \\ b_2 & b_5 & 0 & b_7 \\ 0 & 0 & 1 & b_8 \\ b_4 & b_7 & b_8 & b_9 \end{bmatrix} \begin{bmatrix} m_{21} \\ m_{22} \\ m_{23} \\ m_{24} \end{bmatrix} = 0 \\ a_{13} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \end{bmatrix} \begin{bmatrix} b_1 & b_2 & 0 & b_4 \\ b_2 & b_5 & 0 & b_7 \\ 0 & 0 & 1 & b_8 \\ b_4 & b_7 & b_8 & b_9 \end{bmatrix} \begin{bmatrix} m_{31} \\ m_{32} \\ m_{33} \\ m_{34} \end{bmatrix} = 0 \end{array} \right. \quad (4.38)$$

A equação (4.25) relaciona a imagem A (que é utilizada como imagem de referência) com a imagem i , de modo que a cada relacionamento de A com i gera duas equações na forma de (4.01).

Como cada nova imagem gera duas equações (pelo relacionamento com a imagem A – que serve como referência), serão necessárias no mínimo cinco imagens (quatro relacionamentos mais a imagem de referência) para determinar as sete incógnitas da matriz \mathbf{B} (uma vez que os elementos b_3 e b_6 são nulos).

Depois de determinada a matriz \mathbf{B} , deve-se seguir o procedimento descrito no caso geral, para determinar os parâmetros intrínsecos e os coeficientes do plano no infinito.

Caso 2 – Relação de aspecto unitária, ponto principal na origem e *skew* nulo

Esse é um exemplo da concatenação de três restrições independentes.

Considerar que a relação de aspecto é unitária equivale a dizer que $f_{x_i} = f_{y_i}$.

A matriz \mathbf{A} , nesta situação, fica reduzida a:

$$\mathbf{A} = \begin{bmatrix} f_i^2 & 0 & 0 \\ 0 & f_i^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Percebe-se que, de forma semelhante ao Caso 1, os elementos da matriz a_{13} e a_{23} são nulos, pela imposição da restrição do ponto principal. A imposição das restrições de relação de aspecto unitária e *skew* nulofaz com que o elemento a_{12} seja nulo e que os elementos a_{11} e a_{22} sejam iguais. Matematicamente:

$$\begin{cases} a_{12} = 0 \\ a_{13} = 0 \\ a_{23} = 0 \\ a_{11} = a_{22} \rightarrow a_{11} - a_{22} = 0 \end{cases} \quad (4.39)$$

Utilizando novamente as equações (4.01):

$$\begin{cases} a_{12} = \mathbf{M}_{1_i} \mathbf{B} \mathbf{M}_{2_i} = 0 \\ a_{13} = \mathbf{M}_{1_i} \mathbf{B} \mathbf{M}_{3_i} = 0 \\ a_{23} = \mathbf{M}_{2_i} \mathbf{B} \mathbf{M}_{3_i} = 0 \\ a_{11} - a_{22} = \mathbf{M}_{1_i} \mathbf{B} \mathbf{M}_{1_i} - \mathbf{M}_{2_i} \mathbf{B} \mathbf{M}_{2_i} = 0 \end{cases} \quad (4.40)$$

$$\begin{cases} a_{12} = [m_{11} \ m_{12} \ m_{13} \ m_{14}] \begin{bmatrix} b_1 & 0 & 0 & b_4 \\ 0 & b_1 & 0 & b_7 \\ 0 & 0 & 1 & b_8 \\ b_4 & b_7 & b_8 & b_9 \end{bmatrix} \begin{bmatrix} m_{21} \\ m_{22} \\ m_{23} \\ m_{24} \end{bmatrix} = 0 \\ a_{13} = [m_{11} \ m_{12} \ m_{13} \ m_{14}] \begin{bmatrix} b_1 & 0 & 0 & b_4 \\ 0 & b_1 & 0 & b_7 \\ 0 & 0 & 1 & b_8 \\ b_4 & b_7 & b_8 & b_9 \end{bmatrix} \begin{bmatrix} m_{31} \\ m_{32} \\ m_{33} \\ m_{34} \end{bmatrix} = 0 \\ a_{23} = [m_{21} \ m_{22} \ m_{23} \ m_{24}] \begin{bmatrix} b_1 & 0 & 0 & b_4 \\ 0 & b_1 & 0 & b_7 \\ 0 & 0 & 1 & b_8 \\ b_4 & b_7 & b_8 & b_9 \end{bmatrix} \begin{bmatrix} m_{31} \\ m_{32} \\ m_{33} \\ m_{34} \end{bmatrix} = 0 \\ a_{11} - a_{22} = [m_{11} \ m_{12} \ m_{13} \ m_{14}] \begin{bmatrix} b_1 & 0 & 0 & b_4 \\ 0 & b_1 & 0 & b_7 \\ 0 & 0 & 1 & b_8 \\ b_4 & b_7 & b_8 & b_9 \end{bmatrix} \begin{bmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{14} \end{bmatrix} - [m_{21} \ m_{22} \ m_{23} \ m_{24}] \begin{bmatrix} b_1 & 0 & 0 & b_4 \\ 0 & b_1 & 0 & b_7 \\ 0 & 0 & 1 & b_8 \\ b_4 & b_7 & b_8 & b_9 \end{bmatrix} \begin{bmatrix} m_{21} \\ m_{22} \\ m_{23} \\ m_{24} \end{bmatrix} = 0 \end{cases}$$

Como cada nova imagem gera quatro equações distintas, então serão necessárias ao menos três imagens para que se possam determinar os cinco parâmetros da matriz \mathbf{B} .

Caso 3 – Parâmetros intrínsecos constantes em todas as câmeras

Nesta nova situação, não há uma câmera tomada como referência, mas as imagens podem ser tomadas aos pares para satisfazer a equação (4.26).

As restrições sobre os parâmetros intrínsecos são diferentes das discutidas nos casos anteriores, já que essas relações não serão obtidas apenas pela comparação dos elementos da matriz \mathbf{A} , mas pelos parâmetros constantes em todas as câmeras.

Caso 3.1 – Ponto principal constante

Se o ponto principal for igual em todas as imagens, pode-se escrever:

$$\begin{cases} a_{13_i} = a_{13_j} \\ a_{23_i} = a_{23_j} \end{cases}$$

ou, de forma mais completa:

$$\begin{cases} \mathbf{M}_{1_i} \mathbf{B} \mathbf{M}_{3_i}^T = \mathbf{M}_{1_j} \mathbf{B} \mathbf{M}_{3_j}^T \\ \mathbf{M}_{2_i} \mathbf{B} \mathbf{M}_{3_i}^T = \mathbf{M}_{2_j} \mathbf{B} \mathbf{M}_{3_j}^T \end{cases} \quad (4.41)$$

para $i \neq j$.

Como as imagens devem ser combinadas duas a duas, então se forem conhecidas m imagens, então cada uma das equações (4.41) irá aparecer $m-1$ vezes. Já que existem duas equações independentes, serão geradas $2(m-1)$ equações.

Daí, para que se tenham nove equações que permitam determinar os nove parâmetros de \mathbf{B} , são necessárias pelo menos cinco imagens, mais uma imagem de referência.

Caso 3.2 – Relação de aspecto unitária, ponto principal na origem, *skew* nulo e distâncias focais constantes entre as câmeras

Retomando as equações (4.40) e adicionando essa nova restrição:

$$\begin{cases} a_{1i} = a_{1j} \\ a_{2i} = a_{2j} \end{cases}$$

tem-se um conjunto de equações na forma:

$$\begin{cases} \mathbf{M}_i \mathbf{B} \mathbf{M}_{3_i} = 0 \\ \mathbf{M}_i \mathbf{B} \mathbf{M}_{2_i} = 0 \\ \mathbf{M}_{2_i} \mathbf{B} \mathbf{M}_{3_i} = 0 \\ \mathbf{M}_i \mathbf{B} \mathbf{M}_{1_i} = \mathbf{M}_{2_i} \mathbf{B} \mathbf{M}_{2_i} \\ \mathbf{M}_i \mathbf{B} \mathbf{M}_{1_i}^T = \mathbf{M}_{1_j} \mathbf{B} \mathbf{M}_{1_j}^T \\ \mathbf{M}_{2_i} \mathbf{B} \mathbf{M}_{2_i}^T = \mathbf{M}_{2_j} \mathbf{B} \mathbf{M}_{2_j}^T \end{cases}$$

para $i \neq j$.

Desse conjunto de seis equações, as quatro primeiras valem para cada imagem disponível, enquanto que as demais valem a cada duas imagens, da mesma forma que o Caso 3.1. Assim, para m imagens, serão geradas $4m + 2(m - 1)$ equações.

Para determinar cinco parâmetros de \mathbf{B} (lembrando que neste caso os elementos b_2, b_3 e b_6 são nulos e que $b_1=b_5$) serão necessárias pelo menos duas imagens.

Para todos os casos apresentados, não será possível determinar uma escala geral do objeto, pois todos os sistemas de equações resultantes são homogêneos. Isto significa que o objeto recuperado em coordenadas métricas terá dimensões proporcionais às do objeto original.

4.8.2 Quantas imagens são necessárias para obter uma reconstrução métrica?

De acordo com a Tabela 4.2 e pela análise dos casos apresentados, é possível estabelecer uma relação entre o número de parâmetros conhecidos, o número de parâmetros iguais entre as imagens e o número de imagens necessárias para determinar os nove parâmetros da matriz **B**.

Sendo:

- c o número de parâmetros intrínsecos conhecidos
- i o número de parâmetros intrínsecos iguais entre as imagens e
- m o número de imagens necessárias para obter a homografia que leva a reconstrução projetiva para a métrica,

pode-se escrever:

$$c m + i(m - 1) \geq 9$$

Reescrevendo:

$$m \geq \frac{9+i}{c+i} \quad (4.42)$$

A equação (4.42) valida qualquer um dos casos apresentados anteriormente bem como qualquer outro que possa ser obtido a partir das regras aqui apresentadas (POLLEFEYS et al., 1999).

5. IMPLEMENTAÇÃO DO PROTÓTIPO

Neste capítulo será apresentado um protótipo de *software* que permite ao usuário selecionar um conjunto de imagens que serão utilizadas na reconstrução de um objeto e de uma cena no espaço métrico, sem se conhecer nenhuma informação a respeito da cena ou das câmeras envolvidas no processo de reconstrução.

Na Figura 5.1 encontra-se um fluxograma que representa de forma geral o funcionamento do protótipo. O protótipo foi desenvolvido no ambiente Embarcadero Delphi.

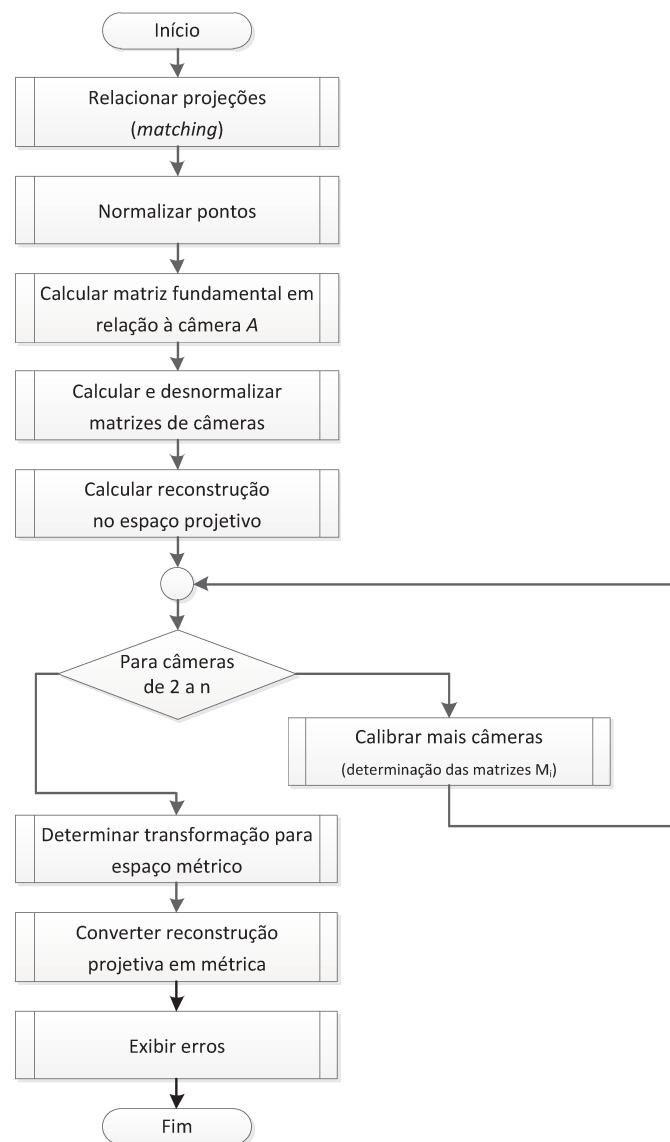


Figura 5.1 - Visão geral do protótipo

5.1 Detalhamento do fluxograma

A primeira parte desse algoritmo consiste em estabelecer uma relação entre as projeções dos pontos em duas imagens, de modo que a primeira imagem será utilizada como referência na reconstrução. Na literatura, existem diversas técnicas que permitem determinar essa correspondência de maneira automática, como os estudados por Nogueira (2007), porém como o foco deste trabalho é o de demonstrar a determinação da reconstrução métrica a partir de várias imagens, optou-se pelo estabelecimento dessa correspondência de forma manual, de modo que uma *interface* computacional foi desenvolvida para que seja possível relacionar as projeções dos pontos manualmente.

A normalização isotrópica aplicada nesse protótipo segue o que foi apresentado no Capítulo 2 deste trabalho.

A partir das imagens escolhidas pelo usuário, o protótipo seleciona automaticamente as duas primeiras imagens para calcular a matriz fundamental. O cálculo é realizado com as projeções normalizadas a partir do que foi exposto na seção 2.1.1 deste trabalho.

Os modelos das duas primeiras câmeras para pontos no espaço projetivo são determinadas por meio das equações (3.02). Entretanto, como os modelos são válidos para as projeções normalizadas é necessário determiná-las diretamente em sua forma desnormalizada por meio da equação (3.09).

A partir dos modelos das duas primeiras câmeras, realiza-se um processo de triangulação a fim de se obter uma reconstrução no espaço projetivo. Essa triangulação para duas câmeras está descrita pela equação (3.06).

O cálculo dos modelos de câmeras M_i para as demais imagens é realizado por meio da calibração realizada a partir da reconstrução no espaço projetivo obtida no passo anterior e das projeções dos pontos da terceira imagem em diante. O procedimento implementado no protótipo baseia-se na determinação dos coeficientes da equação (4.28).

De modo que se possa determinar a transformação que leva o objeto do espaço projetivo para o métrico por meio de métodos lineares, a técnica descrita na seção 4.8 deverá receber simplificações sob a forma de restrições impostas aos parâmetros intrínsecos das câmeras.

No protótipo, foram implementadas três situações diferentes, que representam os Casos 1, 2 e 3.1 descritos na seção 0. Podem-se utilizar até dez imagens em qualquer um dos casos selecionados, gerando sistemas de equações super especificados para cada caso.

O último passo da reconstrução, de acordo com o fluxograma da Figura 5.1, é a aplicação da transformação \mathbf{H} aos pontos obtidos durante a fase de reconstrução no espaço projetivo. Uma vez calculada a transformação, determina-se a reconstrução do objeto no espaço métrico pela equação:

$$P_M = \mathbf{H}P$$


em que P representa um ponto no espaço projetivo e P_M representa um ponto no espaço métrico.

Uma vez determinado o objeto no espaço métrico, o protótipo apresenta uma janela onde são exibidos os erros (mínimo, máximo e médio) nas reprojeções dos pontos do objeto recuperado no espaço métrico.

5.2 Interfaces do protótipo

O programa permite a visualização do objeto reconstruído nas imagens da cena ou em um ambiente tridimensional por onde o usuário poderá navegar.

A interface principal do programa está apresentada na Figura 5.2.

Na Figura 5.2, pode-se ver, à esquerda, uma das imagens selecionadas pelo usuário, com pontos (marcados com +) relacionados manualmente. Para reduzir o erro na medição da projeção, ao selecionar novo ponto (clicando no botão  e selecionando o ponto na imagem), abre-se uma janela de visualização em *zoom*, conforme ilustrado na Figura 5.3. O nível de *zoom* pode ser selecionado pelo usuário com o uso do botão de *scroll* do *mouse*.

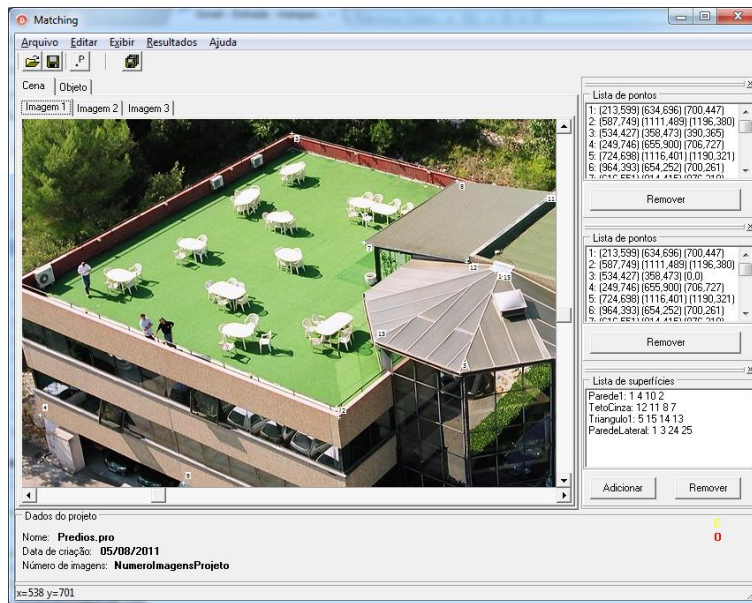


Figura 5.2 – Interface principal do protótipo



Figura 5.3 – Exemplo de zoom para facilitar a tarefa se selecionar pontos na imagem

Para efeitos da reconstrução, o usuário pode criar superfícies que serão renderizadas ao se exibir o objeto no ambiente 3D.

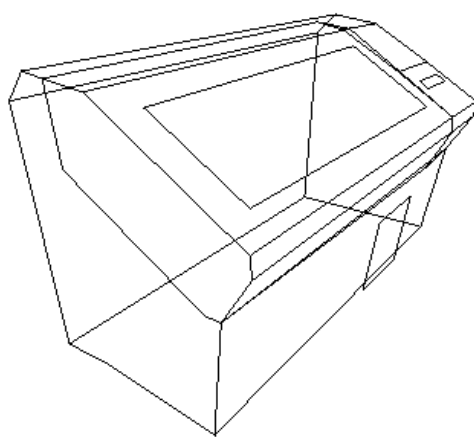
O usuário pode selecionar um conjunto de imagens para a cena e outro conjunto de imagens para um objeto a ser inserido digitalmente na cena. No canto superior esquerdo, pode-se selecionar a visualização do conjunto de imagens da cena ou do objeto. À direita, podem-se exibir as listas de pontos e de superfícies do objeto.

No menu Resultados | Relacionar cena e objeto, o usuário tem a possibilidade de relacionar pontos na cena e no objeto, de modo que o objeto seja reprojetoado nas imagens da cena, criando o efeito de inserção digital do objeto na cena. Há a opção de considerar a posição exata (fazendo o relacionamento entre os pontos da cena e do objeto) ou considerar somente as direções associadas. Se o usuário preferir, é possível escolher duas direções e o programa determina automaticamente a terceira direção. Na Figura 5.4 encontra-se a interface que permite esse relacionamento.



Figura 5.4 – Interface que permite relacionar cena e objeto

Uma vez relacionados os pontos nas imagens da cena e nas imagens do objeto, o usuário pode optar por ver o objeto em ambiente 3D, conforme a Figura 5.5 ou reprojetoado nas imagens da cena, conforme a Figura 5.6. Na Figura 5.6 é mostrado um objeto inserido digitalmente na cena na posição escolhida pelo usuário.



(a) objeto representado em *wireframe*



(b) objeto texturizado

Figura 5.5 – Exemplo de visualização do objeto reconstruído no espaço 3D

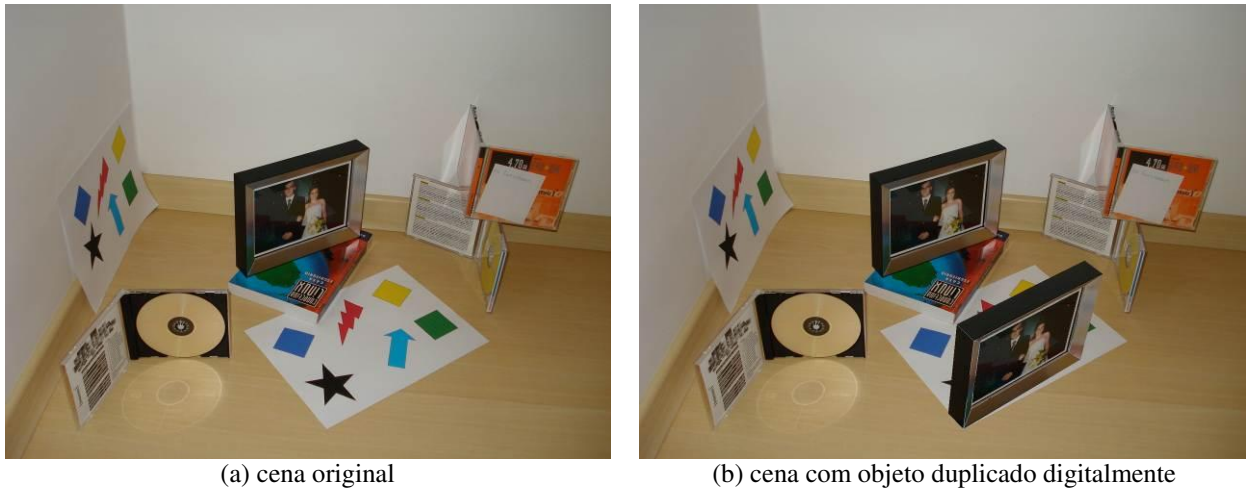


Figura 5.6 – Exemplo de objeto reprojetoado nas imagens da cena

No menu Editar | Alterar ponto principal da cena, o usuário pode alterar a posição do ponto principal da câmera, nos casos em que esses parâmetros sejam conhecidos. Nessa janela, o usuário pode ainda selecionar um dos casos implementados nesse protótipo:

- somente pontos principais constantes;
- somente pontos principais conhecidos;
- relação de aspecto unitária, pontos principais conhecidos e *skew* nulo.

6. AVALIAÇÃO E RESULTADOS

As técnicas para obtenção da reconstrução do objeto no espaço métrico foram aplicadas para imagens sintéticas na plataforma Matlab e para imagens reais no protótipo descrito no Capítulo 5.

A metodologia de avaliação dos resultados para imagens sintéticas baseou-se na medição do erro na reprojeção dos pontos do objeto reconstruído nas imagens originais e na comparação da reconstrução no espaço métrico com o objeto original. Para imagens reais, a avaliação dos resultados baseou-se na medição do erro na reprojeção dos pontos do objeto reconstruído nas imagens originais e na visualização do objeto reconstruído no ambiente tridimensional do protótipo.

6.1 Imagens sintéticas

Foram realizados diversos experimentos com imagens sintéticas, variando-se as restrições e o número de imagens em cada caso, entretanto o resultado de apenas um dos testes está reproduzido a seguir, por ter sido um teste que representa bem o conjunto dos testes realizados.

Como caso de teste representativo, foi escolhida a situação em que são conhecidos somente os pontos principais das imagens. Essa situação, descrita como Caso 1 no Capítulo 4, exige pelo menos 5 imagens. Foram criadas 5 câmeras virtuais com parâmetros apresentados na Tabela 6.1 e geradas as imagens apresentadas na Figura 6.1

Tabela 6.1 – Parâmetros das câmeras utilizadas no teste

Câmera A:	Câmera B:	Câmera C:	Câmera D:	Câmera E:
$T_x = 100$	$T_x = -50$	$T_x = 200$	$T_x = 200$	$T_x = 80$
$T_y = 0$	$T_y = 100$	$T_y = 10$	$T_y = 20$	$T_y = 50$
$T_z = 300$	$T_z = 0$	$T_z = 10$	$T_z = 0$	$T_z = 0$
$view\ up = [0\ 1\ 0]^T$	$view\ up = [0\ 1\ 0]^T$	$view\ up = [0\ 1\ 0]^T$	$view\ up = [0\ 1\ 1]^T$	$view\ up = [0\ 1\ -1]^T$
$look\ at = [0\ 0\ 80]^T$	$look\ at = [0\ 1\ 100]^T$	$look\ at = [10\ 0\ 70]^T$	$look\ at = [0\ 10\ 70]^T$	$look\ at = [0\ 1\ 90]^T$
$f = 1000$	$f = 450$	$f = 800$	$f = 600$	$f = 450$

Na Tabela 6.1, a posição da câmera é definida pela translação $T = [T_x\ T_y\ T_z]^T$ e pelos vetores de $view\ up$ (inclinação da câmera) e $look\ at$ (ponto para o qual a câmera está apontada). A distância focal é definida pelo parâmetro f .

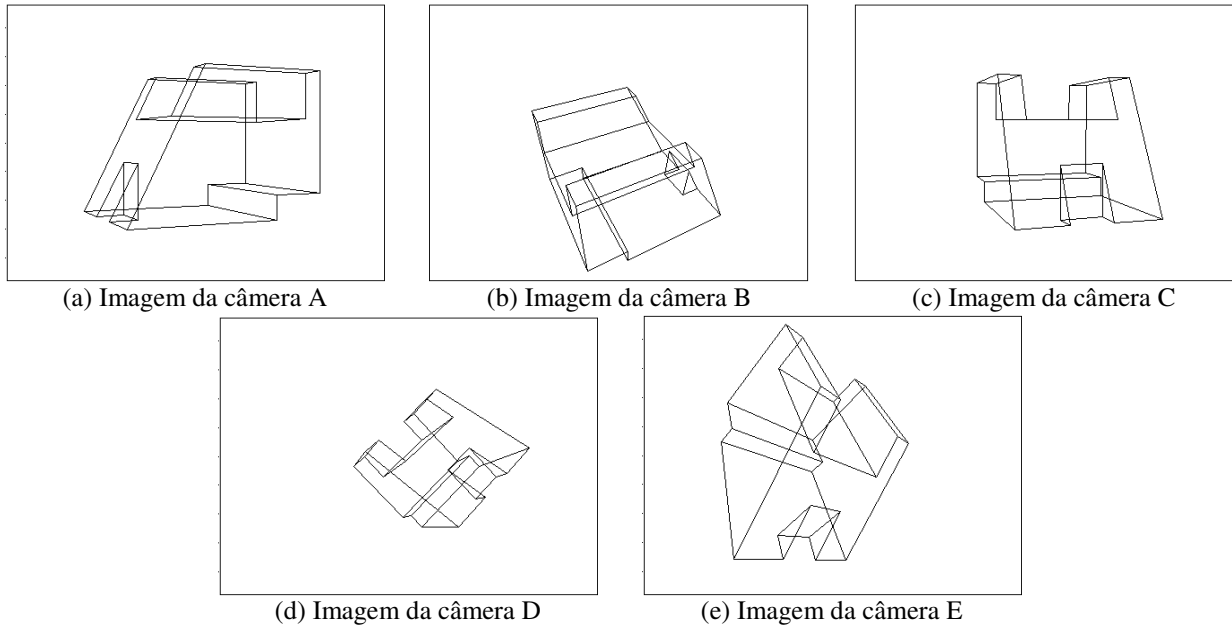


Figura 6.1 – Cinco imagens sintéticas utilizadas no teste

Os testes para imagens sintéticas foram conduzidos da mesma maneira que os testes realizados para a reconstrução no espaço projetivo, apresentados no Capítulo 3.

A reconstrução projetiva foi determinada a partir do cálculo da matriz fundamental entre as imagens A e B, com coordenadas das projeções normalizadas. A câmera A foi definida como referência, de modo que seu modelo apresenta a forma $\mathbf{M}_A = [\mathbf{I}_{3 \times 3} | \mathbf{0}_{3 \times 1}]$ e o modelo da câmera B foi calculada pela equação (3.09). Os pontos da reconstrução foram obtidos por triangulação e os modelos das câmeras C, D e E foram determinadas por calibração implícita, conforme método descrito no Capítulo 4.

A transformação que permite levar a reconstrução do espaço projetivo para o espaço métrico foi determinada conforme o Caso 1, descrito no Capítulo 4. A aplicação dessa transformação aos pontos do objeto reconstruído no espaço projetivo permitiu obter o objeto no espaço métrico.

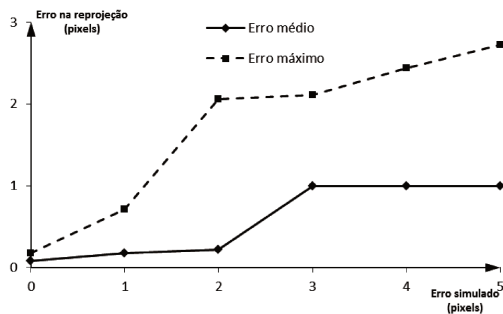
Para cada ponto do objeto reconstruído no espaço projetivo, foram calculadas suas reprojeções nas imagens originais e calculados os erros em relação à projeção original pela equação

$$d = \sqrt{(x_R - x)^2 + (y_R - y)^2} .$$

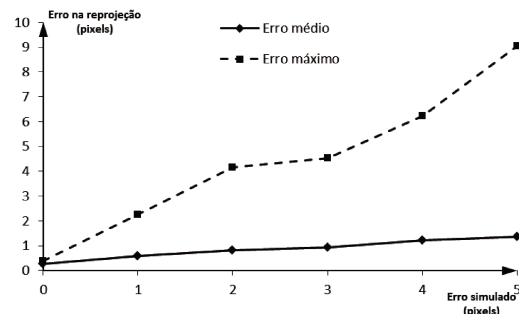
Os resultados da reprojeção a partir dos pontos recuperados no espaço métrico são apresentados na Tabela 6.2 e na forma gráfica na Figura 6.2.

Tabela 6.2 – Erro (em pixels) na reprojeção a partir do espaço métrico

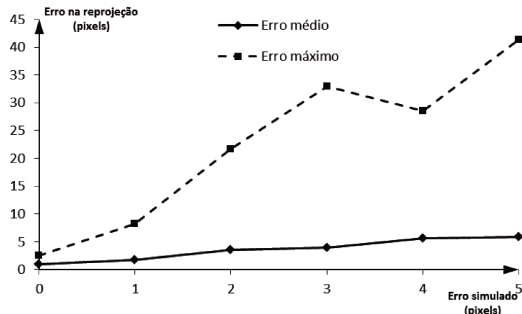
Erro simulado	Imagem A		Imagem B		Imagem C		Imagem D		Imagem E	
	Erro médio	Erro máximo	Erro médio	Erro máximo	Erro médio	Erro máximo	Erro médio	Erro máximo	Erro médio	Erro máximo
0	0	0	0	0	1	2	1	2	1	2
1	0	1	1	2	2	8	2	6	2	8
2	0	2	1	4	4	22	2	13	4	18
3	1	2	1	5	4	33	5	26	4	26
4	1	2	1	6	6	28	3	25	6	29
5	1	3	1	9	6	41	4	24	7	43



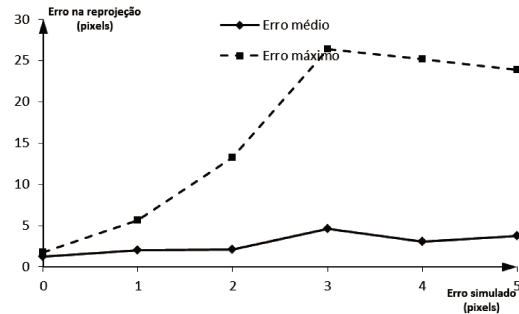
(a) Erro da reprojeção na imagem A



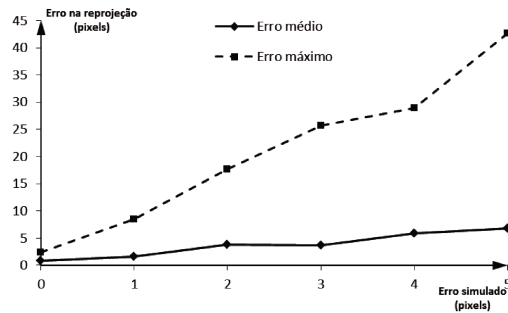
(b) Erro da reprojeção na imagem B



(c) Erro da reprojeção na imagem C



(d) Erro da reprojeção na imagem D



(e) Erro da reprojeção na imagem E

Figura 6.2 – Erro (em pixels) na reprojeção a partir do espaço métrico

A comparação entre as coordenadas dos pontos do objeto original e do objeto reconstruído no espaço métrico resultou nos dados apresentados na Tabela 6.3 e graficamente na Figura 6.3.

Tabela 6.3 – Erro na reconstrução

Erro simulado	Erro médio	Erro máximo
0	1,2	2,5
1	2,6	9,9
2	5,2	108,8
3	10,5	105,0
4	11,8	261,1
5	12,9	340,3

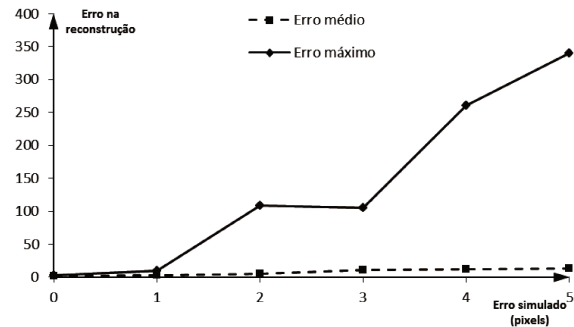
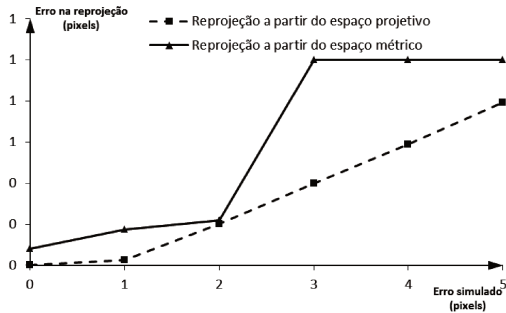
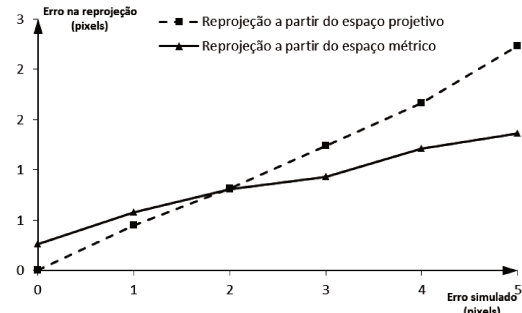


Figura 6.3 – Gráfico do erro na reconstrução

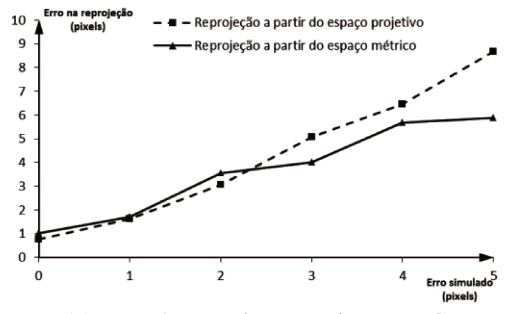
Os gráficos da Figura 6.4 apresentam o erro médio das reprojeções a partir do espaço projetivo e do espaço métrico.



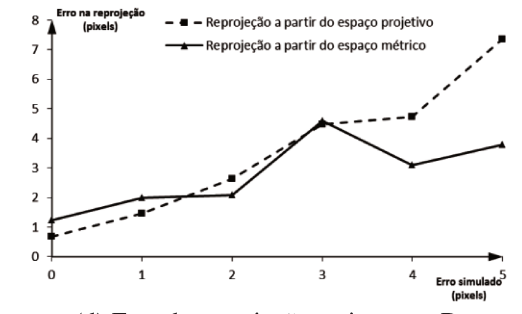
(a) Erro da reprojeção na imagem A



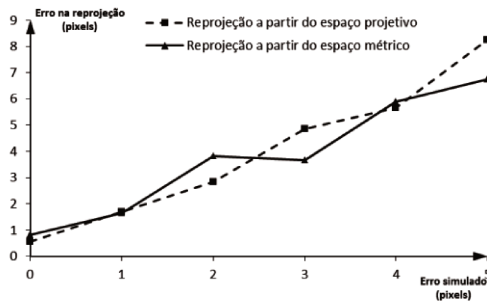
(b) Erro da reprojeção na imagem B



(c) Erro da reprojeção na imagem C



(d) Erro da reprojeção na imagem D



(e) Erro da reprojeção na imagem E

Figura 6.4 – Comparação entre o erro da reprojeção a partir dos espaços projetivo e métrico

A análise dos gráficos da Figura 6.4 permite verificar que o erro da reprojeção dos pontos do objeto reconstruído no espaço projetivo é da mesma ordem de grandeza do erro da reprojeção dos pontos do objeto reconstruído no espaço métrico. Entretanto, a instabilidade da reconstrução no espaço métrico aumenta com o erro de medição, levando a reconstruções por vezes absurdas.

6.2 *Imagens reais*

As imagens apresentadas na Figura 6.5 foram utilizadas como caso de teste demonstrativo do protótipo desenvolvido, com os parâmetros definidos da seguinte maneira:

- *skew* nulo;
- ponto principal no centro da imagem;
- relação de aspecto unitária;
- distâncias focais diferentes entre as imagens.



Figura 6.5 – Três imagens reais de uma cena

Foram estabelecidas 85 correspondências entre projeções nas imagens, que foram utilizados nos cálculos da matriz fundamental e da transformação que leva do espaço projetivo para o métrico.

Os erros na reprojeção das reconstruções projetiva e métrica estão apresentados na Tabela 6.4.

Tabela 6.4 – Erro (em pixels) na reprojeção com imagens reais

Reconstrução	Imagem A		Imagem B		Imagem C	
	Erro médio	Erro máximo	Erro médio	Erro máximo	Erro médio	Erro máximo
Projetiva	0	0	1	5	1	6
Métrica	0	0	1	5	1	6

Como referência, foram escolhidas algumas arestas do porta-retratos para verificar a qualidade da reconstrução, de acordo com a notação adotada para os pontos identificados na Figura 6.6.



Figura 6.6 – Identificação dos pontos utilizados como base para o cálculo do comprimento das arestas

Para avaliar a reconstrução foi calculada e aplicada uma escala global ao objeto reconstruído de modo que a aresta 11-12 tenha a mesma dimensão da aresta correspondente no objeto real, assim os comprimentos das arestas no objeto reconstruído podem ser medidos em milímetros.

A Tabela 6.5 apresenta as dimensões das arestas escolhidas do objeto real e do objeto reconstruído.

Na Tabela 6.5, a coluna “Erro” refere-se à diferença entre o comprimento da aresta do objeto real e do objeto reconstruído no espaço métrico. Por essa coluna, pode-se notar que a reconstrução foi correta, uma vez que os erros são pequenos, ou seja, o objeto reconstruído guarda as proporções do objeto real. Os erros são maiores para as arestas menores, uma vez que menores distâncias entre pontos projetados geram maiores erros na reconstrução.

Tabela 6.5 – Comparação dos comprimentos das arestas do objeto da Figura 6.6

Aresta	Objeto real (mm)	Objeto reconstruído (mm)	Erro (mm)	Erro (%)
11-12	222,0	222,0	-	-
9-10	222,0	223,0	1,0	0,45
49-85	222,0	222,7	0,7	0,31
9-12	172,0	175,2	3,2	1,86
10-11	172,0	173,6	1,6	0,93
9-49	35,0	31,4	-3,6	-10,29
10-85	35,0	33,0	-2,0	-5,71
15-16	182,0	184,2	2,2	1,21

A recuperação correta dos comprimentos das arestas não garante que o objeto tenha sido recuperado no espaço métrico. Para verificar se a reconstrução foi de fato levada ao espaço métrico, foram medidos os ângulos entre algumas arestas conforme apresentado na Tabela 6.6.

Tabela 6.6 – Ângulos (em graus) entre arestas do objeto real e do objeto reconstruído

Ângulo entre arestas	Objeto real	Objeto reconstruído	Erro
9-10 e 9-12	90°	90,2°	0,2°
49-9 e 49-85	90°	85,6°	-4,4°
11-10 e 11-12	90°	90,9°	0,9°
10-11 e 10-85	90°	91,6°	1,6°
49-9 e 86-12	0°	2,4°	2,4°

A análise do erro da Tabela 6.6 permite afirmar que o ângulo entre as arestas foi corretamente recuperado.

O resultado do objeto em ambiente 3D pode ser visto na Figura 6.7. Foram selecionadas as opções de exibir plano de referência (representados pelos pontos pretos) e os eixos do sistema de referência. O objeto representado nessa figura está em perspectiva e permite verificar o correto funcionamento da técnica, pois o objeto guarda a forma do objeto original.

Com o intuito de verificar a qualidade visual do objeto que foi recuperado, o porta-retratos reconstruído no espaço métrico a partir das imagens da Figura 6.5 foi reprojetoado em outra cena, representada pelas imagens da Figura 6.8.

A cena da Figura 6.8 foi reconstruída no espaço métrico utilizando 16 pontos dessa cena, pela mesma técnica utilizada para a reconstrução do porta-retratos.

Depois de reconstruída a cena no espaço métrico, os eixos de referência escolhidos para o objeto foram alinhados aos eixos de referência escolhidos para a cena, de modo a permitir o correto alinhamento entre eles.

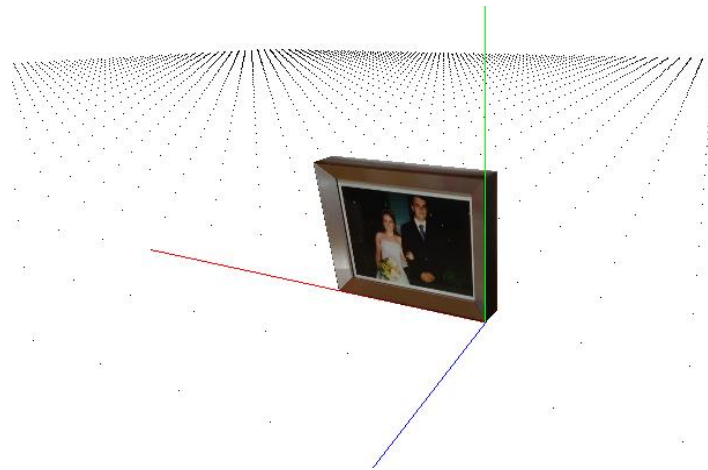


Figura 6.7 – Resultado da reconstrução visto em ambiente 3D

A escala global do objeto nesse experimento foi escolhida de modo que a aresta 11-12, conforme identificado na Figura 6.6, ocupasse a distância entre os extremos dos marcadores rosa e verde fixados na mesa.

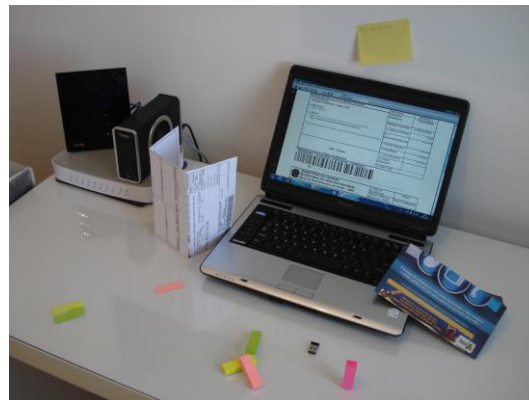
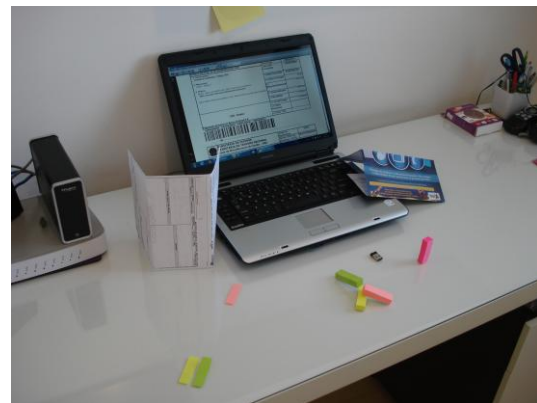
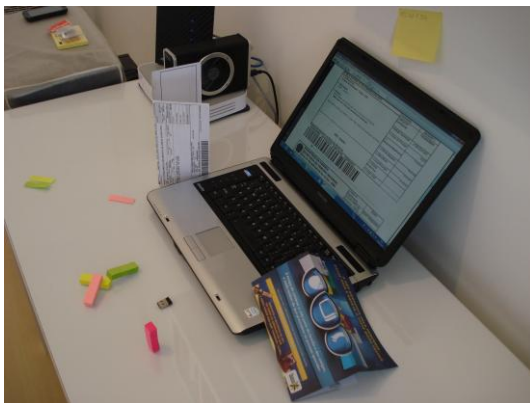


Figura 6.8 – Imagens originais da cena

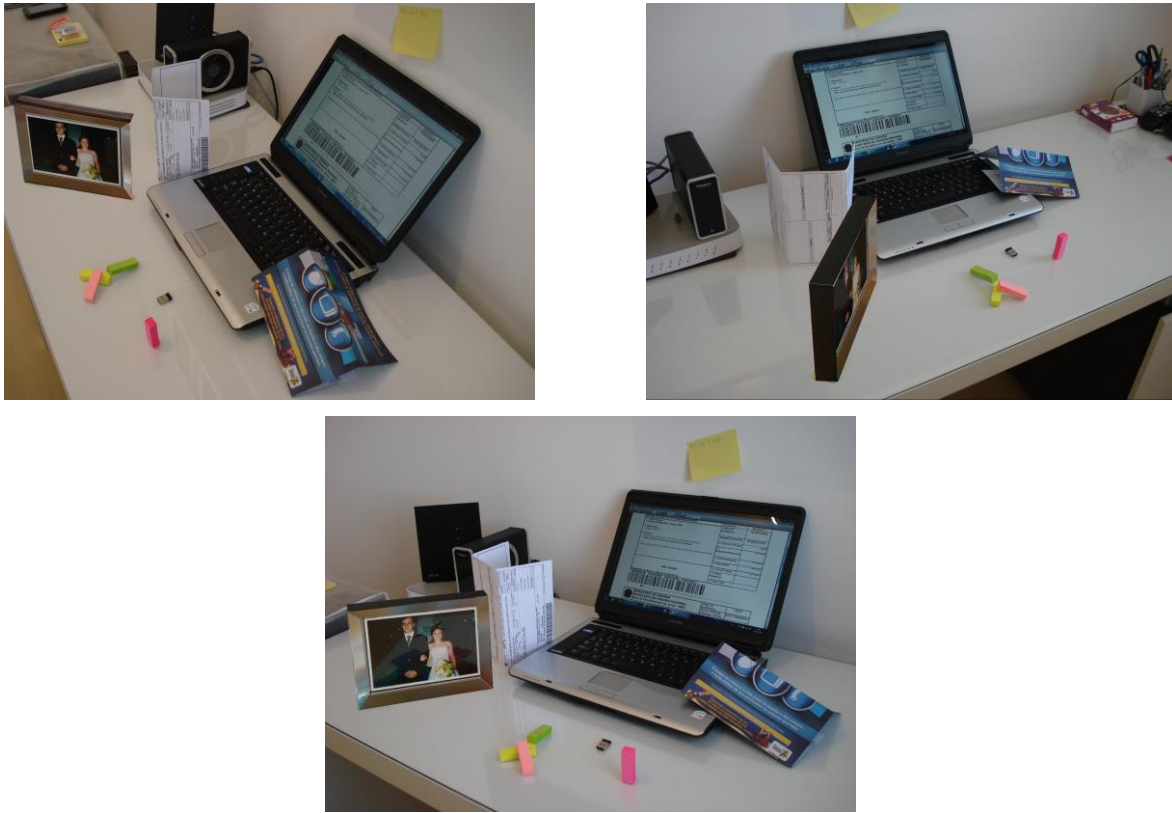


Figura 6.9 – Imagens da cena com objeto inserido

O resultado obtido é mostrado na Figura 6.9. A reconstrução da lateral direita do porta-retratos não foi possível, pois não há correspondência entre projeções nessa superfície, por esse motivo nota-se a ausência dessa superfície na imagem da câmera A. Nas demais câmeras, o aspecto do objeto reprojeto está correto.

Em todas as imagens, observa-se que as dimensões do objeto são condizentes com as dimensões dos demais objetos da cena e que o procedimento de inserção do objeto resultou em imagens de boa qualidade.

7. CONCLUSÃO

O conjunto de técnicas apresentadas mostrou-se suficiente para a recuperação da informação tridimensional considerando-se conhecidas somente imagens da cena ou do objeto. As soluções pela matriz fundamental e pelo tensor trifocal foram capazes de reconstruir um objeto no espaço projetivo e a transformação que leva os pontos do espaço projetivo para o métrico foi capaz de determinar o objeto no espaço métrico variando-se as restrições impostas aos parâmetros intrínsecos das câmeras.

O protótipo desenvolvido também mostrou-se eficaz na solução do problema proposto, inclusive com a recuperação das texturas do objeto a partir das imagens reais.

Os testes descritos no Capítulo 4 para reconstrução no espaço projetivo e os resultados obtidos para imagens reais apresentados no Capítulo 6 resultaram que são fontes de erro na recuperação da informação tridimensional:

- resolução espacial das imagens: nota-se que, quanto maiores forem as resoluções das imagens da cena ou do objeto, melhores são as reconstruções obtidas;
- pequena quantidade de pontos relacionados entre as imagens interfere diretamente na reconstrução projetiva. Essa é uma desvantagem da escolha do *matching* manual, realizada neste trabalho;
- erros de medição interferem diretamente na qualidade da reconstrução, sobretudo na reconstrução métrica. Há situações em que o erro na reprojeção (a partir da reconstrução projetiva) é aceitável, mas não é possível encontrar a transformação do espaço projetivo para o métrico;
- a solução obtida para sistemas lineares por minimização de erro pelo método da matriz pseudoinversa é mais suscetível a instabilidades numéricas que o método de decomposição em valores singulares.

Durante as simulações, observou-se que a reconstrução no espaço projetivo, por vezes, mostrou-se instável para alguns pontos quando se determinam as matrizes projetivas de câmeras pelo tensor trifocal, principalmente para erros de medida nas coordenadas dos pontos maiores que 3 pixels. Quando tal situação acontece, embora seja possível determinar a transformação do espaço projetivo para o métrico, a reconstrução no espaço métrico não é possível.

A reconstrução projetiva, tanto pela matriz fundamental quanto pelo tensor trifocal, mostrou-se robusta o suficiente, mesmo para poucos pontos projetados. O mesmo não ocorreu com a reconstrução métrica, que se mostrou sensível quando poucos pontos foram utilizados.

São sugestões para trabalhos futuros:

- implementação de *matching* semiautomático, em que o usuário estabelece o número mínimo de correspondência manualmente e o software determina mais correspondências, a fim de obter um mapa denso de profundidades;
- implementação dos casos em que os parâmetros de câmeras são constantes entre as imagens no protótipo;
- utilização do tensor trifocal, em conjunto com a matriz fundamental, para determinar a reconstrução no espaço projetivo. Isso permitirá o estabelecimento de correspondência de retas entre trios de imagens, melhorando a reconstrução no espaço projetivo;
- detecção de posições críticas entre as câmeras, que possam levar a uma solução não estável;
- busca de solução para o sistema não linear (4.36), que permite resolver o caso em que somente a restrição de *skew* nulo seja aplicada.

8. REFERÊNCIA BIBLIOGRÁFICA

AUTODESK. Autodesk® ImageModeler. Disponível em: <<http://usa.autodesk.com/adsk/servlet/pc/index?id=11390028&siteID=123112>>. Acesso em: 29 jul. 2011.

AUTODESK. Autodesk® ImageModeler™ 2009 Getting Started: 5-step Tutorial of a Complete Imagemodeler Project. 2009. Disponível em: <http://images.autodesk.com/adsk/files/imagemodeler_getting_started.pdf>. Acesso em: 29 jul. 2011.

BARROSO, L.; ARAÚJO, M.; FERREIRA FILHO, F.; CARVALHO, M.; MAIA M. **Cálculo Numérico (Com Aplicações)**. 2a. ed. São Paulo, Editora Arbra, 1987.

CIPOLLA R.; DRUMMOND, T.; ROBERTSON, D. Camera Calibration from Vanishing Points in Images of Architectural Scenes. **Proc. British Machine Vision Conference**, Cambridge, vol. 2, pp. 382 – 391, 1999.

CREATIVE DIMENSION SOFTWARE LTD. 3D Software Object Modeller Pro. 2011. Disponível em: <<http://www.3dsom.com>>. Acesso em: 29 jun. 2011.

DEVEBEC, P.; TAYLOR C.; MALIK, J. Modeling and Rendering Architecture from Photographs: A Hybrid Geometry- and Image-Based Approach. **Proc. of SIGGRAPH 96**, pp. 11 – 20. 1996.

DUCKE, B.; SCORE, D.; REEVES, J. Multiview 3D Reconstruction of the Archaeological Site at Weymouth From Image Series. **Computers & Graphics**, Volume 35, Issue 2, pp. 375 – 382, 2011.

EOS SYSTEMS INC. Accurate and Affordable 3D Photogrammetry Measuring and Scanning. 2011. Disponível em: <<http://www.photomodeler.com>>. Acesso em: 29 jun. 2011.

FORSYTH, D.; PONCE, J. **Computer Vision: A Modern Approach**. Prentice Hall, 2002.

GIBSON, S.; HOWARD, T. Interactive Reconstruction of Virtual Environments from Photographs, with Application to Scene-of-Crime Analysis. **Proceedings of the ACM Symposium on Virtual Reality Software and Technology**. New York, pp. 41 – 48, 2000.

GONZALEZ, R; WOODS, R. **Processamento de Imagens Digitais**. São Paulo, Edgard Blucher Ltda, 2000.

HARTLEY, R. In Defense of the Eight-Point Algorithm. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Volume 19, Issue 6. 1997.

HARTLEY, R. **Kruppa's Equation Derived from the Fundamental Matrix**. IEEE Transactions on

Pattern Analysis and Machine Intelligence. Volume 19, Issue 2, pp. 133 – 135, 1999.

HARTLEY, R. Lines and Point in Three Views - An Integrated Approach. **Proceedings of Arpa Image Understanding Workshop Monterey**. California, pp. 1009 – 1016, 1994.

HARTLEY, R.; VIDAL, R. The Multibody Trifocal Tensor: Motion Segmentation from 3 Perspective Views. **Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition**. Volume I, Issue C, pp. 769 – 775, 2004.

HARTLEY, R.; ZISSERMAN, A. **Multiple View Geometry in Computer Vision**. Cambridge University Press, 2000.

HEMMLEB, M. Digital Rectification of Historical Images. IAPRS, Amsterdam, Vol. XXXII, pp. 2 – 7, 1999.

HEYDEN, A.; ASTRÖM, K. **Euclidean Reconstruction from Constant Intrinsic Parameters**. IEEE Proceedings of the 13th International Conference on Pattern Recognition. Volume 1, pp. 339 – 343, Viena, 1996.

IZADI, S.; NEWCOMBE, R.; KIM, D.; HILLIGES, O.; MOLYNEAUX, D.; KOHLI, P.; SHOTTON, J.; HODGES, S.; FREEMAN, D.; DAVISON, A.; FITZGIBBON, A. KinectFusion: Real-Time Dynamic 3D Surface Reconstruction and Interaction. SIGGRAPH Talks 2011

KOCH, R.; POLLEFEYS, M.; GOOL, L. V. **Realistic Surface Reconstruction of 3D Scenes from Uncalibrated Image Sequences**. The Journal of Visualization and Computer Animation, Volume 11, Issue 3, pp. 115 – 127, 2000.

KOLMAN, B. **Introdução à Álgebra Linear com Aplicações**. 8^a edição, LTC, São Paulo, 2006.

LUONG, Q.; FAUGERAS, O. The Fundamental matrix: theory, algorithms, and stability analysis. **International Journal of Computer Vision**. Volume 17, pp. 43 – 75, 1995.

MACH, L. Insight3d – Opensource Image Based 3d Modeling Software. 2010. Disponível em: <<http://insight3d.sourceforge.net/>>. Acesso em: 03 out. 2011.

MAYBANK, S.; FAUGERAS, O. A theory of self-calibration of a moving camera, **International Journal of Computer Vision**. Volume 8, Issue 2, pp. 123 – 151, 1992.

NOGUEIRA, F. **Reconstrução Tridimensional por Visão Estéreo e Detecção de Oclusões**, 2007. 139 p. Tese de Doutorado em Engenharia Engenharia Elétrica e de Computação, Faculdade de Engenharia Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, 2007.

POLLEFEYS, M.; KOCH, R.; VAN GOOL, L. **Self-Calibration and Metric Reconstruction in spite of**

Varying and Unknown Intrinsic Camera Parameters. International Journal of Computer Vision, Volume 32, Issue 1, pp. 7 – 25, 1999.

POLLEFEYS, M.; KOCH, R.; VERGAUWEN, M.; VAN GOOL, L. Automated reconstruction of 3D scenes from sequences of images. **ISPRS Journal of Photogrammetry and Remote Sensing.** Volume 55, no. 4 , pp. 251 – 267, 2000.

POLLEFEYS, M.; VAN GOOL, L. A Stratified Approach to Metric Self-Calibration. **CVPR '97 Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition.** Puerto Rico, pp. 407 – 412, 1997.

SHASHUA, A; WERMAN, M. Trilinearity of three perspective views and its associated tensor. **Proc. Fifth International Conference on Computer Vision.** Cambridge, pp. 920 – 925, 1995.

TOMASI, C.; KANADE, T. **Shape and Motion from Image Streams Under Orthography: a Factorization Method.** International Journal of Computer Vision, Volume 9, Issue 2, pp. 137 – 154, 1992.

TOPAY, U. **3D Scene Reconstruction from Uncalibrated Images.** 2002. 127 p. Dissertação (Mestrado em Engenharia Elétrica e Eletrônica) - The Graduate School of Natural and Applied Sciences of The Middle East Technical University, Çankaya Ankara, 2002.

TRUCCO, E.; VERRI, A. **Introductory Techniques for 3D Computer Vision,** Prentice-Hall, Englewood Cliffs, NJ, 1998.

TSAI, R. A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf Cameras and Lenses. **IEEE Journal of Robotics and Automation.** Volume 3, Issue 4, pp. 323 – 344, 1987.

VEBLEN, O.; YOUNG, J. W. **Projective Geometry Volume 1.** Ginn and Co, Boston, 1910.

WIEDEMANN, A.; HEMMLEB, M.; ALBERTZ J. Reconstruction Of Historical Buildings Based On Images From The Meydenbauer Archives. **International Archives of Photogrammetry and Remote Sensing,** Amsterdam, pp. 887 – 893, 2000.

YOON, Y.; IM, J.; KIM, D.; CHOI, J. 3D Reconstruction Using Three Vanishing Points from a Single Image. **The 2002 International Conference on Circuit Systems, Computers and Communication.** Phuket, 2002.

ZHANG Z., KANADE, T. Determining the Epipolar Geometry and its Uncertainty: A Review. **International Journal of Computer Vision.** Volume 27, Issue 2, pp. 191 – 195, 1998.