

Relatório de uma Aula Prática de Visão por Computador sobre Geometria Epipolar

*Departamento de Engenharia Electrotécnica
Faculdade de Ciências e Tecnologia
Universidade de Coimbra*

João Pedro de Almeida Barreto

11 de Junho de 2002

Conteúdo

1	Introdução	2
2	Fundamentos Teóricos	4
2.1	Geometria Epipolar	4
2.1.1	Notação e Nomenclatura	4
2.1.2	Relação entre a Matriz Essencial e Fundamental	6
2.1.3	Dedução da Matriz Fundamental a Partir do Epipolo	7
2.2	Técnicas de Estimação Numérica	9
2.2.1	Estimação dos Mínimos Quadráticos Normais	9
2.2.2	Estimação dos Mínimos Quadráticos Totais	9
2.3	Estimação da Matriz Fundamental	10
3	Trabalho Prático	12
3.1	Carregamento de Dados	12
3.2	A Função de Verificação da Geometria Epipolar	14
3.3	Trabalho a realizar	15
3.3.1	Estimação da matriz fundamental	15
3.3.2	Verificação dos resultados de estimação da matriz fundamental	15
3.3.3	Informação extraída da Matriz Fundamental	16
4	Resolução do Trabalho Prático	17
4.1	Estimação da matriz Fundamental	17
4.2	Extracção de Informação da Matriz Fundamental	18

Capítulo 1

Introdução

Este documento descreve uma aula prática de 2 horas sobre geometria epipolar. A geometria epipolar é ensinada na cadeira de Visão por Computador aos alunos da licenciatura de Engenharia Electrotécnica da Faculdade de Ciências e Tecnologia da Universidade de Coimbra.

O presente relatório é composto por três secções. A primeira secção apresenta os conhecimentos teóricos necessários à realização do trabalho prático. A segunda secção enuncia o trabalho prático propriamente dito. A terceira secção resolve o trabalho prático e discute os objectivos pedagógico dos problemas propostos. Existe ainda uma apêndice onde constam as transparências de apoio à aula e as listagens de código em *MATLABTM* das rotinas fornecidas aos alunos e daquelas que se pretende que eles programem.

Este trabalho prático exige conhecimentos de geometria epipolar e de métodos de estimação numérica linear. Ambos os assuntos são estudados nas aulas teóricas da cadeira de Visão por Computador. Para além do enunciado do trabalho é fornecido um resumo de apoio sobre a teoria necessária à sua realização. Os alunos são aconselhados a estudar este documento de forma a prepararem previamente a lição. A primeira meia-hora da aula destina-se à revisão destes conhecimentos teóricos e ao esclarecimento de dúvidas. Esta introdução é feita com a ajuda das transparências em apêndice.

A aula prática decorre no laboratório de Visão por Computador onde existem 7 postos de trabalho. O número ideal de alunos por turma é de 14 de forma a cada grupo de trabalho ser composto por dois elementos. Em cada posto de trabalho existe, para além de outros equipamentos, um PC com o *MATLABTM* da "MathWorks, Inc." instalado. O presente trabalho prático utiliza quatro imagens calibradas disponibilizadas pelo "Calibrated Imaging Laboratory" da "Carnegie Mellon University" em Pittsburgh, Pensilvânia, Estados Unidos da América (consultar a página "[2](http://www.cs.cmu.edu/cil/cil-</p></div><div data-bbox=)

ster.html”). É ainda disponibilizada a informação de calibração (parâmetros intrínsecos e extrínsecos) e as coordenadas de 28 pontos na imagem com um erro inferior a 0.1 pixels. Esta última informação estabelece a correspondência entre pontos da imagem permitindo a computação das matrizes fundamentais. O primeiro objectivo do trabalho é desenvolver uma rotina em *MATLABTM* que, dadas as coordenadas de pontos correspondentes entre duas imagens, calcule a matriz fundamental. Os resultados obtidos são verificados utilizando as imagens de teste. Existem depois um conjunto de perguntas que visam testar a compreensão da informação contida na matriz fundamental, nomeadamente quanto ao movimento entre as imagens e a posição dos epipólos.

Ao contrário do resumo teórico e do enunciado do trabalho, o material apresentado na terceira secção não se destina aos alunos. Nela consta a resolução do trabalho prático e os objectivos pedagógicos que se pretenderam atingir.

Capítulo 2

Fundamentos Teóricos

Considere um par de imagens de uma mesma cena estacionária em que os respectivos centros de projecção são não coincidentes. Estas imagens podem ser obtidas por duas câmaras ou por uma câmara com movimento de translação. A visão estéreo consiste em ter duas perspectivas diferentes de uma mesma cena estática. A geometria epipolar estabelece as relações geométricas entre as duas imagens descrevendo matematicamente muitos aspectos da visão estéreo. Um dos primeiros resultados da geometria epipolar foi o estabelecimento da matriz essencial por Longuet-Higgins em [1].

Recentemente a aplicação de ferramentas da geometria projectiva clássica [2] à visão por computador permitiu um conjunto de novos desenvolvimentos na área [3, 4, 5]. Uma dessas contribuições foi o estabelecimento da matriz fundamental [6]. Esta matriz estabelece uma relação projectiva entre as duas imagens contendo toda a informação que é possível obter com um par de câmaras completamente não-calibradas. O estudo da matriz fundamental é de grande importância para a visão por computador. Este conceito permite, não só desenvolver técnicas de auto-calibração, como também dispensar a calibração na realização de certas tarefas.

2.1 Geometria Epipolar

2.1.1 Notação e Nomenclatura

$$M_2 = RM_1 + t \quad (2.1)$$

A Fig. 2.1 mostra um par de câmaras estéreo que permitem obter duas imagens em perspectiva de uma mesma cena estática. O ponto do espaço assinalado é projectado em ambas as imagens. \mathbf{M} , \mathbf{C}_1 e \mathbf{C}_2 são, respectivamente, as coordenadas cartesianas do ponto e dos centros ópticos das

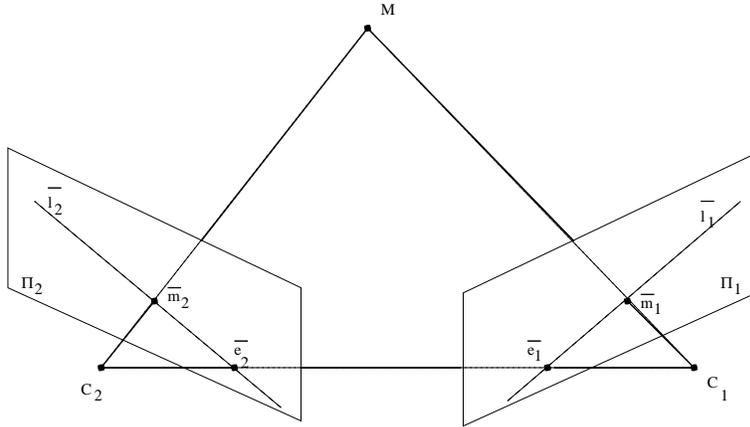


Figura 2.1: Duas vistas em perspectiva de um mesmo ponto no espaço 3D.

câmaras 1 e 2 em relação a um sistema de referência no mundo. Assuma ainda um sistema de eixos cartesianos centrado no centro óptico de cada uma das câmaras. \mathbf{M}_1 e \mathbf{M}_2 representam as coordenadas 3D do ponto no referencial de cada câmara. \mathbf{R} e \mathbf{t} são, respectivamente, a matriz de rotação e o vector de translação entre os referenciais das duas câmaras de forma a que a equação 2.1 se verifique.

$$\tilde{\mathbf{m}}_i = \mathbf{A}_i \mathbf{M}_i \quad (2.2)$$

Considere que as câmaras estão calibradas. O ponto \mathbf{M} é projectado em \mathbf{m}_1 na imagem 1 e em \mathbf{m}_2 na imagem 2. Atendendo a que $\mathbf{M}_1 = (X_1, Y_1, Z_1)^t$ e $\mathbf{M}_2 = (X_2, Y_2, Z_2)^t$, e que \mathbf{m}_1 e \mathbf{m}_2 são as coordenadas homogéneas normalizadas da posição de projecção no plano imagem, então $\mathbf{m}_1 = (\frac{X_1}{Z_1}, \frac{Y_1}{Z_1}, 1)^t$ e $\mathbf{m}_2 = (\frac{X_2}{Z_2}, \frac{Y_2}{Z_2}, 1)^t$. Suponha agora que as câmaras são não calibradas e que \mathbf{A}_1 e \mathbf{A}_2 representam as matrizes de parâmetros intrínsecos de cada uma delas (matrizes 3×3). A equação 2.2 transforma as coordenadas homogéneas normalizadas \mathbf{m}_1 e \mathbf{m}_2 , em coordenadas em pixels na imagem $\tilde{\mathbf{m}}_1$ e $\tilde{\mathbf{m}}_2$ ($i = 1, 2$).

Até ao momento vimos que um ponto \mathbf{M} da cena se projecta nos pontos $\tilde{\mathbf{m}}_1$ e $\tilde{\mathbf{m}}_2$ nas imagens obtidas pelas câmaras com centro óptico em \mathbf{C}_1 e \mathbf{C}_2 . O plano definido pelo ponto e pelos dois centros ópticos $[\mathbf{M}, \mathbf{C}_1, \mathbf{C}_2]$ é designado por plano epipolar. O plano epipolar intersecta cada um dos planos imagem numa linha, também chamada epipolar, que contém a projecção do ponto \mathbf{M} . As linhas epipolares conjugadas são referenciadas na Fig. 2.1 por \bar{l}_1 e \bar{l}_2 e passam respectivamente pelas imagens $\tilde{\mathbf{m}}_1$ e $\tilde{\mathbf{m}}_2$ e os pontos $\bar{\mathbf{e}}_1$ e $\bar{\mathbf{e}}_2$. Estes dois últimos pontos são os epipólos e resultam da intersecção da

linha de base $[\mathbf{C}_1, \mathbf{C}_2]$ com cada um dos planos imagem. Se a posição de \mathbf{M} variar no espaço, os planos epipolares formam um feixe de planos a passar por $[\mathbf{C}_1, \mathbf{C}_2]$, e as linhas epipolares formam dois feixes de linhas através de $\bar{\mathbf{e}}_1$ e $\bar{\mathbf{e}}_2$.

2.1.2 Relação entre a Matriz Essencial e Fundamental

$$\Pi = \mathbf{t} \wedge \mathbf{R}\mathbf{m}_1 \quad (2.3)$$

Assuma que as câmaras estão calibradas e que o ponto \mathbf{M} se projecta em \mathbf{m}_1 na imagem 1 e \mathbf{m}_2 na imagem 2. \mathbf{m}_1 e \mathbf{m}_2 são coordenadas normalizadas expressas no referencial da câmara respectiva. O vector que une o centro óptico da primeira câmara com o ponto de projecção na imagem \mathbf{m}_1 e o vector que une os dois centros ópticos são coplanares (ver Fig. 2.1). As coordenadas cartesianas do plano epipolar Π no sistema de referência da câmara 2 são dados pela equação 2.3. A multiplicação por \mathbf{R} é necessária para transformar \mathbf{m}_1 no sistema de coordenadas da segunda câmara.

$$\mathbf{m}_2^t (\mathbf{t} \wedge \mathbf{R}\mathbf{m}_1) = 0 \quad (2.4)$$

Atendendo a que \mathbf{m}_2 pertence ao plano epipolar Π obtém-se o resultado da equação 2.4.

$$\tilde{\mathbf{t}} = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \quad (2.5)$$

$$\mathbf{E} = \tilde{\mathbf{t}}\mathbf{R} \quad (2.6)$$

$$\mathbf{m}_2^t \mathbf{E} \mathbf{m}_1 = 0 \quad (2.7)$$

Sabe da álgebra que o produto externo entre dois vectores \mathbf{v} e \mathbf{u} é igual ao produto da matriz anti-simétrica do primeiro pelo segundo ($\mathbf{v} \wedge \mathbf{u} = \tilde{\mathbf{v}}\mathbf{u}$ em que $\tilde{\mathbf{v}}$ é a matriz anti-simétrica de \mathbf{v}). A equação 2.5 mostra a matriz anti-simétrica do vector de translação $\mathbf{t} = (t_x, t_y, t_z)^t$. Definindo a matriz \mathbf{E} de acordo com 2.6 pode reescrever a equação 2.4 na forma da equação 2.7. A matriz \mathbf{E} é chamada a matriz essencial.

Note que a equação 2.4 é homogénea em relação a \mathbf{t} . Diferentes módulos do vector de translação originam a mesma matriz essencial. Isto reflecte o facto da escala ser indeterminada, não sendo possível recuperar a escala absoluta da cena sem informação adicional. Assim \mathbf{E} é uma matriz 3×3 singular (característica 2) que só depende de 5 parâmetros (3 de rotação e 2 de

translação). Repare também que no caso de não haver translação então $\mathbf{t} = 0$ e $\mathbf{E} = 0$. Neste caso é impossível calcular informação sobre profundidade no espaço e as duas imagens têm entre si uma relação homográfica.

$$\begin{cases} \mathbf{m}_1 = \mathbf{A}_1^{-1} \bar{\mathbf{m}}_1 \\ \mathbf{m}_2 = \mathbf{A}_2^{-1} \bar{\mathbf{m}}_2 \end{cases} \quad (2.8)$$

Considere agora a situação não calibrada em que as coordenadas em pixels na imagem $\bar{\mathbf{m}}_1$ e $\bar{\mathbf{m}}_2$ relacionam-se com as coordenadas normalizadas \mathbf{m}_1 e \mathbf{m}_2 de acordo com a equação 2.2. A equação 2.8 é deduzida a partir de 2.2.

$$\bar{\mathbf{m}}_2^t (\mathbf{A}_2^{-1})^t \mathbf{E} \mathbf{A}_1^{-1} \bar{\mathbf{m}}_1 = 0 \quad (2.9)$$

Substituindo o resultado de 2.8 em 2.7 obtém-se a equação 2.9.

$$\mathbf{F} \simeq (\mathbf{A}_2^{-1})^t \mathbf{E} \mathbf{A}_1^{-1} \quad (2.10)$$

$$\bar{\mathbf{m}}_2^t \mathbf{F} \bar{\mathbf{m}}_1 = 0 \quad (2.11)$$

Definindo a matriz \mathbf{F} de acordo com 2.10 e substituindo em 2.9 chega-se ao resultado da equação 2.11. A matriz \mathbf{F} designa-se por matriz fundamental. Note que a equação 2.10 define a matriz fundamental a menos de um factor de escala. \mathbf{F} é uma matriz 3×3 , no entanto atendendo ao factor de escala e a que $\det(\mathbf{F}) = 0$, tem só 7 graus de liberdade independentes.

Tanto a matriz essencial com a fundamental descrevem completamente a relação geométrica entre pontos correspondentes num par de câmaras estéreo. A única diferença entre as duas é que a primeira refere-se a câmaras calibradas, enquanto a segunda se aplica à situação não-calibrada. Existem outras formas de deduzir a matriz fundamental [5, 7, 3]. A secção seguinte mostra reproduz a dedução feita em [8, 4].

2.1.3 Dedução da Matriz Fundamental a Partir do Epipolo

Observe a Fig. 2.1. A variação da posição de \mathbf{M} no espaço define um feixe de planos epipolares que definem por sua vez um feixe de linhas epipolares em cada uma das imagens. Note que todas as linhas epipolares na imagem 1 passam pelo respectivo epipólo $\bar{\mathbf{e}}_1$, e que todas as linhas epipolares na imagem 2 se intersectam em $\bar{\mathbf{e}}_2$. A correspondência entre as linhas epipolares das duas imagens é feita por uma colineação projectiva \mathbf{C} definida no espaço dual das linhas.

$$\bar{\mathbf{l}}_1 = \bar{\mathbf{e}}_1 \wedge \bar{\mathbf{m}}_1 = \tilde{\mathbf{e}}_1 \bar{\mathbf{m}}_1 \quad (2.12)$$

Considere $\bar{\mathbf{m}}_1$ como sendo o vector de coordenadas homogéneas do ponto projectado na primeira imagem. A linha epipolar respectiva $\bar{\mathbf{l}}_1$ passa pelo epipólo $\bar{\mathbf{e}}_1$. A equação 2.12 calcula as coordenadas projectivas da linha epipolar. A correspondência entre um ponto na imagem e a respectiva linha epipolar é linear e pode ser representada pela matriz 3×3 $\tilde{\mathbf{e}}$ que tem característica 2 e 2 graus de liberdade (as coordenadas do epipólo).

$$\bar{\mathbf{l}}_2 = \mathbf{C}\bar{\mathbf{l}}_1 = \mathbf{C}\tilde{\mathbf{e}}_1\bar{\mathbf{m}}_1 \quad (2.13)$$

Como já foi referido a correspondência entre as linhas epipolares das duas imagens é uma transformação projectiva \mathbf{C} . A equação 2.13 calcula a linha epipolar $\bar{\mathbf{l}}_2$ que contém a imagem do ponto $\bar{\mathbf{m}}_2$ na segunda câmara. Note que a correspondência entre linhas epipolares não depende da translação entre os centros ópticos. \mathbf{C} é uma matriz com 5 graus de liberdade independentes que é função das matrizes de parâmetros intrínsecos e da rotação relativa das câmaras.

$$\bar{\mathbf{m}}_2^t(\mathbf{C})\tilde{\mathbf{e}}_1\bar{\mathbf{m}}_1 = 0 \quad (2.14)$$

$$\mathbf{F} \simeq \mathbf{C}\tilde{\mathbf{e}}_1 \quad (2.15)$$

Atendendo agora que $\bar{\mathbf{m}}_2$ pertence à linha epipolar $\bar{\mathbf{l}}_2$ vem que $\bar{\mathbf{m}}_2^t\bar{\mathbf{l}}_2 = 0$. Fazendo a substituição de acordo com 2.13 obtém-se o resultado da equação 2.14 de onde se conclui que a matriz fundamental \mathbf{F} é dada por 2.15.

$$\mathbf{F} \simeq (\mathbf{H}^{-1})^t\tilde{\mathbf{e}}_1 \quad (2.16)$$

Assuma que \mathbf{H} é uma transformação projectiva definida no espaço \wp^2 que mapeia um ponto genérico $\bar{\mathbf{m}}$ em $\bar{\mathbf{m}}' = \mathbf{H}\bar{\mathbf{m}}$ preservando a colineariedade. A transformação correspondente no espaço dual das linhas transforma uma linha genérica $\bar{\mathbf{l}}$ em $\bar{\mathbf{l}}'$ de forma a que se $\bar{\mathbf{m}}$ pertence a $\bar{\mathbf{l}}$ então $\bar{\mathbf{l}}'$ passa por $\bar{\mathbf{m}}'$. Assim vem que $\bar{\mathbf{m}}^t\bar{\mathbf{l}} = 0 = (\mathbf{H}^{-1}\bar{\mathbf{m}}')^t\bar{\mathbf{l}} = (\bar{\mathbf{m}}')^t((\mathbf{H}^{-1})^t\bar{\mathbf{l}})$. Daqui conclui que a colineação no espaço dual das linhas é dada por $(\mathbf{H}^{-1})^t$. Considere agora a situação da Fig. 2.1 mas assumo que não existe translação entre as câmaras ($\mathbf{t} = 0$). Existe um homografia \mathbf{H} que mapeia os pontos da imagem 1 na imagem 2. Usando os resultados das equações 2.1 2.2 e 2.8 vem que $\bar{\mathbf{m}}_2 = \mathbf{H}\bar{\mathbf{m}}_1 = \mathbf{A}_2\mathbf{R}\mathbf{A}_1^{-1}\bar{\mathbf{m}}_1$. A matriz \mathbf{C} que mapeia as linhas epipolares é dada por $\mathbf{C} = \mathbf{H}^{-1^t}$. A equação 2.15 pode ser reescrita na forma da equação 2.16.

2.2 Técnicas de Estimação Numérica

A informação extraída das imagens vem geralmente contaminada por ruído. Por isso a aplicação da geometria projectiva à visão por computador exige muitas vezes conhecimentos de estimação e minimização matemática [9, 10]. Esta secção faz uma revisão abreviada de alguns desse conceitos.

2.2.1 Estimação dos Mínimos Quadráticos Normais

$$y^i = f(x_1^i, x_2^i, \dots, x_n^i) = a_1 x_1^i + a_2 x_2^i + \dots + a_n x_n^i, \forall_i \quad (2.17)$$

$$\delta^i = \tilde{y}^i - y^i = \tilde{y}^i - a_1 x_1^i + a_2 x_2^i + \dots + a_n x_n^i \quad (2.18)$$

Considere que tem um número m de pontos $(x_1^i, x_2^i, \dots, x_n^i, \tilde{y}^i)$ em que $i = 1, \dots, m$ e que y^i é uma função linear de $(x_1^i, x_2^i, \dots, x_n^i)$ (ver equação 2.17). Imagine ainda que os m pontos foram obtidos por observação/medição e que \tilde{y}^i vem afectado de ruído δ^i . O erro δ^i em cada ponto é dado matematicamente pela equação 2.18. O seu objectivo é determinar os parâmetros (a_1, a_2, \dots, a_n) de forma a minimizar o erro quadrático $\sum_{i=1}^m (\delta^i)^2$. Isto pode ser feito usando o método de estimação dos mínimos quadráticos.

$$\mathbf{X} = \begin{bmatrix} x_1^1 & x_2^1 & \dots & x_n^1 \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & \vdots & \vdots \\ x_1^m & x_2^m & \dots & x_n^m \end{bmatrix}, \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \tilde{\mathbf{y}} = \begin{bmatrix} \tilde{y}^1 \\ \tilde{y}^2 \\ \vdots \\ \tilde{y}^m \end{bmatrix}, \boldsymbol{\delta} = \begin{bmatrix} \delta^1 \\ \delta^2 \\ \vdots \\ \delta^m \end{bmatrix} \quad (2.19)$$

$$\mathbf{X}\mathbf{a} = \tilde{\mathbf{y}} - \boldsymbol{\delta} \quad (2.20)$$

$$\mathbf{a} = (\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t \tilde{\mathbf{y}} \quad (2.21)$$

O problema de minimização proposto pode ser interpretado como a determinação do vector \mathbf{a} que verifica a equação 2.20 de forma a que a norma $|\boldsymbol{\delta}| = \boldsymbol{\delta}^t \boldsymbol{\delta}$ seja mínima. A solução do problema é dada pela equação 2.21 [9].

2.2.2 Estimação dos Mínimos Quadráticos Totais

$$\mathbf{X} = \mathbf{U} \cdot \mathbf{S} \cdot \mathbf{V}^t \quad (2.22)$$

Considere uma matriz \mathbf{X} $m \times n$. É possível decompor qualquer matriz na forma da equação 2.22 em que \mathbf{U} é uma matriz ortonormal $m \times m$, \mathbf{S} é uma matriz diagonal $n \times n$ semi-definida positiva e \mathbf{V} é uma matriz ortonormal $n \times n$. Esta decomposição é conhecida por decomposição em valores singulares. Os valores diagonais de \mathbf{S} são chamados de valores singulares e o seu quadrado

é igual aos valores próprios de $\mathbf{X}^t\mathbf{X}$. No caso da coluna j de \mathbf{S} ser nula (valor singular nulo), a coluna j de \mathbf{V} faz parte da base do espaço nulo de \mathbf{X} . Obviamente que no caso de \mathbf{S} ser uma matriz definida positiva, não existem valores singulares nulos, e a matriz \mathbf{X} é não-singular. A norma de Frobenius de \mathbf{X} é igual à raiz quadrada do somatório dos quadrados dos seus valores singulares. Este resultados algébricos são úteis para compreender a estimação dos mínimos quadráticos totais que a seguir se explica.

$$(\mathbf{X} - \Delta)\mathbf{a} = \mathbf{0} \quad (2.23)$$

Assuma agora que tem um conjunto de m pontos $(x_1^i, x_2^i, \dots, x_n^i)$ em que $i = 1, \dots, m$ e que existe um vector $\mathbf{a} = (a_1, a_2, \dots, a_n)$ tal que $a_1x_1^i + a_2x_2^i + \dots + a_nx_n^i = 0, \forall i$. Considere ainda que a cada valor x_j^i corresponde um erro δ_j^i . Mais uma vez pretende determinar o vector \mathbf{a} que minimize uma certa métrica do erro.

Considere a matriz X da equação 2.19 e a matriz $\Delta = [\delta_j^i]$ do erro. O objectivo é calcular o valor de \mathbf{a} que verifica a equação 2.23 em que a norma de Frobenius de Δ é mínima (minimização do erro). A solução trivial ($\mathbf{a} = \mathbf{0}$) obviamente que verifica 2.23. Para existir outra solução, e atendendo a que $\mathbf{X} - \Delta$ é uma matriz $m \times n$ em que $n < m$, a característica tem de ser menor que n . Considere a decomposição em valores singulares de \mathbf{X} (equação 2.22). Escolha a matriz de erro de forma a que $\Delta = \mathbf{U}\mathbf{S}_\delta\mathbf{V}^t$ em que \mathbf{S}_δ é uma matriz diagonal em que todos os elementos são nulos excepto um. O único valor singular de Δ não nulo é igual ao menor valor singular de \mathbf{X} . Note que desta forma a matriz $\mathbf{X} - \Delta$ é não singular e a norma de Frobenius de Δ é mínima. Resolvendo agora o sistema obtém-se a solução desejada (não trivial) a menos de um factor de escala. A solução também se pode tirar directamente da matriz \mathbf{V} , escolhendo a coluna correspondente ao menor valor singular de \mathbf{X} .

2.3 Estimação da Matriz Fundamental

Para obter a geometria epipolar de um par estéreo não-calibrado é necessário estimar a matriz fundamental. Existem vários métodos de estimar a matriz fundamental. O algoritmo aqui apresentado é introduzido em [11].

A equação 2.11 mostra que cada par de pontos correspondentes $(\bar{\mathbf{m}}_1, \bar{\mathbf{m}}_2)$ entre as duas imagens proporciona uma restrição linear de \mathbf{F} . Isto permite estimar a matriz fundamental, a menos de um factor de escala, a partir de 8 correspondências. Algebricamente \mathbf{F} é uma matriz 3×3 sendo necessário estimar 9 parâmetros. Devido ao factor de escala um desses parâmetros é fixado a 1 ($f_{33} = 1$). A matriz fundamental é singular tendo de se verificar

que $\det(\mathbf{F}) = 0$ (característica 2). Conclui-se assim que só existem 7 graus de liberdade independentes. A matriz fundamental pode ser estimada a partir de 7 correspondências mais a restrição do determinante. No entanto, dado a última restrição ser uma equação cúbica, existiriam três soluções para o problema. Por isso se utilizam 8 ou mais correspondências no processo de estimação.

$$\mathbf{F} = \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & 1 \end{bmatrix} \quad (2.24)$$

$$x_1x_2f_{11} + y_1x_2f_{12} + x_2f_{13} + x_1y_2f_{21} + y_1y_2f_{22} + y_2f_{23} + x_1f_{31} + y_1f_{32} + 1 = 0 \quad (2.25)$$

A equação 2.24 mostra a matriz fundamental. Para cada par de pontos correspondentes ($\bar{\mathbf{m}}_1 = (x_1, y_1, 1)^t$ e $\bar{\mathbf{m}}_2 = (x_2, y_2, 1)^t$) a restrição linear que obtém é mostrada na equação 2.25. A matriz \mathbf{F} é estimada aplicando o método dos mínimos quadráticos totais.

Note que nada garante que a matriz fundamental estimada tenha característica 2. Para isso faça a decomposição em valores singulares de \mathbf{F} , ponha a zero o valor singular mais baixo e recalcule a matriz. Desta forma obtém uma matriz singular.

Capítulo 3

Trabalho Prático

A geometria epipolar é tudo o que se pode saber sobre um par de imagens estéreo. A matriz fundamental estabelece a correspondência entre um ponto numa imagem e a sua linha epipolar na outra, e permite a reconstrução da cena a menos de uma transformação projectiva 3D. Isto é tudo o que se pode fazer com um par de câmaras não calibradas sem mais nenhuma restrição adicional. Uma aplicação prática da geometria epipolar é auxiliar o estabelecimento de correspondência entre pontos das duas imagens. Dado um ponto numa imagem, se conhecer a matriz fundamental, a área de busca da correspondência na segunda imagem é uma única linha epipolar (em vez de todo o plano imagem). Este trabalho pretende demonstrar que a matriz fundamental entre duas imagens pode ser deduzida a partir da correspondência entre alguns pontos. Para isto não é necessária qualquer calibração das câmaras.

3.1 Carregamento de Dados

Este trabalho é para ser realizado no ambiente do *MATLABTM* com a "images toolbox" instalada. Na sua directoria de trabalho existem dois ficheiros: "loader.m" e "VerifyEpipolarGeometry.m". O primeiro é um "script" e o segundo é uma função. Comece por correr o "script".

O "script loader" carrega para o ambiente do *MATLABTM* os dados com que vai trabalhar disponibilizados pelo "Calibrated Imaging Laboratory" da "Carnegie Mellon University" em Pittsburgh, Pensilvânia, Estados Unidos da América ("http://www.cs.cmu.edu/cil/cil-ster.html"). As imagens a ser utilizadas são um conjunto de perspectivas da cena exibida na Fig. 3.1. O "CIL" seleccionou 28 pontos da cena, que estão devidamente assinalados na figura, e determinou a posição desses pontos em cada uma das perspectivas obtidas. Segundo as especificações o erro na correspondência é inferior a 0.1

Cena com os pontos de referência assinalados

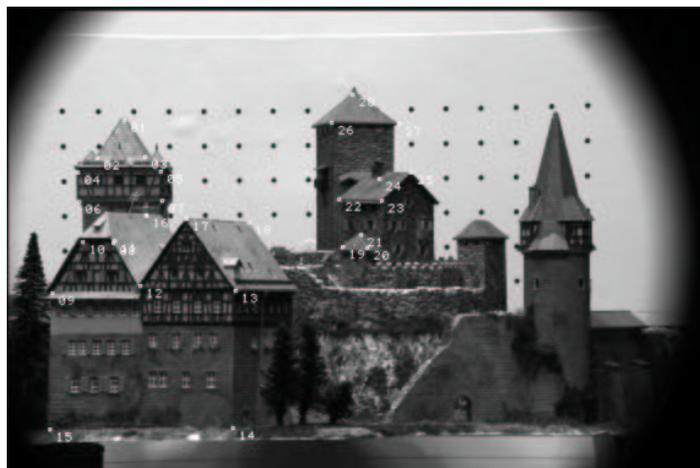


Figura 3.1: Imagens disponibilizadas pelo "Calibrated Imaging Laboratory" da "Carnegie Mellon University" em Pittsburgh, Pensilvânia, Estados Unidos da América. Estão assinaladas as 28 "features" de correspondência.

pixels.

A Fig. 3.2 mostra as quatro perspectivas diferentes que vai utilizar. Cada uma destas perspectivas é uma imagem 576×384 guardada nas tabelas bidimensionais "image1", "image2", "image3" e "image4". Se listar as variáveis correntes no espaço de trabalho (comando "who") encontra ainda "gt1", "gt2", "gt3", "gt4", "A1", "A2", "A3", "A4" e "R34". As variáveis "gt*i*" são tabelas 28×2 com as coordenadas (x,y) em pixels de cada um dos pontos de referência na imagem *i*. Esses pontos são assinalados pelos círculos vermelhos na Fig. 3.2. As restantes variáveis são todas matrizes 3×3 . "A*i*" é a matriz dos parâmetros intrínsecos da câmara que capturou a imagem *i*. "R34" é a matriz de rotação entre a câmara 3 e 4 de tal forma que se \mathbf{M}_3 e \mathbf{M}_4 são, respectivamente, as coordenadas cartesianas dum ponto no sistema de eixos coordenados da câmara 3 e 4 então $\mathbf{M}_4 = \mathbf{R}_{34}\mathbf{M}_3 + \mathbf{t}_{34}$ em que \mathbf{t}_{34} é o vector de translação entre os centros ópticos (desconhecido).

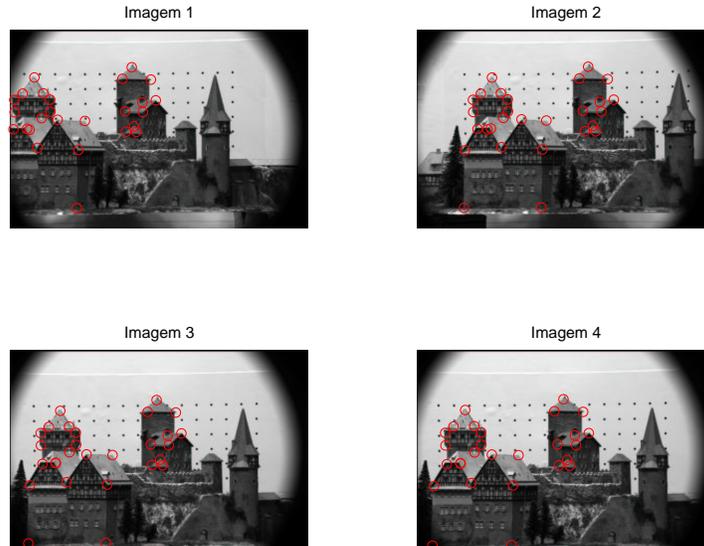


Figura 3.2: As quatro imagens de trabalho com 28 correspondências entre pontos.

3.2 A Função de Verificação da Geometria Epipolar

Para além do "script" tem à sua disposição a função "VerifyEpipolarGeometry". A sua forma de utilização é mostrada digitando "help VerifyEpipolarGeometry".

Imagine que tem duas perspectivas de uma mesma cena e que, de alguma forma, estimou a matriz fundamental. Esta função permite-lhe averiguar a qualidade dessa estimativa. Existem quatro parâmetros de entrada: a matriz fundamental, as duas imagens e o número N de pontos que pretende seleccionar (valor compreendido entre 1 e 6). A função mostra-lhe as duas imagens e pede-lhe para seleccionar com o rato um ponto na primeira imagem. Seguidamente a matriz fundamental é utilizada para estimar a correspondente linha epipolar na segunda imagem. A linha epipolar é assinalada permitindo-lhe inspeccionar visualmente os resultados. A Fig. 3.3 mostra este procedimento. O utilizador seleccionou na imagem da esquerda o topo da última torre (círculo vermelho). Como pode verificar a linha epipolar calculada passa pelo topo da torre na segunda imagem. Isto indica que a matriz fundamental é provavelmente correcta. Caso contrário algo estaria mal. Este processo repete-se N vezes. No final o programa assinala as N

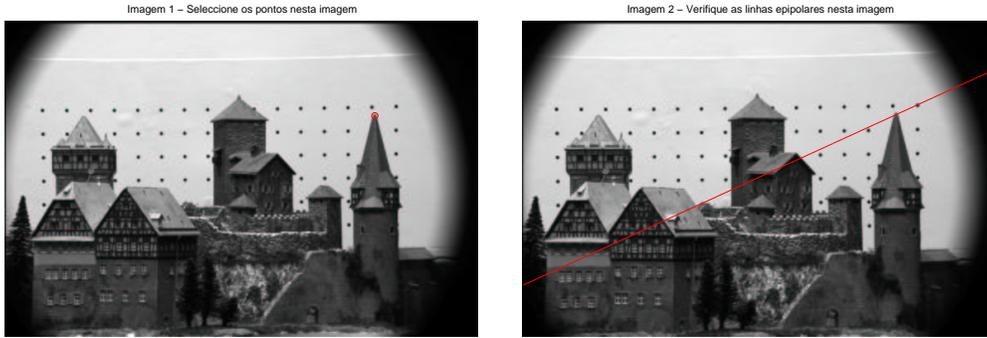


Figura 3.3: Utilização da função "VerifyEpipolarGeometry"

linhas epipolares na primeira imagem e devolve duas matrizes, uma com as coordenadas homogêneas dos pontos seleccionados na primeira imagem, e outra com as coordenadas das linhas epipolares da segunda imagem.

3.3 Trabalho a realizar

3.3.1 Estimação da matriz fundamental

Na primeira parte da aula prática deverá programar uma função em *MATLABTM* que calcule a matriz fundamental entre duas imagens. Para isso assuma que a correspondência entre um conjunto de pontos das duas imagens já foi estabelecida previamente. A sua função deve receber como parâmetros de entrada as coordenadas, numa e noutra imagem, dos pontos de referência. Com essa informação deve estimar a matriz fundamental. A matriz estimada é devolvida à saída. Opcionalmente pode ainda incluir nos parâmetros de saída algum valor ou conjunto de valores que lhe permita ter uma ideia do erro de estimação numérico (erro médio quadrático ou qualquer outra métrica que entenda).

3.3.2 Verificação dos resultados de estimação da matriz fundamental

Utilize a função que programou para estimar as matrizes \mathbf{F}_{12} , \mathbf{F}_{23} e \mathbf{F}_{34} que são, respectivamente, as matrizes fundamentais entre as imagens 1 e 2, 2 e 3, e 3 e 4. Utilize a função "VerifyEpipolarGeometry" para confirmar os seus resultados.

3.3.3 Informação extraída da Matriz Fundamental

Aplique os conhecimentos adquiridos na aula teórica para resolver as seguintes questões.

- Considere as imagens 3 e 4 e a respectiva matriz fundamental. Conhecendo as matrizes de parâmetros intrínsecos das câmaras (\mathbf{A}_3 e \mathbf{A}_4) calcule a matriz essencial.
- Observe a matriz essencial deduzida. O que consegue inferir sobre movimento da câmara? É capaz de o quantificar em unidades métricas? Justifique.
- Considere agora a matriz de rotação entre as câmaras 3 e 4. Determine as coordenadas do epipólo na imagem 3. Assinale na imagem o ponto que obteve e confirme se as rectas epipolares se intersectam nesse ponto.
- Calcule agora as matrizes essenciais entre as imagens 1 e 2, e 2 e 3. Utiliza as matrizes \mathbf{A}_1 , \mathbf{A}_2 e \mathbf{A}_3 para calcular as matrizes essenciais. Que informações consegue extrair sobre a posição relativa das câmaras? A disposição das linhas epipolares pode ajudá-lo neste problema?

Capítulo 4

Resolução do Trabalho Prático

Os dois capítulos anteriores foram textos concebidos para serem disponibilizados aos alunos. O capítulo 2 é uma introdução teórica que deve ser estudada antes da aula prática a fim de preparar o trabalho. Todas as matérias aí abordadas foram objecto de estudo nas aulas teóricas. O texto destina-se a ser um resumo para rever os conceitos necessários à realização do trabalho. O capítulo 3 é o enunciado do trabalho a realizar. A aula começa com uma introdução, nunca superior a meia-hora, em que os acetatos em anexo são utilizados para apresentar o trabalho e esclarecer dúvidas que entretanto tenham surgido. Na restante hora e meia os alunos programam as rotinas pedidas e respondem às perguntas do questionário. A entrega ou não de um breve relatório dependerá do esquema de avaliação em vigor.

Este capítulo não se destina a ser visto pelos alunos. Trata-se da resolução do enunciado com os objectivos que se pretendem que eles atinjam.

4.1 Estimação da matriz Fundamental

Esta é a primeira parte do trabalho estando previsto um tempo de realização de 30 a 45 minutos. A introdução teórica indica de forma muito clara o procedimento. Um exemplo de resposta é a função "ComputeFundamentalMatrix" mostrada em anexo.

Para cada par de imagens existem 28 pontos correspondentes que são os parâmetros de entrada da função. Esses pontos são utilizados para construir a matriz \mathbf{A} de acordo com a equação 2.25. \mathbf{A} é uma matriz 28×9 . Seguidamente aplica-se o método dos mínimos quadráticos totais também descrito na introdução teórica. Os alunos devem compreender que, como os elementos de \mathbf{A} resultam de medições na imagem, existem erros em todos eles. Assim faz-se a decomposição em valores singulares da matriz e

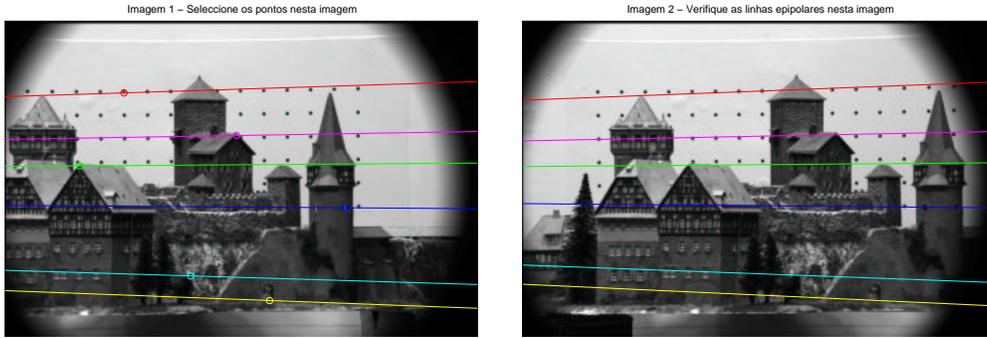


Figura 4.1: Geometria epipolar entre as imagens 1 (esquerda) e 2 (direita). A correspondência entre linhas epipolares é feita utilizando as estimativa de \mathbf{F}_{12} obtida.

escolhe-se a coluna de \mathbf{V} correspondente ao menor valor singular. Este é o vector $\mathbf{f} = (f_{11}, f_{12}, \dots, f_{33})^t$. A matriz \mathbf{F} é construída de acordo com a equação 2.24. No entanto nada garante que a matriz estimada tenha característica 2. Assim, faz-se a decomposição em valores singulares de \mathbf{F} , coloca-se o menor valor singular a zero, e recalcula-se a matriz. O resultado é a estimativa da matriz fundamental. A rotina "ComputeFundamentalMatrix" devolve ainda uma vector com os erros de estimação em cada ponto. Isto permite ter uma noção da eficiência do processo de minimização implementado.

Utilizando a rotina para estimar a matriz fundamental entre as imagens 1 e 2, 2 e 3, e 3 e 4 obtém-se \mathbf{F}_{12} , \mathbf{F}_{23} e \mathbf{F}_{34} . A função "VerifyEpipolarGeometry" é usada para verificar os resultados. As Fig 4.1, 4.2 e 4.3 mostram os resultados dessa verificação.

4.2 Extracção de Informação da Matriz Fundamental

A última parte do trabalho consiste em utilizar as estimativas da matriz fundamental para extrair informação sobre a posição relativa das câmaras.

$$\mathbf{E}_{ij} \simeq (\mathbf{A}_j)^t \mathbf{F}_{ij} \mathbf{A}_i \quad (4.1)$$

Pretende-se estimar a matriz essencial \mathbf{E}_{34} entre a imagem 3 e 4 conhecendo as matrizes de parâmetros intrínsecos. A equação 4.1 é deduzida a partir do resultado 2.10 e permite calcular a matriz essencial entre uma imagem

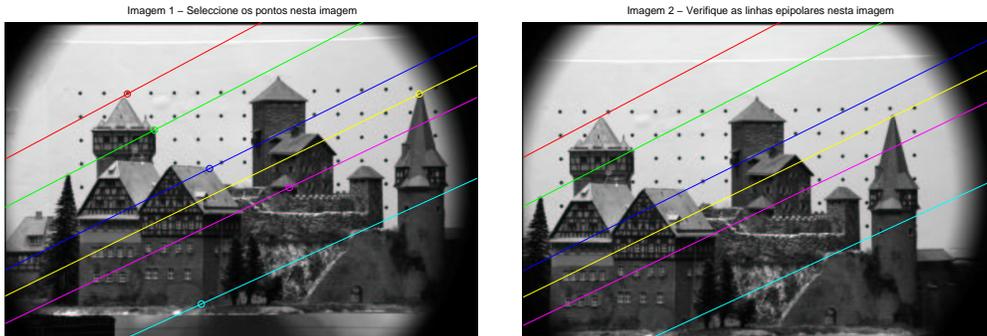


Figura 4.2: Geometria epipolar entre as imagens 2 (esquerda) e 3 (direita). A correspondência entre linhas epipolares é feita utilizando as estimava de \mathbf{F}_{23} obtida.

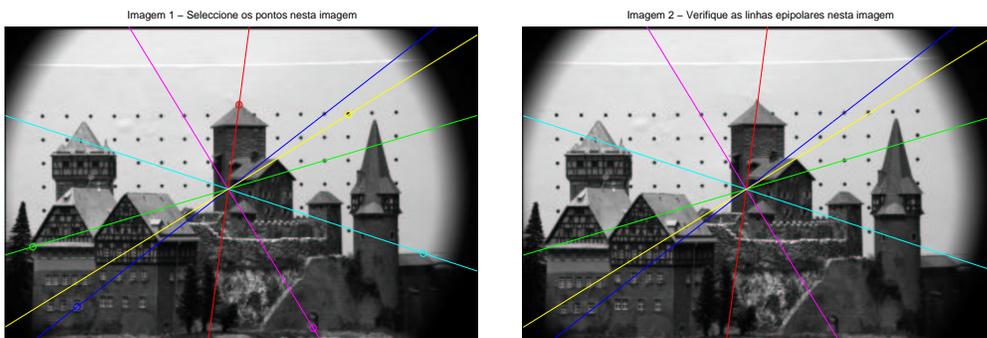


Figura 4.3: Geometria epipolar entre as imagens 3 (esquerda) e 4 (direita). A correspondência entre linhas epipolares é feita utilizando as estimava de \mathbf{F}_{34} obtida.

i e uma imagem j dada a matriz fundamental e os parâmetros intrínsecos. Fazendo $i = 3$ e $j = 4$ obtém-se \mathbf{E}_{34} .

$$\mathbf{E}_{34} = \begin{bmatrix} -0.0000 & 1.2881 & 0.0020 \\ -1.2881 & -0.0001 & 0.0072 \\ -0.0019 & -0.0073 & -0.0000 \end{bmatrix} \quad (4.2)$$

A equação 4.2 mostra a matriz essencial. Note que é uma matriz aproximadamente anti-simétrica. Viu-se na equação 2.6 que a matriz essencial é o produto de uma matriz anti-simétrica, devida à translação, pela matriz de rotação. Daqui se conclui que o movimento de translação é claramente dominante em relação ao movimento de rotação. A componente de translação é sobretudo significativa ao longo do eixo do Z (ver equação 2.5). Como \mathbf{E}_{34} é definida a menos de um factor de escala não se pode estabelecer nenhuma relação com grandezas métricas do mundo real.

Dadas as matrizes de parâmetros intrínsecos e a matriz de rotação \mathbf{R}_{34} , a homografia de rotação entre as imagens 3 e 4 é dada por $\mathbf{H} = \mathbf{A}_4 \mathbf{R}_{34} \mathbf{A}_3^{-1}$. A equação 2.15 estabelece a relação entre a homografia de rotação, a matriz fundamental e as coordenadas projectivas do epipólo. A função "GetHomography" em anexo exemplifica como se pode estimar a posição do epipólo a partir da homografia de rotação e da matriz fundamental.

$$\begin{bmatrix} \mathbf{0} & -\mathbf{h}_3 & \mathbf{h}_2 \\ \mathbf{h}_3 & \mathbf{0} & -\mathbf{h}_1 \\ -\mathbf{h}_2 & \mathbf{h}_1 & \mathbf{0} \end{bmatrix} \mathbf{e} = \mathbf{f} - \delta \quad (4.3)$$

Considere o vector $\mathbf{f} = (f_{11}, f_{12}, \dots, f_{33})^t$ com os parâmetros estimados para a matriz fundamental. Dado tratar-se de estimativas é razoável assumir que cada parâmetro vem afectado de um erro. Seja δ o vector de erro. Assuma ainda que o epipólo é $\mathbf{e} = (e_x, e_y, e_z)^t$ e que \mathbf{h}_1 , \mathbf{h}_2 e \mathbf{h}_3 são as colunas da matriz $(\mathbf{H}^{-1})^t$. A equação 2.15 pode ser reescrita na forma 4.3. Aplica-se agora o método dos mínimos quadrados de forma a calcular o vector $\mathbf{e} = (e_x, e_y, e_z)^t$ que verifica 4.3 minimizando o erro δ .

\mathbf{e} é definido a menos de um factor de escala. Dividindo as três coordenadas por e_z obtém-se $(x, y, 1)$ em que (x, y) são as coordenadas na imagem. A Fig. 4.4 mostra o ponto obtido.

Imagem 1 – Seleccione os pontos nesta imagem

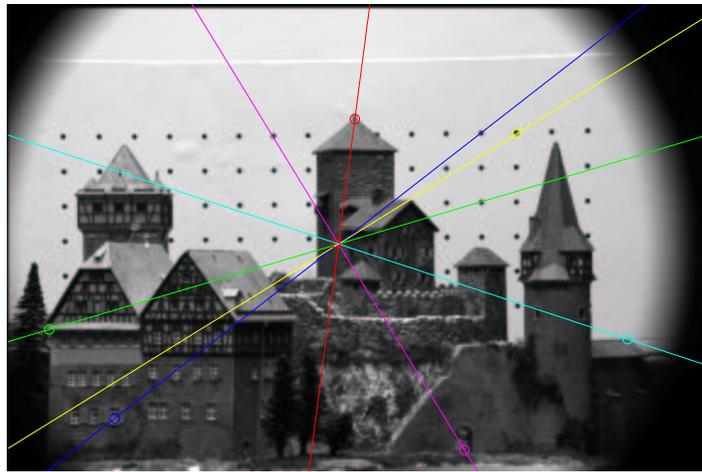


Figura 4.4: Epipolo na imagem 3 (quadrado a vermelho)

Bibliografia

- [1] H. C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, September 1981.
- [2] J. G. Semple and G. T. Kneebone. *Algebraic Projective Geometry*. Oxford University Press, 1998.
- [3] O. Faugeras. *Three Dimensional Computer Vision*. Artificial Intelligence. MIT Press, Cambridge, MA, 1993.
- [4] Roger Mohr and Bill Triggs. Projective geometry for image analysis. In *A tutorial given at ISPRS*. Vienna, September 1996.
- [5] Stan Birchfield. An introduction to projective geometry for computer vision. In *Tutorial*. March 1998.
- [6] Q. T. Luong and O. Faugeras. The fundamental matrix: Theory, algorithms and stability analysis. *International Journal of Computer Vision*, 17(1):43–76, 1996.
- [7] Z. Zhang and G. Xu. *Epipolar Geometry in Stereo, Motion and Object Recognition: A unified Approach*. Kluwer Academic Publishers, 1996.
- [8] R. Mohr, B. Boufama, and P. Brand. Understanding positioning from multiple images. *Artificial Intelligence*, 78:213–238, 1995.
- [9] Athanasios Papoulis. *Probability, Random Variables and Stochastic Processes*. MacGraw-Hill International Editions, 1991.
- [10] Ralph Deutsch. *Estimation Theory*. Prentice-Hall Inc, 1965.
- [11] R. Hartley. In defence of the 8-point algorithm. In *Proc. of the 5th International Conference on Computer Vision*, pages 1064–1070, Cambridge, Massachusetts, USA, 1995.