VISÃO POR COMPUTADOR
2017/2018

ESTIMAÇÃO DA MATRIZ FUNDAMENTAL, DA MATRIZ ESSENCIAL E RECONSTRUÇÃO 3D

Considere-se duas câmaras colocadas em duas posições distintas. Sejam P_e e P_d as coordenadas de um ponto P (no espaço) nos referenciais respectivamente da câmara esquerda e da câmara direita. Sejam p_e e p_d as coordenadas das imagens desse ponto na câmara esquerda e na câmara direita, respectivamente. Note-se que P_e e P_d são as coordenadas do mesmo ponto no espaço expressas em dois sistemas de coordenadas distintos enquanto que p_e e p_d são coordenadas de dois pontos distintos.

Os sistemas de coordenadas associados a cada câmara estão relacionados entre si através de uma transformação de corpo rígido, a qual é definida por um vector de translação $T = (O_d - O_e)$ e por uma matriz de rotação R.

Dado um ponto P no espaço 3D a relação entre as suas coordenadas P_e e P_d é dada por:

$$P_d = R(P_e - T) \qquad (1)$$

A matriz fundamental formaliza as relações geométricas entre as imagens de um par de câmaras que "observa" uma mesma parte do espaço 3D. A geometria correspondente a essa configuração é geralmente designada por geometria epipolar. O plano definido pelo ponto 3D P e pelos centros de projecção das câmaras esquerda (O_e) e direita (O_d) intersecta os planos-imagem das câmaras esquerda e direita em rectas que são designadas por rectas epipolares. Por outro lado a recta definida pelos centros de projecção das câmaras intersecta os correspondentes planos-imagem em pontos que são designados por epipolos. Assim, o epipolo esquerdo é a imagem do centro de projecção da câmara direita e vice-versa.

1. MATRIZ ESSENCIAL

Os vectores P_e, T e $P_e - T$ estão contidos num plano que é o plano epipolar. Estes três vectores são definidos no sistema de coordenadas da câmara esquerda. Pelo facto de pertencerem ao mesmo plano verificam/satisfazem a seguinte equação:

$$(P_e - T)^T (T \times P_e) = 0 \qquad (2)$$

Por outro lado da equação (1) resulta que:

$$P_e - T = R^T P_d \qquad (3)$$

Substituindo na Equação (2) obtém-se:

$$(R^T P_d)^T (T \times P_e) = 0 \qquad (4)$$

Por outro lado o produto vectorial pode ser expresso algebricamente por meio de um produto de uma matriz anti-simétrica (composta pelos elementos de um dos vectores) pelo outro vector ou seja:

$$T \times P_e = S P_e \qquad (5)$$

Onde S é dado por:

$$S = \begin{bmatrix} 0 & -T_z & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{bmatrix} \qquad (6)$$

Consequentemente a equação (4) pode ser re-escrita da seguinte forma:

$$P_d^T E P_e = 0 \qquad (7)$$

onde

$$E = RS$$

Por outro lado podemos substituir nesta equação P_d por p_d e P_e por p_e pois $p_e = \frac{f_e}{Z_e} P_e$ e $p_d = \frac{f_d}{Z_d} P_d$ em que $p_e = \begin{bmatrix} x_e & y_e & f_e \end{bmatrix}^T$ e $p_d = \begin{bmatrix} x_d & y_d & f_d \end{bmatrix}^T$ onde f_e e f_d são respectivamente as distâncias focais da câmara esquerda e direita. Assim pode escrever-se:

$$p_d^T E\, p_e = 0 \qquad (8)$$

A matriz E é designada por matriz essencial. De notar que nesta equação as coordenadas dos pontos nas imagens não são expressas em pixels mas sim numa unidade métrica qualquer (mm por exemplo). A matriz E tem característica 2.

Se interpretarmos as coordenadas dos pontos nas imagens projectivamente, então u_d dado por

$$u_d = E\, p_e \qquad (9)$$

representa as coordenadas de uma recta na imagem direita. Esta recta é a recta epipolar correspondente ao ponto p_e da imagem esquerda. Esta recta epipolar contém o ponto p_d (na imagem direita). A recta epipolar passa também pelo epipolo na imagem direita.

2. MATRIZ FUNDAMENTAL

Sejam K_e e K_d as matrizes dos parâmetros intrínsecos das câmaras esquerda e direita respectivamente. As coordenadas dos pontos p_e e p_d em pixeis são então dadas por:

$$\hat{p}_e = K_e\, p_e \quad \text{e} \quad \hat{p}_d = K_d\, p_d$$

de onde se tira que,

$$p_e = K_e^{-1}\, \hat{p}_e \quad \text{e} \quad p_d = K_d^{-1}\, \hat{p}_d$$

que, substituído na equação (8), dá

$$\hat{p}_d^T F\, \hat{p}_e = 0 \qquad (10)$$

onde

$$F = K_d^{-T} E K_e^{-1} \qquad (11)$$

A matriz F é designada por matriz fundamental. Tal como no caso da matriz essencial

$$\hat{u}_d = F\,\hat{p}_e$$

representa a recta epipolar na imagem direita correspondente ao ponto $\hat{p}_e$. Note-se que, neste caso, as coordenadas dos pontos nas imagens são em pixeis. A matriz fundamental, tal como a essencial, tem característica 2.

3. ESTIMAÇÃO DA MATRIZ FUNDAMENTAL: ALGORITMO DOS 8 PONTOS

Considere-se um conjunto de pontos correspondentes nas imagens esquerda e direita. Cada par de pontos correspondentes satisfaz a equação (10). A matriz F é uma matriz 3x3. Assim temos 9 incógnitas. No entanto a equação (10) é uma equação homogénea o que faz com que possa ser multiplicada por uma constante qualquer diferente de zero. Consequentemente a matriz F é definida a menos de um factor de escala. Esse grau de liberdade reduz-nos o número de incógnitas a 8. Assim se tivermos 8 ou mais pares de pontos correspondentes nas imagens esquerda e direita podemos estimar a matriz F (cada par de pontos permite definir uma equação que tem como incógnitas os elementos da matriz F). No entanto é importante assegurar que os pontos usados não constituem uma configuração degenerada, por exemplo, pertencendo todos a um plano. Para estimar a matriz F no sentido dos mínimos quadrados, impondo a restrição de que a norma do vector das incógnitas (os elementos da matriz F) é unitária, pode usar-a a "singular value decomposition" (svd). Seja A a matriz do sistema de equações obtido a partir da equação (10). A decomposição em valores singulares da matriz A dá

$$A = U D V^T \qquad (11)$$

As matrizes U e V são matrizes ortogonais. A matriz D contém na sua diagonal os valores singulares. A solução do sistema de equações é dada pela coluna da matriz V correspondente ao único valor singular nulo da matriz A. Na prática, devido ao ruído, a matriz A tem característica 9 ou seja, nenhum valor singular é nulo. A solução é então a coluna da matriz V correspondente ao menor valor singular da matriz A.

A matriz F que resulta deste processo de estimação deveria ter característica 2. No entanto devido ao ruído e às imprecisões no processo de estimação só raramente a matriz F tem característica 2. É assim necessário impor essa condição. Para isso faz-se a decomposição em valores singulares da matriz F obtendo-se:

$$F = UDV^T$$

Substitui-se então a matriz D por uma nova matriz D' que é igual à matriz D com excepção do menor valor singular que se faz igual a 0. Recalcula-se então a matriz F de modo a assegurar a característica 2. Assim obtém-se a matriz F' dada por:

$$F^{'} = UD^{'}V^T$$

Um elemento muito importante para uma boa estimação da matriz F é a normalização dos dados. A normalização tem como objectivo tornar o problema mais bem condicionado do ponto de vista numérico. Sem normalização os elementos da matriz A terão grandes variações. Interessa diminuir o intervalo de variações possíveis nos elementos da matriz A, que dependem das coordenadas do pixeis. Para isso faz-se uma translação e uma mudança do factor de escala de cada uma das imagens de modo a que a origem do sistema de coordenadas se situe no centro de massa dos pixeis usados na estimação da matriz F e também de modo a que a distância média quadrática dos pixeis à origem seja de 1. Consideremos os pontos $\hat{p}_i$ de uma imagem com coordenadas $\hat{p}_i = \begin{bmatrix} x_i & y_i & 1 \end{bmatrix}^T$, em que $i = 1,...,n$. O centro de massa tem então coordenadas $\bar{x} = \sum_i x_i / n$ e $\bar{y} = \sum_i y_i / n$. Seja

$$\bar{d} = \frac{\sum_i \sqrt{\left(x_i - \bar{x}\right)^2 + \left(y_i - \bar{y}\right)^2}}{n\sqrt{2}}$$

Com os valores de $\bar{x}, \bar{y}$ e de $\bar{d}$ para cada uma das imagens constroem-se duas matrizes T_e e T_d que permitem obter as coordenada normalizadas nas duas imagens por meio de:

$$\hat{p}_e^{'} = T_e \hat{p}_e \qquad \text{e} \qquad \hat{p}_d^{'} = T_d \hat{p}_d$$

A estimação da matriz F faz-se então com estas coordenadas. A matriz resultante deste processo de estimação é a matriz $\hat{F}$ que é dada por:

$$\hat{F} = T_d^{-T} F T_e^{-1}$$

Estimada a matriz $\hat{F}$ obtém-se a matriz F através de:

$$F = T_d^T \hat{F} T_e$$

Algoritmo dos 8 pontos

Considerem-se n pontos correspondentes (nas imagens esquerda e direita) com $n \geq 8$. Marque os pontos correspondentes à mão ou então determine-os automaticamente primeiro (usando o detector de cantos do trabalho anterior) e estabeleça as correspondências à mão. <u>Atenção: os pontos não devem ser todos co-planares</u>.

1. Normalizar as coordenadas os pontos correspondentes, nas duas imagens usando $\hat{p}_e^{'} = T_e \hat{p}_e$ e $\hat{p}_d^{'} = T_d \hat{p}_d$.
2. Construir o sistema de equações homogéneo definido pela equação 10. Seja A a matriz $n x 9$ do sistema e $A = UDV^T$ a sua decomposição em valores singulares.
3. Os elementos da matriz $\hat{F}$ (a menos de um factor de escala) são dados pelos elementos da coluna da matriz V correspondentes ao menor valor singular de A.
4. Para impor a restrição da característica da matriz $\hat{F}$ decompõe-se a matriz $\hat{F}$ em valores singulares:
$$\hat{F} = \hat{U}\hat{D}\hat{V}^T$$
5. Coloca-se o menor valor singular na matriz $\hat{D}$ a zero e faz-se a matriz D' igual a essa matriz.
6. Calcula-se novamente a matriz fundamental usando:
$$F^{'} = UD^{'}V^T$$
7. "Desnormalização": a matriz fundamental F é então obtida através de,
$$F = T_d^T F^{'} T_e$$

4. ESTIMAÇÃO DOS EPIPOLOS

O epipolo na imagem esquerda $\hat{e}_e$ pertence a todas as rectas epipolares. Assim, qualquer que seja $\hat{p}_d$ obtém-se

$$\hat{p}_d^T F \hat{e}_e = 0$$

O que significa que

$$F \hat{e}_e = 0$$

Como a matriz F tem característica 2 (sendo uma matriz 3x3) , conclui-se que o epiplo esquerdo $\hat{e}_e$ corresponde ao espaço nulo direito da matriz F. Por outro lado o epipolo direito $\hat{e}_d$ corresponde ao espaço nulo direito da matriz F^T . Consequentemente para se determinar os dois epipolos estimam-se os espaços nulos direitos das matrizes F e F^T .

Algoritmo para estimar a localização dos epipólos

Dada a matriz fundamental F,

1. Decompõe-se a matriz F em valores singulares
$$F = UDV^T$$
2. O epipolo $\hat{e}_e$ é dado pela coluna da matriz V correspondente ao valor singular nulo da matriz F.
3. O epipolo $\hat{e}_d$ é dado pela coluna da matriz U correspondente ao valor singular nulo da matriz F.

Objectivos do trabalho

1ª Parte

1—Estimar a matriz fundamental. Para isso deve usar as três imagens captadas por cada câmara, pois a matriz fundamental não pode ser calculada com a imagens de um só plano (lef01.jpg, left02.jpg. left03.jpg assim como right01.jpg, right02.jpg,e right03.jpg). Use o comando “detectCheckerboardpoints” para detectar as coordenadas dos cantos. Estabeleça as correspondências manualmente e calcule a matriz fundamental usando o algoritmo dos 8 pontos.
2—Estimar os epipolos e representá-los nas imagens;
3—Representar nas imagens rectas epipolares correspondentes. O que conclui quanto ao movimento da câmara entre as imagens?
4—Calcule a matriz fundamental usando o comando do Matlab “estimateFundamentalMatrix”. Compare as duas matrizes visualmente, fazendo o display dos epipolos e de rectas epipolares correspondentes.

2ª Parte

Sabendo que as matrizes dos parâmetros intrínsecos das câmaras esquerda e direita são respectivamente:

$$K_{esq} = \begin{bmatrix} 533.0031 & 0 & 341.586 \\ 0 & 533.1526 & 234.259 \\ 0 & 0 & 1 \end{bmatrix} \quad e$$

$$K_{dir} = \begin{bmatrix} 536.9826 & 0 & 326.472 \\ 0 & 536.56938 & 249.3326 \\ 0 & 0 & 1 \end{bmatrix}$$

5—Determine a matriz essencial;

6—A partir da matriz essencial determine a rotação e a translação entre as duas câmaras usando um dos algoritmos descritos na aula teórica;

7—Faça a reconstrução 3D (determine as coordenadas 3D) dos pontos que usou para estimar a matriz essencial. Use o algoritmo descrito na aula teórica ou o comando Matlab “triangulate”. Se usar o comando “triangulate” tem que criar a estrutura “stereoParameters”.