

Lista 7 – Visão Computacional

Aluno: Rennan de Lucena Gaio

DRE: 119122454

Todo o código do trabalho pode ser acessado pelo link do github:

https://github.com/RennanGaio/visao_computacional/tree/master/lista7

Visão Estereoscópica

1. Solução item 1

Dado o conjunto de pontos em 3D no arquivo world.txt e, conjunto com sua projeção na imagem em 2D no arquivo imagem.txt utilizou-se o método de DLT para o calculo da sua matriz de projeção P . Após a obtenção da matriz de projeção foi aplicado o método de decomposição RQ da biblioteca scipy. Como resultado obteve-se as seguintes matrizes:

$$P = \begin{pmatrix} 1.93e + 15 & 3.87e + 15 & 5.81e + 15 & 7.75e + 15 \\ 7.75e + 15 & 5.81e + 15 & 3.87e + 15 & 1.93e + 15 \\ 0.193e + 15 & -2.54 & 1.93e + 15 & 1 \end{pmatrix}$$

$$K = \begin{pmatrix} 7.91e + 15 & 4.47e + 15 & -5.48e + 15 \\ 0 & 6.71e + 15 & -8.22e + 15 \\ 0 & 0 & -2.74e + 15 \end{pmatrix}$$

$$Rt = \begin{pmatrix} -0.40 & 0 & 0.40 & 0.81 \\ 0.28 & 0.86 & -0.28 & 0.28 \\ -0.70 & 0 & -0.70 & 0 \end{pmatrix}$$

$$KR = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 3 & 2 \\ 1 & 0 & 1 \end{pmatrix}$$

2. Solução item 2

Dadas as matrizes P1 e P2, foi executada a decomposição RQ da mesma forma que no exercício anterior com o método de scipy. Para realizar a normalização das matrizes é necessário fazer a multiplicação da inversa da matriz de parâmetros intrínsecos (K^{-1}) com sua respectiva matriz de câmera P. Foram obtidos os seguintes resultados:

$$K1^{-1} \cdot P1 = \begin{bmatrix} -0.82 & 0.13 & -0.53 & -0.07 \\ 0.35 & 0.84 & -0.35 & 0.17 \\ 0.22 & 0 & -0.22 & -0.94 \end{bmatrix}$$

$$K2^{-1} \cdot P2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -0.86 & -0.50 & 0 \\ 0 & 0.35 & -0.61 & -0.70 \end{bmatrix}$$

3. Solução item 3

Dadas as matrizes normalizadas encontradas anteriormente, foi calculada a matriz fundamental que relaciona as duas matrizes. Para realizar tal tarefa, foi necessário fazer o cálculo do centro das duas câmeras com a seguinte fórmula:

$$P = [M | K_c t_w]$$

$$\text{center} = M^{-1} \cdot P[:, 3]$$

Em que M são as 3 primeiras colunas de P normalizado e P[:,3] é a última coluna de P normalizado.

Após o cálculo do centro de câmera é encontrado o epipolo na imagem oposta aplicando a projeção desta imagem no ponto de centro da outra câmera. Para encontrar a Matriz fundamental que relaciona as duas câmeras é utilizada a fórmula:

$$F = [e_{x2}] \cdot P2 \cdot P1^+$$

$$P1^+ = P1 \cdot T @ (P1 @ P1 \cdot T)^{-1}$$

Em que $[e_{x2}]$ é encontrado a partir do epipolo calculado, e $P1^+$ é a pseudo inversa de P1. Foi obtido então a matriz fundamental F igual a:

$$F = \begin{pmatrix} 4.07 & 4.50 & 7.88 \\ -2.72 & -1.41 & -3.37 \\ 8.87 & -6.00 & -1.71 \end{pmatrix}$$

4. Solução item 4

As rotinas descritas neste capítulo foram aplicadas a todos os conjuntos de dados apresentados. Inicialmente será explicado como foi desenvolvido todo o procedimento para a obtenção dos resultados e posteriormente será apresentado e comentado todos os resultados para cada um dos conjuntos de dados.

Primeiramente, para cada conjunto de imagem foi necessário fazer a extração de pontos característicos de forma automatizada. Para tal tarefa, foi utilizado o método SIFT presente no diretório deste trabalho utilizado na lista anterior de SIFT. Dados os `key_points` e suas correspondências nas duas imagens foi utilizado o método de 8 pontos para se calcular os resultados requisitados. Como os resultados iniciais da aplicação pura do método estavam muito ruins, foi implementado o RANSAC em conjunto com o método de 8 pontos para obter a melhor matriz fundamental possível. Neste método foi minimizado a sua distância linear (minimização algébrica) de projeção de forma iterativa para estimar o melhor valor da matriz F .

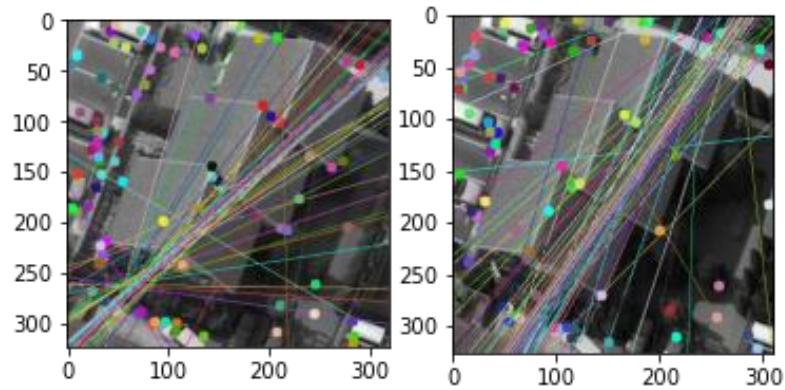
Dentre as variações do método de 8 pontos foi implementado a sua versão sem a condição de posto 2, sua versão assumindo posto 2 (removendo o menor autovalor de sua decomposição SVD) e também a sua versão normalizada com posto 2. Com esses 3 resultados foi possível observar os benefícios de fazer os ajustes para melhores resultados em detrimento de utilizar o método de 8 pontos mais simples.

Para realizar o desenho das retas epipolares em cada uma das imagens utilizou-se uma função já pronta do OpenCV que recebe os pontos da imagem, um indicador de qual das imagens está sendo utilizada e a matriz F encontrada no paço anterior.

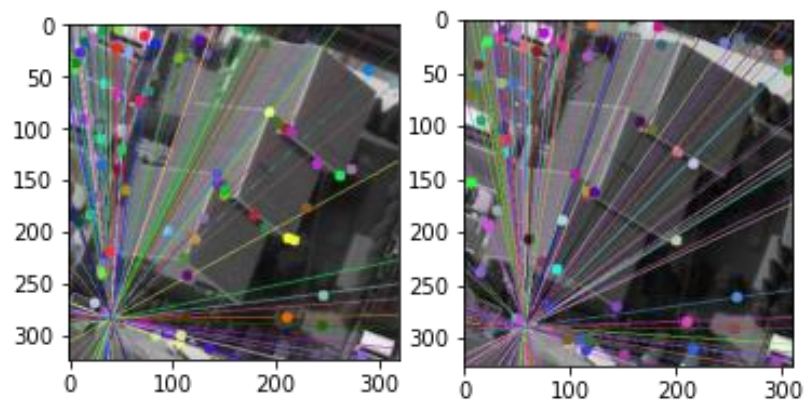
Abaixo estarão explícitos os resultados de cada um dos conjuntos de dados, em que no item 1 é exibido os resultados sem forçar condição de posto 2; no item 2 é exibido o resultado forçando posto 2; no item 3 é utilizado o método normalizado. No final de todos os conjuntos de dados é comentado a importância destes experimentos e a importância dos mesmos para os melhores resultados.

4.1 Conjunto ex-12-F

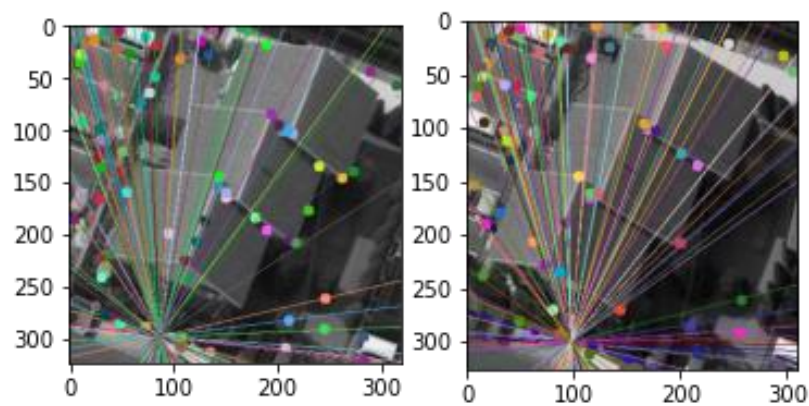
4.1.1 Método de 8 pontos sem posto 2



4.1.2 Método de 8 pontos com posto 2

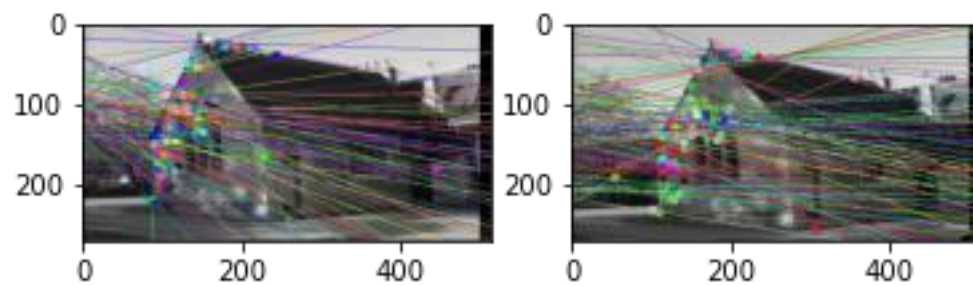


4.1.3 Método de 8 pontos normalizado

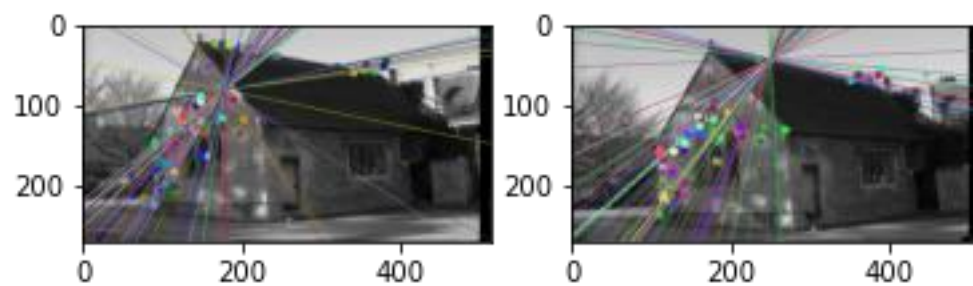


4.2 Conjunto chapel

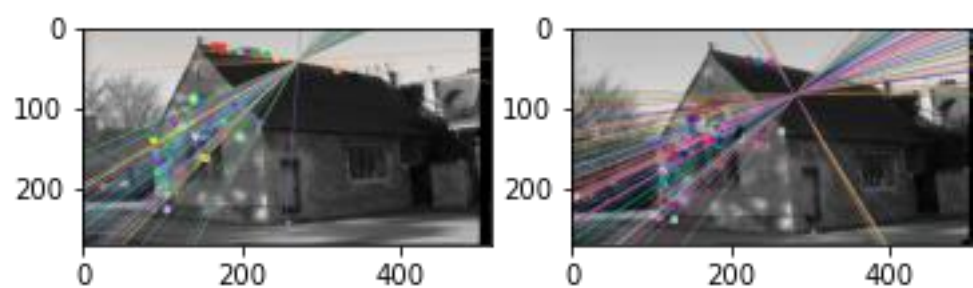
4.2.1 Método de 8 pontos sem posto 2



4.2.2 Método de 8 pontos com posto 2

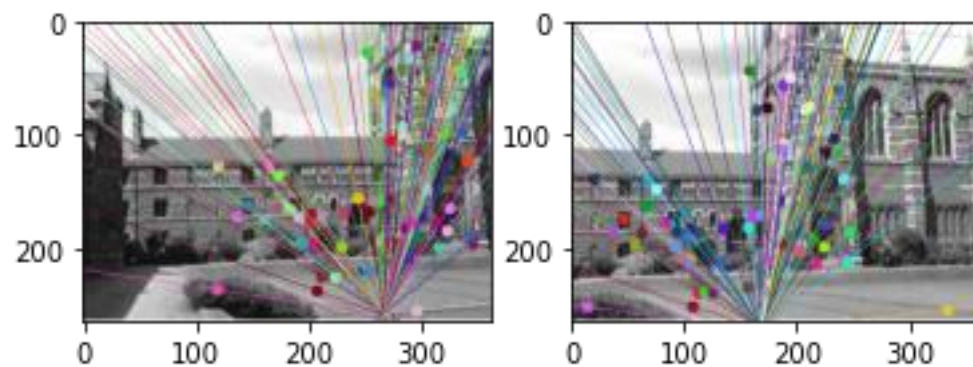


4.2.3 Método de 8 pontos normalizado

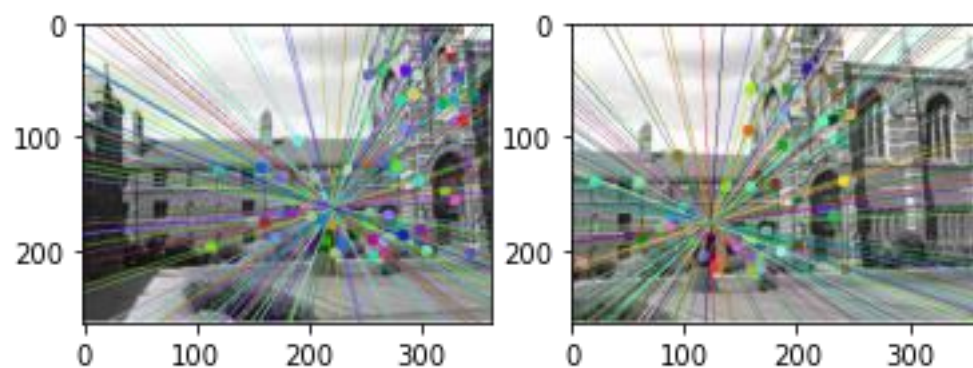


4.3 Conjunto keble

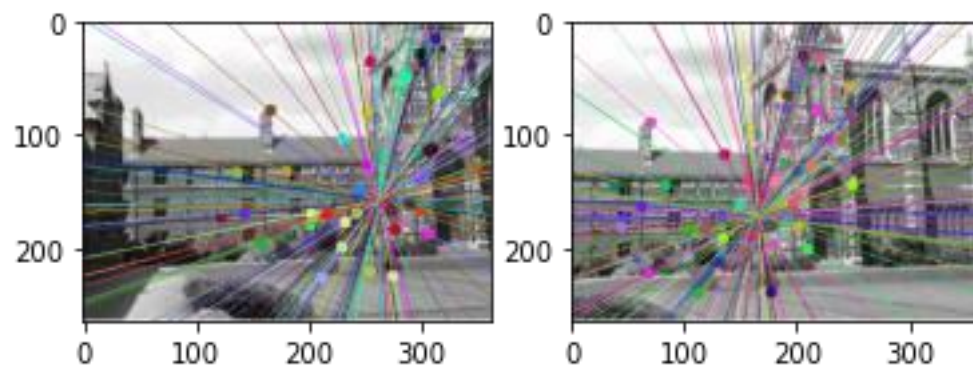
4.3.1 Método de 8 pontos sem posto 2



4.3.2 Método de 8 pontos com posto 2

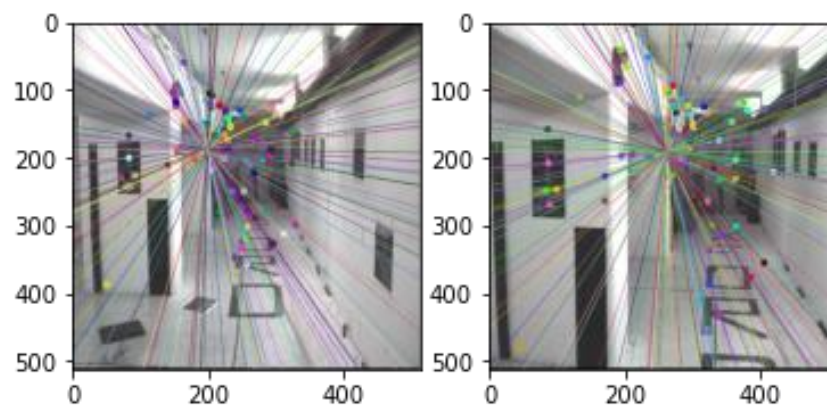


4.3.3 Método de 8 pontos normalizado

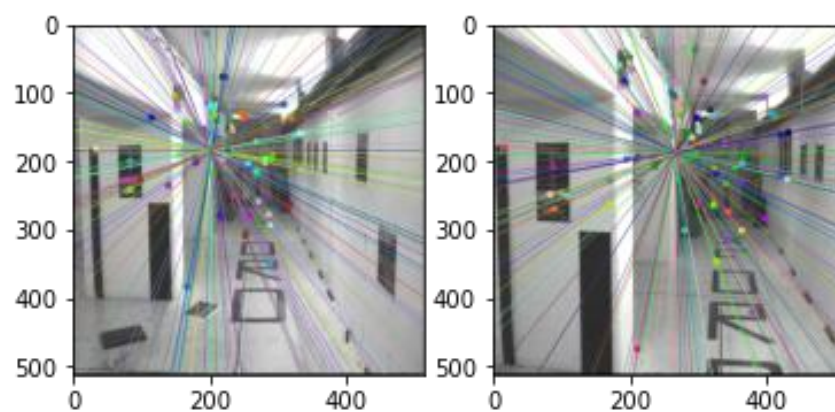


4.4 Conjunto bt00

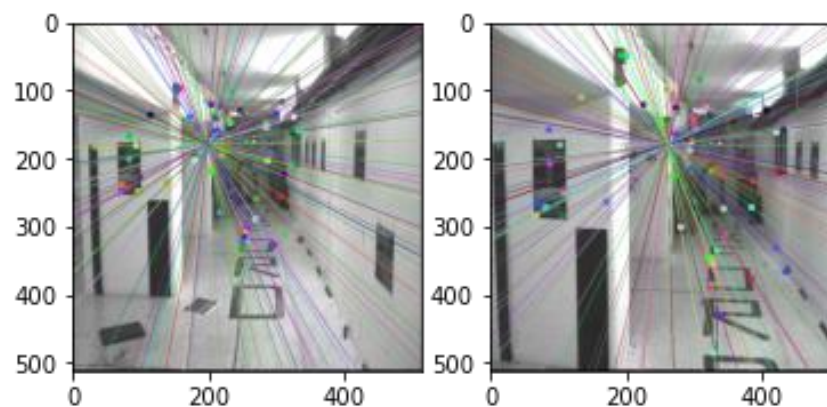
4.4.1 Método de 8 pontos sem posto 2



4.4.2 Método de 8 pontos com posto 2



4.4.3 Método de 8 pontos normalizado



4.5 Comentários

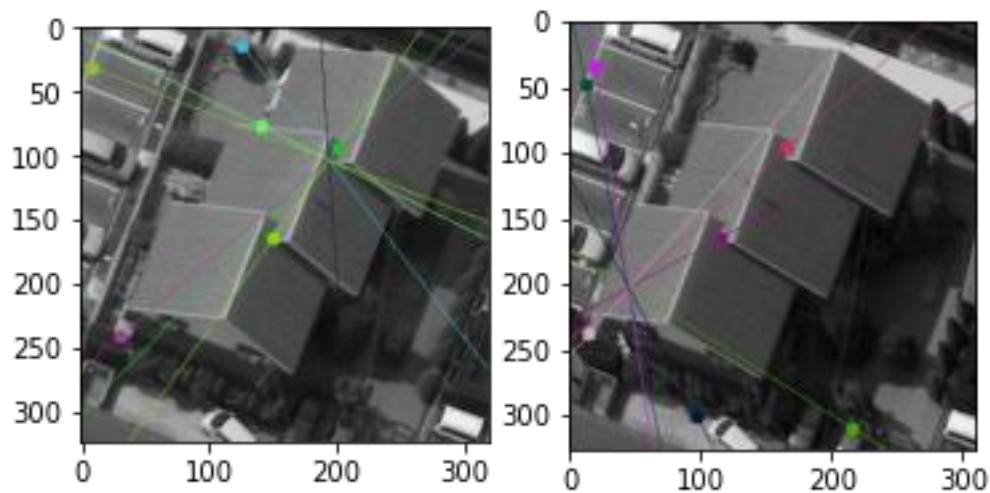
Observando os resultados observa-se 2 tipos de comportamento. O primeiro deles fica evidente a grande importância de se forçar o posto 2 para a realização deste algoritmo (conjuntos 1 e 2). E o segundo, parece que esta etapa não interferiu muito (conjuntos 3 e 4). Porém vale ressaltar que foi utilizado um tratamento com SIFT + RANSAC para obter os resultados. E mesmo depois disto existe ainda um pouco de desvio do epipolo mesmo nas imagens com bons resultados sem a condição de posto 2. Ao não utilizar posto 2 os pontos epipolares de todas as imagens não convergem para apenas 1 ponto, fazendo com que a transformação de imagens e do ajuste das retas epipolares fique muito prejudicado.

Com relação a normalização dos pontos, não fica muito evidente a sua relevância apenas observando as imagens, porém ela estabiliza a computação de F , de tal forma a melhorar a relação entre as duas imagens.

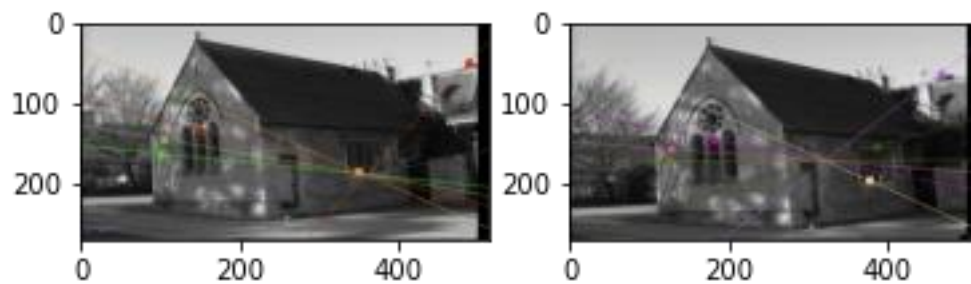
5. Solução item 5

Para cada conjunto de dados da questão anterior foi aplicado o método de 7 pontos, escolhendo, dentre os pontos obtidos pelo SIFT 7 correspondências de pontos aleatórias para este algoritmo. Observa-se que este algoritmo pode ter 2 tipos de retorno, ou ele retorna 1 matriz F ou ele retorna 3 matrizes F s. Nos resultados exibidos são apresentados as imagens correspondentes a obtenção das retas epipolares de apenas 1 das matrizes F obtidas (caso aja mais de uma).

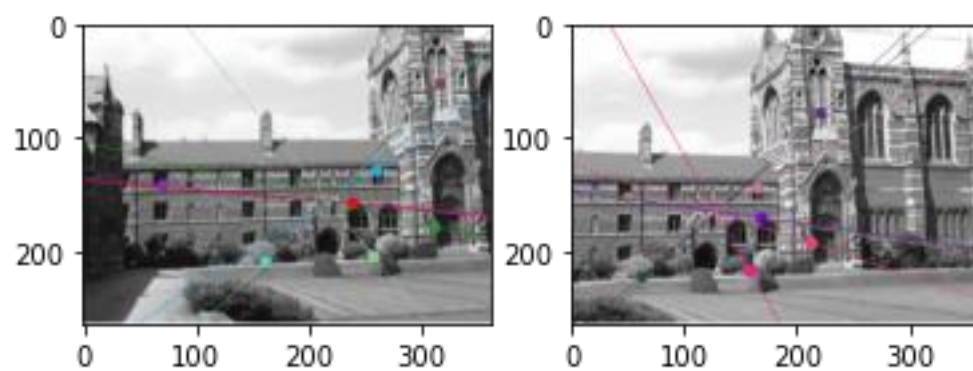
5.1 Conjunto ex-12-F



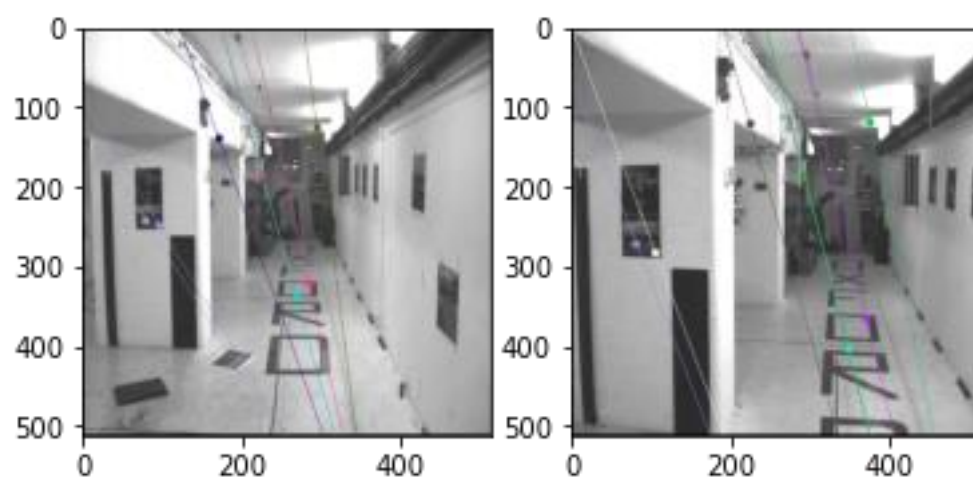
5.2 Conjunto chapel



5.3 Conjunto keble

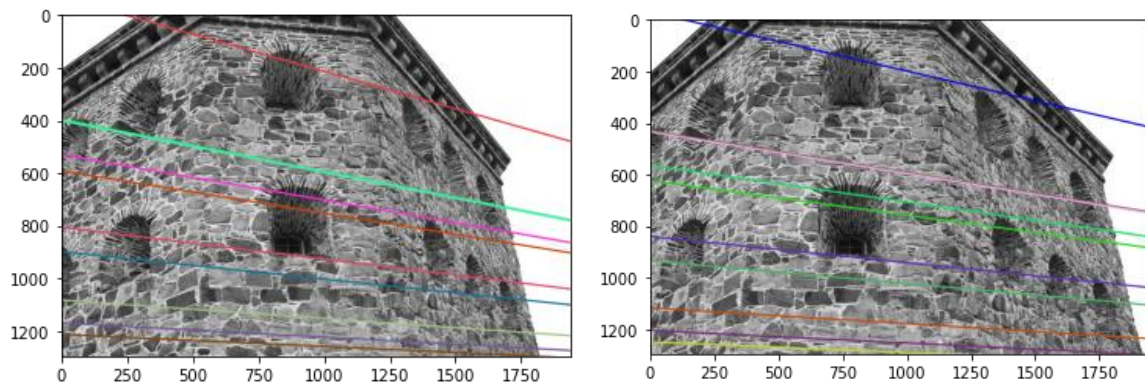


5.4 Conjunto bt00



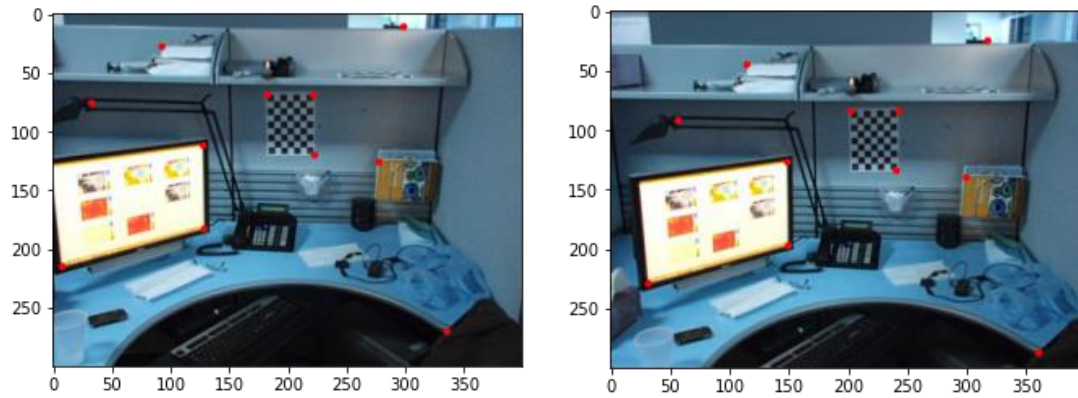
6. Solução item 6

Utilizando o conjunto de imagens kronan e os pontos de interesse já separados presentes no arquivo “kronandata.mat”. Fez-se o calculo da matriz fundamental utilizando o método “least-median” com o algoritmo de 7 pontos provido pelo openCV (foi utilizado este método para teste de outras possíveis soluções do problema). Foi exibido então 10 retas epipolares aleatórias nas imagens como pode ser observado abaixo.



7. Solução item 7

Para a realização desde item foi utilizado o conjunto de dados “mesaTrabalho” em conjunto com os pontos de interesse já destacados em matlab (“pontos_mesa_de_trabalho.mat”). Abaixo podem ser observadas as imagens e seus respectivos pontos destacados.



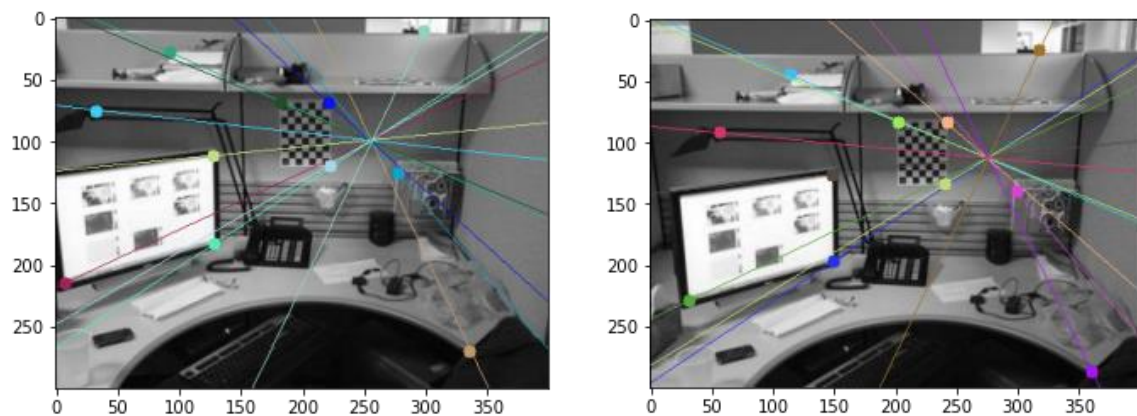
A partir dos pontos, foi feita uma normalização que resultou nas seguintes matrizes de normalização T1 e T2:

$$T1 = \begin{pmatrix} 0.01 & 0 & -1.95 \\ 0 & 0.01 & -1.29 \\ 0 & 0 & 1 \end{pmatrix}$$

$$T2 = \begin{pmatrix} 0.01 & 0 & -2.20 \\ 0 & 0.01 & -1.46 \\ 0 & 0 & 1 \end{pmatrix}$$

Utilizando o método normalizado do algoritmo de 8 pontos (utilizado no item 4 também) e utilizando apenas os 8 primeiros pontos, obteve-se a seguinte matriz fundamental F e as seguintes imagens marcando as retas epipolares:

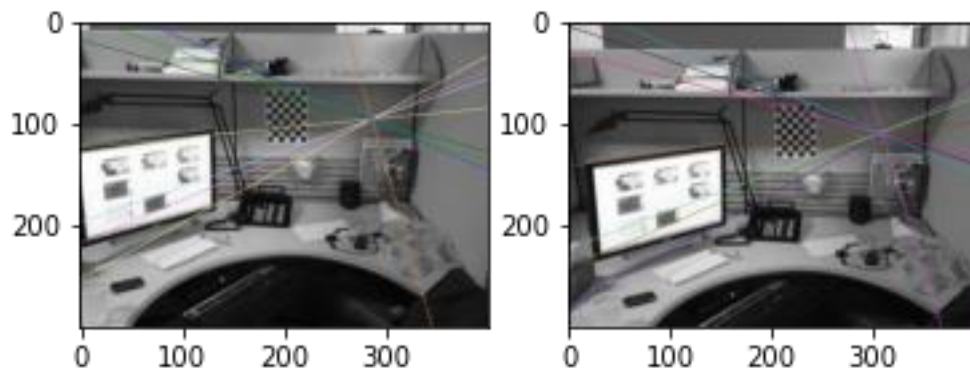
$$F = \begin{pmatrix} 4.75e-06 & -3.64e-04 & 3.51e-02 \\ 3.67e-04 & 3.40e-06 & -9.46e-02 \\ -4.29e-02 & 1.00e-01 & 1 \end{pmatrix}$$



Nota-se que o epipolo das imagens está marcado pela intercessão de todas as retas em ambas as imagens acima.

Para utilizar o método de gold standard é necessário estimar a posição espacial dos pontos de interesse aplicando o método de triangulação a partir das matrizes de câmera das duas imagens P1 e P2. Estas também podem ser calculadas a partir do ponto epipolar e de F inicialmente estimado a partir do método de 8 pontos. Ao obter o ponto espacial, aplica-se as matrizes P1 e P2 respectivamente para encontrar a distancia destes pontos estimados até a distancia dos pontos das imagens fornecidas. Com isto é utilizado o método de Levenberg–Marquardt para a minimização das distâncias.

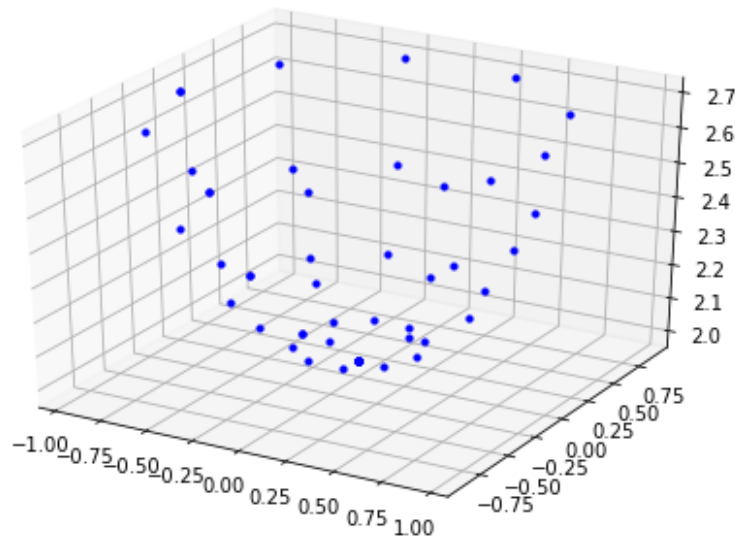
O método do gold standard não foi finalizado a tempo do prazo do trabalho, ele foi apenas iniciado e está presente na função “automatic_fundamental_8_points_gold” porém o mesmo ainda possui alguns erros. Abaixo está a execução do método normalizado com RANSAC da imagem:



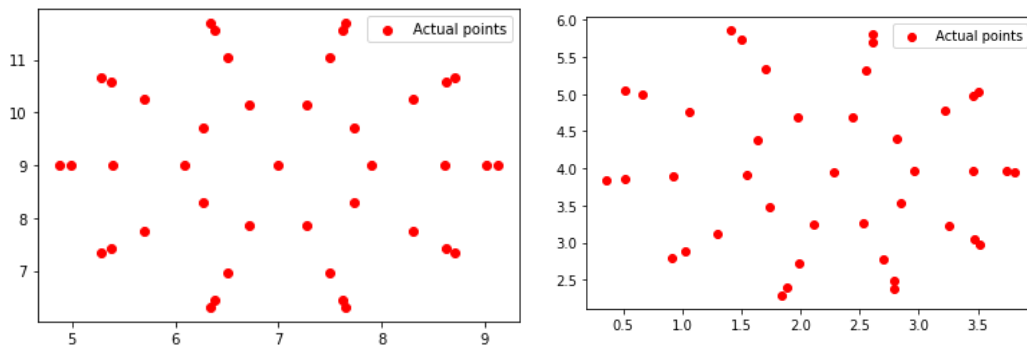
8. Solução item 8

Neste item foram utilizados os pontos 3D sintéticos presentes no arquivo “data_sphere.mat”, das matrizes K1 e K2 relativas aos pontos 2D de duas imagens projetadas e de seus respectivos pontos.

Visualizando os pontos em 3D obtemos a seguinte imagem:



Utilizando os pontos 2D fornecidos pela questão obtivemos as seguintes imagens:



A partir dos pontos 2D das imagens obteve-se a sua matriz fundamental F a partir do método “least-median” com o algoritmo de 7 pontos utilizado no item 6 desta lista. Como resultado obteve-se a seguinte matriz:

$$F = \begin{pmatrix} 0.014 & -0.014 & 0.064 \\ 0.014 & 0.009 & -0.244 \\ -0.142 & 0.017 & 1 \end{pmatrix}$$

A partir da obtenção da matriz F pode-se calcular a matriz Essencial a partir das matrizes K1, K2 e F da seguinte forma:

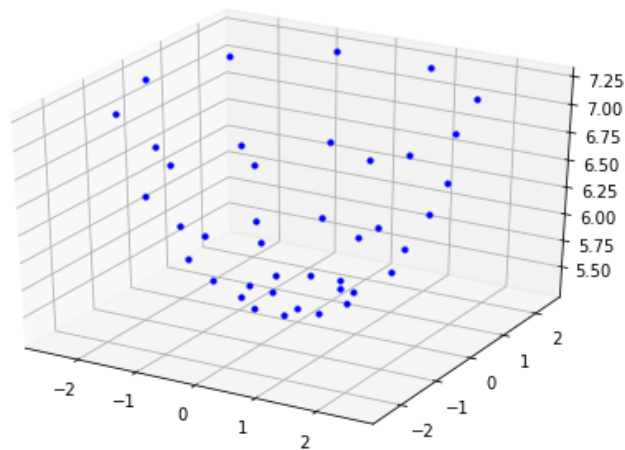
$$E = K2.T @ F @ K1$$

Obtendo a matriz E igual a:

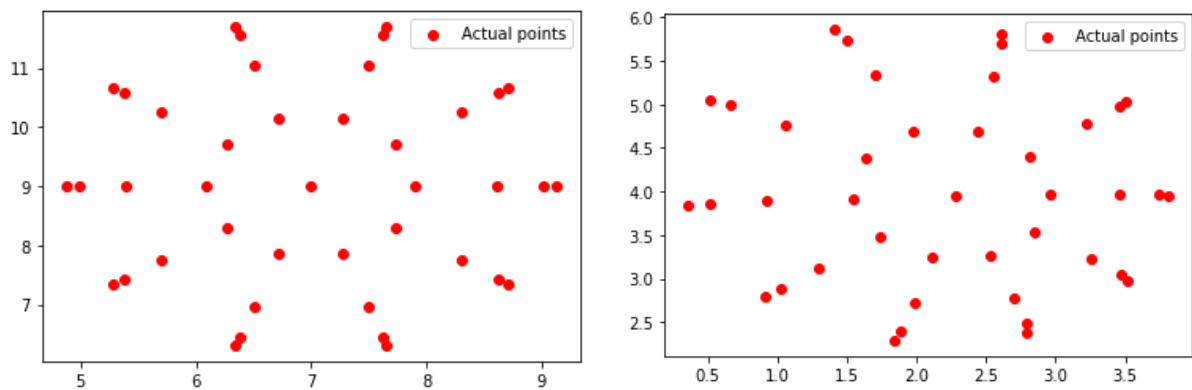
$$E = \begin{pmatrix} 0.42 & -0.56 & 0.18 \\ 0.42 & 0.36 & -0.31 \\ -0.42 & -0.05 & 0.14 \end{pmatrix}$$

Para encontrar as matrizes de câmera das duas imagens utilizou-se a função `recoverPose` do OpenCV, em que são fornecidos os pontos de ambas as imagens e a matriz Essencial obtida anteriormente. Utiliza-se uma das câmeras como central (matriz identidade 3x3 com o último elemento da diagonal 0) e a partir da matriz R_t obtida pelo `recoverPose` é possível reconstruir a matriz P_1 e P_2 .

Para fazer a reconstrução em 3D utilizou-se o método de triangulação de pontos em composição com os pontos em 2D de ambas as imagens e suas respectivas matrizes de câmera P . Como resultado encontrou-se a seguinte configuração de pontos 3D:



Observa-se que o resultado está de acordo com a imagem original fornecida, o que nos indica que o resultado está certo e condizente com o esperado. A partir da reconstrução em 3D foi aplicado as respectivas matrizes P_1 e P_2 para se obter novamente as imagens em 2D. Como resultado obteve-se as seguintes imagens:



Observa-se que as imagens recuperadas também condizem com as imagens originais fornecidas pelo problema.

