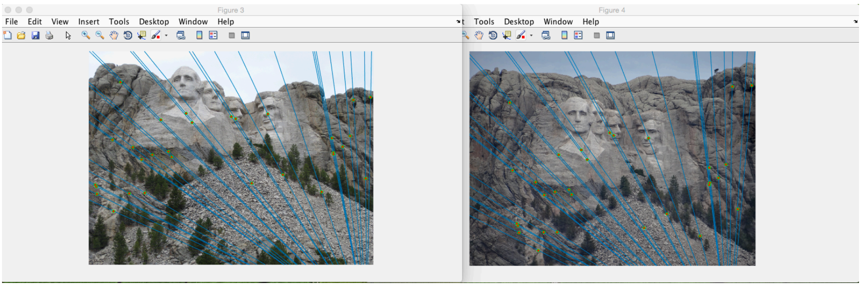
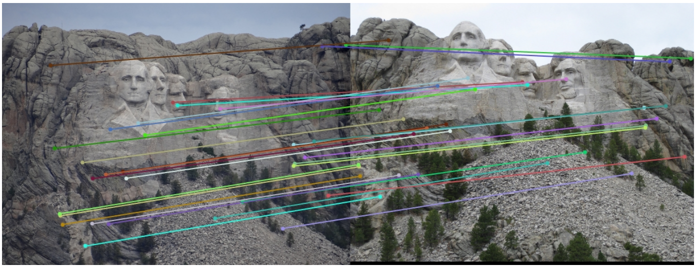
Master MLDM and 3DMT - Computer Vision course

Correction Exam March 2018 - 2h without documents

Part 1 *(5 points)*: **Stereo correspondence**

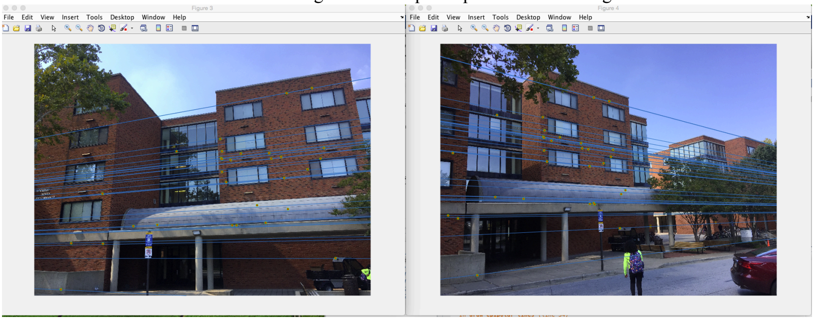


(a) left view (b) right view



(c) Examples of corresponding points

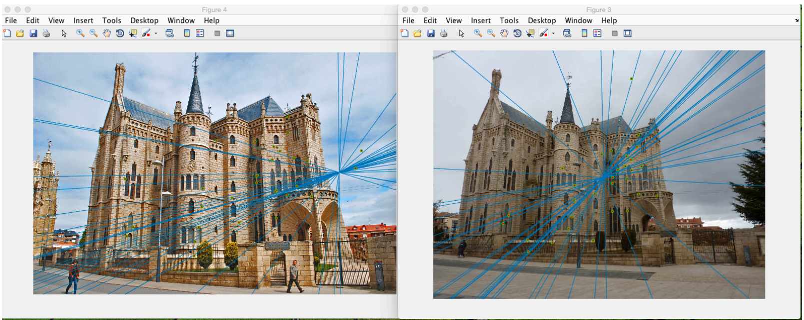
Figure 1: Example of pairs of stereo images.



(a) left view (b) right view

(c) Examples of corresponding points

Figure 2: Example of pairs of stereo images.



(a) left view (b) right view

(c) Examples of corresponding points

Figure 3: Example of pairs of stereo images.

Question 1 *(2 points)***:** In Figures 1 to 3, what do represent blue lines in images (a) and (b)? What these lines tell us?

Answer: they corresponds to the epipolar lines associated to key points. On each figure epipolar lines intersect that means that the stereo images are not rectified. The location of the intersection point inform us on the relative position/viewing direction of each camera of the stereo setup.

Question 2 *(2 points)***:** In Figures 1 to 3, what do represent color lines in images (c)? What these lines tell us?

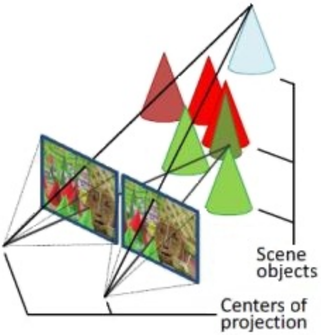
Answer: they inform us about which kind of transformation (rotation and/or translation and/or scaling) was applied to corresponding points. For example, in case of XY translation all these lines should be parallel. The disparity map between left and right views enables us to compute the depth.

Question 3 *(1 point)***:** What is the main interest of the epipolar geometry in stereo views matching?

Answer: it limits the research space (for matching task) to 1 dimension

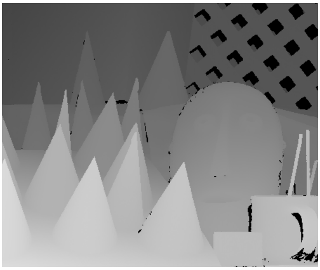
see https://www.cc.gatech.edu/classes/AY2016/cs4476_fall/results/proj3/html/adas62/index.html

Part 2 *(3 points)***: Stereo correspondence**

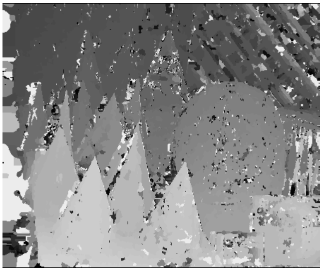


(a) Middlebury *s*tereo image “cones”

(b) left view



(c) Disparity map (ground truth)



(d) Disparity map computed from stereo matching

Figure 4: Example of pairs of stereo images.

Question 1 *(2 points)***:** In Figure 4 (c), what does represent the disparity map? Why stereo matching methods based on local spatial correlation technique cannot perform well on such image (see Fig. 4 (d)).

Answer : the disparity maps inform us about the depth of objects. Stereo matching methods based on correlation cannot perform well on repetitive patterns and textureless areas. Furthermore, when the size of the correlation window is too small the disparity map is noisy.

Question 2 *(1 point)***:** It seems in Fig 4 (b) that all cones have the same size, does this assumption is true?

Answer : The ambiguity that we have with monocular vision regarding the size/height of objects can be solved with stereo vision.
The estimation of the size of an object is inversely proportional to the depth, so it’s possible that the farthest

cones have a higher size than the closest.

Part 3 *(4 points)***: Projective, Affine, Similarity, and Isometric Transformations**

Question 1 *(2 points)***:** Classify each of the following transformations. Fill-in the circle corresponding to the most specific classification.

Answer

i. $H_1 = \begin{bmatrix} 3/4 & -1 & 0 \\ 1 & 3/4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ ○ Projective ○ Affine √ **Similarity** ○ Isometric

ii. $H_2 = \begin{bmatrix} 3/5 & -4/5 & 1 \\ 4/5 & 3/5 & -1 \\ 0 & 0 & 1 \end{bmatrix}$ ○ Projective ○ Affine ○ Similarity √ **Isometric**

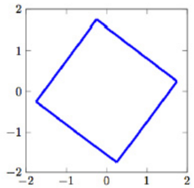
iii. $H_3 = \begin{bmatrix} 3/16 & -1 & -1/4 \\ 1/4 & 3/4 & 1/2 \\ 1/4 & 1/4 & 1 \end{bmatrix}$ √ **Projective** ○ Affine ○ Similarity ○ Isometric

iv. $H_4 = \begin{bmatrix} 3/8 & -5/8 & 0 \\ 1/2 & 5/4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ ○ Projective √ **Affine** ○ Similarity ○ Isometric

Question 2 *(2 points)***:** The figures below shows the outputs of applying one of transformations (projective, affine, similarity, and isometric) to a square with vertices at (1,1), (1,-1), (-1,-1), (-1,1). Fill in the circle corresponding to the most specific transformation used to generate each output.

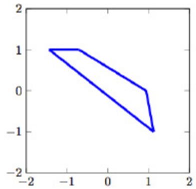
Answer

i.



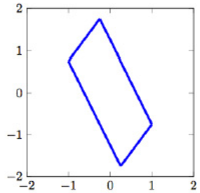
○ Projective ○ Affine √ **Similarity** ○ Isometric

ii.



√ **Projective** ○ Affine ○ Similarity ○ Isometric

iii.



○ Projective √ **Affine** ○ Similarity ○ Isometric

Part 4 *(4 points)*: **Camera model and calibration**

Suppose we want to solve for the camera matrix K and that we know that the world coordinate system is the same as the camera coordinate system. Assume the matrix K has the structure outlined below.

$$\mathbf{K} = \begin{pmatrix} K_{11} & K_{12} & K_{13} \\ 0 & K_{22} & K_{23} \\ 0 & 0 & K_{33} \end{pmatrix}$$

Note that K33 is an unknown. Assume that we are given n correspondences. Each correspondence consists of a world point (X_i; Y_i; Z_i) and its projection (u_i; v_i) for i = 1; ; n.

Question 1 *(1 point)*: What is the minimum number of correspondences needed to solve for the unknowns in the matrix K?

Answer: Given the structure of the camera matrix (K21 = K31 = K32 = 0) and the information that K33 should be treated as an unknown, we have six unknowns. Each correspondence provides two equations and thus we would need at least three correspondences.

Question 2 *(3 points)*: Set up an equation of the form Ax = 0 to solve for the unknowns in K (where A is a matrix, and x and 0 are vectors). Be specific about what the matrix A and vector x are.

Answer: Each correspondence gives us an equation of the form

$$K_{33}z_i \begin{bmatrix} u_i \\ v_i \end{bmatrix} - \begin{bmatrix} K_{11}x_i + K_{12}y_i + K_{13}z_i \\ K_{22}y_i + K_{23}z_i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and we can rearrange to obtain

$$\begin{bmatrix} x_i & y_i & z_i & 0 & 0 & -z_i u_i \\ 0 & 0 & 0 & y_i & z_i & -z_i v_i \end{bmatrix} \begin{bmatrix} K_{11} \\ K_{12} \\ K_{13} \\ K_{22} \\ K_{23} \\ K_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Thus, our linear system has the form $\mathbf{Ax} = \mathbf{0}$ with

$$\mathbf{A} = \begin{bmatrix} x_1 & y_1 & z_1 & 0 & 0 & -z_1 u_1 \\ 0 & 0 & 0 & y_1 & z_1 & -z_1 v_1 \\ x_2 & y_2 & z_2 & 0 & 0 & -z_2 u_2 \\ 0 & 0 & 0 & y_2 & z_2 & -z_2 v_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & z_n & 0 & 0 & -z_n u_n \\ 0 & 0 & 0 & y_n & z_n & -z_n v_n \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} K_{11} \\ K_{12} \\ K_{13} \\ K_{22} \\ K_{23} \\ K_{33} \end{bmatrix}.$$

Question 3 *(1 point)***:** Explain how to solve for the unknowns in the camera matrix K. Make sure K33 = 1.

Answer: We want to minimize ||Ax|| subject to ||x|| = 1. Thus, we use SVD on A and obtain the eigenvector corresponding to the smallest eigenvalue to give us a scaled version of (K11;K12;K13;K22;K23;K33). To ensure that K33 = 1, we divide each element in the eigenvector by the element corresponding to K33 (in our eigenvector). The result is a vector of the form (K11;K12;K13;K22;K23;K33 = 1).

Part 5 *(5 points)***: Epipolar geometry**

The camera projection matrices of two cameras (given in the coordinate system attached to the first camera) are

C = [I 0]   and   C' = [R t]

where R is a rotation matrix and t = (t_x; t_y; t_z) describes the translation between the cameras. Hence, the cameras have identical internal parameters and the image points are given in the normalized image coordinates (the origin of the image coordinate frame is at the principal point and the focal length is 1).

The epipolar constraint is illustrated in the Figure 5 below and it implies that if *p* and *p'* are corresponding image points then the vectors $\overrightarrow{Op}$, $\overrightarrow{O'p'}$ and $\overrightarrow{O'O}$ are coplanar.

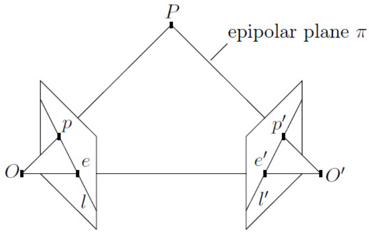


Figure 5: Epipolar geometry. Given a point *p* in the first image its corresponding point in the second image is constrained to lie on the line *l'* which is the epipolar line of *p*. Correspondingly, the line *l* is the epipolar line of *p*'. Points *e* and *e'* are the epipoles.

$\overrightarrow{O'p'} \cdot \left(\overrightarrow{O'O} \times \overrightarrow{Op}\right) = 0$ Consequently: (1)

with x represents the cross product between vectors $\overrightarrow{O'O}$ and $\overrightarrow{Op}$

and . represents the inner product between vectors $\overrightarrow{O'p'}$ and $\overrightarrow{O'O} \times \overrightarrow{Op}$

Let $p = (x; y; 1)^T$ and $p' = (x'; y'; 1)^T$ denote the homogeneous image coordinate vectors of *p* and *p'*.

Questions 1 *(2 points)***:** Show that the equation (1) can be written in the form:

$$p'^T E\, p = 0 \qquad (2)$$

where matrix E is the essential matrix defined by: $E = [t]\, R$

Hints:

In the coordinate system of the second camera we have: $\overrightarrow{O'p'} = \mathbf{p}', \quad \overrightarrow{O'O} = \mathbf{t}, \quad \overrightarrow{Op} = \mathbf{Rp}$

The cross product b x a, of two arbitrary 3D vectors $a = (i; j; k)^T$ and $b = (l; m; n)^T$, can be written as:

$$a \times b = \begin{matrix} 0 & -k & j \\ k & 0 & -i \\ -j & i & 0 \end{matrix} \quad \begin{matrix} l \\ m \\ n \end{matrix} = [a] \,.\, b$$

where [a] is a 3x3 skew matrix associated to the vector a.

Answer: by substituting $\overrightarrow{O'p'} = \mathbf{p}', \quad \overrightarrow{O'O} = \mathbf{t}, \quad \overrightarrow{Op} = \mathbf{Rp}$ into (1) we get: $p'^T (t \times Rp) = 0$

$$\Rightarrow \underset{1x3}{p'^T} (\underset{3x3}{[t]}\ \underset{3x3}{R}\ \underset{3x1}{p}) = 0 \quad = 1x1$$

$$\Rightarrow \underset{1x3}{p'^T} (\underset{3x3}{[t]\, R})\, \underset{3x1}{p} = 0 \quad = 1x1$$

$\Rightarrow p'^T E p = 0$

Question 2 *(1 point)***:** How can the epipolar constraint be utilized when searching point correspondences between two views?

Answer: Given a point in the first image and the epipolar geometry one may directly compute the corresponding line in the second image. Hence, it is sufficient to search the corresponding point in the neighborhood of this line and this significantly reduces the search area.

Question 3 *(2 points)***:** Let

$$R = \begin{matrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{matrix} \quad \text{and } t = \begin{matrix} 1 \\ 1 \\ 1 \end{matrix}$$

be the rotation and translation between two views. Compute the essential matrix and the epipolar line which corresponds to the principal point of the first image (i.e. the point (0; 0) in the normalized image coordinate system).

Answer:

$$[\mathbf{t}]_\times = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}, \qquad \mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

and using the formula from the previous problem we get

$$\mathbf{E} = [\mathbf{t}]_\times \mathbf{R} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & -1 & 0 \\ -1 & 0 & -1 \end{bmatrix}$$

The epipolar line $\mathbf{l}'$ corresponding to the point $(0, 0)$ is

$$\mathbf{l}' = \mathbf{E} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & -1 & 0 \\ -1 & 0 & -1 \end{bmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

Part 6 *(4 points – 15 mn)***: Image Transformations**

Figure 6 : source ; image 1 image 2 image 3 image 4

Question 1 *(2 points)***:** The figures above show the outputs of applying one of the following transformations to Mona Lisa. Fill in the circle corresponding to the transformation used to generate each output.

Image 1	$\begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} S & 0 & 0 \\ 0 & S & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix}$
Image 2	$\begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} S & 0 & 0 \\ 0 & S & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix}$
Image 3	$\begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} S & 0 & 0 \\ 0 & S & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix}$
Image 4	$\begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} S & 0 & 0 \\ 0 & S & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix}$

Question 2 *(2 points)***:** For each of images shown above how many degrees of freedom the corresponding transformation has? Fill in the circle corresponding to the right answer.

Image 1	1 DOF	2 DOF	4 DOF	6 DOF	8 DOF	9 DOF
Image 2	1 DOF	2 DOF	4 DOF	6 DOF	8 DOF	9 DOF
Image 3	1 DOF	2 DOF	4 DOF	6 DOF	8 DOF	9 DOF

Image 4 1 DOF 2 DOF 4 DOF 6 DOF 8 DOF 9 DOF