3D Computer Vision - Final exam
(duration: 2h)

P. Monasse

November 16, 2021

The exercises are independent. You can choose to answer in French or English, at your convenience.

1 Transfer of epipolar lines

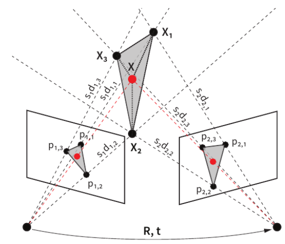
We consider a stereo image pair linked by the epipolar constraint $x^\top F x' = 0$ with epipoles e_L and e_R: $F^\top e_L = F e_R = 0$.

1. Remind the definition of an epipolar line.
2. Explain geometrically why for any point x on a given left epipolar line ℓ_L, its homologous point is on a right epipolar line ℓ_R independent on the choice of x. This shows there is a function H mapping left epipolar lines ℓ_L to their corresponding right epipolar line ℓ_R. We show in the next questions that H is not unique, but that most are homographies:
$$H(v) = F^\top [e_L]_\times + v e_L^\top, \tag{1}$$
with $v \in \mathbb{R}^3$.
3. Show that $x = e_L \times \ell_L$ is a point on epipolar line ℓ_L. Using this point, show that for any v in (1), $H(v)\ell_L = \ell_R$.
4. Let H_0 be a homography mapping left epipolar lines to the corresponding right epipolar lines. Show that there is some scalar λ such that
$$\lambda F^\top = H_0 [e_L]_\times. \tag{2}$$
(Consider the mapping of any $x \in \mathbb{R}^3$ through each member of this equation)
5. Show then that up to scale
$$H_0 = H(-\frac{1}{\lambda} H_o e_L). \tag{3}$$
The remaining questions prove that $H(v)$ is a homography if and only if $v^\top e_R \neq 0$.
6. Consider two independent vectors v_1 and v_2 orthogonal to e_L. Show that $e_L \times v_1$ and $e_L \times v_2$ are two distinct left epipolar lines.
7. Deduce that $H(v)v_1$ and $H(v)v_2$ are two distinct right epipolar lines.
8. Show that $H(v)v_1 \times H(v)v_2 = e_R$ up to scale.
9. Show that $H(v)e_L$ is in the span of vectors $H(v)v_1$ and $H(v)v_2$ if and only if $v^T e_R = 0$.
10. Show that $H(v)$ is a homography if and only if $(H(v)e_L, H(v)v_1, H(v)v_2)$ is a basis.
11. Conclude.

2 Pose estimation with relative depth priors

Some ~~magical~~ deep-learning based methods can achieve the impossible: infer the depth of pixels in a single image, up to an unknown global scale s.

Stereo pair with estimated relative depth maps.



Homography in Question 18.

12. Writing $p_1 = (x_1, y_1, 1)^\top$ and $p_2 = (x_2, y_2, 1)^\top$ matching points in a stereo pair, with relative depth d_1 and d_2 (up to scale factors s_1 and s_2), explain why they are linked through

$$s_2 d_2 p_2 = s_1 d_1 R p_1 + t, \tag{4}$$

with R and t the relative rotation and translation between the two views.

13. Write 3 linear equations satisified by the following 13 coefficients of a vector x composed of: 9 unknowns from $s_1 R$, 3 from t and s_2.

14. Assuming 4 or more matching pairs (p_1, p_2) are known, what is the procedure to recover R, t, s_1 and s_2?

15. Show that we can write six quadratic polynomial constraint equations involving the coefficients of a 3×3 matrix to be of the form $s_1 R$.

16. Assuming only two matching pairs (p_1, p_2) are given, explain how x can be written as a linear combination of 7 orthonormal vectors that we can compute. Show that in theory we are able to recover x up to some scale.[1]

17. Assume that we have only relative depth prior for the first image, not the second one, and that three matching pairs (p_1, p_2) are available.
    (a) For each pair, write two linear equations involving y, the vector composed of the 12 first coefficients of x (excluding s_2).
    (b) Explain how y can be written as a linear combination of 6 orthonormal vectors that we can compute.
    (c) Show that in theory we are able to recover y up to some scale.[2]

18. Suppose we know three pairs (p_1^i, p_2^i) with their relative depths d_1^i and d_2^i.
    (a) Justify that points on the plane spanned by $\{p_1^i\}$ are related to points on the plane spanned by $\{p_2^i\}$ through a homography H.
    (b) Show the centroid of $\{p_1^i\}$ is mapped by H to the centroid of $\{p_2^i\}$.
    (c) Deduce that the homography H can be computed.

19. For each of the above situations, propose an algorithm to recover (R, t) or H from the stereo pair with the estimated relative depths. How is the relative depth useful and how can it lead to faster estimation?

[1] More precisely, x is among a set of up to 16 possible solutions, up to scale.
[2] In this case, there are up to 8 solutions.