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

P. Monasse and R. Marlet

November 13th, 2012

You can choose to answer in French or English, at your convenience.

1 Independent questions close to course ($\sim 25\%$)

1. In $\mathbb{R}^n$, we define the binary relation $\mathcal{R}$ by:

$$x\mathcal{R}y \Leftrightarrow \exists\lambda \neq 0 : x = \lambda y$$

   Show that $\mathcal{R}$ is an equivalence relation (reflexive, symmetric, transitive). Does it remain true if we replace the condition $\lambda \neq 0$ by $\lambda > 0$?

2. Why do we exclude the origin O in the definition of projective space? Why the indice n in $\mathcal{P}^n$ whereas we take a quotient of $\mathbb{R}^{n+1}$?

$$\mathcal{P}^n = (\mathbb{R}^{n+1} \setminus \{0\})/\mathcal{R}$$

3. Let $y \in \mathcal{R}^3$. Show that $[y]_\times [y]_\times = yy^T - (y^T y)Id$.

4. Prove the formula for fundamental matrix, and describe the different components of this formula:

$$F = K'^{-T}[T]_\times RK^{-1}$$

5. Describe the RANSAC algorithm applied to the estimation of a line in the plane approximating most points (x_i, y_i), $i = 1, \dots, n$, some of these points being outliers.

6. How to minimize $\|Ax\|$ under constraint $\|x\| = 1$ with A an $m \times n$ matrix and x a vector of length n ($\|.\|$ means Euclidean norm)? How can it be used to compute the fundamental matrix from n correspondences?

2 Errors for homography ($\sim 35\%$)

The two parts of this exercise are independent.

2.1 Sampson error

1. Given corresponding points $X = \begin{pmatrix} x & y & 1 \end{pmatrix}^T$ and $X' = \begin{pmatrix} x' & y' & 1 \end{pmatrix}^T$ linked by a homography H ($X' = \lambda H X$), recall the expression of the algebraic error.

2. Writing the algebraic error as a function $f(x, y, x', y')$ (H fixed), recall the Sampson error, using the Jacobian matrix J of f.

3. Compute J at a point (x, y, x', y').

4. Show that this Sampson error can be expressed by a closed form formula.

2.2 Gold standard error

5. The "gold standard" error is:

$$\epsilon = \min_{\hat{x}, \hat{y}} d(X, \hat{X})^2 + d(X', H\hat{X})^2$$

where $\hat{X} = \left(\hat{x}, \hat{y}, 1\right)^T$ and d is the Euclidean distance between points represented in homogeneous coordinates. Explain its geometric signification (you can draw a figure).

6. We write $e(\hat{X})$ the term to minimize. Show that it is written:

$$e(\hat{X}) = (x - \hat{x})^2 + (y - \hat{y})^2 + \left(x' - \frac{L_1^T \hat{X}}{L_3^T \hat{X}}\right)^2 + \left(y' - \frac{L_2^T \hat{X}}{L_3^T \hat{X}}\right)^2$$

with L_i vectors depending on H.

7. Show that by an appropriate rotation of the first image, we do not lose any generality assuming that $L_3 = \begin{pmatrix} 0 & L_{23} & L_{33} \end{pmatrix}^T$. Why does it not change the error ϵ?

8. In this case, show that writing $\frac{\partial e}{\partial \hat{x}} = 0$ permits to express $\hat{x}$ in function of $\hat{y}$.

9. Reporting into the equation $\frac{\partial e}{\partial \hat{y}} = 0$, show that it involves that $\hat{y}$ is a root of a polynomial of degree at most 8 (whose exact expression is not asked).

10. Summarizing the above steps, explain the algorithm to compute ϵ numerically.

11. Show that the minimum above is reached (the infimum is indeed a minimum).

12. In the particular case where H is an affine function, show that there is a closed form expression of ϵ.

3 Disparity map estimation with graph cuts ($\sim$ 40%)

Context. Given two rectified images I (left) and I' (right), with respective sets of pixels coordinates P and P', a (left-to-right) disparity map d of I with respect to I' is a function $d : P \to \mathbb{R}$ indicating that a pixel p in I corresponds to pixel $p' = p + (d(p), 0)$ in I'.

We have seen in the course how a discretization $f : P \to \{d_{\min}, \dots, d_{\max}\}$ of d can be seen and computed as a pixel labeling that minimizes an energy function $E(f)$ defined as:

$$E(f) = \underbrace{\sum_{p \in P} D_p(f_p)}_{E_{\text{data}}(f)} + \underbrace{\sum_{(p,q) \in N} V_{p,q}.\mathbf{1}(f_p \neq f_q)}_{E_{\text{smooth}}(f)}$$

where N is a set of neighboring pixel pairs in I.

We have also seen a formulation by Komogorov and Zabih (2001) that takes into account occlusion in the energy as an extra term:

$$E(f) = E_{\text{data}}(f) + E_{\text{smooth}}(f) + E_{\text{occ}}(f)$$

We are going to study here an alternate formulation for taking occlusion into account. We decompose it in a few steps.

Note 1. Do not get stuck at question k; if needed, jump to question $k + 1$ assuming question k is solved.

Note 2. Most questions can be answered in 1 to 4 lines. They require reflexion, not computations. There is no trap.

1. We assume that any pixel p in one image has at most one corresponding pixel p' in the other image. In this context, given an estimated disparity f, if there are two pixels p, q in a same line of I that correspond to the same pixel in I', it means that at least of them is occluded. Define a *simple* function $\text{occ}_p(f_p, f)$ that returns 0 if a pixel p estimated at disparity f_p is possibly occluded according to f, and that returns 1 otherwise.

2. We do not want occluded pixels to influence the energy. Using function $\text{occ}_p(f_p, f)$, define a *simple* new data term $E'_{\text{data}}(f)$ such that pixels that are possibly occluded do not contribute to the energy.

3. What condition for optimizing $E'(f) = E'_{\text{data}}(f) + E_{\text{smooth}}(f)$ using graph cuts is not satisfied and why?

4. This implies that we cannot use graph cut techniques directly. But we can do it indirectly. Suppose that some oracle gives you a sparse reliable disparity map $(r_p)_{p \in P}$, i.e., a function $r : P \to \{\perp, d_{\min}, \dots, d_{\max}\}$ where

$\perp$ indicates that the disparity is undefined. In other words, if $r_p = d \neq \perp$ then the defined disparity d for pixel p is likely to be correct, and if $r_p = \perp$ then the disparity for pixel p is undefined. Define a *simple* new data term $E''_{\text{data}}(f)$ such that:

- pixels p that have a defined reliable disparity according to r tend to remain unchanged in an optimal f,
- pixels p that do not have a defined reliable disparity according to r are assigned in f a disparity in accordance to their visibility condition $\text{occ}_p(f_p, r)$ with respect to r.

5. Still assuming that r is given, show that $E''(f) = E''_{\text{data}}(f) + E_{\text{smooth}}(f)$ satisfies the conditions for being optimized using graph cut techniques.

6. Given an estimated left-to-right disparity map f and an estimated right-to-left disparity map f', recall how the cross-checking technique presented in the course can be used to identify unreliable disparities in f and define a corresponding sparse reliable disparity map.

7. Using the above "building blocks", design an iterative algorithm that computes a good disparity map f. The idea is to use a sparse reliable disparity map estimated from the previous iteration. At first iteration, the sparse reliable disparity map is totally undefined.

8. Explain why this algorithm converges.

Together with an extra filter based on color for discarding extra unreliable disparities, this methods produces quite good results...