Computer Vision Minor 1

Singhal, Madhur

2015CS10235

September 13, 2017

1 Declaration of Originality

This is to certify that to the best of my knowledge, the content of this document is my own work. I worked on this in the Vision lab and had discussions with Aman, Suyash, Ankesh, Makkunda, Ankesh, Krunal, Pranjal sir and Saket. I also consulted the paper 'Surface Orientation and Time to Contact from Image Divergence and Deformation'.

2 Question 5 - Optical flow Approximation

We assume that the Optical flow can be given by the following equation, ignoring the higher order terms.

$$\begin{bmatrix} u \\ v \end{bmatrix} \approx \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} + \begin{bmatrix} u_x & u_y \\ v_x & v_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

This expression assumes affine motion of the object (or the camera). We now use the standard optical flow equation and substitute the approximation for u and v.

$$uI_x + vI_y + I_t = 0$$

$$(u_0 + u_x x + v_y y) I_x + (v_0 + v_x x + v_y y) I_y + I_t = 0$$

$$\begin{bmatrix} I_x & xI_x & yI_y & I_y & xI_y & yI_y \end{bmatrix} \begin{bmatrix} u_0 \\ u_x \\ u_y \\ v_0 \\ v_x \\ v_y \end{bmatrix} = -I_t$$

Now, we assume that that the velocities u_0, v_0 and the gradients u_x, u_y, v_x and v_y are constant over a 5×5 patch of the image around any point (x, y). Then the above equation is valid for all those points and we can collect all the 25 equations formed into a matrix.

$$\begin{bmatrix} I_{x_1} & x_1 I_{x_1} & y_1 I_{y_1} & I_{y_1} & x_1 I_{y_1} & y_1 I_{y_1} \\ I_{x_2} & x_2 I_{x_2} & y_2 I_{y_2} & I_{y_2} & x_2 I_{y_2} & y_2 I_{y_2} \\ & & \text{.......} & & & \end{bmatrix} \begin{bmatrix} u_0 \\ u_x \\ u_y \\ v_0 \\ v_x \\ v_y \end{bmatrix} = - \begin{bmatrix} I_{t_1} \\ I_{t_2} \\ \text{...} \end{bmatrix}$$

We can write the above equation as

$$AX = B$$

Since the row rank is larger than the number of columns, we will use the pseudo-inverse to find the solution that minimizes the least squares error. This is because the least square solution is the projection of B onto the space which is orthogonal to the range space of A, but the null space of $A^\top$ is the space orthogonal to range space of A. Thus the pseudo-inverse is just the matrix which projects B onto $A^\top$.

Thus the solution is given by

$$X = \left(A^\top A \right)^{-1} A^\top B$$

3   Question 6 - Pinhole camera Optical Flow

Let us first express the velocity of a 3D point in terms of the given parameters (assumed constant).

$$V = -U - \Omega \times P$$

$$
\begin{bmatrix} \dfrac{\partial X}{\partial t} \\[2mm] \dfrac{\partial Y}{\partial t} \\[2mm] \dfrac{\partial Z}{\partial t} \end{bmatrix} = - \begin{bmatrix} U_1 \\ U_2 \\ U_3 \end{bmatrix} - \begin{bmatrix} Z\Omega_2 - Y\Omega_3 \\ X\Omega_3 - Z\Omega_1 \\ Y\Omega_1 - X\Omega_2 \end{bmatrix}
$$

Now we express the 2D image coordinates assuming the pinhole camera model.

$$
x = \frac{fX}{Z}
$$

$$
y = \frac{fY}{Z}
$$

Differentiate the 2D image coordinates with respect to time.

$$
u = \frac{fZ\dfrac{\partial X}{\partial t} - fX\dfrac{\partial Z}{\partial t}}{Z^2}
$$

$$
v = \frac{fZ\dfrac{\partial Y}{\partial t} - fY\dfrac{\partial Z}{\partial t}}{Z^2}
$$

Now substitute the expression for derivatives of a 3D coordinate.

$$
u = \frac{fZ\left(-U_1 - Z\Omega_2 + Y\Omega_3\right) - fX\left(-U_3 - Y\Omega_1 + X\Omega_2\right)}{Z^2}
$$

$$
v = \frac{fZ\left(-U_2 - X\Omega_3 + Z\Omega_1\right) - fY\left(-U_3 - Y\Omega_1 + X\Omega_2\right)}{Z^2}
$$

Simplifying and substituting the pinhole model expression into the above equation we will get the desired result.

$$
u = -\frac{fU_1}{Z} - f\Omega_2 + y\Omega_3 - x\left(-\frac{U_3}{Z} - \frac{y\Omega_1}{f} + \frac{x\Omega_2}{f}\right)
$$

$$
v = -\frac{fU_2}{Z} - x\Omega_3 + f\Omega_1 - y\left(-\frac{U_3}{Z} - \frac{y\Omega_1}{f} + \frac{x\Omega_2}{f}\right)
$$

4 Question 7 - Image velocity gradients

From the expressions for u and v from the above question we neglect the second order terms to get the following simplified expressions.

$$u = -\frac{fU_1}{Z} - f\Omega_2 + y\Omega_3 + x\frac{U_3}{Z}$$

$$v = -\frac{fU_2}{Z} - x\Omega_3 + f\Omega_1 + y\frac{U_3}{Z}$$

$$u_x = \frac{fU_1 Z_x}{Z^2} - \frac{xU_3 Z_x}{Z^2}^{0} + \frac{U_3}{Z} + \Omega_3\frac{\partial y}{\partial x}^{0}$$

Here y is independent of x so the last term is zero and since x and Z_x are very less compared to Z^2 the second term can be neglected.

$$u_x = \frac{fU_1 Z_x}{Z^2} + \frac{U_3}{Z}$$

$$u_y = \frac{fU_1 Z_y}{Z^2} + \Omega_3 - \frac{xU_3 Z_y}{Z^2}^{0} = \frac{fU_1 Z_y}{Z^2} + \Omega_3$$

$$v_x = \frac{fU_2 Z_x}{Z^2} - \Omega_3 - \frac{yU_3 Z_x}{Z^2}^{0} = \frac{fU_2 Z_x}{Z^2} - \Omega_3$$

$$v_y = \frac{fU_2 Z_y}{Z^2} - \Omega_3\frac{\partial x}{\partial y}^{0} + \frac{U_3}{Z} - \frac{yU_3 Z_y}{Z^2}^{0} = \frac{fU_2 Z_y}{Z^2} + \frac{U_3}{Z}$$

5 Question 8- Gradient decomposition

The decomposition of the velocity tensor gradient is given as the following.

$$\begin{bmatrix} u_x & u_y \\ v_x & v_y \end{bmatrix} = \frac{div(v)}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{curl(v)}{2}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} + \frac{def(v)}{2}\begin{bmatrix} \cos(2\mu) & \sin(2\mu) \\ \sin(2\mu) & -\cos(2\mu) \end{bmatrix}$$

Simplifying we get four equations.

$$\begin{bmatrix} u_x & u_y \\ v_x & v_y \end{bmatrix} = \begin{bmatrix} \dfrac{div(v)}{2} + \dfrac{def(v)\cos(2\mu)}{2} & -\dfrac{curl(v)}{2} + \dfrac{def(v)\sin(2\mu)}{2} \\ \dfrac{curl(v)}{2} + \dfrac{def(v)\sin(2\mu)}{2} & \dfrac{div(v)}{2} - \dfrac{def(v)\cos(2\mu)}{2} \end{bmatrix}$$

4

Now we can show all of the desired relations by looking at the appropriate linear combinations of these four equations.

$$u_x + v_y = \frac{div(v)}{2} + \frac{def(v)\cos(2\mu)}{2} + \frac{div(v)}{2} - \frac{def(v)\cos(2\mu)}{2} = div(v)$$

$$-(u_y - v_x) = v_x - u_y = \frac{curl(v)}{2} + \frac{def(v)\sin(2\mu)}{2} - \left(-\frac{curl(v)}{2} + \frac{def(v)\sin(2\mu)}{2} \right) = curl(v)$$

$$u_x - v_y = \frac{div(v)}{2} + \frac{def(v)\cos(2\mu)}{2} - \left(\frac{div(v)}{2} - \frac{def(v)\cos(2\mu)}{2} \right) = def(v)\cos(2\mu)$$

$$u_y + v_x = -\frac{curl(v)}{2} + \frac{def(v)\sin(2\mu)}{2} + \frac{curl(v)}{2} + \frac{def(v)\sin(2\mu)}{2} = def(v)\sin(2\mu)$$

Hence all the relations given have been verified.

6 Question 9 - Depth and Orientation from Optical Flow

We are given that A is the translation velocity of object parallel to image plane scale by depth and F is the depth gradient scaled by Z.

$$A = \frac{U - (U \bullet Q)Q}{Z} \quad \text{and} \quad F = \frac{f\nabla Z}{Z}$$

Since we have assumed the pinhole camera model with the camera center at origin and the camera direction as the z-axis. So the camera direction is given by the following expression. This is needed since otherwise our earlier results would not be valid.

$$Q = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Now we use our results from previous questions to find the expressions needed

$$div(v) = u_x + v_y = \frac{U_3}{Z} + \frac{U_1 Z_x f}{Z^2} + \frac{U_3}{Z} + \frac{U_2 Z_y f}{Z^2}$$

$$= \frac{2U_3}{Z} + \frac{f}{Z} \left(\frac{U_1 Z_x}{Z} + \frac{U_3 Z_y}{Z} \right)$$

5

$$= \frac{2}{Z} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \end{bmatrix} \bullet \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + \frac{f}{Z} \begin{bmatrix} Z_x \\ Z_y \\ Z_z \end{bmatrix} \bullet \frac{1}{Z} \begin{bmatrix} U_1 \\ U_2 \\ 0 \end{bmatrix}$$

$$= \frac{2U \bullet Q}{Z} + F \bullet A$$

For the curl part we proceed from the opposite side.

$$|F \times A| - 2\Omega \bullet Q = | \begin{bmatrix} Z_x \\ Z_y \\ 0 \end{bmatrix} \times \frac{f}{Z^2} \begin{bmatrix} U_1 \\ U_2 \\ 0 \end{bmatrix} | - 2\Omega_3$$

$$= \left(-\Omega_3 + \frac{fU_2 Z_x}{Z^2} \right) - \left(\frac{fU_1 Z_y}{Z^2} + \Omega_3 \right)$$

$$= v_x - u_y = -(u_y - v_x) = curl(v)$$

Note that Z_z is zero since our image plane has only two axes.

$$def(v) = \sqrt{(u_y + v_x)^2 + (u_x - v_y)^2} = \sqrt{(\frac{U_1 Z_y f}{Z^2} + \frac{U_2 Z_x f}{Z^2})^2 + (\frac{U_1 Z_x f}{Z^2} - \frac{U_2 Z_y f}{Z^2})^2}$$

$$= \sqrt{\frac{f^2(U_1^2 Z_y^2 + U_2^2 Z_x^2 + U_1^2 Z_x^2 + U_2^2 Z_y^2)}{Z^4}} = \frac{f}{Z^2} \sqrt{(U_1^2 + U_2^2)(Z_x^2 + Z_y^2)}$$

$$= |F||A|$$

For the last part we prove that

$$2\mu = (\angle A + \angle F) \iff \tan(2\mu) = \tan((\angle A + \angle F))$$

$$tan(\angle A + \angle F) = \frac{\tan(\angle A) + \tan(\angle F)}{1 - \tan(\angle A)\tan(\angle F)} = \frac{\dfrac{U_2}{U_1} + \dfrac{Z_y}{Z_x}}{1 - \dfrac{U_2}{U_1}\dfrac{Z_y}{Z_x}}$$

$$= \frac{U_2 Z_x + U_1 Z_y}{U_1 Z_x - U_2 Z_y} = \frac{u_y + v_x}{u_x - v_y} = \frac{def(v)\sin(2\mu)}{def(v)\cos(2\mu)} = \tan(2\mu)$$

6

7   Question 10 - Conclusions of the analysis

a) The divergence observed in the image is related to the motion normal
to the image plane since if a plane moves towards the camera centre,
it's image will expand in all directions. Thus the divergence will pro-
vide information on the time of collision.
Translation motion perpendicular to viewing direction results in image
deformation. The magnitude of the surface slant determines the resul-
tant deformation, provided a given movement of the surface. Thus the
Image deformation can be used to determine the surface orientation.

b) Time to collision will be given by distance(Z) divided by the component
of velocity of object in the direction of camera. If we assume that there
is no motion parallel to image plane then $F \bullet A = 0$ and so from
the divergence relation proved in last question, the time to collision is
$2/div(v)$.
For general motion $F \bullet A$ will be bounded by $\pm def(v)$ and thus the
bound on time to collision is given by the following.

$$\frac{2}{div(v) + def(v)} \leq t_{col} \leq \frac{2}{div(v) - def(v)}$$

c) From the equations for the divergence, curl and deformation of the
image velocity field we see that F and A come together in all places.
We also observe that for a given change in A, we can change F so that
the velocity field is unchanged. Thus a large A and a small F and a
small A and a large F could give the same result. So a nearby shallow
object will give the same field as a far away deep object.

d) Large self motion can be used to determine surface orientation but
small scale ego-motion will not work. If we have a large motion of
the camera, we can just consider the images to be two separate camera
images and apply normal Multiple View Geometry results to determine
the depth and orientation. In case of small ego-motion the depth of
the the surface is undetermined and since the change in image velocity
field also depends on the depth of the surface so we won't be able to
find the surface orientation.

8 Max-Flow Min-Cut Theorem Proof

We will show that the maximum of all flow values (ie, the value of the maximum flow), is equal to the minimum of all cut capacities (ie, capacity of the minimum cut). The proof is stated in plain English.

8.1 Lemma 1

Given a network, for any flow F and cut K on the network, the value of the flow is less than the capacity of the cut.
Proof - The cut separates the graph into two parts, one containing the source and one containing the sink. Thus any flow that originates from the source and goes to the sink has to cross the cut. So the capacity of any cut will be more than any flow value in the graph.
From the above it follows that the maximum flow is less than or equal to the minimum cut.
Consider any initial flow given on a flow network. Now we consider the residual graph and find any path on it from source to sink with positive flow (augmenting path). Then we can augment the initial flow by the augmenting path by just adding or subtracting the flows along the augmenting path edges from the corresponding flow network edges. The resultant flow is just the sum of the original flow and the augmenting path flow.

8.2 Lemma 2

Given a flow network there exist a flow F and a cut K such that the capacity of the cut equals the value of the flow.

Proof - We begin with an empty flow and keep computing augmenting paths and adding them to the flow until there are no more augmenting paths left to get a flow F.
Now we define the cut to be such that all vertices reachable from the source are in one set all all the ones that are not are in the other set (the one with the sink). Now any edge which lies on the cut must have it's flow equal to it's capacity since otherwise we will be able to find an augmenting path by connecting the path joining source to that edge and sink to that edge. Since we assumed there were no augmenting paths left so the capacity of this cut

is saturated with the flow and hence we found a flow F and a cut K with the desired properties.

8.3 Theorem

For any network, the value of the maximum flow is equal to the capacity of the minimum cut.

Proof - Assume that the minimum cut capacity is given by $cap(K^*)$ and max flow value by $val(F^*)$. By Lemma 1 and Lemma 2 there exist a flow F and a cut K such that $cap(K) = val(F) \leq cap(K^*)$, but since K^* is the minimum cut so $val(F) = cap(K^*)$. We earlier proved that max flow is less than or equal to min cut capacity, so $val(F^*) \leq cap(K^*) = val(F)$. Since F^* is the maximum flow so the inequality can again be reduced to $val(F^*) = cap(K^*)$ which is the desired result.