

## Solution 1

(a) For convenience, assume that the coordinate system of the left camera is aligned with the world coordinate system. We can then represent the coordinates in the left image using only the intrinsic matrix  $K$ :

$$p_l = [K0]p \quad (1)$$

For the right image, we can find the projected coordinates of the same point using a Euclidean transformation matrix,  $[R|t]$ . In this case we are not rotating the camera, so the rotation matrix will be equivalent to a 3x3 identity matrix. We can therefore write:

$$p_r = K[I|t] \quad (2)$$

$$= \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \end{bmatrix} \begin{bmatrix} \alpha X \\ \alpha Y \\ \alpha Z \\ \alpha \end{bmatrix} \quad (3)$$

Computing the coordinates in the left and right images, we get

$$x_r - x_l = \frac{f}{Z}(X + t_x) - \frac{f}{Z}(X) = \frac{f}{Z}t_x \quad (4)$$

Since there is no disparity except in the x-direction, it is trivial that  $y_r = y_l$ . Since  $t_x$  is positive, it is also clear that  $x_r \leq x_l$ .

(b) Building off part (a), if we know the displacement  $f$  between cameras we can solve this problem by expressing  $Z$  in terms of the left camera coordinates and the world plane coefficients. From the projection matrix we have:

$$X = \frac{f}{Z}x_l \quad (5)$$

$$Y = \frac{f}{Z}y_l \quad (6)$$

$$(7)$$

Substituting these into the world plane equation, we get:

$$\alpha X + \beta Y + \gamma Z = k \quad (8)$$

$$\frac{Z}{f}(\alpha x_l + \beta y_l + f\gamma) = k \quad (9)$$

$$Z = \frac{fk}{\alpha x_l + \beta y_l + f\gamma} \quad (10)$$

Finally, when we substitute this back into equation (4) for the disparity, we get the linear expression  $d = ax + by + c$  we were looking for.

$$d = x_r - x_l = \frac{f}{Z}t_x \quad (11)$$

$$= \frac{\frac{ft_x}{fk}}{\alpha x_t + \beta y_t + f\gamma} \quad (12)$$

$$= \frac{t_x}{k}(\alpha x_t + \beta y_t + f\gamma) \quad (13)$$

(c) Refer back to the matrix multiplication in equation (3). If we substitute in the new translation vector  $[0, 0, -t_z]$  and calculate the new projection for camera 2, we get

$$x_r = \frac{fX}{Z - t_z} \quad (14)$$

The projection for the first camera remains unchanged:

$$x_l = \frac{fX}{Z} \quad (15)$$

If we multiply each of these fractions by its denominator, we can set them equal:

$$(Z - t_z)x_r = Zx_l \quad (16)$$

$$x_r = \frac{Z}{Z - t_z}x_l \quad (17)$$

Similarly,

$$y_r = \frac{Z}{Z - t_z}y_l \quad (18)$$

This means that for a given point  $(x, y)$  in the first camera, the set of corresponding points in the second camera  $(x', y')$  is a set of scalar multiples  $a * (x, y)$ , where  $a > 1$ . Mathematically, this demonstrates the law of perspective: the image in the nearer camera appears larger.

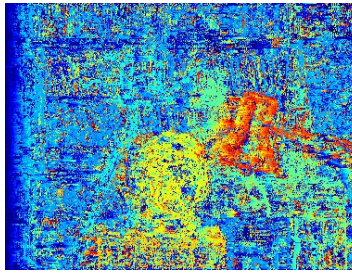
**Solution 2**

Figure 1: Unsmoothed disparity map

(a)

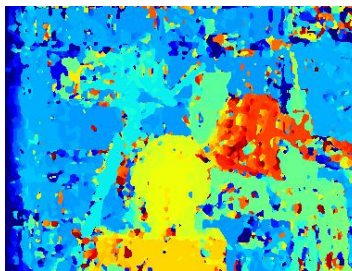


Figure 2: Disparity map smoothed with bilateral filtering

(b)

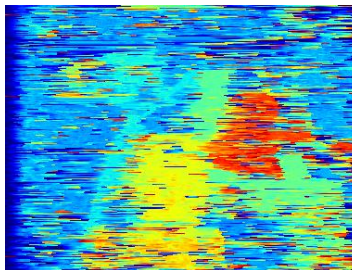
**Solution 3**

Figure 3: Left-to-right viterbi output

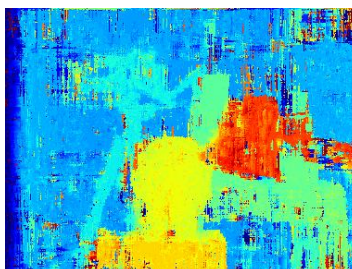
**(a)**

Figure 4: SGM output

**(b)**

## Solution 4

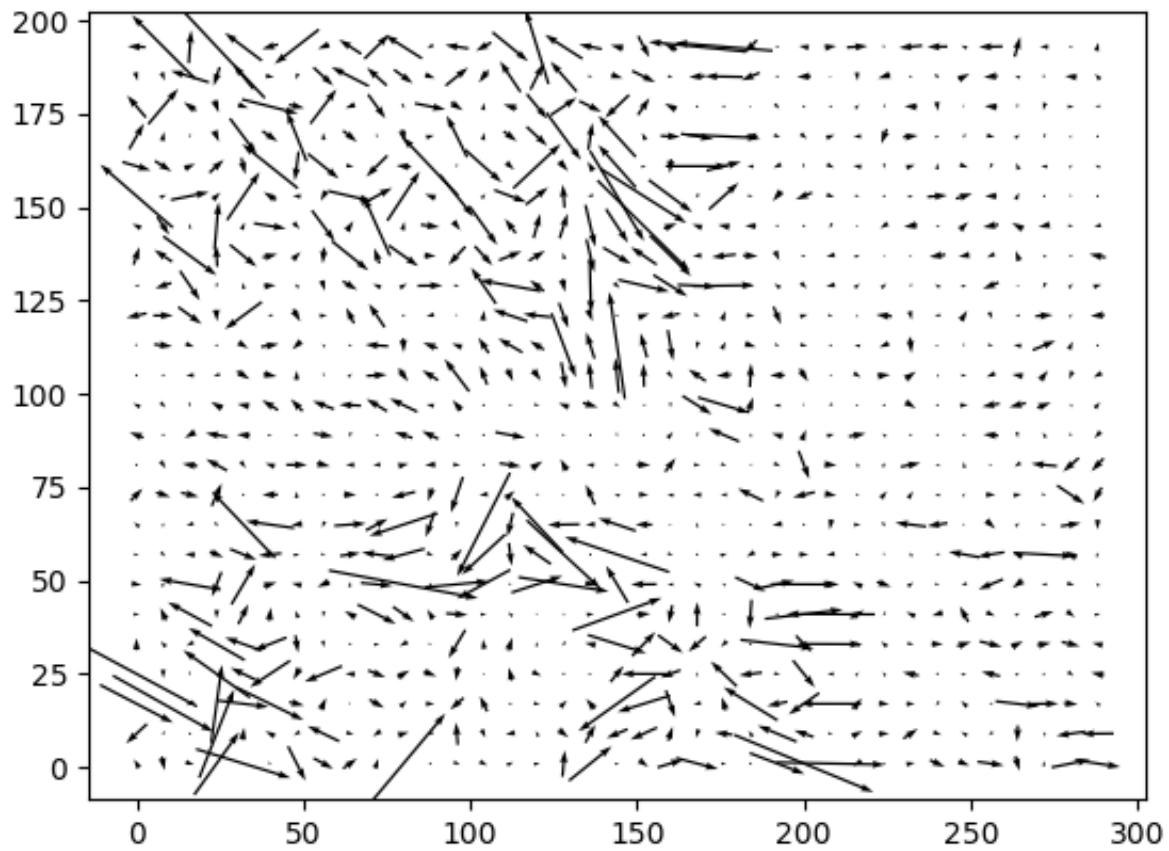


Figure 5: Optical flow map

## Information

This problem set took approximately 16 hours of effort.

I discussed this problem set with:

- Robbie Derber
- TAs

I also got hints from the following sources:

- The Hirschmuller paper on SGM: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4359315>