Assignment 2 Report
Computer Vision EEL803
Single View Metrology & Disparity Maps
Group No 13

Piyush Kaul
Entry No. 2015EEY7544
EE Dept. IITD

Shiva Azimi
Entry No. 2015EEZ8458
EE Dept. IITD

Nilay Pandey
Entry No 2015BSZ8315
EE Dept IITD

***Abstract*—This report covers the ELL803 Computer Vision Assignment 2, on Single View Metrology and Disparity Maps.**

I. Problem Statement 1

Q1.

1) For the above image, estimate the three major orthogonal vanishing points. Use at least three manually selected lines to solve for each vanishing point. In your report, you should:
    - Plot the VPs and the lines used to estimate them on the image plane.
    - Specify the VP pixel coordinates.
    - Plot the ground horizon line and specify its parameters in the form $a*x+b*y+c=0$. Normalize the parameters so that: $a^2+b^2=1$.
2) Using the fact that the vanishing directions are orthogonal, solve for the focal length and optical center (principal point) of the camera. Show all your work.
3) Compute the rotation matrix for the camera, setting the vertical vanishing point as the Y-direction, the right-most vanishing point as the X-direction, and the left-most vanishing point as the Z-direction.
4) Estimate the heights of (a) the building and (b) the tree (spike) assuming that the person nearest to the spike is 5ft 6in tall. In the report, show all the lines and measurements used to perform the calculation. How do the answers change if you assume the person is 6ft tall?
5) Perform additional measurements on the image: which of the people visible are the tallest? What are the heights of the windows? etc.
6) Compute and display rectified views of the ground plane and the facades of the building.
7) Attempt to fit lines to the image and estimate vanishing points automatically either using your own code or code downloaded from the web.
8) Find or take other images with three prominently visible orthogonal vanishing points and demonstrate varions measurements on those images.

Fig. 1. Test Image

9) These vanishing points may lie outside the original image in which case image is shrunk to show the vanishing points. Please see figure. In this figure only the left and right vanishing points are shown as red stars. The vertical vanishing points is at infinity.

II. Solution 1

A. Part1

1) We first convert the image from RGB to grayscale.
2) We then use canny edge detector to find the major edges in the pictures as shown in the figure.
3) In the next step we use the hough transform to find the lines present in the image. We use 30 peaks from the hough transform to find out the prominent lines and use the FillGap parameter so that any discontinuities within the lines are ignored.
4) The final set of lines is as shown in the figure;
5) We manually select 3 lines at a time for generating left, right and top vanishing points respectively

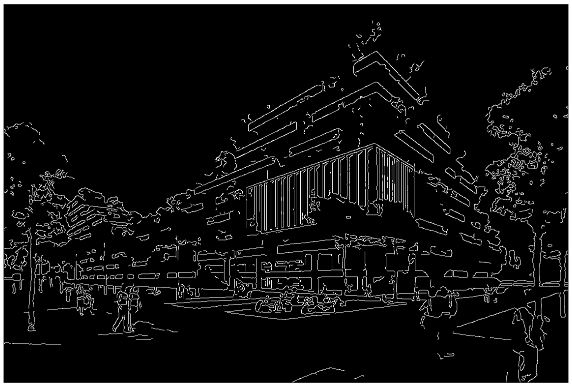

Fig. 2. Detected Edges.

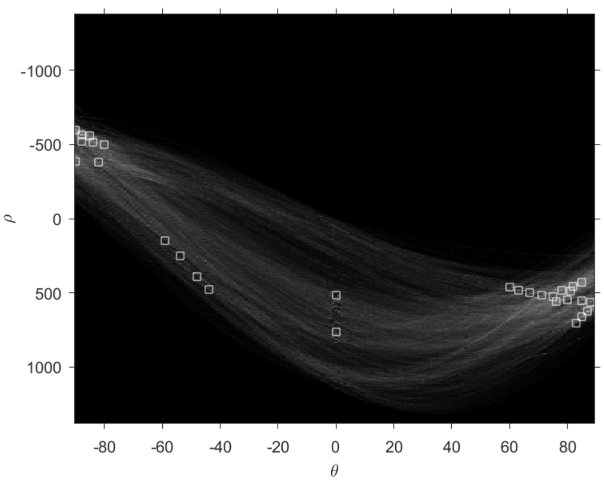

Fig. 3. Hough Transform Results.

6) For the image displayed, only two vanishing points one the left and right are finite. The third vanishing point is at infinity. The locations are as follows

Axis	Kernel	Feature
X	-54.364775	567.916490
Y	1224.976755	578.331715
Z	Infinite	Infinite

The horizon line is specified as line passing through the above points. The equation is given as.
$y = 0.008x + 567.4$

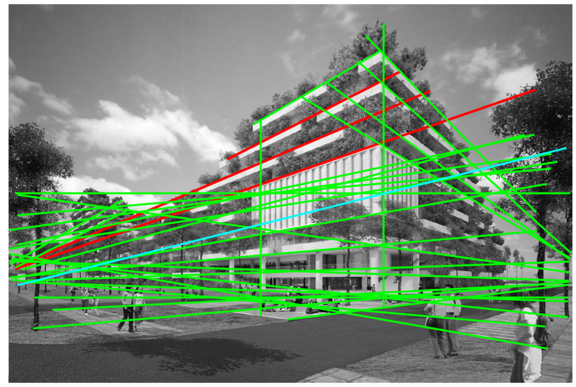

Fig. 4. Final Detected Lines. Manually Selected Lines in Red

Fig. 5. Vanishing Points as Red Starts on Left and Right. Horizon Line in Red

B. Part2: Camera Coordinates

For camera co-ordinates first the projection matrix is calculated as follows

$$M = [v_x v_y v_z I/||I||] \tag{1}$$

The Camera Co-ordinates are given as

$$X_c = [p2 \quad p3 \quad p4] \tag{2}$$
$$Y_c = [p1 \quad p3 \quad p4] \tag{3}$$
$$Z_c = [p1 \quad p2 \quad p3] \tag{4}$$
$$\tag{5}$$

where $p1 = v_x$, $p2 = v_y$, and $p3 = v_z$.

The values obtained for picture above are

$X_c = -505$ and $Y_c = -494$

C. Calibration and Rotation Matrix

For the rotation matrix we first need to obtain the calibration matrix. Let the matrix ω be defined as follows in terms of calibration matrix K.

$$\omega = (KK^T)^{-1} \tag{6}$$

We have the following identities

$$v_x \omega v_y = 0 \tag{7}$$
$$v_x \omega v_z = 0 \tag{8}$$
$$v_y \omega v_z = 0 \tag{9}$$
$$\tag{10}$$

$$\omega = \begin{bmatrix} \omega 1 & \omega 2 & \omega 3 \\ \omega 4 & \omega 5 & \omega 6 \\ \omega 7 & \omega 8 & \omega 9 \end{bmatrix}$$

with constraints. $\omega 2 = 0$ $\omega 1 = \omega$

Utilizing the constraints and utilizing the arbitrary scaling factor to set the element (3,3) to 1, the matrix ω becomes

$$\omega = \begin{bmatrix} \omega 1 & 0 & \omega 4 \\ 0 & \omega 1 & \omega 5 \\ \omega 4 & \omega 5 & 1 \end{bmatrix}$$

So solving the following system of equations utilizing the identities above, we can get the full ω matrix as defined above.

$$\begin{bmatrix} x_3x_1 + y_3y_1 & x_1 + x_3 & y_1 + y_3 \\ x_2x_3 + y_2y_3 & x_2 + x_3 & y_2 + y_3 \\ x_1x_2 + y_1y_2 & x_1 + x_2 & y_1 + y_2 \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_4 \\ \omega_5 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}$$

Now K can be derived from omega by taking inverse and then doing a cholesky factorization.

Once we have K, we can determine the colums of Rotation Matrix R as

$$R1C = K^{-1} * v_x \tag{11}$$
$$R2C = K^{-1} * v_y \tag{12}$$
$$R3C = K^{-1} * v_z \tag{13}$$
$$R = [R1C R2C R3C] \tag{14}$$
$$\tag{15}$$

where R1C, R2C and R3C are colums or Rotation matrix. For image above the following Calibration Matrix was obtained.

$$K = \begin{bmatrix} 493.303113 & 0.000000 & 0.000000 \\ 34.477934 & 999.991524 & 0.000000 \\ 0.080392 & 1.765142 & 1.000000 \end{bmatrix}$$

For image above the following Rotation Matrix was obtained.

$$R = \begin{bmatrix} -0.129382 & 0.488535 & 0.862898 \\ 0.900295 & 0.207211 & 0.382796 \\ 0.890503 & 0.454978 & 0.000000 \end{bmatrix}$$

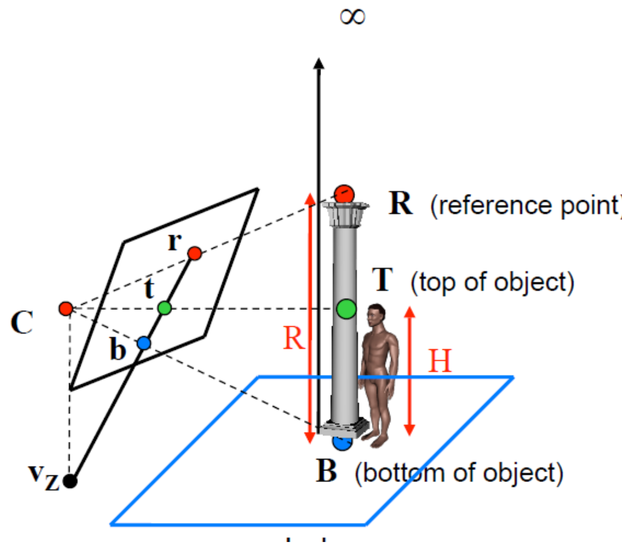

Fig. 6. Height Measurement

Fig. 7. Test Image

D. Height Computation

The height can be measured using the following identity

$$\frac{||t-b|| ||v_z - r||}{||r-b|| ||v_z - t||} = \frac{H}{R}$$

The height of the person nearest to the building is assumed to be 5.5 feet. In that case the height of the building comes out to be 130 feet. In case person is 6 feet tall, the height of building is 142 feet.

E. Other images

Measurements were done for figure shown in figure with 3 finite vanishing points.

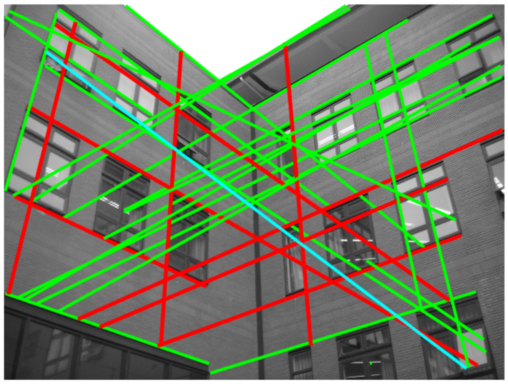

Fig. 8. Final Detected Lines. Selected in Red

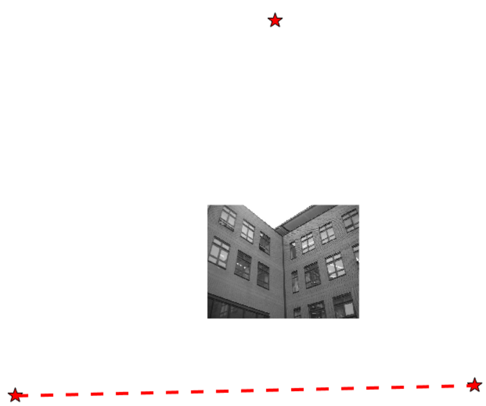

Fig. 9. Vanishing Points as Red Starts on Left and Right

Vanishing Points

$$v_x = \begin{bmatrix} -808.539677 & 807.378585 \end{bmatrix}$$

$$v_y = \begin{bmatrix} 1126.758399764.309588 \end{bmatrix}$$

$$v_z = \begin{bmatrix} 286.297838 - 777.813331 \end{bmatrix}$$

Camera Co-ordinates are $[-362.578 - 131.673 - 505.748]$

For image above the following Calibration Matrix was obtained.

$$K = \begin{bmatrix} 906.500196 & 0.000000 & 0.000000 \\ 83.500954 & 863.929532 & 0.000000 \\ 0.419932 & 0.243848 & 1.000000 \end{bmatrix}$$

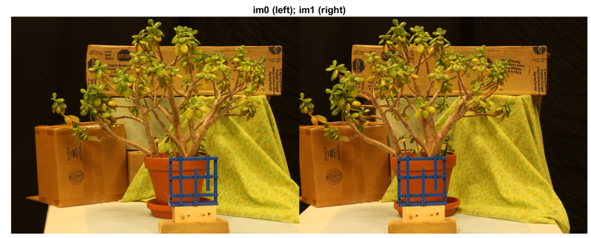

Fig. 10. Input Stereo Images

For image above the following Rotation Matrix was obtained.

$$R = \begin{bmatrix} -0.581610 & 0.555338 & 0.594415 \\ 0.656241 & 0.499880 & 0.565215 \\ 0.283811 & -0.641694 & 0.712517 \end{bmatrix}$$

III. Problem Statement 2

Compute disparity map/ inverse depth map given a stereo image pair. Using stereo data from Middlebury stereo image dataset 4 Marks Marks would be given on basis of method for computation of disparity map, clarity of map, accuracy, speed and optimizations done. Consider that the camera calibration information is not given and only the stereo image pair is given. Clearly demonstrate all the steps used, i.e. rectification, matching, finding depth, optimizing depth. Compare various strategies and see how it can be improved.

IV. Solution

The following steps are taken.

1) We first plot to show both stereo images side by side well as combined.
2) SURF feature are detected in both images. Prominent interest points are detected in those SURF features.
3) We utilize 30 strongest interest point in both images and compare them to find matching interest points.
4) Next we remove the outliers based on those points that do not satisfy epipolar constraints..
5) The inliers are then used to generate the rectification transformation.
6) We next rectify the images based on the rectification matrix and display the rectified images.
7) Lastly we calculate the disparity map from rectified images and display it.

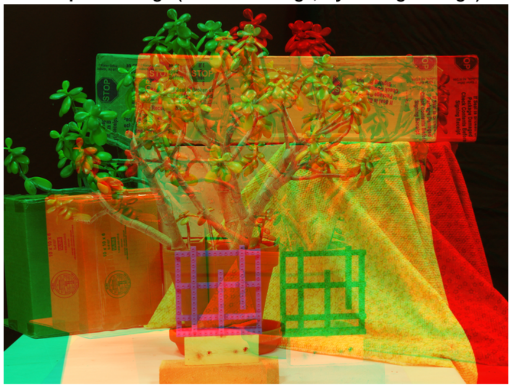

Fig. 11. Input Stereo Images Composite

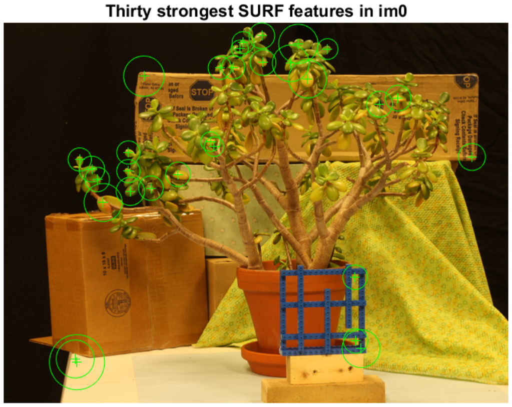

Fig. 12. Surf Features Image 1

Fig. 13. Surf Features Image 2

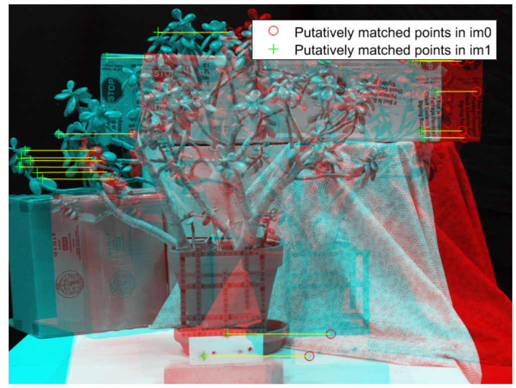

Fig. 14. Matched Points

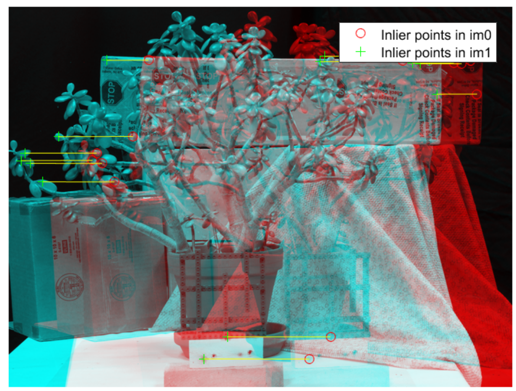

Fig. 15. Matched Inliers

V. Conclusion

The Single View Metrology algorithm was developed and tested on images with 2 or 3 finite vanishing points. Different Calibration, Reflection and Projection matrices were calculated and height calculation also done.

Rectification and Disparity map calculation and display done as required.

References

[1] Matlab http://www.mathworks.com/
[2] Single View Metrology: Criminisi, A., Reid, I. Zisserman, A. International Journal of Computer Vision (2000) 40: 123
[3] Calibration and Single View Metrology: Lecture Slides by Steve Seitz.

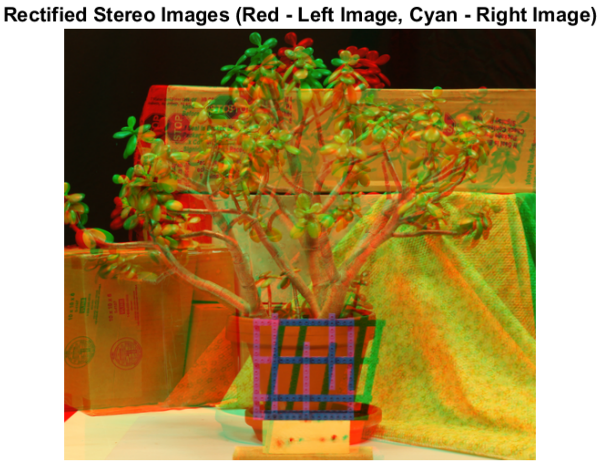

Fig. 16. Rectified Images

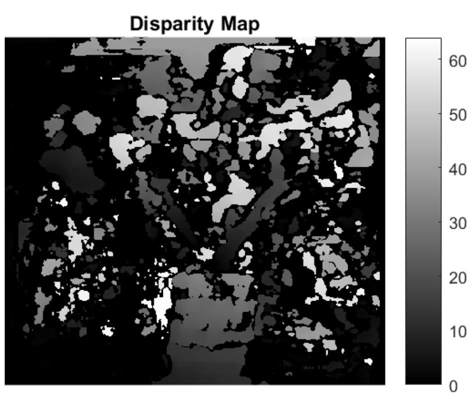

Fig. 17. Disparity Map