A) Matching (Correspondence)

1- Area (correlation)-based Stereo

In correlation based methods the elements to be matched are image windows of fixed size, and a similarity criterion is used to measure the correlation between corresponding windows in the stereo pair. The corresponding element is chosen when the similarity criterion is maximized within a search region. For input stereo pair of images $I_l$ and $I_r$, select a $W \times H$ window and for each point $p_l$ in the left image find its corresponding point $p_r$ in the right image by computing the disparity $d = [d_x, d_y]^T$ that maximizes one of the following similarity measures (where $m = (W-1)/2$, $n = (H-1)/2$):

1- The Normalized Sum of Squared Difference (NSSD)

$$C_1(x, y, d) = \sum_{i=-m}^{n} \sum_{j=-m}^{m} (I_l(x + i, y + j) - I_r(x + i + d_x, y + j + d_y))^2$$

$$\sqrt{\sum_{i=-m}^{n} \sum_{j=-m}^{m} I_l(x + i, y + j)^2 \sum_{i=m}^{m} \sum_{j=m}^{m} I_r(x + i + d_x, y + j + d_y)^2}$$

2- Normalized Cross Correlation (NCC)

$$C_2(x, y, d) = \frac{\sum_{i=-m}^{n} \sum_{j=-m}^{m} I_l(x + i, y + j)I_r(x + i + d_x, y + j)}{\sqrt{\sum_{i=-m}^{n} \sum_{j=-m}^{m} I_l(x + i, y + j)^2 \sum_{i=-m}^{m} \sum_{j=-m}^{m} I_r(x + i + d_x, y + j + d_y)^2}}$$

3- The NSSD with subtracted mean intensity:

$$C_3(x, y, d) = \frac{\sum_{i=-m}^{n} \sum_{j=-m}^{m} \left[I_l(x + i, y + j) - \bar{I}_l(x, y)\right] \left[I_r(x + i + d_x, y + j + d_y) - \bar{I}_r(x + d_x, y + d_y)\right]}{\sqrt{\sum_{i=-m}^{n} \sum_{j=-m}^{m} \left[I_l(x + i, y + j) - \bar{I}_l(x, y)\right]^2 \sum_{i=-m}^{m} \sum_{j=-m}^{m} \left[I_r(x + i + d_x, y + j + d_y) - \bar{I}_r(x + d_x, y + d_y)\right]^2}}$$

4- The (NCC) with subtracted mean intensity defined as:
\[
C_4(x, y, d) = -\frac{\sum_{i=-n}^{n} \sum_{j=-m}^{m} [I_l(x+i, y+j) - \bar{I}_l(x, y)][I_r(x+i+d_x, y+j+d_y) - \bar{I}_r(x+d_x, y+d_y)]}{\sqrt{\sum_{i=-n}^{n} \sum_{j=-m}^{m} [I_l(x+i, y+j) - \bar{I}_l(x, y)]^2 \sum_{i=-m}^{m} \sum_{j=-m}^{m} [I_r(x+i+d_x, y+j+d_y) - \bar{I}_r(x+d_x, y+d_y)]^2}}
\]

where \( \bar{I}_l(x, y) \) and \( \bar{I}_r(x, y) \) are the left and right mean of intensity values within a correlation window centered at \((x,y)\), respectively. For the following stereo pair of images and its corresponding depth map (http://cat.middlebury.edu/stereo/data.html)

Set #1

Set #2

Set #3
Set #4

1. Use the preceding similarity measures to find the depth map images. Compare with the given ground truth depth maps. You may use the root mean square error as a quality measure.

2. For all the above, include in your report the run time taken by the algorithm.

B) 3-D Reconstruction using stereo

Given corresponded pair of point \((x_l, y_l)\) and \((x_r, y_r)\) and assuming calibrated image (\(M_1\) and \(M_2\) are known), the following set of equations can be written to solve for the unknown 3-D point \((X, Y, Z)\):

\[
\begin{align*}
X(x, m_{31} - m_{11}) + Y(x, m_{32} - m_{12}) + Z(m_{33}x_l - m_{13}) &= m_{14} - x, m_{34} \\
X(y, m_{31} - m_{21}) + Y(y, m_{32} - m_{22}) + Z(m_{33}y_l - m_{23}) &= m_{24} - y, m_{34} \\
X(x, m'_{31} - m'_{11}) + Y(x, m'_{32} - m'_{12}) + Z(m'_{33}x_r - m'_{13}) &= m'_{14} - x, m'_{34} \\
X(y, m'_{31} - m'_{21}) + Y(y, m'_{32} - m'_{22}) + Z(m'_{33}y_r - m'_{23}) &= m'_{24} - y, m'_{34}
\end{align*}
\]

These equations can be arranged in the form \(Cl=b\)

\[
\begin{bmatrix}
  x, m_{31} - m_{11} & x, m_{32} - m_{12} & m_{33}x_l - m_{13} \\
  y, m_{31} - m_{21} & y, m_{32} - m_{22} & m_{33}y_l - m_{23} \\
  x, m'_{31} - m'_{11} & x, m'_{32} - m'_{12} & m'_{33}x_r - m'_{13} \\
  y, m'_{31} - m'_{21} & y, m'_{32} - m'_{22} & m'_{33}y_r - m'_{23}
\end{bmatrix}
\begin{bmatrix}
  X \\
  Y \\
  Z
\end{bmatrix}
\begin{bmatrix}
  m_{14} - x, m_{34} \\
  m_{24} - y, m_{34} \\
  m'_{14} - x, m'_{34} \\
  m'_{24} - y, m'_{34}
\end{bmatrix}
\]

\(l\) can be found as:

\[
l = (C^T C)^{-1} C^T b
\]

1. Find the calibration matrices of the left and right cameras using the images \(Calib\_left\), and \(Calib\_right\).
2- Use this reconstruction method to reconstruct the calibration pattern.

3- Using the CardEye system at the CVIP lab (or any stereo system that can capture a stereo pair), generate an image the birdhouse object by the three cameras. From these images that could be named I_cam1 I_cam2, and I_cam3, reconstruct the shape of the birdhouse.

References:


2. [http://cat.middlebury.edu/stereo/data.html](http://cat.middlebury.edu/stereo/data.html)

Acknowledgements: This Laboratory evolved over various iterations of offering of ECE 619(ECE645). Special thanks are due to Dr. Moumen Ahmed and Dr. Ahmed Eid who helped in putting this document together during 2002-2005. Their PhD work is readily available on the CVIP Lab website [www.cvip.uofl.edu](http://www.cvip.uofl.edu). The Laboratory was also implemented on the CardEye vision system by Asem Ali and was maintained throughout 2005-2008.