Georgia Institute of Technology
George W. Woodruff School of Mechanical Engineering
ME6406 Machine Vision (Fall 2017)
Assignment #4
Due December 5, 2017 (Tuesday)

<u>Problem 1 Artificial neural network (ANN)</u>: Design an ANN with 49 inputs and 4 outputs and train using back-propagation learning to recognize ‘M’, ‘E’, ‘1’, ‘7’ on a binary 7x7 square grid.

1. Derive the weight update rule for the ANN (using a **bipolar** sigmoid function for all the processing elements).

2. Write a Matlab program NN_training.m to train the data in training_data.mat (Fig. 1a). Show the converge curve (mean squared error vs. number of epoch). Save the weights of nodes in the file NN_weights.mat.

3. Write a Matlab program NN_test.m to test the data in test_data.mat (Fig. 1b) by reading the weights of nodes in NN_weights.mat. Show the values of four output nodes for each testing pattern.

Note: Download “training and test” data from t-square.

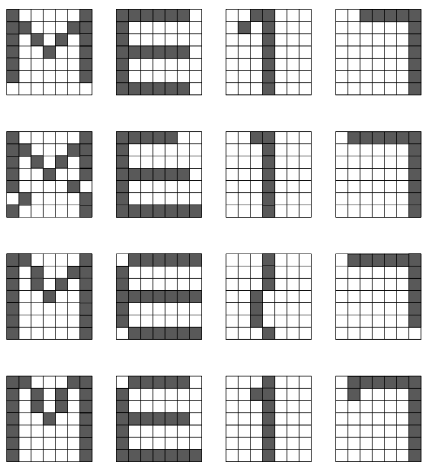

Fig 1(a). Training data

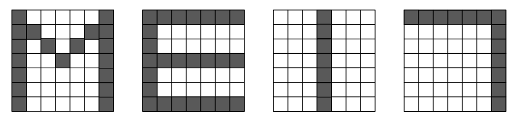

Fig 1(a). Testing patterns

Table 1. Training target

	Output vector
M	[1 0 0 0]
E	[0 1 0 0]
1	[0 0 1 0]
7	[0 0 0 1]

<u>Problem 2 Pose Estimation and Stereo Vision</u>

(a) <u>Camera Model</u>. Write a program (CameraModel.m) to compute Z (depth) from two cameras (Fig. 2a). Transform the image ‘*’ (8 features points in Table 2a) from the 3D world coordinates ($X_wY_wZ_w$) to the 2D undistorted image coordinates (u_1v_1, u_2v_2) for Camera I and II respectively. Use [$\mathbf{R}_x(180°)$], $\mathbf{T}$=[2 1 8]^T, followed by [$\mathbf{R}_y(-40°)$], f=1.3 for camera 1 and [$\mathbf{R}_x(180°)$], $\mathbf{T}$=[-2 1 8]^T, followed by [$\mathbf{R}_y(40°)$], f=1.3 for camera 2. Determine and show these 8 feature points (Table 2) in the u_1v_1, u_2v_2 plane.

(b) <u>Pose Estimation</u> [1]. Write a program (PoseEstimation.m) to calculate the rotation matrix (**R**) and translation matrix (**T**) from world coordinate to Camera I.

1. Determine these 8 feature points (Table 2a) and the centroid and normal vector of the circle (Table 2b) in Camera I coordinate (x_{c1}-y_{c1}-z_{c1}).
2. Form the matrix $\mathbf{W}=\begin{bmatrix}\mathbf{M} & \mathbf{Q} & \mathbf{Q'}\end{bmatrix}^T$ and apply the pseudo-inverse on $\mathbf{WV}=\mathbf{b}$ to reconstruct **R** and **T**
(**M**: point correspondences, **Q**: ellipse-circle correspondences, **Q’**: orthonormal constraints)

(c) <u>Stereo Vision (Parallel)</u>. Write a program (StereoVisionParallel.m) to compute Z (depth) from two parallel cameras (Fig. 2b). Transform the image ‘*’ (8 features points in table 1a) from the 3D world coordinate ($X_wY_wZ_w$) to the 2D undistorted image coordinate (u_1v_1, u_2v_2) for camera 1, 2 respectively (Fig. 2b). Use [$\mathbf{R}_x(180°)$], $\mathbf{T}$=[2 1 8]^T, f=1.3 for camera 1 and [$\mathbf{R}_x(180°)$], $\mathbf{T}$=[-2 1 8]^T, f=1.3 for camera 2. Determine and show these 8 feature points in the u_1v_1, u_2v_2 plane. Determine the depth of each point (Z) by $Z=bf/d$. b is the distance between two cameras. f is the focus of these two cameras. d is the disparity which can be determined by $d=u_1-u_2$.
Note 1: The (u, v) values should be rounded off to a reasonable number of decimal places, say 6; otherwise, numerical errors due to high number of decimals places could cause complex number matrix [**R**].

(d) <u>Stereo Vision (General)</u>. Write a program (StereoVisionGeneral.m) to reconstruct 8 feature points in the world coordinate from u_1v_1, u_2v_2 planes in Fig. 2a.

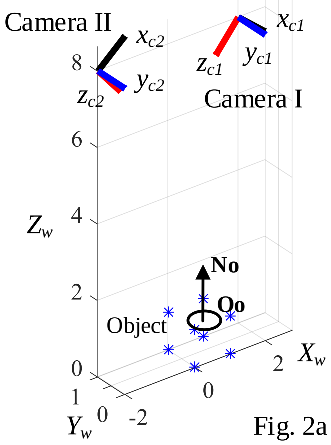

Fig. 2a

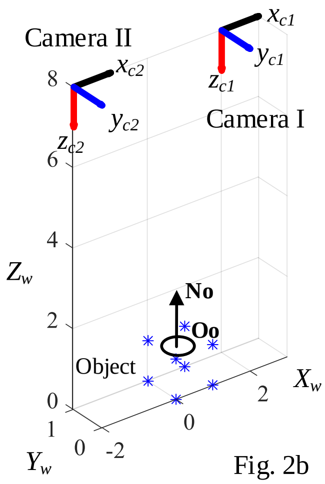

Fig. 2b

Table 2(a) Feature points

X_w	0	1	1	0	0	1	1	0
Y_w	0	0	1	1	0	0	1	1
Z_w	0	0	0	0	1	1	1	1

Table 2(b). Centroid and normal vector of the circle

O_o	0.5	0.5	1
N_o	0	0	1

Problem 3: Color

a) Artificial Color Contrast (ACC) [3]: ACC can be used for color classification. The objective here is to identify fan-bone (Fig. 3) but not the blood stains. This can be done by using *DoG* as discussed in class:

$$h(x,y) = DoG * f_j(x,y) + G_{\sigma s} * (f_j(x,y) - f_k(x,y))$$

Use $f_j(x,y) = R$, $f_k(x,y) = -(R-G)$ on the '*fanbone.bmp*' (Fig. 3) image to filter out blood stains and find fanbone.

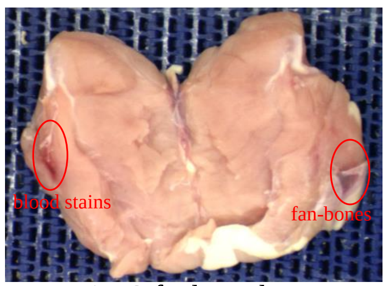

Fig. 3 fanbone.bmp

b) Principle component analysis (PCA): Use the RGB image '*fanbone.bmp*' for the following.

1) Determine the covariance matrix of data.
2) Derive the components (eigenvectors) with eigenvalues arranged in a descending order.
3) Map the R, G, B data on these three component axes. Represent the projected data along three component axes. Show the three single channel component images. Further, represent the projected data by RGB coordinate. Show the three color component images. Compare and discuss these images.

c) Color-based Image Segmentation: Color is an important information. Same objects commonly have their domain color. *L-a-b* color system is the color-opponent space with the *L* lightness dimension and *a-b* color-opponent dimensions. The color-based image segmentation can be performed by applying the clustering method on the points in *a-b* domain. Transfer pixels from *RGB* to *Lab* color system and followed by applying *k-means* clustering on data in *a-b* domain with cluster number (*k*=3) for segmentation of the *RGB* image '*fanbone.bmp*'.

Reference:

[1] Qiang Ji, Mauro Costa, Robert Haralick, and Linda Shapiro, "An Integrated Linear Technique for Pose Estimation from Different Features," *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 13, No. 5, 1999

[2] Simpson, P. (1992). Foundations of neural networks, Chapter 1 in *Artificial Neural Networks*, Lau, C. & Sanchez-Sinencio, E. (Eds.), pp. 2-13, IEEE Press, New York, NY

[3] Lee, K.-M.; Li, Q.; Daley, W, "Effects of Classification Methods on Color-Based Feature Detection With Food Processing Applications," *IEEE Trans. on Automation Science and Engineering*. V. 4, NO. 1, Jan. 2007 pages: 40-51