ME 592X: Signal and Image Pre-Processing (Due on March 2)

Instructions: preferred programming language of choice is, MATLAB and Python. Include proper comments. Submit all files needed to re-run the codes e.g., Makefile and sample results from your simulations

1. **Basic Image Processing**: Write codes to perform the possible operations for pre-processing the three images in top row on slide 1.

   Hint: Use contrast stretching, histogram equalization and/or other basic image-processing algorithms.

2. **Background subtraction**: The goal here is to extract only the Arabidopsis seedlings from the images in slide 2. An example is highlighted on the right hand side for a cropped portion as well as the corresponding segmented result. Write codes for performing the background subtraction using a thresholding based method and compare the optimum thresholds obtained for the two panels shown here.

3. **Metric in image space**: On slide 3, we provide input, output and intermediate features from a deep convolutional selective autoencoder. Write codes to find and plot the distances from the input image to the intermediate features and the output image. Use at least two valid distance metrics (one local, e.g., pixel-wise, one global, e.g., structural) and compare the plots.

4. **Shading correction**: On slide 4, assume I_{min} [m,n] and I_{max} [m,n] are the min and max intensity versions of the original ground truth image respectively;
   - Perform shading corrections for the 3 images in the bottom row. State the assumptions made in the corrections.
   - Calculate the Structural Similarity Index Measure (SSIM) of your final results from the ground truth.

   (Hint: SSIM:Python example - http://scikit-image.org/docs/dev/auto_examples/plot_ssim.html)

5. **Image pre-processing and transformation**: Write codes to:
   - Obtain local patches (of a certain fixed size of your choice, much smaller than the original image size) from all the leaf images
   - Prewhiten the local patches
   - Determine the channel-by-channel distribution of the prewhitened images
   - Determine the channel-by-channel distribution of the original images

   Write short comments on the differences in distributions of the corresponding channels

   Hint: PCA whitening algorithm:
   http://ufldl.stanford.edu/wiki/index.php/Implementing_PCA/Whitening

7. **Time series pre-processing:** The Time_Series.xlsx excel sheet provides the forcing value and the outputs under clean and noisy conditions of a random mix-and-match of a simulated nonlinear Duffing systems and a Vander pol system.

    - Perform pre-processing of the output variable using steps (such as normalization, denoising) as needed
    - Use two suitable time series feature extraction techniques to automatically (without providing any knowledge about the Duffing and Vander pol characteristics to the algorithm) detect changes in time series characteristics (from Duffing to Vander pol and vice versa)
    - Compare the results for the clean and the noisy conditions
    - Show (computationally) that a segment of either Duffing or Vander pol system outputs is *stationary* (i.e., wide-sense stationary, WSS)

8. **Metrics in Time series space**: The .mat files in the electricity dataset contain energy consumption (in Watts) time series for different end uses as well as the main power for a house sampled at 1Hz. Explore the efficacies of different distance metrics and transformations for time series by computing differences among different end uses and the main power. Some of the suggested metrics are – direct Euclidean distance, KL divergence in a discretized space, distance in a frequency domain or a wavelet transformed space. Feel free to use other metrics and transformations. Comment on your findings.