Worksheet 3 (StepScan)

Topic 1: Continue from Worksheet 1

1) The paper entitled “Personal Authentication Using Palm-Print Features” [1] is the main resource on this topic.
2) From Section 2, Step 1, in the paper, the histogram of gray images is analyzed to determine a threshold value for creating a binary image. One of the most common methods to determine a threshold value from histogram is **Otsu's method**.
   “Otsu's thresholding method involves iterating through all the possible threshold values and calculating a measure of spread for the pixel levels each side of the threshold, i.e., the pixels that either fall in foreground or background. The aim is to find the threshold value where the sum of foreground and background spreads is at its minimum.”
   Read http://www.labbookpages.co.uk/software/imgProc/otsuThreshold.html
3) From Section 4.1, in the paper: **Multi-template-matching approach**
   a. Instead of using the minimum distance from *M*-1 templates, we can use the average distance (mean or median) for an individual testing footstep *x* to the other *M*-1 templates.
   b. In Worksheet 1, we apply the leave-one-out cross-validation method over all the available footsteps. It means that we only calculate the training performance (**enrollment** process). In this topic, we should separate some footsteps for testing (**verification** process). The output that we get from the enrollment process is the threshold value to use in the verification process for individual subjects.
   c. From Worksheet 1, you should be able to observe the differences between using Euclidean distance and correlation coefficient as a score. There are several ways to convert the distance values to similarity scores, such as $score = \frac{1}{1+\text{distance}}$ or $score = \frac{1}{e^{distance}}$. If the score value is higher than threshold *t*, then the owner of query sample is claimed to be individual *x*. Otherwise, the query sample is classified as a forged pattern.

Topic 2: Principal Component Analysis

1) The paper entitled “Comparing and Combining Underfoot Pressure Features for Shod and Unshod Gait Biometrics” [2] is the main resource on this topic. This is one of the papers from Worksheet 2.
2) First is to compute COP time series.
3) Second is to compute the PCA coefficients for the training data only. Note that the minimum requirement for normalization is to mean center the data, and the mean values for the training data must be stored for normalizing the testing data.
4) Third is to calculate the COP time series in the new coordinate system using the first principal components explained *X* % of the total variance in the data.
5) Fourth is to use (1) multi-template-matching approach and (2) machine learning approaches (e.g. support vector machine, linear discriminant analysis, and multi-layer back-propagation neural network) for classification.
6) Fifth is to evaluate models using the key performance metrics: (I) the false acceptance rate (FAR), (II) the false rejection rate (FRR), (III) the accuracy graph (which plots the FAR and

FRR as functions of the threshold), (IV) the receiver operating characteristic (ROC) curve (which plots FAR against the FRR), (V) Zero-FAR, and (VI) the equal error rate (EER).

[1] Han et al., “Personal Authentication Using Palm-Print Features”, Pattern Recognition, 36(2), pp. 371-381, Feb. 2003.

[2] P.C. Connor, “Comparing and Combining Underfoot Pressure Features for Shod and Unshod Gait Biometrics,” Proc. IEEE HST, Apr. 2015.

Checkpoint

Please show the key performance metrics for all the subjects for CASIA-D (average results), Left side, and PCA features using (1) a multi-template-matching approach and (2) an SVM approach (with a linear kernel). Please use the Otsu’s method for creating a binary image, the average Euclidean distance (transformed to the similarity score matrix using 1/(1+distance)) for a multi-template-matching approach, and the first 5 footsteps for training and the remaining footsteps for testing for each user.