Supplemental Appendix I

As described earlier, the resulting value of the global similarity statistic is dependent upon the manual selection and annotation of features by fingerprint experts. This supplemental appendix (1) evaluates the precision of feature annotations with respect to their location and angle for a given feature by fingerprint experts as well as (2) describes how such variation is accounted for in the resulting global similarity statistic using simulated variations of feature annotations.

Empirical variability of feature annotations

The variability of feature annotations (for a given feature) was evaluated with respect to the differences in the location and angle of each annotated feature compared to a specified reference point. This was evaluated separately for latent impressions and reference impressions due to the general clarity differences between the two types of impressions.

Latent Impressions:

The variability of feature annotations in latent impressions was evaluated using five practicing latent fingerprint experts employed by a federal crime laboratory in the United States. Each expert was provided five sets of fourteen images of latent fingerprints. Each set contained the same fourteen images. Considering the intent of this evaluation is to capture the reproducibility of annotations for a given feature, a template image was also provided indicating which features the experts should annotate. The template did not, however, indicate exactly where or how to annotate the feature. Although the specific number of features varied across images, the total number of features annotated by each expert per set was 100 (n = 2,500 annotations in total). The overall quality of the images used was subjectively considered representative of the quality of typical latent impressions received during normal casework. Of the available features, those that were subjectively evaluated as “low” or “medium” clarity (on a scale of “low”, “medium”, and “high” clarity) were specified on the template image. Experts were advised to annotate each set during normal business hours using the same software and hardware as they normally would in actual casework. Furthermore, experts were given one week to complete the five sets and were advised to ensure at least four hours lapsed between each set.

The X, Y coordinates and angle for each feature in each image was extracted. The mean X, Y coordinate and angle for each feature was calculated across all experts and sets and served as the “consensus” reference location and angle. The difference between the “consensus” feature location and angle compared to each annotation was calculated. Figure SAI-1 illustrates the empirical distribution of variations in X, Y, and angle annotations, respectively.

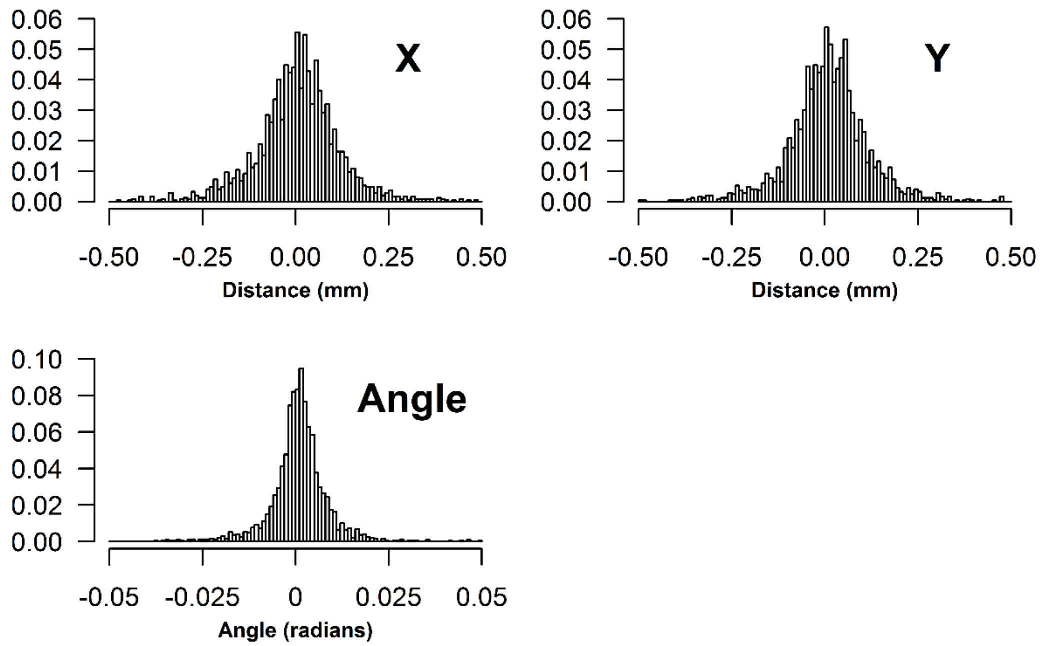

Figure SAI-1. Empirical density distributions of the X, Y, and angle differences as a result of variations in feature annotations on latent impressions (n = 2,500).

<u>Reference Impressions:</u>

The variability of feature annotations in reference impressions was evaluated using ten practicing latent fingerprint experts employed by a federal crime laboratory in the United States. Each expert was provided ten replicate images of a single reference fingerprint. Considering the intent of this evaluation is to capture the reproducibility of annotations for a given feature, a template image was also provided indicating which features the experts should annotate. The template did not, however, indicate exactly where or how to annotate the feature. The template specified ten features to annotate resulting in a total of 100 annotations by each expert (n = 1,000 annotations in total). All specified features were subjectively evaluated as “high clarity” (on a scale of “low”, “medium”, and “high” clarity) and representative of the quality of typical reference impressions received during normal casework. Experts were advised to annotate each image during normal business hours using the same software and hardware as they normally would in actual casework. Furthermore, experts were given approximately one week to complete the annotations and were advised to ensure at least four hours lapsed between annotations for each image.

The X, Y coordinates and angle for each feature in each image was extracted. The mean X, Y coordinate and angle for each feature was calculated across all experts and sets and served as the “consensus” reference location and angle. The difference between the “consensus” feature location and angle compared to each annotation was calculated. Figure SAI-2 illustrates the empirical distribution of variations in X, Y, and angle annotations, respectively.

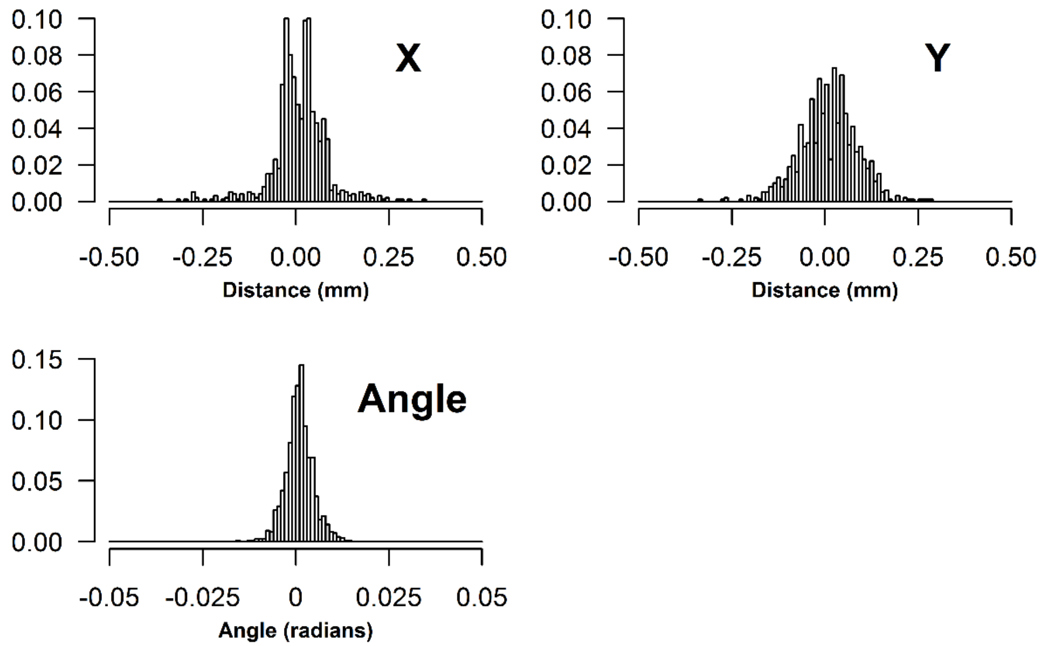

Figure SAI-2. Empirical density distributions of the X, Y, and angle differences as a result of variations in feature annotations on reference impressions (n = 1,000).

Simulated variability of feature annotations

The global similarity statistic value is initially calculated based on the locations and angles of features annotated by the user. To simulate the impact of normal variations by analysts re-annotating the same features, the similarity statistic is recalculated *k* iterations (*k* = 100) using randomly displaced feature locations and angles on both image #1 and image #2 normally distributed around the initial annotation made by the user. The displacements are calculated using a scaled approximation of the inverse of the normal cumulative distribution. The parameter values were estimated using the empirical distributions and manually optimized to minimize the differences between the empirical distribution and a randomly generated sample distribution. For latent impressions, the parameter value for distance variations was estimated with respect to both X and Y-value differences since the two empirical distributions appear very similar to one another. For the reference impressions, however, the parameter value for distance variations was estimated with respect to Y-value differences only since the X-values indicate *less* variation than Y-values. This observation seems to suggest there is some other factor influencing the feature annotations with respect to the X-axis compared to the Y-axis on the reference impressions. As a result of this observation, the template image used to specify which features to annotate was examined to determine whether the feature selection could have biased the variations in one direction vs. another (i.e., X-value vs. Y-value displacements). Indeed, it was observed that 80% of the specific features selected for the expert all occurred in the north-south direction. Accordingly, the differences in variation appear to be related to the uncertainty associated with annotating the specific end-point location of the ridge (spanning north to south) rather than the precise center of the ridge, which is more clearly interpretable. While the explanation seems plausible, the more important consideration is that the parameter value for distance variations be estimated using the Y-value differences, which exhibited greater variation.

The empirical distributions of X, Y, and angle differences were compared to a randomly generated sample distribution. Figure SAI-3 illustrates the comparison between the randomly generated sample distributions and the empirical distributions for X, Y, and angle displacements for latent impressions. Table SAI-1 provides the two sample Kolmogorov-Smirnov (K-S) test statistics as well as the resulting *p*-values under the null hypothesis that the empirical distributions and randomly generated distributions were drawn from the same distribution.

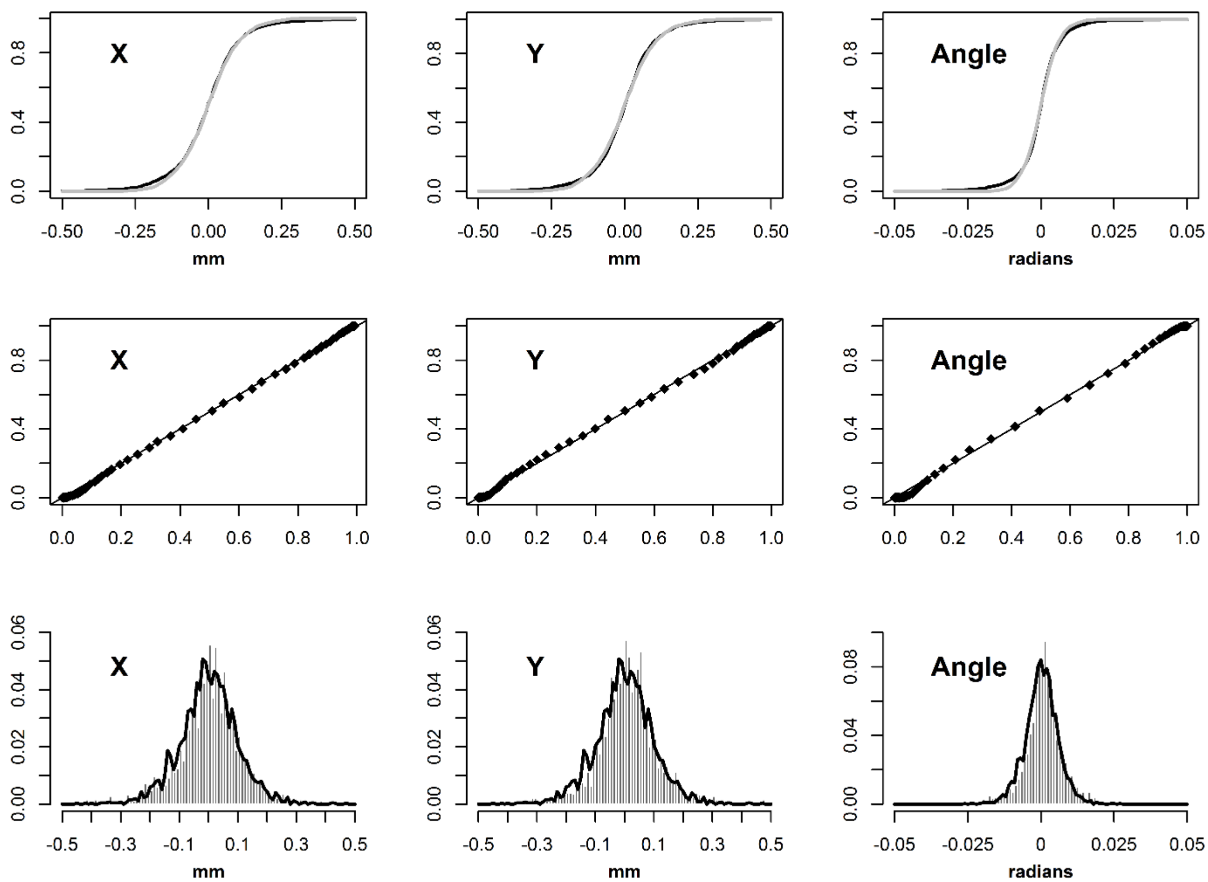

Figure SAI-3. Empirical distributions of the X, Y, and angle differences as a result of variations in feature annotations compared to a randomly generated distribution of X, Y, and angle differences using a scaled approximation of the inverse of the normal cumulative distribution for latent impressions. Top row illustrates overlays of the empirical cumulative distribution (black) and randomly generated dataset (grey). Middle row illustrates the P-P plots between the empirical dataset (X-axis) and randomly generated dataset (Y-axis) (the black dots represent the P-P plot and the grey line represents the ideal slope of 1). Bottom row illustrates overlays of the empirical density distribution (grey histogram) and randomly generated dataset (black line).

Feature quantity	*n* sample 1 (empirical)	*n* sample 2 (randomly generated)	K-S test statistic	*p (null)*
X-value difference	2,500	2,500	0.025	$p >> 0.05$
Y-value difference	2,500	2,500	0.023	$p >> 0.05$
Angle difference	2,500	2,500	0.028	$p >> 0.05$

Table SAI-1. Summary of the Kolmogorov-Smirnov test results between empirical distribution and randomly generated distribution of displacements for the X, Y, and angle for latent prints. Statistical significance is based on a p-value decision threshold of 0.01.

Figure SAI-4 illustrates the comparison between the randomly generated sample distributions and the empirical distributions for X, Y, and angle displacements for reference impressions. Table SAI-2 provides the two sample K-S test statistics as well as the resulting *p*-values under the null hypothesis that the empirical distributions and randomly generated distributions were drawn from the same distribution.

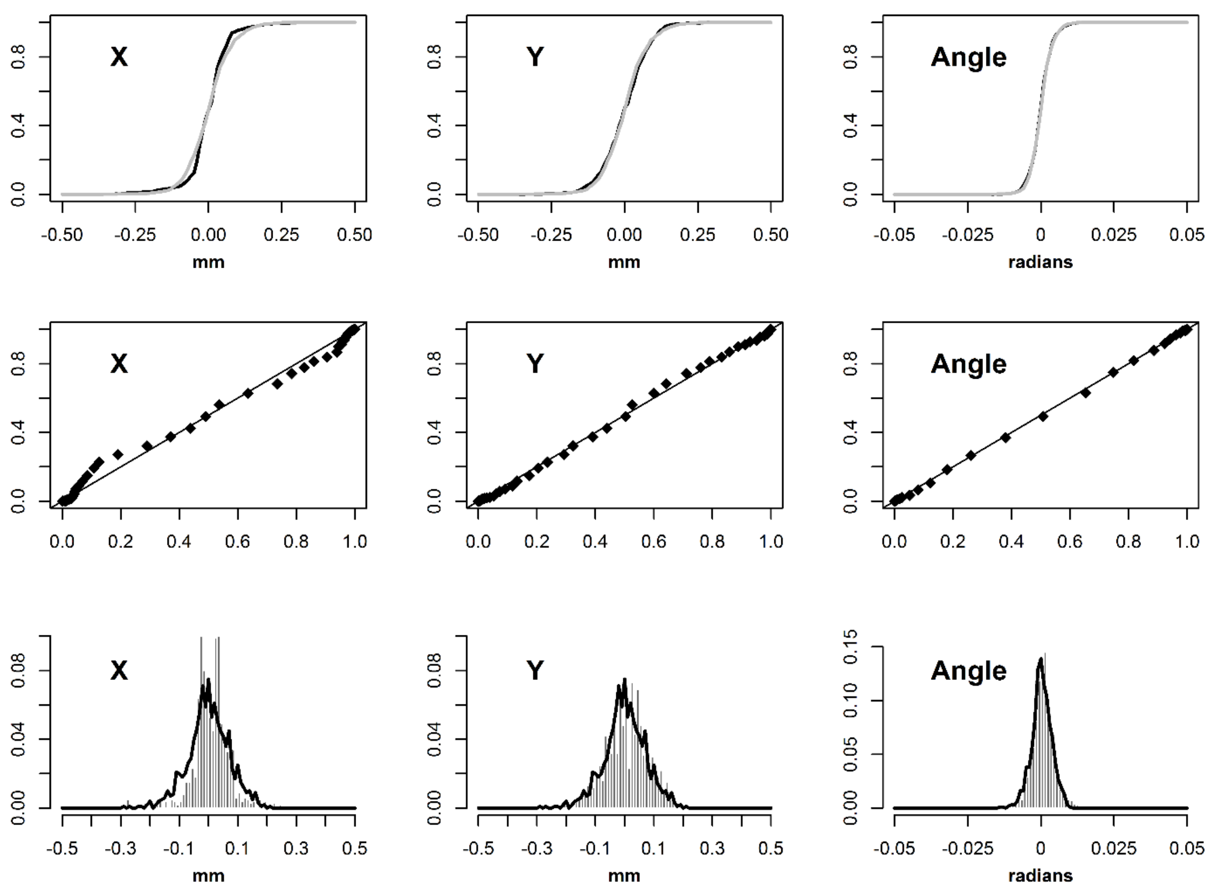

Figure SAI-4. Empirical distributions of the X, Y, and angle differences as a result of variations in feature annotations compared to a randomly generated distribution of X, Y, and angle differences using a scaled approximation of the inverse of the normal cumulative distribution for reference impressions. Top row illustrates overlays of the empirical cumulative distribution (black) and randomly generated dataset (grey). Middle row illustrates the P-P plots between the empirical dataset (X-axis) and randomly generated dataset (Y-axis) (the black dots represent the P-P plot and the grey line represents the ideal slope of 1). Bottom row illustrates overlays of the empirical density distribution (grey histogram) and randomly generated dataset (black line).

Feature quantity	***n* sample 1** (empirical)	***n* sample 2** (randomly generated)	**K-S test statistic**	***p (null)***
X-value difference	1,000	1,000	0.103	$p < 0.001$
Y-value difference	1,000	1,000	0.042	$p > 0.05$
Angle difference	1,000	1,000	0.019	$p > 0.05$

Table SAI-2. Summary of the Kolmogorov-Smirnov test results between empirical distribution and randomly generated distribution of displacements for the X, Y, and angle for reference prints. Statistical significance is based on a p-value decision threshold of 0.01.

Based on these findings, with the exception of X-value differences in the reference impressions (for reasons previously discussed), the distributions exhibit little difference and thus the scaled approximation of the inverse of the normal cumulative distribution is proposed as a sufficient means of simulating the impact of variations in feature annotations by fingerprint experts. Figure SAI-5 illustrates the simulated variations of feature annotations.

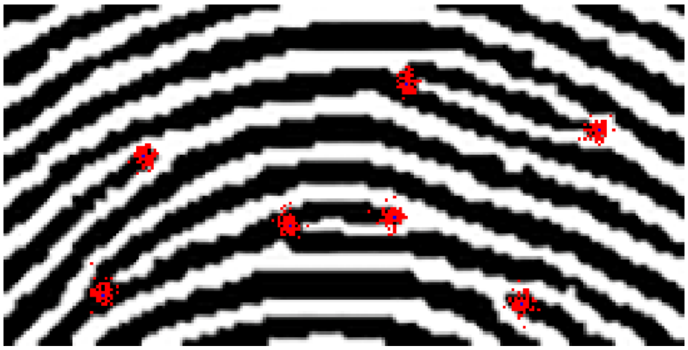

Figure SAI-5. Illustration of the iterative random sampling scheme for the annotated details resulting in random displacements of feature annotations. The blue dot represents the X, Y pixel location of the center of the original annotation by the expert. The red dots each represent separate randomly generated displacements. NOTE: only displacements in terms of Euclidean distance are illustrated in this figure.