Statistical Error Estimation for an Objective Measure of Similarity to a Latent Image

Donald T. Gantz, PhD and John J. Miller, PhD

George Mason University, Volgenau School of Engineering, Fairfax, VA 22030

Innovative Technology for Latent Print Examination

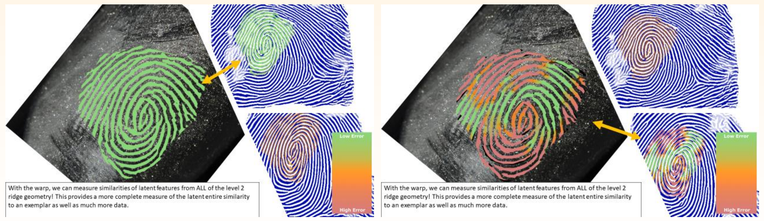
Sciometrics' LatentSleuth Workstation for Latent Fingerprint Examination allows a Latent Print Examiner (LPE) to automatically WARP a Latent Image Region of Interest (ROI) to any fingerprint image.

The WARP automatically finds the best location for the Latent ROI in the fingerprint image and creates a *distortion free* overlay of the Latent ROI onto that location in the fingerprint image.

The LatentSleuth Workstation was evaluated by the Virginia Department of Forensic Science (VA DFS) (in an NIJ Grant) and is being used by VA DFS LPEs for casework.

The Grant being presented here exploits the LatentSleuth WARP technology to create a model of *Random (non-Mate) Level 2 Similarity* to a specific Latent ROI.

Level 2 similarity is quantified by a WARP *continuously and comprehensively across the fingerprint image.*



Quantifying Level 2 Similarity

The Grant Research relates to the following scenario for Latent Fingerprint Examination:

A Region of Interest (ROI) is specified in the Latent Image.

The Latent ROI is consistent with the quality areas of a particular exemplar image of interest.

The scenario for modeling is that it provide support for a Latent Print Examiner's (LPE's) report concerning a Latent Image and a specific exemplar.

An Objective Measure of Similarity of a Candidate Image to a Latent Image ROI

The Similarity of a Candidate Image to the Latent Image ROI is measured relative to the collective measured similarity of a Base Set of known non-mate fingerprint images.

A Base Set of 50 high quality fingerprint images, representative of the full range of Level 1 patterns, is selected as the basis for the definition of an Objective Measure of Level 2 Similarity to a Latent Image.

We independently WARP a Latent Image ROI to all of the Base Set Images as well as to a Test Fingerprint Image. For an image *i*, $d_i(\omega)$ is a non-zero measurement of WARP error at location ω in the Latent Image ROI. WARP error quantifies local curvature correspondence between Latent ROI and Test Image.

		BASE SET						
		1	2	···	*J*	···	49	50
BASE SET	1							
	2							
	•••							
	i				*i,j*			
	•••							
	49							
	50							
Test Image	*k*				*k,j*			

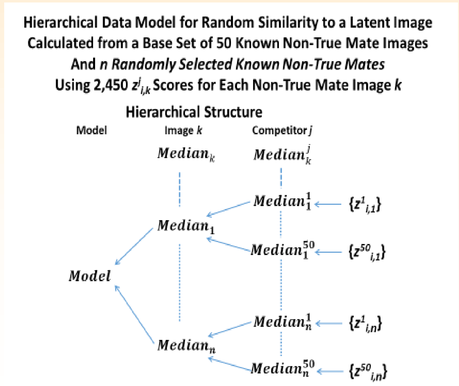
Via the WARPs, all Base Set images compete among themselves for accuracy of coverage of the latent image's Level 2 features at each location ω in the Latent Image ROI. This competition among Base Set images for similarity to the latent image yields the massive amount of data that is the basis for defining the Objective Measure of Similarity to the Latent Image for an additional fingerprint image that competes against the Base Set as a whole.

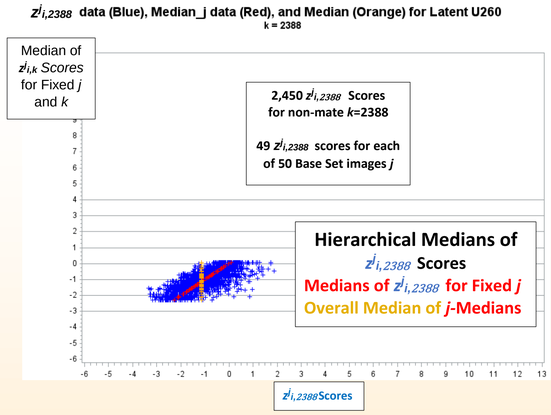
Competitive Similarity Data Structure

For a Test Image *k*, competition with the Base Set yields 2,450 $z^j_{i,k}$ scores. Each $z^j_{i,k}$ score is a summary of the Test Image's similarity to the Latent ROI across the many 1,000's of locations in the Latent Image ROI.

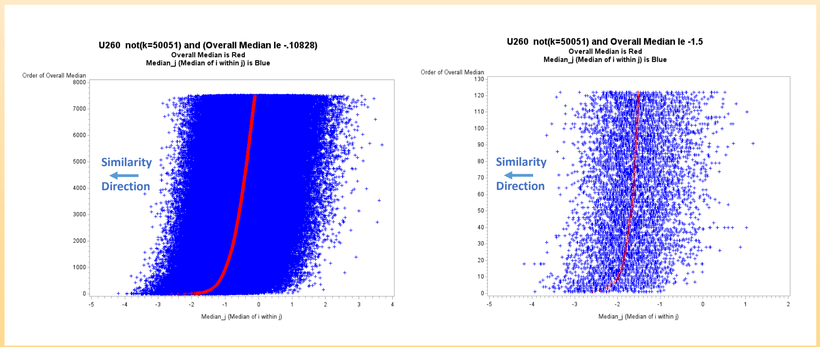
The Hierarchical Median of the 2,450 $z^j_{i,k}$ scores is the statistic used to define, for a Reference Image *k*, an Objective Measure of Similarity to a Latent Image. The Diagram below describes the hierarchical structure of the 2,450 $z^j_{i,k}$ scores.

- **First, 50 medians over *i* values are computed, one median for each fixed Base Set Image *j*.**
- **Second, the Objective Measure of Similarity for Reference Image *k* is computed as the median of the 50 *j*-Medians.**

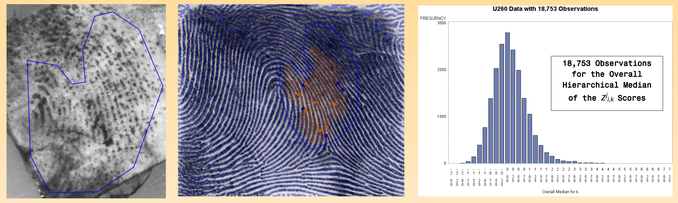




The hierarchical Medians from a large, randomly selected set of known non-mates to the Latent Image are used to create the statistical Model to predict random similarity to the Latent Image.

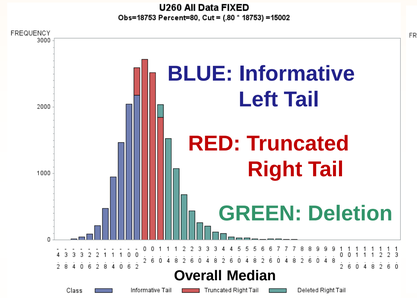


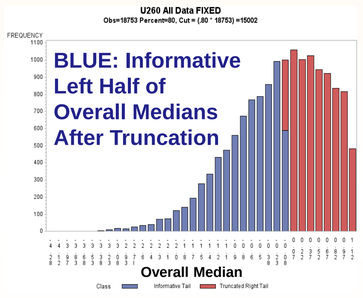
Example: U260 From The NIST SD27 Special Database



Overall Median Scores for a Random Sample of 18,753 non-Mates

The Left-Tail of the Overall Medians comprises the *Informative* Similarity Data concerning *strong* Level 2 similarity to the Latent Image. We use the Left Tail to define a powerful statistic that captures (competitive) Level 2 similarity to the Latent image. This statistic demonstrates a valid 'error' measurement. That is, the statistic fits a well-defined error distribution that we can credibly use to predict *rarity* of an 'outlying' observed Reference Image similarity. That is, we construct a well-defined *Null Distribution* that we use to predict the rarity of an outlying observed similarity. The Null Distribution predicts Level 2 similarity to the Latent Image for non-mate fingerprint images within a very large data base of fingerprint images.

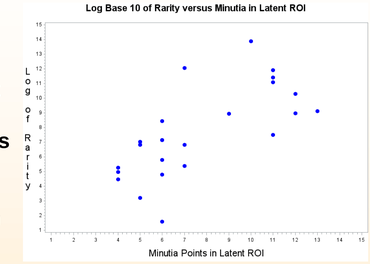




Rarity Predictions

The Objective Measure (Hierarchical Overall Median) calculated for the NIST SD27 ground truth mate image for Latent U260 is - 4.1789. Performing the same calculations on this ground truth score as was done to standardize the data for the Base Set only images, the resulting 'Standardized Score' for the ground truth image is -7.055414037. The p-value for -7.055414037 calculated from the Standard Normal Distribution is 8.604377E-13. The reciprocal of the p-value is 1,162,199,175,291. Therefore, the Predicted Rarity is 1 in 1,162,199,175,291. The base 10 logarithm of this Rarity is 12.0653 which gives the order of magnitude of the Rarity prediction. The graphic on the right is a plot of the Base 10 logarithm of the Rarity prediction vertically versus the number of minutia points matched between the Latent Image ROI and the reported ground truth mate image. Note the Log Rarity variability for fixed minutia values.

The prediction of Rarity is very robust regardless of the images selected as the Base Set or of the particular 20,000 non-mate images that are randomly selected for modeling and estimation of Rarity.



In particular, when the 50 non-mate images most similar to Latent U260 from the modeling described above were subsequently used as a Base Set for a second modeling, the resulting Log Rarity Prediction for the U260 true mate was 12.31.

Researchers are continuing efforts in coordination with statistical researchers from South Dakota State University as well as with LPEs to make Rarity modeling useful for LPE practice.