Detecting Level 3 Features in Fingerprints Using Support Vector Machines

Philip D. Wasserman
Guest Researcher
National Institute of Standards and Technology, Gaithersburg MD
cscience@comcast.net

Objectives

- Discuss the characteristics of fingerprint features.
- Propose a subset of Level 3 features.
- Describe the methods used to identify level 3 features using Support Vector Machines.
- Review characteristics of Support Vector Machines.
- Discuss typical features in the training and test sets.
- Present performance results.

A Preliminary Report

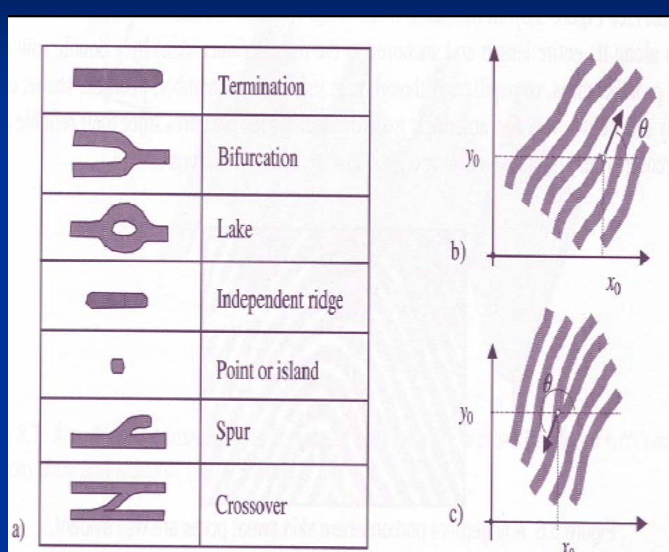
- This is a project in progress.
- Current results are based on a small data set with only a pore feature set collected from 500 dpi live-scanned images.
- Ultimate goal is to reliably detect several different level 3 features in latent, inked, and live-scanned fingerprints.

Strategy

- Difficult to determine how human fingerprint examiner makes decisions
  - Highly intuitive decisions
  - Expressing decisions as rules is probably impossible
- Instead, emulate examiner's decisions by training a learning machine
  - Capture expertise implicitly in examples
  - Train SVM (Support Vector Machine) to duplicate examiners observed behavior

Level 1 Features

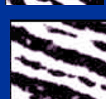
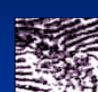
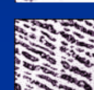
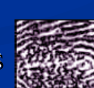
Level 2 Features



Level 3 Features

- In the broadest sense, level 3 features are any not classifiable as Level 1 and Level 2.
- There is no generally agreed upon definition of Level 3 features.
- A NIST working group is in the process of defining Level 3 features.
  - No conclusions as this is written

Some Level 3 Feature Candidates

- Pores 
- Ridge Shapes 
- Incipient Ridges 
- Scars 
- Warts 
- Creases 
- Deformations 

From: BIOMETRICS
Dr. Andrzej Drygajlo

Sweat Pore Chosen As Feature

- The sweat pore feature was selected for this first portion of the study by two criteria:
  - Usefulness to examiners
  - Detectability by Support Vector Machines
- Disadvantage: Sweat pores may not be visible
  - Ink and powder tends to fill pores
- Advantages
  - Numerous
    - 2700 per square inch (approx.)
  - Distinctive
    - Highly variable in:
      - Size: 88 to 220 microns
      - Spacing along ridge is random (9-18 pores/cm or ridge approx.)
      - In any position across ridge
      - Shape: round, oblong, tri

Examples of Sweat Pores at 500 dpi

Image Enhancement

- Conservative enhancement used to preserve information
  - Contrast and brightness enhancement by level adjustments
  - Sharpening (un-sharp mask)
- 500 dpi original image
  - Captured with solid-state fingerprint sensor

Image Enhancement Example

Original

Enhanced

Support Vector Machines

- Support Vector Machines (SVM)
  - Learning machines based on statistical learning theory
  - Trained by examples
  - Classifies previously unseen inputs
- Solid mathematical foundation in Vapnik Chervonenkis theory [Vapnik, 1995a][Smola, 2000]
- Maps training vectors into higher (possibly infinite) dimensional space
  - Using "kernel trick" all computation is done with dot products in low dimensional training vector space.
- All the following were once considered to be different classes of Artificial Neural Networks.
  - Radial Basis Function
  - Sigmoidal Multi layer Perceptron
  - Polynomial
  - Linear
  - Many others
- All the above have been shown to be special cases of an SVM

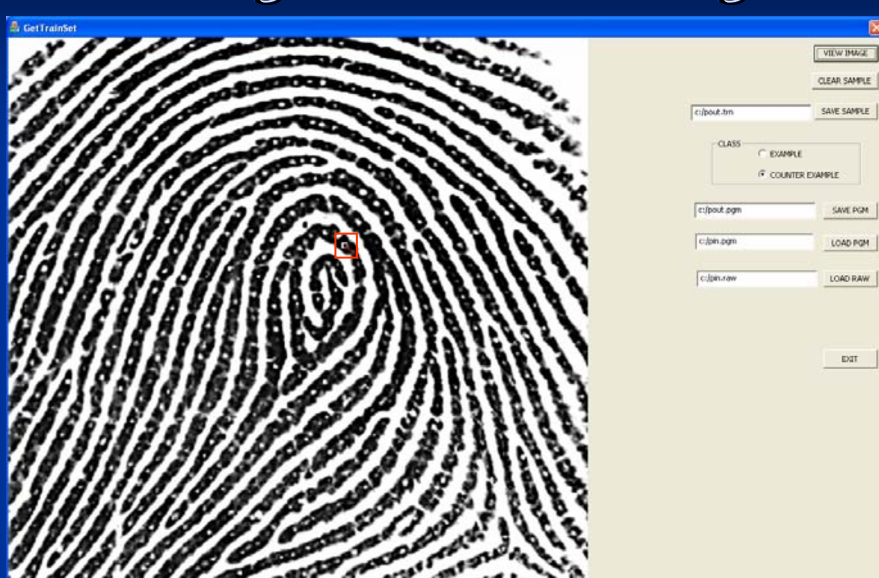
Training and Evaluation Methods

- Trained using SVM-light software
  - Courtesy of Thorsten Joachims [Joachims, 2002a] [Joachims, 2002c] [Klinkenberg, Joachims, 2000a] [Joachims, 2000b] Joachims, 1999a]
    - Available without charge at http://svmlight.joachims.org
  - Another version [CHANG 2001], LIBSVM, also available without charge
- Radial Basis Function Kernel was used
  - $K(x_{i,} x_j) = \exp(-\gamma \mid\mid x_i^T - xj \mid\mid^2)$
- Accuracy evaluated by leave-one-out method

Characteristics of SVMs

- Generalizes from training examples
- Constructs arbitrarily complicated, optimal, non-linear decision surfaces
- Every solution is global; no local minima
- Training is a conventional quadratic programming problem
  - Many different optimizers can be used
  - Specialized optimizers improve performance
- Training complexity is calculable
  - Cubic in number of support vectors
  - Support vectors are typically much fewer than training vectors
- Provides confidence level on decisions
- Accuracy estimate is produced with little additional computation
  - Leave-one-out cross validation

Training Set Selection Program



Training Set Example Selection

- Select correct classification
- Click on an image point
  - Computer program determines training vector components
- Save as training vector
- Components currently based on:
  - Central intensity pattern
  - Radial intensity pattern
- Ridge slope is estimated
  - Will be used for other level 3 features

Estimating Accuracy

- **Cross-validation**, the basic procedure
  - Separate data set into two sub-sets
    - Training set
    - Test set
  - Train classifier on Training Set
  - Measure accuracy on Test Set
- **n-set Cross-validation** improves accuracy
  1. Separate data into n sub-sets
  2. Train on n - 1 subsets, reserving one subset
  3. Measure accuracy on reserved sub-set
  4. Repeat 2 through 3 for all sub-sets
- **Leave-One-Out method**, limit of n-set method, still more accurate
  1. Train on all but 1 example
  2. Classify that example
  3. Repeat steps 1 and 2 for all examples
  4. Calculate error rate as: number of errors / number of training examples
  - Impractical for many types of classifiers: requires re-training for each ex
- SVM performs Leave-One-Out accuracy estimation with ... extra computation

Training Process

- Training set size: 483 samples
- CPU time for training: < .01 seconds
- CPU time for classification: < .01 seconds
- CPU time for leave-one-out cross-validation: .03 seconds.

Estimated Accuracy by Leave-One-Out Method

- No errors found by cross-validation
- Recall: 100% (TAR x 100)
  - Percentage of pores correctly classified (221 pores; 221 correctly classified)
- Precision: 100%
  - Percentage of samples classified as a pores that actually are pores
- Overall accuracy: 100%
  - 483 samples; 483 correctly classified, 0 misclassified
  - 262 pores; 262 correctly classified. 0 misclassified
  - 221 non-pores, all correctly classified

Estimated Accuracy

- TAR (True Accept Rate) = 1.0
- FAR (False Accept Rate) = 0.0

Discussion

- Results are suggestive, but not conclusive
- Sample size is too small to make useful accuracy estimates
  - Because there were no errors, with 95% confidence, the error rate is known to be less than 0.621% (3/sample size) (Rule of 3) [Gamassi, 2004] [Louis 1981] [Jovanovic 1997] [Wayman 2000]
- Errors are too few in number
  - "To be 90% confident that the true error rate is within ± 30% of the observed error rate, there must be at least 30 errors." [Gamassi, 2004] [Doddington, 2000] (Rule of 30)

Future Research

- Expand and evaluate pore training set
- Scan image for pores and display detection regions
- Calculate ROC using confidence levels
- Evaluate performance on other level 3 features
- Expand study to include 1000 dpi fingerprints
- Scan latent fingerprint images and display detection regions

References

[Chang 2001] Chang, C. and Lin, C: *LIBSVM: a library for support vector machines.* http://www.csie.ntu.edu.tw/~cjlin/libsvm/

[Gamassi, 2004] M. Gamassi, et.al, *Accuracy and Performance of Biometric Systems,* IMTC 2004- Instrumentation and Measurement Technology Conference, Como, Italy 18-20 May 2004, http://www.elet.polimi.it/upload/fscotti/pdf/Scotti14.pdf

[Joachims, 2002a] Thorsten Joachims, *Learning to Classify Text Using Support Vector Machines*. Dissertation, Kluwer, 2002.

[Joachims, 2002c] T. Joachims, *Optimizing Search Engines Using Clickthrough Data*, Proceedings of the ACM Conference on Knowledge Discovery and Data Mining (KDD), ACM, 2002. Online [Postscript] [PDF]

Joachims, 2000b] T. Joachims, *Estimating the Generalization Performance of a SVM Efficiently*. Proceedings of the [illegible] onal Conference on Machine Learning, Morgan Kaufman, 2000 Online [Postscript (gz)] [PDF]

[Joachims, 1999a] T. Joachims, 11 in: *Making large-Scale SVM Learning Practic* [illegible] vances in Kernel [illegible] Vector Learning, B. Schölkopf and C. Burges and A. Smola (ed.), MIT Press, 199[illegible] Online [Postscript (g[illegible]

[Joachims, 1999c] Thorsten Joachims, *Transductive Inference for Text Classification using Support Vector Ma*[illegible] tional Conference on Machine Learning (ICML), 1999. Online [Postscript (gz)] [PDF]

References (2)

[Joachims, 1998a] T. Joachims, *Text Categorization with Support Vector Machines: Learning with Many Relevant Features*. Proceedings of the European Conference on Machine Learning, Springer, 1998. Online [Postscript (gz)] [PDF][

[Klinkenberg, Joachims, 2000a] R. Klinkenberg and T. Joachims, *Detecting Concept Drift with Support Vector Machines*. Proceedings of the Seventeenth International Conference on Machine Learning (ICML), Morgan Kaufmann, 2000.
Online [Postscript (gz)] [PDF (gz)]

[Morik et al., 1999a] K. Morik, P. Brockhausen, and T. Joachims, *Combining statistical learning with a k[illegible]edge-based approach - A case study in intensive care monitoring*. Proc. 16th Int'l Conf. on Machine Lea[illegible]ML-99), 1999.

[Smola 2000] Alexander J. Smola, P.Bartlett, B. Scholkopf, D. S[illegible]urmans, *Advance [illegible]gin Classifiers, MIT Press, Cambridge Massachusetts, 2000.*

[Scholkopf 2000] Bernhard Scholkopf, *Statistical Learning and Keernel Methods*, Tech. Rep. MS[illegible] Microsoft Research Limited,

[Vapnik, 1995a] Vladimir N. Vapnik, *The Nature of Statistical Learning Theory*. Springer, 199[illegible]

[Vapnik, 1998] V. Vapnik, Statistical Learning Theory, Wiley, NY, [illegible]

References (3)

- [Louis 1981] A. Louis, *Confidence intervals for a binomial parameter after observing no successes"*, The American Statistician, 1981, 35(3), 154
- [Jovanovic 1997] D. Jovanovic, P. S. Levy, *A look at the rule of three"*, The American Statistician, 1997, 51(2), 137-139
- [Wayman 2000] L. Wayman, *Technical testing and evaluation of biometric identification devices*, Biometrics, Personal identification in networked society, edited by A. K. Jain et al., Kluwer, 2000, 345-368
- [Doddington, 2000] G. R. Doddington, et.al., *The NIST speaker recognition evaluation: Overview methodology, systems, results, perspective*, Speech Communication, 2000, 31(2-3), 225-254.