Online Handwriting Recognition Technology
- State of the Art

Jingtao Wang
Group for User Interface Research
jingtaow@cs.berkeley.edu
Feb 3, 2003
My Background

• Designed and implemented the IBM embedded large vocabulary, multi-lingual (English, simplified and traditional Chinese, Japanese) online handwriting recognizer for mobile devices.

• Team lead - R&D of the IBM full page online handwriting recognizer for Asian languages (simplified Chinese, traditional Chinese and Japanese) on desktop/laptop computers (shipped in ThinkScribe, a.k.a CrossPad, and ThinkPad TransNote) at IBM China Research Lab.
The Proliferation of Handheld Devices
Agenda

• A Brief History
• Handwriting Recognition – Categorization and Related Applications
• Review of Classic On-line Handwriting Recognition Algorithms
• Embedded Handwriting Recognition Technology
• Pen Interfaces for Recognizers
• Conclusions
• Demo, Q&A
A Brief History

• 1914 Hyman Eli Goldberg, U.S. Patent 1,117,184, On-line recognition of hand-written numerals to control a machine in real-time.


• 1957 T. L. Dimond's stylator - the first on-line handwriting recognizer prototype

• Newton, Pen Computer, Palm, Crosspad, Thinkpad TransNote and TabletPC
Handwriting Recognition - Categorization and Applications

• Printed Character Recognition (OCR)
  – Relatively mature these days, key challenges – layout analysis, fonts recovery, robust recognition for low quality, low resolution input
  – Major applications: digital library, document management, digitize legacy paper publications, portable dictionary

• Handwritten Character Recognition
  – Online HWR (with temporal info, i.e. stroke trace info)
    • Character, Word, Sentence Level
    • Text entry, Gesture Command & Control, Annotation & Retrieval
  – Offline HWR (using raster image as input, no temporal info)
    • A rich test bed for pattern recognition research
    • Major Usage - Postal automation, census automation, automatic form processing
Beyond Individual Characters – Word and Sentence Level Recognizer

• Build on top of the character recognizer
• General strategy:
  – Over segmentation
  – Calling a character/component recognizer, getting a list of candidates with scores
  – Applying geometry spatial information (size, component gap) and language information (dictionary, language model etc) to each sub path
  – Using hypothesis search (Dynamic Programming, A*, Beam Search etc) to determine the best possible path
Challenges in Online Handwriting Recognition

• Character set/Dictionary Size
  – 6763 Characters in GB2312 (Simplified Chinese)
  – 13K Characters in BIG5 (Traditional Chinese)
  – 5k Characters in JIS (Japanese)

• Stroke number variations
  – Cursive Writing Styles/Broken strokes/duplicate strokes/omitted components

• Stroke order variations

• Limited memory and Computing Power on Small Devices
A Generic Recognition Workflow

1. Preprocessing
2. Feature Extraction
3. Coarse Classification
4. Detailed Matching
5. Post Processing
Common Recognition Algorithms

- Knowledge/Heuristic Based
  - Decision Tree
  - Hand coded (Fuzzy logic :-)

- Global Feature Vector Matching
  - Parameter or Non-parameter classifier (KNN)
  - Neural Network (MLP, LVQ, PNN..)
  - Support Vector Machine (SVM)

- Structure Based Methods
  - Dynamic Programming/DTW
  - Graph Matching (e.g. DAG matching)
  - Hidden Markov Model
  - Time Delay Neural Network (TDNN)
  - Heuristics based stroke correspondence
Global Feature Vector Matching

• Classifier itself is usually not the most influential factor on accuracy !(less than 10%)
  – However, different classifiers do have different pros and cons in speed, charset switching, largest supported class number, training difficulty.

• Feature extraction/selection usually determines the upper bound of performance

• Common pitfalls
  – Extract multiple feature types and use them directly
  – Extract too many dimensions of feature
Structure Based Methods

• Construct sub-components of handwriting, then use some algorithms to solve the component correspondence problem between the input and the prototypes
• Commonly used sub-components
  – Loop points, Crisp points, Cross points, Curvature points (usually for Western languages)
  – Stroke, Stroke segment (linear approximation of strokes)
  – Start, end points, equal distance/equal number sampling points
  – Radicals (usually for Asian Languages)
Examples of a HMM Based Recognizer

Original Input → After preprocessing, Segmentation and Feature Points Detection → Matching/Correspondence Results after Backtracing
An Example of Stroke Correspondence

Template Input

\[ D(x) = D_{1b} + D_{2A} + D_{3C} + D_{4D} \]

\[ D = \min\{ D(x) \} \]
Commonly Used Recognizer Training Algorithms

- Competitive Learning (SOM, LVQ)
- Linear Programming (SVM)
- Back-Propagation (MLP)
- Expectation-Maximization (HMM, GMM)
- Boosting
- Clustering Algorithms
- PCA, LDA, FA and their variations
- Nonlinear function optimization methods
- Randomized Algorithms (MCMC, Simulated Annealing)
- Brute Force method
Trends in Current HWR research

• Using multiple classifiers to achieve robustness to handle input variations – Classifier Combination
  • Classifier voting, Confusion matrix, Decision classifier, Confidence value transformation
• Obtaining the parameters of a recognizer from large training data set, not from heuristics
• Using transformations to capture delicate shape variations (e.g. SAT)
• Leveraging powerful but time-consuming training algorithms
• User Adaptation Algorithm (Writer dependent recognition)
Major Player on This Arena (Desktop)

- **English**
  - IBM - Used in CrossPad, ThinkPad TransNote
  - Microsoft - In OfficeXP, TabletPC (licensed part of source codes from Paragraph)
  - Motorola
  - Paragraph

- **Chinese/Japanese**
  - IBM
  - Motorola
  - Hanwang - Licensed to Microsoft for TabletPC
  - PenPower
  - Wintone
  - FineArts
Players in this Area
(Embedded/Lightweight)

• English
  – ART - ART Recognizer
  – CIT - Jot
  – IBM - Derived from the Multi-lingual version
  – Microsoft - Transcriber (Licensed version of Calligrapher) & Self developed single character recognizer
  – Motorola
  – Paragraph - Calligrapher

• Chinese/Japanese
  – FineArt – GoGo Pen
  – Hanwang - more than 70% PDA market share in mainland China
  – IBM Embedded HWR
  – Motorola Lexicus - DragonPen
  – PenPower – the most influential vendor in Taiwan
Lessoned Learned

• Collect a large enough data set (at least 200 * charset size) before the start of your research
• Trade-offs are necessary to compromise time & space
• No one-size-fits-all recognition algorithm, classifier combination is usually an effective solution
• Recognition algorithm is not simply a combinational optimization problem - must adapt your recognition model to human cognition model
• Train your recognizer, not code your recognizer!
• Beware of the curse of dimension (CoD)
UI Issues for Embedded Recognizers

• Character Segmentation
  – Uni-stroke
  – Time threshold
  – Spatial Information/ "Ping-pong" window
  – Language Information

• Character Set switching/identification
  – Shift Key/Shift Gesture
  – Specific writing Areas for different character sets
  – Showing guide line in the writing area

• Pen Gesture integration
Prototype Recognizer Interfaces at IBM
Sentence/Page Level Handwriting Recognition

• Support more advanced devices: TransNote, CrossPad, Whiteboard, WebPad, etc.
• Geometry based segmentation combined with a language model

保持咖啡新鲜请储存在干燥的地方
性格を重視する人が多いようだ
133 M Hz 处理器
Pen/Handwriting Technology - Hardware

- **ThinkScribe/CrossPad**
  - Writing on paper
  - Hardware: digitizer, display,
    - memory cpu, IR/Serial, AAA battery
  - Software and API

- **ThinkPad TransNote**
  - 10.4" pivoting touch screen
  - Input by keyboard, touch screen, or handwriting
  - 600-MHz Pentium III processor
  - 10g hard disk
  - folding portfolio

- **InkManager Pro**
  - Ink storage
  - Index of keyword, todos, messages
  - PIM updater connection with lotus notes, organizer, MS-outlook, etc.
  - InkXfer
Collaborative Input via Laptop and PDA

"Handwriting On the Move" - Input solutions for mobile laptop users
Demo
Questions ???

[Image: A meeting scene with a person addressing two others, one of whom is writing with a pen.]