Online handwriting recognition: past, present and future

Masaki Nakagawa

Tokyo University of Agriculture and Technology

ICFHR 2016

October 23-26, 2016

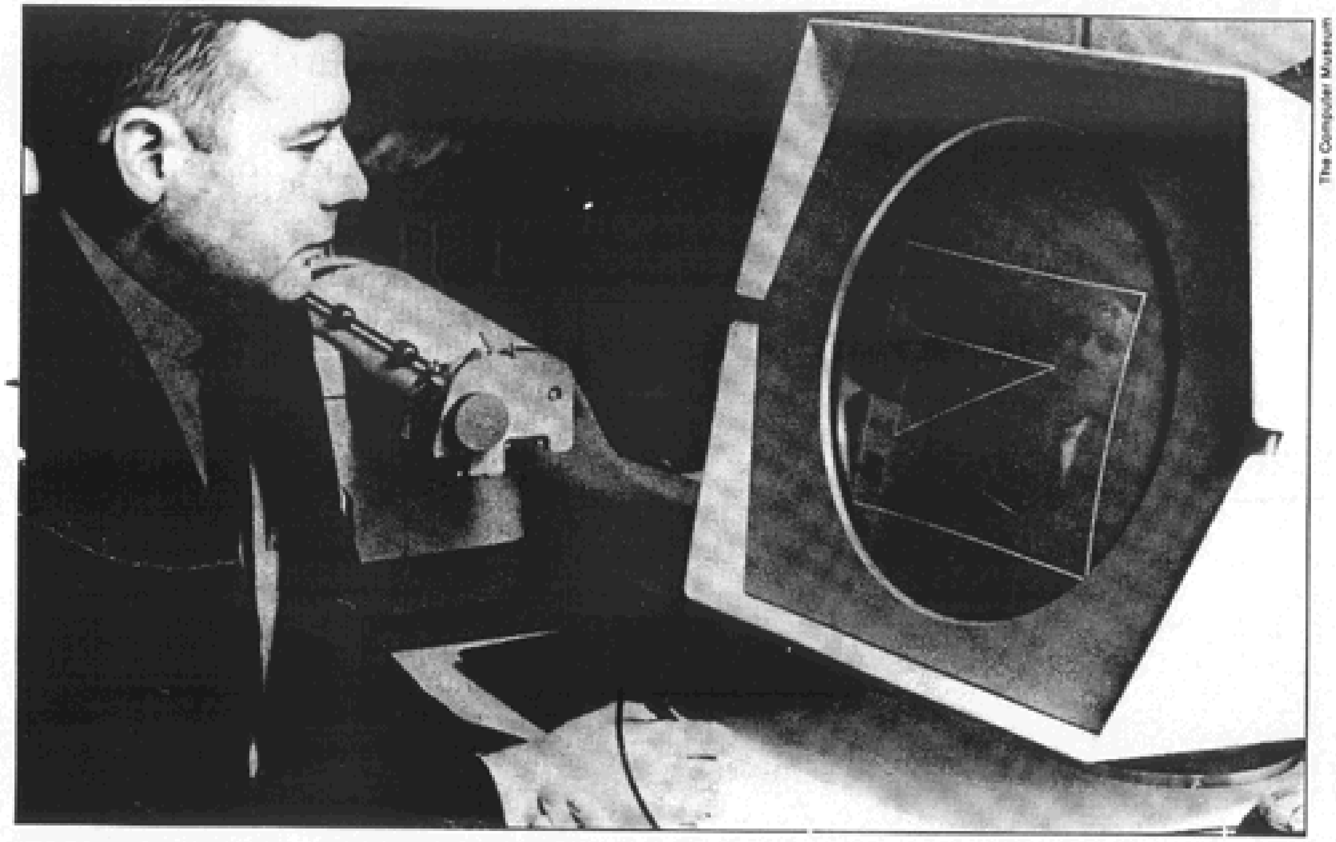
Table of Contents

Brief history of online handwriting recognition

Rand Tablet in 1964

http://itsexplore.blogspot.jp/2013/07/history-of-tablet-pcs-pictorial.html

Sutherland and Sketchpad

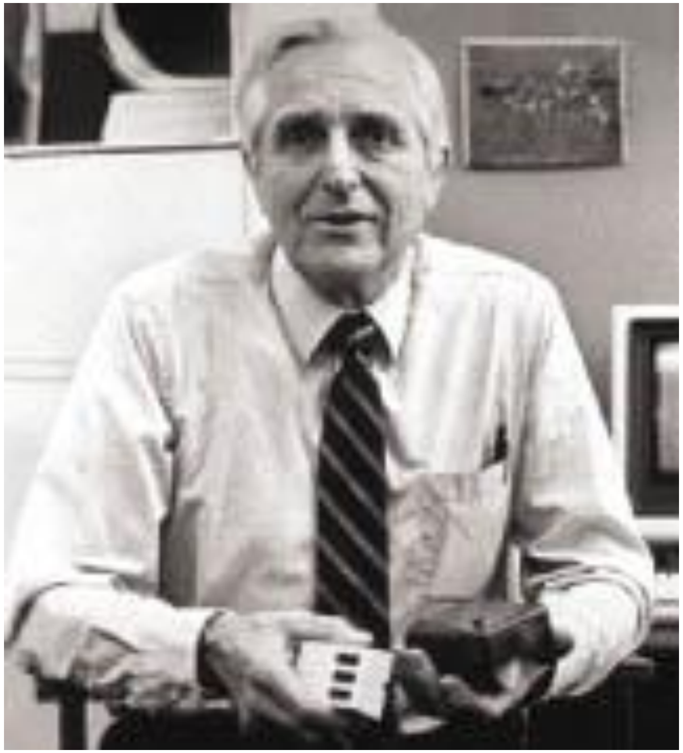


[1] Sketchpad (left), created in 1962 by Ivan Sutherland at Massachusetts Institute of Technology's Lincoln Laboratory in Lexington, is considered the first computer with a windowing interface. More than 1200 of the experimental Alto (right), developed in 1973 by the Xerox Palo Alto Research Center, were distributed to test its windows, menus, and mouse.

ing it at fast speeds—to imp
and speed up larger movem
software incorporates this i
The mouse, still experime
1971. Several members of I
newly established PARC, w

The Computer Museum

Mouse by D. Engelbert



The first mouse made in 1964.

http://www.geek.com/hardware/40-years-of-the-mouse-a-guided-tour-1369967/
http://corefee.blogspot.jp/2016/07/the-first-ever-computer-mouse-ever.html

1965～80

- **Started on RAND tablet from:**
  - G. F. Groner, "Real-Time Recognition of Handprinted Text, " pp. 591-601, Proc. FJSS (1966)
- **Small memory space (64 Kilo bytes) and Slow CPU (**Mini computer and μ processor **～1 Mips)**
- **Heuristics, Decision tree, Syntactic, Structural, Grammar, Primitives and Confusion matrix, Fourier coefficients, Statistical for a small category set, Feature points, etc.**
- **Expected to establish the technology but realized how noisy and distorted human writing patterns are.**
- **Long journey started to relax writing constraints.**

1980～2000

- **Moore's Law**
  - CPU of several Mips and memory of Mbytes is common
- **Development of LCD display: Tablet →**
  LCD display integrated tablet
  - Direct pointing and
    direct manipulation
- **Innovative products**
  Handwriting Input Word Processor from Panasonic in 1984,
  Sony Palmtop**, Apple Newton**,
  Sharp Zaurus in early 1990's.

Parallax free tablet from Hitachi Lab.

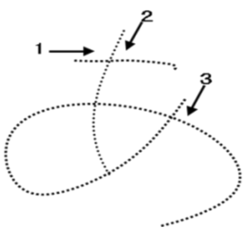
- **Platforms**
  - PenPoint by GO
  - Windows for Pen by Microsoft
- **But, no break until 2010's**

Recognition techniques around 2000

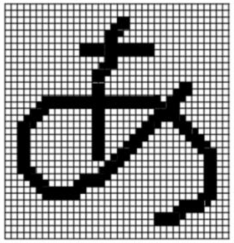
- **Statistic, Stochastic**
  - DP -> HMM, Time-delayed NN, Bayesian Net, MRF, etc.
- **Shape context**
  - S. Jaeger , S. Manke , A. Waibel: “Npen++: An On-Line Handwriting Recognition System,” pp. 249-260, Proc. 7th IWFHR (2000).
- **Combination of online and offline methods**
  - H. Tanaka et al.: “Hybrid Pen-input Character Recognition System based on Integration of On-line and Off-line Recognition,” pp.209-212, Proc. 5th ICDAR (1999).
- **Language model or Linguistic context**
- **Large pattern DB (UNIPEN and TUAT)**

On-line and Off-line Recognitions

- On-line recognition usually works on an on-line pattern: a time sequence of pen-tip coordinates (digital ink).
- Off-line recognition works on a bit-map image.
- Off-line recognition can be also applied by transforming digital ink to a bit-map.
- Both on-line and off-line methods can be employed.



On-line pattern



Off-line pattern

advantage & disad. / **Method**	**advantage**	**disadvantage**
On-line recognition	**robust to stroke connection and distortion**	**not robust to wrong stroke order and duplicated strokes**
Off-line recognition	**Free from wrong stroke order and duplicated strokes**	**not robust to stroke connection and distortion**

No break until beginning of 21st C

- Market did not expand. Companies shrank R & D or even retreated.
- Crash of Bubble Economy and IT bubbles.
- Rubbish stock? But, it might be a blue chip in future. University laboratory can continue.

When you have a big task or difficult task, you can see your mental capacity or toughness.

When you are in favorable circumstance or in unfavorable circumstance, you can see your prudence, mental tolerance or stability.

…

in 呻吟語 by 呂 新吾

Recent 10 Years

Pen and paper devices (storing digital ink without PC)

Conventional LCD integrated tablets

Smart phones and Tablets, …

From Web's of Anoto. Fujitsu, Nintendo and Apple

Research progresses

- **Deep Neural Networks and Deep Learning**

  - **HMM → RNN, LSTM, Bidirectional LSTM, Multi-Dimensional LSTM, …**

  - **QDF, Modified QDF → CNN, …**

Online handwriting Databases

- UNIPEN (English)
- TUAT (Japanese)
- IRONOFF (French and English)
- CASIA (Chinese)
- SCUT (Chinese)
- IAM-OnDB (English)
- IBM-UB (English)
- AltecOnDB (Arabic)
- VN-OnDB (Vietnamese)
- CROHME for (Math)

- Any other? (Please teach me!)

Survey papers

- C-L. Liu, S. Jaeger and M. Nakagawa, “On-Line Recognition of Chinese Characters: the State of the Art,” *IEEE Trans. PAMI*, **26**, 2, pp.198-213, 2004.
- S. Jaeger, C-L Liu and M. Nakagawa, “The state of the art in Japanese on-line handwriting recognition compared to techniques in western handwriting recognition,” *IJDAR*, **6**, 2, pp.75-88, 2003.
- R. Plamondon and S. Srihari, “On-Line and Off-Line Handwriting Recognition: a Comprehensive Survey,” *IEEE Trans. PAMI*, **22**, 1, pp.63-82, 2000.

- **Up to 1990**

- C. C. Tappert , C. Y. Suen and T. Wakahara, “The State of the Art in On-Line Handwriting Recognition,” *IEEE, Trans. PAMI*, **12**, 8, pp.787-808, 1990.
- M. Nakagawa, “Non-Keyboard Input of Japanese Text—On-Line Recognition of Handwritten Characters as the Most Hopeful Approach,” *J. Information Processing*, **13**, 1 pp.15-34, 1990.

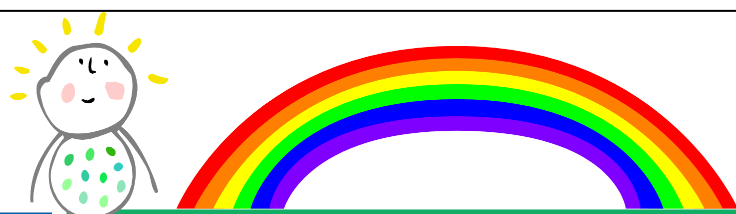
Potentials of handwriting input and user interface

Potentials of Handwriting-based UI

- **Handwriting itself is universal**, although character recognition is language dependent.
  - **European, American, Asian, Arabic, etc. or math, even pictorial languages can be expressed using a single pen.**
- **One can express one's thinking most easily with a pen.**
- **Thinking is not interrupted by the actions for writing.**
- **Thinking and writing form feedback loop to grow one's idea.**

  ⇨ Creative work rather than labor-intensive tasks.

  ⇔ Speech is transient. We cannot interact with spoken words.
- **Much richer information rather than just codes.**
- **One can express graphics inexpressible by phonetics.**
- **Scalable for small, medium and large surfaces.**
  - **PDA, tablet, interactive board, …**
- **Direct pointing and manipulation: Indirect P. & M. → Indirect P. and direct M. → Direct P. & M.**



Creative UI,
UI friendly to children and elder people,
UI for peoples who use large character sets.

v.s. Speech Interface

- **Speech is transient**

Speech and sound can be recorded, but
will you use speech and voice recorder rather than pen and paper to solve mathematical equations?
On the other hand, you can interact with your handwriting.
Write down your idea, interact with it and extend it.

- **Speech is not effective to express diagrams**

Imagine to convey math expressions, diagrams, maps, etc. by phone.
On the other hand, you can write them down.

- **Speech is exclusive**

Multiple people cannot talk at the same time. Noisy and recognition rate is damaged.
On the other hand, they can write at the same time.

But, handwriting is not exclusive to speech. Both speech and handwriting compensate for each other.

Requirements for Recognizer

- **High recognition rates (top N best rate)**
- **Robustness to distortions or deformations**
- **Order variations, delayed or duplicated strokes**
  - Handwriting -> Haandwriiting (in speech)
  - Handwriting -> Hnadwirtign -> Hnaddwwirtiggn (in writing)
- **High speed**
- **Low memory space**
- **Last resort**

➢ You must be able to input a character as far as you write it neatly even if stroke order is wrong and some strokes are repeated.

➢ Dependence on stroke order must be avoided.

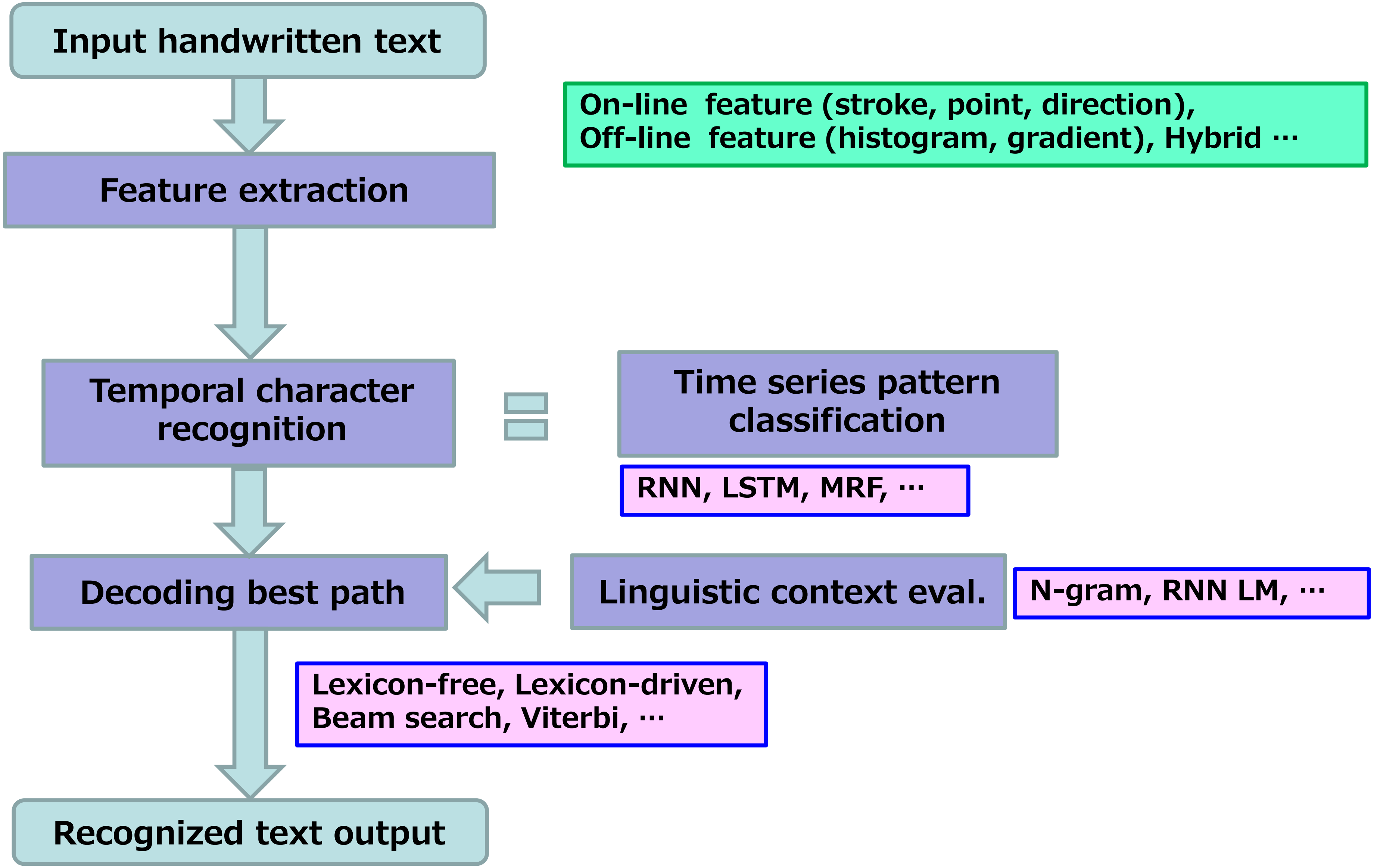
- **Transparency**

➢ not to lose **user confidence**, and to afford **user adaptation to the system**.

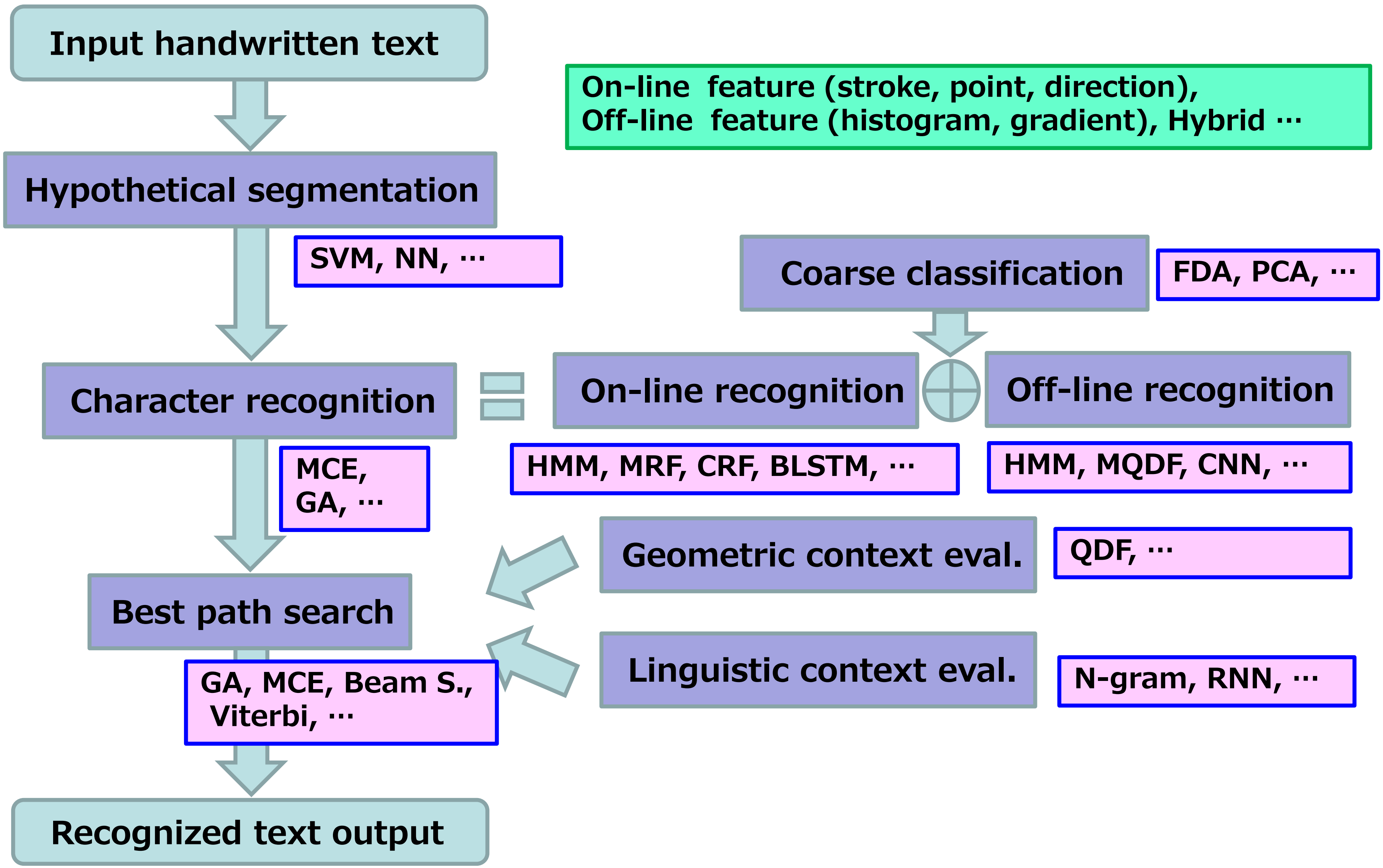
- If not, users verify recognition results every time. Damage the merit of pen interface.
- **Machine adaptation to individual user**

Architecture of recent online handwriting recognition systems

Architecture of On-line Text Recognizer (Segmentation free)

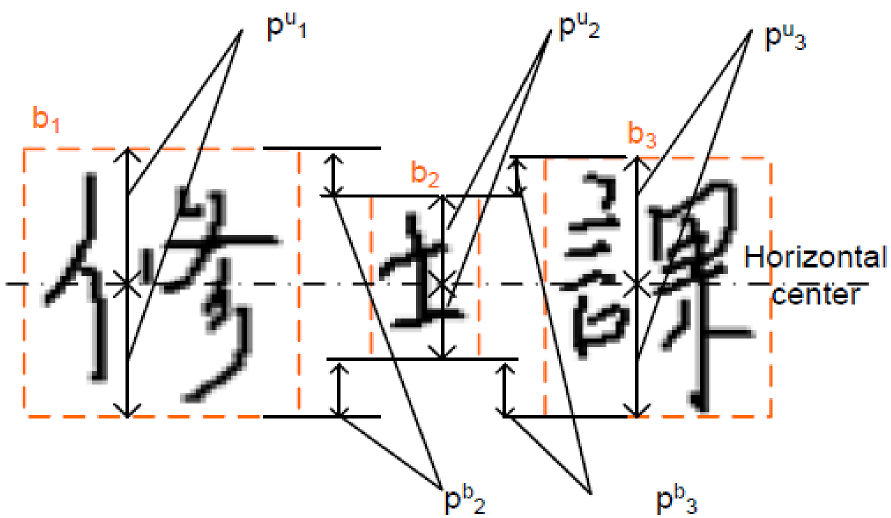


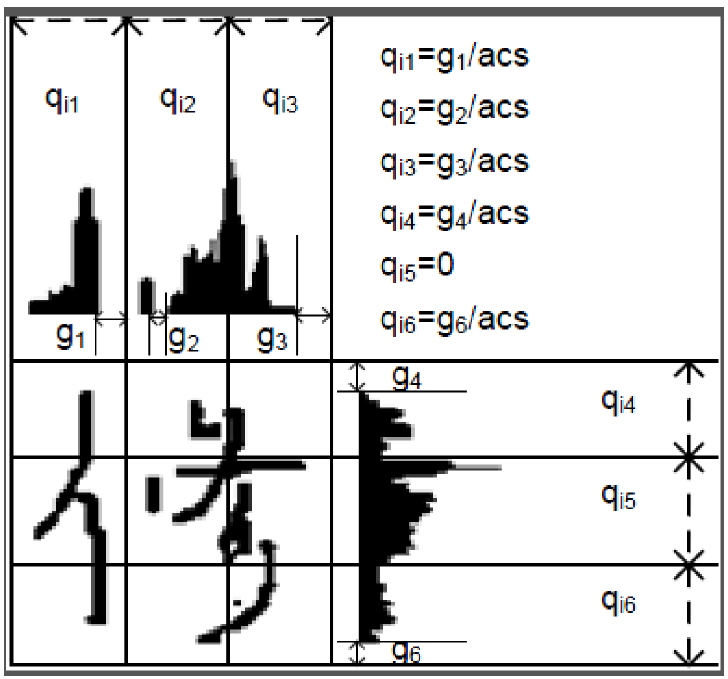
Architecture of On-line Text Recognizer (Over-segmentation)



Evaluation Function

$$
\begin{aligned}
f(X,C) = {} & \boxed{[\lambda_{11}+\lambda_{12}(k_i-1)]\log p(c_i \mid c_{i-2}c_{i-1})} && \Rightarrow \text{Linguistic context} \\
& \boxed{\begin{aligned}&+[\lambda_{21}+\lambda_{22}(k_i-1)]\log p(b_i \mid c_i)+[\lambda_{31}+\lambda_{32}(k_i-1)]\log p(q_i \mid c_i)\\ &+[\lambda_{41}+\lambda_{42}(k_i-1)]\log p(p_i^u \mid c_i)+[\lambda_{61}+\lambda_{62}(k_i-1)]\log p(p_i^b \mid c_{i-1}c_i)\end{aligned}} && \Rightarrow \text{Geometric context} \\
& \boxed{+[\lambda_{51}+\lambda_{52}(k_i-1)]\log p(x_i \mid c_i)} && \Rightarrow \text{Evaluated by character recognizer} \\
& \boxed{+\lambda_{71}\log p(g_{j_i} \mid Sb)+\lambda_{72}\sum_{j=j_i+1}^{j_i+k_i+1}(k_i-1)\log p(g_{j_i} \mid Sw)} && \Rightarrow \text{Segmentation Point Evaluation}
\end{aligned}
$$



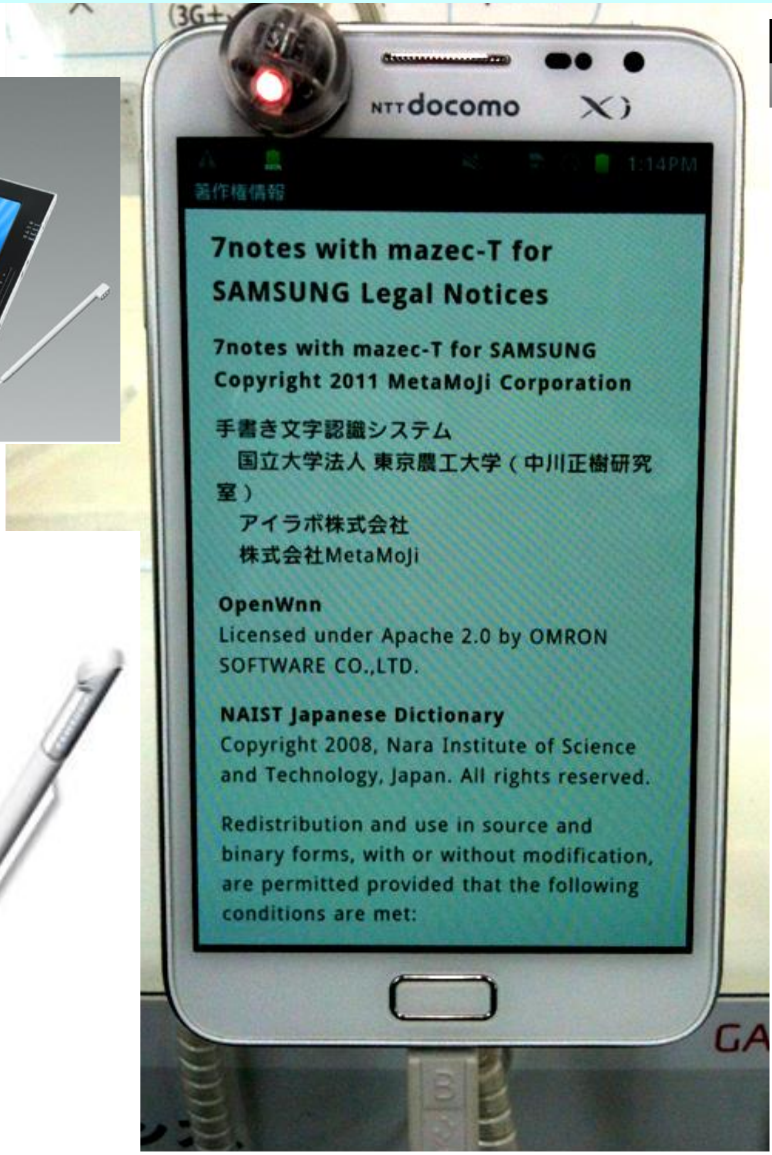


Bilan Zhu, Xiang-Dong Zhou, Cheng-Lin Liu and Masaki Nakagawa: A Robust Model for On-line Handwritten Japanese Text Recognition, *International Journal on Document Analysis and Recognition (IJDAR)*, Vol. 13, No. 2, pp.121-131 (2010)

Jinfeng Gao, Bilan Zhu, Masaki Nakagawa: Development of a Robust and Compact On-line Handwritten Japanese Text Recognizer for Hand-held Devices, *IEICE Trans. on Inf. & Syst.*, Vol. E96-D, No.4, pp.927-938 (2013.4).

Handwriting input to mobile devices

Preinstalled in Smartphone & Tablet




著作権情報

7notes with mazec-T for SAMSUNG Legal Notices

7notes with mazec-T for SAMSUNG Copyright 2011 MetaMoJi Corporation

手書き文字認識システム
　国立大学法人 東京農工大学（中川正樹研究室）
　アイラボ株式会社
　株式会社MetaMoJi

OpenWnn
Licensed under Apache 2.0 by OMRON SOFTWARE CO.,LTD.

NAIST Japanese Dictionary
Copyright 2008, Nara Institute of Science and Technology, Japan. All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

Redistributions of source code must retain the

Worldwide

- More than 10% people use handwriting input in China, and this number is increasing.
- Myscript provides online handwriting recognition for 64 languages.
- Google provides online handwriting recognition for 97 languages.
- …

Overlaid Characters Recognition

- Hand-held devices have become popular
  - Smart phones, mini tablet PCs, smart watches



Writing area is too small to write text.

- To input text messages by overlaid handwriting

Publications

- **Early works in Japanese**

[1] H. Shimodaira, T. Sudo, M. Nakai, and S. Sagayama, "On-line overlaid handwriting recognition based on substroke HMMs," Proc. 7th ICDAR, Edinburgh, Scotland, pp.1043–1047, 2003.
[2] Y. Tonouchi and A. Kawamura, "Text input system using online overlapped handwriting recognition for mobile devices," Proc. 9th ICDAR, Curitiba, Brazil, pp.754-758, 2007 (just for phonetic characters in Japanese).

- **in English**

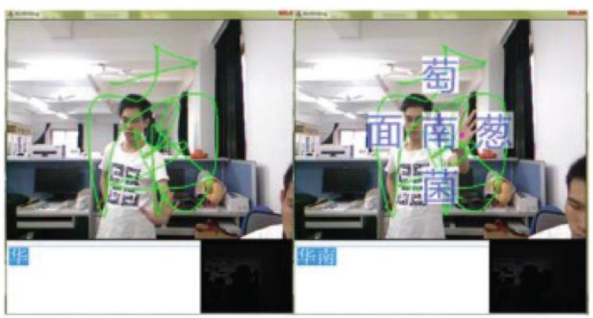
[1] A. Bharath and S. Madhvanath, "FreePad: a novel handwriting-based text input for pen and touch interfaces," Proc. 13th IUI, ACM, Canary Island, Spain, pp.297-300, 2008.
[2] W. Kienzle and K. Hinckley, "Writing handwritten messages on a small touchscreen," Proc. 15th Mobile HCI, Germany, pp. 179-182, 2013.

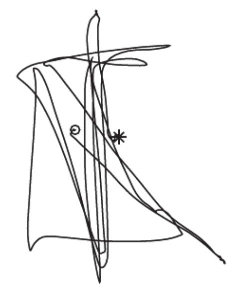
- **in Chinese**

[1] Y. Zou, Y. Liu, Y. Liu and K. Wang, "Overlapped handwriting input on mobile phones," Proc. 11th ICDAR, Beijing, China, pp.369-373, 2011.
[2] X. Wan, C. Liu and Y. Zou, "On-line Chinese character recognition system for overlapping samples," Proc. 11th ICDAR, Beijing, China, pp.799-803, 2011.
[3] Y.F. Lv, L.L. Huang, D.H. Wang and C.L. Liu, "Learning-based candidate segmentation scoring for real-time recognition of online overlaid Chinese handwriting," Proc. 12th ICDAR, Washington, DC, pp.74-78, 2013.

Air-writing as Extension

- Write by finger in the Air (Air-writing)
  - Devices providing a natural way to interact with computers
    - Microsoft Kinect sensor or Leap motion
  - Research works
    - Write Chinese characters in the air [1,2]
    - Write overlaid English characters only for uppercase with the predefined stroke order [3]



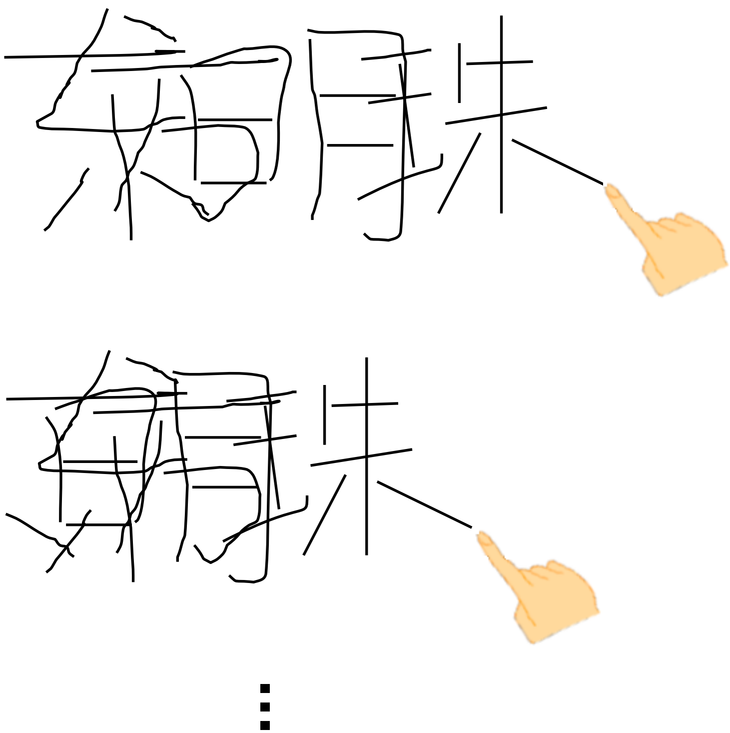


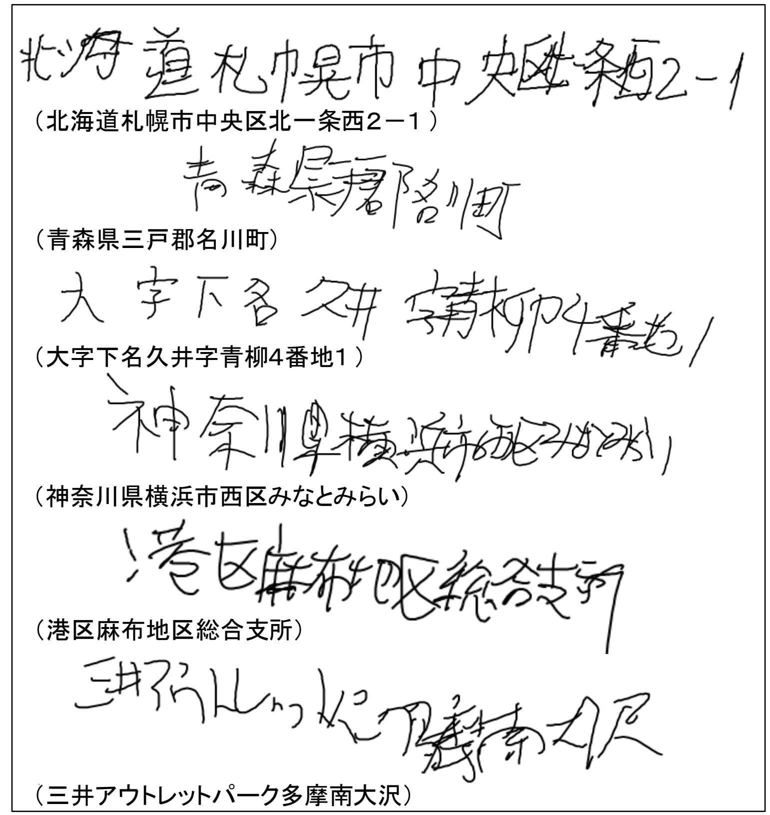
(Text: TITL)

[1] Z. Ye, X. Zhang, L. Jin, Z. Feng, S. Xu, “Finger-writing-in-the-air system using Kinect sensor,” in Proc. IEEE ICME2013, San Jose, USA, 2013.
[2] X. Zhang, Z. Ye, L. Jin, Z. Feng, “A new writing experience: finger writing in the Air using a Kinect sensor,” IEEE on Multimedia, vol. 20, no. 4, pp.85-93, 2013.
[3] M. Chen, G. AlRegib and B.H. Juang, “Air-writing recognition—Part II: Detection and Recognition of writing activity in continuous stream of motion data,” IEEE Trans. Human-Machine Systems, vol. 46, no. 3, pp. 436-444, 2016.

Character-position-free Handwriting

- Writing without **wrist or elbow support and without** visual feedback.

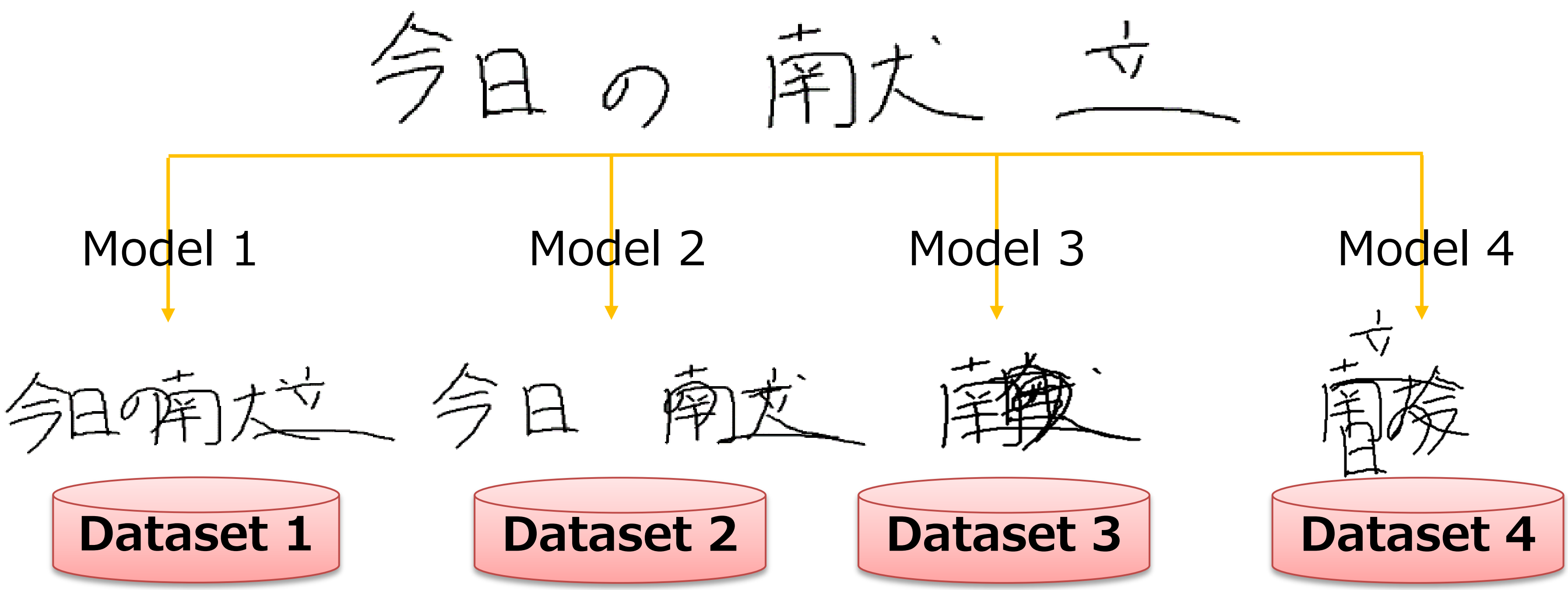




- **Recognized surprisingly well due to context without character pitch information.**

Character-position-free Handwriting

- Database
  - Using **Kondate database**
  - Make 4 models to produce character-position-free handwritten text patterns



0.5 to 1.0 spacing

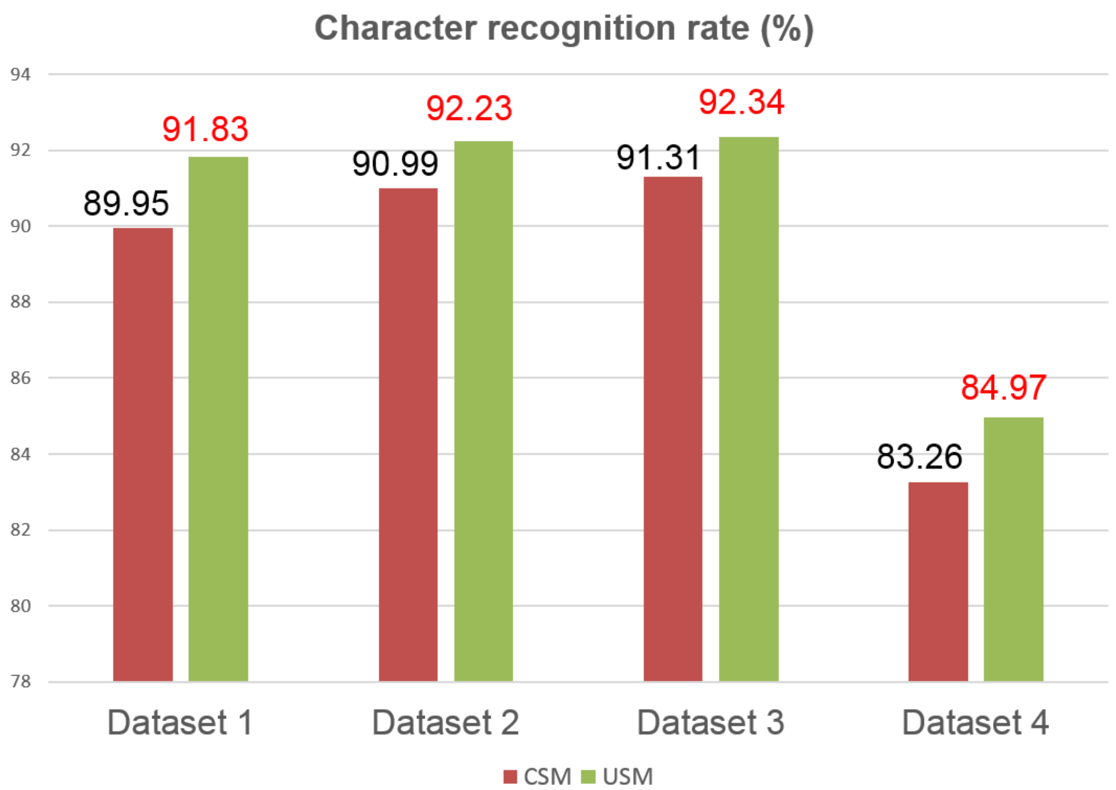
0.4 to 1.5 spacing

Overlaid

-1 to 1 in horizontal & vertical

Character-position-free Handwriting

- Recognition Result: character recognition rate
  - Candidate segmentation method (CSM)
  - Undecided segmentation method (USM)
  - Using 4-fold cross-validation method

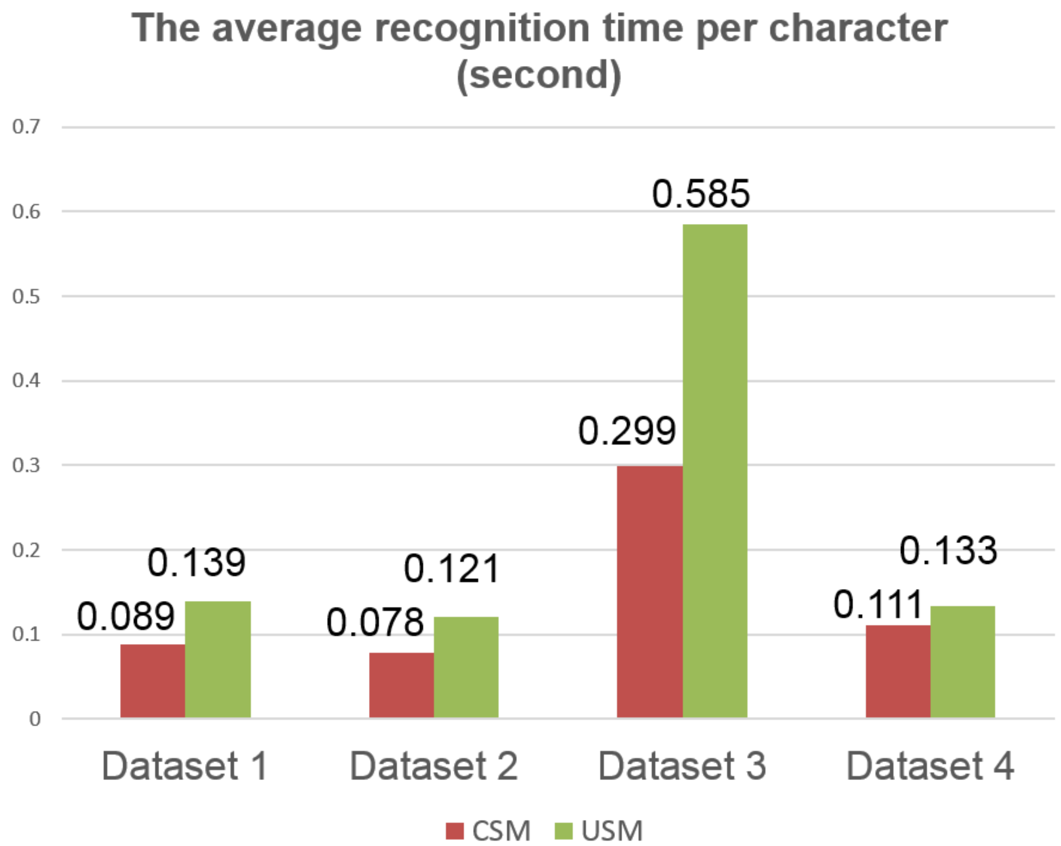


Analysis:

- USM is better than CSM.
- For Dataset 4, the evaluation of geometric context is not effective.

Character-position-free Handwriting

- Recognition Result: speed
  - Candidate segmentation method (CSM)
  - Undecided segmentation method (USM)
  - Using 4-fold cross-validation method



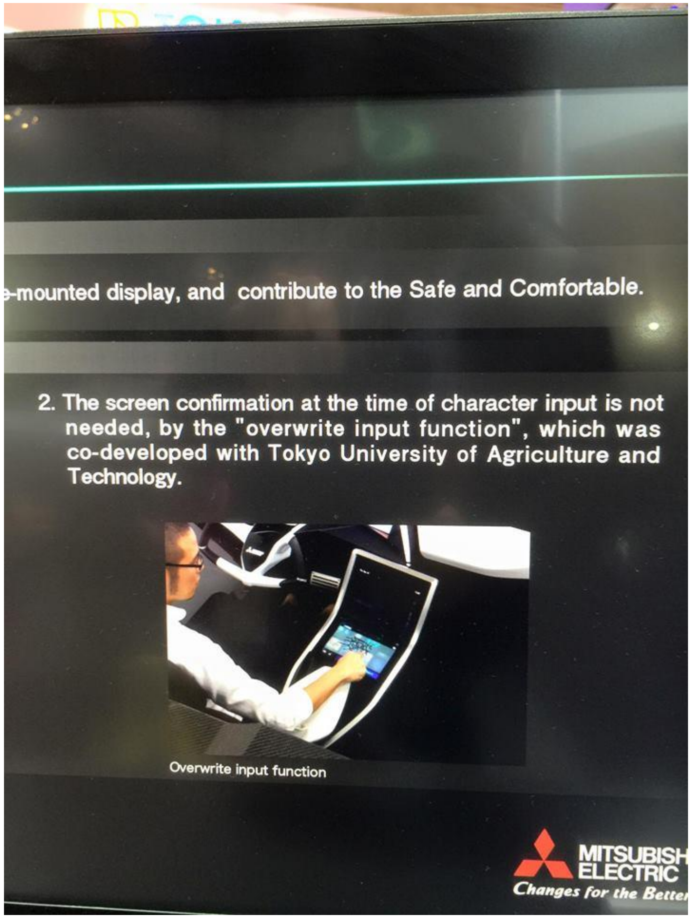
Analysis:

- USM is slower than CSM.
- For Dataset 3 (Overlaid), due to the loss of the spacing, the constructed lattice is bigger than others.

Application to "true" Mobile Device

- Candidate Segmentation Method (CSM) and Undecided Segmentation Method (USM)
- USM is better than CSM
- Proposed USM has been employed for MIRAI 3 xDAS assisted-driving concept car

Demo



Jianjuan Liang, Bilan Zhu, Taro Kumagai and Masaki Nakagawa, "Character-Position-Free On-line Handwritten Japanese Text Recognition", Proc. 3rd ACPR, Kuala Lumpur, Malaysia, Nov. 2015.

Future research topics

computer-assisted and automated marking of handwritten answers

Current State of Examinations

- Examination is essential to evaluate student's learning and ability.
- Marking exams takes large time and effort.
- If it takes time to return the marking result, the effect of review by student decreases.
- Introducing IT: mark sheet, Computer/Web based testing, ··· **Select rather than solve**.

  Are they **compromises?**
- Marking errors are sometimes exposed after the alternative course of an applicant is determined.

  Is it **true** information **disclosure?**

Functional Model of Exam and Marking

problem solving ability and deep understanding
x

Exam g →

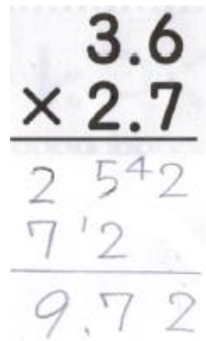
Answer
$g(x)$

Marking f →

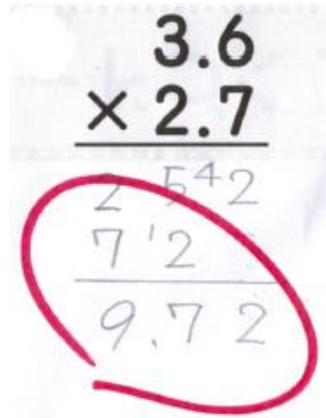
Score
$f(g(x))$

able to multiply decimal numbers

Descriptive
evaluate P.S. ability and deep understanding



Human Marking
Time & Effort, Errors



Selective
inadequate measuring

Q1:3.6 x 2.7 =
① 7.3
② 9.72
③ 97.2

Machine Check
High speed
Few errors

	Answer	Score
Q1	③	0

Mark sheet			
Q1	①	②	●

Correct	
Q1	②

Side effect:
Select rather than solve

Unified Entrance Exam in Japan

- Exam to enter almost all universities
- **Descriptive questions will be included from 2020**
  - To evaluate students' problem solving ability and deep understanding.
  - To foster the ability and attitude to think rather than select.
- **500,000 applicants sit for the exam all over Japan.**
  - It is difficult to employ any electronic devices without troubles at the moment.
  - Only a couple of weeks are allowed for marking.
- **It will make a large impact to education up to universities.**

Handwriting Recognition for Self-learning

- **Tablets are so common in learning**
  - Children are now using tablets to read textbooks.
  - They can write answers to questions in the exercise.

Learner | **System**

writes answers → **recognizes the answers**

verify and correct errors → **marks them immediately**

Immediate Feedback

Comp. assisted/automated marking

- **Computer assists human markers to increase marking reliability for official and large scale exams on paper.**
  - Apply automatic marking before or in parallel with human marking.
  - Human markers mark the portion, train a classifier using the portion and then apply the trained classifier to mark the rest of answers.
  - …
- **Automatically mark with marking confirmation system**
  - Tablet → Electronic paper → Paper for a large number of applicants.
  - A computer marks answers if reliable otherwise rejects ⇒ human markers mark rejected answers and correspond to wrong (false) marking.
  - A examinee can confirm his/her answer and marking after entering ID & password and he/she can claim false marking..
  - By employing multiple recognizers of different characteristics, we can decrease **false negative marking** (which will be claimed by examinees) and especially **false positive marking** (will not be claimed).
  - **If false marking is less than 3 %, the marking time (labor) is shortened (decreased) to 1/3 assuming 20 % rejection and 3 times of effort required for corresponding to false marking claimed.**
  - **It can be shortened more if false marking is processed efficiently.**

Acknowledgements

All the members of Nakagawa Lab.

Especially Dr. Jianjuan Liang who worked on character-position-free recognition.

Fish of the year

Publications

http://www.tuat.ac.jp/~nakagawa/