Man-Machine Communication: of technological dreams and human realism

Lambert Schomaker

presentation held at IBM Watson Center

May 26th 2004

Overview

- Background
- Pattern recognition x User Interfacing
- Characteristics of human cognition
- Lines of research
- Conclusions

Copyright © 2004 L. Schomaker

Speaker background

- physiological psychology, cognitive science, pattern recognition, member IEEE, IAPR
- handwriting process modeling (simulation)
- on-line handwriting recognition
- European projects (IMU,MIAMI,Papyrus)
- Company projects: Olivetti, Tulip computers, Hewlett Packard, Philips, Document Access
- on: handwriting recognition, multimodality and multimediality, information retrieval
- since 2001 full prof & director of AI institute Groningen

Copyright © 2004 L. Schomaker

AI institute Groningen University

- 20 staff, 250 students
- BSc/AI, MSc/AI, MSc/human-machine communication
- part of: Behavioral & Cognitive Neuroscience institute

- Rated 2nd in The Netherlands

Copyright © 2004 L. Schomaker

Research Programmes

- Cognitive Modeling (Taatgen, van Rijn)
- Multi-agent systems (Verbrugge)
- Auditory cognition (Andringa)
- Autonomous Perceptive Systems (Schomaker, de Boer)
- Language & Speech (Hendriks, Wiersinga)

Copyright © 2004 L. Schomaker

Example research line: Social Cognitive Robotics

- Aim: investigate social interactions with robots
  - Focus on the basics
- Themes:
  - Distinction conspecific/other
  - Distinction self/other (mirror experiment)
  - Individual recognition
  - Basic communication/cooperation
- Ideas:
  - Use visual features and motion classification
  - Inspiration from a.o. primate research
- Platform:
  - Sony Aibo

(dr. Bart de Boer)

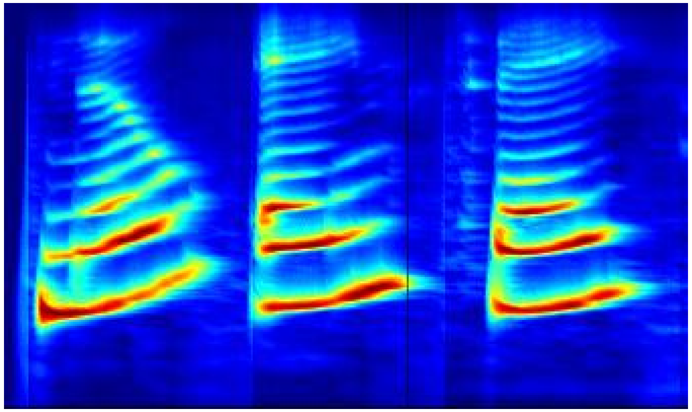
AI RuG

Copyright © 2004 L. Schomaker

bcn

Example research line: Auditory Cognition

- Audio classification on the basis of a model of the cochlea



- Applications: security (shouting detection), traffic classification, environmental measurements

(dr. Tjeerd Andringa)

AI RuG

Copyright © 2004 L. Schomaker

bcn

The tale of the stubborn pattern recognizers

Copyright © 2004 L. Schomaker

The tale of the stubborn pattern recognizers

- *Wouldn't it be great if we could talk to computers?*
- *Wouldn't it be great if computers could read our writings?*
- *Wouldn't it be great if computers would see our world?*
- *(etc.)*

AI RuG

Copyright © 2004 L. Schomaker

bcn

Observations

- From 1984 to 2004, CPU power increased from 6 MHz to 3 GHz, an improvement factor of 500
- Spoken-word recognition error rate went from 25% to 1% (10^4-word lexica), an improvement of factor 25
- Handwritten-word recognition error rate went from 35% to 5% (10^4-word lexica), an improvement of factor 7

Hirsch, Hellwig, Dobler (2001), Eurospeech'01 (Ericsson R&D)

AI RuG

Copyright © 2004 L. Schomaker

bcn

Observations

- From 1984 to 2004, CPU power increased from 6 MHz to 3 GHz, an improvement factor of 500
- Spoken-word recognition error rate went from 25% to 1% (10^4-word lexica), an improvement of factor 25
- Handwritten-word recognition error rate went from 35% to 5% (10^4-word lexica), an improvement of factor 7

➔ ***We must be doing something wrong … (?)***

Copyright © 2004 L. Schomaker

Observations

- From 1984 to 2004, CPU power increased from 6 MHz to 3 GHz, an improvement factor of 500
- Spoken-word recognition error rate went from 25% to 1% (10^4-word lexica), an improvement of factor 25
- Handwritten-word recognition error rate went from 35% to 5% (10^4-word lexica), an improvement of factor 7

➔ ***We have ill-posed goals …(?)***

AI RuG

Copyright © 2004 L. Schomaker

bcn

Four eye openers

- 1. Rudnicky & Hauptman (1989)
- 2. Frankish, Hull & Morgan (1995)
- 3. Goldberg & Richardson (1993)
- 4. Lopresti (1994) & others

AI RuG

Copyright © 2004 L. Schomaker

bcn

Eye opener 1:
User‘s Cost Evaluation & pattern recognition

- Rudnicky & Hauptman (1989): there is a cost associated with user actions, speech, pen or typing
- Cost: #actions to reach goal (example: entering numbers)
- Use explicit Markov modeling to analyze and predict user behavior
- Time (#actions) relates to inverse of transition probability matrix

Copyright © 2004 L. Schomaker

Eye opener 1:
User's Cost Evaluation & pattern recognition

- Rudnicky & Hauptman:
  …There is a keyboard/microphone break-even point which is a function of **recognizer performance**
  and **string length**
  …

- ➔ **with current HWR / ASR word-recognition performances, the keyboard can be expected to win often**

Copyright © 2004 L. Schomaker

Eye opener 2: User perception of classifier performance

- Frankish, Hull & Morgan (1995):
  user acceptance of a pen-based interface is strongly influenced by Pen-UI design.

- Users worked on a dbase application and were asked about the quality of the handwriting recognizer

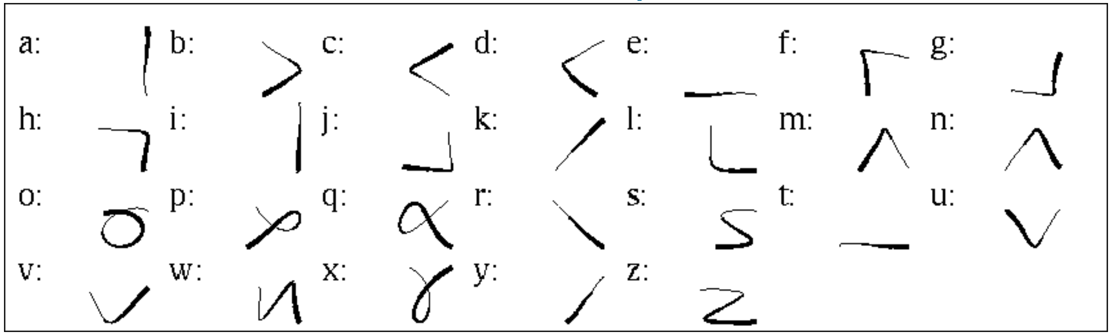
- ➔ A good PUI can make mediocre recognizers look useful

- ➔ A bad PUI can make a reasonably good recognizers look “stupid”

Copyright © 2004 L. Schomaker

Eye opener 3:
If the cpu can't do it, use the brain

- Goldberg & Richardson (1993):
  if pattern recognition does not work,
  let the users adapt their writing style
- Stylized "unistroke", easy on the classifier
- Users like the predictability
- (speed is slower than cursive script)



Copyright © 2004 L. Schomaker

Eye opener 4: Who needs pattern recognition?

- Scribble communication (human-to-human) and personal note taking do not require full-blown pattern recognition
- HP research and Lopresti (1994): leave ink “as is”
- Use an Information-Retrieval paradigm for note search: 100% recognition is not needed at all
- Also ➔ IBM Crosspad concepts, such as *lifelong note taking with a single device*

Copyright © 2004 L. Schomaker

Insights at NICI’s on-line HWR group (1)

- Nobody wants to write an isolated word and wait more than a second for a possibly misrecognized machine-font version of it
- ➔ In writing, users want to fluently produce larger, meaningful text chunks or
- ➔ store exact crisp facts:
  phone numbers, email addresses, URLs
- the user motivation for **post-hoc** annotation is likely to be limited

Copyright © 2004 L. Schomaker

Insights at NICI's on-line HWR group (1)

- Nobody wants to write an isolated word and wait more than a second for a possibly misrecognized machine-font version of it
- ➔ In writing, users want to fluently produce larger, meaningful text chunks or
- ➔ store exact crisp facts:
  phone numbers, email addresses, URLs
- the user motivation for post-hoc annotation is likely to be limited ➔ consistency at **time of input** is required

Copyright © 2004 L. Schomaker

Insights at NICI’s on-line HWR group (2)

- **omni-writer** (free-style) handwriting recognition in an open application is a dream
- … while **user-adaptation** approaches for single-writer recognition require user motivation, knowledge and competence
- … **unsupervised adaptation** is yet another dream

Copyright © 2004 L. Schomaker

Insights at NICI’s on-line HWR group (3)

- Language & **dialog modeling** is very helpful
- …but costly in design
- models are never complete
- which will hold a fortiori for **multimodal** dialog models

Copyright © 2004 L. Schomaker

Performance in Reading Systems (% correct) (4)

Input	Machine	Human
isolated digits	> 99%	> 99% but fatiguable
isolated letters	> 90%	> 95%
cursive **words** in sentence	> 65% (?) (may be higher when using language model)	> 85%
cursive **words, isolated**	> 65% (fixed lexicon, limited size)	~ 77% (huge lexicon)
sloppy cursive **words, unknown context**	<< 20%	~ 54%

Is on-line handwriting recognition a bad idea?

- (maybe)
- Pen-based note taking may be useful
- Pen-based tablet PCs may be very useful

- Counter measures:
  - better classifiers
  - better hardware (tablet resolution, noise)
  - better dialog modeling
  - better feedback
  - better error-correction scenarios

- Multimodality in the UI?

Copyright © 2004 L. Schomaker

Multimodality hypothesis

- the simultaneous or alternating use of different input devices will increase the user-to-system bandwidth

Copyright © 2004 L. Schomaker

Multimodality

- pen and keyboard
- pen and speech
- hand gesture and speech
- visual and audio speech
- mouse and keyboard
- pen, speech and keyboard
- joystick +force feedback, tactile feedback
- etc.

AI RuG

Copyright © 2004 L. Schomaker

bcn

Multimodality

…but: multimodality must be learned, just like driving a car

➔multimodality works well in analog control (F16 fighter planes),

➔multimodality in symbolic communication is difficult!

Copyright © 2004 L. Schomaker

Multi**med**iality hypothesis

- the simultaneous or alternating use of different human sensory systems will increase system-to-user bandwidth

Copyright © 2004 L. Schomaker

Multimediality

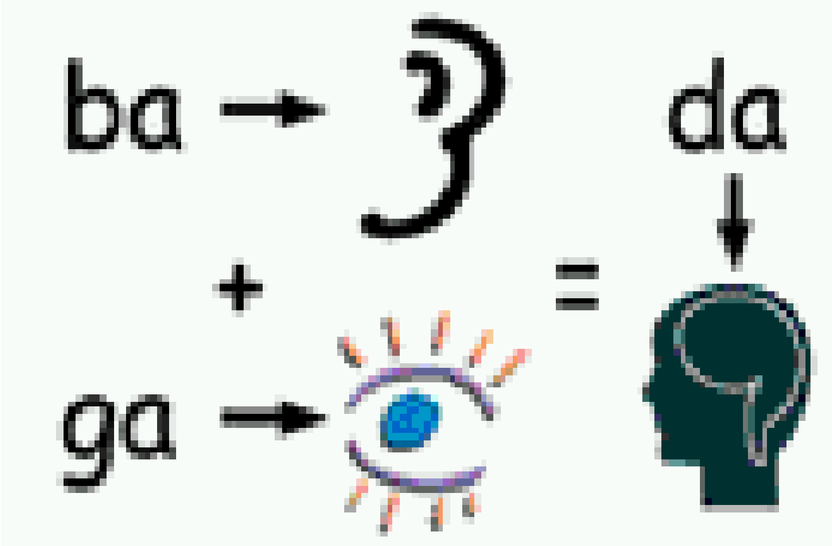
but: badly organized multimediality reduces system-to-user bandwidth.

Examples:

- Text+speech explanations in museal kiosks
- “Flashy” web sites
- Mc Gurk effect

Copyright © 2004 L. Schomaker

McGurk effect



- an example of multimodal fusion
- an example of unpredictable effects of
- inappropriate multimedia combination

Copyright © 2004 L. Schomaker

Back to the drawing board

- We need to reconsider goals

➔ towards a Moore's Law of
User-System Bandwidth !!

Copyright © 2004 L. Schomaker

towards a Moore's Law of User-System Bandwidth

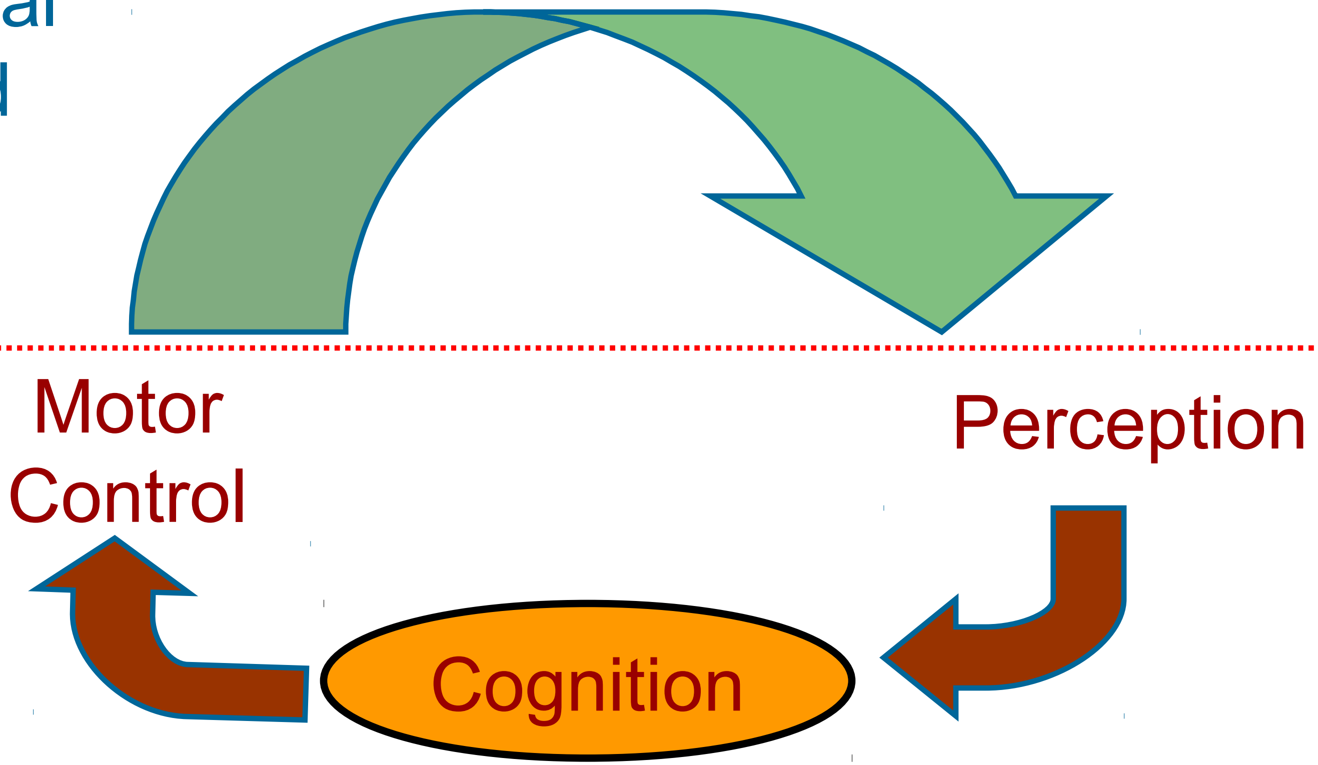
- We need more knowledge on fundamental characteristics of human perception, cognition (language) and motor control
- Sources: neuroscience, cognitive science

Copyright © 2004 L. Schomaker

Human-world interaction

external physical world

perception-action loop

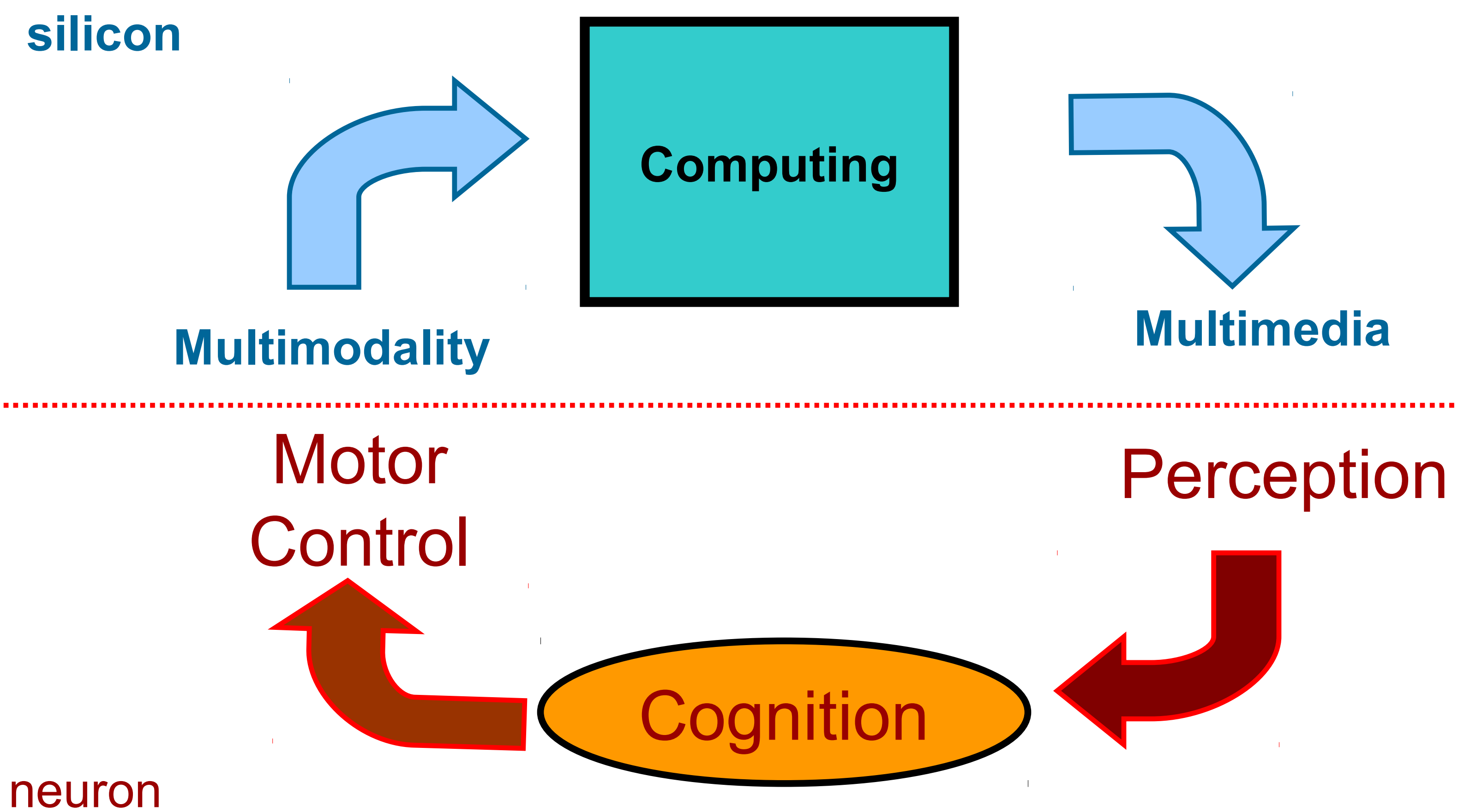


Copyright © 2004 L. Schomaker

Integrated perception-action

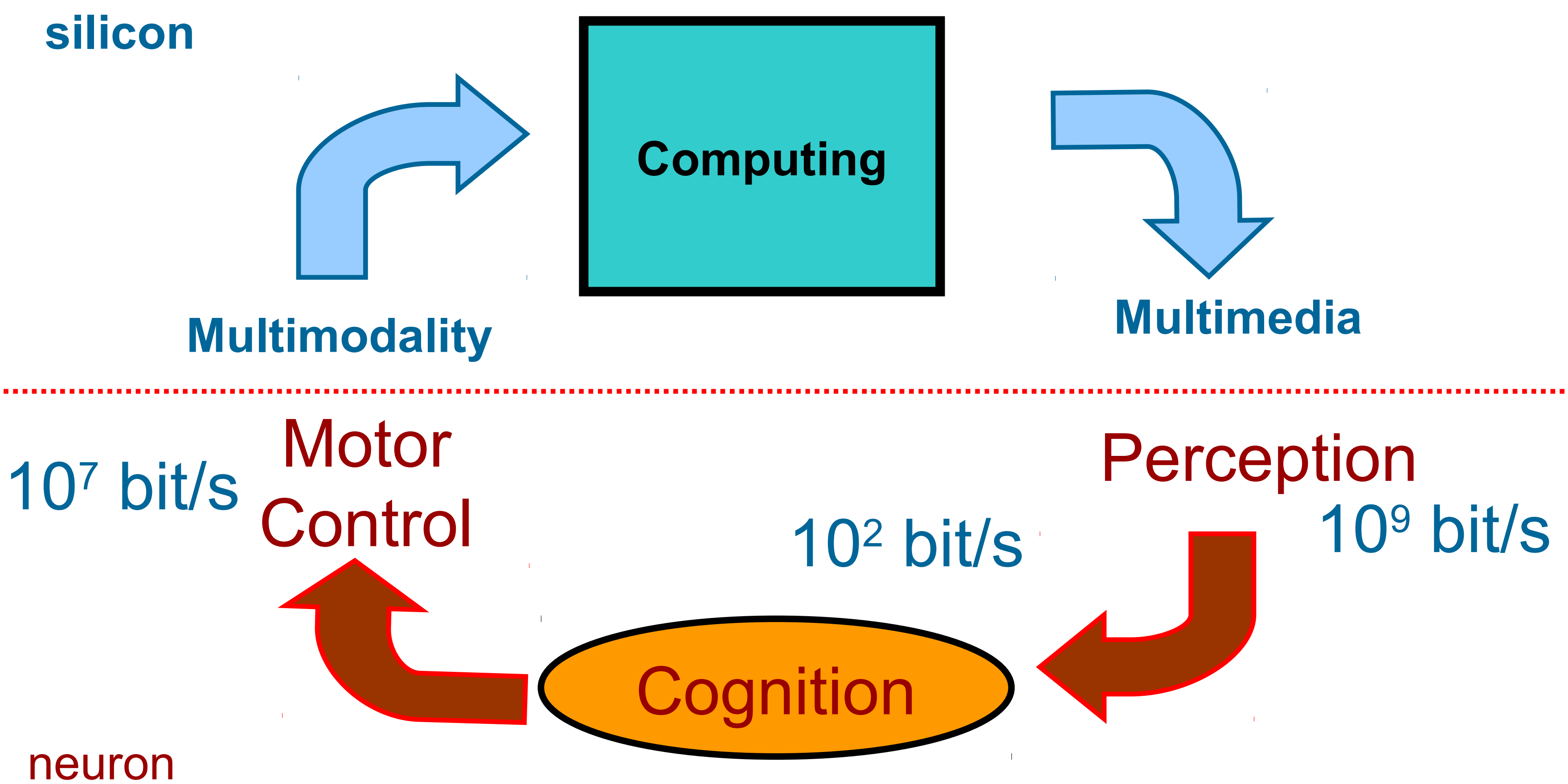
Copyright © 2004 L. Schomaker

Human-computer communication



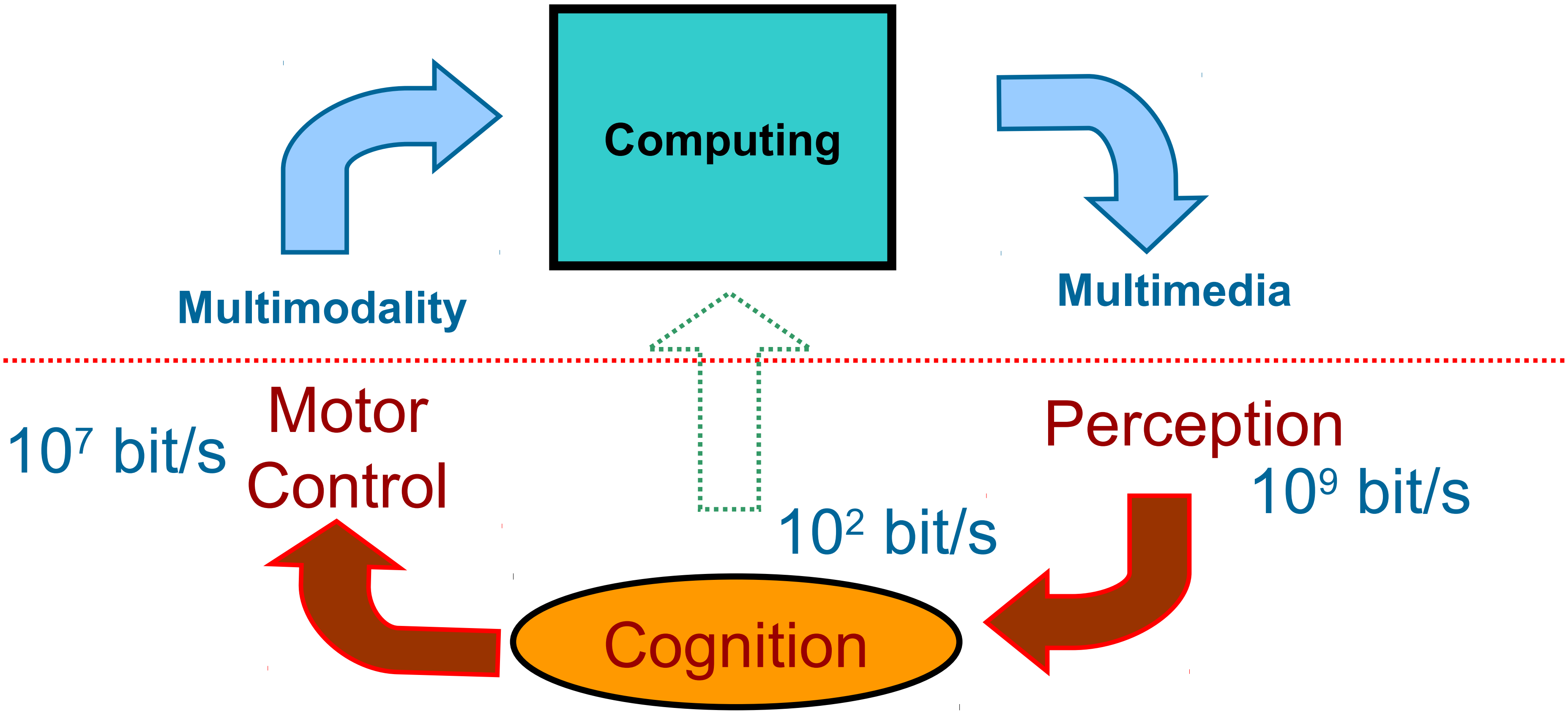
Copyright © 2004 L. Schomaker

Bandwidth



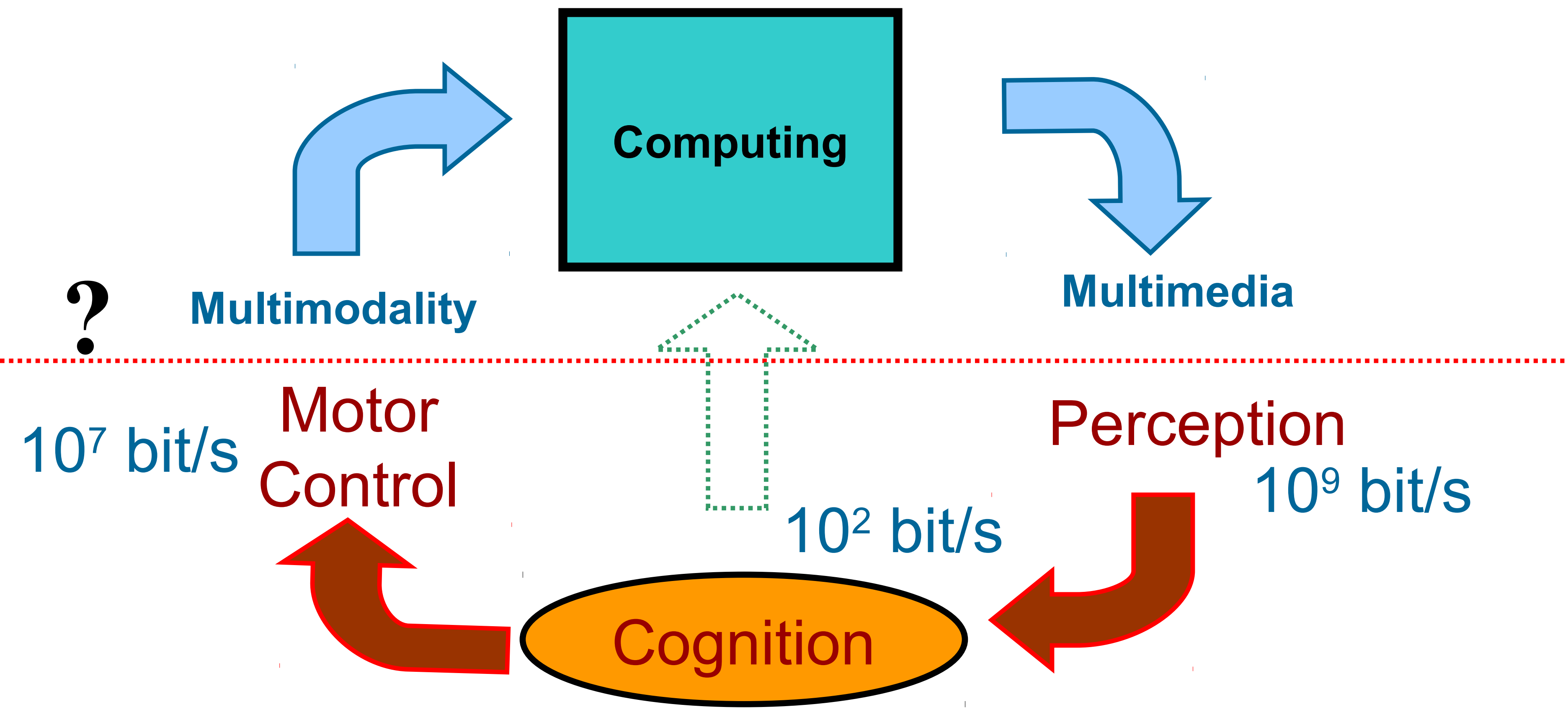
Copyright © 2004 L. Schomaker

Bandwidth



Copyright © 2004 L. Schomaker

Bandwidth



Copyright © 2004 L. Schomaker

apple bits and orange bits

- There is a distinction between the physical signal bits (light intensity, sound-wave pressure, muscle force)
- and intended symbolic bits (language, reasoning)
- between them sits noise and a lot of unknowns

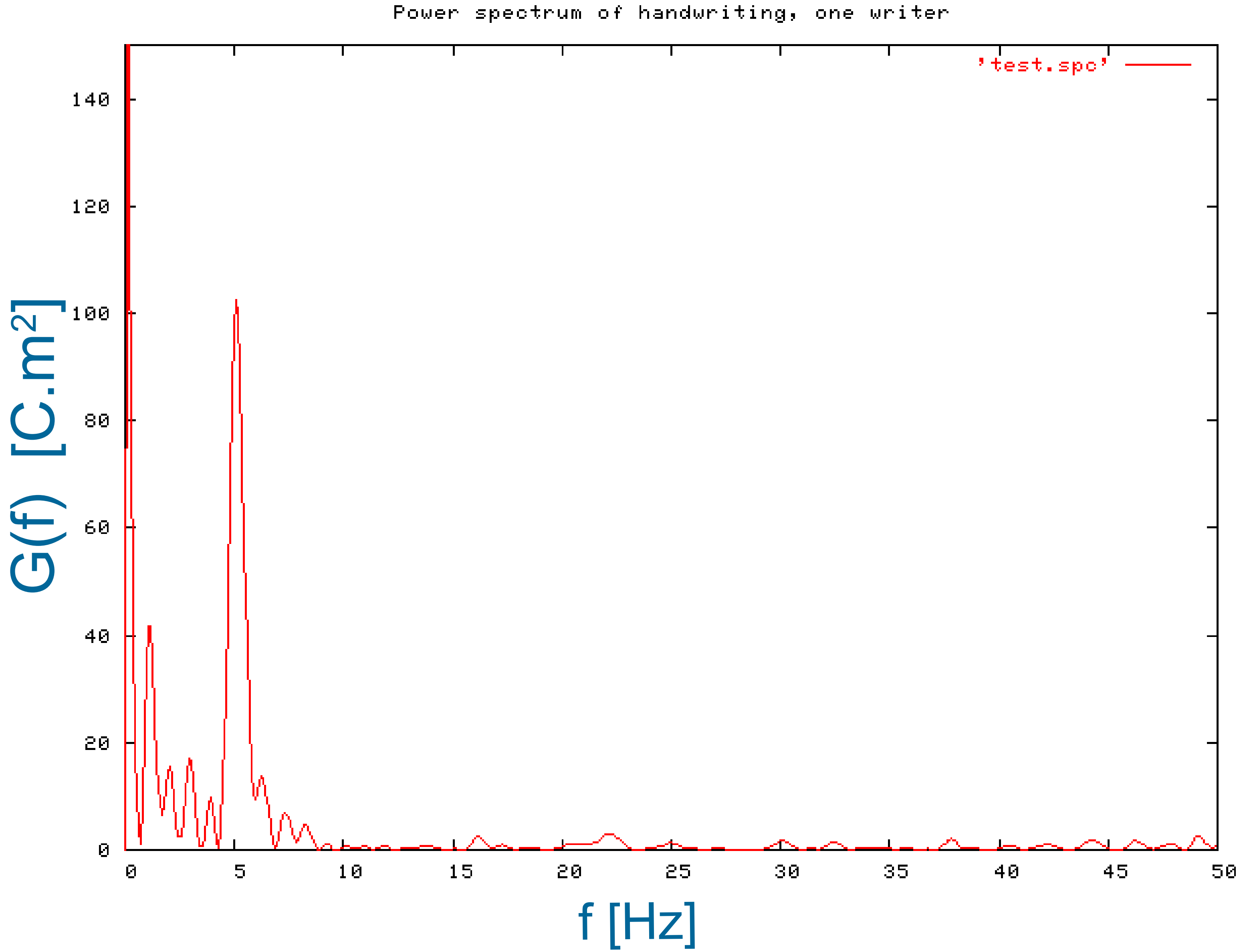
Copyright © 2004 L. Schomaker

Motor-system bandwidth

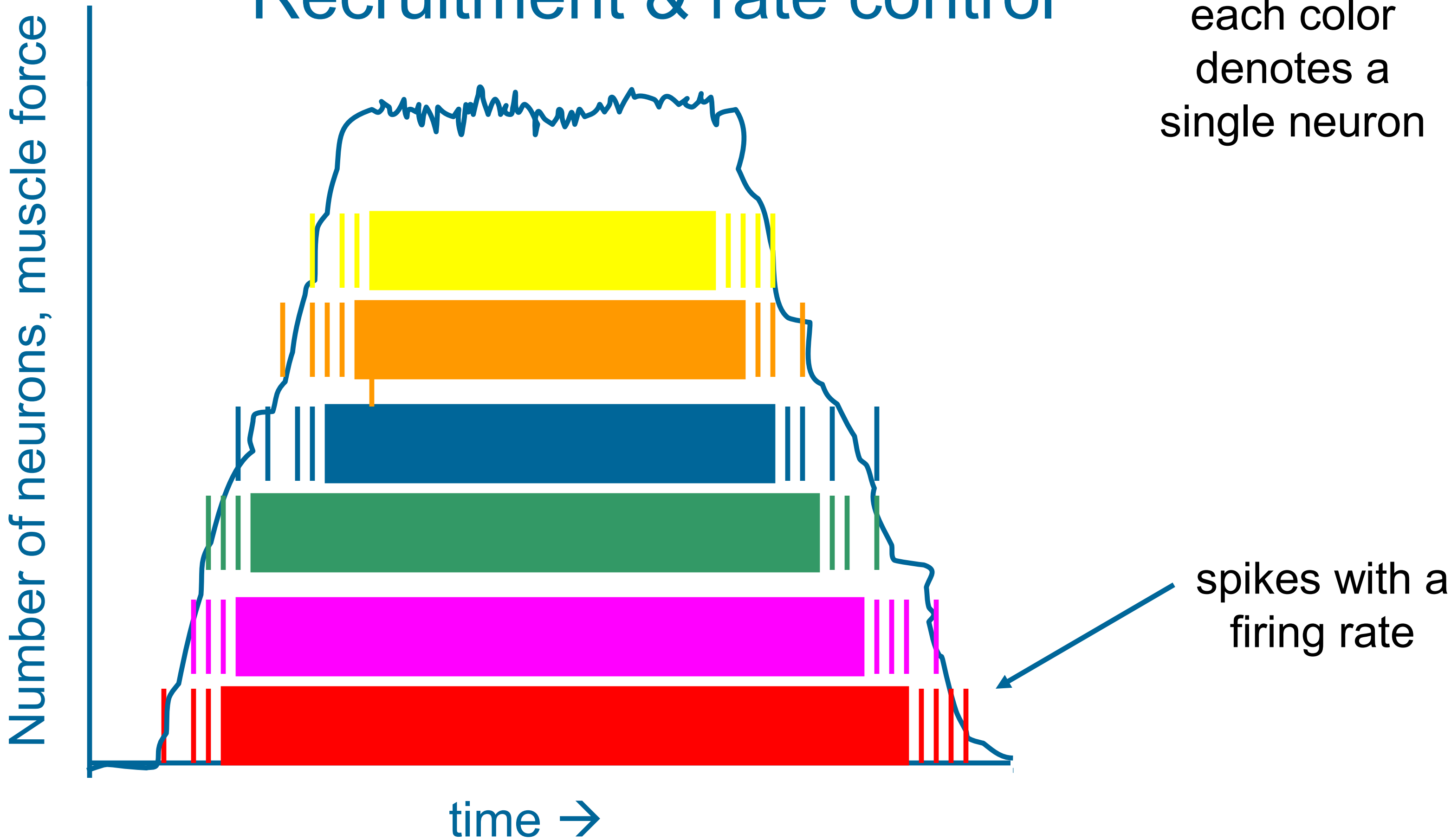
- 10^7 bit/s is peak rate (e.g. modeling tennis player movements)
- it includes all muscles
- of which the *end effector system*
  for a given UI task is just a subset
- per muscle, the bandwidth is limited and there is motor noise (van Galen & Schomaker)

Copyright © 2004 L. Schomaker

Power-spectral density of pen-tip movements



Recruitment & rate control



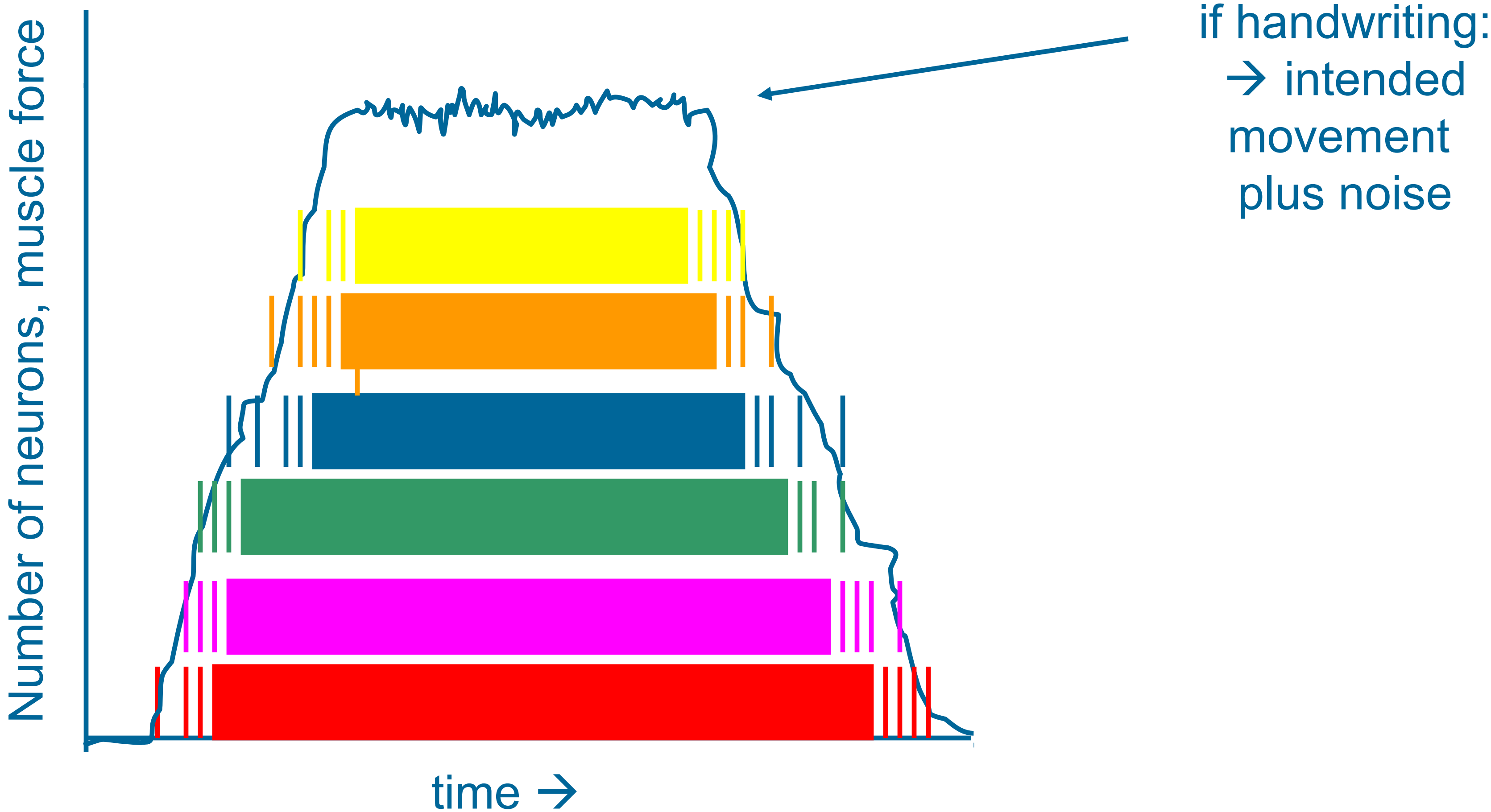
Force = G(#units active

, firing rate)

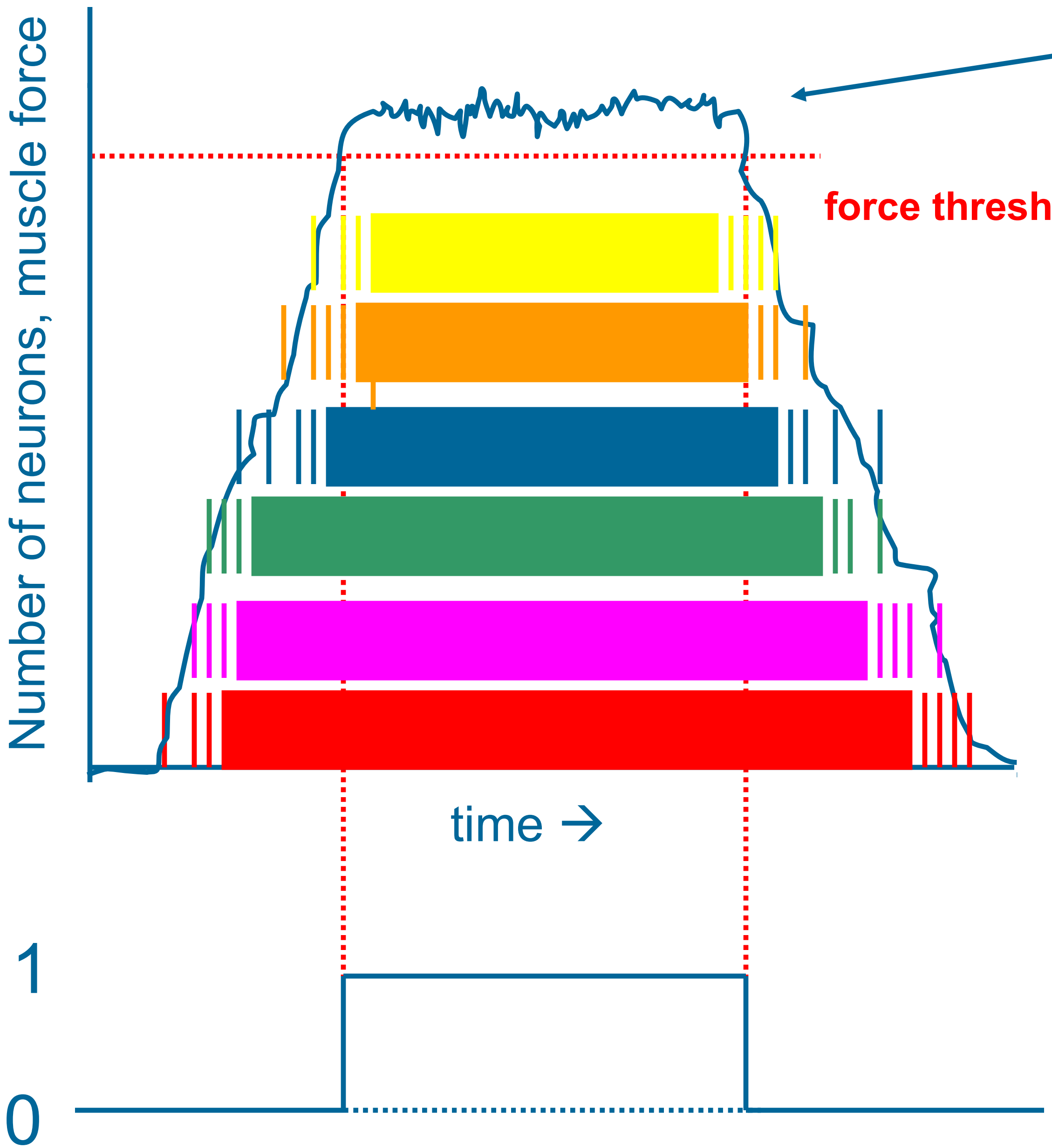
coarse control

fine control

(Fig: after neuromuscular research center)



(Fig: after neuromuscular research center)

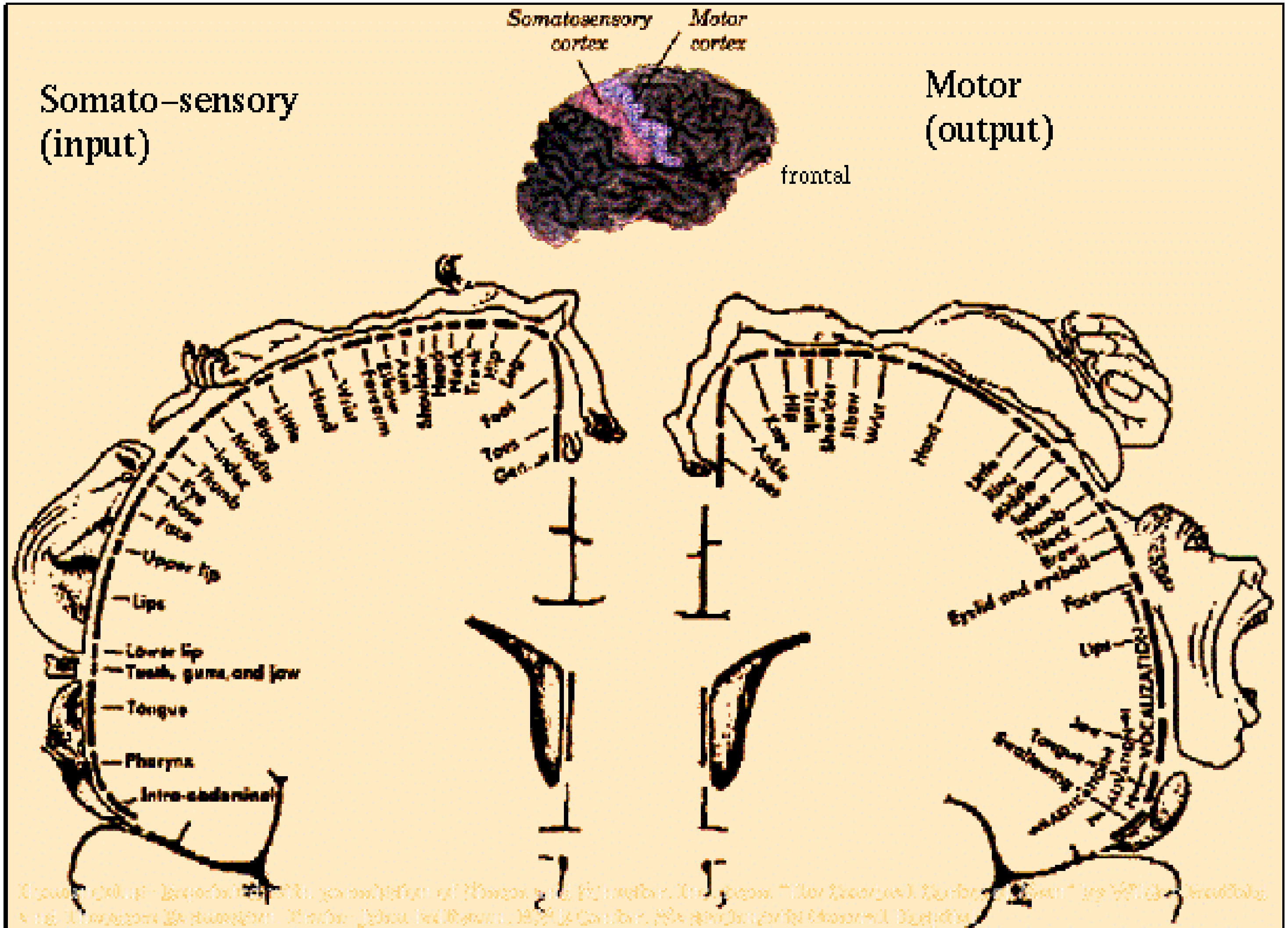


if handwriting:
→ intended
movement
plus noise

if thresholded force, i.e.
keyboard press:
→ closer to the intention

(Fig: after neuromuscular research center)

Where to get your bits?



Adapted from Penfield & Rasmussen

Rule of thumb: 1 neuron, 1 bit/s

Where to get your bits?

- Large motor areas devoted to mouth and tongue
- Large motor areas devoted to fine control of the fingers

- There is intrinsic feedback (propriocepsis) to improve S/N ratio!

Copyright © 2004 L. Schomaker

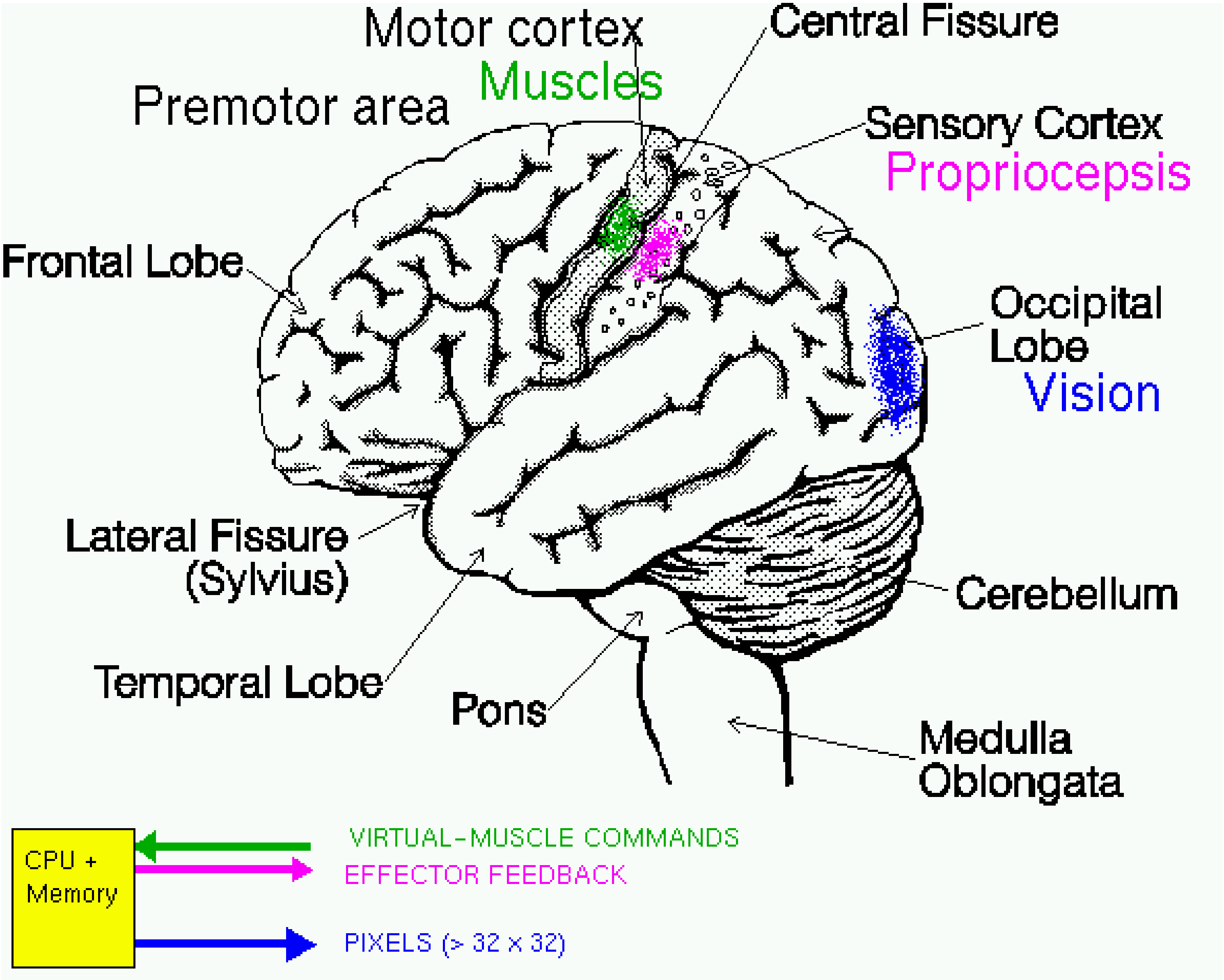
Where to get your bits?

- Tantalizing: keyboard and pen, speech indeed seem to be the way to go:
- these motor-cortex areas are largest
- … neural implants?

Copyright © 2004 L. Schomaker

© 2001 Ted Goff tedgoff@tedgoff.com http://www.tedgoff.com

"It comes with a special tool for pushing buttons on the keyboard."



Motor cortex
Muscles
Premotor area
Central Fissure
Sensory Cortex
Propriocepsis
Frontal Lobe
Occipital
Lobe
Vision
Lateral Fissure
(Sylvius)
Cerebellum
Temporal Lobe
Pons
Medulla
Oblongata
CPU +
Memory
VIRTUAL-MUSCLE COMMANDS
EFFECTOR FEEDBACK
PIXELS (> 32 x 32)

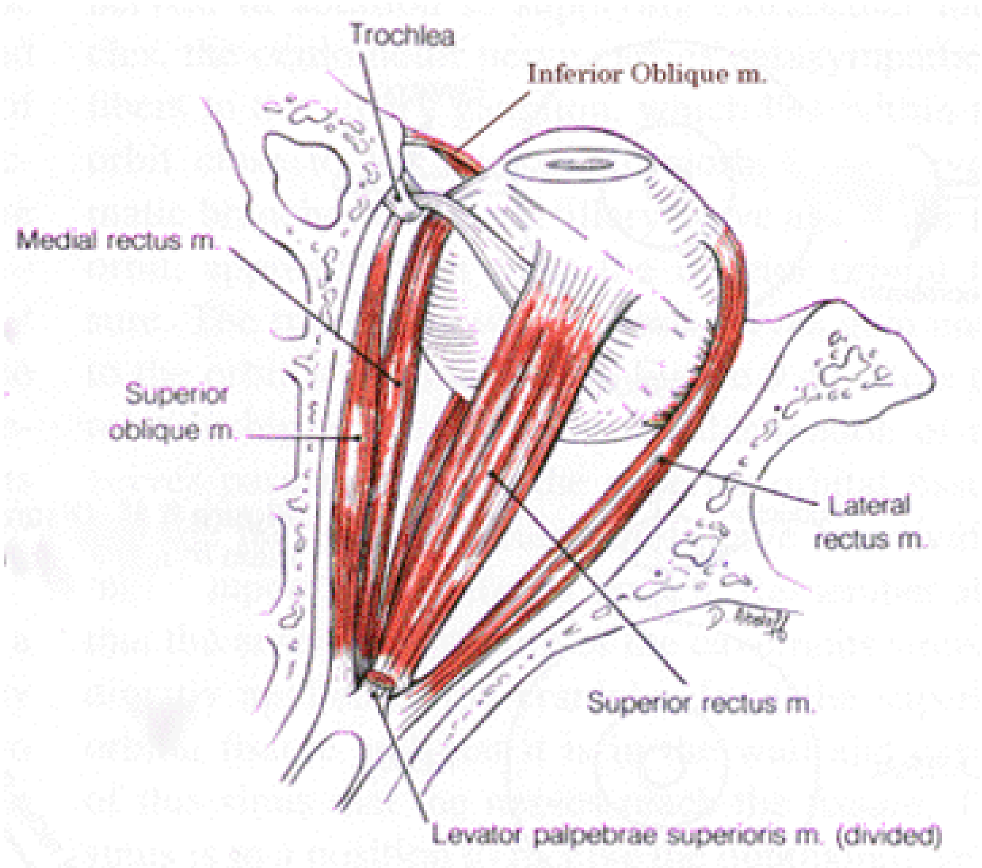
Directions in our research

- Aim at PR applications where the user is motivated and/or expert
- On-line ➔ Off-line (historical documents)
- Writer identification

- Cognitive Modeling (ACT*R) for user agents
- Keyboard-based innovations
- Active vision for camera-based reading

Copyright © 2004 L. Schomaker

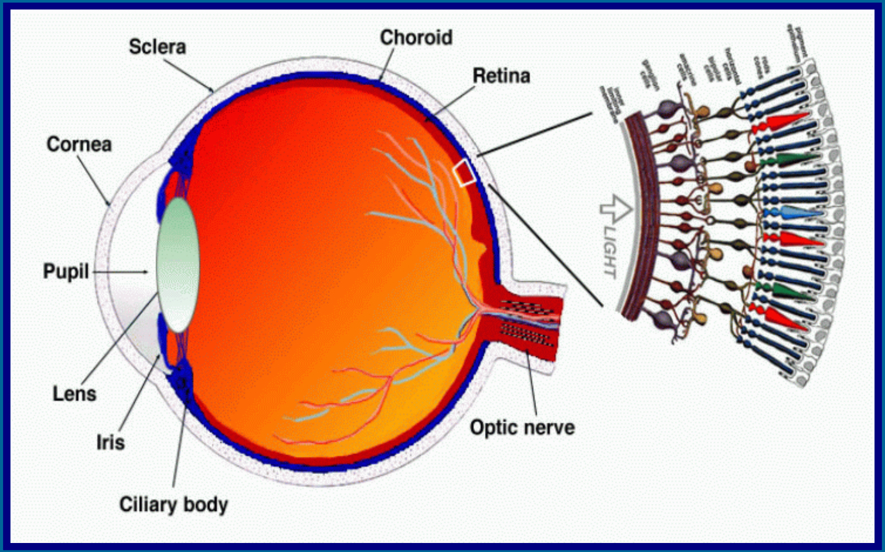
Active Vision

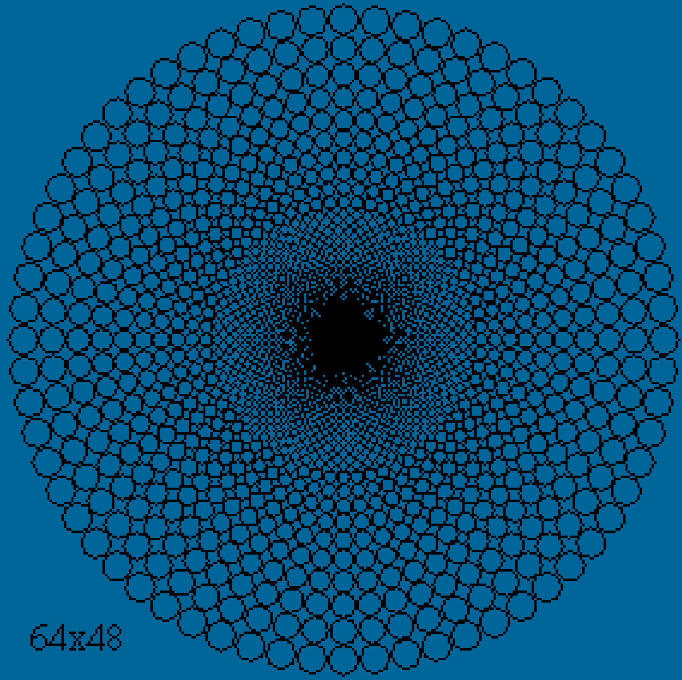


AI RuG

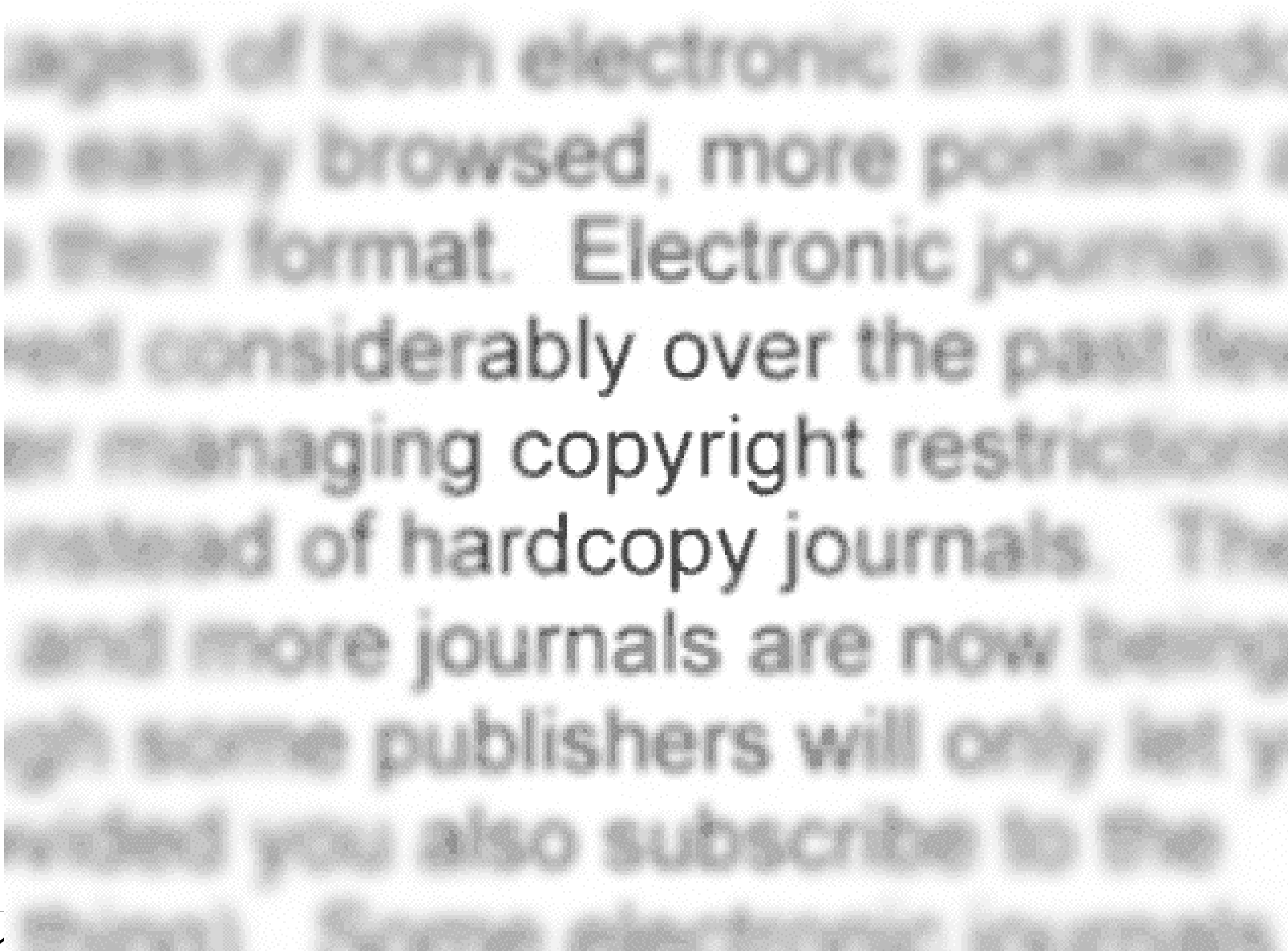
Copyright © 2004 L. Schomaker

bcn



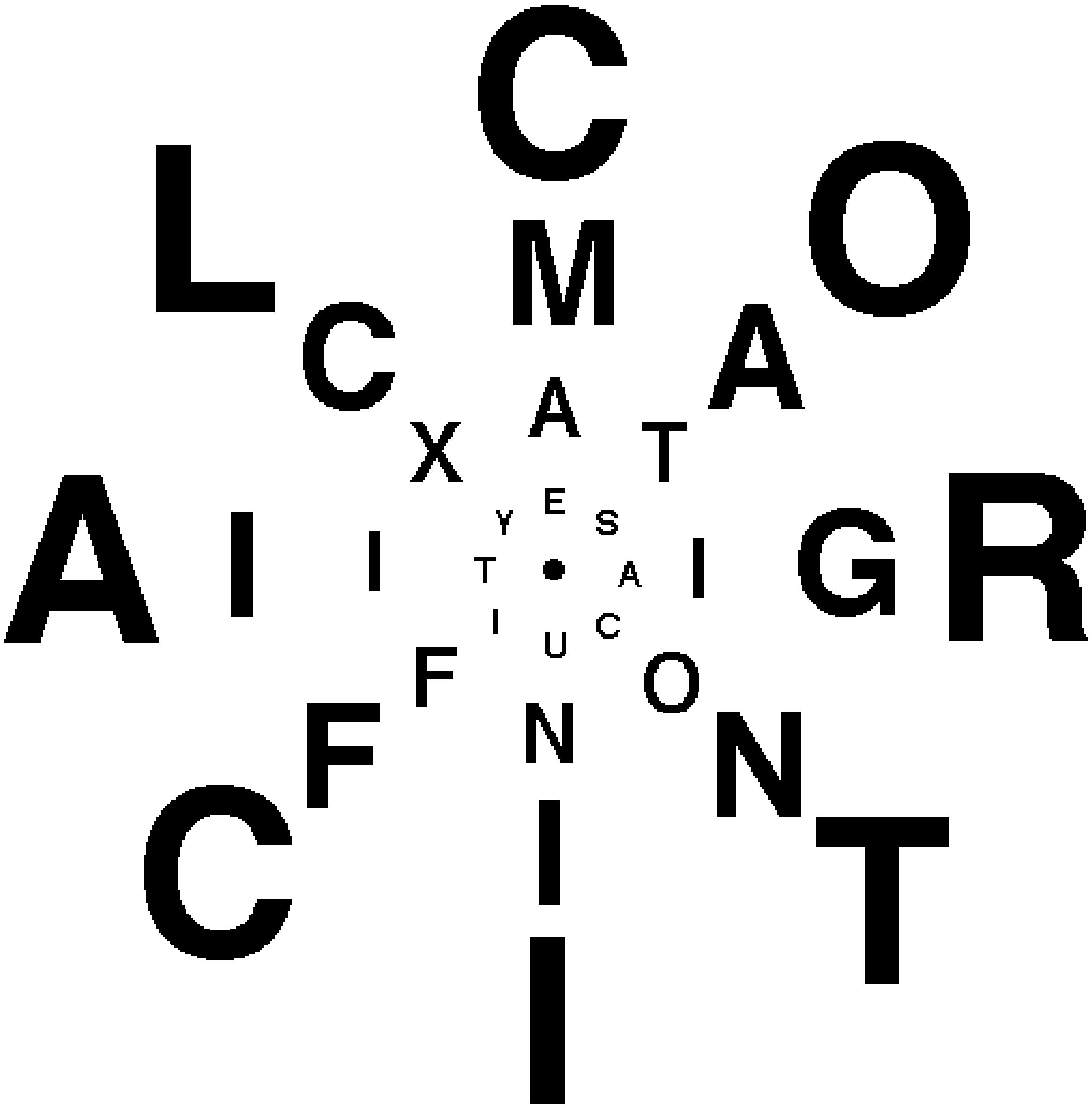


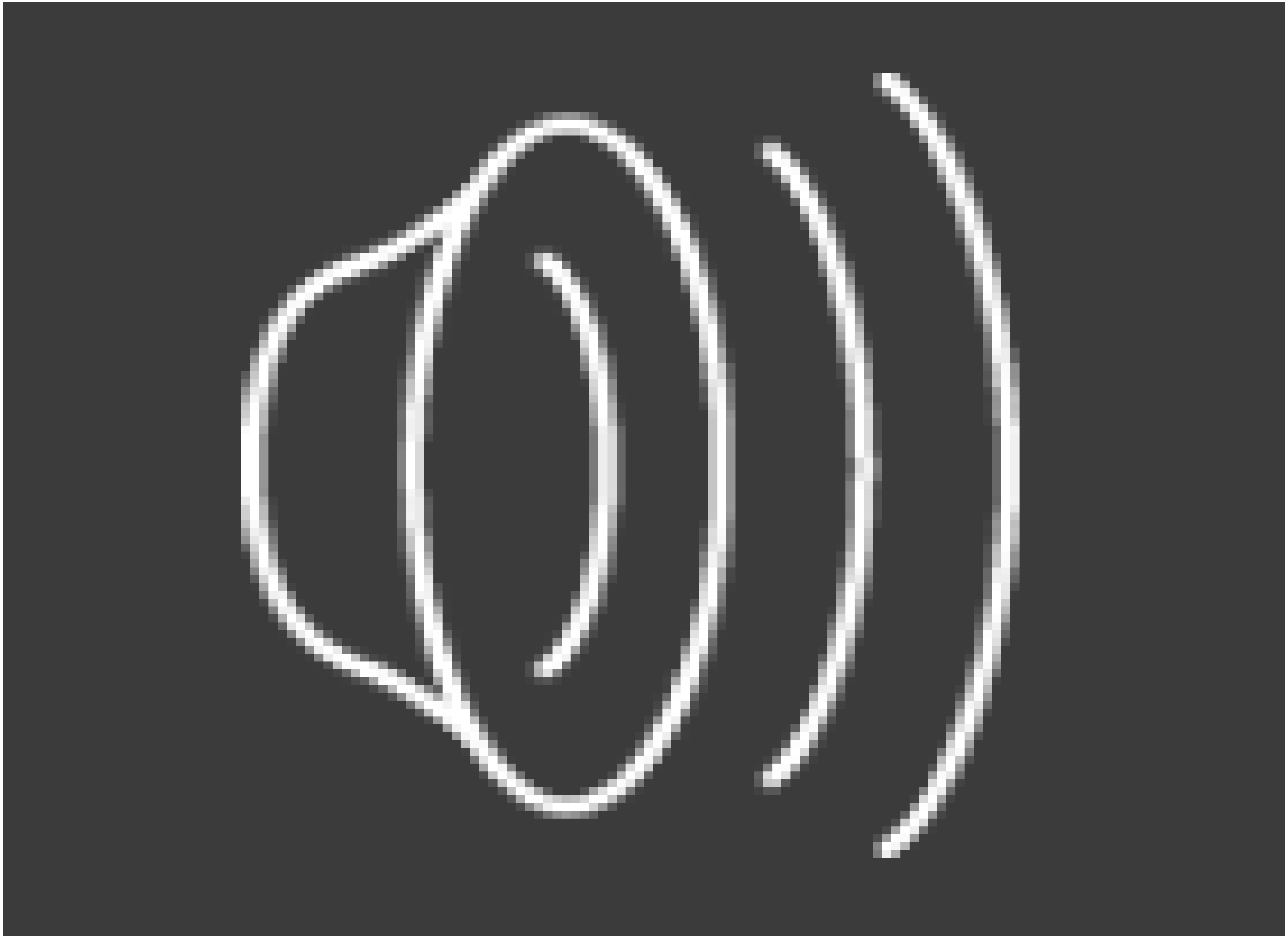
retina: x,y → log(r), phi

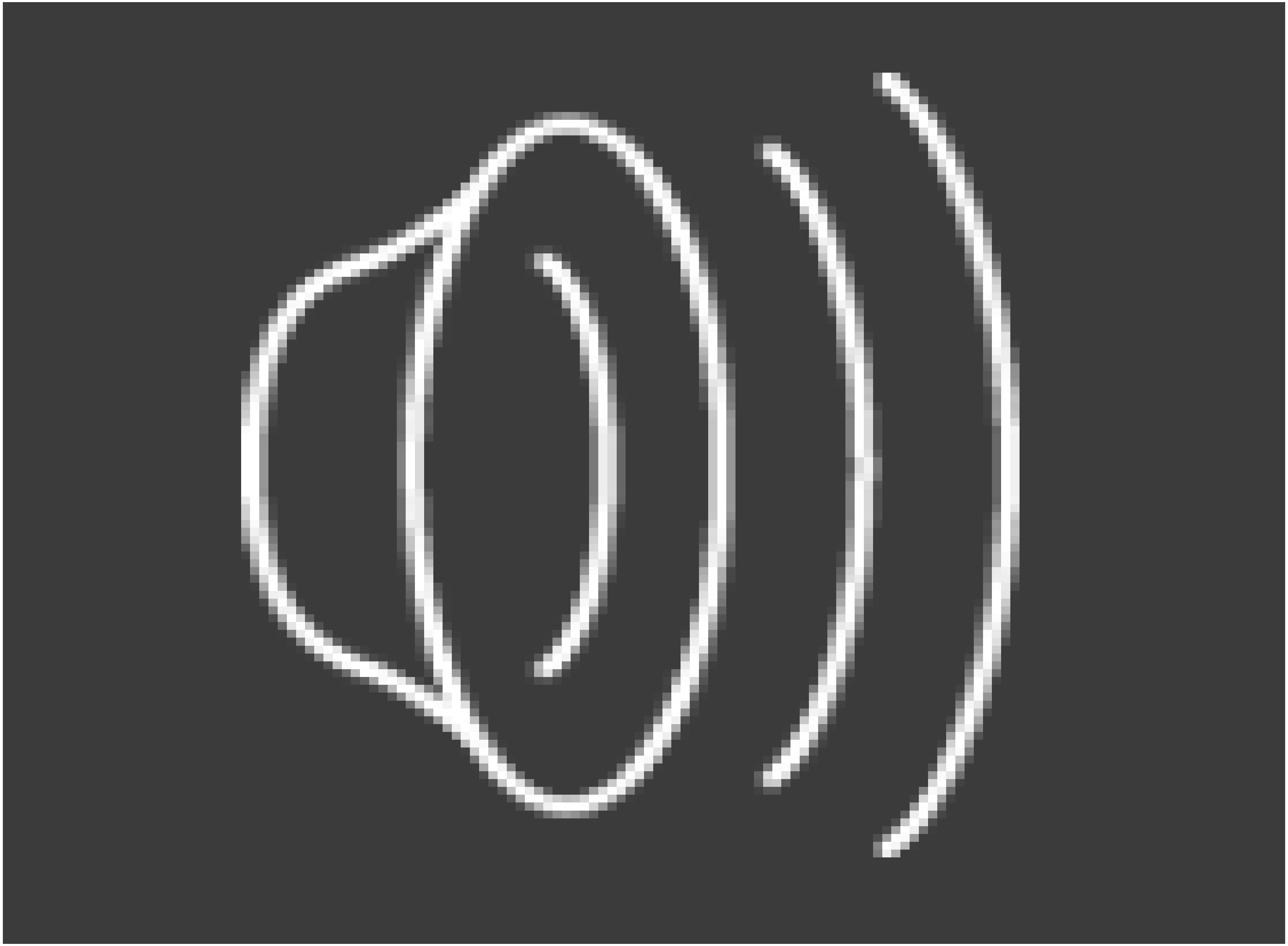


Copyright © 2004 L. Schomaker

equal-density in V1 cortical projection







Lesson from the biological system

- save computation on the image
- …by opportunistic sampling
- selective attention instead of cumbersome full-page segmentation
- dynamic error correction (back tracking)

Copyright © 2004 L. Schomaker

Active Vision

- will be used in a project on robotic reading systems
- robot trying to find its way in man-made environment without electronic navigation (beacons, gps)
- PhD student Marius Bulacu

Copyright © 2004 L. Schomaker

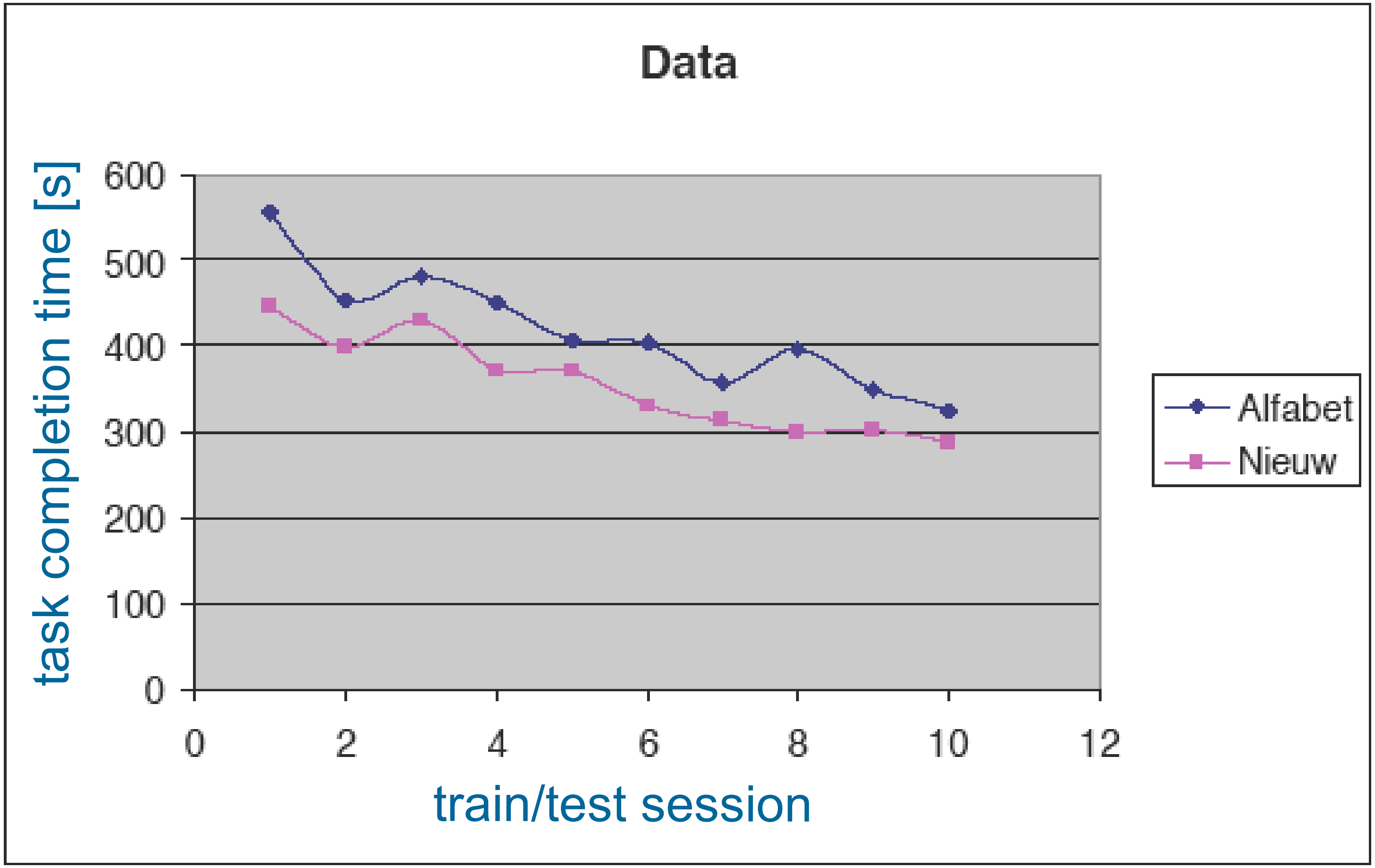
Keyboard methods (student work)

- cell-phone text messaging is popular in Europe
- text entry is expensive:
  - a: press 1x b: 2x   c: 3x
- idea: letter/ngram frequency-dependent ordering
  - cf: Dvorak, Velotype

Copyright © 2004 L. Schomaker

Letter	Frequentie	Percentage
SPACE	2784494	22.48
e	1407151	11.36
n	769818	6.22
a	717151	5.79
t	537183	4.34
i	523371	4.23
o	478815	3.87
d	436487	3.52
r	420218	3.39
l	285381	2.30
g	281465	2.27
s	269930	2.18
h	261449	2.11
k	239107	1.93

SMS text entering (Schaap & Geerdink)



Writer identification

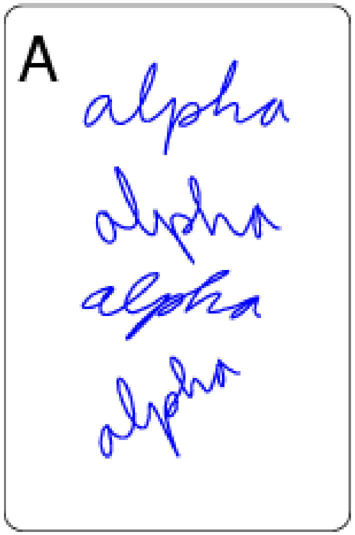
- find the writer identity on the basis of an unknown handwritten sample and a reference set

- applications:
  - automated forensic-writer search
  - mail-address filtering
  - user (style) identification in HWR

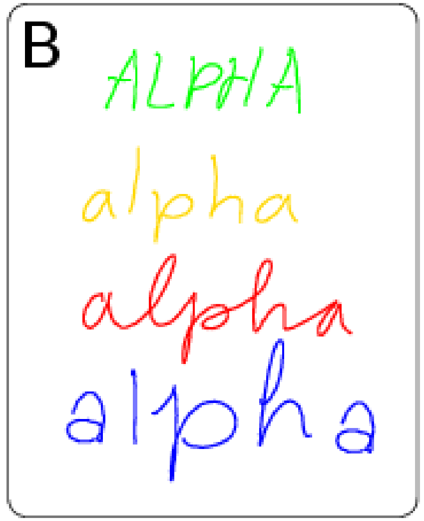
Copyright © 2004 L. Schomaker

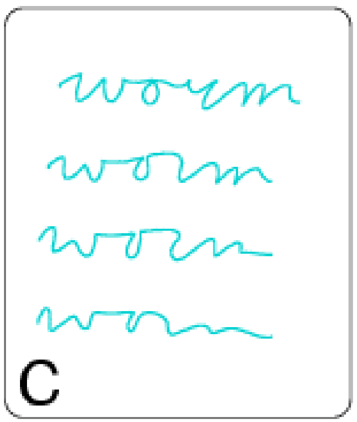
algebra	algebra	algebra	algebra
algebra	Algebra	algebra	Algebra
algebra	algebra	algebra	algebra
algebra	algebra	algebra	algebra
algebra	algebra	algebra	algebra
algebra	algebra	algebra	algebra

affine variation

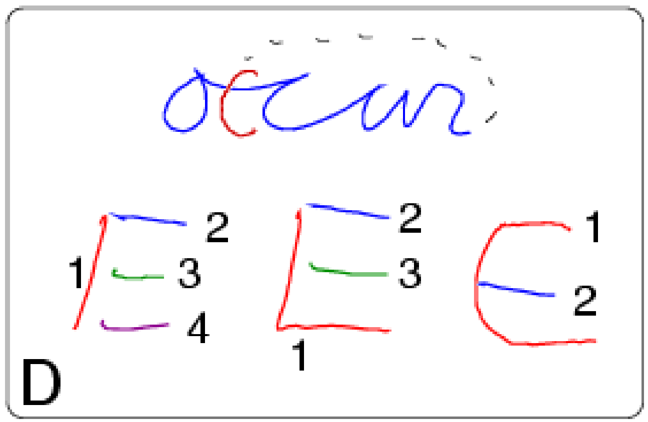


allographic variation



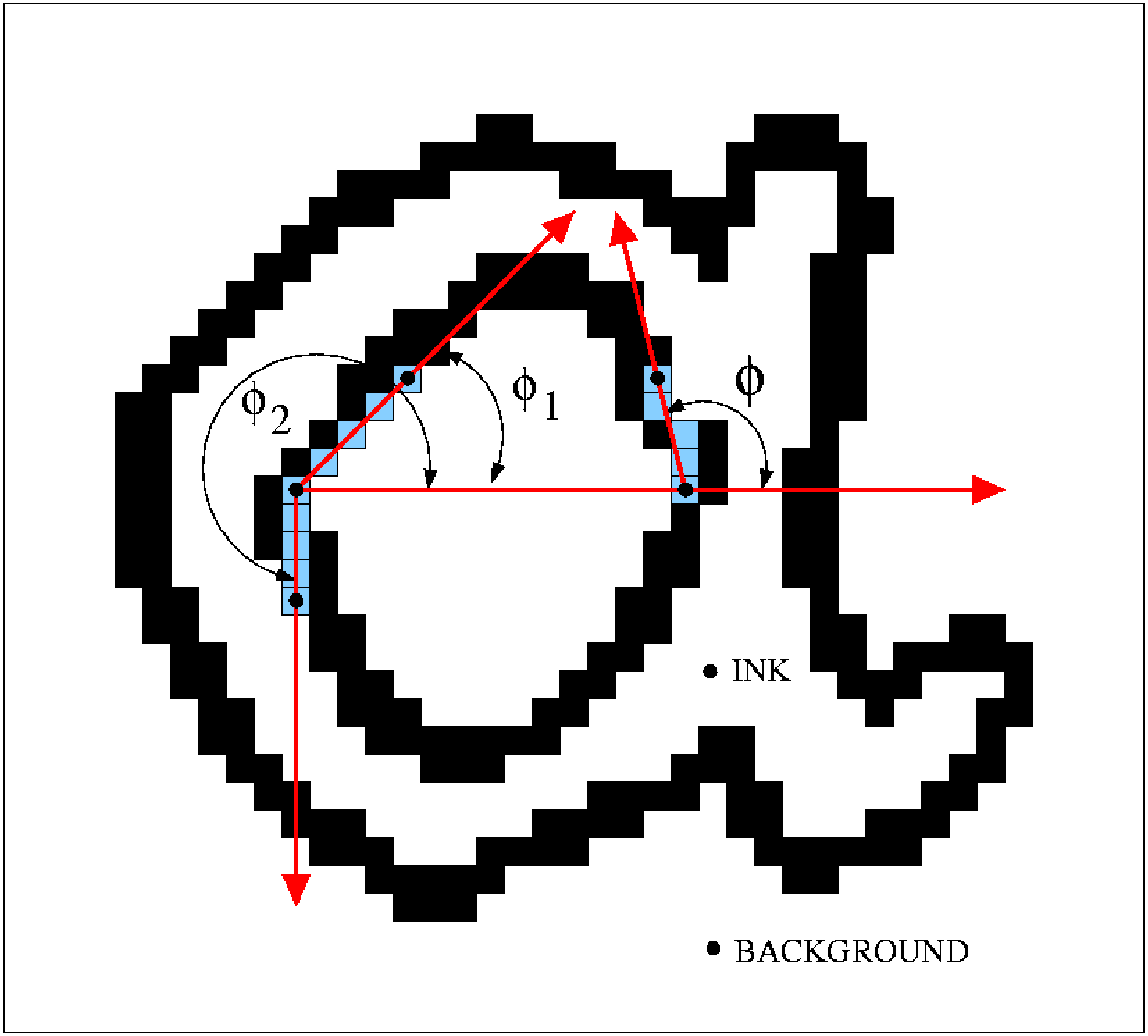


neuro-biomechanical variability

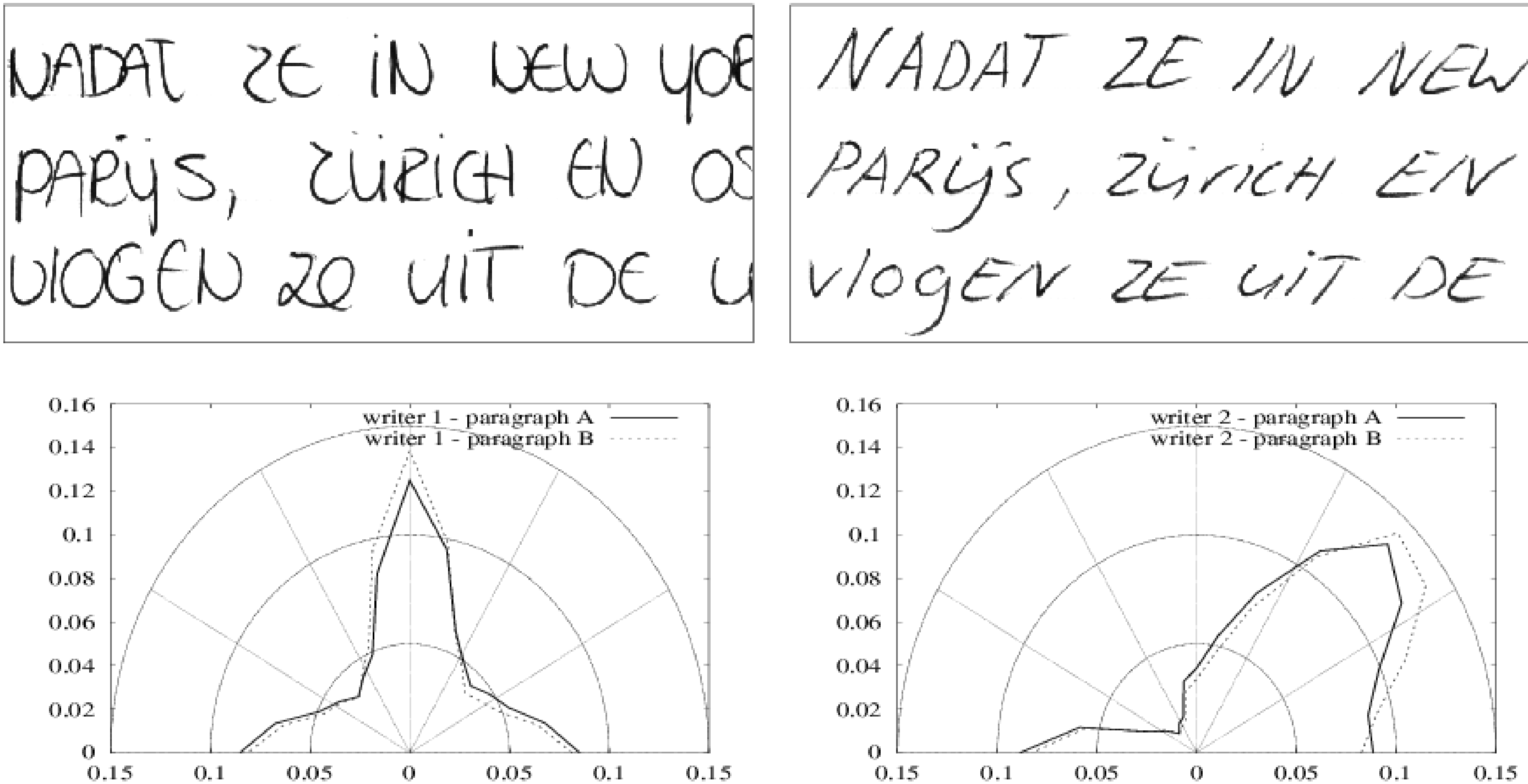


sequence variability

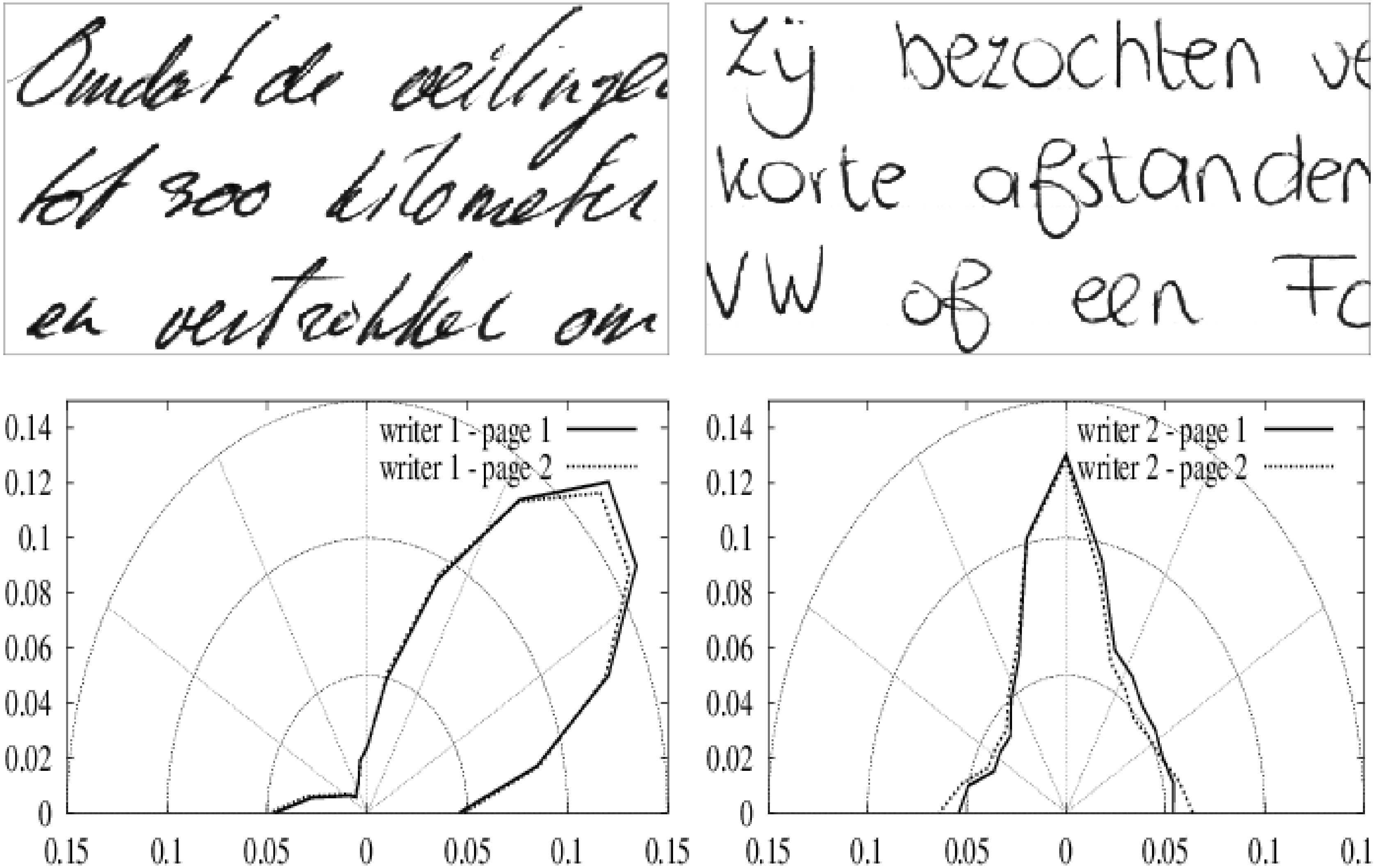
Edge-angular features



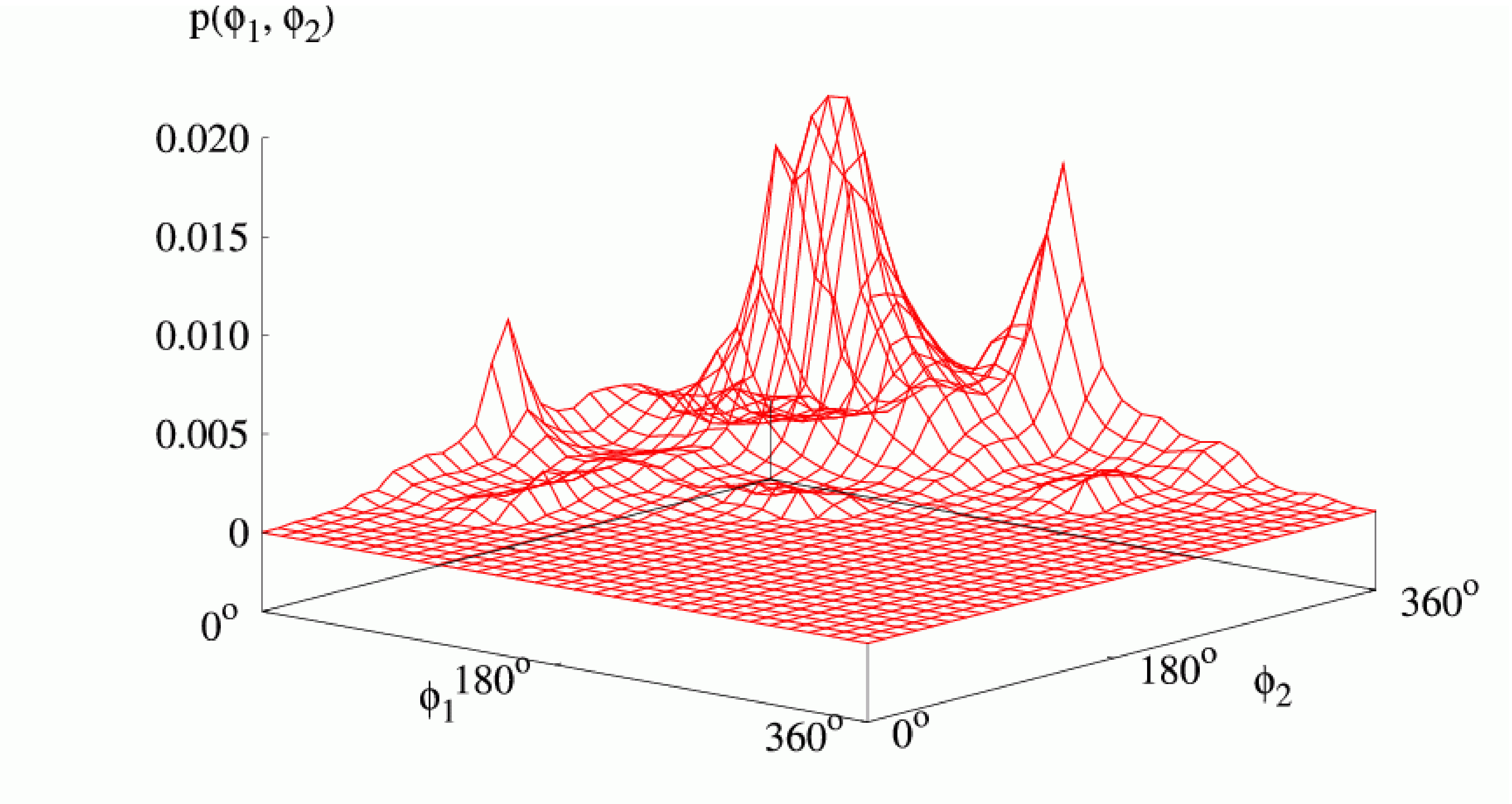
PDF of edge angles (cf Maarse et al., Crettez)



PDF of edge angles (cf Maarse et al., Crettez)

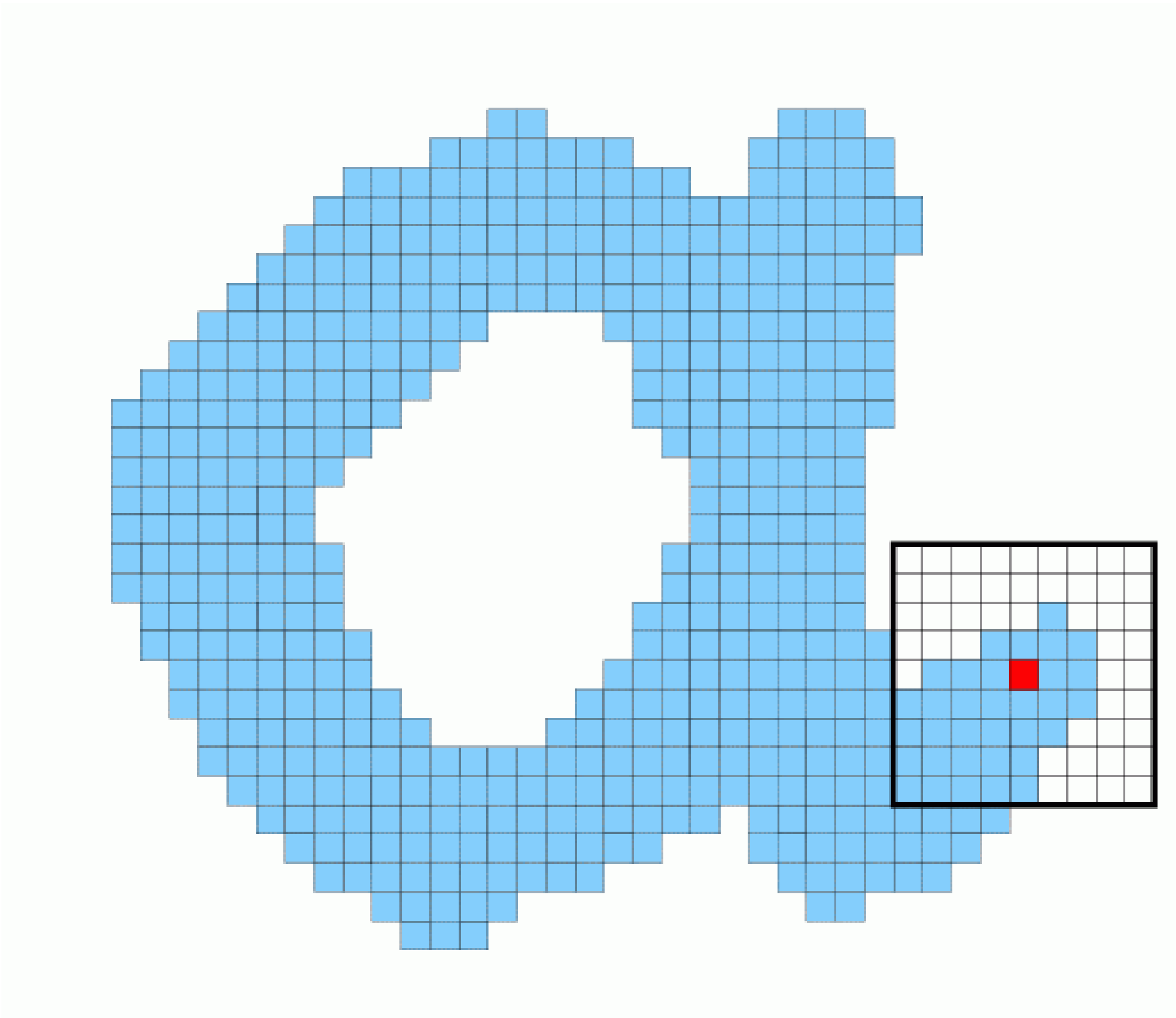


PDF of hinge angles (Bulacu & Schomaker, '03)

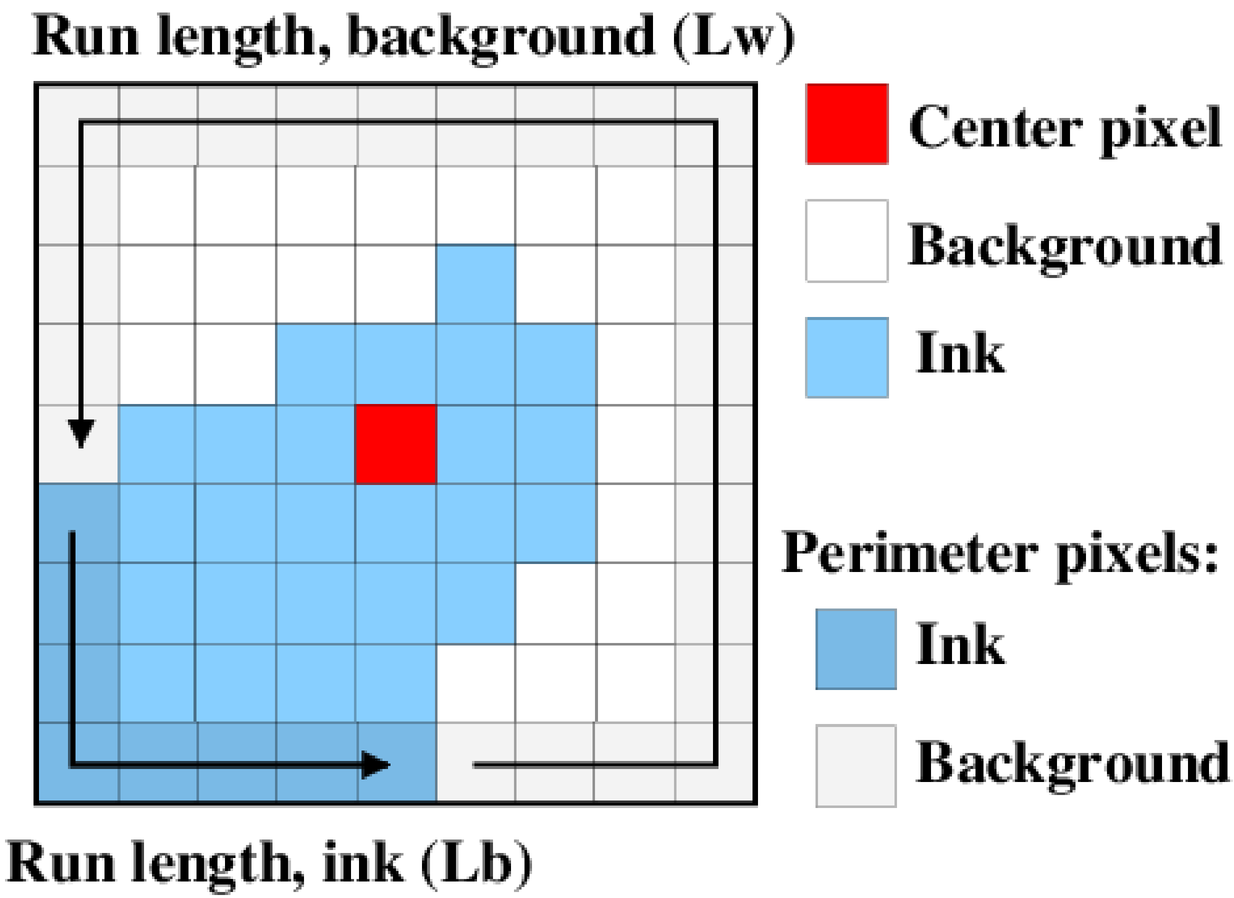


One writer, mixed and lower-case text

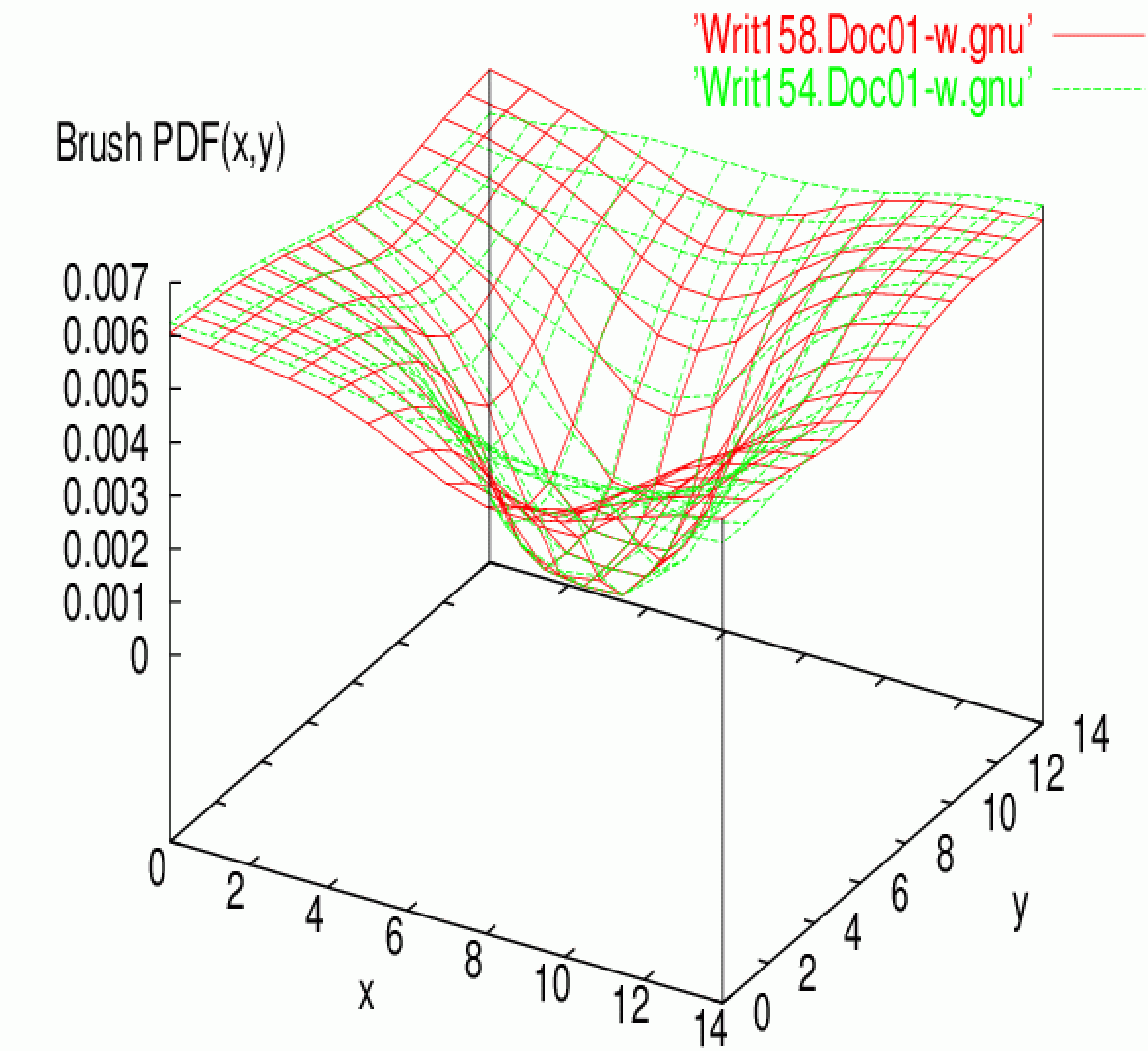
Brush feature (Schomaker & Bulacu, ‘03)

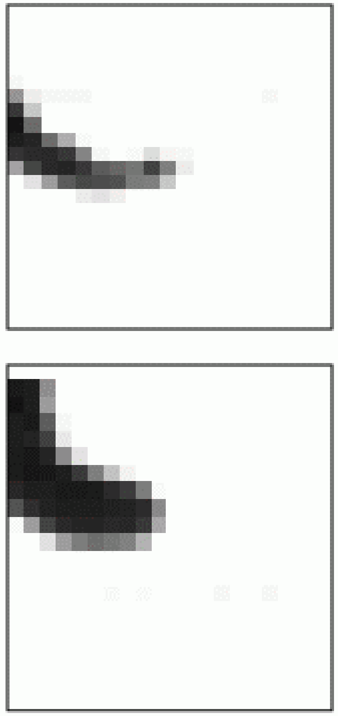


Brush feature (Schomaker & Bulacu, ‘03)



Brush feature (Schomaker & Bulacu, ‘03)





Features in ICIP '03 paper Schomake & Bulacu

f1 autocorrelation of horizontal raster

f2 Vertical run-length PDF of ink

f3 Horizontal run-length PDF of white

f4 Brush (ink density at stroke endings)

f5 Single edge-angle PDF

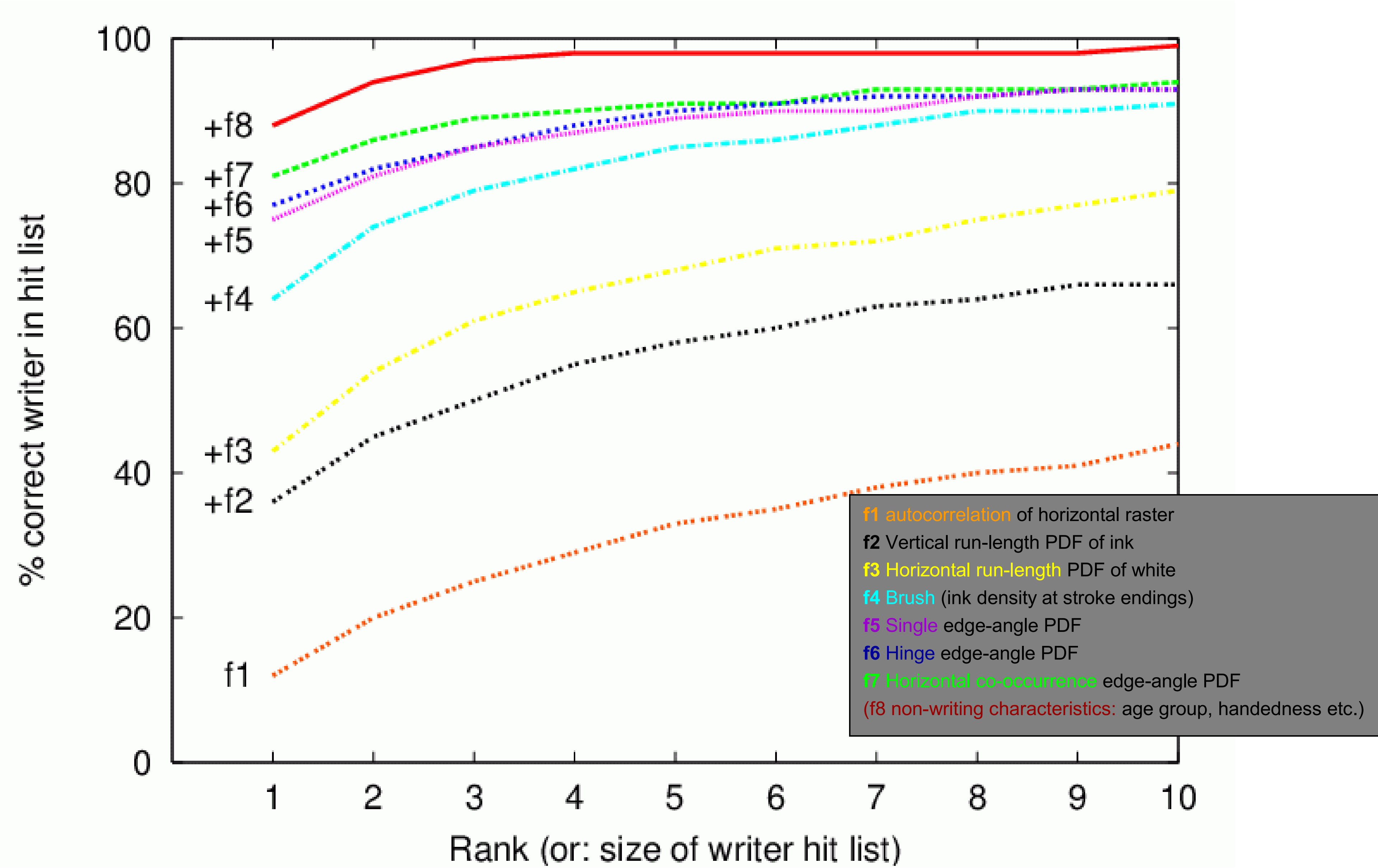
f6 Hinge edge-angle PDF

f7 Horizontal co-occurrence edge-angle PDF

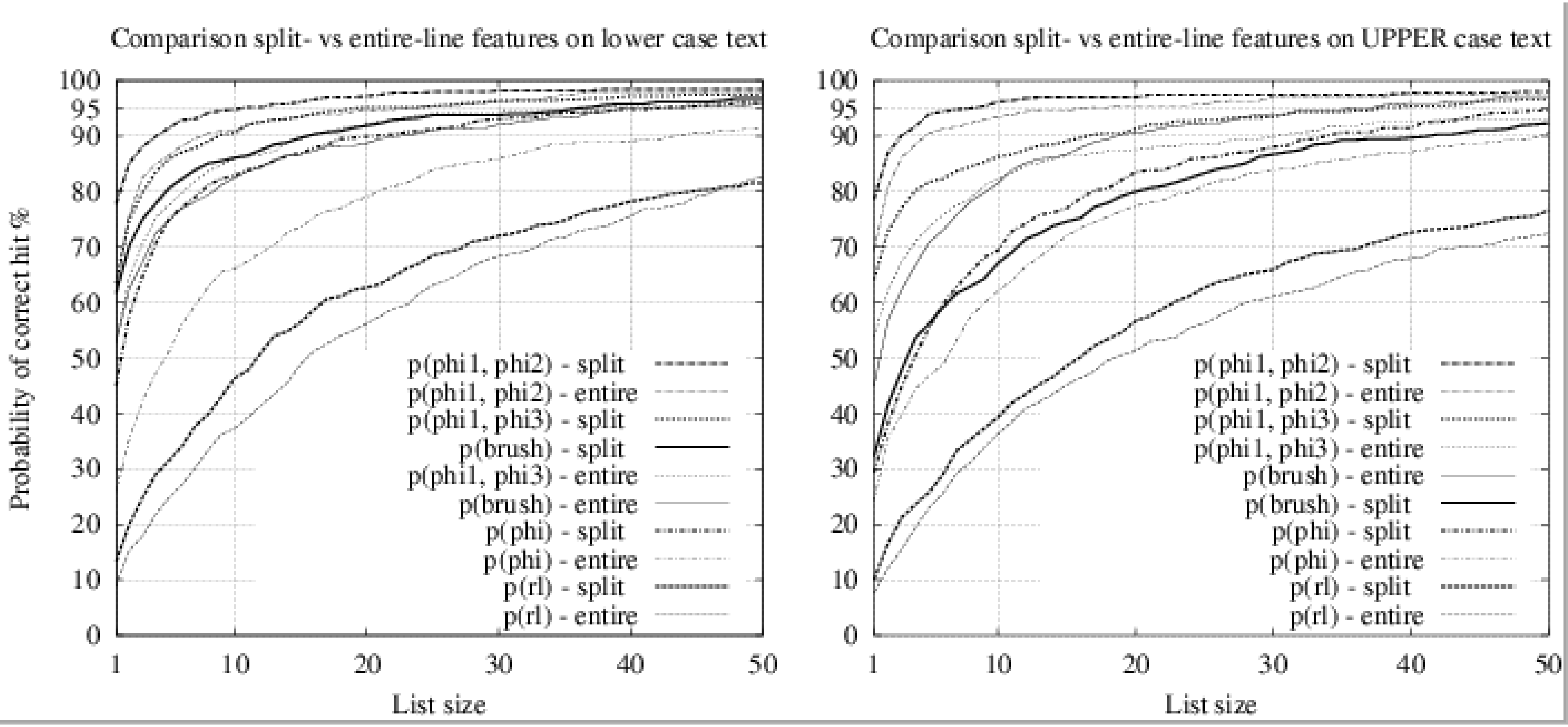
(f8 non-writing characteristics: age group, handedness etc.)

Copyright © 2004 L. Schomaker

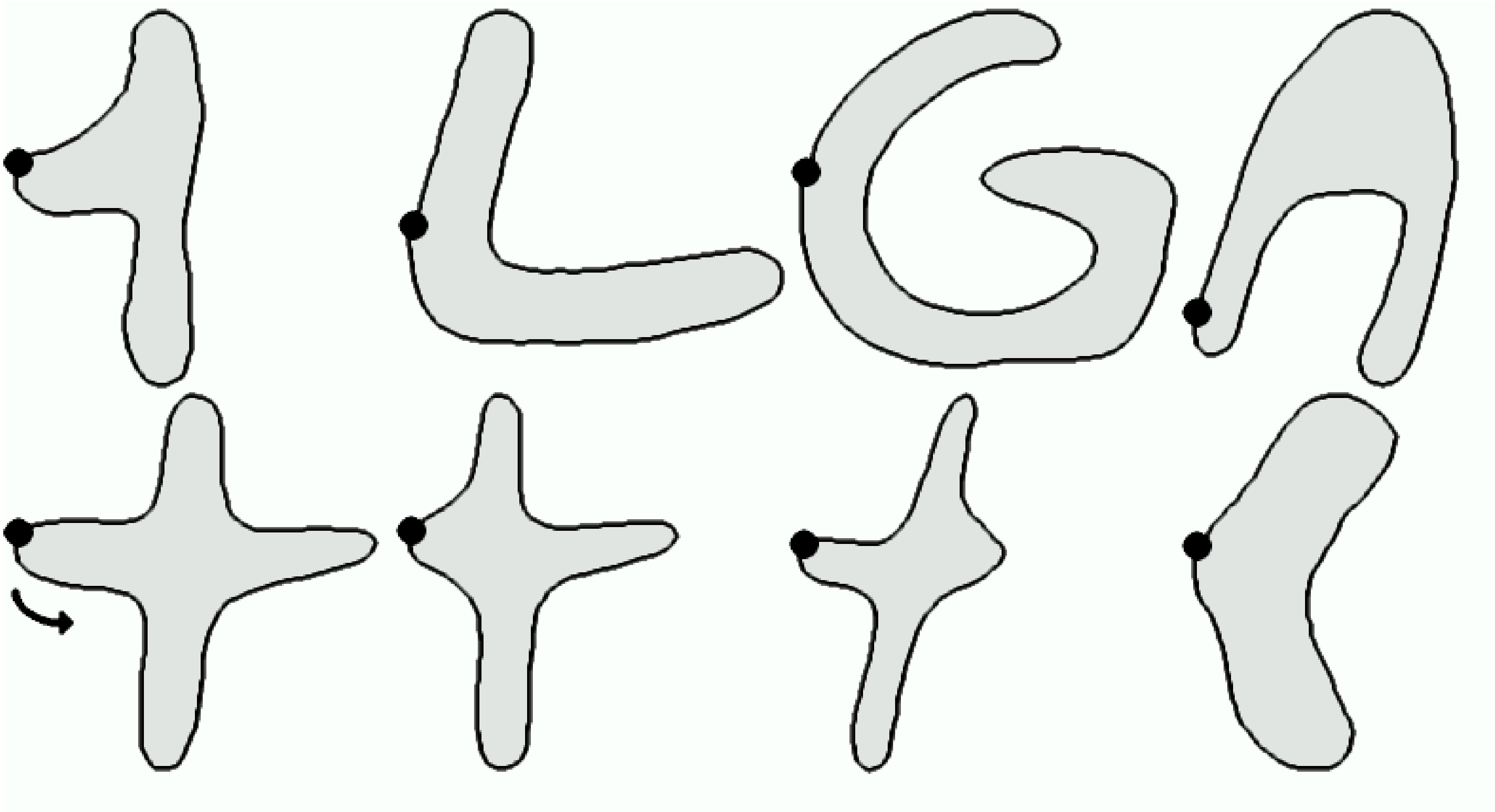
Ordered-Borda combination results



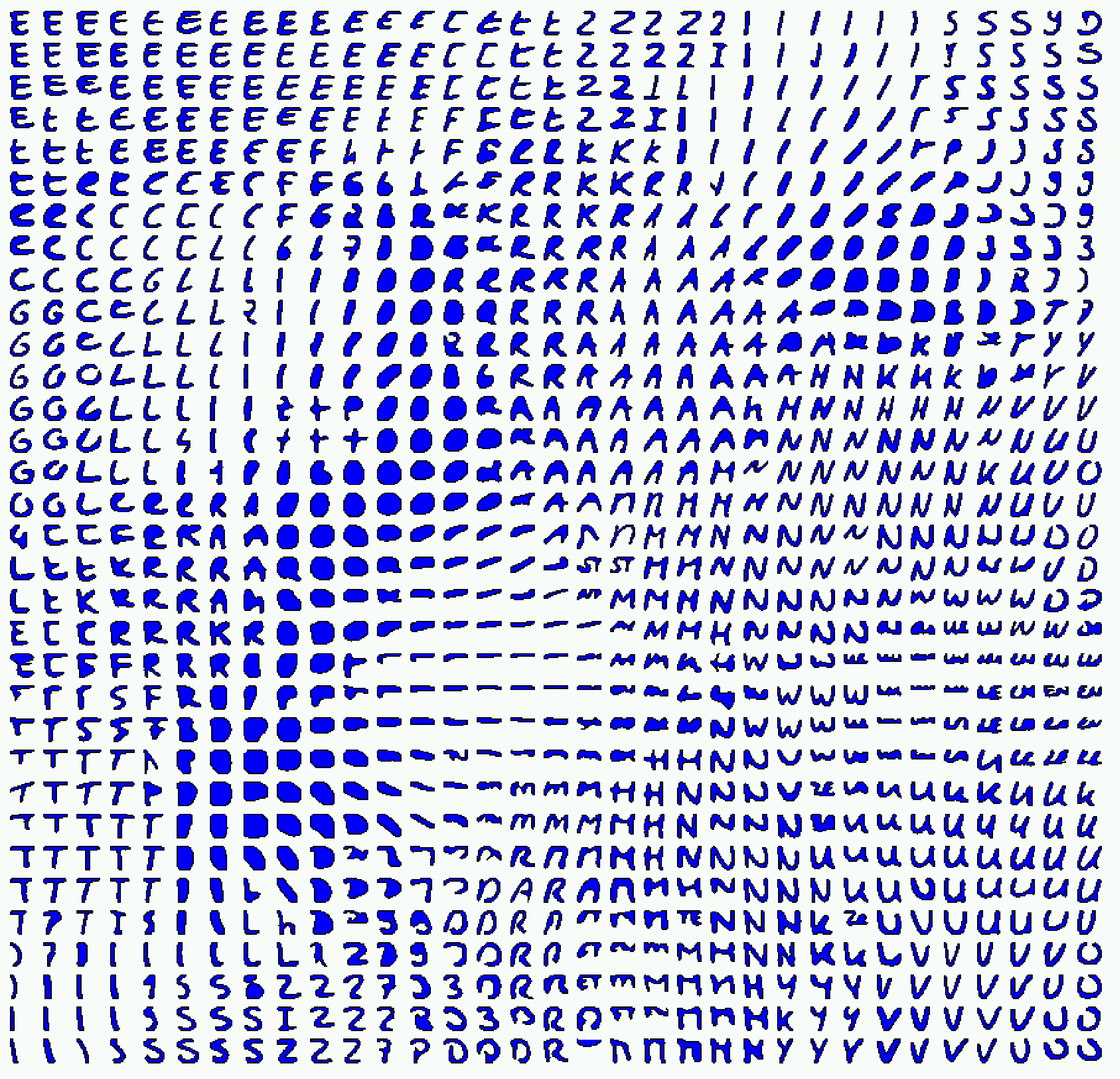
Split-line hinge edge-angle PDF



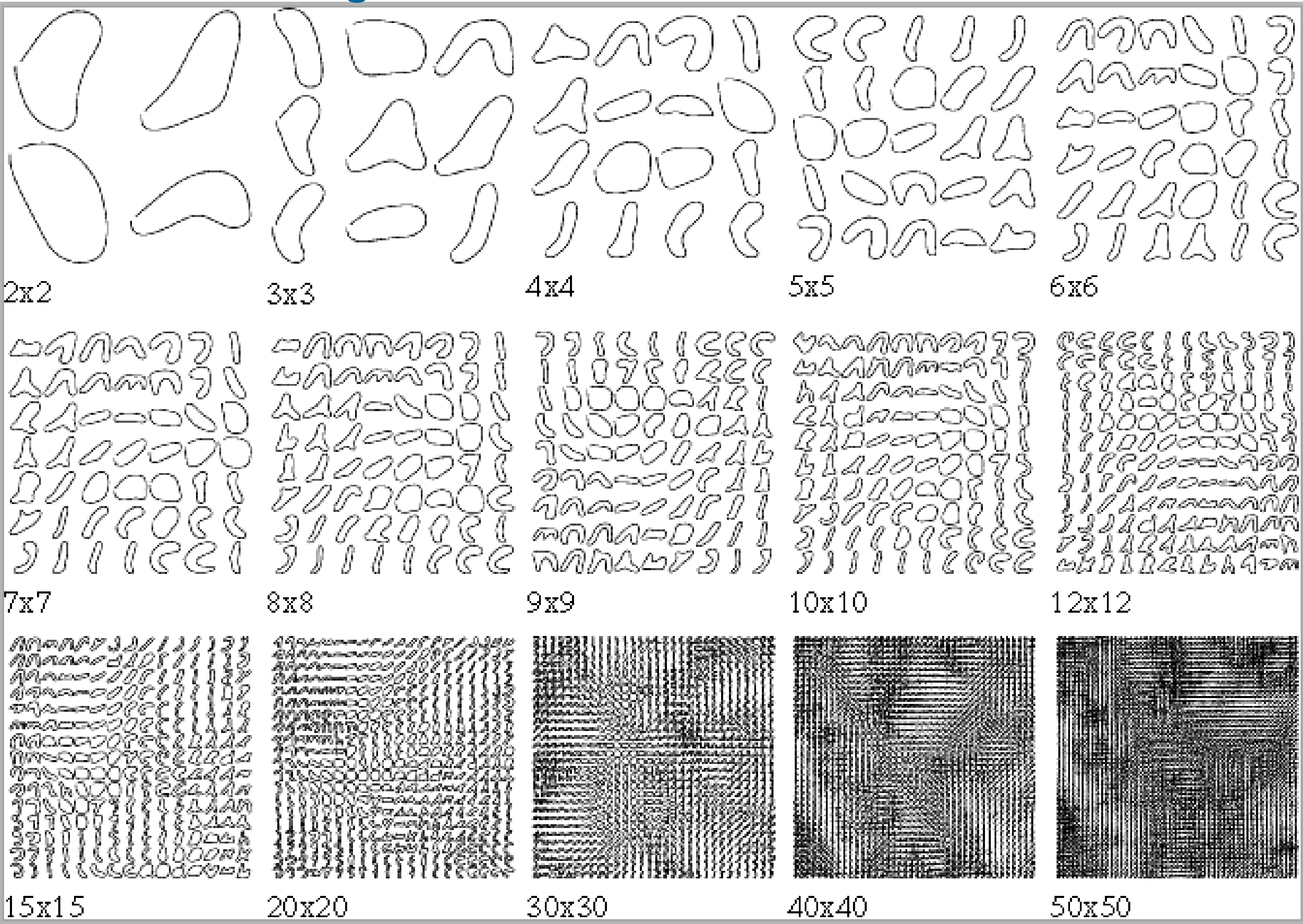
Connected-Component Contours (CO^3)

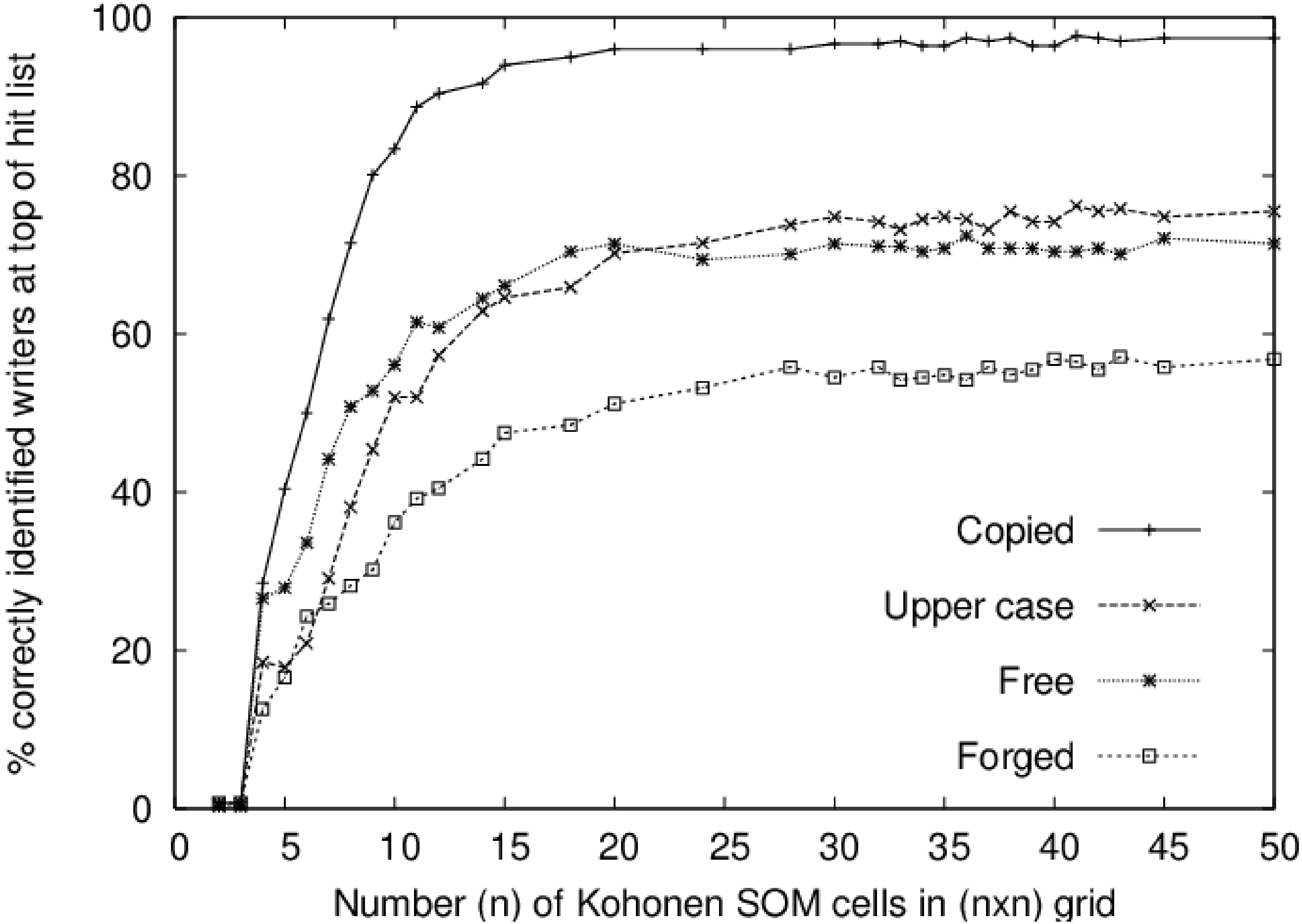


Kohonen SOFM (33x33) of CO^3s



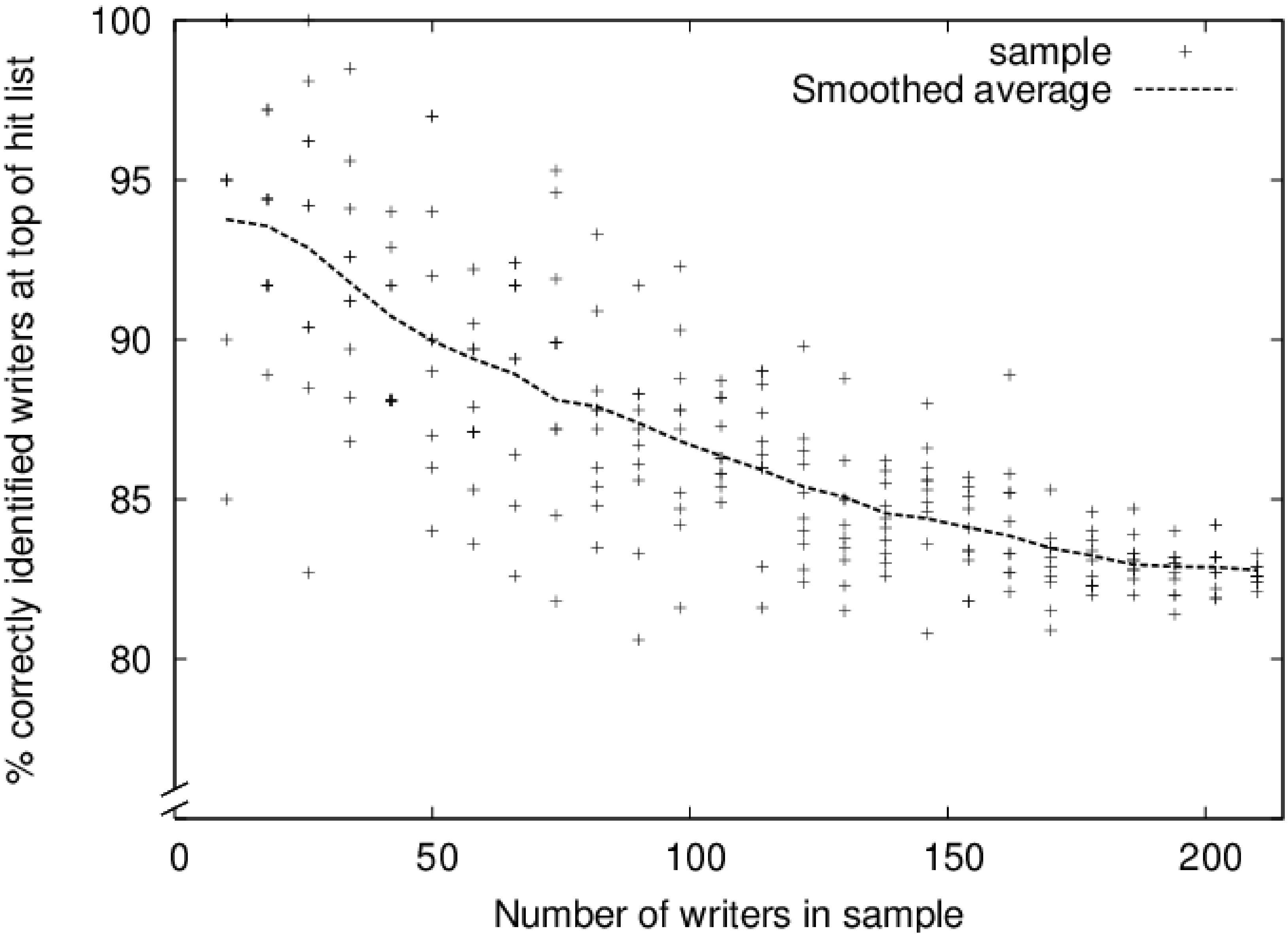
Fragmented CO³ Kohonen SOFMs





% correctly identified writers at top of hit list
Number (n) of Kohonen SOM cells in (nxn) grid
Copied
Upper case
Free
Forged
0
20
40
60
80
100
0
5
10
15
20
25
30
35
40
45
50

FCO³ test on independent ImUnipen data (40x40 net)



Feature vector combination, Hinge and FCO3

- Adjoined PDFs: Hinge and FCO3
- 150 writers, 300 samples, leave one out, 1NN Hamming
- mixed lower case

(% writers correctly identified in top-n list)

Top1	**97.00**
2	**98.00**
3	**98.33**
4	**99.00**
5	**99.00**
6	**99.33**
7	**99.33**
8	**99.33**
9	**99.33**
10	**99.67**

Copyright © 2004 L. Schomaker

The archives of the Cabinet of the Queen at the Dutch National Archive: a challenge for script recognition!

Kunstmatige Intelligentie / RuG

Digitization of the cultural heritage

- scans of handwritten documents
- manually and superficially annotated at the document or page level (not words, characters)

Copyright © 2004 L. Schomaker

Digitization of the cultural heritage

- scans of handwritten documents
- manually and superficially annotated at the document or page level (not words, characters)
- "OCR" of handwriting to generate a "transcription"?

Copyright © 2004 L. Schomaker

Digitization of the cultural heritage

- “OCR” of handwriting to generate a “transcription”?
- ➔ not possible
- ➔ what **IS** possible in **PR**?

Copyright © 2004 L. Schomaker

Digitization of the cultural heritage

- "OCR" of handwriting to generate a "transcription"?

- ➔ not possible
- ➔ what **IS** possible in **PR**?
- Word spotting

Copyright © 2004 L. Schomaker

Handwritten historical documents

- In general, very difficult: documents written by many writers ➔ too much style variation

- Cabinet of the queen:

one archivist, writing hand-written indices and summaries for years

Copyright © 2004 L. Schomaker

A pilot collection

- 48 scanned pages (larger than A4)
- from book: left page vs right page layout
- very strict constraints
  - layout
  - items: imprint, header, footer
  - dates
  - text blocks

Copyright © 2004 L. Schomaker

1903 Dm & No Kon: Kol: Mil: Invalidenhuis.

Jan. 10 43 Rappt KD 7 Jan. No 43, aanbieding van het verslag betreffende het Kon. Kol. Mil. Invalidenhuis op Bronbeek over het 4e kwartaal 1902
— Notificatie

April 16 23 Rappt KD 7 April No 53, als boven over het 1e kw. 1903
— Notificatie

Juli 15 65 Rappt KD 10 Juli No 28, als boven over het 2e kw. 1903.
— Notificatie.

Octb 12 7 Rappt KD 6 Octb No 46, als boven over het 3e kw. 1903
— Notificatie.

Juli 20 53 Rappt KD 14 Juli No 30, tot beschikbaarstelling van een bedrag van f 600.- ter bestrijding der kosten van huisvesting, voeding verpleging enz in het Kon: Kol. Mil. Invalidenhuis op Bronbeek van in West-Indie gediend hebbende Militairen.
— Besluit fiat

Octb 29 21 Rappt KD 20 Octb No 62, om machtiging om twee van de in het museum van het Kon. Kol. Mil. Invalidenhuis te Bronbeek bewaard wordende, in Oost-Indie veroverde, lela's af te staan aan de zich te Arnhem gevormde Commissie tot het oprichten van een gedenkteeken voor wijlen den Luitenant Generaal K. van der Heyden, ten einde aan dat gedenkteeken, dat geplaatst zal worden aan den aldaar nieuw aangelegden Karel van der Heydenweg, op den hooge naam den Berkenheuvel, een met de krijgsverdiensten van den overledene overeenkomend eigenaardig karakter te geven.
— Fiat missive

—

Notificatie

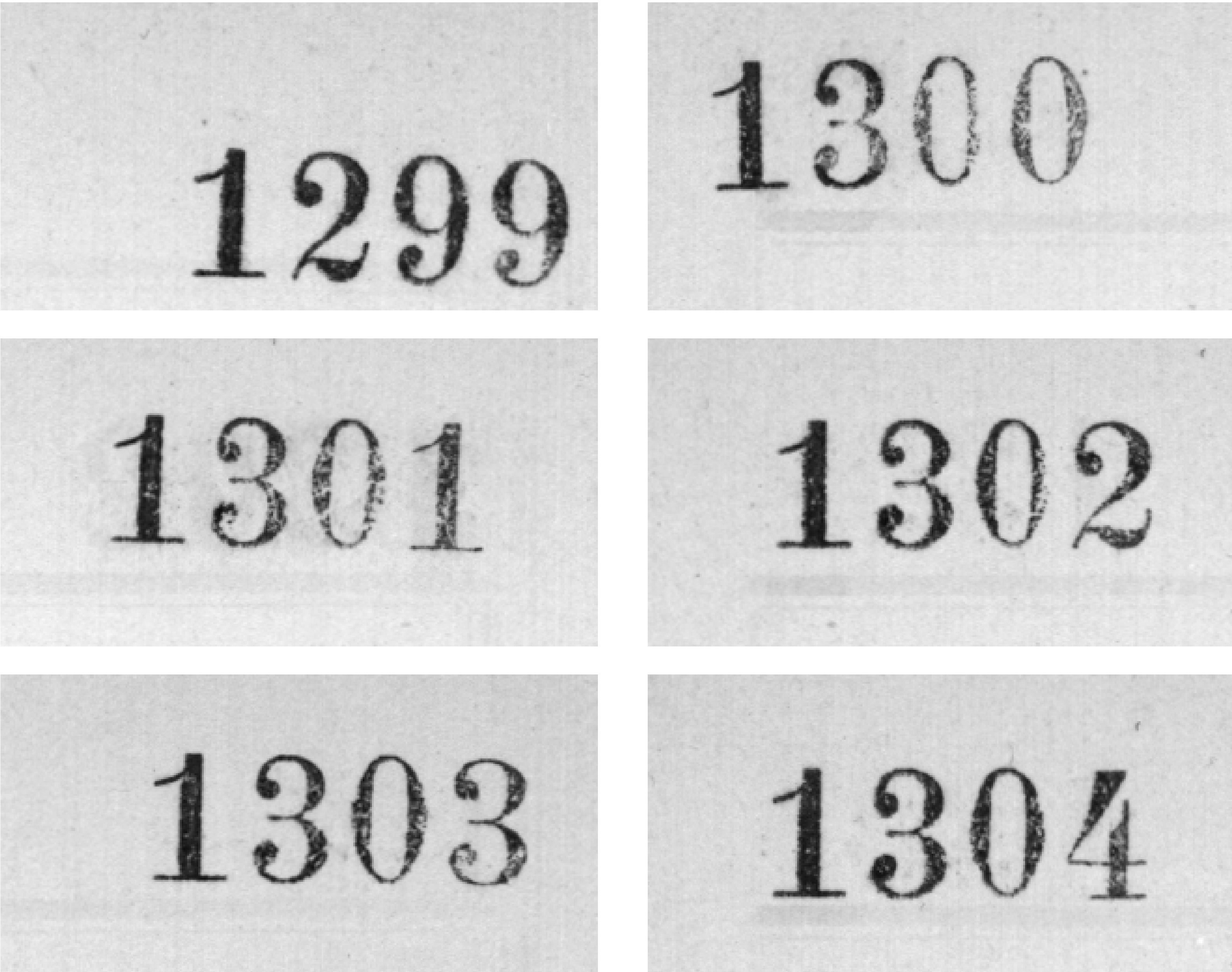
April 16 23 Rappt KD 7 April N: 53, als boven over het 1e kwl 1903
— Notificatie

Juli 15 65 Rappt KD 10 Juli N: 28, als boven over het 2e kwl 1903.
— Notificatie.

Octb 12 7 Rappt KD 6 Octb N: 46, als boven over het 3e kwl 1903
— Notificatie.

Juli 20 53 Rappt KD 14 Juli N: 30, tot beschikbaar stelling van een bedrag van f 600.- ter bestrijding der kosten van huisvesting, voeding verpleging enz in het Kon: Kol. Mil. Invalidenhuis op Bronbeek van in West-Indie gediend hebbende Militairen.
— Besluit fiat

Octb 29 21 Rappt KD 20 Octb N: 62, om machtiging om twee van de in het museum van het Kon. Kol. Mil. Invalidenhuis te Bronbeek bewaard wordende, in Oost-Indie veroverde, lila's af te staan aan de zich te Arnhem gevormde Commissie tot het oprichten van een gedenkteeken voor wijlen den Luitenant Ge



The easy part, imprints: already difficult

Mouth 1642 Rappd

lenes fo

and

to

The transcription

Novb 16 42 Rappt. MD 11 Novb n° 108 tot het toekennen
eener jaarlijksche vergoeding voor bureaukosten
aan den Secretaris van de Commissie ingesteld
tot herziening van het reglement nopens de
burgerlijke werklieden bij de Inrichtingen
der Artillerie, enz.

Besluit fiat

Copyright © 2004 L. Schomaker

Summarizing

- Pattern recognition problems are very interesting, scientifically
- In the user interface, their applicability is limited for standard & generalized input
- Customized applications are always possible, but they will be constrained

Copyright © 2004 L. Schomaker

Summarizing (2)

- There is a need for fundamental cognitive research in the area of human-machine communication…
- …in order to pave the way for innovative interaction and communication patterns.
- Hand, mouth and tongue are still the most likely candidates for user-interaction concepts without surgery

Copyright © 2004 L. Schomaker

Summarizing (3)

- As regards PR & UI:
- we concentrate on what works
- with lowered ambition levels

Copyright © 2004 L. Schomaker

Recent papers

- L. Schomaker, M. Bulacu & M. van Erp (2003). *Sparse-parametric writer identification using heterogeneous feature groups.* ICIP'2003: IEEE International Conference on Image Processing (Vol. I), pp. (I) 545-548.
- M. Bulacu, L. Schomaker & L. Vuurpijl (2003). *Writer identification using edge-based directional features.* ICDAR'2003: International Conference on Document Analysis and Recognition, pp. 937-941.
- M. Bulacu & L. Schomaker (2003). *Writer Style from Oriented Edge Fragments.* In: N. Petkov & M.A. Westenberg (Eds.), LNCS 2756 - Computer Analysis of Images and Patterns, pp. 460-469.
- L. Schomaker & M. Bulacu (2004). Automatic writer identification using connected-component contours and edge-based features of upper-case Western script. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 26(6), June 2004, pp.xx-xxx.
- K. Franke, L. Schomaker, C. Veenhuis, C. Taubenheim, I. Guyon, L. Vuurpijl, M. van Erp, and G. Zwarts (2003). *WANDA: A generic framework applied in forensic handwriting analysis and writer identification*, Design and Application of Hybrid Intelligent Systems (HIS'03), pages 927-938. Amsterdam: IOS Press.

AI RuG

Copyright © 2004 L. Schomaker

bcn

Selected papers

- Plamondon, R., Lopresti, D.P., Schomaker, L.R.B. and Srihari, R. (1999). On-line handwriting recognition. In: J.G. Webster (Ed.). Wiley Encyclopedia of Electrical & Electronics Engineering, 123-146, New York: Wiley, ISBN 0-471-13946-7.
- Schomaker, L., Vuurpijl, L. & de Leau, E. (1999). New use for the pen: outline-based image queries. Proceedings of the 5th International Conference on Document Analysis and Recognition (ICDAR '99). Piscataway (NJ): IEEE. pp. 293-296.
- Schomaker, L.R.B. (1998). From handwriting analysis to pen-computer applications. IEE Electronics Communication Engineering Journal, 10(3), pp. 93-102.
- Schomaker, L.R.B., & Van Galen, G.P. (1996). Computer models of handwriting. In: Dijkstra & De Smedt (Eds.), Computational Psycholinguistics: AI and connectionist models of human language processing (pp. 386-420). London: Taylor & Francis.
- Schomaker, L.R.B, Muench, S. & Hartung, K. (Eds). (1995). A Taxonomy of Multimodal Interaction in the Human Information Processing System. Report of the Esprit Project 8579 MIAMI (187 p.), Nijmegen: NICI.
- Schomaker, L.R.B. (1993). Using Stroke- or Character-based Self-organizing Maps in the Recognition of On-line, Connected Cursive Script. Pattern Recognition , 26(3), 443-450.

AI RuG

Copyright © 2004 L. Schomaker

bcn