

HANDWRITING RECOGNITION A PERSPECTIVE ON TWO DECADES OF INNOVATIONS



Venu Govíndaraju Uníversíty at Buffalo New York, USA











An Abridged History of Writing Instruments



FINGER (IN SAND) OR CLAY)



QUILL



PENCIL



BALL POINT PEN



FINGER (ON TOUCHSCREEN)

Vie





G



Full Circle













Students **WITHOUT** consistent exposure to handwriting **are more likely to have problems** retrieving letters from memory; spelling accurately; extracting meaning from text or lecture; and interpreting the context of words and phrases



Elementary-age students who wrote compositions by hand rather than by keyboarding, one researcher found, wrote **faster**, wrote **longer pieces**, and expressed **more ideas**



Handwriters exhibit better comprehension of the content and were more attentive and involved during the class discussions versus classmates who took notes by computer



During early childhood, writing letters **improves letter recognition**, as shown in brain imaging studies; and teaching handwriting leads to **improved reading**, as shown in instructional studies





Structured handwriting lessons lead to improved writing performance, academic success, and overall student self-esteem

a 2012 study concluded that **one in three** participants had not been required to produce something in handwriting for **more than half a year**





A 2012 study found that 33% of people had difficulty reading their own handwriting



Research shows that you remember information better when you write it by hand than when you type it





A psych professor at Dominican University of California found that people who **wrote** down their goals, **shared** them with others, and **maintained accountability** for their goals were **33% more** likely to achieve them, versus those who just formulated goals



National statistics point to underachievement in writing as a condition that leads to school **dropout**, **grade retention**, and **failure to graduate** from high school or college



Air Writing











E-signatures and Finger Doodles







SEND A NEW DOCUMENT











This AI Algorithm Learns Simple Tasks as Fast as We Do

Software that learns to recognize written characters from just one example may point the way towards more powerful, more humanlike artificial intelligence. *MIT Technology Review, Dec 2015*.



All Things Handwritten

Jee hälte Ihre Bitte genn friher erfällö, men Ihre Briefe micht in der Tille ihrer Geschwirter um meinen Blickefeld verschwurde wären. Dem Graphologie hut mich stets interessiert, wenn ich mich unde nie zystematisch dafür interessiert habe, ich mache es mit Gebriften wie mit Gesichten, ich unteile instinktiv. Hedeachtungsvoll A. Einstein.

S. Winchester: October 28. 1755. Parole Hampton. The officers who came down from Fort bumberland with bolonel Washington, are immediately to go Recruiting, and they are allowed until the S. of December; at which time if they do not punctually appear at the place of Rendezvous afsigned them they will be tried by a bourt Martial, for disobedience of Orders.





Key Innovations





LEXICONS







Lexicon

Postal



10-10-20 Rule

ML Success Story - Highlighted in the CCC Symposium on "Computing Research That Changed the World" (2009); Seminal work of Kim, Govindaraju (PAMI 1997) at the core of the technology





Lexicon

Dynamic Lexicon

ypical Mail-piece Address Block :	Lexicon Entry ((Street Name)	Addon
Meg Knowles 74 Livingston St Bullalo NY	BRADLEY ST COLONIAL CIR DEWITT ST GRANT ST HAMPSHIRE ST HAWLEY ST LAFAYETTE AVE	1024 1467 1534 1905 2012 1053 1347
ZIP and Primary Number Recognition	MANCHESTER PL PO BOX POOLEY PL POTOMAC AVE PUTNAM ST	1251 0074 1128 1146 1650
Primary Number: 74 Query		

30% ZIPs contain less than 100 street names; Max streets returned is 3,071

H. Xue, V. Govindaraju, "On the dependence of handwritten word recognizers on lexicons", *IEEE Transactions Pattern Analysis and Machine Intelligence*, IEEE Computer Society Press, 24(12): 1553-1564 (2002)



Medical Faxed prescriptions







Lexicons

Lexicon

ONE	тмо	THREE.
FOUR	FIVE.	SIX
SEVEN	EIGHT	NINE
TEN	ELEVEN	TWELVE
THIRTEEN	FOURTEEN	FIFTEEN
SIXTEEN	SEVENTEEN	EIGHTEEN
NINETEEN	TWENTY	THIRTY
FORTY	FIFTY	SIXTY
SEVENTY	EIGHTY	NINETY
HUNDRED	THOUSAND	LAKH
LAC	LACS	AND
ONLY	CRORE	LAKHS



- Census forms recognition
 Lexicon of professions
- Prescription forms Lexicon of medicines
- Registry of land ownerships Lexicon of owner

R. Jayadevan, U. Pal, and F. Kimura: Recognition of Words from Legal Amounts of Indian Bank Cheques. ICFHR 2010





HOUSE

SCHOOL



G. Kim, Venu Govindaraju: A Lexicon Driven Approach to Handwritten Word Recognition for Real-Time Applications. IEEE Trans. Pattern Anal. Mach. Intell. 19(4): 366-379 (1997)

21





Interactive Features

Lexicon



Hyderabad

a. Delhi

b. Kolkatta c. Patna d. Dehradun





Lexicon

Interactive Features



- a) Amherst
- b) Buffalo
- c) Boston
- d) None of the above

S. Madhvanath, V. Govindaraju, "The role of holistic paradigms in handwritten word recognition", *IEEE Transactions Pattern Analysis and Machine Intelligence,* IEEE Computer Society Press, 23(2): 149-164 (2001)





Lexicon

Multiresolution features



PJ. Park, V. Govindaraju, and S. Srihari,: OCR in a Hierarchical Feature Space. IEEE Trans. Pattern Anal. Mach. Intell. 22(4): 400-407 (2000)





Lookup table: 2⁹ x 10 x 85 (quad tree, 4 levels)





Interactive Lexicons

Lexicon









IMPACT

Lexicon

|--|

Lex size	%
10	96
100	91
1000 (Top 50)	80 (98)
20000 (Top 100)	62 (94)
Postal Encoding	40

Today

Lex size	%
10	~99
100	~99
1000	~95
Postal Encoding	95



FUSION

- How does one determine x, y ?
- How does one determine the ordering of C1, C2, ... Cn
- Is there a difference based on whether it is an Identification task or Verification task?
- Can it be a generalized principled approach?

Sergey Tulyakov, Venu Govindaraju: Use of Identification Trial Statistics for the Combination of Biometric Matchers. IEEE Transactions on Information Forensics and Security 3(4):719-733 (2008)

Fusion

Lexicon Density

 Lexicon size as feature 	Lexicon 1:	Lexicon 2:
 Different lexicons with the 	1: He	1: Me
same size have inconsistent	2: Me	2: Memo
results	3: No	3: Memory
	4: So	4: Memorize
 Lexicon density as metric 	5: We	5: Memoryless
Recognizer specific		

• Performance is consistent with the same level of density

V. Govindaraju, P. Slavik*, and H. Xue, "Lexicon density as a measure for performance evaluation of handwritten recognizers", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE Computer Society Press, 24(6): 789-800 (2002)

Bootstrapping for Large Lexicons

R. Milewski, A. Bharadwaj, and and V. Govindaraju, "Automatic Recognition of Handwritten Medical Forms for Search Engines", International Journal of Document Analysis and Recognition, Springer, 2009.

Pruning Lexicons

												,	otor
			é	i set	i se							or	7-
		10	305	50	AST						Ca	S.	
	OIO	STI C	10 mil	CUN YO	og Ne	5 80°	* 00°	Nº en	S AL	\$ 250	Ş0		
(bl, no)	0	0	0	.95	0	0	0	0	0	0			
(bl, ns)	.07	0	0	.34	0	.34	0	0	0	2			
(bl, nt)	0	0	0	.95	0	0	0	0	0	1			
(bl, nr)	0	.06	0	.44	0	0	0	.20	0	1			
(bl, mo)	0	0	0	.33	0	.56	0	0	0	1			
(bl, mu)	0	0	0	.41	0	0	0	.51	0	0			
(bl, mt)	0	.05	0	. <mark>4</mark> 4	0	.19	0	0	0	2			
(bl, mh)	0	0	0	.95	0	0	0	0	0	0			
(bo, no)	.18	0	0	.57	0	.25	0	0	0	3			
(bo, ns)	0	0	0	.42	0	.49	0	0	0	1			
	8			9 - S					7	1			
ŏ			١	/ecto	or G	isn	ow	adde	ed to	the			
ě		ŏ		origir	nal	norm	aliz	ed/w	eigl	nted			
-		-				ma	асгіх						

	CLT to RLT	CL to RL	CLT to ALT
HR	↑7.48%	↑7.42%	↑17.58%
Error Rate	↓10.78%	↓10.88%	↓24.53%

Fusion

Principled Approach

Map: M matchers x N classes scores to N combined scores

Score Matrix

Sergey Tulyakov, Venu Govindaraju: Use of Identification Trial Statistics for the Combination of Biometric Matchers. IEEE Transactions on Information Forensics and Security 3(4):719-733 (2008)

Fusion

$$S_i = f_V$$
 $S_i > \theta$? Accept : Reject
 $S_i = f_N$ $S_i \operatorname{arg\,max}_{i \ i=1,...,N}$

$$S_i = f_V > \theta? \qquad \Theta = \prod_{j=trials} \frac{p(s_j^k \mid C^k)}{p(s_j^k \mid \overline{C}^k)}$$

Likelihood Ratio

 $S_i = f_N$?

C1	C2	C1 ^ C2	C1 v C2	LR	Weighted sum	NN
54.8	77.2	48.9	83.0	68.8	81.6	81.7

Independent Matchers

Sergey Tulyakov, Chaohong Wu, Venu Govindaraju: On the Difference between Optimal Combination Functions for Verification and Identification Systems. IJPRAI 24(2): 173-191 (2010)

IMPACT Theoretical

INDEXING AND RETRIEVAL

Retrieval

Historical Documents

I hereby assign I convey to Sam's Madison President Led States the inthing med requant, . An, of Armie called Joh during the remaining term of his server is from the 13 last part the unideration of two hund is for the grand one Do last part then he was delivered to the sac Tamer this 19 " day of april 1809. at mon healton Vorginia. I hereby assign & convey to lames ted States the inthin named servant, Jo. during the remaining term of his service for the consideration of two hundres last part then he was delivered to the This ig " day of april 180g. at Month

Transcript mapping

have duly received your

favor

Retrieval

Name aug. 8 1787 Sur Sir Paris Aug. 8. 1787. There suly received your favor of Sure & and immediately I have duly received your favor of June 6. and immediately communicated at a unlimits to a member of the Academy. He told me communicated it's contents to a member of the Academy. he told me they had received the other igning of your memorial thick you mention they had received the other copy of your memorial which you mention to have unt this another channel, that your ideas and con to have sent thro' another channel, that your ideas were not conway it is suptrictly of to enable them to decide finally -veyed so explicitly as to enable them to decide finally on their merit, but that they had made an entry in their journals to preserve but that they had made an entry in their prismant to you the claim to the original idea. as far as we can conjecture it you the claim to the original class . as far as we can protine it here, we imagine you make a table of Variations of the needle for all the loss, secondance you make a table of deviations of the need for all office different meridians whatever, to apply this table to use in the voiage different meridians chateror to aggily this table to use in the covered between America & Europe, suppose the variation to increase a degree lationen America & Lunope, suppose The variation to inconner a Degra in every 160 miles, two difficulties occur. 1. a ready & accurate method o in many 160 miles two difficulties owner to ready Vournets and of Juday the variation of the place . a an instrument is perfect as that (the the degree on it shall require to miles) it shall give the parts of the degree to minutify as to answer the purposes of the nanigator. the variation of the well at Paris achielly is 20" H. I wake as doubt but you have provided against the doubts earling have and I shall be have that my question may have the borner of permiting the old world that that so long sought in vain. Sam duely no count 2 30 with much request Six main most Readice to hermalote to

Mr. John Churchman Philadelphia.

Catalin I. Tomai, Bin Zhang, Venu Govindaraju: Transcript mapping for historic handwritten document images. IWFHR 2002: 413-418

Т

Documents

Retrieval

Word Spotting

Letters Orders and Instructions Pounts mis 1755: p.12, p.22, ... to so summathe as wedrey pour place the king the proper I repeation and serving that all and need in presented up with dispatch form for think it abarable to infirm your Kinow Express: p.12, p.14, .. fit have chat is will compared with a Good to make aton I was down, that all fing dow the Fort a Major, iron apprented dejetand to any Required there of ours also about long constates in a small tang some had qual find better for the lift on May: p12, p.45, ... for man approvable its of Systember Captain - Houses for as and to camp crows get doubthat, I should be also if your three would fix it in haplain Some of Captain For and in an average of suppose with the approve between prog 5. 18 - Fight time will expend to have the and where Colonet Hopkows in a lo when Indiany this Winker And her down to go sty of be bad any such shought Stehen a a maple - Burry to might be quind wated If your Hear think proper Since 1958, 13 Cabour life Feers and Clustering Labeling Peresses have been created. Most Cabour *1755.* "1755" the de. de. sendiment would # still favour the abolition chat . . . chat chat of the House of Cords, but while it remains May "May" Labour has to have an adequate number of Express

Impact

Retrieval

- Imperfect Segmentation
- Multiple Writers
- Use recognition scores

$$fr\hat{e}q_{i,j} = \sum_{\vec{w}} P(\vec{w} \mid \vec{c}) \cdot \sum_{\vec{\tau}} P(\vec{\tau} \mid \vec{w}) \cdot \#_{t_i}(\vec{\tau})$$

	Lexicon Controlled Historical Medical		Controlled		Historical		Postal	Checks
$\overline{70}$		10K	1K	10K	1K	4K	100	40
	Top 1	57	67	12	28	20	95	99
B14	Тор 10	74	75	32	72	42	99	99

CAPTCHAS, ACCENTS AND SECURITY

MF Swy (Yahoo!) (Google)

Neuromotor Model

ith component trajectory :

$$\varphi_i(t) = \theta_{si} + \frac{(\theta_{ei} - \theta_{si})}{D_i} \int_0^t v_{ti}(\tau) d\tau$$

Chetan Ramaiah, Réjean Plamondonm Venu Govindaraju: A Sigma-Lognormal Model for Handwritten Text CAPTCHA Generation. ICPR 2014: 250-255

Accents

Security

TH many from time to time issue a Variation Order to vary any of the comprehensive Maintenance Services and the contractor shall carry out each Variation, and be bound by the same conditions, insojar as they are applicable, as though the Variation is stated in the contract. If the Variation in the opinion of the contract. If the Variation in the contractor will involve an addition to or deduction from the Contract sum, the Contractor shall, as soon as reasonably possible and in any event not later then fourteen (14) days from the clate of the Variation Order, advise TH in writing to that effect.

मेरा नाम ऋषि जैन है। मैं कहना चार में अध्ययनरत एक उच्च विचारों वाला व्यात्र हूँ। मुझे विज्ञान और रवगोलशास्त्र पढ़ने में काफी रुचि है। मैं बड़ा होकर अलब टें आईस्टाईन की तरह मशहूर होना साहता हूँ। यर पर मेरी माँ मुरो कहती है, बेय, अधक परिश्रम का फल हमेरा मीत्र होता है। मुखे पूर्ण विश्वास है कि एक दिन में महान् वेज्ञानिक बन्द्रा और अपने माता -पिता का नाम विद्वभर में रौशन करुगा।

The furthest distance in the world is not between life and death but when I stand in front of you get don't know that I love

摩泥末解玉骢 骄. 武欲醉眠芽 91省-溪风月, 莫教降碎碌. 解鞍欹兄祝绿杨桥, 科宇-声春晚

Security

خفافتيم تت ختيج

Cauce min el mara ADDRESSES. ADDRESSES NAME. 5 NAME. STREET, NO STREET, NO. CITY. CITY. TELEPHONE NO. TELEPHONE /NO NAME. NAME. STREET, STREET, NO. CITY. CITY. TELEPHONE NO. TELEPHONE NO. NAME. NAME. STREET, NO. STREET, NO. CITY. rethe CITY. TELEPHONE NO. TELEPHONE NO. NAME. NAME. STREET, NO. STREET, NO. CITY. CITY. TELEPHONE NO. TELEPHONE NO. NAME. bodia. NAME. STREET, STREET, NO. CITY CITY. 1922 TELEPHON

Marianne Craig Moore was an American Modernist poet, critic, translator, and editor. Her poetry is noted for formal innovation, precise diction, irony, and wit

Marianne Moore Archive: Notebooks Projects Prototype Viewer

File Description Notebook 07.03.11 Notebook Editor: Cristanne Miller

Notebook Description

Repository: Marianne Moore Collection, Rosenbach Museum and Library, Philadelphia PA Dates: 1929-1940

Note: Yearbook 1921, red leather binding. Includes 8 pages of pre-printed information of various kinds, e.g. Time Differences, Thermometers; Greatest Altitude in Each State; followed by lined calendar pages divided in half so that each leaf contains two dates. First image with MM's writing is the inside front cover. RML number written in pencil: 1250-27

TextLink 🕉 Figure O 07-03-11 0036-recto Chapter headings-The quiet of the written page Few balances no debts Salve next to last A valuable clipping The Leopard Chapter MEMORANDA Illight MEMORANDA the quiet of the written page Crickets Chapter I Critical Standard Poets Protestations of Integrity. Edinburgh Job's balances no debts Stonly from Vagetakes grow & live " wea grow and this and this f. [300 Talay Salve Permanent Salutation vehiable cliffic Beware of Pickpockets (circus So Thopas Sapieto + unforment deluce Crickets chatter - Critica Joba Protestations of Latings: 19- Edice Pretty Name Insperialista the & A. a English invarent Salutate Bravo, Bravado Beware of Rickforchels Hercus My Actions are Open A good Eye Pretty Waine Neves Complain, Neves Expla Pleasure is less than Happiness Brand Bravado dee la Musles It was the month of April, the Bumpkins 9 Tout Elle Pleasure is less then Happinese 7 intereste call it May (Cooper) + malices everything It was the month of afrit, the Burg The one aggravation is war) He who sides one types can never divid (He who rides on a tiger can never dismount A good short alla sorter a born Billiard take green den miter Dean Inge n creation -Assisted to motion at Development Billiard table green an above in the soar grape he Balfled Rat Squeeks The Barowies in Scotland mites of creation-cf Baboon Man P The spint of 76 Dr Popey No thus to is a good words own unless he know how to present fis forgive us for flaws The Baffled Rat Squeaks Phrases from the Bible No hunter is a good woodsman unless he knows how to prevent fire officially noted (Pratt Library book)n 1:1 🚺 👄 💻 33%

Multilingual MADCAT

Transform based method works for both gray scale and binary images

العراق دخلت المح سوريم ٥٠ شاهنه للنغل البري على شكل وتل منتطع المثنية دعد ومن هؤلاء السواة وكانلا لا يمرمون مدا عمين في هذ المناعات المساعد في الشاهنات مسلمان معودل في بعد اد وجلبت الى المساعقية وكان محقع المنا سرات العراض كان قطعوا مساخة معينه موقعهم دجاد المينا سرات

نقاطه ΛJ لبتاء تالمارس العاب الغيبو ، مانا كند فعلا. يح الآغزين حتى لو أختلفت الآداء والدمنكار.

Hand vs Machine vs Graphics

Neighboring patches likely to have the same class label

IES AROUND THE WORLD AND WAS BROUGHT AGAINST THI WAS ALSO CONFIRMED THAT AND INDEED, JUDGE BOGAE LES STATED THAT IF BARCL Since 1958, 13 Calour life Feers and Deresses have been created. Most Caloursendiment would # still favour the abolition of the House of Cords, but while it remains Tabeur has to have an adequate number of

 $\prod_{i \in i} \psi_h(x_i, x_j) \prod_i \psi_o(x_i, y_i)$ (i, j)

 ψ_h : Nbd similarity; ψ_o : Observ dependency

MRF

*y*₁ **y**₆ $\Psi_o(x_i, y_i)$ x_1 x_6 **y**₂ y y_j **y**₅ m_{il} т_{6j} *m* _{5j} m_{i2} m_{ij} x_2 X_{i} X_i x_5 **y**₄ **y**₃ m_{i3} $m_{4j} \quad \Psi_h(x_i, x_j)$ x_{4} X_{2}

$$\psi_o(x_i, y_i) = e^{1/(\lambda D_m(x_i, y_i))}$$

D_m: Distance in Feature Space

$$\psi_h(x_i, x_j) = 1 + \alpha e^{-D_n(i,j)} + \beta e^{-D_e(x_i, x_j)}$$

Distance in Spatial Space (D_n) and Feature Space (D_e :)

Flipped Class

Summary

Handwriting... that action of emotion, of thought, and of decision that has recorded the history of mankind, revealed the genius of invention, and disclosed the inmost depths of the soulful heart. It gives ideas tangible form through written letters, pictographs, symbols and signs Handwriting forms abond across millennia and generations that not only ties us to the thoughts and deeds of our forebears, but also serves as an irrevocable link to our humanity. Neither machines nor technology can replace the contribution or continuing importance of this inexpensive portable skill. Necessary in every age, handwriting remains just as vital to the enduring saga of civilization Michael R. Sull as our next breath.

Thank You

vena Ocabs, baffalo, eda