Making <mark>possible</mark> things <mark>easy</mark>

The \$1 unistroke recognizer (UIST 2007)

UIST 2024 Lasting Impact Award

Jacob O. Wobbrock, Andrew D. Wilson, Yang Li



slow







66 UIST is about making impossible things possible,

and about making possible things easy...

in service of **people**.



– Scott Hudson



VizWiz, Bigham et al. (2010)



OmniTouch, Harrison et al. (2011)



KinectFusion, Izadi et al. (2011)

Automatic clustering generally helps separate different kinds of records that need to be edited differently, but i sin perfect. Sometimes if creates more clusters than needed, because the differences in structure aren't inportant to the user's particular editing task. For example, if the user only needs to edit near the end of each ine, then differences at the start of the line are largely irrelevant, and it isn't necessary to spill based on those differences. Conversely, sometimes the clustering isn't me enough, leaving heterogeneous clusters that must be edited one line at a time. One solution to this problem would be to let the user rearrange the clustering manually, perhaps using drag-and-drop to merge and spill clusters. Clustering and selection generalization would also be improved by recognizing common text structure like URLs, filenames, email addresses, dates, times, etc.

Advanced calebrary generally high separate affiltered holds of concerts that needs to extend disferently, it als only period. Sometimes in structure and microsoft to transmission of dimensions in structure and microsoft to the soft particular diding dimension and dimensions at the start of the line are largely interventia, and intra moreoversity of based on those differences. Conversity, and microsoft based on the soft the line are largely intervential, and that moreoversity of based on those differences. Conversity, calebra to must be outified on the user managing the calebrary amount of the soft the soft of the line are largely intervent. And calebra model be to the user managing the calebrary amount using algorid city walks. Calebrary and selection generalization to this. Iteramise main addresses, data the time the terms that the times. The managine maintervent the term and addresses that the terms in the calebrary and the calebrary and the calebrary and the calebrary based on the term term and the calebrary and the calebrary and the calebrary terms and the term term and the calebrary and the calebrary and the calebrary terms and the calebrary terms and the term terms and the calebrary and the calebrary terms and the terms and the term terms and the Automatic clustering generally helps separate different kinds of records that need to be dield differently, but it isn't perfect. Sometimes A treaters more clusters than needed, because the differences in structure aren't relevant to a specific task. | Corversely, sometimes the clustering isn't fine enough, leaving heterogeneous clusters that must be edited one line enough, leaving heterogeneous clusters that must be edited one line at time. One solution to this problem would be to let the user rearrange the clustering manually using diag-and-drop edits. Clustering and selection generalization would also be improved by recognizing common text structure like URLs, filenames, email addresses, dates, times, etc.

Automatic clustering generally helps separate different kinds of records har need to be odded differently, but it and perform features in studies and the set of th





SearchTogether, Morris & Horvitz (2007)

Soylent, Bernstein et al. (2010)

Interactive Beautification, Igarashi et al. (2007)



How can we make 2-D gesture recognition **easy**?







A variety of gesture-based applications have been created. Coleman implemented a text editor based on proofreader's marks [3]. Minsky built a gestural interface to the LOGO programming language [13]. A group at IBM constructed a spreadsheet application that combines gesture and

handwriting [18]. Buxton's group produced a musical score Permission to copy without fee all or part of this material is granted remnastin to copy without the art on part of this material is get provided that the copies are not made or distributed for direct provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the commercial advantage, the ALM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy

permission of the Association for Computing machinery. 10 Yo otherwise, or to republish, requires a fee and/or specific permiusing GRANDMA, is used as an example. First GDP's @1991 ACM-0-89791-436-8/91/007/0329 \$00.75 Claar ACM-0-89791-436-8-91/007/0329 \$00.75 or to republish, requires a fee and/or specific permission sistion of the Association for Computing Machinery. To copy use, or to resultivible measures a fee and/or specific memorylation. ublication and its date appear, and notice is griven usa-unitation of the Association for Computing Machinery

rapidly trained from a small number of examples of each gesture. Some gestures may vary in size and/or orientation while others depend on size and/or orientation for discrimination. Dynamic attributes (left-to-right or right-to-left, fast or slow) may be considered in classification. The gestural attributes used for classification are generally meaningful. and may be used as parameters to application routines. The remainder of the paper describes various facets of GRANDMA. GDP, a gesture-based drawing program built

built a trainable recognizer specialized toward gestures, as The recognition technology described here produces a small, fast, and accurate recognizers. Each recognizer is

329

 (x_{max}, y_{max}) (x_{P-1}, y_{P-1}) fs. (x_2, y_2) a (x_0, y_0) (x_{p+1}, y_{p+1}) (x_p, y_p) (x_{p-1}, y_{p-1}) (x_{min}, y_{min}) $f_1 = \cos \alpha = (x_2 - x_0) / \sqrt{(x_2 - x_0)^2 + (y_2 - y_0)^2}$ $f_2 = \sin \alpha = (y_2 - y_0) / \sqrt{(x_2 - x_0)^2 + (y_2 - y_0)^2}$ $f_3 = \sqrt{(x_{max} - x_{min})^2 + (y_{max} - y_{min})^2}$ $f_4 = \arctan \frac{y_{max} - y_{min}}{x_{max} - x_{min}}$ $f_5 = \sqrt{(x_{P-1} - x_0)^2 + (y_{P-1} - y_0)^2}$ $f_6 = \cos \beta = (x_{P-1} - x_0)/f_5$ $f_7 = \sin \beta = (y_{P-1} - y_0)/f_5$ Let $\Delta x_p = x_{p+1} - x_p$ $\Delta y_p = y_{p+1} - y_p$ $f_8 = \sum_{n=0}^{P-2} \sqrt{\Delta r_p^2 + \Delta y_p^2}$ Let $\theta_p = \arctan \frac{\Delta x_p \Delta y_{p-1} - \Delta x_{p-1} \Delta y_p}{\Delta x_p \Delta x_{p-1} + \Delta y_p \Delta y_{p-1}}$ $f_9 = \sum_{p=1}^{P-2} \theta_p$ $f_{10} = \sum_{p=1}^{P-2} |\theta_p|$ $f_{11} = \sum_{p=1}^{P-2} \theta_p^2$ Let $\Delta t_n = t_{n+1} - t_n$ $f_{12} = \max_{p=0}^{P-2} \frac{\Delta x_p^2 + \Delta y_p^2}{\Delta t^2}$

 $f_{13} = t_{P-1} - t_0$

Figure 6: Features used to identify strokes

Tappert, IBM 1982

C. C. Tappert

Cursive Script Recognition by Elastic Matching

Ţ

Dynamic programming has been found useful for performing nonlinear time warping for matching patterns in automatic encach economistant there this technique is applied to the peoplant of recognition during series. The partners in automatic Dynamic programming has been found useful for performing nonlinear time warping for matching patterns in automatic speech recognition. Here, this technique is applied to the problem of recognizing cursive script. The parameters used in the warshing are deviaed from time requester of was considered data of words handwritten on an electronic tablet. Chosen for their speech recognition. Here, this technique is applied to the problem of recognizing cursive script. The parameters used in the matching are derived from time sequences of x-y coordinate data of words handwritten on an electronic tablet. Chosen for their matching of the unitime these parameters are found particularly which here to the second particularly which here to the se matching are derived from time sequences of x-y coordinate data of words handwritten on an electronic tablet. Chosen for their properties of invariance with respect to size and translation of the writing, these parameters are found particularly suitable for the adaptive and the adaptive properties is the establichment in a testiming stradius of the resonantion system is the establishment in a testiming stradius of the resonantion system is the establishment in a testiming stradius of the resonantion system is the establishment in a testiming stradius of the resonantion system. properties of invariance with respect to size and translation of the writing, these parameters are found particularly suitable for the elastic matching technique. A salient feature of the recognition system is the establishment, in a training procedure, of protosumes hu each uniter using the system. In this manner, the system is tailored to the trees. Because in a proceeding of the etastic matching technique. A satient Jeature of the recognition system is the establishment, in a training procedure, of prototypes by each writer using the system. In this manner, the system is tailored to the user, Processing is performed on a word-knowed basic office the writing is encroted into word. Using any other for each latter the matching modeling allows and the system of the transfer of the system of th prototypes by each writer using the system. In this manner, the system is failored to the user. Processing is performed on a word-by-word basis after the writing is separated into words. Using prototypes for each letter, the matching procedure allows and basis word-by-word basis after the writing is separated into words. Using prototypes for each letter, the matching procedure allows any letter to follow any letter and finds the letter sequence which best fits the unknown word. A major advantage of this and the second secon any letter to follow any letter and finds the letter sequence which best fits the unknown word. A major advantage of this procedure is that it combines letter segmentation and recognition in one operation by, in essence, evaluating recognition at all processia commentations, thus avoid in the vessal commentations that second the method segmentation with a second term of the second term of terms o procedure is that it combines letter segmentation and recognition in one operation by, in essence, evaluating recognition at all possible segmentations, thus avoiding the usual segmentation-then-recognition philosophy. Results on cursive writing are usual segmentation before to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition converses is over 05 means that are to the investore letter recognition conver possible segmentations, thus avoiding the usual segmentation-then-recognition philosophy. Results on cursive writing are presented where the alphabet is restricted to the lower-case letters. Letter recognition accuracy is over 95 percent for each of the second se

1. Introduction

The advent in the early 1960s of electronic tablets capable of accurately capturing the x-y coordinate data of pen movement precipitated activity on cursive writing recognition [1, 2]. Except for an occasional thesis, activity in this area has decreased in recent years. Almost all efforts have been restricted to noncapital roman script [1, 2]. Two major approaches have been described in the literature. The approach undertaken at MIT by Eden and his students was basically one of "analysis by synthesis," where a model is created for the handwriting process and decoding is performed by fitting the model's production parameters [3]. Several more direct approaches were used by Harmon and his colleagues at Bell Laboratories, the most successful being an analysis on a letter-by-letter basis following explicit segmentation. Letter accuracy for this technique was first reported at 60 percent [4] and later, after improvements, at 90 percent [2] for carefully formed script.

However, with the increased interest in office systems, particularly those that reduce the principal's dependence on secretarial support, there is also a growing interest in communicating with machines in a person's natural modalities,

such as speech and handwriting. For direct written input by principals, perhaps the most difficult technical problem is

The dynamic programming technique of elastic matching (dynamic time warping) was applied to speech recognition problems over a decade ago and has since become widespread [5-7]. Recently, elastic matching has been successfully applied to the recognition of discrete handwritten characters, where an input character is matched against each of a set of prototypes and assigned the name corresponding to the prototype yielding the best match [8]. Here, elastic matching is extended to the recognition of cursive writing. The procedure decodes handwritten words into estimated strings of letters. Operating on a word at a time, using letter prototypes, and allowing any letter to follow any letter, the decoder uses elastic matching to find the prototype sequence which best fits the unknown word, yielding the corresponding letter sequence as the estimated letter string. Elastic matching provides the decoder with the essential feature of being Copyright 1982 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission to republish any other portion of this paper must be obtained from the Editor.











Figure 4 Elastic matching



Figure 5 Lattice of simple, two-prototype model.





Dynamic Time Warping Matching



UIST Lasting Impact Award 2014





4. Compare to templates





Tappert (1982) with "zero look-ahead"

Figure 4 Elastic matching.



specific permission and/or a ree. CHI 2010, April 10-15, 2010, Atlanta, Georga, USA. Copyright 2010 ACM 978-1-60558-929-9/1004... \$10.00.

CHI 2010: Everyday Gestures

Protractor is a novel gesture recognizer that can be easily

Protractor outperforms its peers in many aspects.

based approach, nearest neighbor approach.

ACM Classification Keywords

Gesture-based interaction, gesture recognition, template-

H5.2. [Information interfaces and presentation]: User interfaces, I5.2. [Pattern recognition]: Design methodology

An important topic in gesture-based interaction is

recognizing gestures, i.e., 2D trajectories drawn by users

with their finger on a touch screen or with a pen, so that a

computer system can act based on recognition results. Although many sophisticated gesture recognition algorithms (e.g. [2]) have been developed, simple,

argoniums (c.g., [4]) have been been been been advantages in template-based recognizers [4, 5] often show advantages in the second interaction of a start based interaction of a start based interaction.

personalized, gesture-based interaction, e.g., end users

defining their own gesture shortcuts for invoking

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granned without fee provided that copies are not made or distributed for profit or commercial advantage and that copies the state of the state of

ABSTRACT

Author Keywords

General Terms

INTRODUCTION

Algorithms, performance.

not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republic, to post on servers or to redistribute to lists, requires prior consideration and/or the or reproduing, to post on servers or to redistribute to lists, reo specific paramission and/or a fee. *CHI 2010*, April 10-15, 2010, Atlanta, Georgia, USA, Copyright 2010 ACM 978-1-60558-929-9/10.04...510.00. However, since a template-based recognizer needs to compare an unknown gesture with all of stored templates to make a prediction, it can be both time and space consuming, especially for mobile devices that have limited processing power and memory. In the remainder of this 2169

2169

For personalized, gesture-based interaction, it is hard to for personances, gesture an end user would specify and what the distribution of these gestures will look like. In addition, since an end user is often willing to provide only a small number of training samples, e.g., one sample per gesture category, it is hard to train a parametric recognizer that often has a high degree of freedom with such sparse training data. In contrast, template-based recognizers are

In contrast, recognizers that employ a parametric approach [3] often operate on a highly featurized representation of gestures and assume a parametric model that the target gestures have to fit. For example, the Rubine recognizer [2] extracts a set of geometric features from a gesture such as the size of its bounding box. It uses a linear discriminate any success the source of the second gestures to be linearly separable. These parametric recognizers can perform excellently when the target gestures truly fit the assumed model. However, if not, these

as templates, and at nutime, an unknown gesture is as temptates, and at tumming, an tumulown gesture is compared against these templates. The gesture category (or comparen against mese tempsates. The genue category to the label) with the most similar template is used as the result of recognition and the similarity implies how confident the prediction is. These template-based recognizers perform limited featurization, and a stored template often preserves the shape and sequence of a training gesture sample to a large degree. These recognizers are also purely data-driven, and they do not assume a distribution model that the target gestures have to fit. As a result, they can be easily customized for different domains or users, as long as training samples for the domain or user

Li, CHI 2010

April 10–15, 2010, Atlanta, GA, USA

implemented and quickly customized for different users. implemented and quickly customized for different users. Protractor uses a nearest neighbor approach, which commands. First I offer my insight into why template-based recognizes an unknown gesture based on its similarity to recognizers may be superior for this particular interaction each of the known gestures, e.g., training samples of context. I then focus on Protractor, a novel template-based examples given by a user. In particular, it employs a novel method to measure the similarity between gestures, by Template-based recognizers essentially use a nearest calculating a minimum angular distance between them with a closed form solution. As a result, Protractor is more neighbor approach [3], in which training samples are stored a crosed-torni solution. As a result, riou actor is more accurate, naturally covers more gesture variation, runs significantly faster and uses much less memory than its peers. This makes Protractor suitable for mobile computing, peers, this makes remarked summer to mouse companies, which is limited in processing power and memory. An evaluation on both a previously published gesture data set and a newly collected gesture data set indicates that

Protractor: A Fast and Accurate Gesture Recognizer

Google Research 1600 Amphitheatre Parkway Mountain View, CA 94043

yangli@acm.org

h $\theta_{optimal} = \arctan$

$$b = \sum_{i=1}^{n} \left(x_{ti} y_{gi} - y_{ti} x_{gi} \right)$$

$$S(t,g) = \frac{1}{\arccos \frac{v_t \cdot v_g}{|v_t| |v_g|}}$$

$$v_t \cdot v_g = \sum_{i=1} \left(x_{ti} x_{gi} + y_{ti} y_{gi} \right)$$

$$|v_t| |v_g| = \sqrt{\sum_{i=1}^n (x_{ti}^2 + y_{ti}^2)} \sqrt{\sum_{i=1}^n (x_{gi}^2 + y_{gi}^2)}$$





handwriting recognition: A comprehensive survey. IEEE Trans. Patiern Analysis & Machine Int. 22 (1), 63-84.	Step 3. Scale points so that the resulting bounding box will be of size ² dimension; then translate points to the origin. BOUNDENG-
 Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (1992) Numerical Recipes in C. Cambridge Univ. Press. Rubine D. (1901) Scatificity Gestures in supersymptotic Press 	BOX returns a rectangle according to (mm _x , mm _b), (max _y , max _y). For gestures serving as templates, Steps 1-3 should be carried out once on the raw input points For candidates. Steps 1-4 should be
SIGGRAPH '91. New York: ACM Press, 329-337.	used just after the candidate is articulated.
 Salvador, S. and Chan, P. (2004) FastDTW: Toward accurate dynamic time warping in linear time and space. 3rd Witchn on 	SCALE-TO-SQUARE(points, size) $1 B \leftarrow BOLDEDENG-BOX(points)$
Mining Temporal and Sequential Data, ACM KDD '04.	2 for each point p in points do
Seattle, Washington (August 22-25, 2004).	$\begin{array}{ccc} 3 & q_x \leftarrow p_x \times (size \mid B_{width}) \\ 4 & q_x \leftarrow p_x \times (size \mid B_{width}) \end{array}$
recognition. Proc. IUI '05. New York: ACM Press, 281-283.	5 APPEND(newPoints, q)
 Stojmenović, M., Nayak, A. and Zunic, J. (2006) Measuring linearity of a finite set of points. Proc. CIS '06. Los Alamitos 	6 return newPoints TRANSIATE-TO-ORIGN(noints)
CA: IEEE Press, 1-6.	$1 c \leftarrow \text{CENTROID}(points)$
 Swigart, S. (2005) Easily write custom gesture recognizers for your Tablet PC amplications. Tablet PC Tashwing! Atticles 	2 for each point p in points do 3 $a \leftarrow p = c$
28. Tappert, C.C. (1982) Cursive script recognition by elastic	$\begin{array}{ccc} & & & & \\ & & & \\ 4 & & & q_y \leftarrow p_y - c_y \end{array}$
matching. IBM J. of Research & Development 26 (6), 765-771.	5 APPEND(newPoints, q) 6 return newPoints
of the art in online handwriting recognition. <i>IEEE Trans.</i>	Step 4. Match points against a set of <i>templates</i> . The size variable
Pattern Analysis & Machine Int. 12 (8), 787-808.	on line 7 of RECOGNIZE refers to the size passed to SCALE-To-
 Verman, J.K. (1997) Log-unear Models for Event Histories. Thousand Oaks, CA: Sage Publications. 	SQUARE in Step 3. The symbol φ equals $\frac{1}{2}(-1 + \sqrt{5})$. We use $\theta = \pm 45^{\circ}$ and $\theta_{*} = 2^{\circ}$ on line 3 of RECOGNIZE Due to using
 Wilson, A.D. and Shafer, S. (2003) XWand: UI for intelligent masses. Proc. CHI '02, New York: ACM Proc. 545, 552. 	RESAMPLE, we can assume that A and B in PATH-DISTANCE
 Zhai, S. and Kristensson, P. (2003) Shorthand writing on stylus 	contain the same number of points, i.e., $ A = B $. RECOGNIZE(points, templates)
keyboard. Proc. CHI '03. New York: ACM Press, 97-104.	$1 b \leftarrow +\infty$
APPENDIX A – \$1 GESTURE RECOGNIZER	2 foreach template 1 in templates do 3 $d \leftarrow \text{DISTANCE-AT-BEST-ANGLE}(points, T, -\theta, \theta, \theta_{\lambda})$
Step 1. Resample a points pain into n evenity spaced points. RESAMPLE(points n)	4 if $d \le b$ then
1 $I \leftarrow \text{PATH-LENGTH}(points) / (n-1)$	$\begin{array}{ccc} 5 & b \leftarrow a \\ 6 & T' \leftarrow T \end{array}$
$\begin{array}{ccc} 2 & D \leftarrow 0 \\ 3 & newPoints \leftarrow points_n \end{array}$	7 $score \leftarrow 1 - b / 0.5 \sqrt{(size^2 + size^2)}$
4 for each point p_i for $i \ge 1$ in points do	8 return $\langle T, score \rangle$ DISTANCE-AT-BEST-ANGLE(points, $T, \theta_{\alpha}, \theta_{b}, \theta_{b}$)
5 $d \leftarrow \text{DISTANCE}(p_{i-1}, p_i)$ 6 if $(D + d) > I$ then	$1 x_1 \leftarrow \varphi \theta_a + (1 - \varphi) \theta_b$
7 $q_x \leftarrow p_{i-1_x} + ((I-D)/d) \times (p_{i_x} - p_{i-1_x})$	2 $f_1 \leftarrow \text{DISTANCE-AT-ANGLE}(points, T, x_1)$ 3 $x_2 \leftarrow (1 - \sigma)\theta_1 + \sigma\theta_1$
8 $q_y \leftarrow p_{i-1_y} + ((I-D)/d) \times (p_{i_y} - p_{i-1_y})$	4 $f_2 \leftarrow \text{DISTANCE-AT-ANGLE}(points, T, x_2)$
 APPEND(newPoints, q) INSERT(noints, i a) // a will be the next n. 	5 while $ \theta_b - \theta_a \ge \theta_{\Delta}$ do 6 if $f_1 \le f_2$ then
11 $D \leftarrow 0$	$7 \qquad \theta_b \leftarrow x_2$
12 else $D \leftarrow D + d$ 13 return new Points	8 $x_2 \leftarrow x_1$ 9 $f_2 \leftarrow f_1$
PATH-LENGTH(A)	10 $x_1 \leftarrow \varphi \theta_a + (1 - \varphi) \theta_b$
$\begin{array}{c}1 & a \leftarrow 0\\2 & \text{for } i \text{ from 1 to } A \text{ step 1 do}\end{array}$	11 $f_1 \leftarrow \text{DISTANCE-AT-ANGLE}(points, 1, x_1)$ 12 else
3 $d \leftarrow d + \text{DISTANCE}(A_{i-1}, A_i)$	13 $\theta_a \leftarrow x_1$
4 return d Step 2 Pateta points on that their indication and a is at 0°	14 $x_1 \leftarrow x_2$ 15 $f_1 \leftarrow f_2$
ROTATE-TO-ZERO(points)	16 $x_2 \leftarrow (1 - \varphi)\theta_a + \varphi \theta_b$
1 $c \leftarrow \text{CENTROID}(points)$ // computes (\tilde{x}, \tilde{y})	17 $f_2 \leftarrow \text{DISTANCE-AT-ANGLE}(points, T, x_2)$ 18 return MDN(f, f_2)
2 $\theta \leftarrow \text{ATAN}(c_y - points_{0_y}, c_x - points_{0_y}) // \text{ for } -\pi \le \theta \le \pi$ 2 $\mu = \mu \text{ Britter} + \mu \text{ Botterm Partment} + \theta$	DISTANCE-AT-ANGLE(points, T, θ)
4 return newPoints	1 newPoints \leftarrow ROTATE-BY(points, θ) 2 $d \leftarrow$ PATH-DISTANCE(newPoints, T_{redev})
ROTATE-BY(points, θ)	3 return d
2 foreach point p in points do	PATH-DISTANCE (A, B)
	$1 d \leftarrow 0$
$3 q_x \leftarrow (p_x - c_x) \cos \theta - (p_y - c_y) \sin \theta + c_x$	$\begin{array}{ccc} 1 & d \leftarrow 0 \\ 2 & \text{for } i \text{ from } 0 \text{ to } A \text{ step } 1 \text{ do} \\ \end{array}$
$\begin{array}{l} 5 q_x \leftarrow (p_x - c_y) \cos \theta - (p_y - c_y) \sin \theta + c_x \\ 4 q_y \leftarrow (p_x - c_y) \sin \theta + (p_y - c_y) \cos \theta + c_y \\ 5 \text{AppEnD}(newPoints, q) \end{array}$	1 $d \leftarrow 0$ 2 for <i>i</i> from 0 to $ A $ step 1 do 3 $d \leftarrow d + DISTANCE(A_i, B_i)$ 4 return $d / A $

\$1 Unistroke Recognizer

Jacob O. Wobbrock, University of Washington (contact) Andrew D. Wilson, Microsoft Research Yang Li, University of Washington[†]

[†]Currently at Google

Download

\$1 source code: JavaScript, C# Dynamic time warping: C# Rubine classifier: C# Pseudocode: <u>\$1</u>, <u>Protractor</u> Unistroke gesture logs: <u>XML</u> Paper: <u>PDF</u>

This software is distributed under the <u>New BSD License</u> agreement.

About

The **\$1 Unistroke Recognizer** is a 2-D single-stroke recognizer designed for rapid prototyping of gesture-based user interfaces. In machine learning terms, **\$1** is an instance-based nearest-neighbor classifier with a 2-D Euclidean distance function, i.e., a geometric template matcher. **\$1** is a significant extension of the proportional shape matching approach used in <u>SHARK²</u>, which itself is an adaptation of <u>Tappert's elastic</u> <u>matching</u> approach with zero look-ahead. Despite its simplicity, **\$1** requires very few templates to perform well and is only about 100 lines of code, making it easy to deploy. An optional enhancement called <u>Protractor</u> improves **\$1**'s speed.

Upon its publication at UIST 2007, \$1 was invited for a special reprise presentation at SIGGRAPH 2008.

The <u>SN Multistroke Recognizer</u> extends \$1 to gestures with multiple strokes. The <u>SP Point-Cloud Recognizer</u> performs unistroke and multistroke recognition without the combinatoric overhead of \$N, as it ignores stroke number, order, and direction. The <u>SQ Super-Quick Recognizer</u> extends \$P for use on low-powered mobiles and wearables, as it is a whopping 142× faster and slightly more accurate.

The \$-family recognizers have been built into numerous projects and even industry prototypes, and have had many follow-ons by others. <u>Read about the \$-family's impact.</u>

Demo

In the demo below, only one unistroke template is loaded for each of the 16 gesture types. You can add additional unistrokes as you wish, and even define your own custom unistrokes.



"v"

left square bracket right square bracket

Make strokes on this canvas. If a misrecognition occurs, add the misrecognized unistroke as an example of the intended gesture.

https://depts.washington.edu/acelab/proj/dollar/

delete





\$N

ABSTRACT

Vatavu et al. (ICMI 2012) \$P Gestures as Point Clouds: A \$P Recognizer for User Interface Prototypes Radu-Daniel Vatavu **University Ste** fan cel Mare of Lisa Anthony Suceava 72 Vatavu@ Jacob O. Wobbrock \$P+ Touch Interfaces Vatavu (CHI 2017) ABSTRACT Rapid prototyping platforms require CHI 2017, May 6–11, 2017, Denver, CO, USA ple, and accurate y family of recognize Improving Gesture Recognition Accuracy on current most advan Touch Screens for Users with Low Vision cant memory and e gesture representat member of the 8-fam sidering gestures as to \$1 on unistrokes Specifically, \$P deli Radu-Daniel Vatavu MintViz Lab MANSiD Research Center testing with 5+ trail above 99% for user-10 participants. We Suceava assist developers in a "cheat sheet" to aid Vata \$Q of the 8-family for Categories and Vatavu et al. (MobileHCI 2018) H.5.2 Information \$Q: A Super-Quick, Articulation-Invariant Stroke-Gesture ognition]: Design N Recognizer for Low-Resource Devices General Terms ure 1. Gesture articulations produced by people with low visites measured and an articulation of the second hich negatively affects rey affects recognizers' accuracy rates. Ten (10) super ients (a); three people with congenital nystagmus an MintViz Lab | MANSiD Center University Stefan cel Mare of Suceava tistrokes, Euclidean ABSTRACT ABSTHAC1 We contribute in this work on gesture recognition to impro Lisa Anthony Suceava 720229, Romania 1. INTRODUC we contribute in units work on gesture recognition to impro-the accessibility of touch screens for people with low visi-Department of CISE Jacob O. Wobbrock We examine the accuracy of popular recognizers for gestu We observe the accuracy of popular recognizers for gestu produced by people with and without visual impairments, a vatavu@eed.usv.ro University of Florida Information School | DUB Group The currently incr Gainesville, FL 32611, USA touch input devices Surface, along with University of Washington ABSTRACT produced of people with and without visual impainments, a we show that the user-independent accuracy of SP, the b lanthony@cise.ufl.edu We introduce SQ, a super-quick, arriculation-invariant pointwe show that the user-independent accuracy of sr, the to recognizer among those evaluated, is small for people we Seattle, WA 98195, USA ment, fosters a risin we introduce sQ, a super-quick, articulation-invariant point-cloud stroke gesture recognizer for mobile, wearable, and embedded devices with low computing resources. SQ ran up to 147 v faster than its produces of the new sectors showed weat recognizer among those evaluated, is small for people will low vision (83.8%), despite SP being very effective for g for such platforms. wobbrock@uw.edu ecognition tailored there is not a total of the sector of the se ennecuee uevices with row computing resources. My ran up to 142 x faster than its predecessor SP in our benchmark evaltures produced by people without visual impairments (20.5) By carefully analyzing the gesture articulations produced people with low vision, we inform key algorithmic revisio into existing softwa to 1442x nation into protocoson or in our treminants evaluations on several mobile CPUs, and executed in less than 3% uantons on serveral moone C.F.C.S. and executed in ress train 200 of SP's computations without any accuracy loss. In our most people with row vision, we inform key algorithmic revision for the SP recognizer, which we call SP+. We show significe or ar s computations writiont any accuracy loss, in our most extreme evaluation demanding over 99% user-independent to use ar recognizer, which we can see, we show significance accuracy improvements of \$P+ for gestures produced by p ple with low vision, from 83.8% to 94.7% on average and extreme evaluation demanding over 99% users independent recognition accuracy. SP required 9.4 s to run a single classifi-cation, action SQ completed in Just 191 ms (a 49× speed-up) on a Cortex-A7, one of the most widespread CPUs on the module modes. SC was seen faster on a forward for Attapie with tow vision, irom 83.8% to 94.7% on average and 7 to 98.2%, and $3 \times$ faster execution times compared to SP. 8000 8,362 bear this notice and the full republish, to post on serve permission and/or a fee. ICMI'12, October 22-26 Copyright 2010-ACM Classification Keywords on a Concessory, one or use most wratespread Cross on one mobile market, SQ was even faster on a low-end 600. MHz ACM Classification Keywords H.5.2. [Information Interfaces and Presentation (e.g., HC mobile market. SQ was even taster on a 10%-end 650-56112 processor, on which it executed in only 0,7% of SP's computa-D.S., University of the second strategies, K.4.2. [Construction] processor, on which is executed in only 0.7% of all a computed tions (a 142x speed-up), reducing classification time from two winners to lave then one record. On is the two main test for Copyright 2012 ACM 97 puters and Society Social Issues: Assistive technologies uons (a 144X speco-up), reducing classification une from two minutes to less than one second. SQ is the next major step for minutes to less than one second. SQ is the next major step for the "S-family" of gesture recognizers: articulation-invariant, extremely fast, accurate, and implementable on top of SP with inst 30 overa lines of code. Figure 1. Classification limits for 5P and 5Q (kft) on live (2P) architecture common facilitation limits for 5P and 5Q (kft) on live (2P) architecture for a set of a structure structure structure and the structure structure structure of a structure structure of a structure st Author Keywords Autor roywords Gesture recognition; Touch screens; Touch gestures; Visua Gesture recognition; Touch screens; Touch gestures; Visua Gessure recognition; iouch screens; ioucn gestures; visua impairments; Low vision; S1; SP; SP+; Recognition accura Evaluation; Point clouds; Algorithms; Recognition. ACM Classification Keywords Actin Classification Networks H.5.2. Information Interfaces and Presentation (e.g., HCI): devices are, new smaller devices emerge regularly, and con-tinued miniaturization seems inevitable for some time [14]. Amazingly, ARM hardler de Schillon chips to date (such as power 95% of today's smaller on Figure 1) that are used to and 35% of all electronic devices [4]. 11.5.2. Information interfaces and recommended User Interfaces – Input devices and strategies. Author Keywords cial advantage and that copies bear this opyrights for components of this wor Autura Nymoras Gesture recognition: stroke recognition; S1: SP; SQ; S-family; point-cloud recognizer; mobile devices; low-resource devices. May 06 11, 2017, Denver, CC . -opyright is held by the ownersation(13). Fundaction rights in ACM 978.1-4503-4655.9/17/054_515.00 DOI: http://dx.doi.org/10.1145/3925453.3825941 Input on such small devices has largely been restricted to one In recent years, we have seen the tremendous proliferation of small electronic devices of all kinds [40,41,52,54]. Smart, obvines are the most provalent of these but other smaller Input on such small devices has largely treen restricted to one or two buttons. Some devices, such as smartwatches, have or two buttons. Some devices, such as smartwatches, have touch-sensitive screens, but sophisticated input like text entry is still challenging due to a dearth of screen space. Research rhythms [52] to extend the range of input possibilities, but the convertunity remains to make interpreting with these during of small electronic devices of all kinds [40,41,52,54]. Milari-phones are the most prevalent of these, but other, smaller devices are now widespread, such as smartwatches, activity, monitoring writebands, such as smartwatches, activity, places, Rhashvolt devices are now widespread, such as smartwatches, activity-monitoring writsbands, agemented reality glasses, Biletowh earbudy, Bernory sticks, GPS trackers, nano projectors, and others. All these smart devices embed CPUs that need to ac-evane rode out-it is expectable for applications that records injournes to extend the fullge of input possionities, but the opportunity remains to make interaction with these devices opportunity remains to make interaction with incise devices better. The challenge is exacerbated by the fact that the smaller ouers. An unse smart uevices emoca CFUs that need to ex-ecute code quickly, especially for applications that process human input, such as gesture interfaces. As small as these bener, the channeling is exacerolated by the fact that the sinatu-est devices are often also "low-resource devices," meaning es devices are orient also tow-resolutive devices, incaming they lack the processing power and memory that larger devices have, new input schemes are needed that are suitable for view have revenues durings for more assumption than metals vices nave, new input sciences are needed that are suitance to 100 more source devices, but more expressive than merely a make digital or hard copies of all or part of this work, for provide or particle values for provide that copies are not reached or distributed or copies. Coppers of the copposate provide the transformation of the copposate structure between 4. For composate provide the transformation of the copposate structure between 4. So and the copposate provide the copposate provide the structure between 4. So and the copposate provide the copposate provide the structure between 4. So and the copposate provide the copposate provide the structure between 4. So and the copposate provide the copposate provide the structure of the copposate provide the copposate provide the copposate provide the structure of the copposate provide the copposate provide the copposate provide the copposate structure between 4. So and the copposate provide the copposate provide the copposate structure between 4. So and the copposate provide the copposate provide the copposate structure between 4. So and the copposate provide the co One type of input that has seen significant uptake in recent years is gesture recognition. Although gestures come in many years is gesture recognition. Attaining testime scome in many forms [30], stroke-gestures have gone mainstream, appear torms [30], stroke-gestures have gone mainstream, appear. ing in every touch screen-based product as swipes and flicks, atopy totala keyboards like Swype (som, swype, com) ShapeWriter [25,56], and in the "insert special characters" for-ture in Grounde Dever to some a four A domain on the Si MetalyHCI 'IA September 1.4 A September 2.4 A Disk for the sense of Suggestime (20,50), and in the insert special characters rea-ture in Google Docs, to name a few, A decade ago, the \$1 and the special sector and the special se 34.3229465 sed to ACM ture in Google Docs, to name a rew. A decade ago, ine 51 gesture recognizer [53] provided simple, easy-to-build stroke-



Extended family		Vanderdonckt et al. (ICMI 2018) IFTL IFTL, an Articular	Articulation, Rotation, Scaling, and Translation Articulation, Rotation, Scaling, and Translation Match (ARST) Multi-stroke Gesture Recognize Match Maggor (UCO), Université catholique de Louvain, LouditM. Région Leave Vander Douckt, Université catholique de Louvain, LouditM. & (Ertens, Relevance Review Interruption Interruption Interruption Interruption
<section-header><section-header><section-header><section-header><section-header><section-header><section-header><section-header><section-header><section-header><section-header><section-header><section-header><section-header><section-header><section-header><section-header><section-header><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></section-header></section-header></section-header></section-header></section-header></section-header></section-header></section-header></section-header></section-header></section-header></section-header></section-header></section-header></section-header></section-header></section-header></section-header>	<section-header><section-header><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text><text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></text></section-header></section-header>	<text><text><text></text></text></text>	<section-header><section-header><section-header><text><text><text><text><text><text><text><text><text><text><text></text></text></text></text></text></text></text></text></text></text></text></section-header></section-header></section-header>

Ţ

Other people's projects...





Eden by Kin, Miller, Bollensdorff, DeRose, Hartmann & Agrawala (CHI 2011)



after touch: finger's height controls scrolling

Air+Touch by Chen, Schwarz, Harrison, Mankoff & Hudson (UIST 2014)

F



Expanding the input expressivity of smartwatches... by Xiao, Laput & Harrison (CHI 2014)



Datection Original **Application Output** C monicha vision andTrocker: Hand: right, Finger: index, Form: closed, Code: Unknown AndTrocker: Finished detection at 1443061333924 KeyboordEventListener: Key Code = 106 rameCache: Found frame with difference in timestamp of 9 romeCoche: Index of the frame = 78 149 149 63 ["descriptor":{"code":"Unknown","finger":"index","form":"open","hand":"right"},"event":{" ["alternate":false, "command":false, "control":false, "shift":false}, "timestamp": "1443001334 WandTrucker: Hand: right, Finger: index, Form: open, Code: Unknown andTrocker: Finished detection of 1443061334672

name from the Designation of the Completion of the All Completion of the State

Finger-aware shortcuts by Zheng & Vogel (CHI 2016)

Inkbase Programmable Ink

James Lindenbaum

Szymon Kaliski

Joshua Horowitz

November 2022



Sketchy math example using the <u>\$1Unistroke</u> <u>Recognizer</u> to identify a triangle.

https://www.inkandswitch.com/inkbase/



Land lines by Lieberman & Felsen (2016)

https://lines.chromeexperiments.com/

Takeaways

- UIST is definitely <u>the</u> place for **novel** UI inventions!
- But it also has a long history of contributions that make known things easier to do (e.g., UI toolkits)
- For maximum impact, we must continue to prioritize **both!**
- This means:
 - Make your contributions easy to take up and use by others
 - Enable others to replicate and test your innovations
 - "White box" your prototypes and algorithms
 - Think beyond your UIST publication!

(Yes, some of these are in tension with commercial priorities)

Thank you!

Jacob O. Wobbrock

University of Washington <u>wobbrock@uw.edu</u>

Andrew D. Wilson Microsoft Research awilson@microsoft.com

Yang Li

University of Washington (Now at Google DeepMind) <u>liyang@google.com</u>







Research

https://depts.washington.edu/acelab/proj/dollar/

Result: Thank you! (0.81) in 1 ms.