Automating Tactile Graphics Translation Computer Vision CSE 455 2009

> Richard Ladner University of Washington





Kent Cullers, Ph.D. Physics



Cary Supalo Grad Student Chemistry



Geerat Vermeij, Ph.D. Evolutionary Biologist





Bill Gerrey Electrical Engineering Inventor



Imke Durre, Ph.D. Atmospheric Science



William Skawinski Professor, Chemistry





H. David Wohlers Professor, Chemistry



TV Raman Computer Science Google



Victor Wong EE Grad Student





Chieko Asakawa Computer Scientist IBM



Hideji Nagaoka Computer Scientist Tsukuba U. of Tech



Katsuhito Yamaguchi Physics Nihon University



UW Students



Zach Lattin Math Major

Sangyun Hahn Ph.D. Student CSE





Outline

- Tactual Perception
- Text
- Math
- Graphics
- Problems
- Thanks
- Demo



Tactile Perception

- Resolution of human fingertip: 25 dpi
- Tactual field of perception is no bigger than the size of the fingertips of two hands
- Color information is replaced by texture information
- Visual bandwidth is 1,000,000 bits per second, tactile is 100 bits per second



Braille

 System to read text by feeling raised dots on paper (or on electronic displays). Invented in 1820s by Louis Braille, a French blind man.



Tiger Embosser

- 20 dpi (raised dots per inch)
- 7 height levels (only 3 or 4 are distinguishable)
- Prints Braille text and graphics
- Prints dot patterns for texture
- Invented by a blind man, John Gardner





Outline

- Tactual Perception
- Text
- Math
- Graphics
- Problems
- Thanks
- Demo



Let's use of the less	e this procedure to s on.	olve the applicatio	on presented at the beginning		
Define variables.	Let $x =$ the null Let $y =$ the null	Let $x =$ the number of acres of crop A . Let $y =$ the number of acres of crop B .			
Write inequalities	$x \ge 0, y \ge 0$ $x \le 15$ $x + y \le 20$	Acreage cannot No more than 1 No more than 20	be less than 0. 5 acres of crop A are permitted. 1 acres can be planted in all.		
Graph the system.	20 15- 10- 5- (15, 0- 5- 10- 15- (15, 10- 15- 10- 15- 10- 15- 10- 10- 15- 10- 10- 10- 10- 10- 10- 10- 10	x = 15 (15, 5) (15, 5) (15, 5) (15, 5)	The constraints $x \ge 0$ and $y \ge 0$ tell you to consider only those points that are in Quadrant I.		
	The vertices a	re at (0, 0), (15, 0)	, (15, 5), and (0, 20).		
Write an expression.	Profit equals i equals 600x – B equals 520y function is P(x	Profit equals income less costs. The profit from crop <i>A</i> equals $600x - 120x - 15(5.60)x$, or $396x$. The profit from crop <i>B</i> equals $520y - 200y - 10(5.00)y$, or $270y$. Thus, the profit function is $P(x, y) = 396x + 270y$. P(0,0) = 396(0) + 270(0) = 0 P(15,0) = 396(15) + 270(0) = 5940 P(15,5) = 396(15) + 270(5) = 7290 P(0,20) = 396(0) + 270(20) = 5400			
Substitute values.	P(0,0) = 396(0) P(15,0) = 396 P(15,5) = 396 P(0,20) = 396				
Answer the problem	The maximum plant 15 acres maximum pro	n occurs at (15, 5). of crop A and 5 a fit of \$7290.	Thus, Mr. Washington should acres of crop B to obtain the		
In certai programm at the righ	n circumstances, th ing is not helpful. C t, based on the follo	ne use of linear Consider the graph owing constraints.	y 4 7- 6- 5- 5-		
$ \begin{array}{l} x \geq 0 \\ y \geq 0 \\ y \geq 6 \\ 4x + 3y \leq \end{array} $	12		3 2- 4x+3y≤12 1- +		



Text Translation

Text Image

The constraints do not define a region with any points in common in Quadrant I. When the constraints of a linear programming problem cannot be satisfied simultaneously, then infeasibility is said to occur. This may mean that the constraints have been formulated incorrectly, certain requirements need to be changed, or that additional resources are required before the problem can be solved.

Optical Character Recognition (OCR)

Text

The constraints do not define a region with any points in common in Quadrant I. When the constraints of a linear programming problem cannot be satisfied simultaneously, then infeasibility is said to occur. This may mean that the constraints have been formulated incorrectly, certain requirements need to be changed, or that additional resources are required before the problem can be solved.

Braille Translation (Duxbury) Speech Synthesis (Jaws)

Braille

la electrones e electron o energe a clear subject o local de les estas local local electrones actendes estas estas colo estas en electron



14

Speech

Outline

- Tactual Perception
- Text
- Math
- Graphics
- Problems
- Thanks
- Demo



Math



The constraints do not define a region with any points in common in Quadrant I. When the constraints of a linear programming problem cannot be satisfied simultaneously, then **infeasibility** is said to occur. This may mean that the constraints have been formulated incorrectly, certain requirements need to be changed, or that additional resources are required before the problem can be solved.



Math Translation



Math Translation Examples





Outline

- Tactual Perception
- Text
- Math
- Graphics
- Problems
- Thanks
- Demo







Graphic Translation



Graphic Translation



Finding Text

- Why not just use standard optical character recognition (OCR)?
 - OCR is not effective for graphical images.





© 2003 Elsevier Science (USA). All rights reserved.



ScanSoft OmniPage Pro 14.0

Find Text Letters

- Uses the following principles
 - Text in an image is usually in one font
 - Fonts are designed to have a uniform density at a distance.
 - In the absence of noise an individual letter tends to be connected component of one color. Exceptions are i and j.
- Use machine learning to determine which connected components are letters.



Features

Century Gothic

W = width of bounding box H = height of bounding box A = area of bounding box R_i = i-th radial slice density W

Η



R_i = number of black pixels in i-th slice where a slice is an angle of 360/n. The total number of slices is n.



Center is center of mass of black pixels

Machine Learning

- Training:
 - Sample the connected components and compute their features.
 - Use these features to train a Support Vector Machine (SVM).
- Finding:
 - For a new connected component compute its features.

- Feed these features into the SVM.



Example



© 2003 Elsevier Science (USA). All rights reserved.

Trained on a different images from the same book. About 200 letters in the training set.

Find Text Blocks







29

Group characters logically

- Extracting a set of isolated characters from an image is insufficient
 - Need groups of Braille characters for easier placement
- Challenges
 - Text can be at many angles
 - Individual characters may be aligned along multiple axes



Our approach

- Step 1: User provides training set
 - Software examines defining features
- Step 2: Automatically find similar groups in remaining images
 - A. Minimum spanning tree
 - B. Discard useless edges
 - C. Discard inconsistent edges
 - D. Create merged groups



Minimum spanning tree (1)

Treat the centroid of each connected component as a node



Discard useless edges (2)







33

Discard inconsistent edges (3)





Final merge step (4)

Merge only if the resultant group is consistent





OCR on Text Image

Image of text boxes		Text
14.0 12.0 10.0 8.0 6.0 4.0 2.0 0 Performance relative to AMD Elan SC520 Automotive Office Telecomm © 2003 Elsevier Science (USA). All rights reserved. AMD ElanSC520 AMD K6-2E+ IBM PowerPC 750CX NEC VR 5432 NEC VR 4122	OCR ,	14.0 12.0 10.0 8.0 6.0 4.0 2.0 0 Performance relative to AMD Elan SC520 Automotive Office Telecomm © 2003 Elsevier Science (USA). All rights reserved. AMD ElanSC520 AMD K6-2E+ IBM PowerPC 750CX NEC VR 5432 NEC VR 4122



Braille Placement

- Text boxes of Braille will be of different size than the original text boxes
 - Mode characters
 - Contractions
 - Braille is fixed width





Right justified





















Subtask Pattern



- TGA batch process
- Photoshop and Illustrator scripts
- Omnipage batch manager
- Duxbury command line



Tactile Graphics Assistant





Available Books

• Computer Architecture: A Quantitative Approach, 3rd Edition

Hennessy and Patterson

2002 Elsevier

25 minutes per figure

• Advanced Mathematical Concepts, Precalculus with Applications

Gordon-Holliday, et al.

1999 Glencoe/McGraw-Hill

6.3 minutes per figure

An Intoduction to Modern Astrophysics

Carroll and Ostlie

1996 Addison-Wesley

10.2 minutes per figure

Discrete Mathematical Structures

Kolman, Busby and Ross

2003 Prentice Hall

8.8 minutes per figure



Time per Figure

	Discrete N	Math	Precalcul	us	Astronomy	,
	Min		Min		Min	
SetUp	425	10.3%	660	9.8%	1110	18.3%
Classification	245	5.9%	390	5.8%	270	4.4%
TGA	595	14.4%	570	8.4%	585	9.6%
Omnipage	714	17.3%	660	9.8%	945	15.6%
Photoshop	800	19.4%	975	14.4%	660	10.9%
Duxbury	225	5.5%	630	9.3%	450	7.4%
Illustrator	770	18.7%	1335	19.7%	1845	30.4%
Workflow	350	8.5%	1545	22.8%	210	3.5%
Total	4124	100.0%	6765	100.0%	6075	100.0%
	num figs	467	num figs	1080	num figs	598
	min/flg	8.8	min/fig	6.3	min/fig	10.2

Ave 7.9 min/figure



Work Balance





TGA Workflow

- Advantages
 - Much faster production
 - Batch processing instead of one figure at a time
 - Much tedious work is avoided
- Disadvantages
 - May be of lower quality than custom translation
 - A lot of technology needs to be mastered



One-offs vs. Mass Production



1916 Woods Dual Power



1906 Reo



Model T



Outline

- Text
- Math
- Graphics
- Workflow
- Problems
- Thanks
- Demo



Problem solving

- Each book present a set of unique problems.
- We consider a few today
 - Classification of figures
 - Legends and colors
 - Text at an angle
 - Math in figures
 - Grids





Legends and Colors

- Legends may have to be enlarged.
- Colors may have to be replaced with textures.





© 2003 Elsevier Science (USA). All rights reserved.

consideren onenen annen ellerter orei verlendere.

Angled Text

• Braille should be printed horizontally.





Math – Infty Reader



$$\frac{(y-k)^2}{a^2} - \frac{(x-h)^2}{b^2} = 1$$

$$y = k$$

$$y$$

$$x$$

$$(h, k)$$

$$O$$

$$x = h$$

Extracted Math Image



Grids

• Grids may not work well in tactile form.





спредне за конструктор сложение на секте



TGA Technology

- Tactile Graphic Assistant
 - C++
 - Machine Learning (Support Vector Machine)
 - Learns features of text from positive and negative examples.
 - Computational Geometry
 - Text justification
 - Free executable
 - Licensable source code



Technologies in the Future

Include Audio with Touchpads

Digital Pen and Paper





East



Outline

- Text
- Math
- Graphics
- Workflow
- Problems
- Thanks
- Demo



CSE Undergraduate Students



Terri Moore

Andy Jaya

Eileen Hash



Current Undergraduate Student



Josh Scotland



CSE Graduate Students



Sahngyun Hahn



Chandrika Jayant



Thanks To

- Dan Comden
- Sheryl Burgstahler
- Raj Rao
- Melody Ivory
- Ethan Katz-Basset
- Zach Lattin
- Stuart Olsen
- Many others



Thanks To

















