

Automating Tactile Graphics Translation

Richard E. Ladner, Melody Y. Ivory, Rajesh Rao, Sheryl Burgstahler, Dan Comden, Sangyun Hahn, Matthew Renzelmann, Satria Krisnandi, Mahalakshmi Ramasamy, Beverly Slabosky, Andrew Martin, Amelia Lacenski, Stuart Olsen, Dmitri Groce
University of Washington
Box 352350
Seattle, WA 98195-2350 USA
+1 206 543 9347
ladner@cs.washington.edu

ABSTRACT

Access to graphical images (bar charts, diagrams, line graphs, etc.) that are in a tactile form (representation through which content can be accessed by touch) is inadequate for students who are blind and take mathematics, science, and engineering courses. We describe our analysis of the current work practices of tactile graphics specialists who create tactile forms of graphical images. We propose automated means by which to improve the efficiency of current work practices. We describe the implementation of various components of this new automated process, which includes image classification, segmentation, simplification, and layout. We summarize our development of the tactile graphics assistant, which will enable tactile graphics specialists to be more efficient in creating tactile graphics both in batches and individually. We describe our unique team of researchers, practitioners, and student consultants who are blind, all of whom are needed to successfully develop this new way of translating tactile graphics.

Categories and Subject Descriptors

K.4.2 [Social Issues]: Assistive technologies for persons with disabilities

General Terms

Human Factors

Keywords

Tactile graphics, Braille, user study, image processing, machine learning, disability, accessibility

1 Introduction

Graphical images, such as bar charts, diagrams, and line graphs, are prevalent in mathematics, science, and engineering (MSE) textbooks. For instance, postsecondary com-

puter science textbooks may have images on half of the pages (sometimes up to 500 images; e.g., [10, 13]). Studies have shown that tactile graphics (representations through which the content can be accessed by touch) have the most value for scientific and technical users who are blind, and, furthermore, that tactile perception is the best modality for comprehension of graphical images [3, 9, 12]. Nonetheless, graphical images that are found in many textbooks are not typically available to students in a tactile format.

There are several obstacles to timely translation of images into a tactile format: labor-intensive and slow translation processes, inadequate human and technological resources, and variation in translation expertise [1, 5, 16]. As an example, due to the amount of translation work required, a graduate student (a co-author) had tactile access to only 40 of the 461 graphical images that were in a computer architecture textbook; consequently, the student's academic performance was well below his ability. Without innovations to address this problem over 93,000 visually impaired or blind students in the education system [2] will not have the same access to MSE course materials as their sighted peers.

We want to support the translation of graphical images that exist in print or digital form. Several resources provide some support to tactile graphics specialists who carry out such translation tasks. For example, they can use general image-editing applications, such as Microsoft Visio® and CorelDRAW® or specialized image-editing programs, such as BrlGraphEditor [3], TGD Pro® (Tactile Graphics Designer Pro) [6], and Tiger Designer [15], to create manually an image master (outline with the most important components of the original image) that is suitable for tactile rendering. Tiger Designer also enables the editing of images that exist in digital form (e.g., Illustrator® files) and integration of Braille text into an image. Although there has been an effort to automate the translation of natural images (i.e., pictures of individuals, buildings, scenes, etc.) [17, 18], there have been no such efforts for graphical images.

Our primary objective is to remove a major barrier to equitable access to MSE course materials, by mitigating obstacles to the translation of a textbook's (or other resource's) graphical images into tactile graphics. To accomplish this objective, we conducted an in-depth study on the current work practices of tactile graphics specialists; we describe the study and its implications in Section 2. We propose an automated means by which to improve the efficiency of current practices (Section 3) and describe the implementation

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are 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, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ASSETS'05, October 9–12, 2005, Baltimore, Maryland, USA.
Copyright 2005 ACM 1-59593-159-7/05/0010 ...\$5.00.

of various components of this new automated process (Section 4). In Section 4 we also summarize our development of a tool to support tactile graphics specialists.

A critical component to accomplishing our objective is the unique team of contributors—researchers, practitioners, and student consultants who are blind—that we assembled. We describe the team (Section 5) and conclude with a discussion of future work. Another critical component to accomplishing our objective is having cooperation from textbook authors or publishers to provide textbook images in a digital form. Typically, it is not possible to gain access to these images or there is a delay of a month or more. Mitigating this situation will require intervention at a policy level.

2 Current Work Practices

Before attempting to automate tactile graphics translation, we needed to understand how tactile graphics specialists complete their tasks, as well as what they need and want in a software tool to support their work practices. There are several published studies of transcribers (i.e., people who translate text into Braille); however, these studies examined mainly how many exist (e.g., [1, 16]) and their training [5]. We used these studies to develop domain knowledge and to inform a contextual inquiry (i.e., structured field study; [11]) within the tactile graphics translation community.

2.1 Study Methodology

We administered an on-line questionnaire to 51 people who had some role in tactile graphics translation and, subsequently, conducted in-depth observation sessions within the workplaces of 16 questionnaire respondents. The questionnaire consisted of 34 questions, which explored their training, work practices, hardware used, and software used; we also collected demographic data (age, computer experience, gender, race/ethnicity, and so on). Questions included 11 single-response, 17 multiple-response (i.e., multiple choice or matrix), and 6 freeform items. In total, the questionnaire had 185 response items and took an average of 30 minutes to complete. Examples of questions are listed below.

- Describe the training that you have received in Braille transcription.
- Do you transcribe graphics/illustrations to Braille?
- Do you use Adobe Photoshop for your Braille translation/transcription of graphics?
- Do you have a visual impairment?
- Does transcriber best describe your role?
- Is English your first language?

We recruited questionnaire participants through email discussion lists for the Braille community and via letters and phone calls to organizations that had people who create tactile graphics. From among the 51 questionnaire respondents, we identified participants for in-depth observation sessions. We conducted sixteen three-hour sessions with tactile graphics specialists in the Pacific Northwest, California, and Kentucky. Participants used a broad range of tactile graphics translation practices and worked in different settings (e.g., school system or transcription agency).

We videotaped participants as they completed transcription tasks within their workplaces; we asked them to talk aloud about their thought processes and actions as they

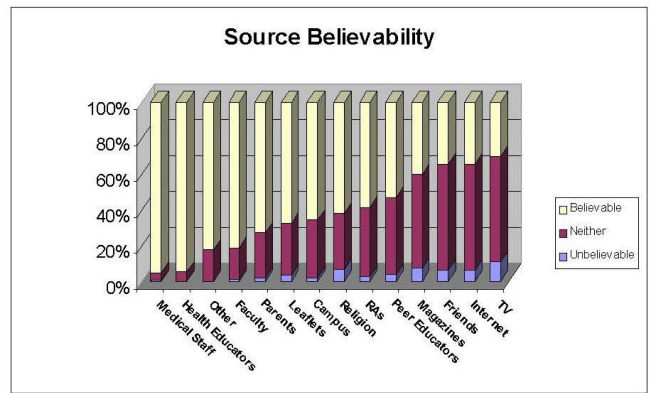


Figure 1: Example Image for Translation.

worked. They completed three control tasks: (1) translate or transcribe a one-page text document, (2) translate a diagram of the human eye into a tactile graphic, and (3) translate a bar chart (Figure 1) into a tactile graphic. We chose complex images to examine how participants currently handle challenging media and to have them to use as many resources (hardware and software) as possible. We supplied them with printed and digital (grayscale and full color) images. The images represented standard diagrams that are found in college-level MSE textbooks.

2.2 Tactile Graphics Specialists

The 51 responses to our on-line questionnaire suggest that tactile graphics specialists are a fairly homogenous group. They tended to be female (93 percent), and, perhaps due to the nature of their jobs, 96 percent had no physical or cognitive impairments. They tended to be Caucasian (80 percent) and in their late forties or early fifties (median age of 48). Some had completed college course work (31 percent) or were college graduates (21 percent). English was their first language (82 percent).

Tactile graphics specialists tended to be full-time employees, but not all of them named tactile graphics translation as their primary role. Most respondents (65 percent) characterized themselves as transcribers (i.e., a person who translates text into Braille) with no other roles. Other roles included: paraprofessionals (20%), teachers of the visually impaired (12%), tactile graphics specialists (12%), educator/trainer (10%), assistive technology specialist (8%), and administrator (2%). (Responses suggest little consistency in how people who translate tactile graphics describe their roles.) They had six or more years of transcribing experience, with 55 percent reporting that they had over 11 years of experience. They had at least 11 years of computer experience, but some respondents mentioned that their equipment was under-funded and had not been updated for many years.

Tactile graphics specialists' compensation is most comparable to transcribers, which was the most common role reported. Full-time, certified transcribers earn about \$18,000-\$50,000 a year, while non-certified professionals earn \$10,000-\$30,000 a year [1]. Respondents tended to not be certified, so their compensation was most likely in the lower salary range. If respondents were certified, their certification was typically through federal agencies (37%), state agencies (27%), and transcription organizations (17%). Most stated that they acquired tactile graphics translation skills from on-the-job training or were self-taught.

Tactile graphics specialists worked mainly within school systems (55 percent); a few also had personal businesses. They worked on text (92%), graphics (90%), and math (82%) translation tasks. They produced: textbooks (90%); graphs, charts, and other graphics (92%); literary works (78%); foreign language materials (57%); and children’s books (63%). They used a variety of methods to render tactile graphics: swellpaper (51%), embossing (40%), and thermoform (39%). Swellpaper is a special paper on which black patterns are created so that the patterns are transformed into raised areas when the paper is heated. Embossing is a process by which a special printer (embosser) punches dots into a sheet of paper to create areas of raised dots. Thermoform is a process by which a sheet of plastic is heated and vacuumed on top of a model of an image.

Ideally, image-editing software like Adobe Photoshop® enables a tactile graphics specialist to use the computer to create a master image (i.e., outline or trace) that they can then render via swellpaper, embossing, thermoform, or other methods. Respondents used various image-editing software applications, with CorelDRAW being the most frequently used (20 percent of respondents). Some respondents (39 percent) did not use any image-editing software. This finding prompted us to examine the software that they use to better understand their tradeoffs (Sections 2.4–2.5).

Respondents used multiple hardware solutions to render tactile graphics, including foil (20%), the Tactile Image Enhancer (16%), Versapoint Embosser (15%), Tiger Embosser (14%), and Juliet Embosser (12%). Foil entails using a stylus to draw an image into a sheet of film or aluminum. The Tactile Image Enhancer is used with paper that is similar to swellpaper. Embossers are special printers for producing dot patterns on paper.

2.3 Observations of Work Practices

We are still in the early stages of analyzing the sixteen observation sessions. Nonetheless, our preliminary analysis revealed the key subtasks that a tactile graphics specialist may complete during a translation task. We describe the ten subtasks below; Figure 2 depicts most of these subtasks.

- **Planning:** At the start of a translation task, participants usually spent time to determine the best method to use based on an image’s characteristics (e.g., the type of image or amount of text) and the intended user’s characteristics (e.g., tactile graphics experience or preferences). Planning also occurred throughout task completion and typically lasted a few minutes.
- **Image Drawing:** Some participants chose to create an image from scratch rather than use the supplied printouts or digital files. The objective was to produce an outline with the most important image elements. They used various methods (e.g., computer software or hand drawing tools) and typically spent the most time on this subtask.
- **Scanning:** Some participants scanned printouts, rather than use the digital files, because they did not receive them in their normal work practices. Scanning lasted a few minutes, but finding scanned files within the computer’s file system was a challenge.
- **Image Tracing:** Using a scanned image, printout, or digital file, some participants created an outline of the graphic by drawing on top of it (e.g., on a blank piece

of paper or a separate image layer within an image-editing application). Similarly to image drawing, they typically spent the most time on this subtask.

- **Image Texturing:** Once participants created an outline, they sometimes needed to add texture (e.g., a fill pattern or line pattern) to certain image areas like bars in a bar chart. Depending on the complexity of the image, texturing can take considerable time. Choosing distinctive textures and using them consistently was a challenge for some participants.
- **Braille Text Creation:** Although participants completed this subtask during the text translation task, they also created Braille text for images, mainly for legends (referred to as keys; see below). Subtask time varied based on the method used (computer software versus Braille typewriter) and the amount of text.
- **Key Creation:** Participants created keys to describe texture patterns and symbols that they used as labels for large blocks of text within an image (e.g., use of the letters ‘Med’ as a label for Medical Staff in Figure 1). They typically spent the most time on this subtask when they used symbols, because they had to determine which symbols to use, determine how to place the symbols within the image, and create a separate document that contained mappings between symbols and their meanings. Depending on the number of elements within the key, this subtask could take more time than image creation or image tracing. They typically created the key on paper first; thus, there were several points during which they made errors.
- **Image Labeling:** Participants needed to add Braille text to images, for instance to encode the key. They may also add lead lines from symbols to specific image areas. Depending on the number of elements, this subtask could take considerable time, but less time than image creation or tracing.
- **Proofing:** At various stages during task completion, participants evaluated their work. For instance, they may print and evaluate several versions of their image master. They typically did the most extensive proofing right before rendering.
- **Rendering:** Using a variety of methods, participants used the image master to create a tactile graphic.

Figure 2 depicts an observed translation task completion sequence, including the resources used and time spent during each step. The figure shows how one participant translated a bar chart (Figure 1) and created the Braille key to accompany it; she performed all subtasks, except for image drawing and scanning. The participant used an assortment of software and hardware and also worked in two different rooms. This behavior was typical during all sixteen observation sessions. In addition, there were embossers and hand tools to do things which computers could not; participants understood that one method could not perform all the required tasks. The figure also shows that translating tactile graphics is labor- and time-intensive. Errors can be made at any point in the process, which was the case in the example.

Automated text translation software such as Duxbury and hardware such as embossers saved participants a lot of time. Without software or hardware support, they would have had to do everything by hand. Current practices are much faster

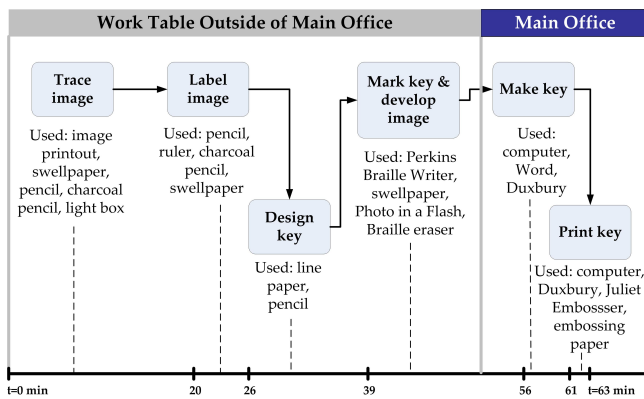


Figure 2: Example Translation Task Flow.

than they used to be; however, it appears that each product is designed to support specific tasks (Figure 2), without consideration of their overall work practices. These work practices need to be streamlined and supported holistically.

2.4 Feedback on Existing Image-editing Applications

We asked respondents about the image-editing software they use currently or attempted to use. Specifically, we asked them if there were ways in which the software could be improved to better support their work practices. We also asked them about additional software features that could be beneficial in the creation of images like charts, diagrams, and maps. There were a wide variety of issues that they mentioned. Comments related to the most frequently used software packages: Microsoft® Word® (9 comments), Adobe Illustrator (2 comments), CorelDRAW (3 comments), Adobe Photoshop (1 comment), and Tactile Graphics Designer Pro Pad (TGD Pro; 1 comment). The latter tool was made specifically for this community. We present some specific software issues below, based on questionnaire responses. We gained additional insight about the software during our observation sessions and subsequent evaluations (Section 2.5).

A surprise from the results of the questionnaire and observations was the prominent role of Microsoft Word in tactile graphics translation. We had not expected that Microsoft Word would be used as a tool for translating tactile graphics. Respondents seemed unwilling to use image-editing software; some tactile graphics specialists perceived full-featured image-editing software such as Photoshop to be too complex or lacking specific features for their needs.

Respondents provided general comments about how software could be improved, although their comments were not associated with a specific application. They wanted more templates and a repository of images from which they could work. They also wanted more texture fills, better translation of text within images, better image detection and auto color contrast, better support for 3D graphics translation, better compatibility across software applications, and more effective documentation and help aids.

2.5 Evaluations of Software Applications

Based on our study findings, we decided to conduct a software evaluation of our own. Commonly used image-editing applications were Adobe Photoshop, CorelDRAW, TGD Pro, and the drawing tools within Microsoft Word. To evaluate the suitability of these tools for a typical graphics translation

task (Figure 2) and tactile graphics specialist, we recreated the bar graph that we used in our observation sessions (Figure 1). We compared the user interfaces and task completion sequences within Adobe Photoshop CS, Adobe Illustrator CS, CorelDRAW 12, and Microsoft Word 2003.

We found that these applications are inadequate for tactile graphics specialists. They do not support adding high-level Braille text (e.g., Grade 2 or Nemeth) to graphics. Each program’s drawing tools require time and effort to learn, and some of their features are inappropriate for translating tactile graphics. As one example, CorelDRAW, Photoshop, and Illustrator only offer fill effects that are too detailed for embossing by a Braille printer. Word, on the other hand, offered a collection of simple geometric patterns that, while less appealing in illustrations for sighted people, would be appropriate for Braille printing. Word’s drawing environment, which specialists used the most, had the best image output and simplest interface, but had minimal export capabilities and suffered from the aforementioned deficiencies.

2.6 Study Implications

Questionnaire responses, observations, and software evaluations provided important information about the work practices and needs of tactile graphics specialists. For instance, we discovered ten key subtasks that a tool needs to support. Below, we summarize additional requirements of an automated tactile graphics translation solution.

- Support straightforward image editing, based on characteristics of the image and intended user. Tactile graphics specialists seem to need the simplicity of Microsoft Word, coupled with the powerful capabilities of full-featured image-editing programs like Photoshop.
- Support Braille text creation (e.g., making keys) and its integration into an image.
- Support multiple methods of image rendering (i.e., creation of a tactile graphic from an image master).
- Support image archiving and reuse.

In addition to these requirements, we created four personas (i.e., detailed descriptions or profiles) of tactile graphics specialists. They describe four types of people—a transcriber within a school system, a tactile graphics specialist within a Braille printing house, an assistive technology specialist within a community college, and a tactile graphics researcher—who may use the Tactile Graphics Assistant that we are developing. Personas are an important user interface design artifact, because they help us to consider the human side of software design. We are using these personas to inform the design of our tool.

3 Vision of Future Work Practices

Our main goal is to automate as much as possible the process of translating tactile graphics. We are interested more specifically in supporting the translation of existing graphical images that are found within textbooks and research papers in MSE disciplines. Our vision is that a tactile graphics specialist could use an automated system to translate an entire textbook’s images into a set of new images that have a simplified structure and contain Braille text instead of ASCII text. These images would be ready for any student who is blind to print-on-demand, either on swell paper, embossed, or thermoformed. We are currently examining embossing.

The system would not be fully automated; rather it would assist the specialist in translating many graphical images at a time. We also envision that the automated system would enable tactile graphics specialists to more quickly translate images into tactile graphics on a one-at-a-time basis than is currently done. Depending on its complexity, a single graphical image could take one to many hours to translate with current practices (see Figure 2). Our aim is to reduce translation time to 15 minutes or less for a single image. Although our current emphasis is on the translation of textbook images and embossing methods, we envision our work being extensible to other application domains such as the Web and rendering methods such as swell paper.

3.1 Challenges

From our observations, the more advanced the textbook, the more likely it is that its images are more abstract and less realistic. K–12 books tend to have numerous realistic images and photographs that are very difficult or impossible to translate. In advanced books, images tend to be abstract, using lines and geometric shapes, which make them easier to translate, except for the text within them. These abstract images tend to have much more text than the images that are found in K–12 books. Automatically finding the text and translating it into Braille is a major challenge in these more abstract images. In summary, automating the process of image translation is difficult at all education levels.

Many MSE textbooks contain hundreds of figures. Publishers take great care to ensure that textbook figures are consistent with each other in terms of overall style. Figures tend to use the same font style and size, line widths, shading, and other features. Figures also tend to come in classes, such as line graphs, bar charts, and diagrams; thus, figures in the same class can be translated in the same way. For example, bar charts in a book might have different colors on the bars. These colors could be transformed into textures (or fill patterns). Automating the translation of an entire book of figures is challenging, but it appears to be feasible if the figures are translated one class at a time.

4 Tactile Graphics Assistant

We learned from our observations of the current work practices of tactile graphics specialists that graphics are translated to tactile form one at a time and that the process is very time consuming. We are in the process of developing the Tactile Graphics Assistant (TGA) that will enable tactile graphics specialists to translate not only individual graphics, but entire textbooks of figures into tactile form.

Figure 3 depicts the overall structure of the TGA. The TGA must be able to perform many tasks including (1) digital image acquisition, (2) image classification, (3) segmentation of the image into text blocks and graphics only, (4) optical character recognition (OCR) and Braille translation, (5) image simplification, (6) image layout, and (7) image output into a tactile form. Once an image master is created, there is a need to evaluate it. The TGA would allow the user to make changes to improve the image master (and, consequently, the rendered tactile graphic) in ways that are similar to current work practices.

Generally, with the TGA, the specialist would not complete the image drawing, image tracing, Braille text creation, or image labeling subtasks as described in Section 2.3, at least not in the ways in which they are completed now. For

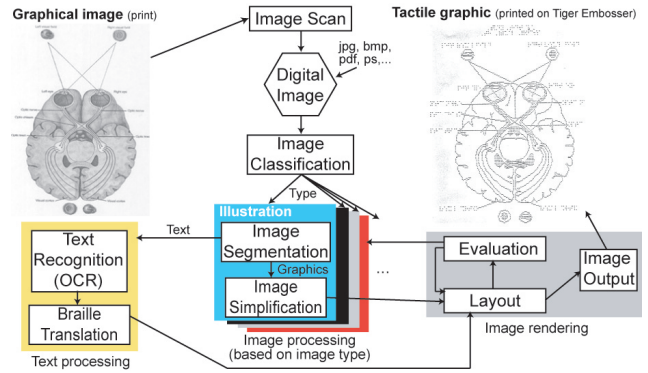


Figure 3: Structure of the TGA.

example, image tracing might be replaced with edge detection, which is the process of automatically tracing edges in an image. The TGA will introduce tactile graphics specialists to a new way of translating graphics into a tactile form.

The disadvantage of the TGA is that some tactile graphics specialists may not be able to adapt to new work practices. In particular, tactile graphics specialists who are uncomfortable with or do not have access to image-editing software may have difficulty. We will take steps to increase the number of tactile graphics specialists who can migrate successfully to new practices. More specifically, we are following a user-centered design process to develop our tool and will create learning resources. On the other hand, the advantage of the TGA is that tactile graphics can be translated much more quickly and in larger volumes (whole books at one time) than with current practices.

Currently, we are separately automating the TGA tasks. When we perfect the components, we will then integrate them into a single product or suite of products. A potential platform for the TGA is Adobe Photoshop, which already has many of the functions needed for image simplification (step 5 above). We can add our components as plug-ins to Photoshop or put them into an alternate job flow using either custom or standard applications like Microsoft Word. For example, the text extraction process might be done on a set of images that are loaded within a separate custom application. The images with the text removed would then be processed in batch to simplify them using a Photoshop script or plug-in. Regardless of the implementation environment, we aim to simplify interfaces and the completion of translation tasks to increase the likelihood that tactile graphics specialists can benefit from using the TGA.

4.1 Digital Image Acquisition

The original image may come in many different digital formats or be on paper. Paper images must be scanned into a digital format. All images need to be converted to color bit maps, which are rectangular arrays of pixels.

4.1.1 Scanned Images

Scanned images are relied upon by tactile graphics specialists, but they present many problems. A major problem is the background noise that is created often in the scanning process when a faint image of the reverse side of the scanned page is captured. Noise removal can be done using a threshold filter that whitens pixels that are nearly white. Another problem is that text often becomes blurry and multicolored. Segmenting the text is much harder than in the clean digital

image case. We are developing techniques to segment text from these kinds of images. In the following discussion, we describe our work with non-scanned digital images.

4.2 Image Classification

Textbooks often contain hundreds of figures. Usually they fall into a number of types, such as bar charts, line graphs, and diagrams. Different types of images might require different types of processing, as we discovered through our observation sessions. For example, the colors in a bar chart need to be converted to textures (e.g., fill patterns).

We developed image classification methods that are suitable for classifying textbook images. Using several image and text features, such as oriented Gaussian filters [8] and simple word and number counts, we used machine learning techniques [19] to classify the 244 figures found in Hennessey and Patterson's *Computer Architecture, A Quantitative Approach* [10]. The accuracy, precision, and recall were respectively: 0.93, 0.89, and 0.91 for bar charts; 0.99, 0.97, 1.00 for diagrams; and 0.94, .095, and 0.83 for line graphs. Accuracy is the fraction of all images in the test set that were classified correctly. Precision is the fraction of images classified to be in the class that they belonged. Recall is the fraction of images in the class that were classified correctly.

This study suggests that automated image classification is possible when the graphics come from a relatively uniform source, such as a single textbook. In a typical setting, the tactile graphics specialist would label some of the figures in a textbook. These figures would form a training set for the classification. The image and text features would be calculated for all the figures, with those of the training set used to train the classifier. Finally, the remaining figures are classified by the trained classifier.

4.3 Image Processing

Image processing consists of: (1) image segmentation where the text in the image is identified and replaced by the background color, and (2) image simplification where the remaining image is simplified to make it suitable for embossing.

4.3.1 Image Segmentation

In image segmentation we want to find each text block in an image and place a bounding rectangle around it. This task is complicated by the fact that text blocks are not always horizontal, but can be vertical (as in labeling the y-axis in a line graph), or at an arbitrary angle (as in labeling an edge in a block diagram). Our experience is that standard OCR (optical character recognition) programs do not do a satisfactory job in finding text within images.

We developed a text block finding algorithm that works for clean images where the letters (except for i and j, and some less common characters such as % and ©) are connected components of black pixels. The algorithm relies on the fact that fonts are designed to have a uniform density of black pixels per letter. Thus, a few statistical features of a connected component approximately determine whether it is a letter. We developed a technique to, first, find all the letters in the figures, then, second, to combine the letters into text blocks. The letter finding method uses training where the tactile graphics specialist identifies a number of individual letters in the one figure or several figures. After the training phase, all those connected components with similar statistical features are declared to be letters also. This

technique is highly accurate for the figures in Hennessey and Patterson's textbook.

Once the letters are identified, the next step is to merge them into text blocks that may be at any angle. To do this operation, the centroid of each letter is computed and the minimum spanning tree [4] connecting these centroids is computed. Letters in words tend to be connected by edges in the minimum spanning tree, but not vice versa. We have a number of criteria to remove edges from the minimum spanning tree that are not likely to be connecting letters in the same word. Remaining portions of the minimum spanning tree are merged to form blocks of text. Resulting text blocks often contain one or more words, and not whole lines of texts. A further merging procedure is used to combine text blocks into complete lines of text. For each text block, its location, angle, and dimensions are recorded. These text blocks can now be replaced by the background color used in the image, leaving just the graphic sans text behind.

4.3.2 Image Simplification

With the text removed, we now have a pure graphic to process. Depending on the image, some image simplification must be done. The eventual tactile graphic will consist mostly of lines and areas filled with texture. A graphic may have many colors, so color reduction can be done using color clustering methods [7]. The graphic may have areas that are shaded to form a 3-D effect. The shading can be removed using edge detection [7]. A small number of colors are often used to convey meaning. For example, a pie chart may use different colors for different parts of the pie. These colors can be replaced with textures (i.e., fill patterns). The graphic may have thin lines that need to be thickened using simple filtering. The vast majority of these tasks can be accomplished quickly using Photoshop. In addition, Photoshop supports batch processing, whereby a set of images can be processed in the same way using a predefined script.

4.4 Text Processing

After the text segmentation and removal, lines of text can be concatenated to form an image that consists only of text lines. Lines that are not horizontal are rotated to be horizontal. This image is then processed by the OCR engine, yielding excellent results compared with using OCR on the original image. For each text line, we captured the actual ASCII text, its location, and its angle. Now that we have this information, it can be translated into Braille using Duxbury, Braille fonts, or similar processes.

4.5 Image Rendering

4.5.1 Image Layout

At this point, we have the text in Braille format and the simplified image. These two components must be brought together to form a final image master for rendering in a tactile form. Braille must be a certain size to be readable, so the image must be scaled in the vertical dimension so that the height of the text boxes allows Braille to fit vertically in them. Scaling in the horizontal dimension may be different than in the vertical dimension, because the width of the Braille font is fixed and because the Braille may have more (or fewer) characters than in the original text. As a result, a Braille text block may have a different aspect ratio than the original corresponding text block. In order to place the Braille text blocks correctly into the image, we

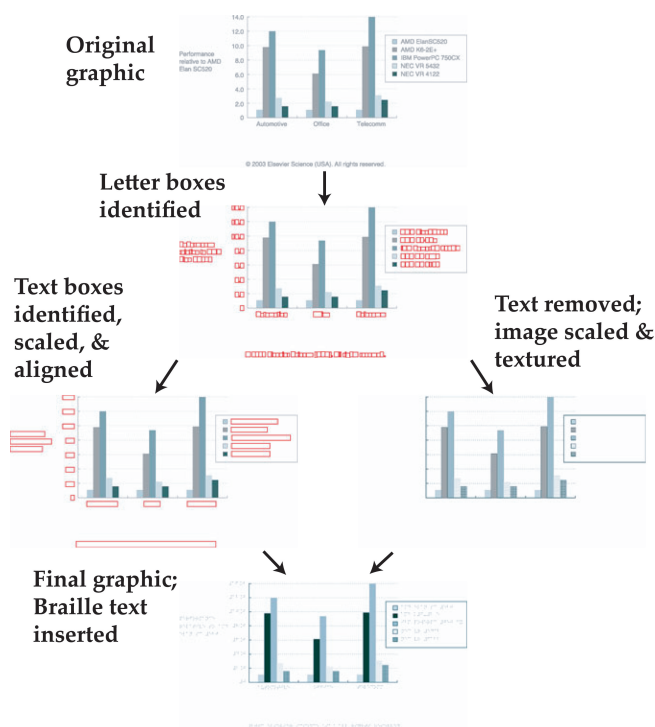


Figure 4: Intermediate Steps of the TGA.

must determine how the text boxes are justified within the original image: left, right, or centered. To determine justification, we developed a method based on the plane sweep algorithm from computational geometry [4]. Braille paper is limited in size, but we can scale the image up to 17 inches by 52 inches. Our experience revealed that this size can accommodate most figures that are found in textbooks.

4.5.2 Image Output

The final step entails rendering the image in an embossed form. For this step, we use ViewPlus’ Tiger Embosser [14], which prints both graphics and Braille as raised dots on thick paper. There are seven heights for the raised dots, but all seven are not easily distinguishable from each other.

4.6 Example Automated Image Translation

Figure 4 shows some of the intermediate steps in our current process—image processing, text processing, and image layout. First, bounding boxes are found for the individual letters. The letters are removed and processed separately into text boxes. We then have two images, one with just text and the other image without text. The image with only text is processed with OCR and Braille translation. The image without text is simplified to improve tactile readability. In this case, colors are replaced with textures. Both images are scaled so that the Braille text will fit. The final image merges the Braille only and pure graphic into one image. The final image can now be rendered using the Tiger Embosser. The sizes of the bottom three images in Figure 4 are actually much larger than the original, but for purposes of the figure, we left them the same sizes.

4.7 Current State of the TGA

As mentioned earlier, we are in the process of developing the Tactile Graphics Assistant. The tasks are now in separate components. We have focused our efforts on tasks that are

not accomplished easily with existing software applications. For the most part, a moderately advanced user of Photoshop or another image-editing application can handle the image simplification steps. Standard OCR software can handle, with some small amount of error, the translation of image blocks of text to ASCII text. Standard Braille translators, such as Duxbury, can handle, with some small amount of error, the translation of ASCII text to some Braille codes. Rendering can then be completed with the Tiger Embosser. Currently, there is no application that integrates all of the required steps to support automated translation of existing images of bar charts, diagrams, and other graphical images.

Our focus has been on classification, text segmentation, and layout. Currently, classification is very slow because the features are costly to compute, with respect to CPU cycles. Text segmentation is very fast, once a training set is created. The layout steps of image scaling and determining text block justification is very fast. All of these tasks have some unavoidable error, which we are working to minimize.

We envision that the tactile graphics specialist will need to spend time finding and correcting errors that arise in all of these steps, irrespective of whether the errors are from standard applications or our own custom application. In addition, the specialist would spend time enhancing some aspects of images, based on automated or personal evaluations. We expect that all the automated steps will take less than a minute per image, while all the steps that require human intervention, such as correction of errors and choice of textures, will take an average of 15 minutes per image. Our goal is that an entire textbook of several hundred images could be completed in 50 hours. Using current practices, several hundred images would take 1 to 5 hours per image, or 200 to 1,000 hours.

5 Research Team and Approach

Developing an effective TGA to support tactile graphics specialists requires expertise from many disciplines—human-computer interaction (HCI), image processing, machine learning, and tactile graphics translation. To accomplish this objective, we created a multi-disciplinary team of eighteen researchers, practitioners, and student consultants from the University of Washington’s Department of Computer Science and Engineering, Information School, DO-IT (Disabilities, Opportunities, and Internetworking Technology), and the Access Technology Lab. Two students from local high schools and one undergraduate student at the University of Washington serve as consultants. The team consists of sighted people, as well as people who have visual impairments (four who are fully blind and one who has low-vision). We describe the roles and respective contributions below.

- **Image processing and classification researchers:** This group consists of faculty advisors, graduate students, and undergraduate students, who are mainly from computer science. The group develops the various image processing and classification algorithms for the TGA. The group is also responsible for integrating components into the final application.
- **HCI researchers:** The human-computer interaction group consists of a faculty advisor, as well as undergraduate, graduate, and high school students, mainly from the Information School. The group conducted the contextual inquiry of tactile graphics translation and

is responsible for informing the design of the TGA, validating the image processing algorithms to ensure that they produce tactile graphics that students who are blind can use, and evaluating the TGA to ensure that tactile graphics specialists can use it to complete their translation work.

- **Practitioners:** This group consists of practitioners at the university who, on a regular basis, translate images for university students, staff, and consultants who are blind. Practitioners are using their current processes and documenting their translation work (i.e., steps completed and feedback received from students). They are also informing the design of the TGA. In the future, they will pilot the TGA.
- **Student Consultants:** This group consists of two high school and one undergraduate student who are blind. They use tactile graphics that the researchers and practitioners create and provide feedback on them.

Each group has a particular focus; however, we work closely together. For instance, we have a weekly meeting with all contributors to discuss issues from all perspectives and to coordinate our efforts. We also interact with vendors, such as ViewPlus Technologies and Adobe Systems, to leverage or help to improve their technologies.

6 Conclusions

We are using a holistic approach to automating the translation of graphical images into tactile forms. We studied and continue to study and take into consideration current work practices, and we aim to enable tactile graphics specialists to be more efficient in producing tactile graphics in ways that are similar to their current practices. In addition, we are developing new tools that will enhance their work practices, for instance by enabling them in the future to translate entire books of figures much more efficiently than is currently done. At every step, we rely on our team of student consultants to help us to evaluate the quality and understandability of the tactile graphics that we produce. We are still evaluating the current work practices of tactile graphics specialists, and we are in the process of building the functional components of the TGA. When the evaluation is complete and the components are ready, we will design work flows and user interfaces to support both batch translation of entire books of figures and translation of individual graphics.

7 Acknowledgments

This research was funded by National Science Foundation grants IIS-0335143 and IIS-0415273, a grant from the University of Washington's Royalty Research Fund, supplementary funds from the University of Washington's Information School, and a software license grant from Adobe Systems Incorporated. We thank ViewPlus Technologies Incorporated for their assistance with the Tiger Embosser. We thank: the study participants; Chelle Batchelor and Marc Shimpeno for helping to conduct the study; Jacob Christensen, Eileen Hash, Andy Jaya, and Terri Moore for early TGA contributions; Ammy Jiranida Phuwanartnurak, Melvin Knox, John Oleinik, Brian Rowe, Aaron Smith, and Raphael Sonnenfeld for their feedback on the paper; and the student consultants.

8 References

- [1] American Foundation for the Blind. Survey on the production of textbooks and instructional materials, 2000.
- [2] American Foundation for the Blind. Statistics and sources for professionals, 2003.
- [3] M. Batusic and F. Urban. Preparing tactile graphics for traditional Braille printers with BrlGraphEditor. In *Computers Helping People with Special Needs 8th International Conference, ICCHP 2002*, Lecture Notes in Computer Science Vol. 2398, pages 535–6, 2002.
- [4] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms*. The MIT Press, Cambridge, MA, 2nd edition, 2001.
- [5] A. L. Corn and R. S. Wall. Training and availability of Braille transcribers in the United States. *Journal of Visual Impairment & Blindness*, April:223–232, 2002.
- [6] Duxbury Systems. Recommended products, 2005. <http://www.duxburysystems.com/recproducts.asp>.
- [7] D. Forsyth and J. Ponce. *Computer Vision: A Modern Approach*. Prentice Hall, 2003.
- [8] W. Freeman and E. Adelson. The design and use of steerable filters. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(9):891–906, 1991.
- [9] J. Gardner. Access by blind students and professionals to mainstream math and science. In *Proceedings of Computers Helping People with Special Needs 8th International Conference, ICCHP 2002*, Lecture Notes in Computer Science Vol. 2398, pages 502–7, 2002.
- [10] J. Hennessy and D. Paterson. *Computer Architecture, A Quantitative Approach*. Morgan Kaufmann Publishers, San Fran., 3rd edition, 2003.
- [11] K. Holtzblatt and S. Jones. Contextual inquiry: A participatory technique for system design. In D. Schuler and A. Namioka, editors, *Participatory Design: Principles and Practice*, pages 180–193. Lawrence Earlbaum, Hillsdale, NJ, 1993.
- [12] W. Schiff and E. Foulke. *Tactual Perception*. Cambridge University Press, Cambridge, Eng., 1982.
- [13] A. S. Tanenbaum. *Computer Networks*. Prentice Hall, Englewood Cliffs, NJ, 3rd edition, 2003.
- [14] ViewPlus Technologies Inc. <http://www.viewplustech.com/>.
- [15] ViewPlus Technologies Inc. Tiger software suite, 2005. <http://www.viewplus.com/products/software/TSS>.
- [16] R. S. Wall and A. L. Corn. Production of textbooks and instructional materials in the United States. *Journal of Visual Impairment & Blindness*, April:212–222, 2002.
- [17] T. P. Way and K. E. Barner. Automatic visual to tactile translation, part I: Human factors, access methods and image manipulation. *IEEE Trans. on Rehabilitation Engineering*, 5:81–94, 1997.
- [18] T. P. Way and K. E. Barner. Automatic visual to tactile translation, part ii: Evaluation of tactile image creation system. *IEEE Trans. on Rehabilitation Engineering*, 5:95–105, March, 1997.
- [19] I. H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. Morgan Kaufmann, San Fran., 1999.