Automatic Visual to Tactile Translation, Part II: Evaluation of the TACTile Image Creation System

Thomas P. Way and Kenneth E. Barner

Applied Science and Engineering Laboratories
University of Delaware/Alfred I. duPont Institute
Newark, Delaware 19711
(302) 651-6830
E-mail: {way,barner}@asel.udel.edu.

Abstract— **This is the second part of a two-part paper that develops a method for the automatic conversion of images from visual to tactile form. In Part I, a variety of topics were reviewed including issues in human factors, access technology for tactile graphics production, and image processing. In this part, the material presented in the first part is used to motivate, develop and support the methods used in the development of a prototype visual-to-tactile translator called the *TACTile Image Creation System* (TACTICS). The specific choices made in the design of the system are discussed and justified, including selection of software platform, tactile output format, tactile image creation procedure, aggregate image processing sequences used, and principles from the discipline of psychophysics. The results of four experiments on tactile image discrimination, identification and comprehension are reported and discussed, and future directions in this area are proposed.**

Keywords— **blindness, image processing, microcapsule paper, tactile graphics, tactile imaging**

I. Introduction

In Part I of this paper, the pertinent human factors, technologies and techniques that are basic to the development of a system for the automatic translation of visual information into comprehensible tactile form are examined. Among the most important considerations are: (1) the lower bandwidth capability of the fingertip as compared to the eye, (2) the hierarchical nature of spatial perception and memory, (3) the state of the art in cost-effective output of tactile graphics, namely microcapsule paper, and (4) image processing techniques that can volumetrically reduce visual information to a level more appropriate for the bandwidth capabilities of the fingertip.

Based on the background given in Part I, we motivate and evaluate the design of our TACTile Image Creation System (TACTICS). We begin with an introduction of the software and hardware components of this system, namely the aggregate image processing algorithms, output media and visual-to-tactile conversion procedure. Results of experiments to evaluate the effectiveness of TACTICS to improve a subject's ability to: (1) discriminate, (2)identify and (3) comprehend tactile images are presented and analyzed. What these results, together with some anecdotal evidence, say about prospects for automatic visual-to-tactile conversion is discussed. Finally, future directions in which these findings may lead are proposed.

II. TACTICS: TACTile Image Creation System

Converting visual information into tactile information in an automatic, timely and ultimately comprehensible fashion is the force propelling development of this prototype system. The lessons learned from the areas of tactual perception, tactile graphic production and the applicability of image processing techniques to tactile graphic generation are extended to and applied in the creation of this system. The details of the system, including the justification for its development, the specific algorithms used for image simplification, the software and hardware utilized, and the complete procedure for acquiring, transforming and tactilizing visual information, are discussed.

The production of tactile graphics can be a time-consuming process of careful translation from visual to tactile form necessitating the involvement of a sighted person. Cost and timeliness prevent most blind persons from having ready access to the abundant high-quality computer images available on the Internet and elsewhere. With an automatic method for performing such translations, increased access to the wealth of computerized graphical information could be provided. Such information is, at present, essentially inaccessible, requiring the intervention of a sighted person to perform conversion from visual to tactile form. Automatic computerized conversion can be accomplished affordably, using readily available or easily adaptable technology, combined with the appropriate image processing techniques.

A technique for the automatic generation of tactile graphics involves acquiring an image, performing some simplifying processing, and displaying the result on a tactile output medium, such as capsule paper or a dynamic, real-time tactile display. The TACTile Image Creation System (TACTICS) is an attempt to further research in the area of automatic tactile graphic generation.

This research is based on a perceived lack of research being performed in addressing accessibility issues related to complex image information. The focus of much of the research in computer access to graphical information for blind persons is restricted to narrow categories of information, such as mathematical formulae, iconic navigation, or better auditory access to text. Our aim is to provide a general method for providing access to photographic and other visual information that is in electronic form.

A. Software Development

The software for the prototype system for automatic generation of tactile images is implemented in the C programming language as an extension to the X-windows image processing application *XV*, developed at the University of Pennsylvania [4]. As of publication, the complete source code for this package is readily available via anonymous `ftp` at `ftp.cis.upenn.edu` in the directory `pub/xv`. The license fee is quite reasonable for this user-friendly software, and it was found to be easily extended to include additional image processing algorithms. The extended version is available via `ftp` at `ftp.asel.udel.edu` in `pub/sem/xv-mod.tar.Z`. Instructions on how to add additional algorithms to the *XV* package are in the file `xvalg.c`.

B. Tactile Output using Microcapsule Paper

Microcapsule paper was chosen as an output medium due to its wide availability, relatively low cost, and ability to render tactile graphics quickly. We compared the two known brands of paper on the market, Repro-Tronics *Flexi-Paper* and a paper imported by the Matsumoto Kosan Company. A comparison of the manufacturers' specifications for the two types of paper reveals that there is very little difference in the vital qualities of resolution, response time, cost and displacement. One laboratory observation, as measured using a mill-meter, is that the displacement achievable with the Matsumoto Kosan paper tends to be more consistent in practice than the Repro-Tronics paper. Measurements reveal this to be the case, but also show that typical displacement is approximately 1mm for both varieties of paper.

The significant difference between the two appears to be the durable nature of the Flexi-Paper, which is highly resistant to folding and crumpling. The stiffer Matsumoto paper is more familiar in feel to the blind community, being similar to the heavy paper used by embossing braille printers, but is prone to cracking and creasing under adverse conditions. For purposes of our experiments, we used both types of paper and discovered that subjects often preferred the slightly stiffer feel of the Matsumoto paper versus the spongier feel of the Flexi-Paper.

Original and processed computerized images were first printed out on a commercial 600*dpi* office laser printer. Next, they were copied onto capsule paper using a typical office photostatic copier machine. Finally, tactile images were developed, or puffed up, with a Repro-Tronics *Tactile Image Enhancer*.

C. Experimental Procedure for Tactile Image Creation

The procedure for producing a tactile image from a visual one is straightforward. The involvement of a sighted person is necessary in the current stage of our research system. Future versions of TACTICS could be made to operate in an unsupervised manner, eliminating the need for a sighted person to be involved. The procedure involves three phases:

1. **Acquisition of Images**: Images were acquired in a fairly random manner from standard image processing benchmark collections, scientific data acquisition, and from a wide array of sources available on the World Wide Web. Every attempt was made to select a representative sampling of the available images. We also looked for candidates from similar classes of images, for example faces, or more generally rounded images, which could prove difficult to distinguish from one another once simplified, as a way to test how ambiguity is dealt with by our prototype system when used by experimental subjects.
2. **Simplification**: The preparation of simplified images was achieved using a number of differing, aggregate, image processing sequences. An image was first loaded into *XV*. Then, the applicable sequence of image processing algorithms was applied. Finally, this processed image was printed on a laser printer in preparation for expansion in the subsequent phase.
3. **Tactilization**: The printed version of the processed image was photocopied onto one of the two types of capsule paper. The capsule paper was then fed through the *Tactile Image Enhancer*, creating the raised tactile image. This procedure was repeated for all images using the variety of image processing algorithm sequences as specified in the experiment protocol.

D. Image Processes

The four experiments conducted made use of the following image simplification techniques consisting of individual and composite processes:

1. **No Processing**: A tactile image is produced directly from the original grayscaled version of the image (Figure 1a). Experimentally, this image serves as a benchmark upon which the effectiveness of further processing can be measured. The unprocessed image represents the visual information in its raw form.
2. **Segmentation**: Performing a segmentation divides an image into regions. In this application, we perform a binary segmentation via adaptive thresholding using the K-means segmentation algorithm, which produces regions of white and black only (Figure 1b). This representation is modeled on the way it is believed that the mind classifies and stores image information, namely in some hierarchical fashion, from general characteristics to specific [3], [10]. In the case of segmentation, general characteristics are emphasized. Note that in some instances negation was applied following an application of segmentation to emphasize content rather than background.
3. **Edge Detection**: Emphasizing the edge information in an image might be all the simplification that is needed. Much of the theory discussed above and in Part I of this article indicates that converting an image into a simpler sketch or line-drawing representation should enhance recognition. The Sobel edge detection operator is used here (Figure 1c), as it is widely used and considered to be effective for general purpose edge detection, although any one of a

number of edge detectors could quite easily be substituted. Note that thresholding is performed on the image, with edge points being set to one intensity value while non-edge points are set to a second value. In this way, a binary edge-only version of the original is produced. Depending on the implementation of the thresholding conducted in association with a given edge detection algorithm, it may be necessary to apply negation to the result.

4. **Blurring, Edge Detection, Segmentation and Median Filtering**: This aggregate process takes into account cognitive and perceptual theory as much as possible to produce a result that, at least visually, appears to be quite simple (Figure 1d) while still clearly resembling the original. The blurring step represents the lower bandwidth capabilities of the fingertip as compared with the eye. The result of this blurring has a potentially beneficial side effect, thicker edges, which appears during the subsequent edge detection. Without the initial blurring step, the resulting lines in the final representation tend to be thinner and sometimes less continuous (Figure 1e). When the edge detector is applied without thresholding to the blurred image, edges appear thicker due to the slight spreading or softening of rapid intensity changes in the original. The segmentation step is used in place of the implicit thresholding normally performed in conjunction with an edge detection. This adaptive thresholding tends to preserve more of the original edge information than an implicit thresholding. The final median filtering step removes any stray noise that was not removed by the segmentation. In fact, there is a proportion of noise that is enhanced rather than removed by the adaptive thresholding of the segmentation step. Median filtering counteracts much of that effect.

III. Evaluation

The primary goal of the procedures used by TACTICS to convert visual information automatically into tactile information is to provide meaningful access to previously inaccessible content. A series of experiments were conducted to evaluate the effects of this prototype system upon a subject's ability to (1) discriminate, (2) identify and (3) comprehend tactile representations of visual information. The selection of the specific aggregate image processes for use in these experiments is discussed and justification is given linking these processes with theories of psychophysics. Descriptions of the subjects, images used and experimental procedures are provided. The results of each experiment, including data comparing results based on types of microcapsule paper and the level of vision of subjects, are reported and analyzed.

The protocol used in these experiments was designed to evaluate the effect of TACTICS upon the accessibility of visual information in a tactile form. Every attempt was made to acquire a diverse sample of subjects and images, and to assure that experimental materials were produced in an automatic and uniform fashion free from the aesthetic biases of a sighted person.

Blind, low vision and sighted subjects were used in the following experiments. As noted in Part I, the tactile acuity of blind and sighted persons, whether male or female, is essentially identical [7], although blind persons tend to have more experience making active use of the sense of touch [11], while sighted subjects generally have a more highly developed visual memory [9]. Any difference in the performance of blind and sighted subjects is noted and discussed.

As mentioned above, images were gathered electronically from a variety of sources and were prepared first by grayscaling any that were color images to achieve uniformity. This homogeneity was necessary because microcapsule paper expands only in response to the color black. Depending on the experiment, one or more image processing algorithms was then applied in a specific order to each image.

Once the images were processed, they were printed out on a standard office laser printer, photocopied onto sheets of microcapsule paper, and expanded using the Tactile Image Enhancer. Both types of capsule paper were used in the production of experimental materials.

A. Psychophysics and Experimental Procedure Justification

To evaluate the effectiveness of this processing for automatic generation of tactile images from visual images, five sets of experiments were performed. These experiments were designed to measure performance on a basic psychophysical level. The field of *psychophysics*, the study of physical and psychological aspects of perception and their interrelationships, identifies four basic perceptual tasks [5]: (1) detection, (2) discrimination, (3) identification, and (4) comprehension. As with all the senses, these four attributes apply to tactual perception.

1. **Detection**: Measuring *detection* using the sense of touch involves designing a task that addresses the question, "Is there anything there?" As discussed in Part I, many limits of the physical detection abilities of the fingertip are known. Since the properties of microcapsule paper produce tactile graphics that are well within the range of such touch perception, any experiment designed here would be trivial. Thus, it is safe to accept as an assumption that TACTICS produces tactile images that are detectable. Thus, no experiments were performed to measure detection of tactile images, as all experiments relied on the implicit ability of subjects to detect the raised tactile images.

2. **Discrimination**: The ability to discriminate, which answers the question, "Is this stimulus different from that one?", is an important perceptual task for any of the senses. For the sense of touch, *discrimination* tells us simply whether two tactile objects are the same or different. The experiments to measure the effectiveness of TACTICS to aid in discrimination involved a task similar to the traditional matching game of *Concentration*. In the study, subjects felt one of a closed set of similarly processed tactile images and then attempted to locate the identical tactile image from among a randomly arranged duplicate set. In further experiments, subjects felt a series of arbitrarily paired tactile images for a period of time and then reported

whether or not each of the pairs felt similar or dissimilar.

3. **Identification**: Being able to identify what something is by its perceived characteristics is another basic perceptual task. *Identification* as it applies to tactile images involves the effectiveness of a representational technique to allow a person to answer, "What is it?" The cognitive load imposed by identification is higher than that for detection or discrimination, so the experiment to measure it is also more involved. In the experiment to assess this factor, subjects felt a series of tactile images, and for each image were given four categories and asked to identify into which category each stimulus belonged.

4. **Comprehension**: *Comprehension* means that questions regarding the content of an image should be answerable. Comprehension is generally accepted as a key consideration it the effectiveness of any perceptual event and therefore is an important factor to explore in the design of an interface to a GUI environment for blind computer users. This experiment measured how well a selected TACTICS aggregate image process affected comprehension of tactile images. Subjects were provided with a brief description of each image and then were asked a number of questions regarding the content of each image.

B. *Subjects*

Ten subjects ranging in age from 22 to 60 were used in this experiment. The subjects participated voluntarily and came from a variety of backgrounds, including college students, homemakers, computer programmers, and a retired chemist. All subjects were educated to at least the four-year college level. Seven subjects were male, three were female. Three subjects were blind, seven were sighted. The two male blind subjects were adventitiously blind, one at age 19, the other at age 39. The one female blind subject was congenitally blind. Additionally, one male subject was classified as low vision. Subjects had little or no experience with tactile images and microcapsule paper.

C. *Images used*

The images collected for use in these experiments represent a diverse sampling (see Appendix for detailed list). Attempts were made to include ambiguity of content and a variety of types of content. The set of images is comprised of these basic groupings: (simple) space shuttles, nuclear mushroom clouds, a tornado, (complex) astronauts on the surface of the moon, electron micrographs of viruses, bacteria, a microscopic worm, (oval) four faces, a hot air balloon, (round) the moon, 2 planets, a chocolate chip cookie, a house fly eyeball, and (squared) desktop and notebook computers, a bridge, a city skyline.

D. *Experiments*

Four experiments were conducted to evaluate the effectiveness of the system. These examined simple and timed discrimination, identification and comprehension of tactile images. Note that approval to conduct these experiments was obtained from the University of Delaware Human Subjects Review Board.

D.1 Simple Discrimination Experiment

Two forms of tactile discrimination experiments were conducted. In this first experiment, subjects were allowed to explore freely the initial and secondary tactile images for a total of one minute per pair. This provided the subjects with enough time to glean some information about both the general shape of the image and some of the more prominent internal details. As described below, a matching task was conducted to measure the effectiveness of the four image processing techniques under consideration when applied strictly for purposes of discrimination. This task is roughly analogous to that of a sighted person leisurely browsing through photographs in a magazine or on the Internet, for instance.

In addition to measuring the effects of processing on discrimination, a study was conducted to determine the effect on discrimination of one form of microcapsule paper versus the other. Results from this comparison of microcapsule papers will be extrapolated to other more complex tactual perception tasks of identification and comprehension.

Materials were produced on both types of microcapsule paper using identical tactile images for each set. Each set consisted of 40 sheets, each with a pair of raised tactile images per sheet, one on either side of a raised line that divided each sheet in half. Each tactile image was limited to four inches in width, which is within the width of one hand span. The height of each image followed proportionally from the scaling of the width and also stayed well within the height of one hand span. Samples were drawn in pairs from the set of original images to prepare the testing materials. Half of the pairs consisted of identical images, and half were not identical. Each pair was prepared using each one of the four processes under consideration, and the same processing was applied to both images on a sheet.

Subjects were asked to perform a discrimination task using one complete set of 40 tactile image pairs. Subjects were seated at a table, blindfolded if sighted, and presented with each of the 40 sheets from a given set in an arbitrary sequence. For each sheet, subjects freely explored the pair of tactile images on the sheet for a period of time totaling one minute and were then asked to report whether the images felt the same or different. Subjects also could reply that they could not say one way or the other, although this reply was rarely used. During this procedure, responses were recorded, as were any unsolicited comments made by the subject in reaction to the materials or procedure. Subjects were given neutral feedback after each matching task.

Overall, each complete set of 40 tactile image sheets was used with five (one-half) of the subjects, so that some comparison could be made of the two forms of microcapsule paper under identical experimental conditions. The same set of 40 sheets of testing materials that was used for a subject in this simple discrimination experiment was randomly reordered and used in the following timed discrimination experiment for the same subject. Note that half of the subjects completed the simple and timed discrimination tasks using the Repro-Tronics paper, and the other half the Matsumoto Kosan paper. With the exception of the type of paper on which the materials were prepared, the two complete sets of testing materials were identical in every respect.

The results of the simple discrimination experiment are summarized in the following three tables. Table I provides an overview of how subjects performed on average for each of the four image processes applied. Analyses of variance indicate significant interaction between the unprocessed original and any of the other processing performed. Compared with results expected by chance, significant interaction was found for all forms of processing[1]. In the case where no processing was used, no interaction was noted.

Table II compares the results of using one type of microcapsule paper versus the other. Table III compares the performance of blind versus sighted subjects. In these tables, there is no indication of interaction between groups of subjects based on the different processes applied, whether compared by output medium or level of vision.

In these tables, the *Mean Pct. Matched* is the computed average percentage of correct responses. The results of analyses of variance are denoted by p and compare results for each of the forms of processing with those for the unprocessed originals, and with chance (50%, in this case).

D.2 Timed Discrimination Experiment

The second experiment imposed a strict time limit of 10 seconds for exploration of each pair of images. This limited-time experiment was designed to measure the effectiveness of TACTICS image processing techniques in a situation reminiscent of a sighted person skimming or quickly scanning through a series of images, making quick determinations. The goal of this experiment was to test how use of these image simplification methods might affect the ability of a blind computer user to perform browsing and navigation tasks using touch in a GUI environment on a level comparable to a sighted computer user.

The same subjects were used for this experiment as in the above simple discrimination experiment. Materials were the same as in the simple discrimination experiment, with identical materials being used for the same subject for both the simple and timed discrimination experiments. Although the identical materials were used for a given subject, they were randomly reordered to counteract possible bias related to ordering. The procedure for this experiment was identical to the above discrimination task, with the single exception being that subjects were limited to 10 seconds per image pair matching task.

The results of the timed discrimination experiment are summarized in the following three tables. Table IV provides an overview for how subjects performed for each of the four image processes applied. Analyses of variance indicate significant interaction between the unprocessed original and any of the other processing performed. Compared with results expected by chance, some degree of interaction for all forms of processing was found. In the case where no processing was used, however, there is no indication of any significant degree of interaction.

Table V compares the results of using one type of microcapsule paper versus the other. Table VI compares the performance of blind versus sighted subjects. Analyses of variance for these two tables show that there is no statistical evidence of interaction between groups of subjects based on processing used, whether compared by output medium or level of vision.

Results of the performance of subjects in the simple and timed discrimination tasks are compared in Table VII. While the trend based on mean performance was for subjects to discriminate tactile images slightly less successfully under time pressure, no interaction was found between the two discrimination modalities. This result is therefore inconclusive in regard to the effect of time pressure on tactile discrimination.

D.3 Identification Experiment

The identification task had subjects explore a series of tactile images and attempt to classify each into one of four categories that varied for each image. This task was designed to provide some insight into the effectiveness of TACTICS to produce a tactile image that resembles the original in such a way that it is identifiable given some small amount of pre-information. This is analogous to the visual task of identifying photographs based on some small amount of textual information, such as a caption.

Used for this experiment were the same subjects as above and a subset of 10 images. These 10 images were selected from the original set of 30 images to reflect a diversity of shape and content. The images were processed using each of the four processes under consideration, with the result of each process placed onto an individual sheet of the Matsumoto Kosan microcapsule paper. The result of this preparation was a set of 40 sheets, each with a tactile image that was processed in one of four ways.

For each of the 10 images, four possible categories were defined. Of these four categories, one correctly identified the content of the image, two identified objects that may closely resemble the content of the image, and one resembled the content of the image less closely.

During the experiment, the 40 sheets were presented in an arbitrary order as the experimenter verbally listed the four associated categories. During this listing the subject freely explored the tactile image, and the categories were listed again per subject request. At the conclusion of a period of no more than 30 seconds, the subject was asked to state which category most closely matched the tactile image that was explored. The responses were recorded, and the procedure was similarly repeated for all tactile images in the set.

The results of the tactile image identification experiment are shown in the following two tables. Table VIII summarizes overall performance of subjects on the identification task for each of the four forms of image processing applied in the production of the tactile images. Analyses of

[1] Analyses of variance comparing results for each form of processing with no processing and with chance differ. This disparity is due to the varied distribution (varying standard deviation) of individual subject performance versus strict chance (standard deviation of 0).

variance indicate significant interaction between the unprocessed original and all of the other processing performed. Significant interaction is also indicated when comparing the various processing with results expected by chance (25%, in this case).

Table IX compares performance of blind versus sighted subjects in the same task and for the same four processes. Comparison of sighted and blind subjects did not indicate interaction between the subject groups.

D.4 Comprehension Experiment

This experiment measured the ability of subjects to comprehend tactile images prepared using the aggregate processing. The assumption was made that unprocessed images would be incomprehensible; and, indeed, this assumption was supported by results of the discrimination and identification experiments. Based on results of the discrimination and identification experiments, the aggregate process was found to be best for improving performance among those considered, and so it was selected for use in this experiment.

For this experiment, the same subjects were used as in the above experiments. The materials used consisted of 10 images selected from the original set of 30, chosen to represent a diversity of content and shape. Each image was processed using the aggregate process described above, and placed onto a sheet of Matsumoto Kosan microcapsule paper for subsequent raising. Due to the results of the prior experiments which indicated no statistically significant interaction based on type of paper used, use of this paper was deemed sufficient.

Associated with each of the 10 processed images was a brief one or two sentence description of the image and four questions designed to test a subject's comprehension of the image's content. Questions were of the "True or False," "Multiple Choice" and "Locate the (fill in the blank)" variety, with the number of choices limited to two. The questions were designed to be of the sort that generally would be easily answered by a sighted person viewing the original image. Some questions asked the subject to locate some feature in the image, such as, "Locate the tail fin of the space shuttle." Other questions concerned understanding some feature in the image. For example, associated with the image of a space shuttle in the process of landing, one question was, "Is the shuttle landing from left to right, or right to left?" Finally, the third and most difficult form of question asked the subject to reason about and draw some conclusion about the content of the image; for example, with an image of a desktop computer was the question, "Is the computer on or off?"

For each of the 10 tactile images, the brief description of the image was read aloud while the subject explored the image. As the subject continued to explore the image freely, each one of the four questions and the two possible answers for each was read aloud. For questions with a verbal response, the experimenter merely recorded the subject's reply on a data collection sheet. For questions in which a subject was asked to locate a specific feature in the image, the experimenter observed the movement of the subject's hand and evaluated both the final location the subject indicated and the subject's verbal reply.

Also recorded were any unsolicited remarks or comments made by the subject and observations made by the experimenter during the experimental procedure. Comments often referred to the difficulty a subject may be having with a particular question or tactile image, some interesting discovery that had been made by the subject regarding the image, or the reasoning used by the subject in reaching a particular conclusion. Observations made by the experimenter included noting initial reactions of the subject, exploratory movements used, and any other reactions that seemed noteworthy.

Results of the tactile image comprehension experiment are shown in the following two tables. The first, Table X, displays subject performance for the three comprehension subtasks as well as overall performance. Analyses of variance comparing these results with chance (50%) indicate significant interaction, suggesting little possibility that the successful performance of subjects occured by random chance.

Table XI compares how blind and sighted subjects performed in this experiment. Comparing subjects based on level of vision indicates probable interaction for the location task. No interaction was found when comparing subjects by level of vision and the understanding and reasoning tasks.

The disparity in performance on the location task may be due to differences in visual memory, with sighted subjects possessing more familiarity with visual material in general than blind subjects [9]. As a result, sighted subjects are more aware of relative size and position of objects as represented in an image. This difference in positional awareness could account for the differences in performance of blind subjects versus sighted subjects for the location task. The lack of significant differences in performance for the understanding and reasoning tasks could indicate that a more developed visual memory is not necessary to these tasks.

D.5 Significance of Results

The analyses of variance for all experiments did not indicate interaction between groupings of subjects based on level of vision. This result is expected based on results of previous studies that found no significant difference between the tactile abilities of blind and sighted persons [7]. There was also no interaction found for groupings based on output medium. Since the characteristics of the two types of papers are similar in most respects, this result is not remarkable. However, the two forms of paper do vary in the property of stiffness, with Matsumoto-Kosan paper being significantly stiffer than the Repro-Tronics Flexi-Paper, which is flexible by design. It appears that stiffness alone is not a significant factor in any of the tactual perception abilities we measured. In spite of the lack of statistical differences, some subjects indicated a preference for the stiffer Matsumoto-Kosan paper over the Flexi-Paper, noting its "better clarity" or "nicer feel." These personal reactions did not appear to translate into differences in the performance of subjects.

Comparing mean performance on the various tasks versus chance performance reveals an apparent trend of improvement based in some measure on the degree of simplification. More formally, analyses of variance based on type of processing showed significant interaction between each form of processing used when compared directly with unprocessed originals. These analyses repeatedly indicated that the application of simplifying image processing techniques in the translation of visual images to tactile images improved performance of subjects in discrimination, identification and comprehension tasks. This result is quite favorable, particularly when compared with subject performance on similar tasks using unprocessed tactile images.

Equally as important as these statistically significant results are the observational and anecdotal evidence gathered during these experiments. The significance of that evidence, in light of the results from these experiments, are discussed below.

IV. Observations, Discussion and Conclusions

In the course of evaluating TACTICS, a number of general observations were made that anecdotally enhance the raw tabulated results. These observations and results are discussed, and conclusions are drawn, regarding the effectiveness of TACTICS as a method for providing blind persons with tactile access to visual information.

A. Observations

While conducting these experiments, the experimenter (the first author) recorded observations in addition to the raw response data. The observations are summarized here for each of the four experiments followed by more general observations.

During the simple discrimination experiment, subjects typically took more time for the first few tasks while they became accustomed to exploration of the tactile images and experimented with different techniques for exploring them. Most subjects developed a two-handed approach to this discrimination task, using one hand for each of the images in a pair and synchronizing the movements of the two hands. While using this technique, subjects usually first attempted to determine the general shape of the image. Then, subjects performed further exploration, again in tandem, to examine details of the image.

One technique developed by many subjects during the timed discrimination experiment was the use of a brushing motion, drawing the fingertips of the hand the length of each tactile image. This technique was fast, and seemed to provide enough basis to discriminate between tactile images. This brushing technique was performed with two hands in tandem, or with a single hand on each image individually, approximately equally as often.

The identification experiment proved to be the most challenging for subjects based on their reactions during the experiment. They often expressed frustration at not being sure about which category was the best match for a given image. Subjects often used a process of elimination to narrow down possible choices from among the four categories given for each tactile image. Another technique subjects used was to explore an image four times, basing each exploration on the assumption that the image was one of the four categories. Guessing was a common strategy used by subjects when they could not determine a category for an image. Guessing occurred most frequently with the relatively feature-free unprocessed images.

Subjects seemed especially to enjoy the final experiment measuring tactile image comprehension. It was not uncommon for subjects in the process of answering one question to make comments and remarks about the contents of the images that turned out to be the answer to later questions. For example, while exploring a tactile image of the space shuttle, a number of subjects indicated the locations of the nose and tail of the vehicle while answering a question about its direction of travel. An image of an astronaut working on the surface of the moon was the most difficult for all subjects. Visual observations made while the subjects were exploring this particular image indicated that the presence of an edge denoting the horizon as well as the busy pattern of edges generated by the texture of the surface of moon made image comprehension difficult. Subjects frequently mistook a U.S. flag in the image for the astronaut's backpack, and that miscalculation caused incorrect answers to other questions about the astronaut's position and activity.

In general, each subject tended to have relative ease or difficulty with the same images and processing that other subjects had ease or difficulty with, respectively. Subjects also tended to gain confidence in performing the various tasks as they gained experience with exploring the tactile images. The general technique that was arrived at by each of the subjects was one of exploring first the overall shape and size of a tactile image, then feeling for details.

Blind subjects had more difficulty with some of the more visual concepts, particularly with images of large objects such as planets, the space shuttle and very small objects, such as the Streptococcus bacteria and Ebola virus. Blind subjects often expressed more apprehension at the outset than sighted subjects, although blind subjects had more experience relying on the sense of touch than sighted subjects.

For subjects who performed the two discrimination tasks using Flexi-Paper, there was an initial reaction to the improved clarity provided in the third experiment which was conducted using the Matsumoto Kosan microcapsule paper. Subjects did not appear to have more difficulty in similar tasks using one variety of paper versus the other. Another comment regarding Flexi-Paper was that some pairs of images seemed to be expanded to differing heights. Interestingly, careful measurements taken in the laboratory using a mill-meter revealed that heights of the tactile images referred to were identical. Possible explanations are that the tactile acuity for the left and right hands may have varied slightly for some subjects, or that the difference in stiffness of the two papers may have affected the outcome.

B. Discussion

Comparison of various results of these experiments provides further insight into the degree of effectiveness of TACTICS for automatic generation of tactile images. Comparing mean percentage of matches for the simple and timed discrimination experiments reveals that performance degraded fairly uniformly, and even then only slightly, when going from the untimed to timed task. For unprocessed tactile images, discrimination for both tasks was about chance (50%). Blind subjects successfully discriminated between tactile images about 10% less frequently than sighted subjects, although it must be noted that various analyses of variance did not indicate a statistical significance for this observation. One possible explanation for this slight difference in performance is a lower level of pre-experiment confidence among the blind subjects, who tended to be somewhat more apprehensive about how well they would perform in the experiments.

Comparison of results from the simple discrimination experiment and analysis based on the two types of microcapsule paper was statistically insignificant yet observationally revealing. Ability to discriminate segmented tactile images was significantly better for Flexi-Paper than Matsumoto Kosan paper. For the timed discrimination experiment, Flexi-Paper produced a significantly higher percentage of successful discriminations than the Matsumoto Kosan paper for segmentation and the aggregate process, and slightly higher for Sobel edge detection. An anecdotal explanation for this result may be that subjects reported a more positive, albeit subjective, reaction to the stiffness of Flexi-Paper over Matsumoto Kosan paper. As mentioned in Part I, specifications and actual measurements comparing the expanded characteristics of the two papers do not provide an empirical explanation for any difference.

In general, simplification to any degree produces improvement in discrimination rate. For the discrimination experiments, when no processing was applied at all, success rates tended to be at about chance. Subjects correctly discriminated between tactile images about 75% of the time, on average, for images processed using Sobel edge detection alone, and slightly better than that for images that were prepared using segmentation. The aggregate process allowed subjects to discriminate correctly from 85% to 90% of the time, and some subjects performed perfectly.

The identification experiment proved to be the most difficult for subjects when compared with the mean percentage of matches for the other experiments. The aggregate process again produced the best rate of success, followed by segmentation and Sobel edge detection, with no processing trailing far behind. Although blind and sighted subjects performed nearly identically on segmented and on edge detected images, sighted subjects performed better when exploring images prepared using the aggregate process, correctly identifying image content more than 88% of the time while blind subjects were successful 77% of the time. This difference may be due to the visual nature of the images and the content therein, and perhaps due to an unintentional visual bias in preparation of the experimental materials.

For the comprehension experiment, subjects generally performed quite well on all tactile images and questions, with the exception of one particular image of an astronaut working on the surface of the moon. The featured content of most images was photographed either straight-on or in profile, producing tactile images that were straightforward to explore and comprehend. The astronaut image was captured at somewhat of a downward angle, producing a confusing horizon line crossing the image at the level of the astronaut's neck. Further research is needed to determine what image processing techniques exist, or can be developed, to handle potentially confusing information, such as horizon lines, adequately, within the framework of automatic conversion to tactile representation.

Generally and informally, blind and sighted subjects performed about the same on all tasks. Blind subjects tended to have more difficulty than sighted subjects in locating specific features within images, perhaps due to a less developed visual memory or lack of experience with characteristics of visual representation from which the tactile images were generated. Blind subjects performed better than sighted subjects with tasks involving understanding the contents of tactile images, perhaps due to more experience in relying on the sense of touch to gather and interpret information.

C. Conclusions

The objective of this work was to provide meaningful access to computer-based visual information to blind persons, and to do so automatically. Image processing techniques were applied to images to produce simplified versions of the originals, appropriate for output as tactile graphics. These image processing algorithms, and the aggregate processes resulting from various combinations, were selected based on effects that were analogous to principles of psychophysics and the science of tactual perception. The result was a system that converts a visual image into a tactile image in an automatic, timely and comprehensible fashion, as supported by results of evaluative experiments.

The significance of the development of this prototype system is that it makes it clear that reasonable and comprehensible access to visual information can be provided to blind persons, and done so without the intervention of a sighted facilitator. Thus, a blind computer user, for instance, could "surf the web", browse a CD-ROM collection of computerized images, or navigate a GUI, unaided and with a better degree of comprehension than is currently possible.

This increased access to visual material can facilitate broader educational and professional opportunities, particularly in areas with a strong tendency toward visual presentation of information. For example, persons with disabilities, including blindness, are currently underrepresented in science-, engineering- and mathematics-related disciplines. The techniques developed in this system can translate the visual information from these fields into tactile form, providing students and professionals with better access to diagrams, graphs and images ranging in scale from the microscopic to the cosmic.

V. Future Directions

The effectiveness of TACTICS at converting visual information into comprehensible tactile information lends credence to the possibility of future investigation in this and related areas. Among the possibilities are:

Development of End User Application: TACTICS can be developed further into a stand-alone application. Such an application would be invokable from the command line, perhaps being called in place of a print routine from a web browser. If a blind computer user desired to explore a tactile version of an image, the application would automatically handle processing of the image.

Extension to Refreshable Tactile Display: The use of image processing appears to be a natural and effective method for production of simplified images suitable for output in tactile form. With such effective pre-processing available, the task of expedient output becomes more important. There is a definite need for real-time, dynamic tactile display technology that could display tactile images efficiently.

The techniques developed in this paper for converting visual information into tactile information lend themselves to use as a front-end to such a real-time, dynamic tactile display device. Such a display would overcome the reliance on a sighted person that a blind person might experience when utilizing microcapsule paper as an output method. One limitation of past technology developed to display tactile graphics was that its effectiveness was determined by the relative simplicity of the material being displayed. Using image processing techniques, as in TACTICS, visual information could be prepared readily for meaningful display in tactile form.

Multimodal Interface: Simplified tactile representations of images, maps and other infrequently changing visual items could be combined with touch-screen technology to create a multimodal interface. With some initial configuration, positions on an image or map could be associated with audio feedback, as with the Nomad. The advantage of this approach would be the speed with which tactile materials could be prepared, and the flexibility offered by the automatic simplification techniques of TACTICS.

Mapping Color to Texture: Segmentation divides a two-dimensional visual representation into regions based on related colors or intensity levels. The result of such a segmentation could be used subsequently to associate the color of each region with a distinct texture, thus providing a blind person with more complete access to the original content of the visual information.

One long-standing problem of graph theory was the *four-color conjecture*, the notion being that any planar graph, for our purposes a two-dimensional visual representation such as a map or photograph, could be segmented into regions and those regions colored using only four colors, and with no two adjacent regions being assigned the same color [6], [8]. Originally posed in 1852 by Francis Guthrie, the four-color conjecture was finally proved in 1977 [1], [2], although finding a four-coloring is not necessarily fast. Given that four colors is sufficient, relaxing the coloring to some reasonably small number (say 10) would allow a very fast coloring to be performed. Thus, a tactile image, simplified using TACTICS, could be segmented and colored quickly using any of a number of simple graph-coloring algorithms. By uniquely mapping textures to colors, it may be possible to preserve more of the original visual information.

Even simpler would be to apply a *K-means segmentation* to an image, with K = *desired number of colors*, and apply the color-texture mapping to the result. This method might not provide as good a texture mapping as a more computationally expensive technique, but it would certainly be fast and may be sufficient for enabling comprehension of tactile images, which is the goal.

Acknowledgments

Work on this project was performed at the University of Delaware Center for Applied Science and Engineering, located at the Alfred I. duPont Institute, in Wilmington, Delaware. As part of the "Science, Engineering and Mathematics" project, funding was provided by the National Science Foundation, grant number HRD-9450019. Additional support was provided by the Nemours Research Programs.

Appendix

Listing of Images

1. Desktop computer
2. Desktop computer (another angle)
3. Notebook computer
4. Astronaut taking soil sample
5. Astronaut planting flag pole
6. Space shuttle landing (left to right)
7. Space shuttle landing (right to left)
8. Double-layer plume nuclear mushroom cloud
9. Single-layer plume nuclear mushroom cloud
10. Micrograph of the eyeball of Drosophiliaeye (house fly)
11. Electron micrograph of a Streptococcus bacteria (96,000x)
12. Planet Saturn
13. Planet Jupiter
14. Moon
15. Chocolate chip cookie
16. Close-up of President Ronald Reagan
17. Close-up of President Bill Clinton
18. Close-up of a researcher (Tom Way)
19. Close-up of Albert Einstein
20. Hot air balloon
21. Two-shot of Beavis and Butthead
22. Two-shot of Bill Clinton and Al Gore
23. Chinese student blocking tanks in Tiananmen Square
24. Chinese student blocking tanks in Tiananmen Square (another angle)
25. Golden Gate Bridge in San Francisco

26. Twin Towers in New York City
27. Tornado funnel cloud in Oklahoma
28. Electron micrograph of a cell shedding HIV particles
29. Electron micrograph of a Pinosyllis Heterocirrata worm
30. Electron micrograph of the Ebola virus

References

[1] K. Appel and W. Haken. Every planar map is four colorable, part I: Discharging. *Ill. J. Math.*, 21, 1977.

[2] K. Appel, W. Haken, and J. Koch. Every planar map is four colorable, part II: Reducibility. *Ill. J. Math.*, 21, 1977.

[3] I. Biederman. Human image understanding: Recent research and a theory. *Computer Vision, Graphics and Image Processing*, 32, 1985.

[4] J. Bradley. *XV Online Documentation.* University of Pennsylvania, 3.10 edition, 1994.

[5] S. Coren and L.M. Ward. *Sensation and Perception (3rd Edition).* Harcourt Brace Jovanovich, San Diego, 1989.

[6] T.H. Cormen, C.E. Leiserson, and R.L. Rivest. *Introduction to Algorithms.* MIT Press, Cambridge, Massachusetts, 1990.

[7] B. Lowenfeld. Effects of blindness on the cognitive functions of children. In B. Lowenfeld, editor, *Berthold Lowefeld on Blindness and Blind People.* American Foundation for the Blind, New York, 1981.

[8] F.S. Roberts. *Applied Combinatorics.* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1984.

[9] J. Sardegna and T.O. Paul. *The Encyclopedia of Blindness and Vision Impairment.* Facts On File, New York, 1991.

[10] B. Tversky and D. Baratz. Memory for faces: Are caricatures better than photographs. *Memory and Cognition*, 13(1), 1985.

[11] M.E. Ward. The visual system. In G.T. Scholl, editor, *Foundations of Education for the Blind and Visually Handicapped Children and Youth: Theory and Practice.* American Foundation for the Blind, 1986.

Thomas Way was born in Rockville, Maryland, in 1962. He received the B.A. degree in Film and Television Production from the University of Maryland at College Park in 1984, and spent the following nine years in Hollywood, California working in the entertainment industry as a writer, producer, director and radio announcer. He received the M.S. in Computer and Information Sciences from the University of Delaware in 1996, and is currently pursuing Ph.D. studies there. His research interests include human-computer interaction, image processing, compiler optimization, and high performance computing. He is married and has a daughter.

Kenneth Barner was born in Montclair, New Jersey, in 1963. He received a B.S.E.E. degree (*magna cum laude*) from Lehigh University, Bethlehem, Pennsylvania, in 1987. He received the M.S.E.E. and Ph.D. degrees from the University of Delaware, Newark, Delaware, in 1989 and 1990, respectively. Dr. Barner was the duPont Teaching Fellow and a Visiting Lecturer at the University of Delaware in 1991 and 1992, respectively. He is currently a Research Engineer at the Applied Science and Engineering Laboratories of the A. I. duPont Institute/University of Delaware, and an Assistant Research Professor in the Department of Electrical Engineering at the University of Delaware. His research interests include signal and image processing, nonlinear systems, speech enhancement, recognition, and synthesis, and pattern recognition. He is a member of Tau Beta Pi, Eta Kappa Nu, and Phi Sigma Kappa.

TABLE I

SUMMARY OF OVERALL RESULTS OF SIMPLE DISCRIMINATION TASK FOR FOUR IMAGE PROCESSES. THE *Aggregate Process* IS COMPRISED OF BLURRING, SOBEL EDGE DETECTION WITHOUT THRESHOLDING, K-MEANS SEGMENTATION, AND MEDIAN FILTERING, APPLIED IN THAT ORDER.

Image process	*Mean Pct. Matched*	p (vs. No proc.)	p (vs. Chance)
No Processing	50.0%	1.00e+00	1.00e+00
K-means Segmentation	83.0%	2.07e-05	3.49e-07
Sobel Edge Detection	81.0%	4.70e-06	1.53e-09
Aggregate Process	95.0%	4.40e-08	1.93e-11

TABLE II

SUMMARY OF PERCENTAGE OF CORRECT RESPONSES COMPARING EFFECTS OF TWO VARIETIES OF MICROCAPSULE PAPER ON SIMPLE DISCRIMINATION TASK.

Image process	*Flexi-Paper*	*Matsumoto-Kosan*	p
No Processing	48.0%	52.0%	6.41e-01
K-means Segmentation	90.0%	78.0%	2.60e-01
Sobel Edge Detection	80.0%	82.0%	3.05e-01
Aggregate Process	96.0%	94.0%	3.59e-01

TABLE III

SUMMARY OF PERCENTAGE OF CORRECT RESPONSES COMPARING RESULTS OF BLIND VERSUS SIGHTED SUBJECTS PERFORMING SIMPLE DISCRIMINATION TASK.

Image process	*Blind Subjects*	*Sighted Subjects*	p
No Processing	53.3%	52.9%	2.94e-01
K-means Segmentation	83.3%	81.4%	6.01e-01
Sobel Edge Detection	63.3%	84.3%	6.43e-02
Aggregate Process	90.0%	95.7%	7.45e-01

TABLE IV

SUMMARY OF OVERALL RESULTS OF TIMED DISCRIMINATION TASK FOR FOUR IMAGE PROCESSES.

Image process	*Mean Pct. Matched*	p (vs. No proc.)	p (vs. Chance)
No Processing	55.0%	1.00e+00	3.82e-02
K-means Segmentation	77.0%	1.80e-03	1.34e-04
Sobel Edge Detection	73.0%	2.90e-03	1.24e-04
Aggregate Process	87.0%	4.67e-05	3.24e-06

TABLE V

SUMMARY OF PERCENTAGE OF CORRECT RESPONSES COMPARING EFFECTS OF TWO VARIETIES OF MICROCAPSULE PAPER ON TIMED DISCRIMINATION TASK.

Image process	*Flexi-Paper*	*Matsumoto-Kosan*	p
No Processing	50.0%	58.0%	1.33e-02
K-means Segmentation	92.0%	68.0%	2.30e-01
Sobel Edge Detection	74.0%	72.0%	8.47e-01
Aggregate Process	96.0%	82.0%	2.30e-01

TABLE VI

SUMMARY OF PERCENTAGE OF CORRECT RESPONSES COMPARING RESULTS OF BLIND VERSUS SIGHTED SUBJECTS PERFORMING TIMED DISCRIMINATION TASK.

Image process	*Blind*	*Sighted*	p
No Processing	43.3%	60.0%	6.54e-01
K-means Segmentation	86.7%	72.9%	4.91e-01
Sobel Edge Detection	73.3%	72.9%	1.96e-01
Aggregate Process	93.3%	84.3%	7.47e-01

TABLE VII

SUMMARY OF PERCENTAGE OF CORRECT RESPONSES COMPARING RESULTS OF ALL SUBJECTS ON SIMPLE DISCRIMINATION VERSUS TIMED DISCRIMINATION TASKS.

Image process	*Simple*	*Timed*	p
No Processing	50.0%	55.0%	2.85e-01
K-means Segmentation	83.0%	77.0%	4.03e-01
Sobel Edge Detection	81.0%	73.0%	1.61e-01
Aggregate Process	95.0%	87.0%	2.26e-01

TABLE VIII

SUMMARY OF OVERALL RESULTS OF IDENTIFICATION TASK FOR FOUR IMAGE PROCESSES.

Image process	*Pct. Identified*	p (vs. No proc.)	p (vs. Chance)
No Processing	7.0%	1.00e+00	1.83e-06
K-means Segmentation	55.0%	3.84e-09	2.24e-07
Sobel Edge Detection	46.0%	6.36e-07	2.02e-04
Aggregate Process	85.0%	5.05e-13	8.93e-13

TABLE IX

SUMMARY OF PERCENTAGE OF CORRECT RESPONSES COMPARING RESULTS OF BLIND VERSUS SIGHTED SUBJECTS PERFORMING IDENTIFICATION TASK.

Image process	*Blind*	*Sighted*	p
No Processing	10.0%	5.7%	4.83e-01
K-means Segmentation	56.7%	54.3%	7.89e-01
Sobel Edge Detection	43.3%	47.1%	7.23e-01
Aggregate Process	76.7%	88.6%	1.13e-01

TABLE X

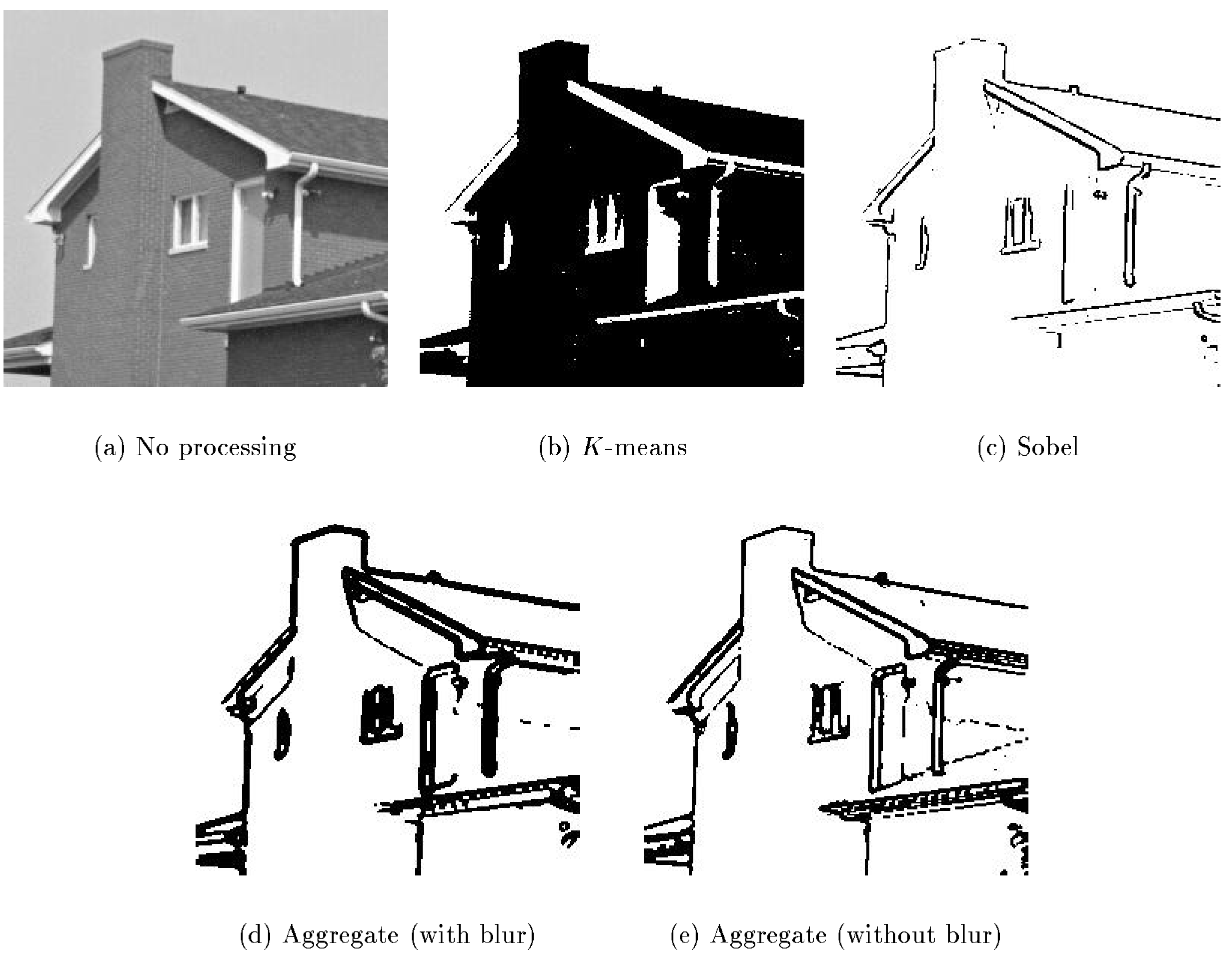
SUMMARY OF RESULTS OF COMPREHENSION TASK FOR THREE SUBTASKS AND OVERALL COMPREHENSIBILITY OF TACTILE IMAGES PREPARED USING AGGREGATE PROCESS.

Comprehension Task	*Correct Replies*	*Pct. Correct*	p
Feature location	108/130	83.1%	6.53e-08
Feature understanding	132/160	82.5%	3.49e-09
Content reasoning	87/110	79.1%	4.45e-14
Overall	327/400	81.8%	1.44e-15

TABLE XI

SUMMARY OF PERCENTAGE OF CORRECT RESPONSES COMPARING RESULTS OF BLIND VERSUS SIGHTED SUBJECTS PERFORMING COMPREHENSION TASK.

Comprehension Task	*Blind*	*Sighted*	p
Feature location	69.2%	89.0%	5.30e-03
Feature understanding	89.6%	79.5%	1.37e-01
Content reasoning	78.8%	79.2%	8.96e-01
Overall	80.0%	82.5%	4.27e-01



(a) No processing (b) K-means (c) Sobel

(d) Aggregate (with blur) (e) Aggregate (without blur)

Fig. 1. Image processes