A Recognizer for Free-Hand Graph Drawings

Hamdi Dibeklioglu
Bogazici University
Dept. of Computer Eng.
Istanbul, Turkey
Hamdi.Dibeklioglu@cmpe.boun.edu.tr

Tevfik Metin Sezgin
University of Cambridge
Computer Laboratory
Cambridge, UK
Metin.Sezgin@cl.cam.ac.uk

Ender Ozcan
Yeditepe University
Dept. of Computer Eng.
Istanbul, Turkey
eozcan@cse.yeditepe.edu.tr

Abstract

Interactive multimedia such as computer simulations and animations received increased attention over the years as supplementary teaching tools and have now become integral components of most engineering and science curriculums. We believe one way to boost the utility of such simulations and animations is to make them easier to use. In this paper, we describe a pen-based interface for constructing weighted and unweighted graphs which functions as a front-end to a shortest path and minimum spanning tree (MST) algorithm simulators that we have developed. We compare the usability of this pen-based interface to that of a WIMP-based interface and a hybrid interface that uses a mixture of pen-based and WIMP-based input methods. We report evaluation results on the usability of the three types of interfaces that we have developed. Our results show that a purely pen-based approach to graph creation may actually suffer from the relatively higher misrecognition rates of the harder-to-recognize symbols and inhibit usability. We conclude that a hybrid approach which combines a soft-keyboard with graph recognition outperforms interfaces that are purely pen-based or WIMP-based.

1. Introduction

Interactive multimedia such as computer simulations and animations received increased attention over the years as supplementary teaching tools and have now become integral components of most engineering and science curriculums.

Although there is a wealth of resources for interactive animations, these applications usually employ traditional WIMP-based interfaces regardless of the appropriateness of such interfaces for the task at hand. For example, a quick web search for shortest path graph algorithms returns over a hundred applets/applications with WIMP-based interfaces despite the fact that drawing a graph – the first step of interaction for all these applications – could naturally be done using a pen-based interface due to its diagrammatic nature.

Pen-based interfaces are not only better suited for this task, but also make it possible to interact with the application in question in a relatively standard diagrammatic language unlike their WIMP counterparts. For example, quick examination of the WIMP-based shortest-path animation applications reveals that they show considerable variation on how the user is expected to construct the graphs. For adding nodes to the graph some applications use buttons with images, text, or both. Some use pull-down menus, some use a drag-and-drop interface while others use a combination of choice menus and buttons. Textual labels associated with the GUI elements also show variation. For example, the words "node/vertex," and "edge," "link," "arc," "line" and "connector" are used interchangeably although they conceptually refer to the same thing and have standard graphical/visual representations.

The above observations give us reasonable support to believe that an intelligent pen-based interface which interprets free-hand pen input would be preferred over traditional WIMP-based interfaces for constructing graphs in our target domain of shortest path graphs. The usefulness of such a system will depend on many factors including the degree to which we can correctly interpret free-hand ink, the perceived ease-of-use for the users, the efficiency of the input method and the overall satisfaction of the users. To

answer these questions, we have built three interfaces for constructing graphs:

1. *a pen-based interface* where the graph structure and weights are entered using free-hand ink
2. *a WIMP-based PowerPoint-like interface* for graph construction and a soft-keyboard for inputting edge weights
3. *a hybrid interface* using pen input for graph construction and a soft-keyboard for inputting edge weights

All of these three interfaces function as front-ends to a shortest path algorithm simulator that we have developed. In the rest of this paper, we describe the architecture of our application including a summary of each front-end interface in detail. We also report evaluation results on the usability of each interface. Our results show that the benefits of a purely pen-based approach to graph creation may actually be undermined by the relatively poor recognition rates of the harder-to-recognize symbols and inhibit usability. We conclude with a discussion of the related and future work.

2 Application Architecture

We have designed our application to have a clear separation between the graph construction module and the algorithm animation module. We use a graph adjacency matrix representation to interface the front-end and the animation modules. This makes it easy to extend the system by adding a new input method or a new animation module.

2.1. Graph Construction Front-Ends

2.1.1 Ink recognition method

The ink-based graph construction method has three components that respectively support graph structure recognition, digit recognition, and editing. The control flow chart in Fig. 1 shows how the graph structure recognition and the editing modules fit together. The digit recognition component is a separate module.

Recognizing the graph structure
As seen in Fig. 1, the graph structure is constructed by recognizing circles and lines (for nodes and edges) and points (for entering the edge-weight editing mode). We only support single stroke objects. So we assume that the users draw nodes and links using single strokes, and also write digits in single strokes.

Node and edge recognition
In our pen-based interface, nodes and edges are recognized by fitting circles and lines to the input stroke, and choosing the interpretation that has a lower least squares error.

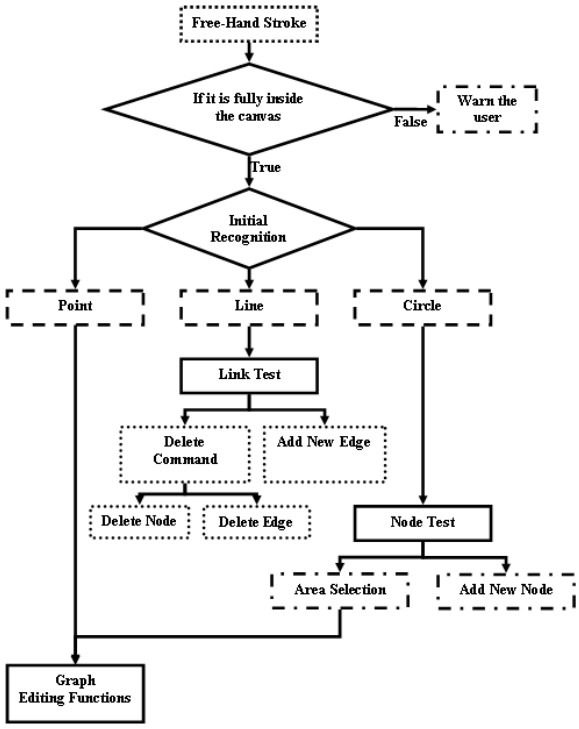


Figure 1. Flowchart for the ink recognition process.

To find a line fit, we use the first and last points in the stroke, (x_1, y_1) and (x_n, y_n), to derive the line equation $ax + by + c = 0$ using the following equation of a line with two known points:

$$(y - y_1)(x_n - x_1) = (y_n - y_1)(x - x_1)$$

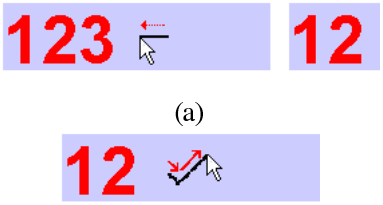
To fit a circle, we find the rectangular bounding box with corner positions (x_{min}, y_{min}) and (x_{max}, y_{max}) for the stroke at hand where $x_{min}, y_{min}, x_{max}$ and y_{max} are the minimum and maximum values for the x and y positions of the points in the stroke. Then the equation $(x - h)^2 + (y - k)^2 = r^2$ gives us the circular fit to the stroke where (h, k) is the center of the circle and r is the radius and

$$h = \frac{x_{min} + x_{max}}{2},$$

$$k = \frac{y_{min} + y_{max}}{2}.$$

After the line and circle fits are computed, we compute the distance of each point (p_x, p_y) in the stroke to the line and circle fits using the following error functions:

$$Error_{line}(p_x, p_y) = \left| \frac{ap_x + bp_y + c}{\sqrt{a^2 + b^2}} \right|,$$



(a)

(b)

Figure 2. Gestures for indicating deletion (a), and the completion of the edge-weight specification.

$$Error_{circle}(p_x, p_y) = \sqrt{(h - p_x)^2 + (k - p_y)^2} - r.$$

We compute the average fitting error for circle and line fits using all points in the stroke and classify it to have the label of the fit with the lowest average error. Circles indicate that the user has drawn a node, and lines indicate an edge. If the end points of the stroke lie on previously drawn nodes, then those nodes become the source and destination nodes. If the user is drawing a directed graph, the direction of the stroke also determines the direction of the edge. If either of the stroke's end-points don't lie inside a node, then the stroke is ignored.

If the recognized shape is a circle that surrounds a set of nodes, then it is interpreted as a selection operation and passed onto the graph editing module.

Handwritten digit recognition

Digit recognition allows the edge weights to be specified by pen input. We have implemented two digit recognition methods. The first one uses Kohonen networks [14]. The other recognition method is based on the iterative closest point [10] and parallel sampling algorithms, and we describe it in detail in [2]. The method based on Kohonen networks has outperformed the latter algorithm and it was our algorithm of choice for the version of the system used in the evaluation.

In addition to digits, we have trained the system to recognize two simple gestures as shown in Fig. 2. A leftward dash deletes the last digit that was recognized, and a checkmark signals that the user is done entering the current weight.

The users can edit their graphs using the pen interface. Nodes are selected by drawing a closed loop around them. Once the nodes are selected, the selection can be dragged around to move the nodes. Drawing a line across the selection deletes the selected nodes and any edges connected to them. Tapping the weight of an edge with the stylus puts the system into edge-weight editing mode where the user can update the weight value using the pen interface.

Figure 3. The soft-keyboard interface used for entering edge weights in the hybrid and WIMP-based interfaces.

2.1.2 Hybrid Method

The hybrid method uses the graph recognition method described above for specifying the graph structure. We use a soft-keyboard for entering the edge weights. Fig. 3 shows the layout of our soft-keyboard. It works much like the soft-keyboard included with the Windows XP Tablet PC version. The user taps in the keys with a stylus or another pointing device to enter the edge weights. Editing in the hybrid method works the same way as in the ink recognition method, except the edge weights are entered using the soft keyboard.

2.1.3 WIMP-based method

Fig. 4 shows our WIMP based interface. The WIMP based method uses image-buttons with textual labels for constructing the graph structure. The image-buttons form a toolbar, and the user selects the operation that needs to be performed by clicking on the appropriate button in the toolbar. This interface is representative of most WIMP-based tools such as MS PowerPoint. As it was the case for the hybrid method, the edge weights are entered using the soft-keyboard interface. The graph editing operations are selected using the toolbar.

2.2 Graph Animation Module

All three interfaces have a "RUN" button that the user can click or tap to launch the graph algorithm simulation module once the graph construction step is completed. Fig. 5 shows a snapshot of the graph animation module simulating Dijkstra's shortest path algorithm on a problem specified by the user. As seen in the figure, the graph is color coded, and there is an English annotation of the algorithm's current status displayed in the bottom of the window. We support all combinations of directed/undirected and weighted/unweighted graphs and have implemented simulation modules for Dijkstra's and Kruskall's shortest path algorithms.

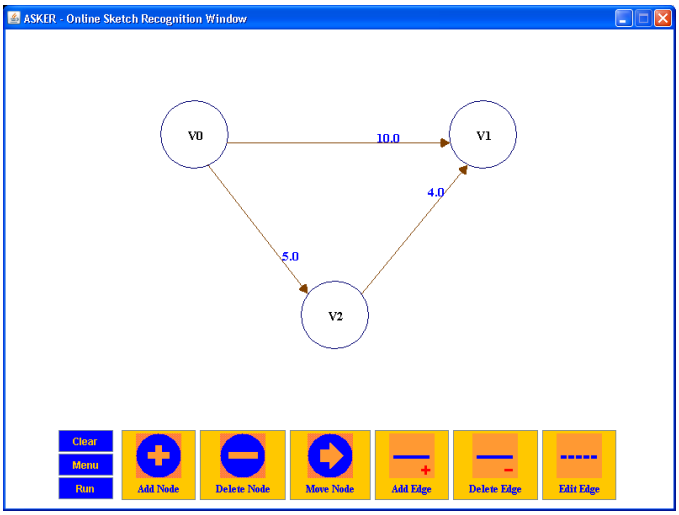


Figure 4. The WIMP-based interface.

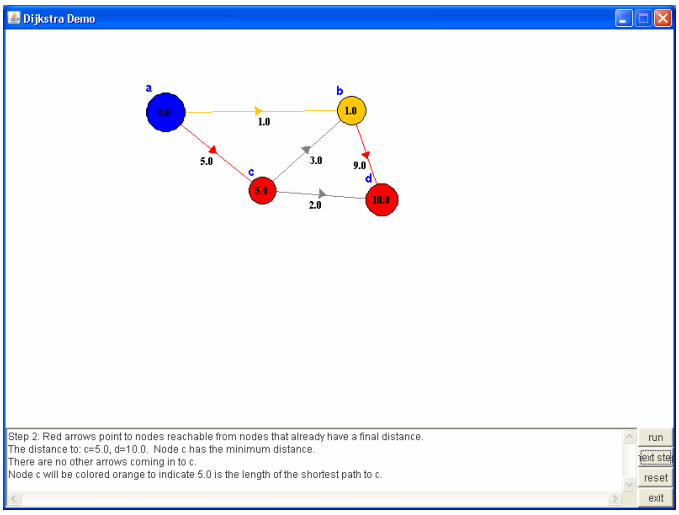


Figure 5. The graph animation module showing the simulation of the Dijkstra's algorithm on a problem specified by the user.

3. Evaluation

To find out how each of the three interfaces compare, we have conducted a user study. Among the goals of the user study were to determine which interface would best suit users who were somewhat knowledgeable in graph theory.

For the user study, we recruited 10 students who had recently taken an undergraduate algorithms course. One subject was left-handed, while the rest were right-handed.

The study consisted of three parts. In part 1, the subjects were shown a weighted directed graph with five nodes and 10 edges. They were asked to:

- construct the graph
- remove three edges
- remove a node and all the edges connected to that node
- edit the weights of three edges
- move two of the nodes

In part 2, the subjects were shown a weighted undirected graph with nine nodes and 13 edges. They were asked to:

- construct the graph
- remove three nodes and all connected edges
- move the nodes around to obtain a specified layout

In part 3, we asked the subjects to draw an undirected graph depicting the relative positions of five cities that the subjects were familiar with. They were then asked to add at least seven edges to their graph labeled by the pairwise distance between the source and destination cities to the best of their knowledge. The goal of part 3 was to have the subjects draw a graph from memory, as opposed to copying one that they were shown.

The subjects were asked to complete all three parts using the pen-based, hybrid and the WIMP-based interfaces in an arbitrary order. For each task in the three parts, we also measured the time spent for the task as well as the number of recognition/input errors and the number of corrections the users had perform. At the end of the study, the subjects have completed a survey.

We have observed that 8 of the participants drew the weighted graphs starting with the vertices first. After drawing all vertices, they filled in the edges. On average, all parts were completed in an hour. Based on the average time spent, the hybrid method ranked first and the pen-based method ranked last.

Fig. 6 lists the questions that the subjects answered at the end of the study. They were asked to assign a score 1-7 for each question, based on how much they agreed with each statement. The scores were defined as: 1 strongly agree, 2 somewhat agree, 3 slightly agree, 4 neural, 5 slightly disagree, 6 disagree, 7 strongly disagree. We applied the Friedman's non-parametric test with a degree of freedom 2 to

	Input Method		
	Pen-based	Hybrid	WIMP
1. The functions I expected to complete the tasks were available	2.40	1.60	2.40
2. The interface was intuitive	2.40	1.50	2.70
3. I was satisfied with how the interface worked	2.40	1.30	2.20
4. The interface was simple to use	2.50	1.30	2.30
5. I could effectively complete the tasks	2.60	1.90	2.60
6. I could complete the tasks quickly	3.10	1.60	2.70
7. I could complete the tasks efficiently	2.90	1.70	2.60
8. I thought the "look" of interface was pleasant	2.20	1.50	2.90
9. Drawing the nodes and the edges was easy	1.80	1.00	2.10
10. Entering the weights was easy	4.90	2.00	2.20
11. I could easily recover from errors	2.40	1.70	1.80
12. I really liked using the GUI.	3.00	1.30	2.80
average score	2.72	1.53	2.44

Figure 6. Aggregated results for the exit survey given to the users. The first column lists a number of statements. The users were asked to assign scores to each interface based on how much they agreed with the statements.

evaluate the consistency of the ratings for the GUIs. The results show that the changes obtained for the statements #6, #7, #10 and #12 were significant within a confidence interval of 95%. In summary, the participants preferred the hybrid interfaces over the other two and they felt it allowed them to complete the tasks more efficiently. The task completion times that we have measured, which are summarized in Fig. 7, agree with this statement. Furthermore, the users found it easier to enter the weights with the hybrid method. Again, the figures summarized in Fig. 7 support this. The table also shows that the pen-based interface had the most number of errors during drawing and weight entry.

Finally, although the ratings for statement #9 in Fig. 6 were not consistent across users, the average rating is better for the hybrid and pen-based interfaces compared to the WIMP-based interface. It seems that with a better digit recognizer, a purely pen-based approach would have been preferred.

4. Related Work

Previous research in interpreting free-hand pen input can be categorized as online methods that interpret ink as it is laid on the drawing surface, and offline systems that interpret previously recorded drawings. Our system uses online recognition.

SILK and DENIM [7] are amongst the earliest tools to demonstrate the use of intelligent processing of free-hand input in enhancing user interfaces. Work in [4, 11, 8, 9, 3, 13, 12, 6, 5] and [1] describe sketch recognition systems for a number of domains including electrical circuit diagrams, UML diagrams, graphs and family trees. Among these systems, some put more emphasis on the usability and user interface issues of their application, while others aim to push forward the state of the art in sketch recognition. From this perspective, our work is closer to the first group of work, because our main contribution is our user study which compares three different user interfaces of varying degrees of intelligent ink processing to achieve the same task. Rather than focusing on developing the best recognition techniques, we have explored the trade-off between the added convenience of doing intelligent ink-processing and the inconvenience created by misrecognitions. We believe we have identified a mixture of pen-based and WIMP-based input methods that significantly outperform either method alone.

5. Discussion and Future Work

We have presented a pen-based interface for constructing weighted and unweighted graphs which functions as a front-end to a shortest path algorithm simulator that we have developed. We compare the usability of this pen-based interface to that of a WIMP-based interface and a hybrid interface that uses a mixture of pen-based and WIMP-based input methods. Our evaluation of the three types of interfaces has revealed that the hybrid interface is preferred by users over only pen-based and only WIMP-based input methods. We believe the main reason for this is the high error rates for digit recognition as indicated by values in Fig. 7. It is very likely that a combination of the relatively low resolution sampling rates for pen-based input and the kinesthetic

	Part 1			Part 2			Part 3		
Input Method	Pen-based	Hybrid	WIMP	Pen-based	Hybrid	WIMP	Pen-based	Hybrid	WIMP
μ	0.8	0.9	0.6	0.4	0.7	0.3	0.3	0.2	0.5
σ	0.63	0.99	0.84	0.70	0.67	0.67	0.67	0.42	0.53

Statistics for the number of errors made during graph structure specification.

	Part 1			Part 2			Part 3		
Input Method	Pen-based	Hybrid	WIMP	Pen-based	Hybrid	WIMP	Pen-based	Hybrid	WIMP
μ	2.90	1.00	1.20	3.40	1.10	1.00	3.00	0.40	0.70
σ	1.29	1.41	1.14	1.51	1.10	1.15	0.94	0.52	1.25

Statistics for the number of errors made during weight specification.

	Part 1			Part 2			Part 3		
Input Method	Pen-based	Hybrid	WIMP	Pen-based	Hybrid	WIMP	Pen-based	Hybrid	WIMP
μ	268.4	217.3	252.7	276.1	191.4	250.1	230.1	140.3	183.1
σ	101.71	56.88	54.70	77.67	40.37	86.07	69.82	54.63	67.23

Statistics for the amount of time spent on each of the three parts in seconds.

Figure 7. Mean (μ) and standard error (σ) for the number of errors made during graph structure specification, weight entry and the amount of time spent for each part by the subjects.

properties of writing on a capturing device as opposed to paper is responsible for the low recognition rates. This suggests that building robust user independent digit recognizers as a critical research problem despite the high recognition rates obtained for scanned images of digits.

References

[1] C. Alvarado. Sketchread: A multi-domain sketch recognition engine. In *Proceedings of ACM Symposium on User Interface Software and Technology*, 2004.

[2] H. Dibeklioglu. A sketch recognizer for graphs. *Engineering Project Report, Yeditepe Univ. Dept. of Computer Eng.*, 2006.

[3] A. S. Forsberg, M. Dieterich, and R. C. Zeleznik. The music notepad. In *ACM Symposium on User Interface Software and Technology*, pages 203–210, 1998.

[4] L. Gennari, L. B. Kara, and T. F. Stahovich. Combining geometry and domain knowledge to interpret hand-drawn diagrams. *AAAI Fall Symposium Series, Making Pen-Based Interaction Intelligent and Natural*, 2004.

[5] M. Gross. Recognizing and interpreting diagrams in design. In *2nd Annual International Conference on Image Processing*, pages 308–311, 1995.

[6] T. Hammond and R. Davis. Tahuti: A geometrical sketch recognition system for uml class diagrams. *Proceedings of the 2002 AAAI Spring Symposium on Sketch Understanding*, March 2002.

[7] M. A. Hearst, M. D. Gross, J. A. Landay, and T. F. Stahovich. Sketching intelligent systems. *IEEE Intelligent Systems*, 13(3):10–19, May-June 1998.

[8] H. Hse and A. R. Newton. Recognition and beautification of multi-stroke symbols. *AAAI Fall Symposium Series, Making Pen-Based Interaction Intelligent and Natural*, 2004.

[9] B. Krishnapuram, C. Bishop, and M. Szummer. Generative bayesian models for shape recognition. *IWFHR '04, Japan*, 2004.

[10] D. Liu and T. Chen. Soft shape context for iterative closest point registration. *IEEE International Conference on Image Processing*, 2004.

[11] E. Saund. Pen-based drawing and editing of node-link diagrams. *Invited Workshop on Pen-Centric Computing*, March 2007.

[12] T. M. Sezgin and R. Davis. Sketch interpretation using multiscale models of temporal patterns. *IEEE Computer Graphics and Applications*, 1(27):28–37, Jan-Feb 2007.

[13] M. Shilman and P. Viola. Spatial recognition and grouping of text and graphics. *Eurographics*, 2004.

[14] B.-J. van der Zwaag. Handwritten digit recognition: A neural network demo. *LNCS 2206*, pages 762 – 771, 2001.