Massachusetts Institute of Technology
Department of Electrical Engineering
and Computer Science

Proposal for Thesis Research in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

TITLE: Combining Sketch Representations for Improved Recognition

SUBMITTED BY: Sonya Cates
32 Vassar St., 32-239
Cambridge, MA 02139

(SIGNATURE OF AUTHOR)

DATE OF SUBMISSION:	September 7, 2007
EXPECTED DATE OF COMPLETION:	August 2008
LABORATORY:	Computer Science and Artificial Intelligence Laboratory

BRIEF STATEMENT OF THE PROBLEM:

Sketching is a common means of conveying, representing, and preserving information, and it has become a subject of research as a method for human-computer interaction, specifically in the area of computer-aided design. Digitally collected sketches contain both spatial and temporal information; additionally, they may contain a conceptual structure of shapes and subshapes. These multiple aspects suggest several ways of representing sketches, each with advantages and disadvantages for recognition. Most existing sketch recognitions systems are based on a single representation and do not use all available information. We propose combining several representations and systems as a way to improve recognition accuracy. Goals of this work are the development of a set of guidelines for using and combining sketch representations and a recognition system based on those guidelines.

Massachusetts Institute of Technology
Department of Electrical Engineering
and Computer Science
Cambridge, Massachusetts 02139

Doctoral Thesis Supervision Agreement

TO: Department Graduate Committee
FROM: Professor Randall Davis

The program outlined in the proposal:

TITLE: Combining Sketch Representations for Improved Recognition
AUTHOR: Sonya Cates
DATE: September 7, 2007

is adequate for a Doctoral thesis. I believe that appropriate readers for this thesis would be:

READER 1: Professor Patrick Winston
READER 2: Professor Pawan Sinha

Facilities and support for the research outlined in the proposal are available. I am willing to supervise the thesis and evaluate the thesis report.

SIGNED: ______________________________
PROFESSOR OF ELECTRICAL
ENGINEERING AND COMPUTER SCIENCE

DATE: ______________________________

Comments:

Massachusetts Institute of Technology
Department of Electrical Engineering
and Computer Science
Cambridge, Massachusetts 02139

Doctoral Thesis Reader Agreement

To: Department Graduate Committee
From: Professor Patrick Winston

The program outlined in the proposal:

Title: Combining Sketch Representations for Improved Recognition
Author: Sonya Cates
Date: September 7, 2007
Supervisor: Professor Randall Davis
Other Reader: Professor Pawan Sinha

is adequate for a Doctoral thesis. I am willing to aid in guiding the research and in evaluating the thesis report as a reader.

Signed: ______________________
Professor of Electrical
Engineering and Computer Science

Date: ______________________

Comments:

____

____

____

____

____

____

Massachusetts Institute of Technology
Department of Electrical Engineering
and Computer Science
Cambridge, Massachusetts 02139

Doctoral Thesis Reader Agreement

To: Department Graduate Committee
From: Professor Pawan Sinha

The program outlined in the proposal:

Title: Combining Sketch Representations for Improved Recognition
Author: Sonya Cates
Date: September 7, 2007
Supervisor: Professor Randall Davis
Other Reader: Professor Patrick Winston

is adequate for a Doctoral thesis. I am willing to aid in guiding the research and in evaluating the thesis report as a reader.

Signed: ______________________________
Associate Professor of Vision
and Computational Neuroscience

Date: ______________________________

Comments:

____

____

____

____

____

____

Contents

Chapter 1

Introduction

Sketching is a critical first step in many design problems; for example, architects and engineers commonly make many rough sketches before selecting and refining a single strategy. Designers often make these initial sketches on paper and must later transfer their work to a computer for further revision. Computer recognition of sketches could streamline this common process, but for many applications, users require an error rate near zero before adopting a new technology, until then preferring to use predictable, if cumbersome methods. Sketch recognition is a relatively new field of study, and while there exists a diverse and advancing body of work, none of the current sketch recognition systems have both the flexibility and accuracy required in a realistic design setting.

The sketches that this work examines are created with a digital pen, which records both position and timing information. These sketches thus have both spatial and temporal aspects; additionally sketches are frequently conceived as a hierarchical structure of shapes and subshapes, for example a square is made up of lines. This multi-faceted quality allows a sketch to be thought of and represented in several ways. In the next section we describe three types of representations: spatial, temporal and conceptual. Most existing sketch recognition systems are based primarily on one of these representations and do not fully use all of the information contained in a sketch.

A goal of this work is to develop guidelines for using and combing all of these representations and systems that are based on them. A second goal is to implement such a combination of several sketch systems in order to achieve greater accuracy based on these guidelines. These objectives are based on our observations that sketch representations and recognition systems have various strengths and weaknesses. These differences often complementary, and we aim to leverage this diversity in order to improve recognition.

Chapter 2

Sketch Representations

We identify three primary aspects of sketches: the spatial, the conceptual, and the temporal. Each aspect provides a way of thinking about a sketch and a means to represent it, and each has advantages and drawbacks for recognition. These positive and negative qualities are often complementary; we propose that sketch recognition can be improved by combining representation and reasoning methods. However, it is not clear how to take full advantage of all of the information contained in these representations or how to combine them.

In the following sections we elaborate on each of these aspects of sketching. We describe corresponding representations, summarize known strengths and weaknesses of each representation, and give rationale for why they are complementary and should be combined. We also relate our taxonomy to existing sketch recognition systems.

2.1 Representation Types

2.1.1 Spatial

By the spatial aspect of a sketch, we mean literally what the sketch looks like: the areas of ink and absence of ink that we see when looking at the sketch on a screen or piece of paper. An obvious spatial representation is simply an array of pixels. This type of representation is appealing both because of its simplicity and because of the large existing body of work in the field of computer vision that uses similar representations.

2.1.2 Conceptual

We define the conceptual view of a sketch as its geometric or symbolic contents and the configuration of the contents within the sketch. A conceptual representation indicates the sketch's geometric primitives, for example line segments and curves, and their spatial relationships, for example locations or whether or not two segments meet. We might alternatively list more complex geometric or symbolic objects as the sketch's contents, such as triangles or resistors. A conceptual representation is

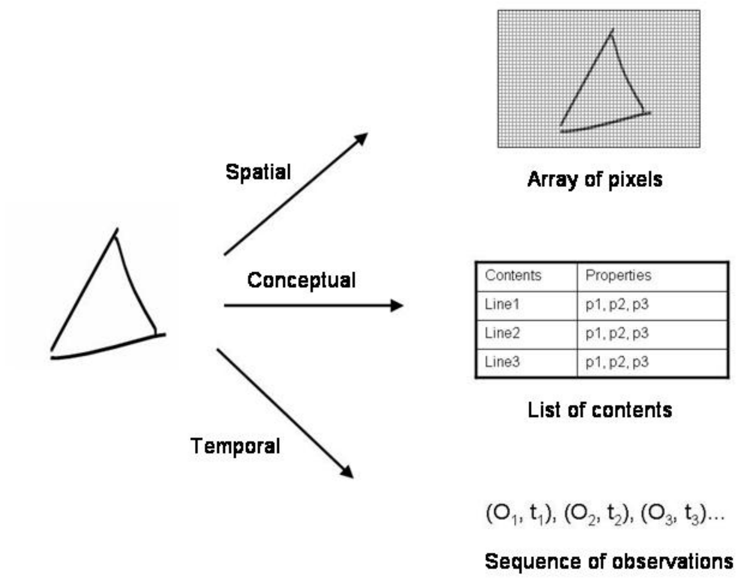


Figure 2-1: The same sketch may be represented in several ways.

attractive because it reflects how the sketch may have been conceived by its author and facilitates high level inferences about what has been drawn.

2.1.3 Temporal

The temporal aspect of a sketch is based on the way the sketch was drawn, including drawing order, pauses, etc. We create the most basic temporal representation with a sequence of time-stamped pen positions. From that higher level abstractions may be created such as time-stamped boolean observations corresponding to whether or not a drawing action occurred, velocity vectors, or other relevant features. The temporal aspect of sketching is an appealing basis for representation because it is a unique quality of online pen-based interaction; without the timing information, we have only a static image, as might be obtained by drawing on paper and scanning the result.

2.2 Comparing Representations

In this section we provide some motivation for the use of multiple representations for sketch recognition. Figure 2-2 presents several simple shapes, the recognition of which is made easier or harder depending on the representation selected.

Figure 2-2(a) contains two squares. These squares might be easily recognized with computer vision pattern matching techniques; however, recognizing these shapes with a conceptual or geometric representation poses a problem. If both squares are drawn with one stroke, the two longest lines must be broken and divided between the two squares. Testing all such possible break points can be time consuming, and it is difficult to define heuristics that are exhaustive and accurate.

Figure 2-2(b) demonstrates the opposite case; here using a spatial representation may make recognition difficult while a conceptual approach is clear. We identify both shapes 1 and 2 as arrows, though they are related by a non-affine transformation. However, a similar non-affine change made to shape 3 results in something that we do not identify as an arrow. Specifying all possible changes to shape 1 that will result in an arrow would be cumbersome with a purely spatial representation. However, recognizing shapes 1 and 2 as arrows without including shape 3 may be done simply with a conceptual definition that specifies an arrow as a line forming the shaft and two lines of equal length forming the head.

The arrows in Figure 2-2(c) present another case where a spatial representation might be cumbersome. In a sketch these arrows are likely to have the same meaning, though they appear different, again through a non-affine transformation. A temporal representation, however could be useful in determining this similarity, as the arrows would likely be drawn with the same temporal pattern. For example, one might consistenly draw an arrow's shaft before drawing its head.

In Figure 2-2(d) the numbers correspond to drawing order, so the two circles were drawn first, followed by the two lines. This interspersing of the parts of different symbols can be problematic for a temporal approach since parts that are relevant to each other are not adjacent temporally. Thus segmenting the sketch, a common first step in recognition, poses a problem if done on purely temporal grounds. However, visually, this interspersing poses no problem since the parts that are relevant to each other are adjacent spatially.

These simple scenarios are representative of common phenomena in hand drawn sketches. Employing an appropriate approach can greatly simplify the problem and improve the accuracy of recognition. However, a domain or a sketch is unlikely to contain only elements that are ideally recognized with a single approach.

2.3 Relationship to Current Sketch Recognition Systems

The classification of sketch representations that we have presented can be extended to group sketch recognition systems as well, according to which representations are used. While most systems do not take purely one approach, many do strongly focus on one of the representations that we have described. Recognition systems presented by Oltmans [19], Sezgin and Davis [27], and Alvarado [2] are each based primarily on one of these representations. Oltmans takes a primarily visual approach to sketch recognition and adapts several computer vision techniques for use with sketches. Sezgin and Davis describe a temporal representation for sketches in which the time-ordered observations include the size and orientation of substrokes. Alvarado uses a conceptual representation in which a domain's symbols are defined hierarchically in terms of more primitive elements.

These systems solve similar problems; all have been tested on circuit diagrams,

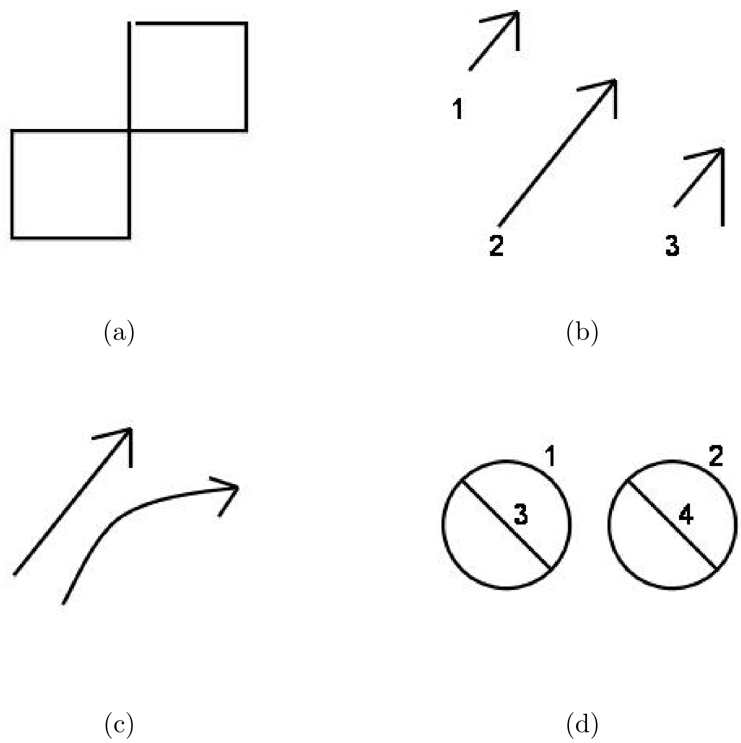


Figure 2-2: Simple shapes whose recognition may be helped or hindered by the choice of representations

each with good results. However, each has limitations that are not shared by all and that parallel the limitations of their underlying representations. For example, Alvarado encounters difficulties when multiple symbols are drawn with one pen stroke. Oltmans addresses this issue specifically with a spatial approach, but has trouble with symbol localization, which is not a problem for the conceptually based system. We suggest that by combining such systems, recognition can more closely approach the human level.

Chapter 3

Combining Representations

In the last chapter we discussed three types of representations for sketches: spatial, conceptual, and temporal. We also provided motivation for combining these representations and recognition systems based on these representations. This chapter elaborates on combinations that we proposed in the last chapter.

3.1 Objectives

The goal of this work is both to develop guidelines for using and combining sketch representations and to apply those guidelines to improve recognition. We identify four main questions that must be answered.

3.1.1 What are the strengths and weakness of a representation?

As we discussed in the previous chapter, different sketch representations have different strengths and weakness. Understanding these differences is necessary first for selecting an appropriate representation to use when addressing a particular problem, but also when attempting to combine representations, as we propose. The previous chapter provides a partial answer to this question through examples and anecdotal evidence; however, a more thorough comparison is necessary to fully answer this question.

3.1.2 How are these strengths related and how might they be combined?

A key premise of this work is the observation that different sketch representations are not equal in the type of recognition that they facilitate. Determining the degree of correlation among the information contained in different sketch representations is an important step, as this relationship should guide combination decisions. Examining the degree to which representations overlap will also provide insight into which errors may be overcome through combination and which errors will require domain knowledge or improvements in lower level processing. While we have provided a partial answer to this question in the previous chapter as well, a more rigorous comparison will be used to guide the development of combination strategies.

3.1.3 Under what circumstances should representations or combination strategies be used?

Research in the field of pen-based computing is varied, and we expect that different applications will require different approaches and may be best served by different representations. Even among recognition applications, the subject of this work, tasks and domains may differ significantly. The third question that we seek to address is how suitable the representations and strategies that we propose are for different tasks, domains, and users. Styles of drawing and recognition requirements may vary substantially between a brainstorming sketch and a final polished version; the nature of the symbols, for example whether they are overlapping or disjoint, may differ amoung domains; novice users draw differently and have different requirements than experts. A single strategy is unlikely to function well in all of these cases.

3.1.4 How can this knowledge of sketch representations be applied to improve recognition?

Answers to the first three questions are not useful unless they can be applied to realistic problems. Thus this work will apply knowledge of the nature of different sketch representations to create a recognition system based on a combination of representations. The goal of such a system would be to improve on previous recognition results and advance current sketch recognition capabilities.

3.2 Constituent Recognizers

It is unclear how to evaluate the usefulness of representations independent of implemented systems, so we will examine recognizers based on the representations that we have described. We propose two approaches, which differ in their constituent recognizers. The first is based on existing sketch recognition systems and the second on novel implementations.

Existing sketch recognition systems, which we described briefly in the previous chapter, perform well on realistic sketches with few restrictions on drawing style, and each is the result of significant development time. However, the systems are also complex and difficult to modify, and so must be treated as partial black boxes, which limits combination possibilities to those that use only the end result of recognition, i.e. a labeled sketch. Additionally, though several systems are based mostly on a single representation, they may also incorporate other information, making it difficult to answer questions about the representations themselves.

By creating new recognition systems we may control the representations used, and thus draw more conclusions about the representations themselves apart from implementations (though the two can not be completely separated in our evaluation scheme). Furthermore, we may access internal information and structures, which will provide more combination possibilities, as intermediate as well as final information may be shared and evaluated. The large drawback to creating new implementations is that they will be less developed and therefore less accurate, less sophisticated, and less realistic than existing systems.

This work will study both types of constituent recognizers. We are interested in sketch representations but also their application, and both approaches are necessary to fully answer the questions posed in the previous section.

3.3 Combination Schemes

Combining experts or classifiers is the subject of much research, described briefly in the next chapter. However, the primary focus of this work is comparing representations rather than comparing combination strategies. Thus we are not looking for the single best combination strategy available, rather we are interested in how combinations strategies relate to sketch representations. We propose three combinations schemes, which we refer to as parallel, multi-stage, and coupled. These strategies reflect some of the diversity in the possible combination strategies and highlight the interesting aspects of sketch representations.

3.3.1 Parallel

A parallel combination of recognizers functions by first applying each recognizer independently. The outcomes, labeled sketches in our case, are then merged by a combination or selection rule. Figure 3-1 depicts a parallel combination scheme. This is a straight forward and common method of combination, and existing recognizers may be used in this way without modification. However, this scheme presents a difficulty when applied to sketch recognition. The output of each recognizer is relatively complex and may consist of many labeled subcomponents, so the combination rule may also be complex. Selecting the one best outcome avoids this problem, but this may not take advantage of different strengths of the constituent recognizers.

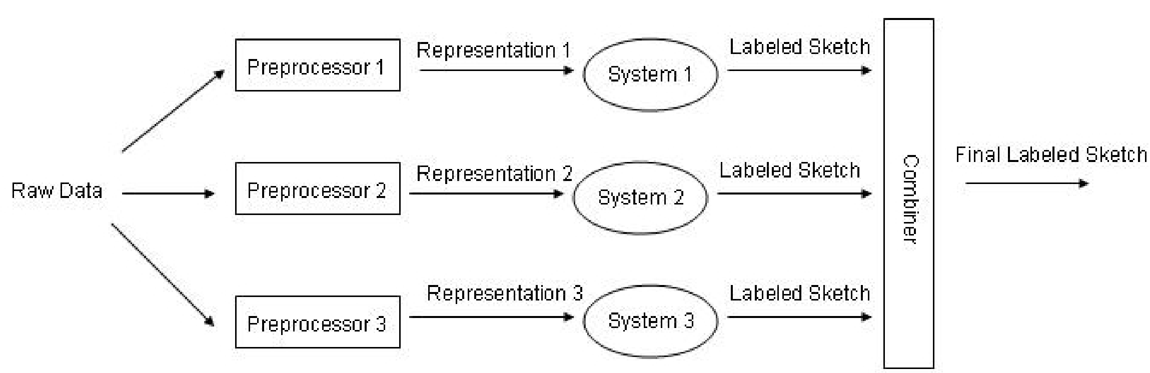


Figure 3-1: Parallel combination scheme

3.3.2 Multi-stage

In a multi-stage system, different representations are used for different parts of the recognition problem (Figure 3-2). This approach is particularly interesting for sketch recognition because the problem may be broken down naturally into several sequential subproblems, for example the identification of low level components and the combination of low level components into higher level structures. We hypothesize that different representations and systems will perform differently on various subtasks. One advantage of this approach is efficiency, since each subproblem is solved only once; however, because each subproblem is solved by only one system, accuracy may be lower than in the parallel case.

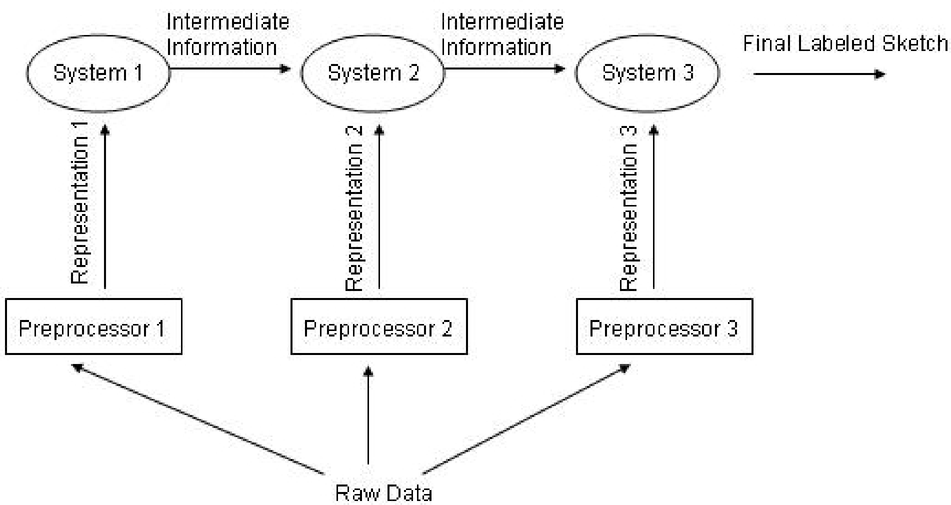
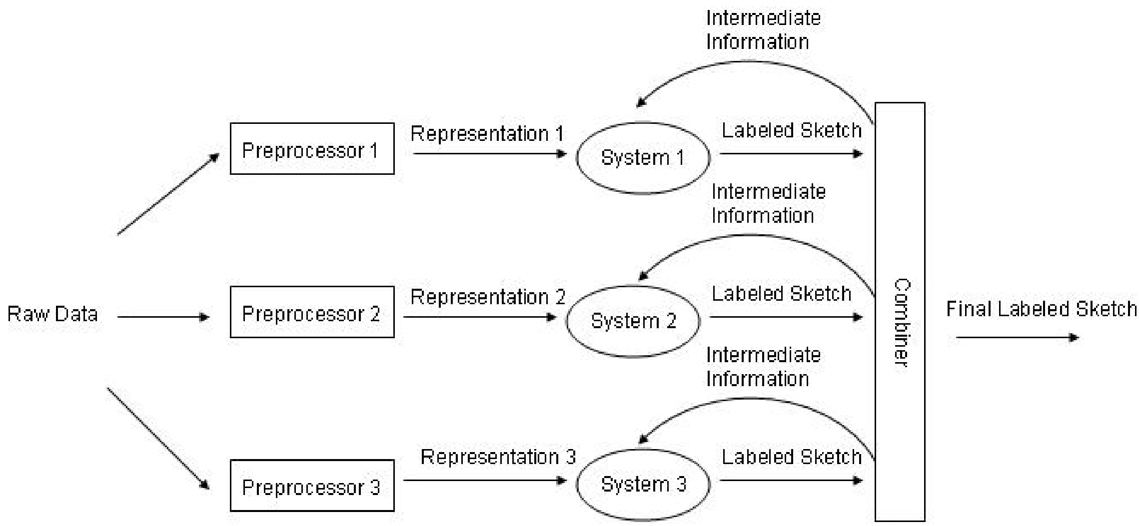


Figure 3-2: Multi-stage combination scheme

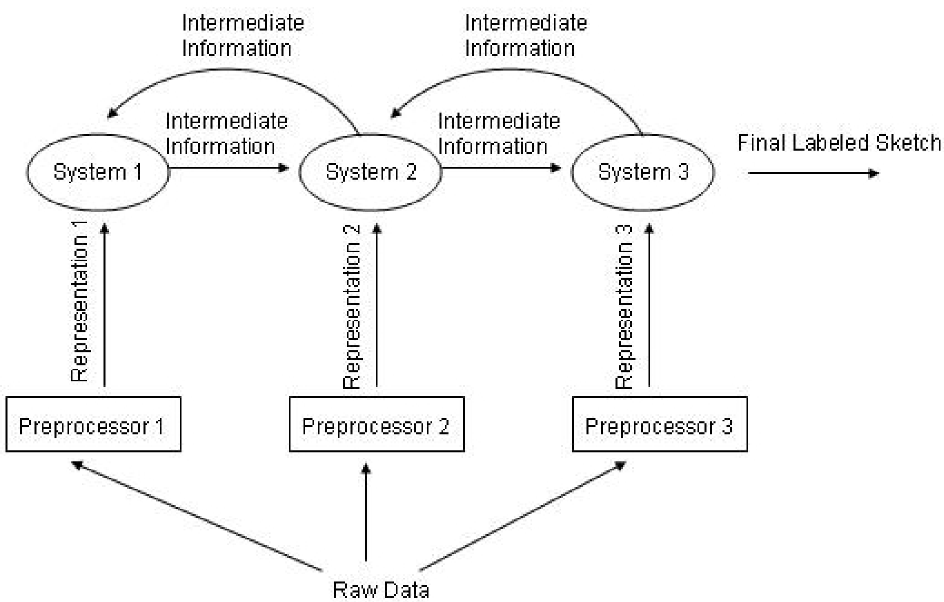
3.3.3 Coupled

A coupled system is one in which the constituent systems may iteratively share intermediate information. This information sharing may be incorporated into either of the above schemes. Figure 3-3 presents two examples of coupled systems. In the first case information is consolidated before being redistributed, and in the second, information is passed only between adjacent stages. Such schemes are particularly relevant for sketch recognition because in many cases, subdecisions affect one another, and improving one subdecision may impact several neighboring decisions and result in a cascading improvement in the final result. For example, changing the interpretation of a stroke from a curve to a line may result in a collection of strokes being reinterpreted as a square, which may then affect the interpretation of a higher level symbol, and if domain knowledge is incorporated, the reinterpretation of a symbol may affect the meaning of a nearby symbol.

The information exchanged between stages can take several forms. It may be partial, for example only information that has high certainty, or it may be complete, for example a fully labeled sketch. The form of the information shared depends on the construction and internal accessibility of the constituent systems. Existing systems are generally not designed to share or accept intermediate information and thus are limited in the way information may be shared.



(a)



(b)

Figure 3-3: Coupled combination schemes

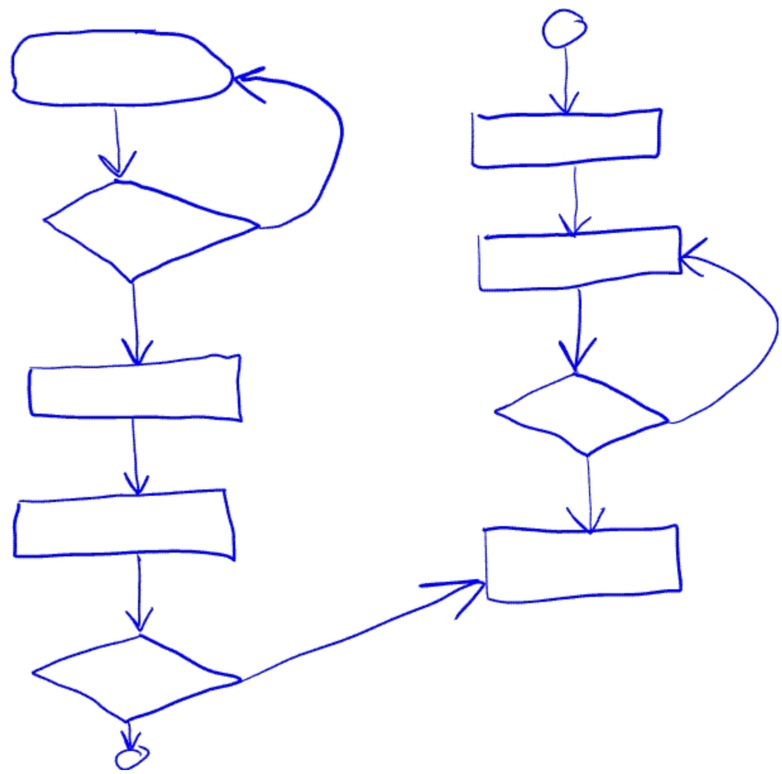
3.4 Sketch Data and Evaluation

Sketches from symbolic domains, for example electrical engineering and flow charts, will be used to develop and test the representations and combination schemes we have described. There is no standard data set of sketches, as in some fields, but several smaller sketch data sets have been collected, including those described in [20], [2], and [5]. The sketches in these data sets were created with digital pens, which collect position, time, and pressure information. Figure 3-4 contains examples from these sketch data sets. We will use some of this previously collected data as well as collect additional sketch data to form training and test sets.

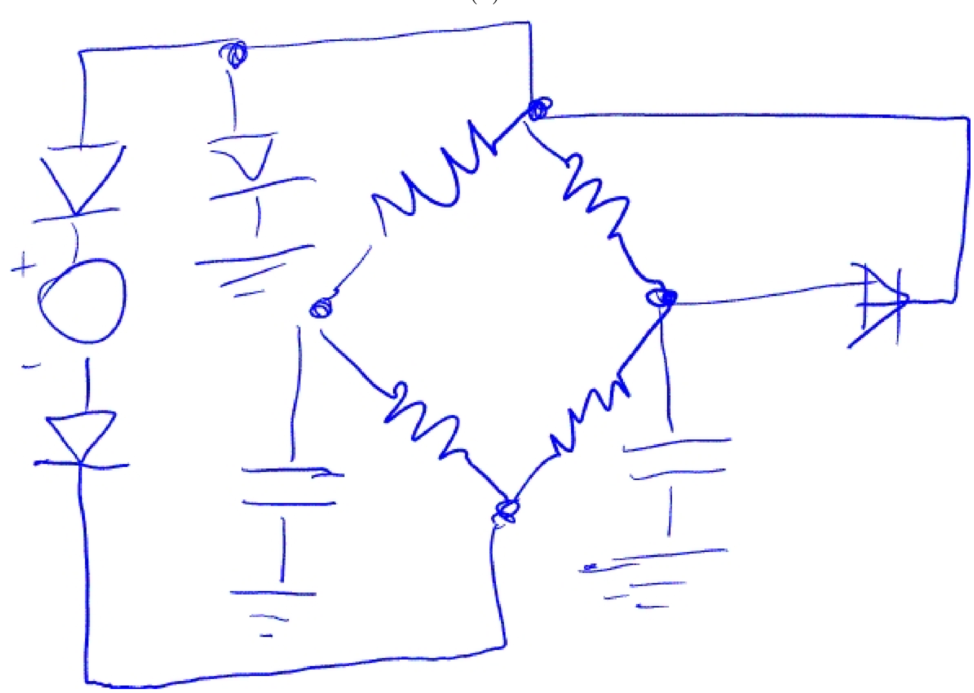
Recognitions systems can be evaluated based on accuracy and on efficiency. These metrics often present a tradeoff. This work will address both; however, the primary focus will be on improving and evaluating accuracy in recognition.

3.5 Preliminary Experiments

Initial combinations of the spatial approach described in [19] and the conceptual approach in [2] demonstrate improvement in recognition accuracy over each constituent system. The output of both recognition systems is a labeled sketch, consisting of the locations of symbols within the sketch, labels for those symbols, and certainty rankings for the labeled symbols. We have tested parallel and multi-stage combination methods. For the parallel combination, the combiner breaks apart each labeled sketch and forms a new labeled sketch by selecting the best group of symbols from the collection of all symbols hypothesized by the constituent recognizers. For the multi-stage approach, we take symbol locations generated by the conceptual system [2] (but not labels) as additional information to be used by the spatial system [19], which then produces a labeled sketch. This combination is based on the observation that the spatial system has high accuracy on isolated symbols but may not locate them precisely in context. Each of these schemes was tested on circuit diagram sketches and found to improve recognition rate; however, we also found that when an oracle was used to select symbols for the parallel combination, recognition could be noticeably improved, suggesting that more distinct information is contained the representations than has been extracted by these relatively simple combination schemes.



(a)



(b)

Figure 3-4: Symbolic sketches

Chapter 4

Contributions

This work contains two main contributions. First, a set of guidelines will be developed for selecting and combining representations for use in sketch recognition. The representations that we propose have all been used for recognition previously; however, they have never been analyzed, compared, or combined. We aim to formalize the intuition and informal observations that determine how and when representations are used and to provide support for the notion of combining multiple representations. The second contribution will be a recognition system based on these guidelines. We aim to improve on the accuracy of existing methods by combining several representations in order to leverage more available information.

Chapter 5

Related Work

5.1 Pen-Based Input Representations

Many applications for digital pen-input have been proposed. These applications place different demands on representations, and to some extent, variation in existing representations reflects the diversity of these applications. However, the variety of representations also reflects the large amount of information available in a sketch; we seek to more fully take advantage of all of this available information.

Handwriting recognition is perhaps the most well studied and commercially successful area of pen-input research. Though more constrained in some ways, handwriting recognition systems share much with sketch recognition approaches, including similarities in representations. Plamondon and Srihari [21] review some of the large body of work on handwriting recognition, and divide the field into online and offline approaches. Offline approaches have only an image available, and thus tend to represent input in visual terms. Senior and Robinson [26] describe a representation which includes a histogram of horizontal ink density. Oltmans [19] takes a similar approach to the more general problem of sketch recognition, in which one may not make assumptions about orientation and size as in handwriting. In this work the visual representation of a sketch includes a bulls-eye shaped histogram of digital ink.

Online handwriting recognition methods deal with input created with digital pens that collect information as a function of time. Hidden Markov Model-based methods are a common approach to handwriting recognition as in [11], which represents handwriting as a time ordered sequence of stroke segments. Sezgin and Davis [27] use a related representation of time ordered substrokes. They apply Dynamic Bayesian Networks, a more generalized version of HMMs to the problem of sketch recognition. This approach handles some properties of unconstrained sketching that are not present in handwriting, such as interspersing of strokes between objects. Alimoglu and Alpaydin [1] combine standard online and offline approaches for the problem of handwritten digit recognition. They find that some of the errors made using each representation alone are uncorrelated and report an improvement in accuracy for the combination.

Systems for recognition with few constraints on what may be drawn are commonly based on conceptual or geometric representations that allow for powerful and often intuitive descriptions. Sezgin et al. [28] describe a method for representing a sketch as a set of geometric primitives, defined as line segments and curves. Work in [29] and [18] describes representations which include similar geometric primitives as well as attributes of primitives and relationships between primitives, such as horizontal, parallel and connecting. These representations facilitate the recognition of complex shapes with little training data. Recognition systems in [3] and [9] are based on similar geometric primitives and also include domain knowledge to improve accuracy. Hammond and Davis [10] formalize the idea of conceptual representation with LADDER, a language for structural description of sketches.

The sketch recognition methods described above impose few constraints on drawing style; however, for many useful applications, unconstrained sketching is unnecessary and may be prone to high error rates. Pen gestures are a constrained type of pen input for specifying objects and commands. Rubine [25] and Long et al. [17] describe methods for recognizing pen gestures, which are typically single strokes with specified directions. A gesture is converted to collection of features that may incorporate spatial and temporal aspects, including speed, size and curviness. Landay and Myers [15] have created SILK, an application for user-interface design based on such gestures.

Some pen-based applications require no recognition at all. Forbus et al. [8] present a digital pen based system for spatial reasoning with maps. Pen input is grouped into glyphs, and no recognition is performed. A representation for this application only requires spatial properties such as size, location, and orientation. Applications related to artistic forms of sketching are another common area of pen-based input research. In these applications sketches are commonly represented by mathematically defined curves. Teddy [13] is a system for designing 3D free form models in which sketched input is converted into polygons. Del Bimbo and Pala [4] describe a method for image retrieval based on a user's sketch, and represent the sketched curves with second order splines.

5.2 Multiple Classifier Systems

Multiple classifier systems are the subject of much research, and it is widely recognized that a combination of classifiers or experts is generally preferred to a single opinion. In spite of a common premise, much of the existing work in multiple classifier systems differs from our problem in two important ways. First, the output of most base classifiers used in combination is relatively simple, for example a class label for a given instance of data. The output of the recognition systems that we are examining is a labeled sketch, which may contain many labeled subcomponents as well as their locations. Second, much of the research in this field is conducted on large, standard data sets; however, sketch data sets are small by comparison, and though there have been attempts to create a standard data set [20], they do not yet compare to data sets in other fields such as speech and handwriting recognition. Despite these differences, several key questions in the field of multiple classifier systems are relevant for us as well.

One question concerns whether the combination mechanism should be fixed, a simple vote for example, or trained, for example a weighted vote with weights related each base classifiers performance on a training set. Duin [7] and Roli et al. [24] examine this question and find that in general a trained approach is preferred but that the answer may vary depending on the circumstances. Duin notes that the preferred strategy is related to the size of the available training set and that a large data set may be required for a trained combination method. Roli et al. find that trained fusion methods are particularly useful in cases where the individual classifiers have different performance and are relatively uncorrelated. Chan and Stolfo [6] examine trained combination approaches further by comparing a combiner, which is trained on the output of the base classifiers, and an arbiter, which is trained on subset of the data to select one of the base classifiers. They find that the combiner generally out performs the arbiter. Kuncheva [14] describes a statistical method for switching between these two approaches; the best single classifier is determined for some regions of the data, and a combination is used where no one classifier excels.

A second relevant area of research in classifier combinations concerns the topology of the combination, which may be parallel, serial, or a combination of the two. Rahman and Fairhurst [22] prefer a hybrid method; however Rahman et al. [23] acknowledge that an ideal combination may require access to internal structures in the base classifiers and describe a serial combination method for constituent classifiers that are black boxes, as many commercial systems are. Alimoglu and Alpaydin [1] compare combination strategies using multiple classifiers based on different representations of handwritten digits, as described in the previous section. They find that serial, multi-stage cascading is the best balance between accuracy and efficiency. In this approach, a simple scheme handles most cases and difficult cases are passed to a more complex classifier. Landgrebe et al. [16] study a different type of multi-stage system. Many recognition systems, including many pen input systems, consist of a detection stage followed by a classification stage. The authors advocate a coupled approach rather than a serial approach for such problems. Also examining the idea of coupling different approaches, Huang et al. [12] use Markov Random Fields to integrate two approaches to image segmentation, one based on a region representation one based on an edge representation. In this way each approach may leverage intermediate information from the other.

Bibliography

[1] Fevzi Alimoglu and Ethem Alpaydin. Combining multiple representations for pen-based handwritten digit recognition. In *Proceedings of the Fourth International Conference on Document Analysis and Recognition*, 1997.

[2] Christine Alvarado. *Multi-Domain Sketch Understanding*. PhD thesis, Massachusetts Institute of Technology, August 2004.

[3] Christine Alvarado and Randall Davis. Dynamically constructed bayes nets for multi-domain sketch understading. In *Proceedings of IJCAI*, pages 1407–1412, August 2005.

[4] Alberto Del Bimbo and Pietro Pala. Visual image retrieval by elastic matching of user sketches. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(2):121–132, February 1997.

[5] Sonya Cates. Using context to resolve ambiguity in sketch understanding. Master's thesis, MIT, 2006.

[6] Philip K. Chan and Salvatore J. Stolfo. Experiments on multistrategy learning by meta-learning. In *Proceedings of Second International Conference on Information and Knowledge Management*, pages 314–323, November 1993.

[7] Robert P.W. Duin. The combining classifier: to train or not to train? In *Proceedings of the 16th International Conference on Pattern Recognition*, pages 765–770, Quebec, Canada, August 2002.

[8] Kenneth D. Forbus, Jeffrey Usher, and Vernell Chapman. Qualitative spatial reasoning about sketch maps. In *Proceedings of the Fifteenth Annual Conference on Innovative Applications of Artificial Intelligence*, pages 85–92, 2003.

[9] Leslie Gennari, Levent Burak Kara, Thomas F. Stahovich, and Kenji Shimada. Combining geometry and domain knowledge to interpret hand-drawn diagrams. *Computers and Graphics*, 29(4):547–562, 2005.

[10] Tracy Hammond and Randall Davis. LADDER, a sketching language for user interface developers. *Elsevier, Computers and Graphics*, 28:518–532, 2005.

[11] Jianying Hu, Michael K. Brown, and William Turin. HMM based on-line handwriting recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:1039–1045, October 1996.

[12] Rui Huang, Vladimir Pavlovic, and Dimitris N. Metaxas. A graphical model framework for coupling MRFs and deformable models. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 739–746, 2004.

[13] Takeo Igarashi, Satoshi Matsuoka, and Hidehiko Tanaka. Teddy: A sketching interface for 3D freeform design. *SIGGRAPH 99*, pages 409–416, August 1999.

[14] Ludmila I. Kuncheva. Switching between selection and fusion in combining classifiers: An experiment. *IEEE Transactions on Systems, Man, and Cybernetics*, 32(2):146–156, April 2002.

[15] James A. Landay and Brad A. Myers. Sketching interfaces: Toward more human interface design. *IEEE Computer*, 34(3):56–64, March 2001.

[16] Thomas Landgrebe, Pavel Paclik, and David M.J. Tax. Optimising two-stage recognition systems. In *Multiple Classifier Systems*, pages 206–215, 2005.

[17] A. Chris Long, Jr., James A. Landay, Lawrence A. Rowe, and Joseph Michiels. Visual similarities of pen gestures. In *Proceedings of the CHI 2000 conference on Human factors in computing systems*, 2000.

[18] Andrew Lovett, Morteza Dehghani, and Kenneth Forbus. Efficient learning of qualitative descriptions for sketch recognition. In *Proceedings of the 20th International Qualitative Reasoning Workshop*, Hanover, NH, July 2006.

[19] Michael Oltmans. *Envisioning Sketch Recognition: A Local Feature Based Approach to Recognizing Informal Sketches*. PhD thesis, Massachusetts Institute of Technology, May 2007.

[20] Michael Oltmans, Christine Alvarado, and Randall Davis. Etcha sketches: Lessons learned from collecting sketch data. In *Making Pen-Based Interaction Intelligent and Natural*, pages 134–140, Menlo Park, California, October 21-24 2004. AAAI Fall Symposium.

[21] Rejean Plamondon and Sargur N. Srihari. On-line and off-line handwriting recognition: A comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:63–84, January 2000.

[22] A.F.R Rahman and M.C Fairhurst. A study of some multi-expert recognition strategies for industrial applications:issures of processing speed and implementability. In *Vision Interface*, pages 569–574, May 1999.

[23] Faud Rahman, Yuliya Tarnikova, Aman Kumar, and Hassan Alam. Second guessing a commercial 'black box' classifier by an 'in house' classifier: Serial classifier combination in a speech recognition application. In *Multiple Classifier Systems*, pages 374–383, 2004.

[24] Fabio Roli, Josef Kittler, Giorgia Fumera, and Daniele Muntoni. An experimental comparison of classifier fusion rules for multimodal personal identity verification systems. In *Multiple Classifier Systems*, pages 325–335, Cagliari, Italy, June 2002.

[25] Dean Rubine. Specifying gestures by example. In *Computer Graphics*, volume 25(4), pages 329–337, 1991.

[26] Andrew W. Senior and Anthony J. Robinson. An off-line cursive handwriting recognition system. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:309–321, March 1998.

[27] Tevfik Metin Sezgin and Randall Davis. Sketch interpretation using multiscale models of temporal patterns. *IEEE Computer Graphics and Applications*, pages 28–37, January/February 2007.

[28] Tevfik Metin Sezgin, Thomas Stahovich, and Randall Davis. Sketch based interfaces: Early processing for sketch understanding. In *The Proceedings of 2001 Perceptive User Interfaces Workshop (PUI'01)*, Orlando, FL, November 2001.

[29] Olya Veselova and Randall Davis. Perceptually based learning of shape descriptions. *Proceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI-04)*, pages 482–487, 2004.

Appendix A

Schedule

1. Combination of existing recognition systems, 2 months
2. Implementation of additional constituent recognizers, 3 months
   (a) Spatial
   (b) Temporal
   (c) Conceptual
3. Combination of new recognition systems, 3 months
   (a) Parallel
   (b) Multi-stage
   (c) Coupled
4. Additional data collection, 1 month
5. Evaluation, 2 months
6. Writing, 3 months