1. Introduction

1.1 An Electronic Scrapbook

The Electronic Scrapbook is a computer system designed to help people to use home video as a creative medium. Segments of video can be placed on "pages" of an electronic scrapbook, just as photographs are placed on pages of a paper scrapbook. Segments of video can be played back individually, or whole pages can be played in sequence to form a story. The system helps the user to describe video and arrange it into stories.

There are two sets of motivations for this work, practical and theoretical. On the practical side, there are millions of people who shoot home video but do not have the time, equipment, or skills to edit their work. Once removed from the camcorder, tapes gather dust on shelves and are viewed rarely if at all. Breaking the barriers between consumers and video editing is largely an issue of interface design. Consumers need a system which is easy to understand, fun to use, and designed to help them gradually learn more about the medium.

While The Electronic Scrapbook can be viewed as an editing system, to do so misses an important point: the interactive, electronic format has desirable features. A scrapbook is not merely a means to generate linear videos -- it is a new media form which is an end in itself.

From a theoretical perspective, this work is motivated by issues of knowledge representation. Computers can not yet look into a video signal and identify a tree or analyze an audio track and recognize the sound of wind in the leaves. These are hard problems and their solution is not imminent. Furthermore, even if it were possible to identify a "tree," the computer still needs a way to represent this information. If computers and video are to work together successfully, then computers will need a way to represent information about motion pictures and audio. The Electronic Scrapbook uses semantic knowledge representation to describe a database of home video. This thesis analyzes the benefits and shortcomings of this approach and raises issues for future research.

Lastly, the Electronic Scrapbook explores what a computational model of a home video "story" might be. How can a computer construct a story? How do people construct stories? The Electronic Scrapbook contains some simple "story models" which begin to explore these questions.

The specific subject matter, home video, was selected to highlight these theoretical issues. There are certain archetypical stories which occur in many

people's lives -- vacations, weddings, and children are common subjects. It is hypothesized that different people might benefit from using the same set of story models to help them to edit their home videos. In computer terms, this is a knowledge-based approach. The computer can use information about the lives of typical camcorder owners to anticipate user's needs. The regularities inherent in the subject matter make this approach practical.

1.2 A Sample Interaction

Here is a fictional narrative of a stereotypical American family using The Electronic Scrapbook:

It's Monday evening at 9 pm. The week with the kids at the beach was fun, and Susan is just adjusting to being back in her normal routine. Now that Kevin and Tim are asleep, she decides to see how the video of the trip turned out.

She plugs her camcorder into her personal computer and starts up The Electronic Scrapbook. She creates a new page in her scrapbook for the trip and calls it "Our Trip to Asparagus Beach." The first clip she decides is interesting is one of the family piling out of the car at the beach. Kevin, who is seven, is wearing his scuba mask and flippers on the asphalt in the parking lot. She digitizes[1] this video clip and describes it, marking the date, place, people present, and the like. She places this clip on the top of the page.

The next clip that looks interesting is one of Tim learning to swim. Tim, who is five, beams proudly at the camera from inside his dragon-headed safety float. A short, lapping wave startles him, and he grabs frantically at his father Jim's swim trunks. Jim lifts Tim out of the water and grabs the dragon with his other hand. They smile at the camera. Susan marks this segment as the events "learning-to-swim" and "trip-to-the-beach" and clicks on the "guess button." A window of similar segments pops up which includes the segment from the parking lot. Susan picks up that segment and drops it down on the new video segment, which causes the system to copy that description. The date and place are correct, so she doesn't need to enter them again. The only difference is that Kevin does not appear in this segment. She removes Kevin's name from the list of people present and clicks on "OK." Describing the segments she chooses takes less time as she proceeds, because she can use these guessed descriptions.

[1]The system currently uses a laser disc to store and play back video. It is in the process of being converted to a digital-video database like the one described in this section. See chapter 3 for more details.

Susan fills a page with segments from the trip, and she and Jim watch the results. The segment of Tim learning to swim reminds Jim of a segment of Kevin from a couple years earlier. He decides to run the story model "important-firsts." The system creates a new page in the scrapbook and fills it with video segments. It alternates shots of the two brothers, beginning with each boy being born, learning to crawl, eating his first solid food, and the like. At the end it shows a shot of Kevin learning to swim from two summers ago, followed by the shot of Tim and the dragon float from the previous week.

Watching this story, Susan decides she doesn't like the segment of Kevin eating his first solid food -- it's boring. She drags that segment to the trash. When she clicks on the "discards" button, the system pops up a window of segments which were candidates for inclusion in this story, but were not used. She finds one of Kevin eating baby food in which he manages smear it all over his face and throw some at the lens of the camera. This shot is much more typical of Kevin, she decides, and puts it into the story. She saves her changes to the scrapbook and shuts down the computer. Tomorrow evening she'll show the video to the boys and let them make a few new pages in the scrapbook with their favorite shots.

1.3 Overview of This Document

A thesis is, regrettably, constrained to be a linear document. This text is organized in a simple progression which follows the development of the system. The reader may wish to jump from the introductory matter in chapters one and two directly to chapter eight, which describes the story models or "knowledge-based templates." Story models are designed to help the user to construct simple stories.

The first two chapters of this thesis introduce the problems being addressed and describe the system developed. Chapter two also introduces relevant background work in artificial intelligence, learning theory, and visual interface design.

Chapters three through eight describe the system in detail. Chapter three explains the system's basic design and details its hardware and software components. Chapter four narrates how I chose my video source material. The specific material used influenced the system's design. Chapters five contains a detailed description of the software and explains design decisions made.

Chapter six describes in detail the underlying knowledge representation developed for the system. Different lessons were learned in developing each of its areas -- people, people in audio only, locations, events, actions, significant objects, date, camera properties, and subjective properties.

Chapter seven describes how inheritance is used to save the user work in logging footage. This is accomplished through use of a "guess button." The guess feature is designed to make the system's behavior easily understandable, making the concept of inheritance accessible to a naive user.

As mentioned above, chapter eight introduces story models . The results of sample story models are compared and evaluated.

Chapter nine places the work in the context of other work done in artificial intelligence. This knowledge-based approach was inspired by research in case-based reasoning and large knowledge-based systems like the Cyc project. Chapter ten introduces a variety of topics for future research. Lastly, chapter eleven draws conclusions about interface design and knowledge representation.