MAN-COMPUTER PROCEDURES FOR IMAGE RETRIEVAL

PRELIMINARY FINDINGS

PAUL S. LICKER
BELL-NORTHERN RESEARCH, OTTAWA

ABSTRACT

The research described in this paper is intended to uncover design parameters critical for user interaction with a data base consisting of digitally - encoded images. Experiments have been performed to establish and standardize tasks for users of image retrieval systems. Measures of the performance of these tasks are used to compare the usability and usefulness of a set of features and facilities for procedures of image retrieval. A mini-computer-based image retrieval system to be used in further experimental research is described. This system is designed to enable performance of the set of standardized retrieval tasks previously mentioned as well as to provide for automatic subject data collection during experiments.

PROCÉDÉS HOMME-ORDINATEUR POUR DOCUMENTATION D'IMAGES

CONSTATATIONS PRÉLIMINAIRES

PAUL S. LICKER
RECHERCHES BELL-NORTHERN, OTTAWA

ABRÉGÉ

Cet article décrit la recherche destinée à découvrir les paramètres de conception critiques à l'action réciproque d'un usager avec un fichier central constitué d'images encodées numériquement. On a fait des expériences pour créer et normaliser les tâches d'usagers de systèmes de documentation d'images. Les mesures de rendement de ces tâches sont employées pour comparer les possibilités d'utilisation et l'utilité d'un ensemble de dispositifs et d'équipements pour des procédés de documentation d'images. Un système de documentation d'images centralisée dans un mini-rodinateur qu'on utilisera pour de la recherche expérimentale complémentaire est également décrit. Ce système est conçu pour permettre l'exécution de l'ensemble de tâches de documentation mentionnées plus haut et pour fournir une collection automatique de données traitées durant les expériences.

MAN-COMPUTER PROCEDURES FOR IMAGE RETRIEVAL PRELIMINARY FINDINGS

Information retrieval sysems for images rather than text, pose unique human factors design questions. Existing systems are intended primarily for trained or highly task-motivated individuals, yet there is no baseline data for comparing the usefulness of different software features for this group, let alone data for unspecialized users of image retrieval systems. In fact, the behaviour of any human retriever is less well-understood than assumed, often on the basis of such behaviour-irrelevant factors as job description, computer system facilities or existing manual procedures.

The research reported on here is a systemmatic attempt to find out which factors are critical for the kinds of tasks that might be encountered by non-specialized individuals solving a general class of problems involving retrieval of images by computer.

Image composition per se is not of interest in this research. We are interested in this question: which features of the software are likely to increase the usability of a basic image retrieval system for which tasks and for which users? If the features do not facilitate, why not?

The salience of the tasks to problem solutions and the pervasiveness and necessity of these problem solutions in real business envioronments will of course bear heavily on the ultimate value of the features to marketable image-retrieval systems.

This paper will discuss the needs for a task taxonomy for image retrieval and present examples. Based on this taxonomy a number of tasks have been selected for study in a laboratory. The study of these abstracted tasks is explained and some preliminary findings are presented. These standardized tasks, with baseline performance data, will then be transferred to a computer-based image retrieval system now under development. The facilitation in performance of these standardized tasks by certain software features can be measured and statements concerning the effectiveness of these facilities made.

IMAGE RETRIEVAL PROBLEMS

A study by interviews of a number of business users of images in Ottawa and a Yellow Pages compilation of potential users of image retrieval turned up a wide range of problems that are amenable to treatment on computer systems. Basically they fall into these eight classes: 1. Acquisition: ordering, creation, examination, judgment; 2. Cataloguing and filing; 3. Retrieval by intrinsic (content) and extrinsic (identifiers) labels; 4. Maintenance of files and indices; 5. Editing for presentation, of sequences; 6. Formatting of individual images; 7. Presentation, display, distribution; 8. Non-pictorial adjuncts: captioning, layout, binding, copying, physical handling of originals.

The Features

The kinds of features we are looking at relative to problem-solving of the kind outlined above are the following: 1. Word indexing and retrieval; 2. Manual indexing and retrieval by pointing, light pens, etc.; 3. Split screen display of many images; 4. Whole data-base scanning techniques; 5. Iconic indexing (using images as indices); 6. Subsetting a grouping; 7. Textual adjunct (captions, content analysis of text); 8. Store and forward communication vs. realtime; 9. Decay or growth of data structures based on time and usage.

Task Taxonomy For Visual Information Retrieval

Although the ultimate description of 'what people do' during visual information retrieval is heavily determined by the categories of classification adopted, a taxonomy is a good first approximation if it leads to useful classification, meaning that descriptions in terms of the classification provided by the taxonomy, have some validity over a wide range of activities.

The taxonomy is based upon a multilevel analysis of the term 'doing'. On one level, we can look at the problem-solving activities individuals engage in while retrieving information of a visual nature to satisfy some need. For instance, LOOKING FOR A MOVIE in the newspaper is a problem-solving activity for satisfying the need to locate a movie to attend. CHOOSING THE MOST ATTRACTIVE LAYOUT for creating a magazine article.

On another level, each problem-solving activity consists of the performance of a number of tasks. These tasks relate both to the user's needs (the reason why he is engaged in problem-solving behavior in the first place) and the medium which is available to him (a visual information retrieval system). A task can be thought of as the pursuit of one goal and a chain of such tasks comprise the problem-solving activity.

The completion of a task does not correspond to the satisfaction of the user's original need; that need is satisfied only through a chain of task completions. For instance, in the problem-solving activity of looking for a movie, there are tasks such as scanning all the entries (for a newspaper these would be ads; on a computer system they would be frames depicting the movies), remembering the interesting ones, rescanning the interesting ones, and picking the 'best' one.

Some of these tasks are easy on some media and more difficult on others. In the newspaper, one can see all the ads at once (at least in Ottawa this is possible). Scanning is rapid, and can be performed many times in a short period. On a computer system in which pictures can be viewed only one-at-a-time this task becomes more difficult and tedious ; are some tasks are unperformable on some systems?

At the third level, a task is seen composed of a series of physical interactions of human with medium, or in the case under discussion, a retriever with a visual information retrieval system. Each interaction consists of a single act performed by the retriever utilizing one or more functions of the medium. Task performance is then seen as the simultaneous systematic performance by a retriever of a series of acts and systematic invocation of corresponding functions of the computer system.

In the layout problem-solving activity, the task of scanning for attractive layouts involves successive acts of requesting the next picture, noting whether or not it is attractive, and when it is attractive, making a notation for later reference. Let us call these three acts 'Request next', 'Note' and 'Name for reference'.

To each act there correspond information system functions necessary for the performance of these acts. Some acts are obviously unperformable on some systems. For instance, one cannot view two pictures simultaneously in a video-output system without some sort of special effects (split-screen is one technique).

In practice, each act will have a corresponding set of functions specified in the taxonomy. For the above three acts, the functions necessary to effect them are sequential display and sequential structuring (for 'request next'), description input and lacation saving (for 'note attractive') and labeling, index storage and dictionary entry (for 'name for reference').

The benefits of a well-designed task taxonomy are six in number: 1. System comparison in terms of task facilitation; 2. Performance norms for problem-solving by untrained individuals; 3. Ideas of the training that will become necessary; 4. New tasks from the set of acts; 5. A taxonomy for problem solving based on tasks; 6. New structures of tasks and procedures for prpblem-solving.

The Tasks

Thus far, the following tasks have been abstracted; we are in the process of standardizing them for manual systems. 1. Choosing by pictorial criterion (search for match); 2. Composition based on descriptions; 3. Replication of viewing order (memory task); 4. Placing in chronological order; 5. Standard IQ tests with images; 6. Process completion (criterion is "what comes next?"); 7. Data description for filing and communication; 8. Finding exemplars in sets of images; 9. Data hiding and uncovering (game for two subjects); 10. Division of data into classes.

The Measures

We have chosen to measure all three levels of problem-solving. Physical activity and verbal behaviour in act performance can be measured by items like viewing time per frame, number of frames viewed, protocol of features used, time to completion, request rate. Task performance is measured by cognitive measures such as memory tests, examination of the types of errors, confidence in performance, communication value of results. And finally, affective variables such as believability of materials, applicability of tasks to real problems, frustration and preference reflect potential rewards and drawbacks in whole-problem solving.

Preliminary Results of Experimentation

Some measures for performance of a memory task have been made. This task involves three subtasks: 1. Viewing a data base of images in sequence (slides); 2. Remembering the items in order of viewing; 3. Placing prints of the images in the viewed order.

Analysis of the data from three experiments shows that memory for order is more related to memory for pairs of neighbors than for such factors as absolute position. This finding and the theoretical development necessary before beginning the experiment illustrates the paucity of good data on mediated-viewing memory tasks. In addition, the research has pointed out that a memory/sorting task

will require some comparison mechanisms, quick, nonverbal storage and recall and rapid serial display procedures. Directly or indirectly, these attempts to derive baseline performance data for non-computer systems will lead to a set of design features for testing on a computer-driven system.

The Computer Facility

A computer facility is being constructed for testing some of the features mentioned. This system, operated by the University of Quebec, is built around a PDP-11/45 with a 1/4 million byte directly addressable digital frame memory and disk drives for 4-second access to any of 150 images as well as A-D and D-A converters for image transmission. The software developed specially for this experimentation includes a flexible facilities manager for defining experimental tasks by the experimenter, a file manager, built around a key-word retrieval system, and an interaction manager for communicating with the subject and recording his communication (ie., his acts). Utimately the hardware will accommodate 2400 frames of full-colour image at a 30 fp display rate.

It should be stressed that this is not a graphics system although the hardware could be used in that fashion but a system for observing storage and retrieval behaviour of subjects in experiments which setup differing tasks and provide differing storage and retrieval features.