ARMY RESEARCH LABORATORY

Advanced Displays and Interactive Displays Report Compendium II

Paul Rose, editor

ARL-SR-95

July 2000

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

Abstract

This report contains citations and abstracts of papers and presentations produced during the fourth year of the U.S. Army's Federated Laboratory (FedLab) program. The program was formed in 1996 to establish partnerships among the Army, industry, and academic research communities. The Advanced Displays and Interactive Displays consortium seeks to provide innovative, cost-effective solutions to information access, understanding, and management for the soldier of the future.

The research encompasses a range of topics. Some work deals with the representation of uncertainty and imprecision in databases, or with the representation of relationships in multimedia databases, in ways that are compatible with human cognitive-processing capabilities. Other work adopts the means of human communication (such as speech, gesture, eye gaze, and lip-reading) for human-computer interaction. Additional work explores methods for incorporating information in virtual-reality displays that support decision making without distracting or overwhelming the soldier. Although diverse, the research is linked by its overriding goal: the presentation of information in a form that allows effective human understanding and decision making in complex battlefield situations.

Introduction

The U.S. Army's five-year Federated Laboratory (FedLab) program was created in 1996 to establish partnerships among the Army, industry, and academic research communities. Three consortia make up the FedLab program: the Advanced Sensors Consortium, the Advanced Telecommunications and Information Distribution Consortium, and the Advanced Displays and Interactive Displays Consortium. The Displays consortium focuses on cognitive- and perception-related aspects of human-computer interaction, seeking to provide innovative, cost-effective solutions to information access, understanding, and management for the soldier of the future. The partners of the Displays Consortium are led by Rockwell Science Center (RSC), a company with a wide-ranging experience in designing and developing displays for military and commercial aircraft, in integrating complex systems, and in managing complex R&D programs. Academic institutions associated with the consortium include the University of Illinois at Urbana-Champaign (UIUC) and North Carolina Agricultural & Technical State University (NC A&T). Much of the work at the University of Illinois is conducted by researchers affiliated with the Beckman Institute for Advanced Science and Technology, known for its extensive program in human-computer intelligent interaction, and the National Center for Supercomputer Applications (NCSA), an institution focused on information visualization questions. Other industrial partners include SYTRONICS, Inc., a small business located in Dayton, Ohio, possessing a strong background in human factors research, and MCNC (now known solely by its initials, but formerly as Microelectronics Center of North Carolina), a private, nonprofit corporation that provides advanced resources in electronic and information technologies to support education and industry, and to enhance technology-based economic development in North Carolina.

This report contains citations and abstracts of papers and presentations produced by researchers affiliated with the Displays consortium, from late 1998 through 1999, which is generally the period since the third FedLab symposium. The research encompasses a range of topics. Some work deals with the representation of uncertainty and imprecision in databases, or with the representation of relationships in multimedia databases, in ways that are compatible with human cognitive-processing capabilities. Other work adopts the means of human communication (such as speech, gesture, eye gaze, and lip-reading) for human-computer interaction. Additional work explores methods for incorporating information in virtual-reality displays that support decision making without distracting or overwhelming the soldier. Although diverse, the research is linked by its overriding goal: the presentation of information in a form that allows effective human understanding and decision making in complex battlefield situations.

For those papers that lacked them, abstracts were supplied by the Army Research Laboratory (ARL).

Atchley, P., & Kramer, A. F. (1998)

Spatial cueing in a stereoscopic display: Attention remains "depth-aware" with age

Journals of Gerontology: Series B: Psychological Sciences & Social Sciences, **53B**, P318–P323

Previous research has demonstrated that spatial attention is "depth-aware." Reaction times (RTs) are greater for shifts in both depth and two-dimensional (2-D) space than for shifts in 2-D space alone. This experiment examined whether the ability to focus attention at a depth location is maintained with advanced age. Twelve 18- to 25-year-old and twelve 62- to 85-year-old observers viewed stereoscopic displays in which one of four spatial locations was cued. Two of the locations were at a near-depth location, and two were at a far-depth location. When the focus of visual attention was shifted to a new location in space (because of an invalid cue), the cost in RT for switching attention (measured as the difference between RT on valid cue and invalid cue trials) was greater when observers had to switch attention between different depth locations and different locations in 2-D space than for shifts in 2-D space alone. This effect was observed for both younger and older observers, suggesting that the ability to orient attention to a depth location is maintained with age.

Azoz, Y., Devi, L., & Sharma, R. (1998)

Reliable tracking of human arm dynamics by multiple cue integration and constraint fusion

Proceedings of the 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 905–910

The use of hand gestures provides an attractive means of interacting naturally with a computer-generated display. In a setup using one or more video cameras, the hand movements can potentially be interpreted as meaningful gestures. One key problem in building such an interface without a restricted setup is the computer's limited ability to localize and track the human arm robustly in image sequences. This paper proposes a multiple-cue-based localization scheme combined with a tracking framework to reliably track the human arm in unconstrained environments. The localization scheme integrates the multiple cues of motion, shape, and color for locating a set of key image features. These features are tracked by a modified extended Kalman filter that uses constraint fusion and exploits the articulated structure of the arm. We also propose an interaction scheme between tracking and localization for improving the estimation process while reducing the computational requirements. The performance of the framework is validated with the help of extensive experiments and simulations.

Azoz, Y., Devi, L., & Sharma, R. (1998)

Tracking hand dynamics in unconstrained environments

Proceedings of the Third International Conference on Automatic Face and Gesture Recognition, 274–279

A key problem in human-computer interaction via hand gestures is the computer's limited ability to localize and track the human arm in image sequences. This paper proposes a multimodal localization scheme combined with a tracking framework that exploits the articulated structure of the arm. The localization uses the multiple cues of motion, shape, and color to locate a set of image features. These features are tracked by a modified extended Kalman filter that

uses constraint fusion. An interaction scheme between tracking and localization is proposed in order to improve the estimation while decreasing the computational requirement. The results of extensive simulations and experiments with real data are described including a large database of hand gestures involved in display control.

Bangayan, P. T., & Chen, S. L. (1999)

Noise reduction techniques for speech recognition in the military environment

Proceedings of the 3rd Annual Federated Laboratory Symposium, 141

We have developed noise-reduction algorithms in an effort to improve speech recognition in noisy environments. We constructed a discrete speech recognition engine using the Entropic Hidden Markov Model Toolkit (HTK) and trained it using isolated and spelled word data from the DARPA-funded RMI database. The speech samples were corrupted with additive noise obtained from personnel at the ARL Hostile Environment Simulator (HES) facility and from the commercially available NOISEX database of military sounds. Results indicate that spectral subtraction reduces the error rate for stationary noise sources at signal-to-noise ratios ranging from 20 to 0 dB. However, applying spectral subtraction to nonstationary sources, which constitute many battlefield noises, resulted in an increased error rate. To mitigate the problem of nonstationary noise, a dual-microphone approach has been taken. A signal consisting of both speech and noise is filtered with a second signal consisting of noise alone. Data were collected at Rockwell Science Center; the noise mix was provided by ARL HES personnel. The data collection constituted a first step towards an audio-visual database for bimodal speech recognition planned for FY99. Samples of the audio-only data collection are presented.

Bargar, R., & Choi, I. (1999)

Sonification of dynamic data representation networks to reduce visual overload and enhance situational awareness

Proceedings of the 3rd Annual Federated Laboratory Symposium, 21–25

We describe a working sonification system for design and implementation of real-time data-driven auditory display. Sonification is applied to enhance the visual display of an interactive decision support system consulting a Bayesian Belief Network (BBN). The sonification case presented in this paper employs the concept of auditory signature. The auditory signature is attributed to the nodes that observers wish to keep track of, particularly for monitoring the dynamics of internal nodes. The objective is to provide fine gradients of auditory information to help observers be aware of the relative contribution of internal nodes to the outcome. For implementation of the prototype system, we developed a task-based model of BBN dynamics. This model provides criteria for the design of sonification architecture. The early development of prototype architecture allows the research team to identify constraints presented by the visual display and interactivity of the BBN, and to develop alternatives early in the project cycle.

Bargar, R., Choi, I., & Betts, A. (1999)

Scoregraph: A software architecture for rapid configuration of multimodal interaction in distributed virtual environments

Proceedings of the 3rd Annual Federated Laboratory Symposium, 41–45

This paper presents software architecture for rapid configuration of multidimensional and multimodal interactions in virtual environments. The architecture is currently in active use in the Integrated Support Laboratory, Beckman Institute, UIUC. Observation is described as interaction with a virtual environment to extract information in a time-critical manner. In the present research, an observer's multimodal capacity is supported by time scheduling techniques for parallel processing of sensors and displays to provide synchronous perceptual feedback. This modality is coupled to multidimensional numerical simulations. The ScoreGraph software architecture facilitates a temporal framework for dynamic interplay in virtual environments. A temporal framework is complementary to the static spatial organization of geometric graphical objects. Design criteria that encompass both include the management of computing resources, a configuration of an observation space, and virtual reality (VR) authoring. The temporal criteria in VR authoring have to do with efficient reconfiguration of interactive capacity in a virtual scene and the dynamics of services exchanged among parallel processes.

Barnes, M. J., & Fichtl, T. (1999)

Cognitive issues for the intelligence analyst of the future

Proceedings of the 3rd Annual Federated Laboratory Symposium, 15–19

The purpose of the paper is threefold: (1) identify important trends that affect the future analyst, (2) discuss the cognitive implications of these trends, and (3) suggest empirical and theoretical issues for further research. Four important cognitive areas are discussed in detail: knowledge acquisition, situation awareness, prediction, and intuitive processes. The conclusion is that the 21st century analyst will face radically new technology and a variety of unconventional intelligence missions. Research and decision support are discussed as possible amelioratives.

Behringer, R. (1998, October)

Improving the precision of registration for augmented reality in an outdoor scenario by visual horizon silhouette matching

paper presented at *International Workshop on Augmented Reality,* San Francisco, CA

A system for enhancing situational awareness in an outdoor scenario is being developed. The goal of such a system is to provide information through an overlay superimposed onto a video stream or directly into a head-mounted display; the superposition is done by augmented reality techniques. In an outdoor scenario, the registration between the overlay and the real world can be obtained by a combination of Global Positioning System (GPS), digital compass, and inertial sensors. However, these methods lack the precision that is required for a convincing augmented reality overlay. A means to increase the registration precision, if the terrain is well structured, is to exploit the known position of visual terrain features or man-made objects. If visible, the horizon silhouette provides cues for observer orientation. In a first step towards a system for visual

outdoor registration, visual registration through horizon silhouettes has been demonstrated on single-image snapshots. The theoretical 360° horizon silhouette could be computed from USGS digital elevation maps, which provide a grid of elevation data. The best match of the extracted visible silhouette segment onto the predicted 360° silhouette provides orientation (elevation, azimuth) and calibration of the observer camera. The system runs on a PC (200 MHz) and is being ported to a wearable platform (TREKKER).

Behringer, R. (1999)
A system for inertial stabilization of a video display
Proceedings of the 3rd Annual Federated Laboratory Symposium, 127–131

In the future, soldiers will operate equipment while riding over rough terrain in the U.S. Army's Command and Control Vehicle (C2V). The motion-induced vibration in this environment causes a massive reduction in the readability of the displays. To mitigate this problem, we have developed a system that can compensate for computer monitor motion by projecting the information on the display in an inertially stabilized window. The window is shifted on the monitor in the opposite direction as the monitor motion. A three-axis linear accelerometer measures the acceleration at the display. The acceleration data are used to shift the display window so that it appears at a fixed spatial location, although the monitor itself is moving. The system is implemented on a standard PC (Pentium Pro, 200 MHz, Windows NT 4.0) using commercial off-the-shelf components. In the paper, we present an overview on the algorithm, the system implementation, and results from vibration simulation.

Behringer, R. (1999)
A hybrid registration system for outdoor augmented reality
Proceedings of the 3rd Annual Federated Laboratory Symposium, 117–120

Using augmented reality (AR) to enhance the soldier's situational awareness requires registration of the displayed information with the real world. A hybrid registration system has been developed for registration in an outdoor environment. The system consists of the following components: a magnetometer for determining magnetic northing (digital compass) and an inclinometer for obtaining the user's head tilt and roll angle. An additional visual silhouette registration system using a camera, aligned with the user's view, improves the accuracy of the orientation. The system is prepared for later integration with a Global Positioning System (GPS) receiver for obtaining location. The registration system is being ported to a van, which will allow it to be tested at arbitrary locations. It is also being ported to a mobile wearable PC (TREKKER), which can provide simple AR functionality. The AR system will be capable of providing remote AR to a central command post. The paper describes the system architecture and presents first results of the overlay.

Berry, G. A., Pavlovic, V., & Huang, T. S. (1998, November)
BattleView: A multimodal HCII research application
paper presented at *Workshop on Perceptual User Interfaces*, San Francisco, CA

To demonstrate some of our research topics in Human-Computer Intelligent Interaction (HCII), we employ two modes of natural human-computer interaction to control a virtual environment. By using speech and gesture recognition,

we outline the control of a virtual environment research test bed (BattleView) without the need for traditional virtual reality interfaces such as a wand, mouse, or keyboard. The use of features from both speech and gesture creates a unique interface where different modalities complement each other in a more "human" communication style.

Cantú-Paz, E. (1999, July)
Migration policies, selection pressure, and parallel evolutionary algorithms
paper presented at *Late-Breaking Papers of the 1999 Genetic and Evolutionary Computation Conference*, Orlando, FL

This paper investigates how the policy used to select migrants and replacements affects the selection pressure in parallel evolutionary algorithms (EAs) with multiple populations. The four possible combinations of random and fitness-based emigration and replacement of existing individuals are considered. The investigation follows two approaches. The first is to calculate the takeover time under the four migration policies. This approach makes several simplifying assumptions, but the qualitative conclusions that are derived from the calculations are confirmed by the second approach. The second approach consists of quantifying the increase in the selection intensity. The results may help to avoid excessively high (or low) selection pressures that may cause the search to fail, and may offer a plausible explanation to the frequent claims of superlinear speedups in parallel EAs.

Cantú-Paz, E. (1999)
Topologies, migration rates, and multi-population parallel genetic algorithms
Proceedings of Genetic Algorithms and Classifier Systems, 91–98

This paper presents a study of parallel genetic algorithms (GAs) with multiple populations (also called *demes* or *islands*). The study makes explicit the relation between the probability of reaching a desired solution with the deme size, the migration rate, and the degree of the connectivity graph. The paper considers arbitrary topologies with a fixed number of neighbors per deme. The demes evolve in isolation until each converges to a unique solution. Then, the demes exchange an arbitrary number of individuals and restart their execution. An accurate deme-sizing equation is derived, and it is used to determine the optimal configuration of an arbitrary number of demes that minimizes the execution time of the parallel GA.

Cepeda, N. J., & Kramer, A. F. (1999)
Strategic effects on object-based attentional selection
Acta Psychologica, **103**, 1–19

The same-object benefit—that is, faster and/or more accurate performance when two target properties to be identified appear on one object than when each of the properties appears on different objects—has been a robust and theoretically important finding in the study of attentional selection. Indeed, the same-object benefit has been interpreted to suggest that attention can be used to select objects and perceptual groups rather than unparsed regions of visual space. This article reports and explores a different-object benefit—that is, faster identification performance when two target properties appear on different objects than when

they appear on a single object. Participants in all three experiments included 7 male and 37 female 18- to 31-year-old college students. The results from the three experiments suggest that the different-object benefit was the result of mental rotation and translation strategies that participants performed on objects in an effort to determine whether two target properties matched or mismatched. These image manipulation strategies appear to be performed with similar but not with dissimilar target properties. The results are discussed in terms of their implications for the study of object-based attentional selection.

Chakrabarti, K., & Mehrotra, S. (1999)

Efficient concurrency control in multidimensional access methods

SIGMOD Record, **28**(2), 25–36

The importance of multidimensional index structures to numerous emerging database applications is well established. However, before these index structures can be supported as access methods in a "commercial-strength" database management system (DBMS), efficient techniques to provide transactional access to data via the index structure must be developed. Concurrent access to data via index structures introduces the problem of protecting ranges specified in the retrieval from phantom insertions and deletions (the *phantom problem*). This paper presents a dynamic granular locking approach to phantom protection in Generalized Search Trees (GiSTs), an index structure supporting an extensible set of queries and data types. GiSTs provide a set of interfaces using a new multidimensional index structure that can easily be integrated into a DBMS. The granular locking technique offers a high degree of concurrency and has a low lock overhead. Through our experiments, we show that the technique scales well under various system loads. Since a wide variety of multidimensional index structures can be implemented with GiST, the developed algorithms provide a general solution to concurrency control in multidimensional access methods. To the best of our knowledge, this paper provides the first such solution based on granular locking.

Chan, M. T. (1999)

Tracking lip motion at video rate for bimodal speech recognition

Proceedings of the 3rd Annual Federated Laboratory Symposium, 47–50

In support of the development of a vision-assisted speech recognition system, we have developed a video-based algorithm that can track movements of the speaker's lips during speech utterances. The method takes advantage of prior knowledge that we have about the shape of the speaker's lips and their color in contrast to that of the skin. Because it (a) uses an explicit coarse-to-fine local search strategy, (b) constrains deformation of the model from its reference shape in an affine subspace, and (c) monitors errors and ignores outlier measurements as necessary, the algorithm is robust but still runs at a real-time video rate. Using a fast lip localization algorithm based on clustering analysis that uses the hue and saturation images, our system can also self-start without requiring user intervention at run time. We plan to incorporate the tracking subsystem into a real-time bimodal speech recognition system.

Chan, M. T. (1999)

Visual speech interface: Apparatus and algorithms

1999 World Aviation Congress, Society of Automotive Engineers (Report No. 99WAC-150)

To make speech recognition a viable input modality in the cockpit, we propose to include visual speech input to improve robustness of the approach in the presence of noise. The visual speech interface includes a head-mounted lip imaging apparatus and algorithms to recognize spoken words visually. Our algorithms are based on a few components that address all issues related to lip localization, lip shape model extraction, tracking, and feature extraction and recognition. We demonstrate the practicability of the concept with a visual speech recognizer for a discrete-word recognition task that is relatively simple but achievable in real time.

Chan, M. T., Zhang, Y., & Huang, T. S. (1998)

Real-time lip tracking and bimodal continuous speech recognition

Proceedings of the 1998 IEEE Second Workshop on Multimedia Signal Processing, 65–70

We investigate a bimodal approach to improve the accuracy of an automatic speech recognition system by augmenting acoustic speech features with visual features derived from the lip movement of the speaker. Our initial test bed includes a system that tracks in real time the positions of color markers placed on the speaker's lips while utterances are simultaneously recorded. By combining both features, we train a context-dependent hidden Markov model-based recognizer using continuous speech data that we collected based on a confined vocabulary useful for our application area. Our preliminary results show that the experimental bimodal recognizer has a higher recognition accuracy than the acoustic-only counterpart, especially at low signal-to-noise ratios. We are currently incorporating into our recognizer a new algorithm for lip tracking so that markers would not be needed. Currently the algorithm can track the outline of the lips in real time under some moderate assumptions about the speaker.

Chernyshenko, O., & Sniezek, J. A. (1998, November)

Priming for expertise and confidence in choice: Evaluating the global improves calibration for the specific

paper presented at annual meeting of *Judgment and Decision-Making Society,* Dallas, TX

Two experiments investigated the relationship between expertise priming and subjects' over- or underconfidence in their judgments. Judgment about an event is based on an individual's subjective estimate of an event's probability of occurrence. Under high uncertainty, for example, subjective probabilities often exceed the actual probability of an event, leading to overconfidence in one's judgment. Overconfidence was reduced when decisions were difficult and underconfidence was reduced when they were easy, if subjects were guided through an exercise that focused attention on their beliefs about their expertise (i.e., when the subjects were "primed" for expertise).

Cibulskis, M. J., & DeJong, G. (1999)

Interfaces that learn: Path planning through minefields

Proceedings of the 3rd Annual Federated Laboratory Symposium, 143

An approach is described for studying the problems involved in implementing an adaptable human-computer interface. To provide useful information and guidance, an adaptable interface must be sensitive to the expertise level of the user and to the user's tolerance to system interference, which may not be predictable from a user's level of expertise. Further complications arise if user preferences change over time. The authors describe a system that begins with a simplified Bayesian network that predicts what a user would like done, growing the network toward increasing accuracy of the predictions. A task in which subjects must find a route through a mine field is used to study the problems that arise with adaptable interfaces. [Abstract provided by ARL.]

Colmenarez, A. J., & Huang, T. S. (1998)

Face detection and recognition

in H. Wechsle (Ed.), *Face recognition: From theory to applications* (pp 174–185). New York: Springer

Two of the most important aspects in the general research framework of face recognition by computer are addressed here: face and facial feature detection, and face recognition—or rather face comparison. The best reported results of the mugshot face-recognition problem are obtained with elastic matching using jets. In this approach, the overall face detection, facial feature localization, and face comparison are carried out in a single step. This paper describes our research progress towards a different approach for face recognition. On the one hand, we describe a visual learning technique and its application to face detection in complex background and accurate facial feature detection/tracking. On the other hand, a fast algorithm for two-dimensional template matching is presented, as well as its application to face recognition. Finally, we report an automatic, real-time face recognition system.

Darkow, D. J., & Marshak, W. P. (1998)

In search of an objective metric for complex displays

Proceedings of the Human Factors & Ergonomics Society 42nd Annual Meeting, **2**, 1361–1365

Advanced displays for military and other user-interaction intensive systems need objective measures of merit for analyzing the information transfer from the displays to the user. A usable objective metric for display interface designers needs to be succinct, modular, and scalable. The authors have combined the concepts of weighted signal-to-noise ratio and multidimensional correlation to calculate a novel index of display complexity. Preliminary data are presented supporting the development of this metric for complex visual, auditory, and mixed auditory and visual displays. Analysis of the human subject data indicates that the coefficients for the algorithm are easily determined. Furthermore, the metric can predict reaction times and accuracy rates for complex displays. This combination of semi-automated reduction of display information and calculation of a single complexity index makes this algorithm a potentially convenient tool for designers of complex display interfaces.

Dunn, R. S. (1999)

Visualization architecture technology

Proceedings of the 3rd Annual Federated Laboratory Symposium, 145

The goal of the Crewstation Technology Laboratory is to develop and demonstrate the power of visualization architecture technology (VAT) to depict, in a command and control environment, tactically relevant information during complex operations. To this end, advanced, three-dimensional, stereoscopic display systems must be integrated with high-resolution geo-referenced imagery running on a real-time communications network managed by an executive scenario controller. As a component of VAT, the Force Operational Readiness Combat Effectiveness Simulation (FORCES) controls tactical scenarios illustrating a variety of information visualization concepts. The development of flexible VAT architecture permits, during the design phase, the evaluation of information handling and processing and of decision aids for future systems. The ultimate goals include intensifying command situation awareness and increasing the tempo of operations, as well as improving mission planning and control.

Ellis, C. D., & Johnston, D. M. (1999)

Qualitative spatial representation for situational awareness and spatial decision support

in C. H. Freksa & D. M. Mark (Eds.), *Spatial information theory: Cognitive and computational foundations of geographic information science: COSIT '99.* Berlin: Springer-Verlag

This paper summarizes research on the effectiveness of qualitative spatial representation (QSR) in two- and three-dimensional displays for improving situational awareness and decision making. The study involved (1) creating spatial query functions based on QSR that capture knowledge about objects in space; (2) building these query functions into a graphical user-interface environment as simulated user accessible support functions; and (3) testing the utility of these support functions by evaluating the performance of human subjects in solving sets of spatial decision-making and information-retrieval tasks.

Fiebig, C. B. (1999)

Designing experience-centered planning support systems

unpublished doctoral dissertation, University of Illinois, Urbana-Champaign

A design methodology, DAISY (Design Aid for Intelligent Support Systems), is used to develop computer planning support systems that meet the special needs of users at specified levels of experience. In this iterative methodology, the designers observe experts and nonexperts to develop models of the planning tasks and to identify the information and knowledge used by each group. Focusing on differences between the groups, the designers identify specialized system requirements needed to meet the information and display needs of users at a given level of experience. The effectiveness of DAISY was illustrated by its application to the design of the planning support system called Fox, a software application that generates friendly courses of action (FCOAs). Two evaluations showed that Fox significantly increased the range of FCOA options considered by expert users. [Abstract supplied by ARL.]

Fiebig, C. B., & Hayes, C. C. (1998)

DAISY: A design methodology for experience-centered planning support systems

IEEE International Conference on Systems, Man, and Cybernetics, **1**, 920–925

Designing systems to effectively assist planners in grasping a situation quickly and in making high-quality decisions is very difficult, even within a single problem-solving domain. Different types of users have very different needs, and a system designed to assist one group of users may frustrate others with differing amounts of experience. In this paper, we present DAISY, a methodology intended to enable systems designers to identify, before the system is designed, the system features required to meet the information and display needs of users at a given level of experience. These required features are identified through user problem-solving studies that result in the development of a model of the task and the identification of user information requirements and typical user errors. The DAISY methodology is unique in that it identifies the needs of planners with varying levels of experience and allows these specialized user needs to be incorporated into the software design. Unlike other approaches, DAISY provides concrete methods that are specific to the design of decision support systems for planning. We illustrate the use of this methodology in the design of an intelligent agent and human-computer interface, called Fox, for the military planning task of generating courses of action. This is a complex and difficult decision-making task in which users make life and death decisions while they are under extreme time pressure and overloaded with information.

Fiebig, C. B., Hayes, C. C., & Winkler, R. P. (1999)

What's new in Fox-GA?

Proceedings of the 3rd Annual Federated Laboratory Symposium, 9–13

It is very difficult to design planning assistants that are truly effective in helping planners to create high-quality plans quickly. In this paper, we present the results of a series of usability assessments that were conducted to determine how Fox-GA affects military planners' problem-solving behavior, and what changes needed to be made to the Fox-GA system to make it a more effective tool.

Fijalkiewicz, P. (1999)

An intelligent guidance architecture for definition and preparation of the battlefield

Proceedings of the 3rd Annual Federated Laboratory Symposium, 147

IGUANA (Intelligent Guidance and User-Adapted Interaction Agent) is a software application that provides context-sensitive, self-adapting assistance to staff planners as they define and prepare the battlefield using interactive computer controls. The battlefield definition can then be used as input for a course-of-action generator. IGUANA is distinct from previous intelligent user interfaces in that its guidance rules are not static, but evolve based on its interpretation of data about the current application, the system's hardware, the user, the user's task, and the user's environment. The IGUANA guidance agent architecture can also provide support in the form of debriefings that summarize relevant actions of past users and by providing configuration management suggestions that assist the user in adapting the presentation of information. The IGUANA architecture can also provide decision scripts to enable a user to understand the reasoning

behind other users' actions. By providing context-sensitive support, the IGUANA framework enables systems to be developed that improve user understanding of the system and user task performance.

Ghelani, D. (1998, July)

Hand tracking in video using active contour models

master's thesis, The Pennsylvania State University, State College, PA

Active contour models have attracted considerable interest in recent years. Many kinds of active contours and surfaces as well as energy-minimization schemes have been presented. One example is a snake, an energy-minimizing spline that is influenced by external forces as well as image forces that pull it toward features such as lines and edges. Snakes are used in a number of computer vision applications, such as detection of edges and lines, and in motion tracking and stereo matching. This paper presents an approach in motion (hand) tracking and analysis of deformable objects. The method is based upon modeling and extracting the boundary of an object as a generalized active contour model (snake), and then tracking the object boundary in image frames by minimizing the energy function of the contour model. We present an analysis of the contour model (snake), and discuss how the various parameters and forces of the model are selected. The proposed method has been applied to the analysis of a hand tracking experiment. In this method a snake is used to track a continuous sequence of images captured by video. Results for tracking are presented. Possible failures of the method are also presented.

Goodwin-Johansson, S., Palmer, D., Mancusi, J., Nwankwo, H., Wesier, M., & Marshak, W. (1999)

Tactile interface on a mobile computing platform

Proceedings of the 3rd Annual Federated Laboratory Symposium, 51–55

Tactile devices can be used by dismounted soldiers to augment the traditional visual and auditory communication channels. We conducted experiments to investigate the use of an experimental first-generation system of tactile devices controlled by a portable computer (DASHER) to convey directional information to the dismounted soldier. Two experiments were performed. The first experiment investigated the ability of a subject to correctly identify which vibratory tactile device was actuated out of five spatially separate devices. The second experiment investigated the ability of a subject to use vibratory tactile inputs from five devices to identify 18 different directions. The results of the first experiment indicated that subjects could correctly identify which device was actuated between 82 and 98 percent of the time for the strong vibration level. The results of the second experiment indicate that if we use combinations of actuators operating at different vibration levels, five actuators are sufficient to communicate to a soldier 18 different directions.

Goldberg, D. E. (1998)

A meditation on the application of genetic algorithms

Tech. Rep. No. 98003, University of Illinois at Urbana-Champaign, Illinois Genetic Algorithms Laboratory

An argument is presented that genetic algorithms, as search procedures, are not ephemerae, even though they exhibit limitations when shifted from simple, small-scale problems to more complex, real-world ones. Rather than describe successful applications of genetic algorithms, the author accounts for researchers' persistence in employing genetic algorithms by emphasizing the overriding importance of natural selection as an explanatory account of life in the natural environment and the ineffectiveness of traditional optimization and operations research methods.

Goldberg, D. E. (1999)

Using time efficiently: Genetic-evolutionary algorithms and the continuation problem

Tech. Rep. No. 99002, University of Illinois at Urbana-Champaign, Illinois Genetic Algorithms Laboratory

This paper develops a macro-level theory of efficient time utilization for genetic and evolutionary algorithms. Building on population sizing results that estimate the critical relationship between solution quality and time, the paper considers the trade-off between large populations that converge in a single convergence epoch and smaller populations with multiple epochs. Two models suggest a link between the salience structure of a problem and the appropriate population-time configuration for best efficiency.

Goldberg, D. E., & Voessner, S. (1999)

Optimizing global-local search hybrids

Tech. Rep. No. 99001, University of Illinois at Urbana-Champaign, Illinois Genetic Algorithms Laboratory

This paper develops a framework for optimizing global-local hybrids of search or optimization procedures. The paper starts by idealizing the search problem as a search by a global algorithm *G* for either (1) acceptable *targets*—solutions that meet a specified criterion—or (2) *basins of attraction* that then lead to acceptable targets under a specified local search algorithm *L*. The paper continues by abstracting two sets of parameters: probabilities of successfully hitting targets and basins and time-to-criterion coefficients. With these parameters, equations may be written to account for the total time of search and for the probabilistic success (reliability) in reaching an acceptable solution. Thereafter, optimization problems are formulated in which the division of local versus global search time is optimized so that solution time to acceptable reliability is minimized, or reliability under specified solution time is maximized. A two-basin optimality criterion is derived and applied to important representative problems. Continuations and extensions of the work are suggested, but the theory appears to be immediately useful in better understanding the economy of hybridization.

Gupta, M. P. (1999)
Reservation-based distributed resource management
master's thesis, University of Illinois, Urbana-Champaign

An architecture is described that allows a process to reserve resources on a remote host. The architecture incorporates a resource agent on all hosts involved in a distributed application. These agents are connected and provide for transfer of reservation information among themselves. A request for distributed reservation is made with one of the agents. The request is split according to process locations, and individual components are sent to corresponding agents. The agents in turn interact with various brokers and reserve resources. A broker specializes in management of a single resource in a single end system. The prototype implementation provides reservation for CPU cycles. As brokers for other end-system resources are developed, they can be easily incorporated into the architecture. [Abstract supplied by ARL.]

Hahn, S., & Kramer, A. F. (1998)
Further evidence for the division of attention among noncontiguous locations
Visual Cognition, **5**, 217–256

An investigation was made of the boundary conditions on the ability to divide attention among different locations in visual space. In each of five studies, undergraduates (aged 18 to 33 years) performed a same-difference matching test with target letters that were presented on opposite sides of a set of distractor letters. Experiments 1, 2, and 3 provide further support for the proposal that subjects can concurrently attend to noncontiguous locations as long as new objects do not appear between the attended areas. Experiment 4 examined whether the disruption of multiple attentional foci was the result of the capture of attention by new objects per se, or by task-irrelevant objects. Multiple attentional foci could be maintained as long as distractor objects did not appear between target locations. Experiment 5 examined whether attention can be divided among noncontiguous locations within as well as between hemifields. Hemifield boundaries did not constrain the subjects' ability to divide attention among different areas of visual space. The results are discussed in terms of the nature of attentional flexibility and putative neuroanatomical mechanisms that support the ability to split attention among different regions of the visual field.

Harik, G., Cantú-Paz, E., Goldberg, D. E., & Miller, B. L. (1999)
The gambler's ruin problem, genetic algorithms, and the sizing of populations
Evolutionary Computation, **7**, 231–253

A model is presented to predict the convergence quality of genetic algorithms (GAs) based on the size of the population. The model is based on an analogy between selection in GAs and one-dimensional random walks. Using the solution to a classic random-walk problem (the gambler's ruin), the model naturally incorporates previous knowledge about the initial supply of building blocks (BBs) and correct selection of the best BB over its competitors. The result is an equation that relates the size of the population with the desired duality of the solution, as well as the problem size and difficulty. The accuracy of the model is verified with experiments using additively decomposable functions of varying difficulty. The paper demonstrates how to adjust the model to account for noise present in the fitness evaluation and for different tournament sizes.

Hayes, C. C., Schlabach, J. L., & Fiebig, C. B. (1998)

Fox-GA: An intelligent planning and decision support tool

Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, **3**, 2454–2459

Fox-GA is described, an intelligent planning decision support tool for assisting military intelligence and maneuver battle staff in rapidly generating and assessing battlefield courses of action (COAs). The motivations behind Fox stem from the need to plan and replan rapidly so as to allow users flexibility and control over planning objectives and options. The environment in which plans are executed (the battlefield) is one that is inherently uncertain and rapidly changing, demanding frequent replanning during execution. To help meet these rapid replanning needs, we designed Fox to rapidly generate and evaluate a broader variety of high-quality COAs faster than military staff could do themselves. Fox then evaluates the COAs and presents only the best few to users, allowing users to reassess those options according to their own judgment, and to either edit or select the ones they feel are best. Early evaluations indicate that users explore a wider variety of COAs with Fox than without.

Huang, T. S., Ramchandran, K., Smith, M.J.T., & Farvardin, N. (1999)

Image and video compression: Meeting the Army needs

Joint Proceedings of the 3rd Annual Federated Laboratory Symposium, 7–14

The Army needs compression technologies for multispectral and multisensor images and video that are high-performance, low-complexity, scalable, interoperable, and robust to noise. Performance includes not only compression ratio, but also good target recognizability and ease of manipulation in the compressed domain. This paper highlights work in data compression under study in the three Army Federated Laboratory Consortia.

Huang, J., & Zhao, Y. (1997)

Energy constrained signal subspace method for speech enhancement and recognition

IEEE Signal Processing Letters, **5**, 283–285

An improved signal-subspace-based speech-enhancement algorithm is proposed for automatic speech recognition under an additive noise environment. The key idea is to match the short-time energy of the enhanced speech signal to the unbiased estimate of the short-time energy of the clean speech; this technique has proven very effective for improving the estimation of the low-energy segments of continuous speech under low-noise conditions. Experimental results show significant improvement in both the segmental signal-to-noise ratios (SNRs) and the word-recognition accuracy of the enhanced speech under SNRs of 10 to 20 dB.

Iskarous, K., & Morgan, J. (1999)

Speech synthesis in a virtual environment

Proceedings of the 3rd Annual Federated Laboratory Symposium, 149

A method is described that increases the intelligibility of synthesized speech by focusing on the synthesis of stop consonants like *t*, *d*, *k*, *g*, and *n*, which occur frequently enough to hinder understanding of synthesized speech. As a solution

to this problem, tongue movement produced during consonant-vowel frequency transitions is modeled by a cubic bezier spline curve whose shape is specified completely by four control points. Complex tongue motion during a transition is modeled by the movement of only two of these four points, which can be represented by a change in a very small number of parameters sampled at 5 to 8 points. This is an improvement over current systems, which synthesize speech by transitioning between concatenated speech sounds by linear or higher order frequency interpolation. [This abstract was supplied by ARL and is based on a poster summary appearing in the conference proceedings.]

Iskarous, K. (1999)
Patterns of tongue movement
Proceedings of the 14th International Congress of Phonetic Sciences, 429

This paper discusses the pivot pattern of tongue movement. In this pattern, there seems to be a point in the vocal tract where there is no motion, but there is motion at points of the vocal tract anterior and posterior to the pivot point. Based on tongue edge tracings of frames from ultrasound and x-ray dynamic imaging of the vocal tract, I show that the pivot pattern is used in a variety of sequences, and I discuss the possible causes of the pattern.

Jog, K. (1998)
Stereoscopic calibration of a see-through head-mounted display
unpublished master's thesis, The Pennsylvania State University, University Park, PA

Abstract not available.

Johnston, D. M., & Ellis, C. D. (1999)
The effectiveness of qualitative spatial representation in supporting spatial awareness and decision making
Proceedings of the 3rd Annual Federated Laboratory Symposium, 71–75

This paper summarizes elements of research on the effectiveness of using qualitative spatial representation (QSR) in two- and three-dimensional display modes to determine its usefulness for spatial awareness and decision making. The study involved (1) creating spatial query functions based on QSR that capture knowledge about objects in space; (2) building these query functions into a graphical user interface environment as simulated user-accessible support functions; and (3) testing the utility of these support functions by evaluating the performance of human subjects in solving sets of spatial decision-making and information-retrieval tasks.

Jojic, N., Gu, J., Shen, H. C., & Huang, T. S. (1998)
3-D reconstruction of multipart self-occluding objects
in R. Chin & T. C. Pong (Eds.), *Lecture Notes in Computer Science* (pp II-455–II-462). Springer: New York

In this paper, we present a method for reconstruction of multipart objects from several arbitrary views using deformable superquadrics as models of the object's parts. Two visual cues are used: occluding contours and stereo (possibly aided by projected patterns). The object can be relatively complex and can exhibit numerous self-occlusions from some or all views. Our preliminary experiments on a

human body and a tailor's mannequin show that the reconstruction is more complete than in purely stereo or structured light-based methods and more precise than the reconstruction from occluding contours only.

Jojic, N., & Huang, T. S. (1998)
On analysis of cloth drape range data
in R. Chin & T. C. Pong (Eds.), *Lecture Notes in Computer Science* (pp II-463–II-470). Springer: New York

In this paper, we present an algorithm for analyzing the range data of cloth drapes. The goal is the estimation of parameters for modeling and the geometry of the underlying object. In an analysis-by-synthesis manner, the algorithm compares the drape of the model with the range data and searches for the best fit. It can be applied to any physics-based cloth model. The motivating application is fashion design using CAD systems, but the ability of the algorithm to estimate the shape of the object supporting the scanned cloth indicates the possibility of using cloth models to overcome problems in human tracking algorithms caused by clothing.

Jones, P. M., Hayes, C. C., Wilkins, D. C., Bargar, R., Sniezek, J., Asaro, P., Mengshoel, O., Kessler, D., Lucenti, M., Choi, I., Tu, N., & Schlabach, J. (1998)
CoRAVEN: Modeling and design of a multimedia intelligent infrastructure for collaborative intelligence analysis
Proceedings of the 1998 IEEE International Conference on Systems, Man, and Cybernetics, **1**, 914–919

Intelligence analysis is one of the major functions performed by an Army staff in battlefield management. In particular, intelligence analysts develop intelligence requirements based on the commander's information requirements, develop a collection plan, and then monitor messages from the battlefield with respect to the commander's information requirements. The goal of the CoRAVEN project is to develop an intelligent collaborative multimedia system to support intelligence analysts. Key ingredients of our design approach include (1) significant knowledge engineering activities with domain experts, (2) representation of an explicit model of reasoning and activity to drive design, (3) the use of Bayesian belief networks as a way to structure inferences that relate observable data to the commander's information requirements, (4) collaborative graphical user interfaces to provide flexible support for the multiple tasks in which analysts are engaged, (5) sonification of data streams and alarms to support enhanced situation awareness, (6) detailed psychological studies of reasoning and judgment under uncertainty, and (7) iterative prototyping of candidate designs with domain experts. This paper presents our recent progress on all these fronts.

Kramer, A. F., Larish, J. L., Weber, T. A., & Bardell, L. (1999)
Training for executive control: Task coordination strategies and aging
in D. Gopher & A. Koriat (Eds.), *Attention and Performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp 617–652). The MIT Press: Cambridge, MA

The authors studied the ability to successfully coordinate the performance of multiple tasks as a function of two multitask training strategies, variable priority (VP) training and fixed priority (FP) training. The acquisition, retention, and transfer of task coordination skills was investigated in adults, both young (aged

18 to 29 years) and old (60 to 75 years). After training on two tasks (a canceling and a tracking task), each of which possessed both repeating and random sequences, the authors asked subjects to perform several novel versions of the two tasks in an effort to evaluate learning of the repeated patterns in the single- and dual-task conditions. The authors then had the subjects perform two novel tasks in an effort to examine the generalizability of task coordination skills acquired during VP and FP training. Finally, retention of the original training tasks was assessed, in single- and dual-task conditions, 45 to 60 days after the training intervention. Results indicated that subjects trained with the VP procedure learned the training tasks more quickly and exhibited a higher level of mastery of the tasks than did subjects trained with the FP technique. Furthermore, the decrement in dual-task performance usually found in older adults (and observed before training in the older adults in this study) was substantially reduced for the VP- but not for the FP-trained subjects. Finally, subjects trained with the VP procedure exhibited better transfer to novel tasks as well as higher levels of retention than did FP-trained subjects.

Kramer, A. F., & Weber, T. A. (1999)

Applications of psychophysiological techniques to human factors

Proceedings of the 3rd Annual Federated Laboratory Symposium, 85–89

This paper provides a brief overview and critical review of two different potential applications of psychophysiological techniques to important issues in human factors: the assessment of fluctuations in alertness and the use of psychophysiological measures in online adaptive algorithms. The advantages and disadvantages of using psychophysiological measures in these domains are described, and the potential for further development of psychophysiologically based assessment of mental processing and operator state is discussed.

Kramer, A. F., Weber, T. A., & Watson, S. E. (1997)

Object-based attentional selection—Grouped arrays or spatially invariant representations? Comment on Vecera and Farah (1994)

Journal of Experimental Psychology: General, **126**, 3–13

S. P. Vecera and M. J. Farah addressed the issue of whether visual attention selects objects or locations. They obtained data that they interpreted as evidence for attentional selection of objects from an internal spatially invariant representation. Kramer, Weber, and Watson question this interpretation on both theoretical and empirical grounds. First, the authors suggest that there are other interpretations of the Vecera and Farah data that are consistent with location-mediated selection of objects. Second, they provide data, using the displays employed by Vercera and Farah along with a post-display probe technique, suggesting that attention is directed to the locations of the target objects. The implications of the results for space- and object-based attentional selection are discussed.

Kramer, A. F., Hahn, S., Irwin, D. E., & Theeuwes, J. (1999)

Attentional capture and aging: Implications for visual search performance and oculomotor control

Psychology and Aging, **14**, 135–154

Two studies were performed that examined potential age-related differences in attentional capture. Subjects were instructed to move their eyes as quickly as possible to a color singleton target and to identify a small letter located inside it. On half of the trials, a new stimulus (i.e., a sudden onset) appeared simultaneously with the presentation of the color singleton target. The onset was always a task-irrelevant distractor. Response times were lengthened, for both young and old adults, whenever an onset distractor appeared, despite the fact that subjects reported being unaware of the appearance of the abrupt onset. Eye-scan strategies were also disrupted by the appearance of the onset distractors. On about 40 percent of the trials during which an onset appeared, subjects made an eye movement to the task-irrelevant onset before moving their eyes to the target. Fixations close to the onset were very brief, suggesting parallel programming of a reflexive eye movement to the onset and goal-directed eye movement to the target. These data are discussed in terms of age-related sparing of the attentional and oculomotor processes that underlie the phenomenon of attentional capture.

Kramer, A. F., & Weber, T. A. (1999)

Object-based attentional selection and aging

Psychology and Aging, **14**, 99–107

Two studies were conducted that examined potential age-related differences in object-based attentional selection. In both studies, subjects were briefly presented with pairs of wrenches and asked to make one response if two target properties (i.e., an open end and hexagonal end) were present and another response if only a single target property was present in the display. The critical manipulation was whether the target properties were present on one wrench or distributed between two wrenches. Space-based models of selective attention predict no difference in performance between these conditions. However, object-based attentional selection models predict better performance when both target properties appear on a single object. The results from both studies were consistent with object-based models of attentional selection. Furthermore, both young and old adults showed similar performance effects, suggesting that object-based attentional selection is insensitive to normal aging.

Li, Y., & Zhao, Y. (1998)

Recognizing emotions in speech using short-term and long-term features

Proceedings of the 5th International Conference on Spoken Language Processing, **6**, 2255–2258

The acoustic characteristics of speech are influenced by speakers' emotional status. In this study, we attempted to recognize the emotional status of individual speakers by using speech features extracted from short-time analysis frames as well as speech features representing entire utterances. Principal component analysis was used to analyze the importance of individual features in representing emotional categories. Three classification methods were used, including vector quantization, artificial neural networks, and a Gaussian mixture density model. Classifications using short-term features only, long-term features only,

and both short-term and long-term features were conducted. The best recognition performance (of 62-percent accuracy) was achieved when the Gaussian mixture density method was used with both short-term and long-term features.

Loschky, L. C., & McConkie, G. W. (1999)

Gaze contingent displays: Maximizing display bandwidth efficiency

Proceedings of the 3rd Annual Federated Laboratory Symposium, 79–83

One way to economize on bandwidth in single-user head-mounted displays is to put high-resolution information only where the user is currently looking. This paper describes a series of six research projects investigating spatial, resolutional, and temporal parameters affecting perception and performance in eye-contingent multiresolutional displays. Based on the results of these projects, suggestions are made for the design of eye-contingent multiresolutional displays.

Ma, J., & Ahuja, N. (1998)

Dense shape and motion from region correspondences by factorization

Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 219–224

In this paper, we propose an algorithm for estimating dense shape and motion of dynamic piecewise planar scenes from region correspondences using factorization. Region correspondences are used since they are easier to establish and more reliable than either line or point correspondences. The image measurements required are the centroid and area for each region. We use singular value decomposition to find the basis of range space of the motion, shape, and surface normal matrices. By imposing model constraints, we can recover motion, shape, and surface normal only from region correspondences.

Ma, J., & Ahuja, N. (1999)

3-D reconstruction from video sequences

Proceedings of the 3rd Annual Federated Laboratory Symposium, 151

A process is described to estimate three-dimensional structure from two-dimensional video sequences. In contrast to existing methods that use only pixel- or line-based features, the process presented here was a multi-feature-matching algorithm. Image frames are independently segmented at multiple scales, and salient regions are identified across successive video frames based on characteristics such as region area, moments, intensity values, shape compactness, and adjacency. The three-dimensional (3-D) motion and structure of these matched regions are estimated from the established correspondences with a region-based structure-from-motion algorithm. In a second step, the 3-D estimates are used to guide pixel-level matching of the unmatched areas. Candidates for pixel matches are selected in part on the basis of the 3-D motion and structure estimates, and matching is performed in terms of intensity, edgeness, and cornerness. Finally, the 3-D structure for each pixel is calculated. From matches of the first three frames, a trilinear tensor can be recovered, which describes the relations between pixels in three images and can be used to predict locations of pixels in subsequent frames. The trilinear tensor provides a general warping function between the pixels in different frames, and is used as a measure of confidence for matching in subsequent frames. [This abstract was supplied by ARL and is based on a poster summary appearing in the conference proceedings.]

Marshak, W. P., & Darkow, D. J. (1998)

Prototype depth-separated coincident transparent (true depth) display

Proceedings of the 42nd Annual Meeting of the Human Factors & Ergonomics Society, **2**, 1151

Getting the "big picture" from computer displays is a critical problem for user interface designers. Traditional solutions layer information on a single display or make multiple displays available simultaneously. These strategies fragment the information and require the user to integrate information across displays. A new display strategy under development uses depth-separated coincident transparent displays that we call "True Depth." The True Depth Display (TDD) employs two display surfaces in the same visual space but separated in depth. Users may read either surface by refocusing their eyes or by focusing between the displays to see both. Display formats can be organized to exploit their spatial coincidence, making integration across displays easy. Information density can be increased without the debilitating effects of clutter. A compact hardware prototype of the TDD was shown along with a variety of example formats to demonstrate the capabilities of this new display technology interface.

Marshak, W. P., & Darkow, D. J. (1998)

Objective measurement of display formats: Multidimensional and multimodal user perception models

Proceedings of the IEEE 1998 International Conference on Image Processing, **2**, 505–509

Comparing the effectiveness of display formats, especially displays set in different sensory modalities and containing complex combinations of dimensions, can be like comparing the proverbial apples and oranges. Dissimilar displays can be compared if a "unitless" dimension can be found that describes how well critical information is expressed, compared to other information contained in the display. Signal-to-noise ratio (SNR) is such a measure. Fourier power spectra can be computed for energy imparted by the display of critical information (signals) and the remainder of the display (noise). By computing SNRs for each feature channel (modality or dimension), one can obtain complex SNRs to describe the salience of the signal. Also considered is the similarity of signal and noise as expressed in the Pearson product-moment correlation coefficient. Computational examples of such display SNRs are presented and discussed.

Marshak, W. P., Winkler, R., Fiebig, C., Khakshour, A., & Stein, R. (1999)

Evaluating intelligent aiding of course of action decisions using the Fox genetic algorithm in 2-D and 3-D interfaces

Proceedings of the 3rd Annual Federated Laboratory Symposium, 27–31

Intelligent aiding to improve decision processes and reduce support staff will become increasingly important in future Army Tactical Operations Centers (TOCs). Federated Laboratory researchers have developed the Fox genetic algorithm (GA) decision aid to increase the number and quality of alternative courses of action (COAs) considered by the commander. Eleven Army officers at Ft. Leavenworth, Kansas, used both traditional paper-based briefing and the Fox-GA COA generator to determine a course of action in three different combat scenarios. Presentation of the Fox-GA COAs was made either within a two-dimensional (2-D) interface based on the ARL's Combat Information Processor (CIP) or within the National Center for Supercomputer Applications' battleview

three-dimensional (3-D) visualization system. The findings indicate that Fox-GA significantly increased (by two to three times) the number of alternatives considered over the paper condition, and that the 2-D visualization with Fox was both preferred and led to the best performance. These results indicate that an improved GA-based COA generation system can significantly increase the number of alternatives considered in the military planning process.

McCormick, E. R., Wickens, C. D., Banks, R., & Yeh, M. (1998)

Frame of reference effects on scientific visualization subtasks

Human Factors, **40**, 443–451

Performance measures for three frames of reference (full egocentric, full exocentric, and tethered) were contrasted across four different scientific visualization subtasks: search, travel, local judgment support, and global judgment support. Participants were instructed to locate and follow a designated path through 15 simple virtual environments and answer simple questions about that environment. Each participant completed five trials in all three frame-of-reference conditions. The results revealed that frames of reference that use egocentric or tethered viewpoints support better travel performance, especially when participants were nearing the target. However, the exocentric frame of reference supported better performance in the search subtasks and in the local and global judgment subtasks. Actual or potential applications of this research include proper uses of virtual reality to support certain scientific visualization subtasks.

Mengshoel, O. J. (1999)

Evolutionary computation in Bayesian networks

in J. R. Koza (Ed.), *Late Breaking Papers at the Third Annual Genetic Programming Conference on System Sciences* (p 159). Madison, WI: Omni Press

Genetic algorithms (GAs) are stochastic algorithms for search, optimization, and machine learning. In this research, the focus is on using a Bayesian network (BN) as the GA fitness function. More formally, a Bayesian network is a tuple (V,W,P_r), where (V,W) is a directed acyclic graph with nodes $V = \{V_1, ...,V_n\}$ and edges $W = \{W_1, ...,W_m\}$; P_r is a set of conditional probability distribution tables. The nodes correspond to random variables, and the edges to conditional dependencies between these random variables. For each node $V_i \in V$, there is one conditional probability table that defines a conditional probability distribution over V_i in terms of its parents $P_a(V_i)$: $P_r(V_i \mid P_a(V_i)) \in P_r$.

Mengshoel, O. J., Goldberg, D. E., & Wilkins, D. C. (1998)

Deceptive and other functions of unitation as Bayesian networks

in J. R. Koza (Ed.), *Genetic Programming* (pp 559–566). San Francisco, CA: Morgan Kaufmann

In trying to understand which fitness functions are hard and which are easy for genetic algorithms to optimize, researchers have considered deceptive and other functions of unitation. This paper focuses on genetic algorithm fitness functions represented as Bayesian networks. We investigate onemax, trap, and hill functions of unitation when converted into Bayesian networks. This paper shows, among other things, that Bayesian networks can be deceptive.

Mengshoel, O. J., & Goldberg, D. E. (1999, July)

Probabilistic crowding: Deterministic crowding with probabilistic replacement

paper presented at *1999 Genetic and Evolutionary Computation Conference*, Orlando, FL

This paper presents a novel niching algorithm—probabilistic crowding. Like its predecessor, deterministic crowding, probabilistic crowding is fast and simple, requiring no parameters beyond that of the classical genetic altorithm. In probabilistic crowding, subpopulations are maintained reliably, and we analyze and predict how this maintenance takes place. This paper also identifies probabilistic crowding as a member of a family of algorithms that we call integrated tournament algorithms. Integrated tournament algorithms also include deterministic crowding, restricted tournament selection, elitist recombination, parallel recombinative simulated annealing, the Metropolis algorithm, and simulated annealing.

Mengshoel, O. J., & Wilkins, D. C. (1998, March)

Abstraction for belief revision: Using a genetic algorithm to compute the most probable explanation

presented at *AAAI Spring Symposium Series*, Stanford University, Menlo Park, CA

A belief network can create a compelling model of an agent's uncertain environment. Exact belief network inference, including computing the most probable explanation, can be computationally difficult. Therefore, it is interesting to perform inference on an approximate belief network rather than on the original belief network. This paper focuses on approximation in the form of abstraction. In particular, we show how a genetic algorithm can search for the most probable explanation in an abstracted belief network. Because belief network approximation can be treated as noise from the point of view of a genetic algorithm, this topic is related to research on noisy fitness functions used for genetic algorithms.

Mengshoel, O. J., & Wilkins, D. C. (1998)

Genetic algorithms for belief network inference: The role of scaling and niching

in V. W. Porto, N. Saravanan, D. Waagen, & A. E. Eiben (Eds.), *Proceedings of the 7th International Conference on Evolutionary Programming* (pp 547–556). Berlin, Germany: Springer-Verlag

Belief networks encode joint probability distribution functions and can be used as fitness functions in genetic algorithms. Individuals in the genetic algorithm's population then represent instantiations, or explanations, in the belief network. Computing the most probable explanations (belief revision) is thus cast as a genetic algorithm search in the joint probability distribution space. At any time, the best fit individual in the genetic algorithm population is an estimate of the most probable explanation. This paper argues that joint probability distribution functions represented by belief networks typically are multimodal and highly variable. Thus the genetic algorithm techniques known as sharing and scaling should be helpful. It is shown empirically that this is indeed the case, in particular, that niching combined with scaling significantly improves the quality of a genetic algorithm's estimate of the most probable explanations. A novel scaling approach, root scaling, is also introduced.

Merlo, J. L., Wickens, C. D., & Yeh, M. (1999)

Effect of reliability on cue effectiveness and display signaling

Tech. Rep. No. ARL-99-4/FED-LAB-99-3, Urbana-Champaign: University of Illinois, Aviation Research Lab, Institute of Aviation

The effects of automation failure on trust and of visual cueing precision on attention were investigated in a target detection task. Twenty military subjects searched a simulated mountainous terrain for military-relevant targets while performing a secondary monitoring task on either a hand-held display (HHD) or a helmet-mounted display (HMD). Both displays had target cueing present for half the trials, with the precision of the target cues varied across blocks. Cued trials were either precise (a cueing reticle always circumscribed a target) or imprecise (the target was outside the reticle by 22.5° or 45°). Imprecise cueing simulated degraded sensor resolution. Cue precision and imprecision were conveyed to subjects by solid or dashed lines, respectively. A high-priority target was presented twice each block, once with a precisely cued target and once with an imprecisely cued target. Target cueing induced an attention cost (as revealed by the low detection rate of high-priority uncued targets), when a cue occurred simultaneously with a low-priority target. During the last experimental block, the automated target cueing failed on some trials, resulting in attention and trust costs, with subjects initially showing signs of overtrust of the cueing information, and then on subsequent trails tending to undertrust the cueing information, with trust seemingly restored after a few reliable trials. Failures in automation also seemed to mediate the effects of attention costs, as the detection rate of the higher priority but uncued target increased. [Abstract provided by ARL.]

Mountjoy, D. N., & Marshak, W. (1999)

Impact of nonlinear mapping on mileage estimation

Proceedings of the 3rd Annual Federated Laboratory Symposium, 97–101

Nonlinear mapping is a display technique that can be applied to situation maps to maintain detail in the commander's area of interest while displaying more peripheral land area to convey contextual information. A series of studies has been undertaken to explore the perceptual advantages and limitations of this technique in an effort to produce a more efficient tactical mapping system. The first of this series (the effect on mileage estimation) is discussed here, along with directions of future research.

Munetomo, M., & Goldberg, D. E. (1999)

Identifying linkage groups by nonlinearity/nonmonotonicity detection

Proceedings of the 1999 Genetic and Evolutionary Computation Conference, 433–440

This paper presents and discusses direct linkage identification procedures based on nonlinearity/nonmonotonicity detection. The algorithm we propose checks arbitrary nonlinearity/nonmonotonicity of fitness change by perturbations in a pair of loci to detect their linkage. We first discuss the condition of the LINC (linkage identification by a nonlinearity check) procedure and its allowable nonlinearity. Then we propose another condition of the LIMD (linkage identification by nonmonotonicity detection) and prove its equality to the LINC with allowable nonlinearity (LINC-AN). The procedures can identify linkage groups for problems with, at most, order-k difficulty by checking $O(2^k)$ strings; the computational cost for each string is $O(l^2)$, where l is the string length.

Ntuen, C. A. (1999)

An ecological model of situation awareness: What does it mean to battlefield awareness?

Proceedings of the 3rd Annual Federated Laboratory Symposium, 153

Most studies of situation awareness (SA), especially those designed for decision aiding, rely on the theories and models of cognition and perception. Theories developed by Endsley and by Pew conceptualize SA as the interaction of product and process. *Product* refers to the state of our knowledge about the environment, and *process* to the perceptual and cognitive activities that update our knowledge. The author discusses how these ideas pertain to designing decision-aiding software applications. [This abstract was supplied by ARL and is based on a poster summary appearing in the conference proceedings.]

Ntuen, C. A., Chi, C.-J., McBride, M. E., & Park, E. H. (1998)

Decision support display modeling for digital battlefield

Proceedings, Fourth Annual Symposium on Human Interaction with Complex Systems, 155–159

A decision support display (DSD) was developed as a cognitive aiding tool to support the decision maker in an unstructured, dynamic, uncertain, and information-intensive environment. Battlefield information is modeled as a context-dependent and action-oriented object that adapts to a defined system goal or mission statement. The DSD philosophy is applied to a graphical display of alternative courses of action designed to amplify the decision maker's knowledge and experience levels.

Ntuen, C. A., Park, E. H., Chi, C., Yarborough, L. P., & Mountjoy, D. N. (1999)

Effect of information presentation mode on condition monitoring of battlefield events

Proceedings of the 3rd Annual Federated Laboratory Symposium, 155

The goal of this study was to determine the most effective method of presenting critical battle information to the commanders to ensure the rapid detection of potentially disastrous conditions. Electronic map displays containing unit symbols and course-of-action arrows—which were drawn with bands across them—served as stimuli. Four methods of presentation were tested: color band changes; color band changes and flashing unit symbols; color band changes and an auditory alarm; color band changes, flashing unit symbols, and an auditory alarm. Results indicated that performance was faster on the second and fourth conditions than on the first and third. [This abstract was supplied by ARL and is based on a poster summary appearing in the conference proceedings.]

Nwankwo, H. E., Urquhart, R., Goodwin-Johansson, S., & Mancusi, J. (1999)

Tactile communication interface design: Efficacy of euphemistic terms as interface location cues

Proceedings of the 3rd Annual Federated Laboratory Symposium, 159

To apply tactile interface communication for the purpose of increasing human information processing, we must address the issue of how the interface device should be designed to ensure meaningful information transfer. In this paper we examine the relationship between a set of military communications (e.g., danger

area, stop) and associated body locations (e.g., armpit) and gestures (e.g., "cut throat"). Eighty subjects indicated on a questionnaire how intuitive the relationship between a military communication and body location was. For example, "danger area" was strongly related to "armpit," and "stop" to a "cut throat" gesture. The body locations identified could become interface locations for receiving tactile messages. Experiments are under way to validate findings gleaned from subjects' questionnaire responses. [This abstract was supplied by ARL and is based on a poster summary appearing in the conference proceedings.]

Ortega, M., Chakrabarti, K., Porkaew, K., & Mehrotra, S. (1998, June)
Cross media validation in a multimedia retrieval system
paper presented at *3rd ACM Conference on Digital Libraries*, Digital Library Metrics Workshop, Pittsburgh, PA

The increasing size of document databases has prompted a change from manual indexing and querying to automated methods. This switch necessitated a performance metric for the automated systems; however, performance measurement of automated systems was and still is performed manually. Ever-increasing collection size makes manual evaluation progressively more difficult, and this difficulty is compounded by the addition of multimedia. In this paper we describe an automated method for measuring the retrieval performance of a new arbitrary retrieval algorithm suited to a particular media type.

Ortega, M., Rui, Y., Chakrabarti, K., Porkaew, K., Mehrotra, S., & Huang, T. S. (1999)
Supporting ranked Boolean similarity queries in MARS
IEEE Transactions on Knowledge and Data Engineering, **10**, 905–25

To address the emerging needs of applications that require access to and retrieval of multimedia objects, we are developing the Multimedia Analysis and Retrieval System (MARS). In this paper, we concentrate on the retrieval subsystem of MARS and its support for content-based queries over image databases. Content-based retrieval techniques have been extensively studied for textual documents in the area of automatic information retrieval. This paper describes how these techniques can be adapted for ranked retrieval over image databases. Specifically, we discuss the ranking and retrieval algorithms developed in MARS based on the Boolean retrieval model and describe the results of our experiments, which demonstrate the effectiveness of the developed model for image retrieval.

Pavlovic, V., & Huang, T. S. (1999)
Multimodal prediction and classification of hand gestures and speech
Proceedings of the 3rd Annual Federated Laboratory Symposium, 161

The authors propose a novel framework for multimodal feature prediction and classification based on multimodal hidden Markov models (MHMMs). Previous approaches employed loosely coupled unimodal techniques in which feature estimation, prediction, and lower level classification are performed independently within each of the modality domains. MHMMs model the redundancy among co-occurring modalities such as speech, hand gestures, lip motion, etc. In this report, the test bed application was a joint audio-visual interpretation of speech and unencumbered hand gestures for interaction with virtual environments. The setup allowed a user to interact with a three-dimensional virtual

environment using hand gestures (such as pointing and simple symbolic motions) and spoken commands. Bimodal HMMs were employed to model the influence of speech on gestural actions. MHMM parameter learning was performed on a set of 39 bimodal commands. The test set was a different sequence of 31 commands performed by the same user. Two experiments compared the performance of bimodal with unimodal models on the test data. In the normal visual noise environment, recognition performance of bimodal HMMs significantly exceeds the performance of unimodal HMMs (62 versus 35 percent). High visual noise reduced the recognition performance of both models. However, bimodal HMMs retained a relatively significant recognition ratio of 52 percent, while the unimodal approach failed almost completely (10 percent). Results of the test indicated that the bimodal HMMs significantly improved the recognition performance in two different gestural speech classification tasks. Future work is aimed at further examination of the robustness of classification as well as the online implementation of the algorithms. [This abstract was supplied by ARL and is based on a poster summary appearing in the conference proceedings.]

Pavlovic, V. I., Sharma, R., & Huang, T. S. (1997)

Visual interpretation of hand gestures for human-computer interaction: A review

IEEE Transactions on Pattern Analysis and Machine Intelligence, **19**, 677–695

The use of hand gestures is an attractive alternative to cumbersome interface devices for human-computer interaction (HCI). In particular, visual interpretation of hand gestures can help provide the ease and naturalness desired for HCI. This has motivated active research in computer vision-based analysis and interpretation of hand gestures. In our review of the literature on visual interpretation of hand gestures in the context of its role in HCI, we organize our discussion according to the method used for modeling, analyzing, and recognizing gestures. Important differences in approaches to gesture interpretation arise depending on whether a three-dimensional model or an image appearance model of the human hand is used. Three-dimensional hand models allow more elaborate modeling of hand gestures, but also lead to computational hurdles that have not been overcome, given the real-time requirements of HCI. Appearance-based models lead to computationally efficient "purposive" approaches that work well under constrained situations, but seem to lack the generality desirable for HCI. We also discuss implemented gestural systems as well as other potential applications of vision-based gesture recognition. Although the current progress is encouraging, further theoretical as well as computational advances are needed before gestures can be widely used for HCI. We discuss directions of future research in gesture recognition, including its integration with other natural modes of human-computer interaction.

Pelikan, M., Goldberg, D. E., & Cantú-Paz, E. (1999)

BOA: The Bayesian optimization algorithm

Tech. Rep. No. 99003, Urbana-Champaign: University of Illinois, Illinois Genetic Algorithms Laboratory

We propose an algorithm that uses an estimation of the joint distribution of promising solutions to generate new candidate solutions. The proposed algorithm, based on the concept of genetic algorithms, is called the Bayesian optimization algorithm (BOA). To estimate the distribution of promising solutions, the

algorithm exploits techniques for modeling multivariate data by Bayesian networks. The proposed algorithm identifies, reproduces, and mixes building blocks up to a specified order. It is independent of the ordering of the variables in the strings representing the solutions. Prior information about the problem can be incorporated into the algorithm, but it is not essential. Preliminary experiments show that as the problem size grows, the BOA outperforms the simple genetic algorithm, even on decomposable functions with tight building blocks.

Poddar, I., Sethi, Y., Ozyildiz, E., & Sharma, R. (1998)

Toward natural gesture/speech HCI: A case study of weather narration

Proceedings of the Workshop on Perceptual User Interfaces (PUI '98), 1–6

For human-computer interaction to be more natural, computers must be able to recognize continuous natural gestures and speech. To this end, previous researchers, using hidden Markov models (HMMs), have reported high recognition rates for gesture recognition; however, these gestures were defined precisely and were bound with syntactical and grammatical constraints. Natural gestures neither string together in syntactical bindings nor are amenable to strict classification. By recording the hand gestures and speech of a reporter standing before a weather map, we have studied the interaction between speech and gesture in the context of a display. We have implemented a continuous HMM-based gesture-recognition framework. To understand the interaction between gesture and speech, we conducted a co-occurrence analysis of different gestures with some spoken keywords. We also demonstrated the possibility of improving continuous gesture recognition results based on the co-occurrence analysis. Fast feature extraction and tracking is accomplished by the use of predictive Kalman filtering on a color-segmented stream of video images. The results in the weather domain should be a step toward a natural gesture-and-speech computer interface. [Abstract provided by ARL.]

Poddar, I., & Sharma, R. (1999, November)

Continuous recognition of natural hand gestures for human computer interaction

paper presented at *12th Annual ACM Symposium on User Interface Software and Technology (UIST '99)*, Asheville, NC

The use of hand gestures is an attractive alternative to cumbersome interface devices for human-computer interaction (HCI), particularly within a multimodal system, such as a speech and gesture interface. In particular, visual interpretation of hand gestures can help achieve the ease and naturalness desired for HCI. To exploit this potential, we need to develop recognition techniques that can handle continuous natural gesture inputs. Natural gestures are usually embedded in speech with no fixed, predefined meanings, and they do not string together in any syntactic bindings. In this paper, we propose techniques for the recognition of natural gestures that occur in the context of controlling and interacting with spatial maps through speech and gesture. We first present a study of a "parallel" domain using data from the weather narration in broadcast TV. This gives us a way to bootstrap the development of a gesture/speech system for interacting naturally with a graphical display of a spatial map.

Porkaew, K., Chakrabarti, K., & Mehrotra, S. (1999)

Query refinement for multimedia similarity retrieval in MARS

Proceedings of the 7th ACM International Multimedia Conference, 235–238

A new method for refining queries in the Multimedia Analysis and Retrieval System (MARS) was compared with a method already incorporated in MARS. The researchers posit a two-step process for multimedia searches. Users create initial queries by providing examples of objects similar to those they wish to retrieve; then, in a step called relevance feedback, they modify their queries by indicating which of the returned objects is most like the objects they seek. An object is represented as a collection of features, which in turn are represented by vectors in an object space. A query is represented as the sum of several object spaces. During the relevance feedback step, a clustering technique called *query expansion* is used to modify a query by identifying a set of objects to be added to the query representation. Experimental results show that query expansion significantly outperforms an older query modification technique in MARS (query point movement), both in terms of retrieval effectiveness and execution costs. [Abstract furnished by ARL.]

Porkaew, K., Mehrotra, S., & Ortega, M. (1999)

Query reformulation for content-based multimedia retrieval in MARS

IEEE International Conference on Multimedia Computing and Systems, **2**, 747–751

Unlike traditional database management systems, content-based multimedia retrieval databases make it difficult for a user to ask for information in a direct, precise query. A typical multimedia interface allows a query to be based on examples of objects similar to the ones users wish to retrieve. Such an interface, however, requires mechanisms for the system to learn the query representation from the examples. In this paper, we describe the query refinement framework implemented in the Multimedia Analysis and Retrieval System (MARS) for learning query representations using relevance feedback. The proposed framework uses a query expansion approach to modifying the query representation, in which relevant objects are added to the query. Furthermore, query reweighting techniques are used to adjust similarity functions.

Porkaew, K., Mehrotra, S., & Yu, H. (1999)

Continuous query in moving object databases to support efficient visualizations

Tech. Rep. No. TR-MARS-99-13, University of California, Irvine

Increasingly, application domains require database management systems to represent mobile objects and support motion-specific queries. An important type of query in such domains is a continuous query, which consists of a sequence of instantaneous queries, one for each point of time $t' > t$, where t is the time the query is initially posed to the database. An example of a continuous query is monitoring objects within a specified distance of an object x, which itself may be mobile, starting at a given time t. A naive approach to evaluating continuous queries is to repeatedly submit instantaneous queries to the database, one for each point of time $t' > t$. Since subsequent queries have a high degree of overlap with previous ones (because of the continuity of motion), much computation is wasted. This paper proposes two alternative mechanisms that attempt to reuse the answers returned by previous queries in evaluating subsequent queries,

thereby optimizing the evaluation of continuous queries. Experiments conducted over a real-life dataset consisting of mobile objects—AHAS data containing Army battle exercises—are used to validate the efficiency of the developed approaches.

Porkaew, K., Mehrotra, S., Ortega, M., & Chakrabarti, K. (1999)
Similarity search using multiple examples in MARS
Lecture Notes in Computer Science, **1614**, 68–75

Unlike traditional database management systems, content-based multimedia retrieval databases make it difficult for a user to ask for information in a direct, precise query. Typically, content-based retrieval systems allow users to ask for information using examples of objects similar to the ones they wish to retrieve. Such an interface, however, requires mechanisms for the system to learn the query representation from the examples provided by the user. In our previous work, we proposed a query refinement mechanism in which a query representation is modified by the addition of new relevant examples based on user feedback. In this paper, we describe query processing mechanisms that can efficiently support query expansion using multidimensional index structure.

Pringle, H. L., Kramer, A. F., Irwin, D. E., & Atchley, P. (1999)
Detecting changes in real-world scenes: The role of change characteristics and individual differences in attention
Proceedings of the 3rd Annual Federated Laboratory Symposium, 121–125

Recent research suggests that humans are surprisingly poor at detecting changes in scenes that occur during the course of eye movements. Indeed, this research has indicated that even large and apparently salient changes to scenes take a substantial amount of time to detect. In the present research, we examine the influence of several change characteristics (i.e., salience, meaning, and eccentricity) and individual differences in visual attention (i.e., the useful feld of view) on perceptual change detection in the context of detailed driving scenes. These data are discussed in terms of how displays might be designed to help users to rapidly and accurately detect task-relevant changes.

Raghavan, V., & Molineros, J. (1999, June)
Interactive evaluation of assembly sequences using augmented reality
IEEE Transactions on Robotics and Automation, **15**, 435–449

This paper describes an interactive tool for evaluating assembly sequences using the novel human-computer interface of augmented reality. The goal is to enable the user to consider various sequencing alternatives of the manufacturing design process by manipulating both virtual and real prototype components. The augmented reality-based assembly evaluation tool would allow a manufacturing engineer to interact with the assembly planner while manipulating the real and virtual prototype components in an assembly environment. Information from the assembly planner can be displayed superimposed directly on the real. A sensing technique is proposed that uses computer vision along with a system of markers for automatically monitoring the assembly state as the user manipulates the assembly components. An implemented system called AREAS (augmented reality system for evaluating assembly sequences) is described. Also discussed is

the advantage of using mixed prototyping and augmented reality as a means of capturing human intuition in assembly planning.

Rozenblit, J. W., Nugyen, H., & Barnes, M. J. (1999)
Effects of computer-displayed color characteristics on individuals
Proceedings of the 3rd Annual Federated Laboratory Symposium, 163

The Advanced Battlefield Architecture for Tactical Information Selection (ABATIS) is described: ABATIS is a means of presenting battlefield information that facilitates understanding the process of the battle rather than simply the current location of various forces. The design of this system would reflect how the user assimilates battlefield-state information into a process-centered viewpoint. A key concept in the design of ABATIS is the process-centered display (PCD), a construct that can display complex, evolutionary processes, as well as simple, repetitive changes. For PCD to be effective, its architecture must support dynamic change, since battlefield processes (e.g., maneuver, attack) evolve and change as the battle unfolds, and must be flexible enough to permit the quick creation of new battlespace objects from old ones. A secondary goal would be to use motion, color changes, morphing, or other types of animation to convey information. Some uses of animation are obvious, such as moving a symbol from one location to another. However, abstract quantities can also be tied to motion. A simple example would be representing the strength of a ground force by the speed of rotation of its symbol. When representation matches the intuitive notions of the user, the result is a metaphor that correlates familiar experiences with the actions of symbols. [This abstract was supplied by ARL and is based on a poster summary appearing in the conference proceedings.]

Rudmann, D. S., & McConkie, G. W. (1998, April 30–May 2)
Acquiring spatial knowledge under varying field of view sizes
paper presented at *Midwestern Psychological Association Seventieth Annual Meeting,* Chicago, IL

No abstract available.

Rudmann, D. S., & McConkie, G. W. (1999)
Eye movements in human-computer interaction
Proceedings of the 3rd Annual Federated Laboratory Symposium, 91–95

The potential benefits of incorporating eye movements into the interaction between humans and computers are numerous. For example, knowing the location of a user's gaze may help a computer to interpret a user's request, aid natural language processing, speed up interaction by allowing the eyes to serve as a pointing device, and possibly enable a computer to ascertain some cognitive states of the user, such as confusion or fatigue. This paper details the problems encountered in previous attempts to use eye movements in human-computer interaction, and evaluates current technology for its ability to overcome these limitations. An assessment of the accuracy and reliability of the ISCAN eye-tracking system (manufactured by Iscan, Inc.) and the pcBird head tracker (manufactured by Ascension Technology) is provided for two-dimensional displays. Recommendations are made for the design of eye-controlled display systems based on these technologies.

Rui, Y., Huang, T. S., & Chang, S.-F. (1998)

Digital image/video library and MPEG-7: Standardization and research issues

Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, **6,** 3785–3788

Much research activity and interest has emerged in two closely related areas: the digital image/video library (DIVL) and MPEG-7. We review the critical research issues in DIVL from a signal processing viewpoint, the objectives and scope of MPEG-7, and the relationships between these two.

Rui, Y., Huang, T. S., & Mehrotra, S. (1998)

Browsing and retrieving video content in a unified framework

IEEE Second Workshop on Multimedia Signal Processing, 9–14

In this paper, we first review the recent research progress in video analysis, representation, browsing, and retrieval. Motivated by the standard mechanisms for accessing book content (that is, tables of contents and indexes) we then present novel techniques for accessing video content by constructing video equivalents. We further explore the relationship between video browsing and retrieval, and propose a unified framework to seamlessly incorporate both entities. Preliminary research results justify our proposed framework for providing access to videos based on their content.

Rui, Y., Huang, T. S., & Mehrotra, S. (1998)

Exploring video structure beyond the shots

Proceedings of the IEEE International Conference on Multimedia Computing and Systems, 237–240

While existing shot-based video analysis approaches provide users with better access to the video than do raw data streams, they are still not sufficient for meaningful video browsing and retrieval, since (1) the shots in a long video are still too numerous to be presented to the user, and (2) shots do not capture the underlying semantic structure of the video, the basis upon which the user may wish to browse/retrieve the video. To explore video structure at the semantic level, this paper presents an effective approach to extracting the underlying video scene structure and grouping shots into semantically related scenes. The output of the proposed algorithm provides a structured video that greatly facilitates the user's access. Experiments based on real-world movie videos validate the effectiveness of the proposed approach.

Rui, Y., Huang, T. S., Ortega, M., & Mehrotra, S. (1998)

Relevance feedback: A power tool for interactive content-based image retrieval

IEEE Transactions on Circuits and Systems for Video Technology, **8,** 644–655

Content-based image retrieval (CBIR) has become a highly active research area in the past few years. Many visual feature representations have been explored and many systems built. While these research efforts establish the basis of CBIR, the usefulness of the proposed approaches is limited. Specifically, these efforts have generally ignored two distinct characteristics of CBIR systems: (1) the gap between high-level concepts and low-level features, and (2) the subjectivity of human perception of visual content. This paper proposes an interactive retrieval approach, based on relevance feedback, which effectively takes into account the above two characteristics in CBIR. During the retrieval process, the user's high-

level query and perception subjectivity are captured by dynamically updated weights based on the user's feedback. Experimental results for more than 70,000 images show that the proposed approach greatly reduces the user's effort in composing a query, and captures the user's information need more precisely.

Schlabach, J. L., Goldberg, D. L., Hayes, C. C. (1999)

Fox-GA: A genetic algorithm for generating and analyzing battlefield courses of action

Evolutionary Computation, **7**, 45–68

This paper describes Fox-GA, a genetic algorithm (GA) that generates and evaluates plans in the complex domain of military maneuver planning. Fox-GA's contributions are to demonstrate an effective application of GA technology to a complex, real-world planning problem, and to provide an understanding of the properties needed in a GA solution to meet the challenges of decision support in complex domains. Previous obstacles to applying GA technology to maneuver planning include the lack of efficient algorithms for determining the fitness of plans. Detailed simulations would ideally be used to evaluate these plans, but most such simulations typically require several hours to assess a single plan. Since a GA needs to quickly generate and evaluate thousands of plans, these methods are too slow. To solve this problem, we developed an efficient evaluator (wargamer) that uses coarse-grained representations of this problem domain to allow appropriate yet intelligent tradeoffs between computational efficiency and accuracy. An additional challenge was that users needed a set of significantly different plan options from which to choose. Typical GAs tend to develop a group of "best" solutions that may be very similar (or identical) to each other. This may not provide users with sufficient choice. We addressed this problem by adding a niching strategy to the selection mechanism to ensure diversity in the solution set, providing users with a more satisfactory range of choices. Fox-GA's impact will be in providing decision support to constrained and cognitively overloaded battlestaff to help them rapidly explore options, create plans, and better cope with the information demands of modern warfare.

Servetto, S. D., Ramchandran, K., & Orchard, M. T. (1999)

Image coding based on a morphological representation of wavelet data

IEEE Transactions on Image Processing, **8**, 1161–1174

An experimental study of the statistical properties of wavelet coefficients of image data is presented, as well as the design of two different morphology-based image-coding algorithms that use these statistics. A salient feature of the proposed methods is that, by a simple change of quantizers, the same basic algorithm yields high-performance embedded or fixed-rate coders. Another important feature is that the shape information of morphological sets used in this coder is encoded implicitly by the values of wavelet coefficients, thus avoiding the use of explicit and rate-expensive shape descriptors. These proposed algorithms, while achieving nearly the same objective performance as state-of-the-art zerotree-based methods, can produce reconstructions of a somewhat superior perceptual quality, because they exhibit a property of compression and noise reduction.

Servetto, S. D., Rui, Y., Ramchandran, K., & Huang, T. S. (1999)

A region-based representation of images in MARS

Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology, **20,** 137–150

We study the problem of representing images within a multimedia Database Management System (DBMS) to support fast retrieval operations without compromising storage efficiency. To achieve this goal, we propose new image-coding techniques that combine a wavelet representation, embedded coding of the wavelet coefficients, and segmentation of image-domain regions in the wavelet domain. A bit stream is generated in which each image region is encoded independently of other regions, without the need to store information describing the regions. Simulation results show that our proposed algorithms achieve coding performance that compares favorably, both perceptually and objectively, to that achieved by state-of-the-art image/video coding techniques, while additionally providing region-based support.

Sethi, Y. (1998)

Multimodal analysis of gesture and speech in video sequences

unpublished master's thesis, The Pennsylvania State University, University Park, PA

A gesture recognition system, based on hidden Markov modeling, was developed to make possible machine recognition of gestures, and thus enable more natural human-computer interaction. A hidden Markov model was trained to recognize a few natural gestures produced by weather reporters during weather forecasts, and then validated on television weathercasts. Gesture recognition was highly accurate (100 percent) with discrete gestures isolated from a continuous stream of gestures, but less accurate (about 56 percent) when the targeted gestures were part of a stream of movements. Accuracy for streamed gestures increased (by about 12 percent) when a speech recognition system was used in combination with the gesture-recognition system to detect words that co-occurred with the targeted gestures.

Sharma, R., & Hutchinson, S. (1997)

Motion perceptibility and its application to active vision-based servo control

IEEE Transactions on Robotics and Automation, **13**, 607–617

We address the ability of a computer vision system to perceive the motion of an object (possibly a robot manipulator) in its field of view. We derive a quantitative measure of motion perceptibility, which relates the magnitude of the rate of change in an object's position to the magnitude of the rate of change in the image of that object. We then show how motion perceptibility can be combined with the traditional notion of manipulability into a composite perceptibility/manipulability measure. We demonstrate how this composite measure can be applied to a number of different problems involving relative hand/eye positioning and control.

Sharma, R., Pavlovic, V., & Huang, T. S. (1998)

Toward multimodal human computer interaction

Proceedings of the IEEE, **86**, 853–869

Recent advances in various signal processing technologies, coupled with an explosion in available computing power, have given rise to a number of novel

human-computer interaction (HCI) modalities: speech, vision-based gesture recognition, eye tracking, electroencephalograph, etc. Successful incorporation of these modalities into an interface could potentially ease the HCI bottleneck that has become noticeable with the advances in computing and communication. It has also become increasingly evident that the difficulties encountered in the analysis and interpretation of individual sensing modalities may be overcome by their integration into a multimodal human-computer interface. We examine several promising approaches to achieving multimodal HCI. We consider some of the emerging novel input modalities for HCI and the fundamental issues in integrating them at various levels, from early signal level to intermediate feature level to late decision level. We discuss the different computational approaches that may be applied at the different levels of modality integration. We also briefly review several demonstrated multimodal HCI systems and applications. Despite all the recent developments, it is clear that further research is needed for interpreting and fitting multiple sensing modalities in the context of HCI. This research can benefit from many disparate fields of study that increase our understanding of the different human communication modalities and their potential role in HCI.

Sharma, R., Poddar, I., Ozyildiz, E., Kettebekov, S., Kim, H., & Huang, T. S. (1999)
Toward interpretation of natural speech/gesture: Spatial planning on a virtual map
Proceedings of the 3rd Annual Federated Laboratory Symposium, 35–39

Hand gestures and speech are the most important modalities of human-to-human interaction. Accordingly, there is considerable interest in incorporating these modalities into "natural" human-computer interaction (HCI), particularly within virtual environments. An important feature of such a natural interface would be an absence of predefined speech and gesture commands. The resulting bimodal speech/gesture HCI "language" would thus have to be interpreted by the computer. While some progress has been made in the natural language processing of speech, the inclusion of gestures is even more challenging. This challenge ranges from the low-level signal processing of bimodal (audio/video) input to the high-level semantic interpretation of natural speech/gesture. In this paper, we consider the design of a speech/gesture interface in the context of a set of spatial tasks defined on a virtual map of an urban area. The task constraints then make it feasible to study the critical components of the bimodal interpretation problem and define an agent-based architecture for implementing the interface. An experimental test bed is also described, where free hand gestures and spoken words are used for spatial planning tasks defined on a virtual two-dimensional map. Such tasks would also be involved in crisis management, mission planning, and briefing.

Sistla, A. P., Wolfson, O., Chamberlain, S., & Dao, S. (1998)
Querying the uncertain position of moving objects
in O. Etizon, S. Jajodia, & S. Sripada (Eds.), *Temporal Databases: Research and Practice* (pp 310–337). Berlin, Germany: Springer-Verlag

The authors propose a data model for representing moving objects with uncertain positions in database systems: the Moving Objects Spatio-Temporal (MOST) data model. They also propose Future Temporal Logic (FTL) as the query lan-

guage for the MOST model, and devise an algorithm for processing FTL queries in MOST.

Sistla, A. P., Wolfson O., & Huang, Y. (1998)

Minimization of communication cost through caching in mobile environments

IEEE Transactions on Parallel and Distributed Systems, **9**, 378–390

Users of mobile computers will soon have online access to a large number of databases via wireless networks. Because of limited bandwidth, wireless communication is more expensive than wire communication. In this paper, we present and analyze various static and dynamic data allocation methods. The objective is to minimize the communication cost between a mobile computer and the stationary computer that stores the online database. Analysis is performed on two cost models. One is connection (or time) based (as in cellular telephones), where the user is charged per minute of connection. The other is message based (as in packet radio networks), where the user is charged per message. Our analysis addresses both the average case and the worst case for determining the best allocation method.

Sistla, A. P., Wolfson, O., Yesha, Y., & Sloan, R. H. (1998)

Towards a theory of cost management for digital libraries and electronic commerce

ACM Transactions on Database Systems, **23**, 411–452

One feature that distinguishes digital libraries from traditional databases is new cost models for client access to intellectual property. Clients will pay to access data items in digital libraries, and we believe that optimizing these costs will be as important as optimizing performance for traditional databases. We discuss cost models and protocols for accessing digital libraries, with the objective of determining the minimum cost protocol for each model. We expect that in the future, information appliances will come equipped with a cost optimizer, in the same way that computers today come with a built-in operating system. We make the initial steps toward a theory and practice of intellectual property cost management.

Sundareswaran, V., & Chen, S. (1999)

Hand-held displays for control and communication with large-format displays

Proceedings of the 3rd Annual Federated Laboratory Symposium, 165

In a demonstration, a hand-held personal computer (HPC) was used to control the view on a large display. The large display on a desktop computer (Windows 95, DirectX) showed only a portion of a prerendered isometric view of a battlefield. Animated units were controlled through a stylus, a graphics tablet, and the HPC. Troop movement and identified red-unit positions were displayed on both the large and the HPC displays. Control was achieved through stylus interaction and speech commands in a multimodal fashion. Control of wireless integrated network sensors is demonstrated, as well as a display of the situation reported by the sensors. [This abstract was supplied by ARL and is based on a poster summary appearing in the conference proceedings.]

Sniezek, J. A., & Chernyshenko, O. S. (1999)

Psychological evaluation of Co-RAVEN technology for battlefield decision making: Probabilistic reasoning by Army intelligence experts

Tech. Rep. No. 99-1, Urbana-Champaign, IL: University of Illinois, Department of Psychology

Co-RAVEN is a Bayesian-based decision aid that generates probabilities of the occurrence of high-level events from detailed data. The Bayesian decision net is created by the encoding of probability statements from actual military intelligence experts on real-world intelligence problems. However, a common finding of decision research is that decision makers are often overconfident in their judgments. This paper found that intelligence officers exhibit overconfidence in their decisions and that they do not agree in their probability estimates. The implications of these findings for creating Bayesian decision nets was discussed.

Tang, H., & Beebe, D. J. (1999)

An ultra-flexible electrotactile display for the roof of the mouth

Proceedings of the First Joint BMES/EMBS Conference, **1**, 626

No abstract available

Tang, H., & Beebe, D. J. (1999)

Tactile sensitivity of the tongue on photolithographically fabricated patterns

Proceedings of the 3rd Annual Federated Laboratory Symposium, 167

Previous psychology and neuroscience studies suggest that the oral cavity is a sensory-rich location. Recent advances in miniaturization technologies make it possible to build tactile devices that can operate within the oral cavity. In order to design optimal tactile interfaces for the mouth, we must understand the perceptual characteristics of the mouth. This report describes preliminary work aimed at measuring several perceptual parameters of the tongue's tip and anterior dorsal surface. [This abstract was supplied by ARL and is based on a poster summary appearing in the conference proceedings.]

Tao, H., & Huang, T. S. (1999)

Facial motion synthesis and analysis using a free-form deformation model

Proceedings of the 3rd Annual Federated Laboratory Symposium, 169

Capturing real facial motions from video sequences is a powerful approach for automatic generation of a facial deformation model. In this paper, a three-dimensional piecewise Bézier volume deformation (PBVD) model is proposed for both facial animation and facial motion analysis. Because this model is independent of the mesh structure, the resulting deformation model can be used for animating different geometric face models. The more important linear property of this model also implies an efficient and robust analysis algorithm, from which a customized facial deformation model can be derived. Experimental results on facial animation and video analysis are demonstrated.

Tao, H., & Huang, T. S. (1998)

Bézier volume deformation model for facial animation and video tracking

in N. Magnenat-Thalmann and D. Thalmann (Eds.), *Modeling and Motion Capture Techniques for Virtual Environments* (pp 242–253). Berlin, Germany: Springer-Verlag

Capturing real motions from video sequences is a powerful approach for automatically building a facial deformation model. In this paper, a three-dimensional Bézier volume deformation model is proposed for both synthesis and analysis of facial movements. Since this model is independent of the mesh structure (provided that the feature points are given), it can animate geometric facial models of different shapes and structures. Of equal importance, the linear property of this model implies a simple and robust analysis algorithm, from which a customized facial deformation model is derived. Experimental results on animation and video analysis are demonstrated.

Tayeb, J., Ulusoy, O., & Wolfson, O. (1998)

A quadtree-based dynamic attribute indexing method

Computer Journal, **41**, 185–200

Dynamic attributes are those that change continuously over time, making it impractical for explicit updates to be issued for every change. In this paper, we adapt a variant of the quadtree structure to solve the problem of indexing dynamic attributes. The approach is based on the key idea of using a linear function of time for each dynamic attribute, so that we can predict its value in the future. We contribute an algorithm for regenerating the quadtree-based index periodically, which minimizes CPU and disk access costs. We also provide an experimental study of performance, focusing on query processing and index update overheads.

Uckun, S., Tuvi, S., Winterbottom, R., & Donohue, P. (1999)

OWL: A decision-analytic wargaming tool

Proceedings of the 3rd Annual Federated Laboratory Symposium, 133–137

OWL is a decision-analytic wargamer that is used to evaluate the benefits and risks of multiple friendly courses of action. OWL uses stochastic simulation principles to evaluate alternative outcomes of a battle, given uncertainty in the information available about friendly forces, the enemy, mission, weather, and the terrain. OWL is designed as a postprocessor for Fox, a tool that evaluates thousands of potential courses of action and selects a small number of plausible ones.

Theeuwes, J., Kramer, A. F., & Atchley, P. (1998)

Attentional control within 3-D space

Journal of Experimental Psychology: Human Perception and Performance, **24**, 1476–1485

Four experiments investigated whether directing attention to a particular plane in depth enables observers to filter out information from another depth plane. Observers searched for a red line segment among green line segments in stereoscopic displays. Results showed that directing attention to a particular depth plane cannot prevent attentional capture from another depth plane when the colors of the target and distractor are identical. However, attentional capture by a singleton from another depth plane is prevented when the colors of the target and distractor are different. These results indicate that only when both color and

depth information are selective in guiding attention to the target singleton can attentional capture by irrelevant singletons be prevented. The results also suggest that retinal disparity does not have the same special status as location information in two dimensions and should be considered just another feature along which selection may occur.

Theeuwes, J., Kramer, A. F., & Atchley, P. (1998)

Visual marking of old objects

Psychonomic Bulletin and Review, **5**, 130–134

D. G. Watson and G. W. Humphreys presented evidence that selection of new elements can be prioritized by on-line, top-down attentional inhibition of old stimuli already in the visual field (visual marking). The experiments on which this evidence was based always presented old elements in green and new elements in blue; selection, therefore, could have been based on color. The present experiment, which does not contain this confound, showed that visual marking is a strong and robust process that enables subjects to visually mark at least 15 old elements, even when these elements are the same color as the new ones. The results indicate that preview of the elements is critical—not the fact that those elements contained a common feature.

Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998)

Our eyes do not always go where we want them to go: Capture of the eyes by new objects

Psychological Science, **9**, 379–385

Observers make rapid eye movements to examine the world around them. Before an eye movement is made, the observer's attention covertly shifts to the location of the object of interest. The eyes will typically land at the position at which attention is directed. Here the authors report that a goal-directed eye movement toward a uniquely colored object is disrupted by the appearance of a new, but task-irrelevant, object, unless subjects ($n = 15$) have enough time to focus their attention on the location of the target before the appearance of the new object. In many instances, the eyes started moving toward the new object before gaze started to shift to the color-singleton target. The eyes often landed for a very short period of time (25 to 150 ms) near the new object. The results suggest parallel programming of two saccades: one voluntary, goal-directed eye movement toward the color-singleton target and one stimulus-driven eye movement reflexively elicited by the appearance of the new object. Neuroanatomical structures responsible for parallel programming of saccades are discussed.

Vassiliou, M. S., Sundareswaran, V., Chen, S., & Wang, K. (1999)

Multimodal HCI integration

Society of Automotive Engineers, 1999 World Aviation Congress, Report No. 99WAC-149

A multipurpose test bed for integrating user interface and sensor technologies has been developed, based on a client-server architecture. Various interaction modalities (speech recognition, three-dimensional audio, pointing, wireless handheld-PC-based control and interaction, sensor interaction, etc.) are implemented as servers, encapsulating and exposing commercial and research software packages. The system allows users to interact with large and small displays

using speech commands as well as pointing, spatialized audio, and other modalities. Simultaneous and independent speech recognition for two users is supported; users may be equipped with conventional acoustic or new body-coupled microphones.

Weber, T. A., Kramer, A. F., & Miller, G. A. (1997)

Selective processing of superimposed objects: An electrophysiological analysis of object-based attentional selection

Biological Psychology, **45** (1–3), 159–142

A study investigated whether object-based attentional selection occurs from grouped-array or spatially invariant representations. Eighteen college students were presented with colored objects and asked to judge whether a particular color/shape conjunction was present, regardless of whether the color and shape were part of a single object (same-object condition) or occurred on two different objects (different-object condition). Reaction times (RTs) and accuracies were recorded for subjects' judgments. Event-related brain potential components, in particular the P1 and N1, were elicited both from the presentation of the target objects and from a postdisplay probe that was used as an index of spatial attention. Consistent with predictions of object-based selection models, RTs and accuracies were faster on same- than on different-object trials. N1s elicited by the target objects and P1s elicited by the postdisplay probes discriminated between same and different object trials when the two target objects were superimposed. These data are consistent with the proposal that object-based selection is spatially mediated, even for partially overlapping objects. The data are discussed in terms of space- and object-based models of visual selective attention.

Wickens, C. D., Pringle, H. L., & Merlo, J. (1999)

Integration of information sources of varying weights: The effect of display features and attention cueing

Tech. Rep. No. ARL-99-2/FEDLAB-99-1, Savoy, Illinois: University of Illinois at Urbana-Champaign, Aviation Research Laboratory Institute of Aviation

This report reviews research in which multiple sources of variable reliability information are integrated for making diagnostic judgments or allocating resources. A framework for considering these experiments is provided, and some evidence is presented regarding the extent to which humans are "calibrated" in allocating processing proportionately to the ideal weights (i.e., reliability or importance) of information channels. Two generic sources of bias are identified. Attentional biases occur when more processing is given to less important channels, at the expense of more important ones (i.e., a failure to allocate attention optimally). Trust biases occur when less than fully reliable information is offered more processing than is warranted (i.e., "overtrust"). The report also reviews and integrates the conclusion from a smaller number of specific studies that examined how multisource information processing is modulated by properties of the display of those sources. Two sources of display information are considered: attentional guidance (e.g., cueing), which directs attention to certain regions of the display, and reliability guidance, which explicitly displays the level of reliability of the information sources. Each type of information can induce the appropriate behavior from the user, either explicitly (e.g., by highlighting the

important feature) or implicitly (by placing the important feature in the center of the display). Generalizations regarding the effectiveness of these display features are sought from the studies reviewed.

Wickens, C. D., Thomas, L., Merlo, J., & Hah, S. (1999)

Immersion and battlefield visualization: Does it influence cognitive tunneling?

Proceedings of the 3rd Annual Federated Laboratory Symposium, 111–115

Thirty officers at the U.S. Military Academy participated in a study in which a three-dimensional exocentric display was compared with a three-dimensional immersed display, as a means of supporting situation awareness regarding an evolving battlefield scenario. The immersed viewpoint allowed 360° panning and was coupled with a small plan view inset. A series of questions were asked on successive scenes as the movement to contact progressed. Results revealed that users of the immersed display demonstrated a form of "cognitive tunneling" in which they were overly influenced by information in the initially presented forward view, failing to adequately pan views behind them. The data speak to the advantage of three-dimensional exocentric displays.

Wilkins, D. C., Mengshoel, O, J., Chernyshenko, O., Jones, P. M., Hayes, C. C., & Bargar, R. (1999)

Collaborative decision making and intelligent reasoning in Judge Advisor Systems

Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences, 1–9

This paper examines the Raven and CoRaven decision-making tools, which are used to filter, interpret, and visualize large amounts of uncertain data. Raven and CoRaven are multimodal advisory decision aids that base their inferential reasoning on Bayesian networks. Human decision makers and information sources interact with these decision-making systems in many ways during their design, construction, refinement, and use. The collaborative aspects of using Raven and CoRaven are analyzed with the Judge Advisor System model.

Wolfson, O. (Ed.) (1997)

Data management issues in mobile computing

Mobile Networks and Applications, **2** [Special section 1]

This special section contains four articles that address some of the most important issues in adapting databases to a mobile computing environment.

Wolfson, O., Chamberlain, S., Dao, S., & Jiang, L. (1997, October)

Location management in moving objects databases

paper presented at *Second International Workshop on Satellite-Based Information Services,* Budapest, Hungary

The authors first introduce moving-objects databases and their related research problems; they then concentrate on a particular problem, namely, reducing the information cost associated with a trip taken by a moving vehicle. The information cost of a trip consists of the overhead of position-update messages, average uncertainty, and the deviation of the database position from the actual position of the object. Three position update policies are introduced: immediate linear policy

(ILP), plain dead reckoning (PDR), and adaptive dead reckoning (ADR). ADR is shown to have a lower information cost than PDR.

Wolfson, O., & Huang, Y. (1998)
Competitive analysis of caching in distributed databases
IEEE Transactions on Parallel and Distributed Systems, **9**, 391–409

The contributions of two models to distributed databases are described. The first is a model for evaluating the performance of data allocation and replication algorithms in distributed databases. The model is comprehensive in the sense that it accounts for I/O and communication costs and, because of reliability considerations, it accounts for limits on the minimum number of copies of the object. The model captures existing replica-management algorithms, such as read-one-write-all, quorum-consensus, etc. These algorithms are static in the sense that in the absence of failures, the copies of each object are allocated to a fixed set of processors. The second model is concerned with the fact that in modern distributed databases (particularly in mobile computing environments), processors dynamically store and relinquish objects in their local database. An algorithm is introduced for automatic dynamic allocation of replicas to processors. Using the new model, the authors compare the performance of the traditional read-one-write-all static allocation algorithm with the performance of the dynamic allocation algorithm. The relationship between the communication cost and I/O cost for static allocation is superior to that for dynamic allocation. [Abstract provided by ARL.]

Wolfson, O., Lelescu, A., & Xu, B. (1999, September)
Retrieval of collaborative work from multimedia databases using relevance feedback
paper presented at *Symposium on String Processing and Information Retrieval,* Cancun, Mexico

In this paper we address the problem of retrieving stored multimedia presentations using relevance feedback. We model multimedia presentations using a crisp relational or object-oriented database, augmented with a text attribute. We also introduce a language for retrieval by content from such databases. The language is based on fuzzy logic. We also introduce a method for query refinement that uses relevance feedback provided by the user.

Wolfson, O., Sistla, P., Xu, B., Zhou, J., Chamberlain, S., Yesha, Y., & Rishe, N. (1999)
Tracking moving objects using database technology in DOMINO
Lecture Notes in Computer Science, **1649**, 112–120

Methods are discussed for overcoming the limitations of computerized database management systems (DBMSs) when they contain information about moving ground or air vehicles. DBMs have problems managing large amounts of continuously changing data (e.g., changes in the location of a large number of vehicles), representing spatial data (e.g., vehicles near a common destination), and dealing with imprecise information on a vehicle's location. The authors discuss how their database for moving objects (DOMINO) project will resolve these issues. [Abstract provided by ARL.]

Wolfson, O., Xu, B., Chamberlain, S., & Jiang, L. (1998)

Moving objects databases: Issues and solutions

Proceedings of the 10th International Conference on Scientific and Statistical Database Management, 111–122

The authors report on research into the tracking of moving objects and their locations in a database, such as the location of moving taxicabs in a city. Currently, moving-objects database applications are being developed in an ad hoc fashion. Database management system (DBMS) technology provides a potential foundation upon which to develop these applications; however, DBMSs are currently not used for this purpose because a critical set of capabilities needed by moving-objects database applications is lacking in existing DBMSs. The objective of the current project, called DOMINO (databases for moving objects), is to build an envelope containing these capabilities on top of existing DBMSs. Problems and proposed solutions are discussed. [Abstract supplied by ARL.]

Wolfson, O., Xu, B., Chamberlain, S., & Jiang, L. (1998)

Challenges and approaches in motion databases

Proceedings of the 14th International Conference on Advanced Science and Technology, 182–194

Abstract not available.

Wright, S. (1999)

Effects of computer-displayed color characteristics on individuals

Proceedings of the 3rd Annual Federated Laboratory Symposium, 171

Fifty participants subjectively rated five-color samples for pleasantness, arousal, and dominance. Twenty-five different color samples were used, based on combinations of five different hues (blue, green, red, yellow, and purple), three saturation levels (low, medium, and high) and three brightness levels (low, medium, and high). These combinations were varied in a methodical manner along a predetermined scale. The colors were specified by RGB (red, green, blue) values, HSV (hue, saturation, value) values, and Munsell notation. Based upon results from these ratings, numeric models were developed through regression analysis to predict the pleasantness and arousal levels of screen background colors based on the color's characteristics. These models may be used to determine choice of background and foreground colors for information displays that require the user to experience a predetermined level of arousal. [This abstract was supplied by ARL and is based on a poster summary appearing in the conference proceedings.]

Wright, S. (1999)

The impact of color characteristics on visual search patterns

Proceedings of the 3rd Annual Federated Laboratory Symposium, 173

The primary purpose of this research was to determine whether visual search patterns were affected by different color combinations of brightness and saturation. A secondary purpose was to determine whether individuals usually start their scanning of graphic information in the same position. Colors with high levels of brightness and saturation were expected to draw the eye, thus changing the visual search pattern. An eye scanner was used to examine the search patterns of 15 subjects while they scanned an array of 16 variously colored icons for

a previously designated icon. Results show that eye-scanning patterns do change based upon the color combinations of surrounding icons. Results from this experiment should influence the color characteristics of icons and symbols requiring immediate attention on a display. [This abstract was supplied by ARL and is based on a poster summary appearing in the conference proceedings.]

Wu, Y., & Huang, T. (1999)

Capturing articulated human hand motion: A divide-and-conquer approach

Proceedings of the Seventh IEEE International Conference on Computer Vision, **1**, 606–611

The use of the human hand as a natural computer interface device has inspired research in the modeling, analyzing, and capturing of the motion of the articulated hand. Model-based hand-motion capturing can be formulated as a large nonlinear programming problem, but this approach is plagued by local minima. An alternative is to use analysis by synthesis in searching a huge space, but the results are inexact and the computation expensive. In this paper, articulated hand motion and finger motion are decoupled, and a new two-step iterative model-based algorithm is proposed to capture articulated human hand motion. A proof of convergence of this iterative algorithm is given. In our proposed work, the decoupled global hand motion and local finger motion are parameterized by the three-dimensional hand pose and the state of the hand. Hand pose determination is formulated as a least median of squares (LMS) problem rather than the nonrobust least squares (LS) problem, so that three-dimensional hand pose can be reliably calculated even if there are outliers. Local finger motion is formulated as an inverse kinematics problem. A genetic-algorithm-based method is proposed as an effective method of finding a sub-optimal solution to the inverse kinematics problem. Our algorithm and the LS-based algorithm are compared in several experiments. Both algorithms converge when local finger motion between consecutive frames is small. When large finger motion is present, the LS-based method fails, but our algorithm can still successfully estimate the global and local finger motion.

Yeh, M., & Wickens, C. D. (1999)

Visual search and target cueing with augmented reality: A comparison of head-mounted with hand-held displays

Proceedings of the 3rd Annual Federated Laboratory Symposium, 105–109

We conducted a study to determine the effects of target cueing and conformality with a hand-held or head-mounted display on visual search tasks requiring focused and divided attention. Eleven military subjects were asked to detect, identify, and give azimuth information for targets hidden in terrain presented in a simulated far domain environment, while concurrently monitoring a nearby domain using either a helmet-mounted or hand-held display. The results showed that the presence of cueing aided the target detection task for expected targets but drew attention away from unexpected targets in the environment. This effect was reduced when subjects used the hand-held display. Additionally, the results showed that the presence of cueing hindered performance on the secondary task.

Yeh, M., Wickens, C. D., & Seagull, F. J. (1999)

Conformality and target cueing: Presentation of symbology in augmented reality

Proceedings of the 42nd Annual Meeting of the Human Factors and Ergonomics Society, **2**, 1526–1530

We conducted a study examining several issues in the design of see-through helmet-mounted displays (HMDs) to determine their effects on tasks of focused and divided attention. These issues are frame of reference (world referenced versus screen referenced), target expectancy, target cueing, and viewing condition (i.e., one eye versus two). Sixteen subjects (eight civilian, eight military) were asked to detect, identify, and give azimuth information for targets hidden in terrain presented in the far domain (i.e., the world) while performing a monitoring task in the near domain (i.e., the display). The results showed that the presence of cueing aided target detection for expected targets, but drew attention away from unexpected targets in the environment. However, analyses support the observation that this effect can be mitigated by the use of world-referenced symbology. Displaying symbology to two eyes provided a slight benefit for target detection when the target was cued.

Yu, H., Mehrotra, S., Winkler, R., Ho, S. S., Gregory, T. C., & Allen, S. D. (1999)

Integration of SATURN system and VGIS

Proceedings of the 3rd Annual Federated Laboratory Symposium, 59–63

The Spatiotemporal Uncertainty Reasoning (SATURN) system, currently under development, is being integrated with the Virtual Geographic Information System (VGIS) system in an effort to improve VGIS performance and scalability to complex dynamic environments, as well as to enhance its functionality as a collaborative planning tool. We added three new components to VGIS—a spatiotemporal object manager, a performance monitor, and a task database. The spatiotemporal object manager uses SATURN techniques for indexing dynamic multidimensional (spatiotemporal) objects to support effective and efficient object traversal during visualization. The performance monitor adjusts the resource allocation between VGIS components and adaptively adjusts image quality to guarantee bounded visualization performance. The task database extends VGIS as a tool for collaborative planning. Performance results illustrate that the SATURN techniques for object management and the performance monitor significantly improve VGIS performance, allowing it to scale to complex scenarios with a large number of dynamic objects.

Zeller, M., Phillips, J. C., Dalke, A., Humphrey, W., Schulten, K., Sharma, R., Huang, T. S., Pavlovic, V. I., Zhao, Y., Lo, Z., & Chu, S. (1997)

A visual computing environment for very large scale biomolecular modeling

IEEE International Conference on Application-Specific Systems, Architectures and Processors, 3–12

Knowledge of the complex molecular structures of living cells is being accumulated at a tremendous rate. Key technologies enabling this success have been high-performance computing and powerful molecular graphics applications; however, the technology is beginning to lag in the face of challenges posed by the size and number of new structures and by the emerging opportunities in drug design and genetic engineering. For interactive modeling of biopolymers, a visual computing environment is being developed that links a three-dimensional (3-D) molecular graphics program with an efficient molecular dynamics simula-

tion program executed on remote high-performance parallel computers. The system will be ideally suited for distributed computing environments, because it uses both local 3-D graphics facilities and the peak capacity of high-performance computers for interactive biomolecular modeling. For creating an interactive 3-D environment, various input methods are possible. Three are explored: (1) a six-degree-of-freedom "mouse" for controlling the space shared by the model and the user, (2) voice commands monitored through a microphone and recognized by a speech recognition interface, and (3) hand gestures, detected through cameras and interpreted with computer vision techniques. Controlling 3-D graphics connected to real-time simulations and using voice with suitable language semantics, as well as hand gestures, promise great benefits for many types of problem-solving environments. Our focus on structural biology takes advantage of existing sophisticated software, provides concrete objectives, defines a well-posed domain of tasks, and offers a well-developed vocabulary for spoken communication.

Zeller, M., Schulten, K., & Sharma, R. (1997)

Learning the perceptual control manifold for sensor-based robot path planning

Proceedings of the IEEE International Symposium on Computational Intelligence in Robotics and Automation, 48–53

The perceptual control manifold is a concept that extends the notion of the robot configuration space to include sensor feedback for robot motion planning. In this paper, we propose a framework for sensor-based robot motion planning that uses the topology-representing network algorithm to develop a learned representation of the perceptual control manifold. The topology-preserving features of the neural network lend themselves to yield, after learning, a diffusion-based path-planning strategy for flexible obstacle avoidance. Simulations on path control and flexible obstacle avoidance demonstrate the feasibility of this approach for motion planning and illustrate the potential for further robotic applications.

Zhuang, Y., Rui, Y., Huang, T. S., & Mehrotra, S. (1998)

Applying semantic association to support content-based video retrieval

Fifth Very Low Bit-Rate Video Workshop (pp 45–48), University of Illinois, Urbana-Champaign

In the traditional approach to video retrieval, queries base their search on textual information (titles and keywords) annotated to the video. Since automated annotation is not yet available, generating keyword descriptors requires a great amount of labor and has proved to be unrealistic in applications. An approach that seems to be at the other extreme is using the low-level video content, such as color, texture, shape, and motion features, in an attempt to eliminate the necessity of keyword annotation. A preferable query form should include both key words and video content. In this paper, we explore the semantic aspect based on video table of contents structuring. Closed-captioning is used to extract a basic keyword set. Word-Net, an electronic lexical system, is used to provide semantic association. The approach has shown that retrieval performance is greatly improved.

Zhuang, Y., Rui, Y., Huang, T. S., & Mehrotra, S. (1998)

Adaptive key frame extraction using unsupervised clustering

Proceedings of the IEEE International Conference on Image Processing, **1**, 866–870

Key frame extraction has been recognized as an important research issue in video information retrieval. Although progress has been made in key frame extraction, the existing approaches are either computationally expensive or ineffective in capturing salient visual content. We first discuss the importance of key frame selection and then review and evaluate the existing approaches. To overcome the shortcomings of the existing approaches, we introduce a new algorithm for key frame extraction based on unsupervised clustering. The proposed algorithm is both computationally simple and able to adapt to the visual content. The efficiency and effectiveness are validated by a large number of real-world videos.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2000	Final,

4. TITLE AND SUBTITLE Advanced Displays and Interactive Displays Report Compendium II

5. FUNDING NUMBERS
DA PR: —
PE: —
worked performed under DAAL01-96-2-003

6. AUTHOR(S) Paul Rose, editor

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Laboratory
Attn: AMSRL-HR-I email: prose@arl.mil
Aberdeen Proving Ground, MD 21005-5425

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-SR-95

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Laboratory
Aberdeen Proving Ground, MD 21005-5425

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
ARL PR: —
AMS code: —

12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This report contains citations and abstracts of papers and presentations produced during the fourth year of the U.S. Army's Federated Laboratory (FedLab) program. The program was formed in 1996 to establish partnerships among the Army, industry, and academic research communities. The Advanced Displays and Interactive Displays consortium seeks to provide innovative, cost-effective solutions to information access, understanding, and management for the soldier of the future.

The research encompasses a range of topics. Some work deals with the representation of uncertainty and imprecision in databases, or with the representation of relationships in multimedia databases, in ways that are compatible with human cognitive-processing capabilities. Other work adopts the means of human communication (such as speech, gesture, eye gaze, and lip-reading) for human-computer interaction. Additional work explores methods for incorporating information in virtual-reality displays that support decision making without distracting or overwhelming the soldier. Although diverse, the research is linked by its overriding goal: the presentation of information in a form that allows effective human understanding and decision making in complex battlefield situations.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Federal Laboratory, FedLab, abstracts	53
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102