CONTRACTOR'S PROGRESS STATUS AND MANAGEMENT REPORT

Intelligent Collaboration and Visualization

for the period 1 January to 31 March 1999
Report #8
CDRL A001

Contract N66001-97-C-8517

20 April 1999

SUBMITTED TO

Receiving Officer
e-mail address: "icvnrad@nosc.mil"

Rich Laverty	Steve Murray
619-553-2918	619-553-6350
laverty@nosc.mil	murrays@spawar.navy.mil

SUBMITTED BY

Carnegie Mellon University
Pittsburgh, PA 15213

Principal Investigator	Administrative Contact	Contract/Financial Contact
Howard Wactlar	Dale James	Karen M. Faber
(412) 268-2571	(412) 268-7689	(412) 268-5838
fax: (412) 268-5576		(412) 268-5841
hdw@cmu.edu	dale+@cs.cmu.edu	faber+@andrew.cmu.edu

Quarterly Status Report

Intelligent Collaboration and Visualization
for the period 1 January to 31 March 1999
Cooperative Agreement N66001-97-C-8517
CDRL A001

1.0 Purpose of Report

This status report is the quarterly contract deliverable (CDRL A001) which summarizes the effort expended by the Carnegie Mellon University team in support of Intelligent Collaboration and Visualization (IC&V) on Contract N66001-97-C-8517.

2.0 Project Members

Wactlar
Christel
Faloutsos
Gong
Hauptmann
Reddy
Stevens
Other technical staff

3.0 Project Description (last modified 2/97)

This work will develop tools, techniques, and systems allowing people to capture a complete record of their personal experiences, and to share them in collaborative settings. Users may range from rescue workers carrying personalized information systems in operational situations to remote crisis managers in coordinating roles. Personal Informedia Experience-on-Demand (EoD) units record audio, video, GPS and other sensory data, which can be annotated by human participants. The EoD environment synthesizes data from many EoD units into a "collective experience" – a global perspective of ongoing and archived personal experiences. Distributed collaborators are thereby brought together over time and space to share meaning and perspectives.

Each constituent EoD unit captures and manages information from its unique point of view, making this information available to others in the EoD environment. Each operates as a portable, interoperable, Informedia system, allowing search and retrieval by both its human operator and remote systems. The EoD environment thus enables integration of the multiple points of view to provide more details for local decision-making and superior event coverage in support of real-time collaboration. The EoD capability will significantly improve situation awareness and analysis, both in real-time and retrospectively. The indexed and summarized information also enables "remembering" analogous past experiences and "projecting" into future simulated ones. Techniques will be developed to manage the vast quantities of information and to search, summarize, and visualize video, audio, and text content and annotations from multiple perspectives. The foundation for this work, the Informedia Digital Video Library (DVL) Project, has demonstrated the successful application of speech, image, and natural language processing in automatically creating a

rich, indexed, searchable multimedia information resource. We will build on these technologies, moving beyond a DVL into new information spaces by addressing collaboration and summarization of multiple simultaneous information generators integrated across people, time, and space.

4.0 Performance Against Plan

Spending proceeded as planned during this reporting period.

5.0 Major Accomplishments to Date

- Integrated VOCR (video optical character recognition) into EoD's underlying Informedia system.
- Improved our image-retrieval method by exploiting color-cluster and image-region characteristics.
- Tested image-retrieval performance using a larger image set (10,000 images) from the Informedia database.
- Modularized our image-retrieval method to facilitate integrating it into the EoD system.
- Integrated a preliminary data-visualization technique, VIBE, into the EoD system.
- Developed MindReader, a method that improves query by image content.
- Developed a method to synthesize a basic panoramic view from multiple perspectives that might be provided by concurrent EoD users.
- Developed a technique for motion detection in video, by comparing against reference frames.

6.0 Artifacts Developed During the Past Quarter

- We built a new, smaller and more reliable, EoD system. The new package is much cleaner with fewer failure points, and should allow better recording of field data.

7.0 Issues

7.1 Open issues with no plan, as yet, for resolution:

- None.

7.2 Open issues with plan for resolution:

- Client does not allow two concurrent users at the time. We must build capability of visualization of two or more users.
- Incorporate into the system the ability to acquire and process digital-compass data, in addition to GPS. This new data type will allow tracking view direction.
- Detection of incidental "scene text." Our VOCR can detect some scene text (signs on a truck, for instance) if it has certain properties (horizontal, stable, etc.).

7.3 Issues resolved:

- Initiated transition of underlying Informedia system to Oracle for greater speed and reliability.
- Rudimentary filtering of redundant video and composition of basic panoramic views are now possible.

8.0 Near-term Plan

- Complete move of IDVLS to an Oracle database.
- Conduct more experiments with two or more concurrent EoD users in more realistic situations.
- Experiment with distribution of the IDVLS across a number of machines to determine what configuration provides greater utility.
- Continue exploring query-space visualization for the user.

9.0 Completed Travel

None.

10.0 Equipment Purchases and Description

None.

11.0 Summary of Activity

We've begun investigating how the IDVLS architecture might be redesigned to make it fully distributed, available to several concurrent users, and persistent. Currently, it's a huge client running against a database. Our transition to an Oracle database allows us to consider a thin client using browser technology so that it becomes universally useable, without being constrained to having to download a monolithic, multi-megabyte client and install it on each user's machine. Informedia is not now scalable to multiple, concurrent EoD users because if the client crashes in the middle of a transaction, data is lost. By distributing access to the system we hope to provide some persistence in the server, and if the client (browser) dies, rerunning the client will allow the user to reconnect to his old instance of service in the server domain.

We've also continued our work on 3D reconstruction of multiple views from multiple EoD users. Data must be gathered manually in a very painstaking way, frame-by-frame, with GPS readings for each frame. We currently cannot do it real time. Our next step is to explore exploiting data from, for instance, ten frames with one single GPS reading. Given that incomplete data, how well can we generate a composite view?

Research continues on detecting "scene text". Our current algorithms can detect some scene text that has the proper properties (horizontal, stable, etc.), but do not as a rule deal well with signs, writing on moving vehicles, stylized text, etc. We're starting to experiment with a neural network approach, similar to that used for faces, to detect such text.

We've done more work with mobile, audio-only experience recording, and begun filtering long silences automatically.

Presentation of more, and more relevant, information to the user becomes a substantial challenge as the dataset available to the user grows. Our interface now incorporates a tabbed Visualizer form to house the timeline and VIBE interfaces, and the timeline has been updated to support existing VIBE functionality.

11.1 Significant Events:

- None this period.