

CONTRACTOR'S PROGRESS STATUS AND MANAGEMENT REPORT

Intelligent Collaboration and Visualization

for the period 1 January to 31 March 2000

Report #12
CDRL A001

Contract N66001-97-C-8517
27 July 2000

SUBMITTED TO

Receiving Officer
Email address: icvnrad@nosc.mil

Rich Laverty
(619) 553-2918
laverty@nosc.mil

Steve Murray
(619) 553-6350
murrys@spawar.navy.mil

SUBMITTED BY

Carnegie Mellon University
Pittsburgh, PA 15213

Principal Investigator
Howard Wactlar
(412) 268-2571
fax: (412) 268-7458
wactlar@cmu.edu

Administrative Contact
Colleen Everett
(412) 268-7674
(412) 268-7458
everett@cs.cmu.edu

Contract/Financial Contact
Karen M. Faber
(412) 268-5838
(412) 268-5841
faber+@andrew.cmu.edu

Do not distribute to DTIC or other data depositories.

Distribution authorized to DOD components only; premature dissemination (date). Other requests shall be referred to Naval Command, Control and Ocean Surveillance Center (NCCOSC), RDT&E Division, San Diego, California 92152-5000.

Quarterly Status Report

Intelligent Collaboration and Visualization
for the period 1 January to 31 March 2000
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
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 VOOCR (video optical character recognition) into EoD's underlying Informedia system, allowing EoD to "read" some naturally occurring scene text.
- Improved our image-retrieval method by exploiting color-cluster and image-region characteristics.
- Tested image-retrieval performance using a larger image set (20,000 images) from the Informedia database.
- Modularized our image-retrieval method to facilitate integrating it into the EoD system.
- Developed a map-based display and query mechanism enabling better access to and visualization of library data.
- 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.
- Automatic silence filtering integrated into video processing.
- Location names can now be disambiguated automatically in geocoding.

6.0 Artifacts Developed During the Past Quarter

A library of realtime recorded EoD experience data integrated with historical and archived footage.

7.0 Issues

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

None.

7.2 Open issues with plan for resolution:

- More useable, real-world detection and reading of incidental "scene text." Our VOOCR can detect some scene text (signs on a truck, for instance) if it has certain properties (horizontal, stable, etc.)
- Collect real-time multiple EoD views from several units.

7.3 Issues resolved:

- Improved scene text detection with backward propagating neural nets.
- Designed a more fault-tolerant EoD architecture.
- Reduced bandwidth requirements.

8.0 Near-term Plan

- Refine and optimize distribution of the IDVLS across a number of machines to determine what configuration provides greater utility.
- Develop new strategies for deployment and coordination of multiple, concurrent EoD units.
- Develop new visualization techniques

9.0 Completed Travel

Wactlar - January 25-26, 2000 - NIST Pervasive Computing Conference, Gaithersburg, MD

10.0 Equipment Purchases and Description

<omitted>

11.0 Summary of Activity

Developed an algorithm for on-demand panoramic view reconstruction that reduces dependency on potentially inaccurate and occasionally unavailable GPS information. This new approach uses matched images on consecutive “frames” as markers, computes the transformation between the frames, and then joins them together into a panoramic representation.

Conducted a user study to measure the effectiveness of Multimodal map and text queries and relevance feedback strategies, included in a prototype EoD system. Our various EoD implementation to date have demonstrated that system effectiveness is heavily dependent upon the robustness of the underlying search engines and the feature extraction process. While improving the accuracy of named entity extraction and other representative relevance criteria can greatly improve the precision of relevance feedback, user feedback shows that the new system already facilitates user knowledge discovery and contributes to higher system usability and interactivity.

Began recording GPS data onto one channel of the audio track of the digitally recorded video stream. By sacrificing stereo sound, we are able to eliminate the requirement for a separate GPS data stream that must be transmitted back to a central base, synchronized with the video, indexed, and archived.

Changed video segment titles to be more practically useful. The titles now reflect recording time and location, in addition to the name of the EoD field user. GPS coordinates are looked up in the gazetteer, and the appropriate location name is automatically added to the title.

Collaborated with the local public television station, WQED, to develop a library of realtime recorded EoD experience data integrated with historical and archived footage from their in-house stock.

Began investigating new ways of training neural nets for VOICR. Previously, the neural nets were trained on *tiled* images, creating greater possibility of spurious partial-character recognition. The new approach is to train the neural nets on a character-by-character basis, much as one would teach a human to recognize and read characters.

11.1 Significant Events:

- None.