CareMedia:
Automated Video and Sensor Analysis for Geriatric Care

NSF Cooperative Agreement No. IIS-0205219
Annual Progress Report
March 2003

Carnegie Mellon University
School of Computer Science
Pittsburgh, PA 15213-3890

Principal Investigator: **Howard Wactlar**
School of Computer Science

Co-PI's:

Alexander Hauptmann
Computer Science Dept.

Takeo Kanade
Robotics Institute

Scott Stevens
Computer Science Dept.

Ashok Bharucha
Western Psychiatric Institute

1 Project Participants

What people have worked on the project?

Participant's Name	Project Role	>160 Hours?
Howard D. Wactlar	Principal Investigator	Yes
Takeo Kanade	Co-Principal Investigator	Yes
Ashok Bharucha	Co-Principal Investigator	Yes
Scott M. Stevens	Co-Principal Investigator	Yes
Alexander G. Hauptmann	Co-Principal Investigator	Yes
Sonya Allin	Graduate Student	Yes
Christopher Atkeson	Project Scientist	Yes

Are any other persons involved with the project?

None.

What other organizations have been involved as partners?

University of Pittsburgh Medical Center (UPMC).

Western Psychiatric Institute and Clinic (WPIC).

Have you had other collaborators or contacts?

Longwood at Oakmont Retirement Resort, Dementia Care Facility

2 Project Activities and Findings

Overview

This research will create a continuously recorded, digital history of patients' activities in a nursing home by capturing all that is heard, seen and experienced. Our research challenge is to transform a voluminous amount of captured video, audio and sensor data into a meaningful information resource that enables more complete and accurate assessment, diagnosis, treatment, and evaluation of behavioral problems for the elderly. The proposed work in (i) information extraction and retrieval, (ii) behavior recognition, analysis and summarization, and (iii) secure, efficient visual information access will let geriatric care specialists obtain greater insights into problems, effectiveness of treatments, and environmental and social influences on patient behavior. Prototype systems for activity monitoring and behavior analysis will be deployed at local area nursing homes and dementia wards to be utilized by medical professionals in trials conducted by our project partners from the University of Pittsburgh Medical Center (UPMC), Western Psychiatric Institute and Clinic (WPIC).

Activities

Low funding at the outset of the CareMedia project greatly affected our approach to the research problems posed. We have been able to gather some preliminary "surveillance-type" video at a nursing care facility, Longwood at Oakmont's Dementia Care Facility, under the supervision of our psychiatric colleague, Dr. Ashok Bharucha. We have exploited this set of data extensively, developing techniques to detect and classify behavior in video, and to digitally "remove" persons from video who do not want to participate in this research. These techniques, described in further detail below, will be submitted for publication.

2.1 Pilot Study

We conducted a pilot study to: (a) explore the feasibility of implementing CareMedia technologies in a nursing home setting, and (b) provide the medeical investigators with exposure to the potential challenges and advantages of using this approach for behavioral assessment.

A one-week naturalistic observational study was designed in conjunction with the nursing home administration and staff. The rationale for the study approach evolved from Van Haitsma et al's (1997) strategy of collecting data on "streams of behavior" using hand-held computers. The authors described the theoretical underpinnings of their study thus: "…the functioning of a holistic unit (in this case, a treatment area) is studied in terms of the behaviors of individuals in transaction with their social and physical context. Persons and environment form "behavior settings," the dynamics of which are studied with the use of "behavior maps," where behaviors are cataloged in terms of their locations, and the recording of a single individual's behavior over time and space ("behavior stream")." This study was utilized to address the following specific aims.

Methods to achieve digital data privacy and confidentiality:

Ethical concerns loomed large in the initial administrative meeting where the objectives and methods of the study were reviewed. The investigators decided to address the data privacy and confidentiality issues by offering a multi-tiered level of study participation. The options, designed to minimize infringement on resident privacy and autonomy, included the following categories: (a) permission to record video and audio, and include resident as a study participant, (b) permission to record, but resident not a study participant, (c) permission to record on the unit but digital deletion of the particular resident requested. A sign was placed on the entrance door to the dementia unit alerting staff and visitors that the study was in progress. Private spaces such as bedrooms and bathrooms were excluded from the study. Digital deletion was not requested by any consenting legal proxy or staff member; the only person to decline study participation had no objection to the actual recording on the unit. As a long-term goal, technologies will be developed as part of the project to prevent non-consenting persons from being recorded.

Instrumentation:

The single main hallway, the dining and living rooms of the unit were instrumented with four clearly visible, but unobtrusive, pole-mounted surveillance caliber cameras and eight

audio-phones that stored data directly onto computer hard drives. The residents' activities were recorded during four two-hour blocks of time from 8am-10am, 11am-1pm, 2pm-4pm and 5pm-7pm. The periods of assessment were chosen based on staff input regarding the times the residents are most active on the unit. Also, per staff input, all three mealtimes were recorded each day.

Data collection:

In addition to the video/audio capture, the primary nurse for each resident completed the Pittsburgh Agitation Scale (PAS) twice daily. The PAS is a brief, user-friendly instrument that rates the severity of disruptive behaviors from 0 (none) to 4 (severe) based on direct observations of the patient (Rosen et al, 1994). The PAS covers four broad domains: Aberrant Vocalization, Motor Agitation, Aggressiveness, and Resistance to Care.

2.2 Findings

2.2.1 Summarizing Human Activity in Video

We have developed a method for detecting and summarizing human activities from a large amount of video. A video is divided into short segments (called stories), and a set of simple motion/image features are computed for each image frame. Prototypical elementary action patterns (called action events) are constructed by vector quantization on these features. A bipartite graph is constructed by taking stories and action events as nodes, and the co-occurrence relationship between them as the graph edges. We compute an optimal reordering of the video story and the action event nodes so that they are in maximum correlation with each other. The optimal mapping between the stories and action events leads to a grouping and ordering of the video stories into activity patterns. The corresponding ordering on the action events reveals their characteristic features.

2.2.2 Protecting Privacy in Recorded Video

In many surveillance situations, a large amount of video gets recorded, but depending on the local laws and the perspective of the camera view, as well as the nature of the setting that is being recorded, researchers and others interested in utilizing or distributing the video must secure the permissions of the recorded persons to use that video data. These permissions can sometimes be secured, but it has been our experience that frequently only a subset of the recorded individuals can be tracked down to for us to obtain the required permissions. Furthermore, some people choose not to have their recordings become part of the archive collection.

We have developed algorithms to automatically extract and label individuals in video, using integrated tracking and event detection. We used a unifying multi-class active learning framework in order to reduce the human labeling effort. Our experiments demonstrate that an active learner with careful sample selection can achieve remarkably good performance (5.7% labeling error) with much less human labeling effort (50

examples, which translates into only 5% of the labeling effort) compared to supervised learning. Also, an active learner with the proposed sample selection strategies can do much better than one with only a random sampling strategy, which yields an over two-fold error reduction. Using these techniques, we have successfully demonstrated "removal" of specific individuals from a video stream.

2.2.3 Hygiene Activity Analysis

We have prototyped a system that can observe individual personal activities conducted in front of the bathroom mirror, and provide summary statistics over time of personal hygiene activities. It aims to be useful for older people, nursing home patients, children, etc. We used a miniature camera and monophonic microphone to record video clips. After extracting features including face position, arm position, audio pitch, etc., we perform a supervised classification of different hygiene activities including hair combing, face washing, etc., based on multiple cues. We then create a summary for these activities, providing a statistical record across activity type and over time.

References:

[Rosen94] Rosen, J., Burgio, L.O., Kollar, M., Cain, M., Allison, M., Fogleman, M., Michael, M., and Zubenko, G. "The Pittsburgh Agitation Scale (PAS): A User-Friendly Instrument for Rating Agitation in Dementia Patients," *The American Journal of Geriatric Psychiatry*, Winter 1994, **2**(1), 52-59.

3 Training and Development

Graduate student support has enabled research into object tracking, and activity and behavior classification in video.

Clinical psychiatrist Dr. Ashok Bharucha has been introduced to computation techniques for capturing behavior in a clinical setting.

Undergraduate students have been exposed to techniques for video analysis during the analysis phase of our pilot study.

4 Outreach Activities

Our work at a nursing home during the pilot study has exposed staff and patient families to the potential clinical value of computational technology.

5 Publications and Products

"Summarizing Human Activity in Video", H. Zhong, J.Shi, submitted to IEEE Computer Vision and Pattern Recognition Conference, June 2003

"Finding (Un)Usual Events in Video", J. Shi, H. Zhong,
The Learning Workshop, Snowbird, Utah, April 1-5, 2003

6 Contributions

6.1 Contributions within discipline

Techniques for auto-elimination of specified human subjects from video content.

Semi-automated system for rapid (10% of real-time) viewing, annotating and truthing of continuously captured video.

6.2 Contributions to other disciplines

Initial field studies were exploratory, but the subsequent planned CareMedia/nursing home collaboration will result in the world's most comprehensive study of the ecology of behaviors in a dementia ward to date.

6.3 Contributions to human resource development

Nursing home field studies have educated computer scientists with respect to issues of privacy and sensitivity in automated analysis of observational data.

6.4 Contributions to resources for research and education

Nothing significant yet.

6.5 Contributions beyond science and engineering

Nothing significant yet.

7 Special Requirements

7.1 Objectives and Scope

A brief summary of the work to be performed during the next year of support if changed from the original proposal.

7.2 Special Reporting Requirements

Do special terms and conditions of your award require you to report any specific information that you have not yet reported?

No.

7.3 Unobligated Funds

Do you anticipate that more than twenty percent of the funds under your NSF award will remain unobligated at the end of the period for which NSF currently is providing support?

No.

7.4 Animals, Human Subjects, Biohazards

Has there been any significant change in animal care and use, use of human subjects, or biohazards, from what has previously been approved?

No.