Groupscope: A Microscope for Large Dynamic Groups Research

Sandeep Puthanveetil Satheesan, Marshall Scott Poole, Rob Kooper, Kenton McHenry
National Center for Supercomputing Applications
University of Illinois at Urbana-Champaign
Urbana, IL, USA
{sandeeps, mchenry}@illinois.edu

***Abstract*—Social science research on large dynamic groups of people involves many challenges of which a prominent one is working with large amounts of video and audio data. Lack of automation has resulted in a methodological gap on how research is conducted in this area. For example direct observation and annotation becomes a tedious and manual task as the size of the group and the duration of interaction increases. We propose to leverage an extensible and scalable content aware data repository called Medici to incorporate content based analysis tools from computer vision as well as manual annotations tools to provide a more practical solution for storing, managing, processing, and analyzing data for researchers enabling them to make firsthand observations and verification of theories.**

Keywords—cyberinfrastructure; computer vision; large dynamic groups; large-scale data analysis;

I. INTRODUCTION

We define a large dynamic group (LDG) as an assembly of people of size ranging from 8 to 200 usually interacting in dynamic subgroups in indoor or outdoor spaces. Understanding the behavior of LDGs is important because they perform many critical functions in the modern-day society like disaster and emergency reponse, military operations, and medical care. A lot of research has been conducted on groups smaller than these due to managable data and hence ease of making direct observations from interactions. In case of LDGs there are gaps between theory, research and methodologies due to unavailability of analysis tools for such large scales. For e.g. a hypothetical study on a group of 40 emergency responders moving through a disaster simiulation scene for 6 hours recorded using 20 cameras and 40 microphones generates a total of 120 hours of video and 240 hours of audio data. Total size of all this data could be anywhere in the range 600 – 800 GB. To find substantial evidence for theories many number of such studies will need to be conducted and the resulting data will need to be analyzed as a whole. This increases the scale of the problem multiple times. At these scales it becomes extremely difficult if not impossibile to manage data, do manual annotations, identify dynamic networks of people, obtain different views of data which will inturn give hints on the behavior of the group. Currently most of the research on LDGs are done using ethnography and participant observation, analysis of transactional records, and network analysis based on survey data, but what is missing is direct observation and analysis of multilayered data captured over time. The proposed system aims to solve this problem and fill the above mentioned gaps using a content aware data repository.

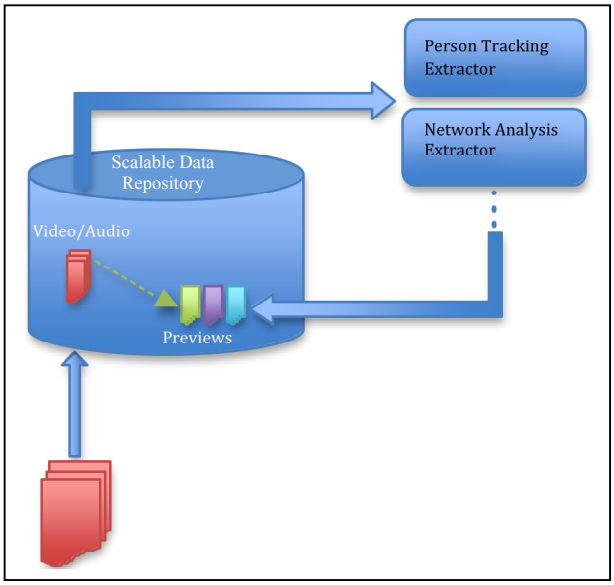

Fig. 1: Dataflow diagram depicting the process of metadata extraction and preview generation.

II. PROPOSED SYSTEM

We propose a scalable content aware data repository [1] based solution for conducting studies on LDGs and analyzing the data. The repository leveraged, NCSA Medici, supports hetergenous data, RESTful services, and abstraction of the underlying storage, which helps in scaling to larger storage repositories as per the studies' requirement. Medici has two major components, namely the extractors and previewers. Extractors are software components which generate new data from existing data stored in the repository. This can be metadata derived from a files' content or a new data file itself. For example, an extractor may identify and track people in a video and generate a new video file with each person labelled in all the frames. Previewers fetch previews associated with a data item and renders it to the user. For example, a previewer would obtain an image preview which is zoomable and displays it for user interaction within the browser.

Through extractors we intend to generate metadata about video and audio files which will be used either directly or indirectly in answering qustions like who is present in the current scene, who is interacting with whom etc.

ELAN [2] is a tool used by researchers to annotate and record dialogues and actions in video and audio files. Integration of ELAN with the data repository helps in partially automating the process of annotation. We use a WebDAV [3] interface to Medici's contents in order to achieve this as it provides a more intuitive desktop user interface for the files uploaded in the repository. When annotation of a video or audio is completed it is uploaded to the repository via a WebDAV drive. Extractors listening for annotation files can now extract information from it.

Person identification and tracking is an important component of Groupscope. Localizing persons in space and time across video files captured from different cameras using computer vision algorithms reduces the effort needed in manual annotation to a great extent. Since person tracking is never 100% perfect due to factors like occlusion and varying lighting conditions, manual intervention may be needed at times but nevertheless it reduces the annotation effort by a big margin. Here we consider tracking solutions discussed in [4] and [5].

Another important step is the identification of communication links between people and representing this information in a form which enables the researchers to "zoom in" to subgroups and "zoom out" to get the big picture. We explore the methods suggested in [6] to create extractors which identifies communication links. A map based preview is used to represent the data generated by the extractor.

III. Applications

There are a number of imporant applications identified for the proposed system where multi-team interactions are involved. Emergency response is a critical task where efficiency in planning and execution could be the difference between life and death for many. Groupscope will enable the analysis of many audio and video files where by researchers can propose and verify better theories for improving the efficiency of emergency responders. Another application area invloves studies of aggression and bullying in classroom and playground. Here in addition to data quantity, there are additional challenges related to person tracking from long distances and noise in the playground due to activity. Data is being collected for the aforementioned applications. Another application could be for conducting studies to improve the efficiency of medical care professionals interacting within a hospital. Other application areas might be military operations and product design.

IV. Conclusion

We presented our efforts towards a scalable content aware data repository based solution for conducting research and anlayzing data on Large Dynamic Groups. The proposed approach aims at alleviating much of the manual and subjective

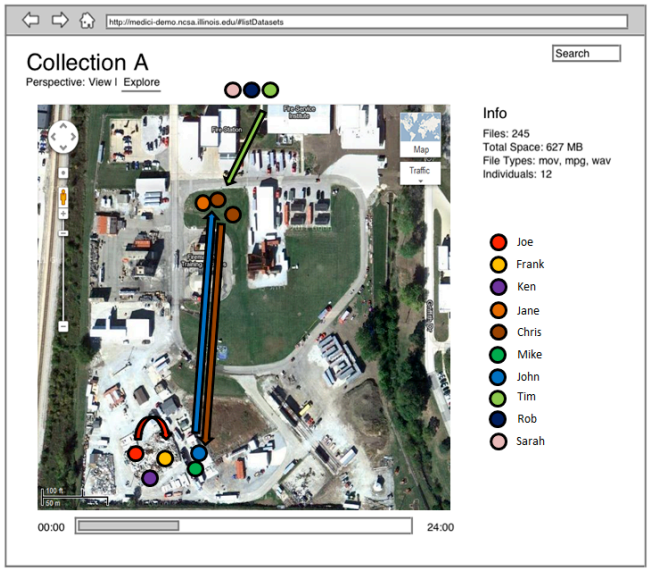

Fig. 2.: A sample preview of communication links laid over a map. Persons are represented by dots and arrows represent their communications directions.

elements of such research while also enabling new discoveries in Social science research regarding LDGs.

Acknowledgements

This work at NCSA/UIUC has been supported by the National Science Foundation under grant BCS-0941268.

References

[1] L. Marini, R. Kooper, J. Futrelle, J. Plutchak, A. Craig, T. McLaren, and J. D. Myers, "Medici: A Scalable Multimedia Environment for Research", Microsoft Research eScience Workshop, Berkeley, CA, October 2010.

[2] P. Wittenburg, H. Brugman, A. Russel, A. Klassmann, and H. Sloetjes, (2006). "ELAN: a Professional Framework for Multimodality Research", In Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation. url: http://tla.mpi.nl/tools/tla-tools/elan

[3] B. McEvoy, Milton Webdav, [online] 2013. http://milton.io/downloads/index.html (Accessed: 3 September, 2013)

[4] D. Ramanan, D. Forsyth, and A. Zisserman, "Tracking People by Learning their Appearance", Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 29(1), pp. 65-81, 2007.

[5] Z. Kalal, J. Matas, and K. Mikolajczyk, "P-N Learning: Bootstrapping Binary Classifiers by Structural Constraints", In Computer Vision and Pattern Recognition, 2010 IEEE Conference on, pp. 49-56, 2010.

[6] S. Mathur, M. S. Poole, F. Pena-Mora, M. Hasegawa-Johnson, and N. Contractor, (2012). "Detecting interaction links in a collaborating group using manually annotated data". Social Networks, vol. 34(4), pp. 515-526, October 2012.