**REPORT DOCUMENTATION PAGE**

1. **AGENCY USE ONLY (Leave blank)**
2. **REPORT DATE**
3. **REPORT TYPE AND DATES COVERED**
   - Final
4. **TITLE AND SUBTITLE**
   - Distributed Multimedia Reasoning
5. **FUNDING NUMBERS**
   - DAAH04-95-1-0174
6. **AUTHOR(S)**
   - V.S. Subrahmanian
7. **PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)**
   - University of Maryland
     - College Park, MD 20742
8. **PERFORMING ORGANIZATION REPORT NUMBER**
9. **SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
   - U.S. Army Research Office
     - P.O. Box 12211
     - Research Triangle Park, NC 27709-2211
10. **SPONSORING / MONITORING AGENCY REPORT NUMBER**
    - ARO 34061.2-MA
11. **SUPPLEMENTARY NOTES**
    - The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.
12a. **DISTRIBUTION / AVAILABILITY STATEMENT**
    - Approved for public release; distribution unlimited.
12b. **DISTRIBUTION CODE**
13. **ABSTRACT (Maximum 200 words)**
    - The research performed under this grant led to the development of abstract models of multimedia data, concrete ways of representing and indexing such multimedia data on a computer, querying such multimedia data using a single unified query language, assembling presentations comprised of a variety of multimedia objects, and delivering such presentations across a distributed network. In effect, the work developed a firm theoretical foundation for reasoning about multimedia databases.

**19990616 129**

14. **SUBJECT TERMS**
15. **NUMBER OF PAGES**
16. **PRICE CODE**
17. **SECURITY CLASSIFICATION OR REPORT**
18. **SECURITY CLASSIFICATION OF THIS PAGE**
19. **SECURITY CLASSIFICATION OF ABSTRACT**
20. **LIMITATION OF ABSTRACT**
Distributed Multimedia Reasoning

ARO Grant No. DAAH-04-95-1-10174

Principal Investigator: V.S. Subrahmanian
Department of Computer Science
University of Maryland
College Park, MD 20742

This proposal started in June 1995 and ended in Nov. 1998.

1 Accomplishments

The research performed under this grant has been on a variety of topics related to the efficient storage and retrieval of massive amounts of multimedia data located at one or more sites across the Internet. During this time period, we obtained the following important results.

Conversion and Delivery of Multimedia Objects: When a group \( \{I_1, \ldots, I_n\} \) of individuals wishes to collaboratively construct a complex multimedia document, the first requirement is that they be able to manipulate media-objects created by one another. For instance, if individual \( I_j \) wishes to access some media objects present at participant \( I_k \)'s site, he must be able to: (1) retrieve this object from across the network, (2) ensure that the object is in a form that is compatible with the viewing/editing resources he has available at his node, and (3) ensure that the object has the desired quality (such as image size and resolution). Furthermore, he must be able to achieve these goals at the lowest possible cost. In this research, we develop a theory of media objects, and present optimal algorithms for collaborative object sharing/synthesis of the sort envisaged above. We then extend the algorithms to incorporate quality constraints (such as image size) as well as distribution across multiple nodes. The theoretical model is validated by an experimental implementation that supports the theoretical results.

Collaborative Multimedia Documents: Authoring and Presentation Multimedia documents are composed of different data types such as video, audio, text and images. Authoring a multimedia document is a creative exercise. Unlike traditional computer supported collaborative work where documents are composed of static objects, multimedia documents have temporal, spatial and quality of service (QoS) requirements that must be supported by any collaborative multimedia platform. In this work, we show that most requirements (including temporal, spatial, and QoS requirements) for collaborative multimedia systems can be expressed in terms of a highly-structured class of linear constraints called difference constraints that have
been well-studied in the operations research literature. As a consequence, well known algorithms for solving difference constraints may be used as a starting point for creating multimedia documents. Based on our difference-constraint based characterization, we develop efficient, incremental algorithms for creating and modifying multimedia documents so as to satisfy the required temporal, spatial and QoS constraints. We further develop methods to identify inconsistent requirements, and show how such inconsistencies may be removed through constraint relaxation techniques.

**CHIMP System:** We developed part but not all of the Collaborative Heterogeneous Interactive Multimedia Platform (CHIMP) system under this grant. CHIMP has the goal of studying the technical aspects of collaborative multimedia document authoring and presentation, as well as building a system based upon these results.

Suppose we consider a team of individuals jointly authoring a multimedia document. In order to successfully author such a document, the authors must:

- Identify the objects (e.g. audio objects, video objects, text objects, etc.) that will be part of the authored multimedia document, and
- Specify how these objects should be presented to an end-user wishing to view the final multimedia document. This specification may involve:
  1. Temporal constraints, e.g. “Start showing object a.mpg and b.avi at the same time, and ensure that a.mpg finishes at least 15 seconds before b.avi. Start showing the text document c.txt as soon as a.mpg finishes and stop showing c.txt when b.avi finishes.”
  2. Spatial constraints, e.g. “Ensure that the lower left corner of the window in which object a.mpg appears coincides with the upper right corner of the window in which object b.avi appears.”
  3. Other constraints, e.g. quality of service constraints, delay constraints, delay jitter constraints, and cell loss probability constraints.

The above constraints, collectively termed *Presentation Constraints*, specify the spatial, temporal and QoS constraints associated with a specification, and they have been studied extensively by many authors [?, ?]. We have started on implementing efficient algorithms for incrementally solving such *presentation constraints*.

**Multimedia Databases:** Though numerous multimedia systems exist in the commercial market today, relatively little work has been done on developing the mathematical foundations of multimedia technology. We attempt to take some initial steps towards the development of a theoretical basis for multimedia information system. To do so, we develop the notion of a structured multimedia database system. We begin by defining a mathematical model of a media-instance. A media-instance may be thought of as “glue” residing on top of a specific physical media-representation (such as video, audio, documents, etc.) Using this “glue”, it is possible to define a general purpose logical query language to query multimedia data. This glue consists of a set
of "states" (e.g. video frames, audio tracks, etc.) and "features", together with relationships between states and/or features. A structured multimedia database system imposes a certain mathematical structure on the set of features/states. Using this notion of a structure, we are able to define indexing structures for processing queries, methods to relax queries when answers do not exist to those queries, as well as sound, complete and terminating procedures to answer such queries (and their relaxations, when appropriate). We show how a media-presentation can be generated by processing a sequence of queries, and furthermore we show when these queries are extended to include constraints, then these queries can not only generate presentations, but also generate temporal synchronization properties and spatial layout properties for such presentations. We developed an architecture for prototype multimedia database system based on these principles.

**Similarity Algebras:** The need to automatically extract and classify the contents of multimedia data archives such as images, video, and text documents has led to significant work on similarity based retrieval of data. To date, most work in this area has focused on the creation of index structures for similarity based retrieval. There is very little work on developing formalisms for querying multimedia databases that support similarity based computations. Given the fact that most feature extraction and identification algorithms in media data are very expensive, the need for query optimization to such databases is critical — yet, it is impossible to create principled query optimization algorithms without a declarative data definition and query language.

In this work, we introduce a similarity algebra that brings together relational operators and results of multiple similarity implementations in a uniform language. The algebra can be used to specify complex queries that combine different interpretations of similarity values and multiple algorithms for computing these values. We then state the conditions on equivalence and containment relationships between similarity algebraic expressions and develop query rewriting methods based on these results. We then provide a generic cost model for evaluating cost of query plans in the similarity algebra and query optimization methods based on this model. We supplement the work with experimental results that illustrate the use of the algebra and the effectiveness of query optimization methods using the Integrated Search Engine (I.SEE) as the testbed. (I.SEE was developed at RPI not at Maryland).

**Video Queries:** There is now growing interest in organizing and querying large bodies of video data. In this work, we develop a simple SQL-like video query language which can be used not only to identify videos in the library that are of interest to the user, but which can also be used to extract, from such a video in a video library, the relevant segments of the video that satisfy the specified query condition. We investigate various types of user requests and show how they are expressed using our query language. We also develop polynomial-time algorithms to process such queries. Furthermore, we show how video-presentations may be synthesized in response to a user query. We show how a standard relational database system can be extended in order to handle queries such as those expressed in our language. Based on these principles, we have built a prototype video retrieval system called VIQS. We describe
the design and implementation of VIQS and show some sample interactions with VIQS.

2 Contact with Army Personnel

During the 1995—1996 period, we interacted extensively with John Benton of US Army Topographic and Engineering Center (TEC) until June 1997. In June 1997, Mr. Benton moved from TEC and joined the University of Maryland. In addition, have worked extensively with LTC Jack Marin and MAJ Joseph Schafer of the US Military Academy in West Point, NY on applying many of our theoretical results to practical needs of the US Army Logistics Integration Agency. More recently, we have been involved in working with LTC. George Stone of STRICOM.

3 Educational Impact

The following students received PhDs that were funded (in part) by this grant.

1. Sibel Adali (now at Rensselaer Polytechnic Institute)

2. K.S. Candan (now at Arizona State University)

3. Eenjun Hwang (now at Bowie State University)

In addition, the grant supported (in part) the research conducted by several other PhD and undergraduate students.

4 Relevant Publication:


15. K.S. Candan, E. Lemar and V.S. Subrahmanian. Multimedia View Management, Currently being revised according to referee comments, VLDB Journal.


