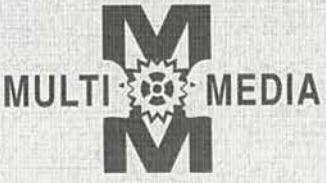
Centrum voor Wiskunde en Informatica

Centre for Mathematics and Computer Science

Multi-Media Research at CWI: Goals and Objectives



Multi-Media Research at CWI: Goals and Objectives

D. C. A. Bulterman
Department Head, Computing Systems and Telematics

CWI: The Center for Mathematics and Computer Science
PO Box 4079, 1009 AB Amsterdam, The Netherlands

1. Introduction

This document will serve as a brief introduction to the motivation and goals of the multi-media research at CWI. We provide this introduction outside of the context of any particular project description; we do this to highlight our general intentions without needing to motivate individual applications. Companion documents are available that detail individual research proposals for several multi-media projects under development at CWI. Additional information is available from Mr. H. P. Dijkhuis, Multi-Media Project Coordinator, at the address above.

This report is divided into two sections: the first provides an overview of the motivations for multi-media research, and the second provides a focus for the types of multi-media systems we plan to investigate.

2. Motivation for CWI's Multi-Media Research

One of the particularly appealing aspects of multi-media research to CWI is its inherent interdisciplinary nature. At a time when collaborative research is being stimulated, we feel that we are in a unique position to bring together researchers with experience in many of the key areas of multi-media systems design. Of the areas that we have current expertise in, we feel the following are especially useful to multi-media research in general:

(1) **User interfaces:** one of the single most important aspects of a system that provides extensive user facilities for collecting and manipulating various types of data is clearly the interface presented to the user. The user interface group at CWI has experience with two types of systems that provide the foundations for multi-media work: a programming language (*ABC*) that allows novice users to quickly manipulate various data types in a user-friendly manner, and a general user interface framework (*VIEWS*) that allows for the definition of an ''interface-independent'' application environment that shields individual applications from detailed knowledge of the user/system interface.

(2) **Interactive data manipulation:** given sets (or *streams*) of multi-media data, it is important to manipulate and transform these streams in a variety of ways to provide different mappings of the data to a particular user. This problem is similar to the transformation of graphical and image data that our Interactive Systems graphics group has been considering during the last decade. This group has gained broad experience in modelling and manipulating various streams of data, from the development of high-level graphics languages such as GKS, down to the development of special-purpose VLSI devices for the high-speed manipulation of independent graphical objects.

(3) **Database design:** a fundamental element of multi-media systems is the ability to manipulate different types of data. This implies a need for a data storage and retrieval system that provides storage- and time-efficient methods for saving and accessing components of multi-media streams. Our database group, through its participation in the *TROPICS* Esprit-II project, has developed data models that provide for the encapsulation of flexibly-defined data attributes within a general structure that can be adapted for the high-speed access required by multi-media applications.

(4) **Distributed systems:** the computational needs of a multi-media system require the coordinated use of many nodes in a computer network. To support the access of multi-media objects, we see the need for an underlying operating system structure that provides communication, synchronization, and security services among the network nodes. We have gained extensive experience in this area through the development of the *Amoeba* distributed operating system, which was jointly developed by CWI and the Free University in Amsterdam. The Amoeba system provides a fully distributed foundation that already functions in a multi-vendor environment, and support high-speed communication and storage facilities, along with object protection, naming, and synchronization primitives.

(5) **Wide-area networking:** one of the important aspects of multi-media systems is the desire for broadly-based systems, both in terms of the functionality of a single node and in terms of the ability to communicate across a wide range of nodes over various layers of networks. CWI has gained extensive experience in the latter area by our leading role within the European networking community. This experience has included the development of networking protocols for efficient wide-area information exchange, and the development of network management expertise that can be use to coordinate the development of multi-media systems.

(6) **Performance analysis:** given the projected intensive use of a number of different logical and physical subsystems to support multi-media application, we feel that performance analysis can play a major role in evaluating different systems options at the design state. Performance analysis, as an area of research, is concerned with the description and prediction of the behavior of service systems under varying load conditions. This discipline entails modelling of the relevant systems, model analysis by mathematical and experimental techniques, evaluation of the results so obtained and finally their implementation in systems design, routing and flow control. One area of interest in multi-media applications is the use of cyclic scheduling of various traffic streams (as would be found in ISDN-style networks); CWI, in partnership with IBM, has been involved in the intensive study of these algorithms during the last five years, with a focus on the analysis of IBM's token ring network.

(7) **Image processing:** Image analysis requires the close integration of computer graphics functionality (polygon draw /fill, template deformation, stencilling) with image processing (linear filtering, masking, distance transform, global average). For applications such as feature recognition, document scanning and video compression, the system needs fast computation and data transfer rates for both image processing and draw/fill/stencil tasks. Fundamental research in this area makes an additional demand for 'debugging' features, so that all of the computations described above can be displayed on a workstation during execution. Research on image analysis at CWI is under way in BS 6 (Statistical Methods in Image Analysis), AM 5 (Image Reconstruction) and in the Departments AA and IS. The CWI also has close links with the image processing group of the University of Amsterdam Faculty of Mathematics and Computer Science, sited at WCW, which is producing commercial image processing software; and with the CSIRO Division of Mathematics and Statistics, Australia, who are producing an interactive statistical package with combined image/graphic displays.

We view our leadership role in each of the areas above as providing a foundation upon which coordinated multi-media research can be built. By integrating the expertise of existing groups, we feel that we can bring an important mix of experience to the development of new concepts and models that will support a significant effort in defining and prototyping multi-media systems. We also feel that this combined effort will also provide a useful base for expansion into related areas, such as data compression, but providing a directed focus upon which such expansion can be built.

3. CWI's View of Multi-Media Research

At CWI, as at many research and development organizations, the area of multi-media research is becoming increasingly prominent. The area is seen as an important next-step in the development of computing systems, offering the potential for vastly enhanced computer access and human-to-human communication via machine. Unfortunately, as in many rapidly growing fields, there is not yet a general agreement on the nature of multi-media systems. While this is not rare—consider, for example, that after over twenty years of research, there still is considerable disagreement over the nature of distributed computing systems—it does mean that any discussion of multi-media research needs to start with a description of the scope of the work to be undertaken. This section provides such a description by considering two aspects of a multi-media environment: its external capabilities and its internal structuring.

3.1. External characteristics of multi-media systems.

If one takes a general view of that multi-media systems are those systems that can manipulate mixed forms of sound, pictures and text, then multi-media is hardly a new research area. Telecommunications companies have long been able to support mixed voice and video in a single telephone-like device. In addition, many computer companies have developed highly integrated user interface systems that all have their roots in the seminal work done on the Xerox Star system in the early 1970's. In practical terms, it is possible to purchase personal computers that already offer impressive demonstration packages supporting innovative mixes of pictures, sound, and text. Finally, any visitor to video arcades is quickly confronted with both the benefits and excesses of multiple streams of information presented to (and required from) various types of users. Still, few (if any) of these systems and applications would fall under a definition of a true multi-media system.

One way of defining multi-media systems is to classify their external capabilities in terms of their inherent input and/or output functionality. Consider, for example, the following three classes of multi-media systems:

- *Output-oriented:* this class focuses on the presentation of various types of information to a user; this information may be dynamically synthesized from information within the system or it may be obtained from one or more special input sources, such as video disks or digital audio tape. Examples include mixed text and pictures or mixed pictures and sound used to illustrate the contents of, say, a department store catalogue.
- *Input/output-oriented:* this class focuses not only on the presentation of data, but also on its capture and (possibly) its archiving for later presentation. Examples include systems that time-stamp voice mail on a workstation or ones capable of recording of mouse movements over a piano-keyboard icon, with the tone corresponding to the pointed-at note being played by the computer.
- *Globally-oriented:* this class includes those systems in which *all* types of data can be entered and presented from/to a user, plus it includes all possible types of data integration, migration, and and interrogation as well. Systems of this type extend the notion of internal data transformation, so that data can be converted from one type to another, as well as being captured, archived and presented. One example domain could be scientific visualization; here, different transformations of information are dynamically created to illustrate different relationships that can be detected long after the information is captured.

While an input/output-based definition of multi-media is sufficiently broad to define a wide range of systems, it is incomplete because it does not describe two essential aspects of multi-media systems: (1) the ability to perform *discrete manipulations* on data and (2) support for the notion of multi-media *authoring*. For example, while newspapers routinely provide the presentation of mixed forms of pictures and text, and while even inexpensive greeting cards offer the ability to

use text, illustrations and sound in presenting birthday greetings, neither of these can be included as multi-media systems because the encoded data can only be consumed as a whole, rather than as a collection of discrete parts. As for system authoring, consider a popular flight simulator available for personal computers: while this program supports mixed text, graphics and sound, it is not a multi-media system because all of its data streams are hard-wired. In this respect, flight simulators are applications that make use of multiple output media, but they provide no ways of authoring the relationships among the media they manipulate.

For our purposes, then, we consider multi-media systems to be defined as systems in which multiple streams of information can be captured and subsequently discretely edited, stored, shared, and presented, all under the control of the producer and/or consumer of the information streams. Note that this definition does not take into account the nature of the individual information streams supported by the multi-media system, or the manner in which that data is internally structured. Rather, it concentrates on the notions of collecting "manipulatable" information that can be altered and merged, with an added implied ability to synchronize data in the various information streams.

3.2. Internal characteristics of multi-media systems.

While the internal structuring of information is not considered in the definition of multi-media systems, we do consider it as a measure of the extensibility of the system itself. In order to classify the nature of the internal characteristics of multi-media systems, we use a "generation-based" approach:

- *First-generation:* this generation consists of systems in which the internal representation of information is in raw or unstructured form. This may include raw text, raw images, raw voice, etc. First generation systems are characterized by an "availability" attribute rather than an "flexibility" attribute.
- *Second-generation:* this generation consists of systems that have structured internal data, and thus allow not only binary-style inclusion/exclusion or on/off operations on information, but also the potential for data decomposition and recomposition as well as information editing and information compression.
- *Third-generation:* this generation consists of systems that allow for content-based operations on data, rather than only structure-based operations on data. These operations could include information filtering, information type-conversion (such as speech-based to text-based or image-based to object-based, etc.).

Early multi-media implementations can be described as first-generation systems: information streams in these systems are stored as unstructured data, with the result that only limited manipulation of the data is possible. An example of a first-generation system might be one that uses FAX-like data images to augment a text-based information stream. Although systems of this type certainly qualify under our more restricted multi-media definition, they have the drawback of being inflexible: information can usually be manipulated only via "cut-and-paste" type operations, with only limited content-based filtering and processing available. In contrast, second-generation multi-media systems offer the potential for manipulating the contents of an information stream rather than simply controlling the stream itself. The benefits of using structured data can include: higher reliability by having error detection/correction information carried with the data, increased flexibility by having the structure information provide the basis for searching and synchronizing of data, and better performance by providing special-purpose encoding mechanisms for each of the types of data used. By extension, one can consider third-generation multi-media systems ones in which the contents of the information stream can be further utilized. In third-generation systems, it might be possible to create links between and among streams that are based on information content instead of (or in addition to) information structure, such as analyzing the speech content of a stream to find a cue to an accompanying diagram.

4. Overall Objectives of CWI's Multi-Media Research

In the context of the discussion above, CWI is interested in investigating second and third generation multi-media systems that combine a global view of data with the ability to segment and process aspects of each data stream individually. This goal is our long-term vision of the ultimate utility of multi-media systems. As one of our first steps in understanding the problems associated with processing multiple streams of data in a coordinated fashion, we have defined two projects that investigate different aspects of multi-media research. The first project has as its goal the development of a multi-media kernel that can support the authoring and manipulation of multi-media mail and messages in a heterogeneous environment. The second project has as its goal the investigation of consistent multi-media user interface systems based on image understanding techniques. Each of these projects will be briefly described in the following paragraphs.

In the multi-media messaging project, we plan to investigate techniques for the efficient collection and dissemination of multiple streams of data in a heterogeneous environment. The emphasis of this project is to define a multi-media kernel system based on the Amoeba distributed operating system, with extensions to support the development of multi-media databases and high-speed communication links. In defining the project's goals, we decided on a short-term pilot project to bootstrap our understanding of multi-media issues. In so doing, it was clear that a choice needed to be made among the various types of information that could be carried and in the data structuring options available. While it was tempting to consider third-generation systems supporting content-based links among information streams, we felt that this required a level of investigation that was incompatible with the goals of a pilot project. By the same token, a first-generation multi-media system, while potentially impressive, would not address what we consider to be fundamentally important issues in manipulating information in a heterogeneous environment.. Therefore, the project was designed to be a second-generation system with some third-generation extensions that would deal with audio, still-video, and text data.

The prime purpose of the second multi-media project is to investigate aspects of multi-media systems not directly addressed in the messaging system project. This was done both to extend the range of potential applications being considered at CWI and to verify the usefullness of the multi-media kernel developed for the messaging system. The second project will aim to integrate the use of intelligent understanding of the multi-media data being processed with a sophisticated user interface. This should provide insight into the characteristics of new generation multi-media systems. Specifically, the project will make use of interactive time varying image processing techniques, together with sound output and animated graphics for the visualization and analysis of multi-dimensional data. The major research effort would be to develop a sufficiently broad user interface abstraction that can be consistently applied over a wide range of media. This project would be framed in the context of specific problems associated with the particular application domains of medical image processing and the interpretation of architectural drawings, maps or blueprints.

In summary, both of these projects have been defined to investigate different aspects of the multi-media spectrum, with the goal of supporting a broad, inter-disciplinary research effort at CWI.