Signal Processing for Multimedia
J.S. Byrnes (Ed.)
IOS Press, 1999

Multimedia Signal Processing: Past, Present and the Future

Andrew G. Tescher

LMT Interactive Technology Center and LMMS Palo Alto Advanced Technology Center
Lockheed Martin Corporation

ABSTRACT. Multimedia, as an application, has at its very core the field of signal processing technology. Although multimedia has leveraged on numerous disciplines, signal processing is the most relevant. Some of the basic concepts, such as spectral analysis, sampling theory and the theory of partial differential equations, have become the fundamental building blocks for numerous applications and subsequently have been reinvented in such diverse areas as transform coding, display technology and neural networks. The latter, most recently, lead to a fast implementation of vector quantization. It is evident that the diverse signal processing algorithms, concepts and applications are interconnected and in numerous instances appear in various reincarnated forms (note that, in this instance, reincarnation may not mean immortality, rather ignorance on the part of the reinventor). For example, sub-band coding existed for many years before wavelets became fashionable. In this paper, an attempt will be made to provide a historical overview of signal processing through the present, followed by a highly personal speculation for the future.

1. Introduction

The term multimedia represents many different concepts. It includes basic elementary components such as different audio types. These basic components may originate from many diverse sources. They may come from individuals and they can be synthetically generated. For audio, the synthetics may be traditional musical presentation. One may also argue that multimedia is based on the extended visual experience, which includes representation of the real world as well as its model through a synthetic representation.

In this chapter, a brief historical perspective is attempted to cover the underlying signal processing associated with the multimedia experience. The fundamental difficulty remains to define what multimedia is before we can address issues associated with it.

At one extreme, multimedia is just a label! Conversely, multimedia is a form of human experience: it includes not just accurate representation of the real world but also a world that does not exist if it was not for the multimedia experience. In a special issue of the IEEE Signal Processing Magazine on the topic of this chapter, the contributors spent considerable effort defining the related signal processing issues without really defining what multimedia is. Although this special issue contains several excellent articles, the organizers had the same difficulty in defining the basic issues [1]. Regarding multimedia signal processing the "points we converged are:

- There is not much past for multimedia signal processing since not much has been done.
- There is a lot going on *now*…
- There is a bright future for all of us to explore."

Without really defining what multimedia is, the same article discusses the relevant set of labels as defined by the Technical Committee (TC) on Multimedia Signal Processing of the

Signal Processing Society. The self-defined charter of this group, MMSP-TC is SP-7. It consists of eight major categories, which most individuals would agree are relevant to multimedia, although not necessarily multimedia signal processing. For example, the categories include “Multimedia Applications” as well as “Standards and Related Issues.” The purpose of this present discourse is not to criticize a collection of outstanding researchers, rather to emphasize the difficulty in a meaningful definition of multimedia.

It is all too easy to identify a new field, even if not well defined, with an emerging new technology. This type of definition automatically avoids the need to look back to the past. On the other hand, if we identify the field by examples, an approach this author prefers, the relevant history may immediately appear. It is, however, fundamental that when attempting to write an article on a field as broad as multimedia signal processing a serious attempt be made to define the underlying concept. The “label” is recent, but how new is it?

Leonardo Chiariglione gives the closest definition this author found for multimedia in his recent presentation [2]. Chiariglione argues that the basic concept one needs to consider is the human experience of the real world through his senses. The human experience of the real world projected through time and space lead to Chiariglione’s concept of “virtual world.” This observation further leads to the definition of multimedia: the combination of processes involved in a meaningful interaction of human beings with this virtual world. The principal components of this interaction are creation, delivery, consumption, and access. This definition of multimedia is creative, imaginative and consistent with our experience of what we think the field of multimedia is. Chiariglione uses his definition to project where the field is going without stopping and looking back (of course, he needs not do that as his self-assigned task is to lead us to a future realization of virtual worlds). Interestingly, however, it becomes obvious that a virtual world based definition of multimedia includes a long history of applications even though the term multimedia is recent.

The multimedia concept, based on human experience and perception via virtual worlds, may be new, but what it identifies, in fact, has a significant history. Obviously, the underlying technologies have dramatically changed and will keep changing, but has the human experience always been enhanced by technology? We, being technologists, might strongly argue in the affirmative. It is erroneous, however, to claim advances simply because the final product is based on better technology. For example, it would be futile to convince the producer of a commercially developed movie of high artistic quality that a digitally produced and manipulated video is superior, regardless of its technical sophistication and interactivity. Clearly, in representing a virtual world in realism, the latest digital video has a long way to go to achieve the quality of film-based representations even as far back as thirty or more years.

A comprehensive entertainment “package” fits the virtual world based definition of multimedia well. Was a theatrical presentation of a classical tragedy in ancient Greece multimedia? What about a silent movie with an accompanying pianist in the early part of this century? We will abandon the Greek tragedy in ancient Greece. However, the silent movie with the live music has many of the elements we associate with modern multimedia. It has excellent sound, it is interactive and provides a good synchronized environment.

Consider the most representative of today’s virtual world entertainment. The movie “Titanic” not only was among the most expensive to produce, it represents an example of how real a virtual world experience can be. Figure 1 emphasizes some of the processing issues in the commercial entertainment business.

The Power of Simulations

- Some “Titanic” Facts (SJ Mercury News, March 22, 1998): The “plunge” combined 75-100 layers:
  - computer generated
  - miniature modeled
  - photographic
  - live action
- Virtually every frame was composited (Digital Equipment Corporation)
- 300 SGI computers did the “rendering”
- FX world (www.fxhq.com): 75% of US films have digital work
- Silicon Graphics: 15% of business is entertainment related

FIGURE 1. Computers in the multimedia business

While this movie format clearly outperforms anything we might consider in the pure digital video field of the conventional multimedia domain, the actual production used several of the new technologies associated with modern multimedia. The movie uses real persons and a realistic model to recreate its “virtual world,” the doomed ship, yet many synthetics are used to generate realism. The background, including the ocean and sky are digitally produced. In many of the scenes depicting crowds, the human figures are digitally generated. Obviously we still have a virtual world presentation in a classical format, the film, yet modern multimedia technologies have a profound impact. However, the “consumer” is totally unaware that what he perceives as real in the virtual world is in fact highly artificial.

2. Definition of Multimedia

Based on the previous discussions, we could argue that multimedia is basically a label whose fundamental concept is old, if not ancient. It could be argued that such examples as

silent movies with live music, fireworks together with music and a laser display, and Greek tragedy might well be covered by the multimedia label. These examples may be somewhat unusual to the casual reader, because one traditionally associates the multimedia concept with "new" technologies. On the other hand, if one applies a somewhat more philosophical approach to the definition, these "old" examples quickly fall in place.

The "world" is a perception by the observer (which is a grand collection of all individuals), utilizing his senses for visual impulses, sound, tactile processes and fragrance! Consequently, we propose the following definition as the basic concept: multimedia is the process of perception (real or virtual) through the delivery of components effecting the human sensory process. Today's multimedia is virtually identical to the definition of the concept of an audiovisual system. It is limited by the technology to visual and audio effects. The sense of smell is completely lacking, although it is fundamental to the human sensing process, and little is available through "touching." These comments reinforce the observation that limitations in today's multimedia are not the result of improper definition of the term, rather it is the reality that the technology, in spite of enormous progress, is basically rather immature.

3. Multimedia Elements

The sensed world is perceived by the observer through the composition of fundamental impulses, which are associated with basic human senses. As stated above, current technology limits these components to visual and audio elements.

Although a multimedia presentation is observed in its entirety, generation of the various visual and audio components is fundamentally different. In the following a somewhat simplified hierarchy of the component generation is discussed.

3.1. Generation Of The Real Scene

Although the prevalent definition of multimedia presupposes a digital representation, this is not a true requirement. In fact, recording of a real three-dimensional "analog" scene is inherently analog. Some basic terms associated with representation of the visual component of the perceived world are listed in Table 1.

Audio components can be grouped into two categories:

1. The first is directly associated with the visual part, as for example, the voice of the displayed speaker. Obviously, close synchronization between the visual and audio components in this case is absolutely critical.
2. For audio elements that are basically independent, close synchronization may not be necessary. On the other hand, the audio part itself is a composite of different components with varying degrees of synchronization requirements.

Modern multimedia is likely to be rather complex, consisting of many different components, visual as well as audio. Even in the conventional case, where composition of the final product is completed prior to any possible transmission, maintaining an accurate synchronization among the various elements is critical.

TABLE 1. Multimedia concepts and definitions

Basic Concept	Definition
Film	Even today most professional recordings are made on film. Understanding of the final scene representation, including those represented digitally, is a basic process in creating high quality multimedia
Camera and other recording instruments	Together with film the basic recording medium, the actual image generation system has been the fundamental device for generating visual perception.
Stereo Composition	Although the concept of stereo — including associated devices — has been around since the beginning of the century, its utilization in most multimedia delivery systems is limited. Again, one can argue that today's presentations are constrained by the basic limitations of current technologies.
Color	The ability of the human eye in sensing the range of colors is poorly approximated in current multimedia technologies. Considerable evolution will be required before the digital implementations will catch-up with the rich film based presentation of many years past.
Hierarchy of Composition	The perceived scene can always be decomposed in certain hierarchical representations. An effective multimedia presentation process is able to utilize and optimize interrelationships among the relevant components
Models and "Augmented Reality"	Multimedia systems are not restricted by the recording of real scenes. Visual components, based in part or in their entirety on virtual models, are inherent in an effective multimedia presentation

3.2. Multimedia System Issues

It is evident that more recent versions of multimedia generation differ from the simpler versions of the past, in that a careful procedure must be developed for handling "system" issues. These system issues can be categorized as follows:

1. *Relationship of elementary components and scene composition.* Maintenance of proper synchronization as well as all intercommunication among the individual elements is the basic role performed by the underlying multimedia system. For more complex multimedia composition, involving extensive communication infrastructures, system issues become the primary drivers in designing the multimedia product.
2. *Delivery mechanism.* Unlike simple systems of the past, a multimedia composition involves transmission to the customer at a remote location. The actual delivery process covers numerous different categories: a) real and non real-time delivery; b) communication channels of numerous types, quality and bandwidths; c) single and composite channels. The delivery mechanics is highly dependent on the supporting technology. Evolution of digital communication technology has an obvious profound impact on multimedia system design. Also, evolving storage components such as the compact disc (CD) and its more recent enhanced version, the digital video disc (DVD), have significant influence on multimedia technologies.

3. *Human interaction and feedback.* Although not a strict requirement, it is assumed that a modern multimedia application is likely to include active participation by the end user. For the underlying multimedia system operation this additional dynamic process, which allows feedback yet remains consistent with the composition and delivery mechanisms, generates a highly demanding scenario. Perhaps, it is this interaction, that fundamentally identifies modern multimedia.

4. Technology Issues

Modern multimedia is inherently digital. The maturity, limitations and expectations of this digital technology have a profound impact on all aspects of multimedia. Although all significant advances of this technology are relevant to our topic, we must restrict the discussion to key categories:

4.1 Recording Technology

The information throughput relevant to recording real scenes is high. Digital processing of visual information requires stable and high speed electronic processing. A principal component is the 8-12 bit dynamic range of the required A/D converter for typical visual applications. Special applications, such as digital x-ray recording require even higher dynamic range. Representation of high quality color introduces additional demands on the digital recording process. It is an interesting observation that the A/D conversion process introduces an inherent mapping of multidimensional information into one dimension.

A secondary effect of digital recording of raw digital information is the generation of potentially unmanageable amounts of data. While the purist (here we refer to anyone with extreme interest in the original information) dislikes compression, it rapidly becomes the "necessary evil" if modern multimedia technology is to become practical.

4.2. Delivery Technology

Some of the basic delivery concepts were introduced in section 3.2., under system issues. Fundamental ideas include transmission of information via communication channels, utilization of storage media, and maintenance of an ongoing high degree of synchronization with or without observer interaction.

From a communication perspective, the key is cost-effective availability of reliable high bandwidth. It is obvious even to the most casual observer, that limited availability of such bandwidth represents the fundamental hindrance in the evolution of multimedia technology. A separate but closely associated technology issue is networking. Implementation of effective networks over the available physical channels is another critical component of the overall delivery scenario.

One may argue that bandwidth/network concerns are likely to be eliminated, or at least minimized, based on the significant progress being made in these fields. Both networking technologies and physical communication channels are evolving with various, and sometimes competing, new methodologies. If we are to see a true convergence of advanced multimedia technology delivery systems that includes mature video-based entertainment systems and computer technologies, enormous technological improvements are required.

4.3. Delivery Of The Final Product

The actual multimedia presentation is accomplished through a physical display/audio system, which may contain additional processing functions. As with other elements of the delivery chain, the visual element is the critical component. Traditional display terminals have undergone significant changes in recent years. Besides the improving economics associated with display technologies, critical parameters of the physical display devices have significantly improved. These parameters include both temporal and spatial resolution, color representation, size, and dimensionality for (visual) stereo. Visual stereo presentations are significantly less common even though their role in an advanced multimedia presentation could be very effective.

The role of pre and post processing is likely to become increasingly common as the receiving terminals become more specialized for handling digital information [3]. Certain functionality has been simply adopted from conventional analog video and represents various calibration procedures. However, the ability of the receiver to perform significant amounts of digital processing considerably enhances the overall performance of the receiving terminal.

It should be noted that proper preprocessing could play a significant role in the enhancement of the underlying compression process. The role of the human visual process for image coding has been considered for some time. It is evident that, with the availability of increased processing power at all levels of the delivery chain, pre and post processing could play a significant role in optimizing the entire multimedia system.

5. Signal Processing Elements

Many, if not most, classical signal processing procedures have become deeply embedded in multimedia. A key driver is the optimization of multimedia components such as representation, storage and delivery. Procedures range from the very simple to sophisticated optimization methods. Some of the principal techniques are:

5.1. Non-Linear Analog (Visual And Audio) Mapping

This procedure may be purely analog. Its motivation may be the desire to enhance the delivery process. It could also be introduced to mask the limitations of various components of the overall multimedia chain. Typical constraints are introduced by bandwidth limitations and constrained dynamic range in the display terminal.

5.2. Quantization Of The Analog Signal

This procedure is fundamental to any digital representation of a signal which originated in the analog world. The quantization process is inherently a lossy procedure and fundamentally non-invertible. This classical signal processing element, although not very exciting compared with other multimedia issues, remains a basic constraint in limiting performance [4]. Quantization techniques comprise a whole field by themselves. The major relevant issues include linear and non-linear techniques, and adaptive and non-adaptive procedures. Finding general optimization procedures to minimize the associated distortion in the quantization process is still a fundamentally unsolved problem.

5.3. Statistical Characterization

Statistical concepts and applications are, directly and indirectly, strongly embedded in processing components associated with multimedia. This field is part of classical signal processing. We can only highlight the major categories:

1. Spectral analysis is fundamental to the entire range of image modeling, filtering and algorithm design. The procedures are critical to both visual and audio components.
2. Statistical redundancy is the basic concept upon which the entire field of data compression is built. Mathematical extension of the concept leads to an "optimum transform" for decorrelation, which in turn leads to the entire field of modern transform coding.
3. Model-based representations, primarily for compression techniques, are derived from assumed or derived statistical models. Entire classes of transform coding algorithms are based on this technology [5]. The technique includes evaluation of elementary model components by the human observer.
4. Utility of the Fourier transform and its discrete extensions (DCT, wavelets and others) are based on the principal that these transforms "asymptotically" approach the "optimum" transform assuming a reasonable "statistical" behavior [6].

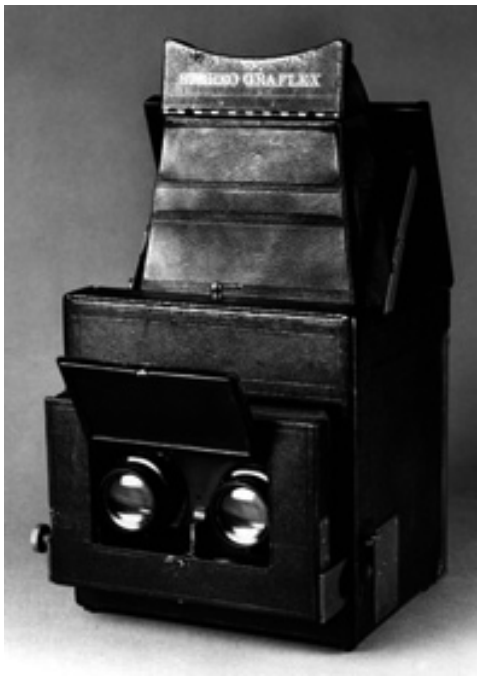
5.4. Motion Representation And Models

Efficient representation and development of the relevant algorithm is probably the most critical element of modern multimedia. Effective utilization of algorithms associated with motion has a significant impact for two separate application areas.

1. Implantation of motion detection and associated compensation in subsequent image frames can significantly reduce the required bandwidth. Successful prediction of image segment locations in future frames reduces the required information update to the required motion vectors. Thus the associated update information under this assumption is dramatically reduced. It is not surprising that ongoing development of effective motion prediction and compensation algorithms remains a high priority. This technology has been consistently embedded in the standardized video compression systems.
2. Effective combination of the presence of motion in video segments and limitations of the human visual systems provides additional bandwidth reduction potential. Since human vision deteriorates when observing moving areas, image blur associated with those regions becomes significantly less noticeable. Consequently, additional image compression can be introduced in segments that contain image motion with minimal noticeable effect. Although video compression systems indirectly utilize this observation, direct research activity in this area is significantly smaller than for the previous application area.

5.5. Color Processing

With regard to color processing, the traditional signal processing community has not fully adapted the significant amount of results achieved either in the vision sciences or in several important applications areas such as printing.



- From the California Museum of Photography (CMP/UCR). Introduced in 1904. Two 6 inch Bausch and Lomb f/6.3 Tessars with iris diaphragms and a focal plane shutter with 1/10 to 1/1000 of a second shutter speeds.
- Folmer's Patent, April 5, 1904. The stereoscopic form of the regular 5x7 Graflex Camera. Using many of the same film holding accessories the Stereo Graflex creates two images on a 5x7-film format. Unique design features for this camera include a rising front that allowed the photographer to crop the foreground when desired. Focusing through the lenses created the stereoscopic effect when focusing.

FIGURE 2. “Early” Stereo Camera

Traditionally, correlation properties among color planes are utilized in image and video compression algorithms. An ad hoc approach to color was implemented in the original analog NTSC television standard, which is likely to stay with us for some time. On the other hand, major problem areas remain, such as matching the real world to the available color space of the display medium. The derivation of optimum quantization models represents a relatively modest effort by the signal processing community to attack such problems.

5.6. Three-Dimensional Representations

Human vision is basically three-dimensional. We see a three-dimensional world through our eyes. The stereo process, the principal aspect of human vision, is one of several ways we may replicate presentation of the three-dimensional aspect of scenes in multimedia. Efficient representation of three dimensions is a major challenge of multimedia. It is also an extremely active development area in numerous applications. The following primary approaches to three-dimensional representation should be considered.

Stereo projection is the best known. Cameras for stereo recording (Figure 2) have been available commercially since the beginning of the century. The basic principal has not changed. The same three-dimensional scene is recorded from two slightly different perspectives, essentially replicating our eyes. The two separate recordings are subsequently presented to each of our two eyes individually. The actual technology includes numerous different viewers, from the inexpensive, to the latest commercial version presented in the IMAX theaters, which utilize infrared controlled shutter mechanisms. Unlike the early stereo film based stereo recordings, modern techniques are heavily dependent on digital processing, which corrects for camera projection inaccuracies and results in significantly enhanced stereo display.

1. Projection techniques exist to effectively recreate multidimensionality from individual projections of an original object. Although this technology has been successfully used in medical applications, its application to multimedia is not likely to occur in the near term. The primary limitations are complexity, lack of easy real-time implementation, and the availability of less complex alternatives.
2. Most common three-dimensional techniques are essentially based on two-dimensional displays. The three-dimensional scene is projected onto two dimensions in the "rendering" phase of the multimedia chain. The correct hierarchy of object elements and behavior maintains the three-dimensional illusion. The relevant processes include shadowing and preserving the proper "hidden body" behavior. The required processing resources are still significant. However, they more easily lend themselves to a standardized set of components. A substantial industry produces various processing components, such as chip sets and graphics boards, to develop solutions for many diverse applications including desktop computing. The associated technology is very effective in high-end applications. "Virtual reality" models utilizing large screens are impressive even though the presentation remains two-dimensional.

6. Technology Resource and Maturity Constraints

Modern multimedia is possible because of the maturity of numerous technical components. The underlying technologies, which make multimedia feasible, are in fact driving this major commercial applications field. Although prioritization of the various individual technology elements may be debatable, some of them clearly stand out as being markedly significant. Some of the principal categories are:

6.1. Transition From Analog To Digital Technologies

Although we take digital technologies for granted, particularly for signal processing applications, in fact the transition is far from over. Even today, the major element of multimedia video and audio in commercial television depends on analog delivery. It is likely to be several years before the transition to homogeneous digital technology occurs in all multimedia application areas.

6.2. Bandwidth Limitations

This issue remains a significant concern for multimedia technologies. The reality is that the average customer is not likely to have any effective network connection beyond the hybrid analog phone connection in the near term, limiting the incoming transmission to 56 Kbps. New technology is available and new approaches are being developed continuously, but wide scale implementation is lagging. ISDN, widely used in Europe for many years, has not been fully utilized in the United States.

6.3. "Cheap" Processing Power

Processing power is critical for advanced multimedia applications. This element of the multimedia chain continues to progress significantly. Today's desktop processors exceed mainframe computers of the recent past in processing power. And the end of progress is nowhere in sight!

TABLE 2. Prominence of the Internet

Recent Internet Highlights
• Netpulse Communications is to provide Internet access to Stairmasters and Lifecycles at 24 Hour Fitness • AOL market capitalization is 13.8 B > Washington Post + NY Times + Dow Jones Co. • On-weeknights active AOL members ~ 650,000, same as MTV or CNN • Ref. March 22, 1998 SJ Mercury News and www.pathfinder.com/fortune/1998/980330/aol.html

6.4. Robust Networking Environment

This resource is a must for effective multimedia. Although major advances have been made, reliability limitations remain a major hindrance (justifiably so) to the "convergence" of broadcast and computer technologies. Competing high bandwidth protocols have not yet reached an optimum solution. Although multiple protocols should not hinder application development, in a practical sense they do. Investors are concerned in committing the major funds needed for the necessarily huge infrastructure (e.g., satellite networks) while the experts argue about the optimality of the underlying protocols.

6.5. The Internet

The Internet (Table 2.) has probably become the most important factor for the entire multimedia application infrastructure. Will it be the primary mechanism to deliver multimedia in the coming years? Any answer may be arguable! However, many companies (Microsoft, IBM, America Online, etc.) have considered the Internet as the principal part of their strategic plan, and their investments reflect that.

6.6. Display Technology

Display technology is the sleeping giant. Although the display is the critical element in the multimedia chain, it seems to receive less attention than components associated with processing and communication technologies. Nonetheless this field has produced significant advances. High quality flat panel displays, although still expensive, now appear in numerous applications. They have become indispensable in mobile computing and they find application in the latest aircraft video displays. The utilization of large size flat panel displays for consumer entertainment devices is only held back by the high manufacturing cost.

6.7. Specialized Components

Specialized components can have major influence on implementation. Among the many new hardware components, DVD deserves a special mention. Although the full impact is yet to be felt, this technology permits the "cheap" recording of up to 20 Gbytes of information on a single disc with current technology and will have a profound impact on multimedia. Besides its utilization as a recording medium for traditional multimedia content, its importance as a reference medium is also likely to be critical until delivery channels of reasonable cost, bandwidth, and capacity are available, which is not likely to be soon. It can be postulated that many new applications based on new multimedia processing scenarios, such as MPEG-4, will appear on DVD before appearing in a regular communication infrastructure.

6.8. Standardization

"Standardized" processing is becoming routine for multimedia applications. Standardization usually refers to a formal process through various standardization bodies, although some proprietary systems (e.g., the Microsoft Windows operating system) have such dominant influence that they have become de facto standards. Efficient delivery procedures have been commonly accepted for most elements of multimedia components. The best-known standards, MPEG-1, MPEG-2 and JPEG, cover a broad range for visual (static as well as dynamic) and audio components as well as the associated delivery mechanisms. The success of these standards, particularly MPEG, extends the process to significantly broader areas. The currently finalized MPEG-4 includes considerable synthetic modeling capability available only in high prized proprietary software (Figures 3,4 and 5). Through its "structured audio" this synthetic modeling is extended to audio as well. The currently formulated MPEG-7 will extend common processing technology to visual databases. The underlying technology is only partially available and needs to be further developed. As a separate comment, it is significant that the key standards organizations have become leaders in technology development, unlike their earlier role, where they primarily endorsed the "optimum" practical implementation based on available technology

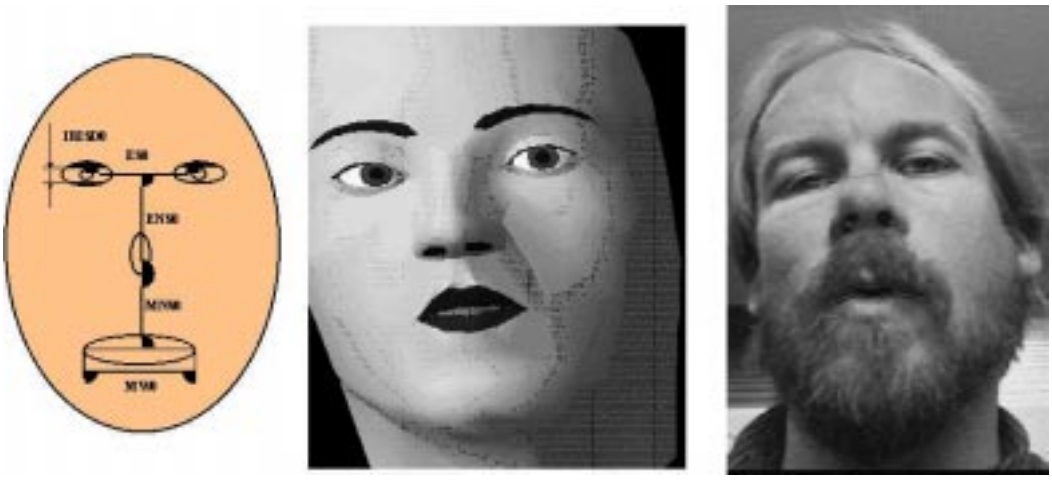


FIGURE 3. Face Animation (from the presentation by P. K. Doenges, "Synthetic/Natural Hybrid Coding Mixed Media Content in MPEG-4". MPEG-4 Seminar, San Jose, CA., January 31, 1997)

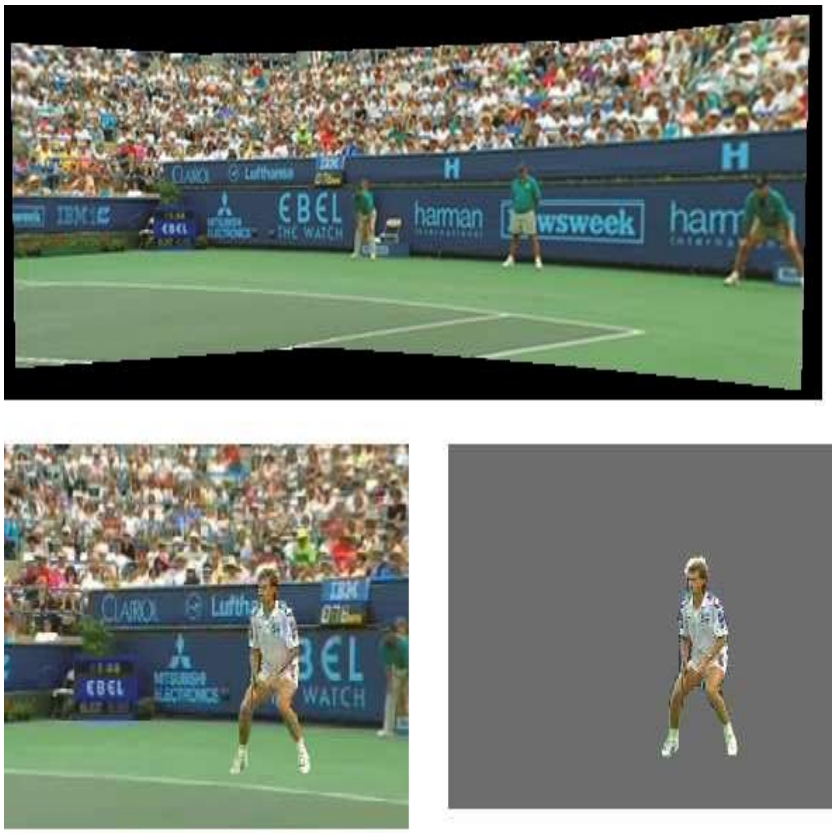


FIGURE 4. "Sprite Coding" (from the presentation by T. Ebrahimi, "Natural Video Tools in MPEG-4." MPEG-4 Seminar, San Jose, CA., January 31, 1997)

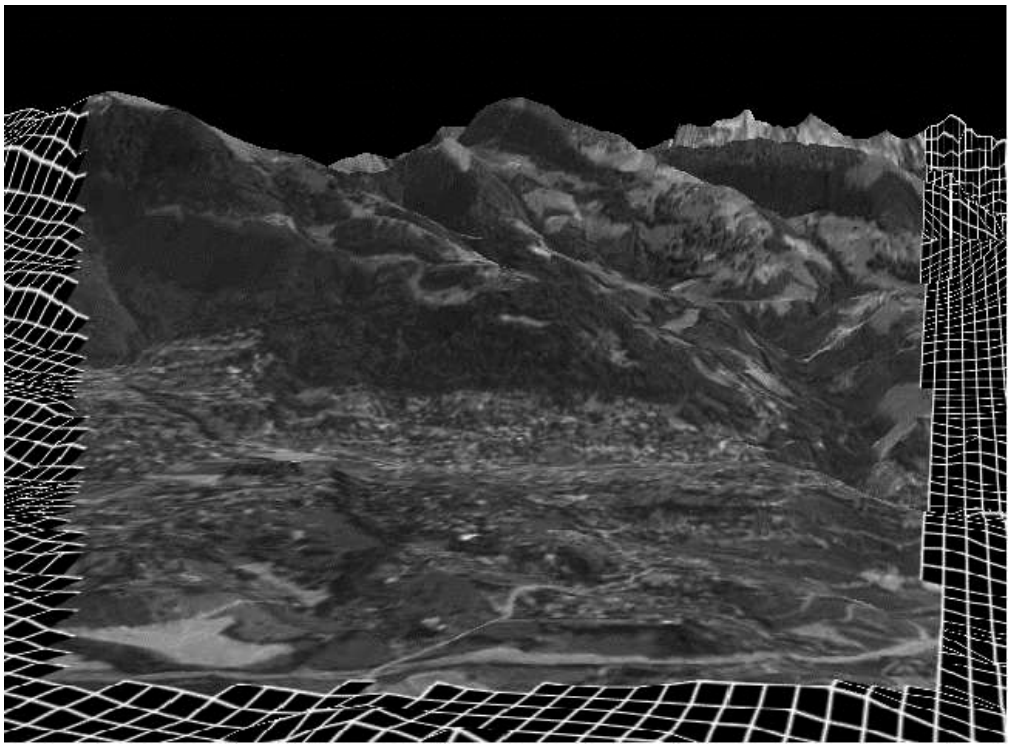


FIGURE 5. Map on Mesh (from the presentation by G. Rajan, "SNHC Visual." MPEG-4 Seminar, San Jose, CA., January 31, 1997)

6.9. Alternate New Delivery Mechanisms

Alternate new delivery mechanisms under consideration are likely to have a major impact on multimedia technologies. Some of these implementation scenarios, although requiring major investments, are obvious. Perhaps the most obvious is the development of satellite based networking infrastructures. These systems, once operational, will provide flexible, reasonable-cost bandwidth with global availability. Direct impact on multimedia implementations is likely to be significant because of the homogeneous infrastructure of these systems as well as the associated global economic resources. A less well-known yet important implementation scenario involves the power companies utilizing their very extensive power line grid for communication networking. Equally important is the observation that these newer delivery implementations will provide significantly enhanced competition in the market place, which is still dominated by the traditional communication companies not known for their aggressiveness in inserting new technologies.

7. Application Drivers

A wide variety of applications are likely to be developed based on evolving multimedia technologies. It is somewhat premature to predict what the “killer” applications will be. It is expected, however, that applications not currently in use will emerge in the coming years. Based on current applications, the following major fields can be extrapolated to the future with reasonable certainty:

7.1. Entertainment

Entertainment is likely to dominate multimedia applications. The critical requirement is the growth in technology resources: increased bandwidth, display technology enhancement, and the associated networking and software support. Consumers will rapidly expect and demand significant improvement in various entertainment products. In addition to improved technical quality, continually increasing interactivity is expected to be part of the home video field.

7.2. Electronic Games

Electronic games are already popular and their popularity will further increase, as applications may be presented over enhanced delivery media.

7.3. Tele-Medicine

Although this field exists today to a limited extend, it is an obvious candidate to use the tools of multimedia. It is further driven by the increasing pressure to reduce medical costs.

7.4. Distance Learning

This application is an another existing field which is likely to expand as the supporting technologies mature. There are interesting social implications, as students will have the technical options to take courses at a large number of educational institutions at numerous diverse geographic locations, virtually simultaneously.

7.5. Electronic Commerce

This application is becoming a major economic driver on the Internet. Wider availability of enhanced multimedia tools will further fuel the explosion of this application. In fact, one can expect increasing numbers of business categories to migrate to the electronic media. Real estate, travel, and automobile sales are among those areas where in-person shopping will continuously diminish.

7.6. Digital Libraries

The “digital libraries” initiatives, converting traditional archives to digital media, will significantly expand with the maturing of multimedia. Many additional standards initiatives for visual database description and applications, such as MPEG-7, will provide significant additional resources.

7.7. Telepresence

The term “telepresence” is the current focus of several multimedia studies. It is the conspicuous evolution of the “old” video teleconferencing. The primary objective is to enable participants at spatially different locations to interact as if present at the same place. New multimedia tools are required for this technology. Previously listed applications, such as tele-medicine and distance learning, will probably become sub-categories of this field when it matures.

7.8. High-End Applications

For high-end applications, such as “virtual sets,” the movie making business is likely to be the continuing beneficiary of modern multimedia (Figure 6). Digital technology has already become a common processing step in making professional movies. The general evolution of this field together with the availability of the required resources will permit significant additional techniques in movie making. Replacing the “real set” with a virtual one is likely to become much more common with the maturing of this technology.

FIGURE 6. The "virtual" set (from the presentation by P. K. Doenges, “Synthetic/Natural Hybrid Coding Mixed Media Content in MPEG-4”. MPEG-4 Seminar, San Jose, CA., January 31, 1997)

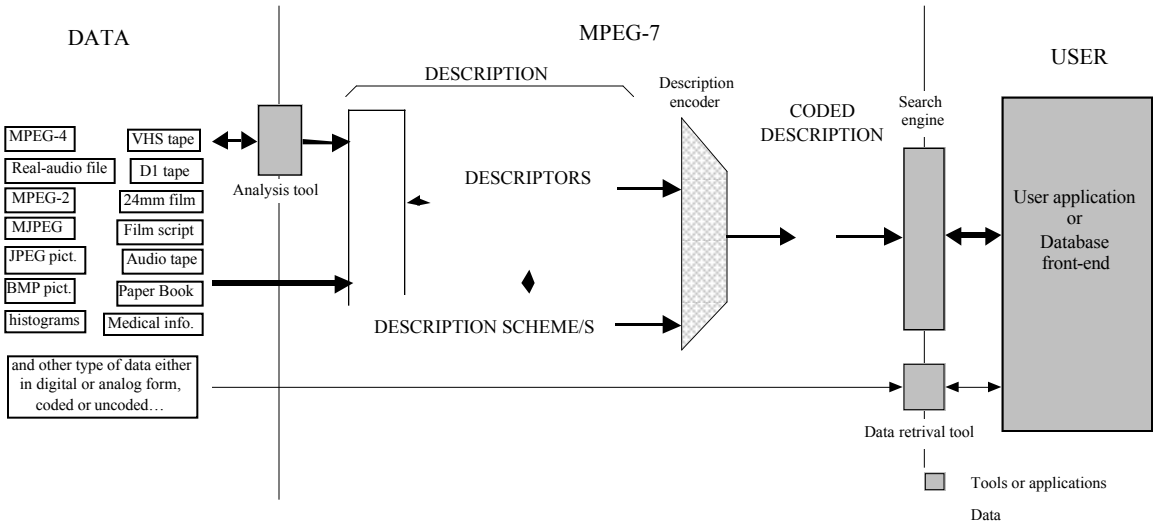


FIGURE 7. MPEG-7

8. Multimedia Databases

The term “database” significantly underestimates the tremendous challenge that lies ahead for multimedia applications. By its nature, multimedia is complex. Its components are partially or fully independent and it requires extensive resources for storage and transmission. A major contribution to this complexity is that recent trends in multimedia management systems favor object oriented manipulation (unlike conventional systems such as broadcast television.) The currently finalized MPEG-4 embodies this trend. This approach views the final multimedia software product as a composition of basic elementary components. An associated database management system must handle the individual components. The complete information management problem becomes significantly more complex.
However, for owners of multimedia “content” to permit what amounts to control of their property by third parties, this problem must find a credible solution. MPEG-7 is a new initiative in the standards community, which is likely to address many of the relevant issues. Figure 7 provides a graphical overview of the connection between the data sources and the user [7]. However, MPEG-7, similarly to the preceding MPEG projects standardizes the decoder. Although this project is likely to be of significant value, it will not provide the complete solution. The complete process to be developed on the “transmitter” side, similarly to other MPEG systems, will reside in proprietary implementations. The following specific challenges need to be addressed:

8.1. Characterization

Elements of multimedia information need to be characterized. The challenge is considerable and just beginning to be addressed by ongoing research. The requirement is to identify meaningful features associated with multimedia information. Simple examples include conventional characterization of imagery: color, component information (description of specific objects), and spectral description of the spatial content. For speech, it may include language and speaker characterization (age, sex, etc.). Development of even a partially automatic process generating credible decompositions is likely to be the subject of ongoing research for many years to come.

8.2. Indexing

Indexing of historical information will become an obvious requirement for multimedia delivery systems to achieve the level of efficiency required for extensive (and profitable) commercial implementations.

8.3. Information Management

Associated information management systems must be developed for advanced multimedia scenarios. These information systems are likely to evolve from current familiar Internet search engines. When (and if) they become successful, these systems could have a major impact on the entire range of delivery of modern multimedia information.

8.4. Creative Reconstruction

Tools for decomposition and reconstruction based on elementary components can add significant efficiency enhancement to the delivery mechanism. In a “server-client” model, text reconstruction is based on the “nearest” font. Equivalent approaches may be considered for more complex audiovisual composition. MPEG-4 has introduced the “sprite” concept. For this implementation, only the foreground is transmitted and continuously updated. The background is derived from the original scene and it is transmitted only once. Also, it may be generated from other available information.

9. Conclusions

In this opening chapter, an attempt was made to assess multimedia technologies. In particular, the relevance of signal processing to multimedia was addressed. As important as multimedia technology is today, its further dramatic growth seems assured. Technology constraints, such as communication bandwidth, processing components, availability of open standards, and display technologies are being successfully addressed by industry. Signal processing technologies have made important contributions to all areas of multimedia and will continue to do so. Signal processing has been particularly important in all areas of compression, modeling, and in the entire field of digital representation of complex signals.

References

[1] The Past Present and Future of Multimedia Signal Processing, IEEE Signal Processing Magazine. Vol. 14, no 4, pp. 29, July, 1997.
[2] Leonardo Chiariglione, Invited Plenary Presentation to the 1998 SPIE Symposium on Electronic Imaging, San Jose, CA. January, 1998.
[3] T. N. Cornsweet, Visual Perception, Academic Press, New York, NY, 1970.
[4] A. N. Netravali and B. G. Haskell, Digital Pictures, Plenum Press, New York, NY, 1988.
[5] J. L. Mitchell, W. B. Pennebaker, C. E. Fogg and D. J. LeGall (editors), MPEG Video Compression Standard, Chapman and Hall, New York, NY, 1996.
[6] W. B. Pennebaker and J. L. Mitchell, JPEG Still Image Compression Standard, Van Nostrand Reinhold, New York, NY, 1993.
[7] “MPEG-7: Context and Objectives,” ISO/IEC JTC1/SC29/WG11 N2082, February, 1998.