Multi Resolution Video Compression System

Venson M. Shaw
School of Applied Science
Nanyang Technological University
Nanyang Avenue, Singapore 639798
asmhshaw@ntu.edu.sg

Samuel S. Y. Sung
Department of Information Systems and Computer Science
National University of Singapore
Kent Ridge, Singapore 119260

Abdul Wahab
School of Applied Science
Nanyang Technological University
Nanyang Avenue, Singapore 639798

I. Abstract

This paper relates to a general purpose system architectural method for multimedia communications. The object of this paper is to improve the quality and efficiency for human communications. Our architectural method allow for the access of a plurality of computing, consumer, and communication equipment, e.g., PC and workstations, camera, television, VCR, telephone, etc, and allow for conveying multiple types of media information, e.g., sound, image, animated graphics, and live video.

Despite of the real-time constraints and resource limitation to store, retrieve, and exchange these massive media data information, an efficient architectural method was developed to make multimedia communications system a final reality.

II. Introduction

This paper relates to a general purpose architectural method suitable for most conceivable combinations for multimedia communications. PC workstations are widely available at most offices and homes today, yet due to their processing and storage limitations, they were never considered for complex image / live video applications. Alternatively, existing methods employee single media communications. Namely, telephone for human voice communications, fax for text communications, or PC workstations for data communications.

Noticeably all of these single-media communications use existing analog telephone lines connecting through the central office (CO) switch, only one of the media types can be selected at a time, and the fax and PC's use dial-up modem for analog transmission of the digital data.

Meanwhile, various coding techniques are available today so that source media (image, live video, sound, and animated graphics) can be reduced (coded or compressed) into lesser quantity
to ease the storage and transmission constraint, and the destination media can be restored (decoded or decompressed) and playback without quality degradation, then such digital coded media information can find wide applications for remote database retrieval, teleconferencing, messaging, distance education and other

applications to complement traditional single media (voice, data, and text) communications.

We now turn to the reviewing of existing product and technologies. Various single-media codec (compression and decompression) techniques has matured in recent years to allow the high reduction (compression) of the source media and the quality playback (decompression) of the destination media. Individual international standards (CCITT and ISO) will soon be established to facilitate the worldwide communications of still image, quality sound, live video, and animated graphics.

The multimedia products we have searched to-date are either video conferencing systems (i.e. CLI, PictureTel) using dedicated systems and complex algorithms for quality video and audio only, or incorporate desktop PC workstation for a one-way, decode only (playback and display) mixed media presentation (DVI, CDI et.al). Video phones (Sony, Panasonic, et.al.) have been the only communications product which utilize real-time coder and decoder for image and voice transmission through traditional analog or digital transmission, However, their quality are poor, and effects are limited. In conclusion, the present products and technologies involve either real-time playback of the precoded compressed data (live video, sound, and graphics) for a multimedia presentation, or the real time coding and decoding of live video and voice for a live conferencing applications.

Accordingly, we feel it is superior to provide digital media communications in conjunction with the traditional voice and data communications because it combines the use of live video, graphics, and audio media, therefore make up a much more effective means for human to communicate with each other. Since "single picture worthies a thousand words", it is conceivable that pictorial information such as image and live video can definitely enhance and complement the traditional communications.

II. System Overview

The initial objective of our paper is to allow for PC/WS (PC or workstation) as a single platform technology and to define an integrated architectural method which accommodate communications (remote transmission and retrieval) for all types of digital coded (compressed) multiple-media information.

Our paper next provide a flexible architecture which allow for management and control of the variable communications bandwidth and address the flexible combinations of the digital coded mutiple-media information for a wide variety of application requirements. Some of the applications examples are distance education (teaching and learning), teleconferencing, messaging, videophone, video games, cable TV decoders, and HDTV. Our paper further provide the application of digital coding techniques for reducing the storage and transmission requirements for multiple media information, we also suggest the conversion of digital compressed media to analog form for convenient interface with the traditional analog storage or transmission techniques.

Our system further make use of the combinatorial use of animated graphics and motion estimation / compensation for regeneration of the live video. Namely, animated graphics techniques will be applied for the playback of estimated motion effects.

One of the important features of our system is the interactive use of multiple media types. Namely, the user has the control to program and select the appropriate media combination for specific application needs either before or during the communications session. For examples, the user can decide to select the live video with voice quality audio before the session starts, but during the session, he can choose instead to use the high quality audio with slow motion and still freeze pictures for more effective communications.

We designed our system to leverage with all of the available international standard codec technologies, and evolve into a human interactive communications model, and conclude with a low cost, high quality, highly secured, interactive, yet flexible, and user friendly method for desktop, handheld, or embedded media communications.

We have developed within our system a cost effective method for transmission bandwidth

and local storage. Coding techniques have been used to conserve storage and transmission bandwidth since the media information data can be greatly reduced. These coded information still preserve the original quality and allow for presentation at selective quality levels at users request. Since these information are coded according to selective algorithms, without the corresponding decoder, information can not be properly decoded and used, this allow for high degree of security for special applications.

When we designed our system, we feel it is essential to provide implementation for selecting one of a plurality of multiple quality levels for live video, graphics, audio, and voice. Depending on the application requirement, user can select the appropriate media quality as desired. For example, high quality audio and high quality image and graphics may be suitable for collage education, voice combine with live video will be suitable for K-12 education, face to face video and voice will be effective for business negotiations.

Since the regular phone line is constrained in bandwidth, our system can conserve transmission bandwidth, still image can be blended with locally generated live background video or animated graphics. User can instaneously adjust the quality levels during the sessions to make the meeting or presentation more effective.

III. Performance Evaluation

The significant difference between our process and the traditional video conferencing is that only photo images of the conferees (talking heads) have been shown on a traditional video conferencing/videophone setup. In our method, the conferees are allowed to substitute the conferee photo images with other important pictorial information retrievable form the database and present (broadcast) to others for better illustrations. The conferees also have the control to select the appropriate quality level that he or she wants in order to conserve bandwidth.

As an example, for a product presentation, it is better to provide coarse quality live video with high fidelity audio as a introduction. Once specific interests are generated, fine quality video without audio can be presented to facilitate further discussions. The other example is an international meeting while different languages are used, live video can always make ease the verbal explanation, and quality audio can harmonize the atmosphere during tense moments. To further conserve the bandwidth, live coarse video can overlay with locally generated fine quality still background image to provide acceptable video presentation (Notice that the fine quality video will be locally generated therefore doesn't consume any communications bandwidth).

Finally since all coded multimedia information will require proper decoder to expand back to the original presentable forms, therefore it is highly secured, furthermore, different security level can be assigned to each conferee, therefore appropriate information will only be shown to various audience without any concerns on security.

In summary, television only facilitate an traditional analog video and audio session, since it is one-way non-interactive communication, receiver can only observe and listen, they can not make comments or edit (remark) a media message, not to mention the ability to control (select and edit)
the appropriate media massage and return to the sender. These interactive capabilities will be extremely beneficial for distance learning, or remote classroom applications.

IV. Conclusions

The multimedia products to-date are either video conferencing systems using dedicated systems and complex algorithms for quality video and audio only, or incorporate desktop PC workstation for a one-way, decode only (playback and display) mixed media presentation. Videophones have been the only communications product which utilize real-time coder and decoder for image and voice transmission through traditional analog or digital transmission, However, their quality are poor, and effects are limited. In conclusion, the present products and technologies involve either real-time playback of the precoded compressed data (live video, sound, and graphics) for a

multimedia presentation, or the real time coding and decoding of live video and voice for a live conferencing applications.

We have developed a digital media communications in conjunction with the traditional voice and data communications because it combines the use of live video, graphics, and audio media, therefore make up a much more effective means for human to communicate with each other. It is conceivable that pictorial information such as image and live video can definitely enhance and complement the traditional communications.

References

1. A. N. Netravali & J. D. Robbins, "Motion Compensated Television Coding" Bell System Tech. J. , Vol 58, 1979, pp 631-670
2. S. Sivakumar,E. Kappos, S. Y. Sung, and V. Shaw, "A Multimedia Database for Computer based Learning,", ISMM-International Conference on Distributed Multimedia Systems and Applications, Hawaii, Aug., 1994.
3. S. Sivakumar,E. Kappos, S. Y. Sung, and V. Shaw, "A Client Server Multimedia Communications System for ISDN", ISMM-International Conference on Distributed Multimedia Systems and Applications, Hawaii, Aug., 1994.
4. H. Radha, "A Multiresolution Approach to Binary tTree Representations of Images, Proceeding IEEE ICASSP, 1991, pp 2653-2656