Published in the Proceeding of the International Conference on Computer Assisted Radiology, CAR'96, June 1996, pp579-584.

Focus Driven Radiological Image Transmission For Tele-Diagnostics

Javed I. Khan and D. Y. Y. Yun

Laboratories of Intelligent and Parallel Systems
Department of Electrical Engineering, University of Hawaii at Manoa
Holmes# 492, 2540 Dole Street, Honolulu, HI-96848, USA
javed@hawaii.edu

This paper presents our research on a perceptual encoding scheme for fast high resolution medical video image transmission based on Advanced Satellite Network. In this scheme to satisfy the bandwidth/ quality requirement of diagnostic medical video, the focal information of a viewer in an image is traced by an infrared eye-glance tracer or by direct interactive input from the viewer. The focal information is then sent back to the encoder at the transmitting end. The encoder, depending on the location of focus performs a spatially variable resolution encoding, and transmits the image to the decoder at the receiving end which regenerates the image with appropriate distribution of resolution.

1. INTRODUCTION

The principal challenges of next generation life-critical medical information system are (i) rapid inter-connectivity across geographical barrier and distance, (ii) high speed and high volume data networking, and (iii) rapid, interactive and intelligent information access and utilization.

This research is a part of our initiative [YuGa96] which attempts to facilitate on demand long haul high performance medical information and resource sharing with high-bandwidth advanced communication network. The center-piece of this initiative is a new satellite communication technology ACTS (Advanced Communication Technology Satellite). ACTS provides a number of unique advantages those are particularly important in life-critical medical services of teleradiology, diagnostics and distributed treatment planning.

In this initiative, we are demonstrating a number of medical resource sharing experiments involving supercomputing power (at the Ohio Supercomputer Center and the Maui High Performance Computer Center), 3D volumetric modeling technology (at OSC and University of Hawaii, Image Information Lab) and medical image data sharing (at Georgetown University Medical Center and Tripler Army Medical Center of Hawaii) to validate a suite of deployable digital radiology technologies for image-based cooperative medical diagnostics and treatment planning. NASA, ARPA, and State of Hawaii are jointly providing the support to connect these resource sites through a combination of ACTS-HDR and ground-based high performance communication links.

Application-wise, a core part of such resource sharing is comprised of interactive sharing/transmission of still and video image data between remote sensor locations and physician connected though high-bandwidth network. Despite, the availability of high-performance communication links through ACTS-HDR technology, the efficient utilization available bandwidth is always a key concern. Also the efficiency of medical diagnostics depends significantly on the resolution of images. In this particular paper we briefly present an overview of our scheme for focus based sharing of medical video data over ACTS satellite network.

In section 2, we first provide a brief-overview of the ACTS-HDR technology. Then section 3 presents the novel focus based transmission scheme which provides a flexible technology to obtain a dynamic balance between the bandwidth capacity and quality requirement for video based remote diagnostics.

2. ACTS-HDR TECHNOLOGY

Satellite communication offers a number of unique advantages which makes it a crucial network medium for life critical medical applications. Although satellite communication suffers from weaknesses such as as transmission latency, medium stability, and limited bandwidth capacity, it has several unique advantages over wire/fiber-based communication. Most important of all, satellites can provide on demand connectivity to any part of the world no matter how remote and to highly mobile sites (such as island nations, ships, or troops) no matter what surrounding, so long as a ground terminal can be made available. In some cases, satellite linkage may be the only means of providing telecommunication services. Since medical services may be required any where and any time, satellite communication can be a critical link to these remote /disadvantaged areas and mobile sites.

Recently satellite communication has matured from the initial stage of bent pipe operations through the stage of multiple beams and switch routing to a new stage, where scan beams can be synchronized with dynamic on-board routing. ACTS is pioneering the key technologies required for the ground control of on-orbit switching that can dramatically enhance the capabilities of communication satellites. Specifically, ACTS embodies essential advanced communication technologies to rival ground fiber networking. Some of its enhancements include:

high gain, hopping, spot beams - A rapidly reconfigurable pattern of narrow hopping and fixed spot beams. The high gain antenna used to produce the spot beams enables communications with small, low cost earth stations;

on-orbit switching - A high speed, digital processor on-board the satellite which routes individual circuits to provide single hop mesh interconnectivity. This single hop interconnectivity is in contrast to today’s VSAT systems, which require a double hop employing a central hub ground station. It also reduces the delay.

Published in the Proceeding of the International Conference on Computer Assisted Radiology, CAR'96, June 1996, pp579-584.

Ka-band, wide bandwidth transponders - Use of the Ka-band opens up a new portion of the radio frequency spectrum. It uses 2.5 GHz bandwidth allocation for Ka-band communication.

HDR (High Data Rate) is an earth station jointly developed by NASA and ARPA to exploit the wide bandwidth (900 MHz) capabilities of ACTS. It is fitted with a 3.5 m antenna. HDR earth station provides multiple OC-3 (155.52 Mb/s) or a single OC-12 (622.08 Mb/s) interconnections for point-to-point or point-to-multipoint operations. These earth stations have full duplex capability using satellite switch TDMA and support the SONET protocols.

ACTS services are designed for seamless interconnection into the terrestrial network The ACTS network has the capability to provide both ISDN and SONET transmissions. The architecture of the network is composed of a transportable HDR like Gigabit Earth Stations (GES), and tried-and-true T1/VSATs, with fiber-optic SONET interfaces. These stations communicate directly over satellite using the antenna beams and the on-board uplink-to-downlink beam switching capabilities of ACTS. The network control and management functions are distributed in the various GESs. The operator’s interface is centralized in a Network Management Terminal (NMT). The NMT can be located at any GES site or, even at any location with terrestrial Internet connectivity to a GES designated as a reference station. The network control and management functions are done with standard SNMP protocol. Authorized operators can access control and management functions from console interfaces local to the GES or even remotely through the Internet and dial-up modems. The end user services of this network are configurable point-to-point or point-to-multipoint 155 Mb/s and 622 Mb/s SONET “links” over satellite.

Transmissions to the satellite are performed using Satellite-Switched Time Division Multiple Access (SS-TDMA) techniques. The on-board space-time-space switching is performed by the High-Data-Rate section of the ACTS Multibeam Communications Package. Up to three uplink and three downlink antenna beams can be active simultaneously, that combined with 696 Mb/s burst rates per antenna beam, results in an aggregate system bit rate in excess of 2 Gb/s. Forward Error Correction (FEC) and overhead functions use approximately 10% of the total system bandwidth, resulting in end-user aggregate throughput in excess of 1.8 Gb/s (3 x OC-12). In summary, ACTS technology offers the connectivity as well as realistic bandwidth to generate on demand vital medical communication links across geographical distance and barriers.

3. TRANSMISSION SCHEME

Conceptually a transmission scheme involves the (i) source, (ii) the channel, and (iii) the sink. The principal emphasis of current research in information compression emphasizes extensive analysis of source and channel characteristics. The source of “compression” for most of today’s schemes are in the redundancy that exists in the content. Various compression schemes depend on the extraction of one or more forms of redundant components of a given source information. Also, a significant amount of research has been performed in analyzing the matching channel characteristics with the transmitted information. Extensive research in these areas over last two decades has dramatically improved the state of the art in information coding and transmission. However, as more ambitious applications have evolved, the demand for quality and the competition for bandwidth has far outgrown the advancement.

In this research, we now look into the very characteristics of the sink of information [Wong92]. Human eye probably is not physically capable of processing all the information that a regular video screen wants to pump into it. It is very likely that as a physical device we selectively process only a fraction of the information what it provided to our eye in terms of digital estimate. In this research, our objective is to develop a synchronized information transmission scheme, which tries to co-ordinate the entire image information transmission process with the perceptual focusing process.

In this particular scheme we demonstrate a video encoding scheme, which given a "focal-center" and a "focal-spread" encoded a spatial image region with variable resolution. The region at the "focus" is encoded with full resolution. The resolution of region decreases (with increased compression factor) with inverse proportions of their distance from the "focal-center". Both the "focal-center" and the "focal-spread" can be dynamically varied.

This scheme is very different from most of current encoding schemes. Current encoding schemes (such as DCT, JPEG, VQ) does not make distinction between the varied perceptual importance among the various regions in the scene. The acceptable encoding resolution of the entire scene is therefore dictated by the reproduction quality of the most sensitive region (i.e., it always has to depend on the worst case). The proposed technique allows the encoding resolutions to adapt spatially. The higher compressibility is achieved by such local adjustments [Wong92, Mall89].

3.1 Encoding/Decoding:

In the proposed scheme, the encoder and decoder is connected by two way links. The *control link* from decoder to encoder is used to specify (i) the quantization window, (ii) the focus window location, and (iii) the size of the focus window. The *data link* from the encoder to decoder carries the compressed and encoded image data.

We here describe our first experimental scheme which is based on MPEG-2 (ISO/IEC 13818) standard. The compression stream carries 3 types of frames (i) Intra Frame "I" (ii) Predictive Frame "P" and (iii) Bi-directional Interpolated Frame "B". I frames are coded without using any history and using still image compression technique. Each I frame is divided into 8x8 blocks. DCT's are then computed for each block. The DCT coefficients are then "quantized". This quantization means each of the DCT coefficients are divided by a fixed constant so that bits are dropped off the bottom end of the coefficients resulting in many zeros. Generally a predefined matrix of constants are used to quantize the 64 coefficients. The constants appropriate for quantization of each coefficient are determined from perceptual tuning. However, this matrix can be scaled by a factor Q for each 16x16 macro-block. A large Q makes the constants larger, and results in more zeros in the coefficients, resulting in low bit rate. On the other hand, a small Q retains more bits resulting in higher resolution encoding. This scale constant Q can be varied for each 16x16 macro-block, and plays the key role in our encoding scheme in varying the spatial resolution.

In P frames, instead of coding each of the blocks directly, for each 16x16 macro-block in the current frame, a close match is searched in the past frame. If a close match is found then instead of the actual block data, the difference between this block and a close match is DCT encoded and quantized as before. B frames are used when the current block, instead of being close to a past frame block, is close to a future I or P frame block, or is close to the average of a past and future I and P frame block. In our scheme we use a low Q value for window of focus and high Q value for background region on the image. The macro block Q values are automatically determined using the following Q distribution model followed by the encoder and decoder.

3.2 Conserving Distribution Model (CDM):

This model assumes that the total *Q* over the image space is conserved. Let *a* is the *Q* factor for a high resolution, and *b* is the *Q* factor for a region of low resolution, and *C* is the total accuracy distribution over all the macro-blocks in the frame, and *n* is the fraction of macro blocks in the region of high focus:

$$C = a.n.P_{base} + b.(1-n).P_{base}$$

$$P_{base} = \frac{C}{a.n + b(1-n)}$$

$$Q_{window} = f(a.P_{base}), \quad Q_{backgnd} = f(b.P_{base})$$

The values of *a*, *b*, and *C* remain constant over a transmission session. What varies is the size of the focus window, and locations. These values are transmitted over the control-link. However, since, both encoder and decoder follow the CDM, the dynamic variations of the *Q* values are not required to be transmitted with each macroblock. This reduces the usual MPEG overhead of *Q* variation. The function *f()* depends on the actual effect of *Q* on the bit rate or resolution.

Depending on the type of frame, each macro block code contains quantized DCT coefficients of the actual data, or difference from a past, future or average block, the motion vectors, and quantization parameters. This macro block code is then further Huffman coded using fixed tables and run-length encoded.

4. CONCLUSIONS

We have recently implemented a transmission scheme based on this general scheme, and are in the process of characterizing its performance. A typical high resolution 1024x1024 frame with 24 bit/pixel medical video stream in uncompressed format requires about 1 Gb/s channel capacity. High resolution encoding schemes (with +40db accuracy) can offer a compression ratio in the order of 10. While, low resolution schemes (with +20db accuracy) can offer a compression ratio in the order of 100. Our objective is to maintain near 100 times compression and still offer diagnostically acceptable quality image with prudent management of focus information, so that a bidirectional voice, text, and medical-video (control and data link) stream can be accommodated within 10-20 Mb/s link.

The other principal research challenge in this initiative is the adaptive determination of the perceptual significance of the regions. The resolution distribution surface is computed from the (i) anatomical and optical geometry of eye and ratina (ii) the direction of glance and distance of ratina from the screen. The first part is relatively static (at least for a particular viewer), and the second part of the information is traced from a dynamic eye-glance tracer. In addition to the spatio-geometric characteristics of perceptual focus also is important is the temporal properties of eye movement. Such as the speed to tracking, the fixation time, perceptual response to loop-back time.

The overall success of the scheme is critically dependent on the real-time synchronous adaptability of the encoder and the decoder to the dynamic parameters, which are characteristically perceptual in nature. This imposes some very specific constraint on the encoding/decoding scheme.

REFERENCES

[Mall89] Mallat, S., "A theory of Multiresolution Signal Decomposition: The Wavelet Representation", IEEE Trans. on Pattern Analysis and Machine Intelligence", Vol. 11, no 7, July, 1989.

[Wong92] Wong, P. W., "A Multiscale Image Coder", SPIE proc. on Image Processing Algorithms and Techniques, v.1657, 1992, pp46-57.

[YuGa96] D. Y. Y. Yun, H.M. Chen-Garcia"Sharing Computational Resources and Medical Images Via ACTS-Linked Networks", *Proceedings of Pacific Conference on Telecommunications, PTC'96*, February 1996, pp752-759.