Wireless Progressive/Scalable Multimedia Communications: Current Trends & Challenges

Suayb S. Arslan[1] *Student Member, IEEE*

IEEE Potentials Surveys 2009

AS the next century of technology goes wireless, the multimedia transmissions over wireless links attracts a tremendous amount of interest from the research community. The ongoing research is usually focused on efficient transmission schemes, optimal system designs and their hardware implementations. Multimedia transmissions necessitate demanding communication systems especially after the ubiquitous use of High Definition(HD) multimedia content. In addition, HD multimedia transfer is one of the promising services of 3G and beyond technologies in the market. However, the transmission of multimedia data over band limited wireless networks is a challenging task in which one needs to maintain a decent quality of multimedia at all times.

In traditional multimedia source encoders, the user end is either able to reconstruct the source at full quality or receives no reliable information to reconstruct the source. Given the random nature of wireless channels, this property may lead to excessive quality fluctuations and thereby user dissatisfaction. In later sections of this article, we will see that *Progressive* image and *scalable* video encoding techniques are introduced to alleviate the undesirable effects of wireless links and allow graceful degradations of the source at the expense of little coding inefficiency.

Progressivity and *Scalability* are the terms used for image and video coding mechanisms when the source encoders allows decoding at various source rates. Both terms refer to the mechanism of the source encoder-decoder pair to provide progressive reconstruction of the source. In other words, as more and more bits are reliably received, they are used to refine decoded images or videos at any decoding time instant. Although output bitstream of such encoders has desirable properties such as progressivity, any error due to the channel may render the whole bitstream useless. Therefore, the transmission problem does not include just the wireless channel impairments and bandwidth limitations but also due to the progressive encoding nature of the source, any error in the bit stream leads to an unrecoverable error propagation throughout the bit stream. Different protection mechanisms are proposed in the literature to overcome the challenging problems of current multimedia transmissions.

This article starts off by giving basic information about some of the well known compression techniques (but not in detail), progressive source encoders and wireless channels. Then, it introduces progressive multimedia transmission over wireless links and mention design tools, challenges and methods up to date to overcome the present problems.

[1]Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093, E-mail:sarslan@ucsd.edu

Why do we need compression?

Before we begin, let us take a look at the figures in Table I. It shows uncompressed information of various different types of data and necessary communication resources: disk space, raw data rate, and transmission time using a $28.8Kb/sec$ modem.

TABLE I

MULTIMEDIA DATA TYPES AND UNCOMPRESSED STORAGE SPACE, TRANSMISSION BANDWIDTH, AND TRANSMISSION TIME REQUIRED.

Multimedia type	Size/Duration	Bits/Pixel-Sample	Uncomp. Size	Raw data rate	Trans. Time (using 28.8K modem)
A page of text	11" x 8.5"	Varying	4-8 KB	32-64 Kb/page	1.1 - 2.2 sec
Telephone speech	10 sec	8 bps	80 KB	64 Kb/sec	22.2 sec
Grayscale Image	512 x 512	8 bpp	262 KB	2.1 Mb/image	1 min 13 sec
Medical Image	2048 x 1680	12 bpp	5.16 MB	41.3 Mb/image	23 min 54 sec
Full-motion Video	640 x 480, (30frames/sec,1min)	24 bpp	1.66 GB	221 Mb/sec	5 days 8 hrs

KB:Kilo Bytes, Source: http://www.acm.org/crossroads/xrds6-3/sahaimgcoding.html

It is easy to see the need for large storage spaces and huge bandwidth along with long transmission and waiting times. Communication channels, wireless channels in particular, are in scarce of those communication resources. Every resource is considered very valuable and one resource can be more important then the other depending on the nature of the application. Therefore the compression of data by exploiting the redundancies at the transmitter is the clear cut solution. Reducing the amount of information by multiple factors relaxes the conditions on the natural constraints which results in efficient utilization of these communication resources.

MULTIMEDIA COMPRESSION

Multimedia compression can be classified into two major encoding strategies: Lossy and Lossless. The details of these methods are beyond the scope of this article. However, it is apparent that lossless compression compresses the multimedia without any loss so that the decoder can reconstruct the source perfectly. Lossless compression on the other hand does not provide much of a compression to be used in our multimedia communication scenario. Lossless compression is usually used for data compression (ex: text, executables) where any loss of data is intolerable. We need to note that progressive encoding is classified under lossy compression.

IEEE Potentials Surveys 2009

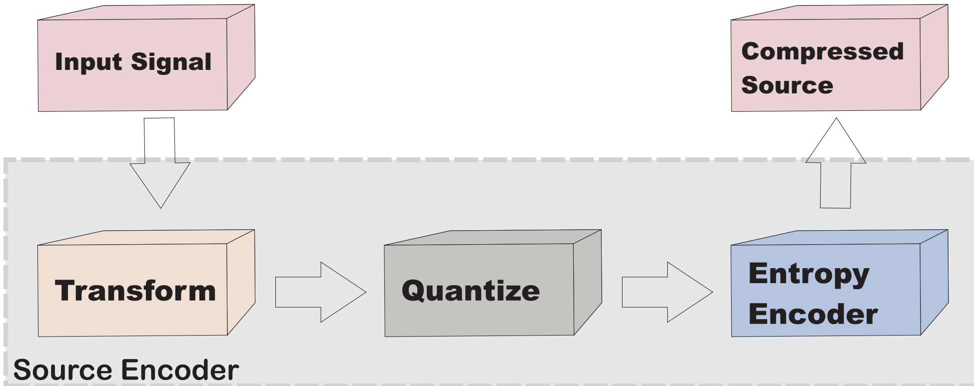

Fig. 1. General source encoder block diagram.

A practical source encoding process consists of a transform, a quantization and an entropy block. General source encoder block diagram is roughly shown in Fig. 1. One of the main building blocks of a generic multimedia encoder is the transform operation. The transform block decorrelates the multimedia information and achieve energy compaction to eliminate the redundancies of the multimedia source. In other words, the energy is randomly distributed among the source samples and there is no any ipriori information about the underlying distribution. The main objective of the transform block is therefore to assign more energy to fewer transform coefficients. The energy is usually concentrated on a few transform coefficients so that they can fit the quantization better (minimization of loss due to quantization) which is the next step before entropy coding. In lossy encoded image or video, the main source of loss comes from the quantization block in the encoder. quantization maps the real valued coefficients to predetermined fixed real values. This mapping is not one-to-one and thus the unrecoverable error results. Vector quantizers are especially designed to optimize the quantization process based on the nature of the source such that the loss due to quantization is minimized. Those quantizers can be optimal but they may become very complex implementation wise. Instead, practical image and video encoders use simpler scalar quantizers with different linear transformation techniques. Lastly, entropy coding block is used to efficiently store the quantized values as a stream of bits. In that, each real valued coefficient is mapped to binary bits without loss. Then based on the bit patterns redundancy among the bits are exploited to reduce the amount of bits to be transmitted. This entropy compression can further introduce loss in the decoding process.

There are basically two very commonly used transform techniques in literature to encode still images and video sequences : *Discrete Cosine Transform (DCT)* and *Wavelet Transform (WT)*. DCT is real-valued technique for converting a signal into elementary frequency components, thus packing energy into fewer coefficients. Advantages of DCT are its practical simplicity, satisfactory performance, and availability of special purpose hardware for implementation. Wavelets, are on the other hand, finite length time domain functions having zero mean and special shape. Mother wavelet is the original function from which other wavelets are derived. The main goal is to represent any time domain function as the sum of scaled and shifted version of

IEEE Potentials Surveys 2009

the mother wavelet. Wavelet transform is used extensively in recent applications of multimedia technologies that have greater compression performance compared to other compression methods that use different transform techniques. The advantage of wavelets is that the window size vary such that it gives better resolution in both frequency and time. This is achieved by having infinite number of different basis functions contrary to sine and cosine functions of DCT. More information about Wavelets can be found in "*Wavelet Basics*" by Chan.

DCT and wavelet transform based codecs and progressive encoding of multimedia

Prior to the discussion of *progressive/scalable* source encoding, let us take a look at some of the basic non-scalable encoders. There are various types of multimedia codecs in the literature and only some of the popular ones will be considered in this introductory article.

Joint Photographic Expert Group (JPEG) lossy image compression, is one of the very basic block-based (it divides the image into 8×8 blocks) transform image encoders. Although the original standard defines JPEG in sequential, progressive and hierarchical mode of operation, the baseline JPEG is the sequential type. JPEG uses DCT as the transform technique on each 8×8 block and quantizes the transform coefficients before what is called zig-zag ordering of coefficients and entropy coding (Huffman encoding). This particular type of ordering helps the entropy coding compress the data by placing low frequency coefficients at the beginning of each zig-zag ordered coefficient sequence. JPEG is a very well known image compression technique and the detail of its operation can be found in any standard image processing book.

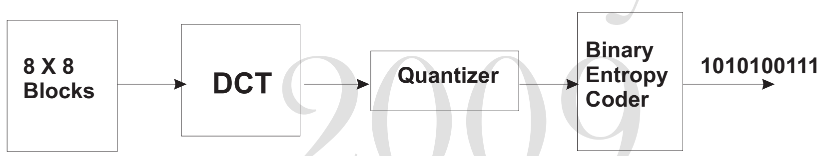

Fig. 2. JPEG compression block diagram

Progressive JPEG on the other hand, divides the file into a series of scans instead of one. The first scan enables a very coarse description of the image at a low quality setting. Following scans gradually improve the image quality. Each scan adds extra information to the already provided data, so that the equivalent quality of the last scan is the same as the baseline JPEG image. The advantage of progressive JPEG is that one can observe the whole image at early stages of the decoding but at a low quality setting whereas the baseline system display the image by decoding the blocks from top to bottom, the whole image can be seen after the whole decoding process ends.

IEEE Potentials Surveys 2009

Baseline JPEG's output contains information of equal significance. In other words, no part of the total bit stream bears more important information than other parts. In case of a loss or an error, because of its block based nature, JPEG suffers the blocking artifacts (it is the blocky look, loss of edge clarity and fuzziness at the boundary of each block. Blockiness is more common at busy regions of the image i.e. where the high frequency of the texture is concentrated) especially visible at low compression ratios. This is because blocks are assumed to be independent at the time of transformation. Since the codec does not eliminate the correlation among the blocks, blocking artifacts are inevitable.

Most modern image coding algorithms use WT in their transform block. For example, Embedded Zero-tree Wavelet (EZW) and Set Partitioning In Hierarchical Trees (SPIHT) algorithms are WT based algorithms. One of the advantage of WT over DCT is that WT based coders do not suffer from blocking artifacts. Yet, those coders can suffer from ringing artifacts. EZW or SPHIT will not suffer from block based artifacts because they do not separate the source into blocks, they rather treat the source as a whole and construct the spatial trees of wavelet coefficients in order to eliminate redundancy both in frequency and spatial domain. However, at lower source rates, high frequency WT coefficients are effected more and that creates ringing artifacts near the edges. Secondly, WT based image coders naturally allows for hierarchical and progressive images in different resolutions. The final output is also progressive to better match the requirements of various applications. In other words, they encode the bit stream naturally in a progressive fashion so that the bits that appear first in the total encoded bit stream are more important than the bits that appear later. An example of progressively encoded image is shown in Fig. 3. In general, wavelet based encoding schemes make use of the dependencies between tree structures of wavelet coefficients aligned in a parent-childhood relationship instead of dividing the source into blocks and removing the redundancies. More information can be found in **CIPR at RPI web site** (Details of these algorithms are beyond the scope of this article. EZW and SPIHT are very well known image encoders. Through some web search, a decent amount of information can easily be obtained). Finally, WT leads to better energy compaction properties and thereby better compression potential for multimedia.

Scalability is usually used within video encoding terminology (Scalable Video Encoding (SVC)) and refers to the rate scalability of the video encoders. However, it is also used in image compression terminology as well. JPEG2000 is one of the examples where we see different types of scalability mentioned in the standard. Among them are spatial, SNR and temporal scalability. Rate scalable multimedia encoding provides very useful properties such as the ability of decoding at any given bit rate. Especially in wireless packet networks, the rate changes during the transmission. Thus, scalable source encoders are efficient choice for the graceful degradation of the source at the receiver. Spatial scalability (also known as resolution scalable) allows the decoding at different resolutions so that depending on the channel conditions multiple resolutions of the same video source can be obtained. Therefore, for example, users with good channel conditions will be able to get higher resolution video

IEEE Potentials Surveys 2009

whereas the users with worse channel conditions will be able to playback only the baseline video.

Progressivity in video coding has a different meaning. It usually refers to the type of scanning the pixels of the frames of the video sequence. Progressive video is usually the one that has non-interlaced scan displaying of the frames. Yet, *Progressivity* term in image coding finds a correspondence in scalable video coding.

Video encoders such as MPEG-2 divides the whole source block into layers of different significance. The idea behind layered encoding is to divide the source into different priority groups so that the significance of each layer will be unequal. Basically, important information such as header or motion vectors (Motion vectors contains information about the movements of any point in the current frame with respect to the reference frame. So that we do not need to send the each successive image in a given video stream.) are called the *Base Layer*(BL) information and similarly the *Enhancement Layer*(EL) is said to be less important information such as the residual texture data in the successive frames of a video sequence. On the other hand, as found in Fine Granular Scalable (FGS) MPEG-4 standard, EL information is bit-wise scalable in nature to achieve smooth rate-distortion characteristics compared to older standards. In other words, the receiver can truncate EL bit stream at any place to reconstruct the video. However, the BL information is in no way progressive or scalable and usually used to construct the video up to some acceptable quality.

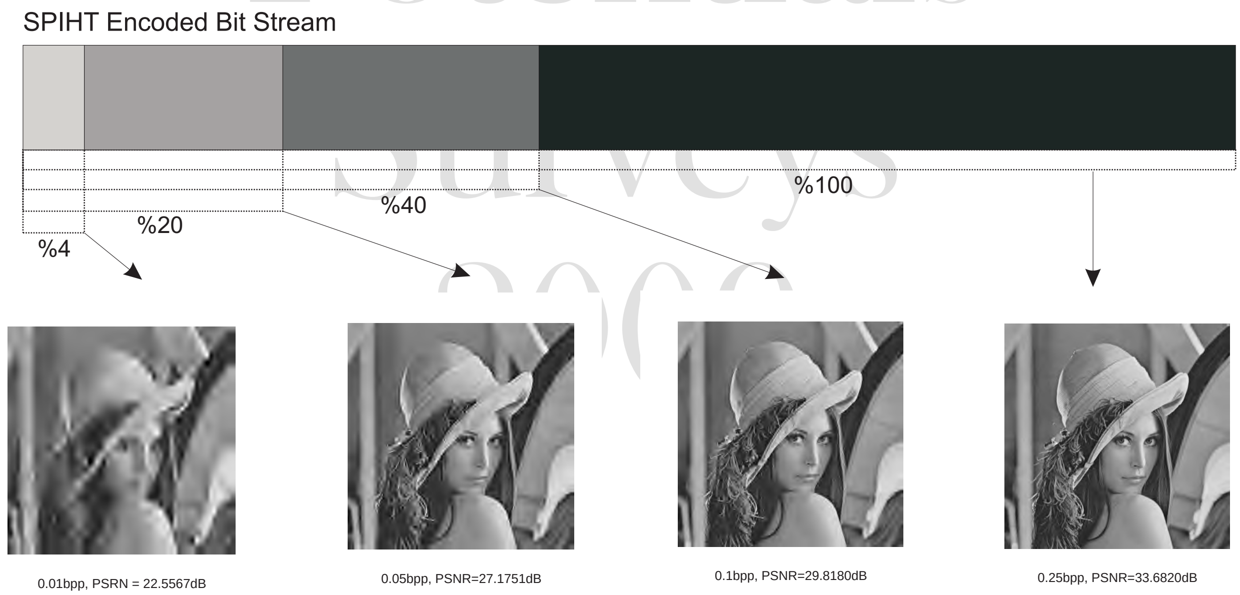

Fig. 3. Image Lena. Progressive Image Encoding using SPIHT without arithmetic encoding. Compression rate is usually expressed in bits per pixel(bpp). SPIHT is one of the best progressive image encoding technology giving very good quality compressed images.

Bit stream of both approaches is even sensitive to a single bit error in case of which may result in catastrophic outcomes. In JPEG, every transformed 8×8 block has 64 coefficients. The first coefficient is usually called DC coefficient representing

IEEE Potentials Surveys

the average value of pixels in the block. For block based compression techniques such as JPEG, differential encoding for DC coefficients and Huffman encoding of transformed coefficients are the main operations sensitive to error events. The remaining 63 coefficients are named AC coefficients. Similarly, an error in Huffman encoded AC coefficients can lead to loss of synchronization among the 8×8 blocks of the image. For EZW or SPIHT, the error propagation is because of the way zero-trees quantize the wavelet coefficients and produce the encoded bitstream. Output bit stream bears not only the information of texture or image color but also the orientation and position of these zero trees in the corresponding subband. Therefore if an error occurs somewhere in the bitstream which includes information about, say, the orientation, the resultant bit stream can be misinterpreted by the decoder. This will yield loss of synchronization and thus the distortion will go up although the decoder could have obtained correct bits after the first error.

Robust Approaches: Robust approaches are introduced in order to somewhat eliminate the error propagation for both DCT and WT based image and video encoders. Adaptive Discrete Cosine Transform (ADCT) based JPEG, which is also standardized, is one of the robust image encoders. It is robust because the output packets are fixed size i.e. not variable length coded (Huffman encoded) and error in one block does not effect the other. Also some robust SPIHT algorithms exist in the literature. They use what is called as wavelet coefficient partitioning. In this way separate trees are encoded individually and hence spatial dependence is removed to provide robustness. However, additional robustness will eventually cause loss in rate-distortion performance of the encoder. In other words, the new compressed bit stream can only perform worse than the original algorithm but it provides some robustness against errors.

The scalable design of image and video encoders are not arbitrarily chosen as described. The main idea behind it is to allow sequential data reception over unreliable transmission environments. In other words, the terminal in a network with good channel conditions will be lucky enough to get all the information reliably, on the other hand one other terminal with bad channel conditions will hopefully get only BL information reliably and thus an acceptable multimedia quality is guaranteed at the receiving end. Scalable encoding paradigm is a very useful tool for distributed networks in particular, utilizing web browsing applications and any wireless device deemed to communicate video or image sources over wireless channels. For example, imagine that you are surfing the internet on your laptop, for example a web site that provides HD quality images. HD images are usually large multimedia files and comes in high resolutions. At first, the user will have the coarse description of the source at early stages of transmission. As soon as he realizes that it is not what he is looking for, he can terminate the transfer through a simple click. By doing so, he would ultimately yield less congestion and improved efficiency in the network.

All in all, scalable multimedia files have high rate demanding, loss-tolerant, and mostly delay and error-sensitive nature. Unlike randomly generated computer data, the nature of the encoded bit stream is an important parameter in designing the over all communication system. Another equally important design parameter is the nature of the channel which poses significant

IEEE Potentials Surveys 2009

challenges in the design process by introducing random and often correlated errors into the error-prone multimedia bitstreams.

WIRELESS EFFECTS, MOBILE COMMUNICATIONS

In a scalable multimedia transmission scenario, the major effect on the performance is the noisy channel. In fact, the nature of mobile environment poses a different set of difficulties and challenges on the performance of any wireless communication system, yet its effect on a progressive/scalable source can be detrimental because of the error propagations mentioned in the previous section. Performance degradation due to the channel can be attributed to inherent shadowing, fading or interference of typical wireless environments. Unlike wired channels, wireless transmission medium inserts predominantly random effects. Depending on how rapid the channel varies or how long it has been in a deep fade or not, it could be a major source of severe distortion in the signal waveform. Based on the frequency and time variation characteristics, there has been some formulation of the radio channel in order to analyze such wireless behavior.

IEEE Potentials Surveys 2009

- **Time Dispersion Effects**

  Reflections from other objects cause multiple replicas of the transmitted signal at the destination. If objects are placed far apart, then the time for the reflected signal to be received will be large enough so that it can interfere with other transmitted symbols causing what is called Intersymbol Interference(ISI). In technical terms, if bandwidth(BW) of the transmitted signal is less than the BW of the channel (also known as *coherence bandwidth*), it is usually referred as a flat fading channel. Otherwise it is called frequency selective fading channel and the output is distorted because different frequency components undergo different attenuation and phase characteristics.

- **Frequency Dispersion Effects**

  On the other hand, If the channel happens to vary rapidly, i.e. faster than the rate of communication, then the same symbol will experience different fades (random amplitude changes) and phase shifts. The rapid changes in the amplitude and phase gains will distort the signal and widen the frequency response of the original signal. This is usually called Doppler spread and the frequency of change is inversely related the *Coherence time* which is technically defined as the amount of time during which the samples of channel impulse response are highly correlated. If coherence time is less than symbol period it is called fast fading otherwise it is named slow fading.

In many cases, flat slowly varying rayleigh fading channel is assumed. It is not only the simplest wireless channel model but also applicable to some real scenarios. Also, the analysis is tractable. However if the user is mobile and motion becomes dominant and objects are placed at a distance, then this widely used model is no longer applicable. For example, a man talking to his wife while driving in San Diego downtown might be a good example of where we no longer have this model.

Transmission of progressively encoded sources over wireless links

As noted, any error in a progressively encoded bit stream should be truncated unless there is some synchronization methodology adapted. Even a single bit error may lead to an error propagation in the bit stream. This is illustrated in Fig. 4. For a 512×512 monochromatic Lena image, $0.25bpp$ corresponds to 65536 bits. The figure shows two images when a single bit error is in 1500th and 20000th bit position. It is clear that a single bit error leads to decoder to misinterpret the remaining correct bits after the first error. The same figure compares the truncated version of the letter image by labeling them (1) and (2). Thus, disregarding the rest of the bit stream after the first error help improve the decoding performance of the source.

Transmission of multimedia files is usually performed by packetizing the compressed bit stream. In other words, each information chunk is concatenated with a Cyclic Redundancy Check (CRC) code that is instrumental in error detection and truncation process. Once the packet is error-free, it is included in the decoding process. Otherwise, whenever a CRC flags for an error, the decoding process ceases and incomplete decoding of the image is declared. Finally, received packet stream included in the decoding process is used to reconstruct the multimedia source. Either packet size or information block size is assumed to be fixed based on the particularity of the application.

IEEE Potentials

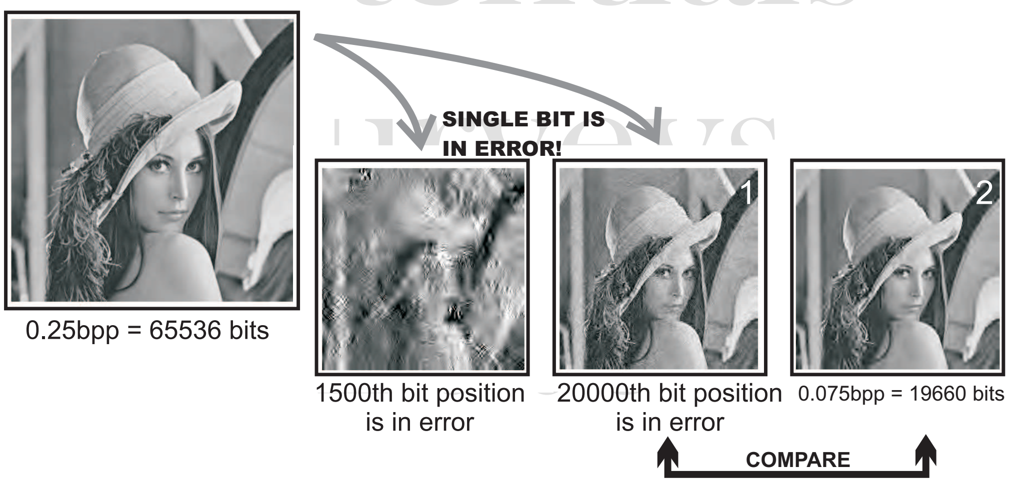

Fig. 4. **In case of an error, truncation process is a must for a progressively encoded multimedia. An Image example is shown where we have only single bit errors in different locations of the bit stream. Compare the two images as shown that the first image has 65536 bits to reconstruct the image with a single error in 20000th bit position. Second picture has only 19660bits without any error. As it can be seen, second picture is a better quality product.**

Transmission over interference-limited wireless channels

There are also interference dominated systems such as Wi-Fi and Wimax based multimedia access. Possible interferers are microwave ovens, cordless phones, other devices that use Wi-Fi or Wimax access and bluetooth headsets in a typical office environment. In the presence of interferers, the throughput of the system usually decreases. Also the closer the interferer, the greatest is the impact on the multimedia reconstruction quality. General interference mitigation techniques can directly be applied to improve the throughput performance. In interference dominated transmission scenarios, reduced capacity of channel is usually used for high priority data transmission. When the channel capacity increases, bits that are used to refine the multimedia source is transmitted to enhance the multimedia reconstruction quality. As the interference is time varying, less important data may be used at various frequencies depending on the location of the interferers. At the receiver, an acceptable video or image quality can still be obtained in a progressive image or scalable video transmission scenario through optimal scheduling of information layers tailored to specific channel conditions at each transmission period. As can be seen, the progressive nature of the bitstream alters the way the system design is handled in an interference dominated transmission scenario.

LATENCY IN PROGRESSIVE/SCALABLE SOURCE TRANSMISSIONS

IEEE Potentials Surveys 2009

Unlike the computer generated data, multimedia files are commonly considered delay sensitive sources. Usually, latency is not a consideration when the progressive source is an image. This is usually because progressive image gradually improves the quality as more and more bits are received. Any delay or packet loss in transmission will not bother the end user as long as the encoder-decoder pair does not cause propagation of error. This is usually achieved by forward error control coding where only the correct packets are used at the decoder. Wrong packets can also be retransmitted through some Automatic Repeat reQuest (ARQ) mechanisms and the end user will be satisfied with the delayed version of the image which is still high quality product. Yet in ARQ, because of some congestion in the network, quality refinement of the image can take more than expected. The user might be interested in the details of some part of the image and therefore has to hold on untill some threshold quality is obtained. If the channel is feedback-free like found in broadcast applications, we have to note that with ever-increasing transmission data rates and improved image compression algorithms in today's technology, download of a high-resolution image takes only seconds and a delay of several milliseconds will not be considered an issue.

However in a video transmission scenario, latency is and will be a serious issue partly because of the dependency structures among the video frames exploited by the video encoder for efficient compression and because of the ever-dynamic random nature of typical wireless channels. In this particular transmission scenario, transmission may lead to unrecovered lost packets. ARQ mechanisms might not be viable solution here because excessive delay may cause dropped frames by the decoder which may hurt the video quality.

A practical video encoder-decoder pair is usually equipped with a finite buffer to store the incoming transmitted bits before the decoding operation and the display. If the decoding rate of the decoder is higher than the transmission rate, then the received video will incur unexpected delays in the video (*buffer underflow*). On the other hand, if the decoding rate of the decoder is lower than the transmission rate, then some of the packets will automatically be dropped even if they are received correctly. This is usually called *buffer overflow*. Unless they are retransmitted, the video quality will have time varying defects which might be very annoying. In addition to varying channel capacity and bit rates, the image complexity in a typical video sequence changes in time, thereby the compression algorithms operates at time varying source rates. Therefore, some rate-control mechanism should be adopted to save the codec from *buffer underflow* and *buffer overflow*. Practical SVCs adopt such a rate-control mechanism to account for unexpected delays for a robust system design.

In scalable video codecs, different priority layers will have different latency tolerances. Therefore, rate-control optimization tools & methods to minimize the delay in video transmission can be applied to each layer in order to meet the end-user video quality and latency requirements. For example, for a particular delay-sensitive application, scalable video base layer will typically incur the least of amount of delay in order to increase the user satisfaction. Allocation of resources based on the delay sensitivity of the multimedia source will also lead to efficient use of scarce communication resources.

IEEE Potentials Surveys 2009

UNEQUAL ERROR PROTECTION (UEP)

The nature of compressed scalable multimedia data has unequal significant information distribution. This unequal sensitivity inspired the idea of unequal protection to protect heavily the more important data more than the less important information.

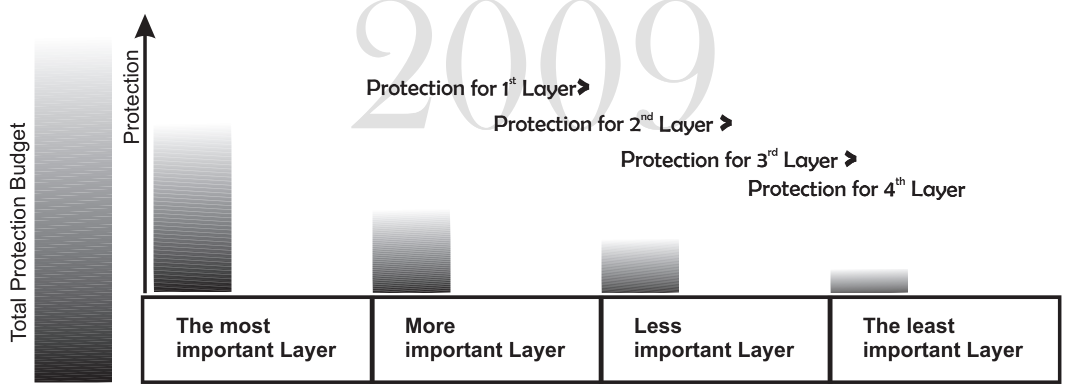

Fig. 5. Unequal Error Protection (UEP) for a multi layer encoded multimedia transmission.

Unequal error protection is provided in many different ways. The main idea is to allocate the communication resources unequally among the different layers of the source. For example, power can be allocated between different kinds of source symbols. This method allows more power to be used to send the critical information to combat against wireless channel impairments. In analogy to Fig. 5, you can think of the protection levels corresponding to different power levels. Thus, more important packets are assigned more power then less important content paving the way for more reliable transfer of the coarse description of the source (Ex: a basic quality image or video). In practical communication systems, for spectrum efficiency, bits are modulated before transmission. Hierarchical Modulation (HM) is a UEP technique that allocates different powers among the bits that make up the modulated symbol. In other words, UEP is accomplished by playing with the relative distances of symbols in a given signal constellation so that some of the bits gain more protection. In this method, symbols are assigned a binary representation (ex: gray coding is one of them) in such a way that for example the first bit (being the most important bit) falls far apart from the decision boundary and has the lowest probability of bit error (BER). Similarly, the hierarchical parameter adjusts the distances of other symbol points so that we have high and low priority bits assigned different BERs. A simple hierarchical constellation with gray coding is depicted in Fig. 6. As seen, the first bit is the high priority (HP) bit and the y-axis is the decision region for that bit. As we increase the hierarchical parameter α, two symbols in the right hand side of the constellation fall apart decreasing the BER of the HP bit and increasing the BER of the low priority (LP) bit. Exact reverse effect shall be observed on HP and LP BERs should we decrease α.

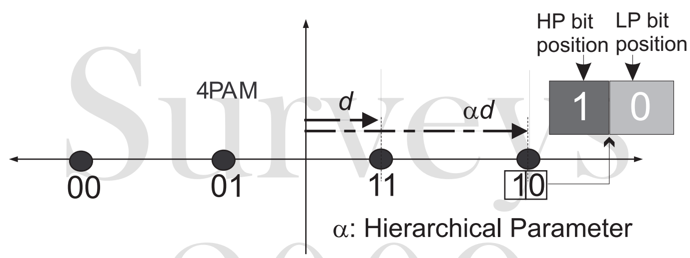

Fig. 6. Unequal Error Protection (UEP) using Hierarchical Modulation. Ex: Hierarchical 4PAM.

IEEE Potentials Surveys 2009

In multi antenna systems, space time block code (STBC) is an extremely powerful tool and provides spatial diversity increasing the possibility of more reliable transmission compared to point-to-point communication. In contrast, spatial multiplexing techniques are used to increase communication data rates at the expense of reduced diversity and BER performance. The details of these techniques are beyond the scope of this article, yet their hybrid use in multi antenna systems provides UEP and will be briefly mentioned here. This technique is usually found application in multi input multi output (MIMO) channels where an array of antennas are available at the transmitter and receiver. Base layer information are typically sent through STBCs to benefit from spatial diversity, and hence ensured reliable transfer of the core layer of bits whereas less important layers (low priority stuff) are sent using various spatial multiplexing techniques to increase the rate of communication. This way a decent multimedia source reconstruction can be obtained at the receiver with graceful degradation as the channel gradually gets worse. More information can be found in “Video transmission over MIMO-OFDM system" by Zheng *at al*.

Last but perhaps the most frequently used method is called joint source channel coding(JSCC) where more important bits are protected using stronger channel codes where as the refinement information is barely protected using weaker channel codes. JSCC is usually most effective when no channel feedback or ARQ is permissible.

Fig. 7. Carphone sequence transmissted over a noisy channel. The effect of dropped frames due to buffer overflow and the decoder concealment of transmission errors are shown in pictures in the middle and on the right. Source encoder is scalable extension of H.264 which is based on block based DCT. I: Intra frame, B:Bipredicted frame (a frame predicted using two other frames in the sequence), UEP: Unequal error protection, EEP: Equal Error protection.

All these approaches need some optimization to achieve the best reconstruction quality in order to minimize some performance metric such as distortion. Using UEP techniques, it is already known that better image and video quality can be obtained even in very mild channel conditions compared to equal error protection (EEP) case. An example for commonly used *Stefan* sequence is shown in Fig. 7. As seen EEP has blocky artifacts and low quality content. Unequally protected data enables significant information to be received more reliably compared to refinement information. Hence the end user will have an approximate description of the source at early stages of transmission before even the buffer is full. Especially in case of severe channel degradations, source reconstruction will still be possible and the user can differentiate what kind of image or video is transferred. However, especially in video transmissions, once critical information (BL information) is not reliably received, then the quality degradation can be unforeseeable.

Joint Source Channel Coding(JSCC)

When the retransmissions are prohibitively time consuming leading to increased network inefficiency and there are excessive multimedia playback delays at the receiver, JSCC is the most commonly used UEP mechanism. Even if the application is

Potentials Surveys 2009

moderately tolerable to latency, JSCC can still be used with popular ARQ designs to create hybrid mechanism to improve the system performance.

In coding community, *Source coding* refers to the compression by exploiting the redundancy within the source samples whereas *Channel coding* works by adding redundancy for reliable transmission subject to some bandwidth constraint. To deal with wireless impairments, there have been studies in search of optimal source and channel coding schemes for a variety of applications. According to Shannon's theorems where block sizes and encoder/decoder characterizations were not realistic (Ex: infinite block size and complexity), separate optimization of source and channel coding blocks will optimize the overall system. However, real systems with limited processing power and finite block sizes undergoing time varying channels should be optimized by considering source and channel coding problem jointly. This idea is very well known and called as Joint Source and Channel Coding(JSCC) problem.

In JSCC, the main objective is the optimal rate allocation between source and channel codings such that some performance criteria is optimized. Performance criterion is usually application specific: it can be either the overall multimedia distortion& quality or useful source rate. Total transmission is usually constrained to use limited communication resources such as bit budget or bandwidth. For a given bit budget, allowing less bits to represent the source will decrease the quality of the reconstruction process due to quantization errors. On the other hand few bits to protect the bit stream will cause channel errors to dominate and therefore the signal is will be destroyed and become useless. Thus, this demonstrates that some compromise is needed between these two extremes. In Fig. 8, parity and information bits within each packet is chosen depending on the significance of information content. Through some optimization, best allocation of source and channel coding bits can be found. Often times, channel code rate set is discrete i.e., there are only finite number of code rates in the set. When the optimal code rate turns out to be some code rate that is not in the set, closest available code rate is used to protect the bit stream.

IEEE Potentials Surveys 2009

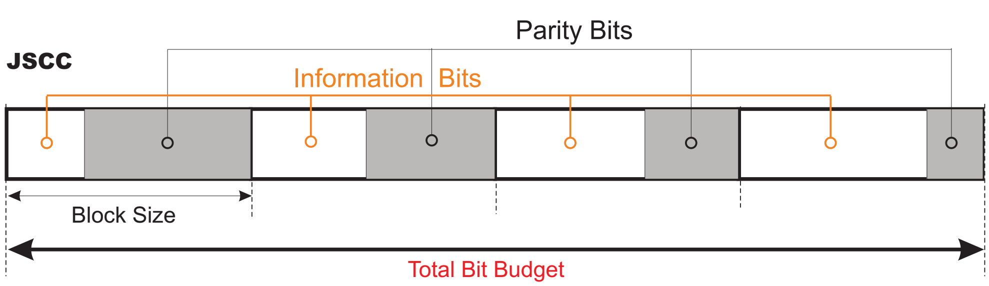

Fig. 8. Unequal Error Protection(UEP) provided with JSCC for different layers of a coded multimedia source

As observed in Fig. 8, more important layers or packets are assigned more powerful codes to ensure their reliable transfer.

Also, total bit budget is determined by the band width limitations of the system and the modulation format being used.

Further Challenges

The number of solutions for the efficient transmission of scalable multimedia over wireless channels is ever increasing and attracts more attention than it did in the past. However, most of the system models currently in consideration assume ideal cases such as perfect synchronization or perfect channel state estimation. In fact, it is not hard to verify that all the optimization criteria mentioned in this paper are based on the accuracy of these estimator outputs. In case of imperfection of any subset of these parameters, performance improvements can be dramatically lessened and such models might not even be efficient solutions anymore. In addition, when the transmitter is required to have current channel state information, then a feedback mechanism from the receiver will become mandatory. This is actually not always possible for every system design especially for delay-sensitive multimedia applications.

IEEE Potentials Surveys 2009

Conclusion

Although there are number of constraints of UEP designs developed for scalable multimedia transmissions over wireless channels, emerging novel techniques secure more efficient solutions to meet our future multimedia transmission needs. As the quality of multimedia rises up with increasing user demand, video and image communications over wireless channels is considered a lingering challenge for multimedia communications technology today. Better compression and transmission schemes are needed to enable a robust HD multimedia transfer respecting the communications resources. The fact that the source encoding output is becoming extremely sensitive to errors make this problem even more interesting. However, great advances are being made in both transmission and channel encoding structures to overcome those problems. As the current circuit technology gets even faster in nano scale, new complex designs have the potential for near future implementations enabling internet users to browse, download and play in real time the perfect HD quality multimedia sources easily on their hand held devices or desktop computers.

Read More About it

- M. Mrak, N. Sprljan, E. Izquierdo "An Overview of Basic Scalable Video Encoding," *46th International Symposium Electronics in Marine,* ELMAR-2004, Zadar, Croatia.
- Center for Image Processing Research (CIPR) at the Rensselaer Polytechnic Institute web page. Available online: http://www.cipr.rpi.edu/research/SPIHT/
- H. Zheng, C. Ru, C. Chen and L. Yu, "Video transmission over MIMO-OFDM system: Multiple Description Coding (MDC) and space-time coding-based approaches", *Hindawi Publishing Corporation Advances in Multimedia*, Vol. 2007.

- Chan, Y. T. Wavelet Basics, Kluwer Academic Publishers, Norwell, MA, 1995.
- A. Ortega, K. Ramchandran, “Rate-Distortion Methods for Image and Video Compression: An Overview" *IEEE Signal Processing Magazine*, pp. 23-50, Nov. 1998.
- Y.-Kheong Chee, “Survey of progressive image transmission methods," *International Journal of Imaging Systems and Technology*, Wiley, Vol 10, Issue 1, pp. 3-19, 1999.
- Leger, A. Mitchell, J.L. Yamazaki, Y. “Still picture compression algorithms evaluated for international standardisation" *IEEEGlobal Telecommunications Conference and Exhibition.* vol.2 pp.1028 - 1032 Dec. 1988
- Bell-Labs Innovations. Multimedia Communications Research Lab. Available online: http://www.bell-labs.com/org/1133/
- Stuber, G.L., Barry, J.R., McLaughlin, S.W., Ye Li, Ingram, M.A., Pratt, T.G. “Broadband MIMO-OFDM wireless communications" *Proceedings of the IEEE* Vol.92, Issue 2, Feb 2004 Page(s):271 - 294.

IEEE Potentials Surveys 2009

About the Author

Suayb S. Arslan earned his B.S. degree with high honors in Electrical and Electronics Engineering Department from Bogazici University, Istanbul , Turkey and M.S. degree in Electrical and Computer Engineering Department from University of California, San Diego (UCSD), La Jolla, USA in 2009. During the summer of 2009, He was with imaging research group in Mitsubishi Electric Research Lab (MERL), Boston, MA. He is currently a Ph.d. candidate in UCSD and has been with Center for Wireless Communications (CWC) since 2006. He was awarded TUBITAK fellowship and UCSD departmental fellowship in the past. He is a Research Assistant in Center for Wireless Communication (CWC) lab at UCSD and interested in efficient coding schemes for wireless multimedia transmissions and cross layer optimizations for multimedia systems. He is a student member of IEEE and associate member of Sigma Xi.