

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Source Coding and Channel Coding for Mobile Multimedia Communication

Hammad Dilpazir<sup>1</sup>, Hasan Mahmood<sup>1</sup>, Tariq Shah<sup>2</sup> and Hafiz Malik<sup>3</sup>

<sup>1</sup>Department of Electronics, Quaid-i-Azam University, Islamabad

<sup>2</sup>Department of Mathematics, Quaid-i-Azam University, Islamabad

<sup>3</sup>Department of Electrical and Computer Engineering,  
University of Michigan - Dearborn, Dearborn, MI

<sup>1,2</sup>Pakistan

<sup>3</sup>USA

## 1. Introduction

In the early era of communications engineering, the emphasis was on establishing links and providing connectivity. With the advent of new bandwidth hungry applications, the desire to fulfill the user's need became the main focus of research in the area of data communications. It took about half a century to achieve near channel capacity limit in the year 1993, as predicted by Shannon (Shannon, 1948). However, with the advancement in multimedia technology, the increase in the volume of information by the order of magnitude further pushed the scientists to incorporate data compression techniques in order to fulfill the ever increasing bandwidth requirements. According to Cover (Cover & Thomas, 2006), the separation theorem stated by Shannon implies that it is possible for the source and channel coding to be accomplished on separate and sequential basis while still maintaining optimization. In the context of multimedia communication, the former represents the compression while the latter, the error protection. A major restriction, however, with this theorem is that it is applicable only to asymptotically lengthy blocks of data and in most situations cannot be a good approximation. So, these shortcomings have tacitly led to devising new strategies for joint-source channel coding.

This chapter addresses the major issues in source and channel coding and presents the techniques used for efficient data transmission for multimedia applications over wireless channels. The chapter highlights the strategies and notions of the source, channel, and joint source-channel coding. The remaining chapter is divided into two parts. The first part provides a brief description of digital image storage and transmission process, and the second part provides a brief description of digital video storage and transmission process.

### 1.1 Source coding

The process by which information symbols are mapped to alphabetical symbols is called *source coding*. The mapping is generally performed in sequences or groups of information and alphabetical symbols. Also, it must be performed in such a manner that it guarantees the

exact recovery of the information symbol back from the alphabetical symbols otherwise it will destroy the basic theme of the source coding. The source coding is called *lossless compression* if the information symbols are exactly recovered from the alphabetical symbols otherwise it is called *lossy compression*. The Source coding also known as compression or bit-rate reduction process. It is the process of removing redundancy from the source symbols, which essentially reduces data size. Source coding is a vital part of any communication system as it helps to use disk space and transmission bandwidth efficiently. The source coding can either be lossy or lossless. In case of lossless encoding, error free reconstruction of source symbols is possible, whereas, exact reconstruction of source symbols is not possible in case of lossy encoding. The minimum average length of codewords as function of entropy is restricted between an upper and lower bound of the information source symbols by the Shannon's source coding theorem (Shannon, 1948). For a lossless case, entropy is the maximum limit on the data compression, which is represented by  $H(\chi)$  and is defined as

$$H(\chi) = \sum_i p_x(i) \log p_x(i) \quad (1)$$

where  $p_x(\cdot)$  is the probability distribution of the information source symbols.

The Huffman algorithm is basically used for encoding entropy and to compress data without loss. In order to choose a particular representation for each symbol, Huffman coding makes use of a particular method that leads to a prefix code also called as *prefix-free code*. This method uses a minimum number of bits in the form of strings to represent mostly used and common source symbols and vice versa. Furthermore, Huffman coding uses a code table of varying length in order to encode a source symbol. The table is made on the basis of the probability calculated for all the values of the source symbol that have the possibility to occur. This scheme is optimal in terms of obtaining code words of minimum yet possible average length.

The coding concept of Shannon-Fano (Shannon, 1948) is used to construct prefix code using a group of symbols along with the probabilities by which they are measured or calculated. This technique, however, is not the optimum as it does not guarantee the code word of minimum yet possible average length as does the Huffman coding. In contrast, this coding technique always achieves code words lengths that lie within one bit range of the theoretical and ideal  $\log P(x)$ .

Shannon-Fano-Elias coding, on the other hand, works on cumulative probability distribution. It is a precursor to arithmetic coding, in which probabilities are used to determine codewords. The Shannon-Fano-Elias code is not an optimal code if one symbol is encoded at time and is a little worse than, for instance, a Huffman code. One of the limitation of these coding techniques is that all the methods require the probability density function or cumulative density function of the source symbols which is not possible in many practical applications. The universal coding techniques are proposed to address this dilemma. For instance, Lempel-Ziv (LZ) coding (Ziv & Lempel, 1978) and Lempel-Ziv-Welch (LZW) coding. The LZ coding algorithm is a variable-to-fixed length, lossless coding method.

A high level description of the LZ encoding algorithm is given as,

1. Initializing the dictionary in order to have strings/blocks of unity length.
2. Searching the most lengthy block  $W$  in the dictionary that matches the current input string.

3. Encode  $W$  by the dictionary index and remove  $W$  from the input.
4. Perform addition of  $W$  which is then followed by the a next symbol as input to the dictionary.
5. Move back to step 2.

The compression algorithm proposed in (Ziv & Lempel, 1978), shows that this technique asymptotically achieves the Shannon limit (Shannon, 1948). More details on existing source coding methods and their properties can be found in (Cover & Thomas, 1991) and references therein.

The rate distortion theory highlights and discusses the source coding with losses, i.e., data compression. It works by finding the minimum entropy (or information)  $R$  for a communication channel in order to approximately reconstruct the input signal at the receiver while keeping a certain limit of distortion  $D$ . This theory actually defines the theoretical limits for achieving compression using techniques of compression with loss. Today's compression methods make use of transformation, quantization, and bit-rate allocation techniques to deal with the rate distortion functions. The founder of this theory was C. E. Shannon (Shannon, 1948). The number of bits that are used for storing or transmitting a data sample are generally meant as *rate*. The distortion is the subject of discussion in these lines. The Mean Squared Error (MSE) commonly used distortion measure in the rate-distortion theory. Since most of lossy compression techniques operate on data that will be perceived by human consumers (listening to music, watching pictures and video), therefore, the distortion measure should be modeled on human perception. For example, in audio compression, perceptual models are comparatively very promising and are commonly deployed in compression methods like MP3 or Vorbis. However, they are not convenient to be taken into account in the rate distortion theory. Also, methods that are perception dependent are not considered promising for image and video compression so they are in general used with Joint Picture Expert Group (JPEG) and Moving Picture Expert Group (MPEG) weighting matrix.

Some other methods such as Adam7 Algorithm, Adaptive Huffman, Arithmetic Coding, Canonical Huffman Code, Fibonacci Coding, Golomb Coding, Negafibonacci Coding, Truncated Binary Encoding, etc., are also used in different applications for data compression.

## 1.2 Channel coding

The channel coding is a framework of increasing reliability of data transmission at the cost of reduction in information rate. This goal is achieved by adding redundancy to the information symbol vector resulting in a longer coded vector of symbols that are distinguishable at the output of the channel.

Channel coding methods can be classified into the following two main categories:

1. *Linear block coding* maps a block of  $k$  information bits onto a codeword of  $n$  bits such that  $n > k$ . In addition, mapping of  $k$  information bits is distinct, that is, for each sequence of  $k$  information bits, there is a distinct codeword of  $n$  bits. Examples of linear block codes include Hamming Codes, Reed Muller Codes, Reed Solomon Codes, Cyclic Codes, BCH Codes, etc. A Cyclic Redundancy Check (CRC) code can detect any error burst up to the length of the CRC code itself.

2. *Convolutional coding* maps a sequence of  $k$  bits of information onto a codeword of  $n$  bits by exploiting knowledge of the present  $k$  information bits as well as the previous information bits. Examples include Viterbi-decoded convolutional codes, turbo codes, etc.

The major theme of channel coding is to allow the decoder to decode the valid codeword or codeword with noise which means some bits would be corrupted. In ideal situation, the decoder knows the codeword that is sent even after corruption by noise. C. E. Shannon in his landmark work (Shannon, 1948) proposed a framework of coding the information to be transmitted over a *noisy channel*. He also provided theoretical bounds for reliable communication over a noisy channel as a function of the channel capacity.

### 1.3 Joint source-channel coding

The joint source-channel coding corresponds to the process of mapping (or encoding) the source data such that it can be compressed and protected against transmission errors over a noisy channel. Shannon's separation theorem states that source coding (compression by redundant bit removal) and channel coding (bit interleaving for error protection) can be performed separately and sequentially, while maintaining optimality. However, when operating under a *delay-constraint* on a *time-varying* channel, sequential optimality of two coders does not guarantee joint optimality. In order to achieve joint optimality, the source coding and the channel coding can be achieved through Joint Source-Channel Coding (JSCC). Thus, considerable interest is developed in various schemes of JSCC. For example, Gray and Ornstein (Gray & Ornstein, 1976) proposed sliding-block codes based on discrete-time-invariant nonlinear filtering to achieve JSCC. Gray and Ornstein provided theoretical bound for overall symbol error probability  $\epsilon > 0$  for the proposed time-invariant finite-length sliding-block encoder and decoder in terms of channel capacity  $C$  and entropy rate  $H(U) < C$ , where  $U$  denotes discrete ergodic source.

A technique of using joint source channel coding was proposed in (Modestino et al., 1981), where the author exploits the trade off between compression and error correction. The performance of the proposed framework was tested using 2D Discrete Cosine Transform (DCT) for source coding and convolutional codes for channel coding together. In (Ayanoglu & Gray, 1987), a design of joint source and channel trellis waveform coding was proposed. Effectiveness of the proposed method has been experimentally demonstrated. Another important result of the article was to show that the jointly optimized codes achieve performance either near or better than separately optimized codes. Farvardin and Vaishampayan (Farvardin & Vaishampayan, 1987) have also proposed a scheme for joint source-channel coding optimization, however, this scheme is based on the optimal quantizer design. Similarly, Goodman (Goodman et al., 1988) have also modeled the joint source-channel coding as a joint optimization problem. Sayood (Sayood et al., 1988) have provided a Maximum A Posteriori (MAP) based approach for jointly optimizing the problem of communication over an unreliable link. Effectiveness of their scheme has been demonstrated using image transmission over wireless channel. Wyrwas and Farrell (Wyrwas & Farrell, 1989) have shown that the joint source channel coding is more efficient for low resolution graphics transmission.

## 2. Image

Still images contain both spatial and psycho-visual redundancy. For efficient transmission or storage, all redundancy should be removed. However, transmission of compressed images over a noisy channel makes it very sensitive to the channel noise. Channel coding is generally used to combat these errors. This section provides an overview of image compression and its transmission over a wireless channel.

### 2.1 Image compression using JPEG

JPEG is among the most popular and widely used image compression methods. JPEG supports both lossy and lossless compression.

The JPEG compression scheme is divided into the following stages:

1. Transform the image into an optimal color space.
2. Down sample chrominance components by averaging groups of pixels together.
3. Apply a forward 2D DCT to blocks of pixels to remove redundancies from the image data.
4. Quantize each block of DCT coefficients using weighting functions which are optimized for the human eye.
5. Encode the resulting coefficients (image data) using entropy coding based on Huffman variable word-length algorithm to further remove redundancies in the quantized DCT coefficients.

Fig. 1 shows the block diagram of the JPEG encoder.

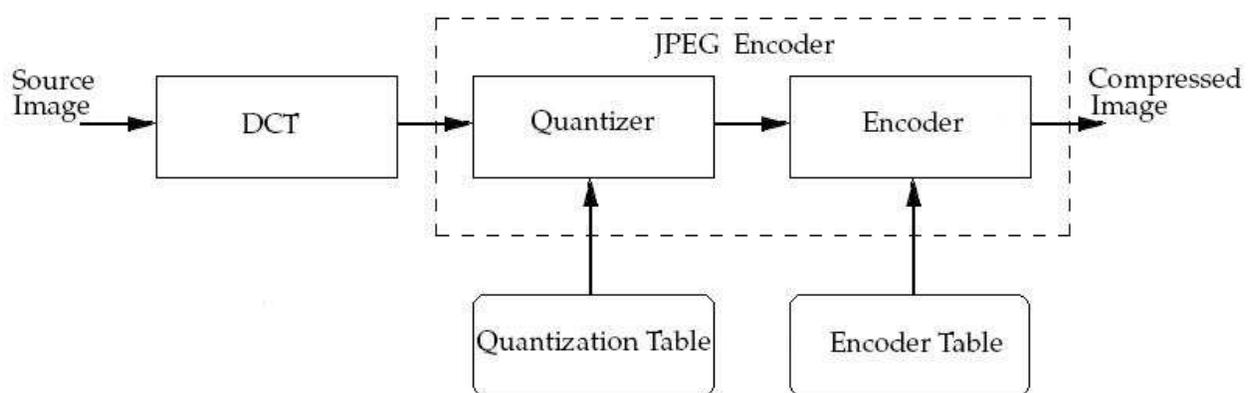


Fig. 1. JPEG encoder.

The gray-scale image contains a 2D set of pixels that correspond to the intensities (shades). There are 255 intensity levels, so pixel values are between 0 to 255. Since JPEG works on pixel values between -128 to 127, the values of source image are shifted to this scale. The image after scaling is decomposed into  $8 \times 8$  block. DCT is performed using standard method. In case of color images, the method is the same, however, color images are in the beginning decomposed into either three sets, for example, Red, Green, Blue (RGB) or Brightness, Hue, Saturation (BHS) or in four components Cyan, Magenta, Yellow, Black (CMYK). Hence, without the loss of generality, we can proceed with gray scale image.

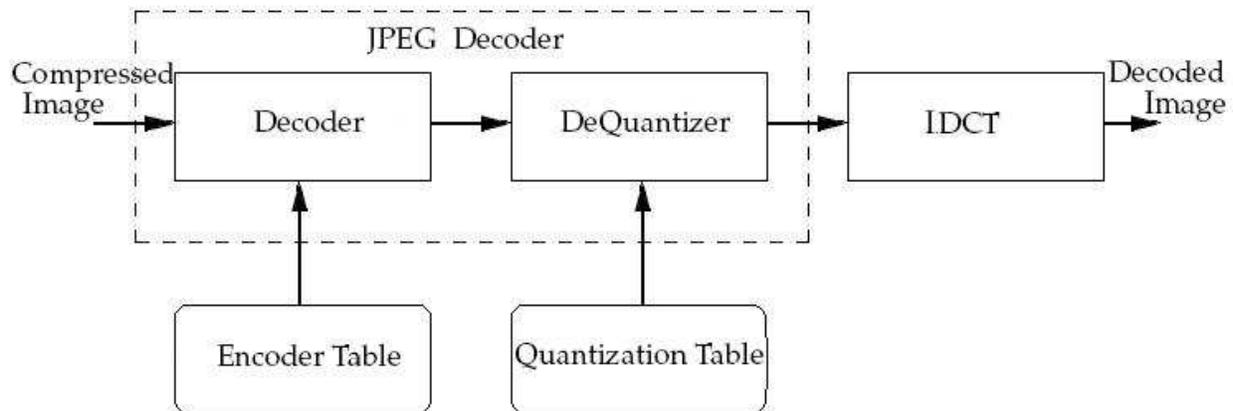


Fig. 2. JPEG decoder.

The image reconstruction, that is, the decoding of the compressed image is achieved by following encoding steps in reversed order. Fig. 2 shows the block diagram of a the decoder.

Up till now, the baseline requirements for JPEG were discussed. However, few extensions of JPEG standard also exists, such as attaining the progressive image buildup, arithmetic encoding and decoding for improved compression ratio and lossless compression scheme etc. Though these features are not essential for many JPEG implementations and hence are called extensions of JPEG standard.

The baseline JPEG image data is first received and decoded and then is reconstructed and displayed, therefore, baseline mode cannot be used for applications that demand for receiving the data stream and displaying it in a run. Instead of transmitting the image in lines the progressive buildup supports the layered based transmission. The layered based transmission is performed by sending a succession of image buildup, starting with an approximation of the image to the original image. The first scan is low quality JPEG image. Subsequent scans improve the quality gradually to refine the image. It can be observed that the first image appears to be rough but recognizable estimate of the original image but within an instance a refine image appears, this is because we need less data to produce the original image from the first estimate. Higher computational complexity is one of the limitations of progressive JPEG mode as with the progressive nature, every scan needs a full JPEG decompression cycle. For a fixed data transmission rates, faster JPEG decoder is required for efficient use of progressive transmission for real time applications.

Higher computational complexity is one of the limitations of progressive JPEG mode as each scan takes essentially a full JPEG decompression cycle to display. Therefore, with typical data transmission rates, a very fast JPEG decoder is needed to make effective use of progressive transmission to real time application.

Other JPEG extensions include arithmetic coding, lossless JPEG compression, variable quantization, selective refinement, image tiling, etc.

The JPEG standard became very popular after its publication, however, it has few shortcomings. This standard does not provide significant improvement for SNR scalability.

To overcome this problem, JPEG2000 standard was developed which contains some extra features like object and region based representation.

The JPEG2000 compression scheme is divided into the following stages:

1. *Color components transformation*: Transform the image into an optimal color space.
2. *Wavelet transform*: These color tiles are then transformed via a Discrete Wavelet Transform (DWT) to an arbitrary depth. The JPEG2000 uses two different wavelet transforms:
  - **irreversible**: The CDF 9/7 wavelet transform is said to be "irreversible" because it introduces quantization noise that depends on the precision of the decoder.
  - **reversible**: a rounded version of the bi-orthogonal CDF 5/3 wavelet transform. It uses only integer coefficients, so the output does not require rounding (quantization) and so it does not introduce any quantization noise. It is used in lossless coding.
3. *Quantization*: The wavelet coefficients are scalar-quantized to reduce the number of bits to represent them, at the expense of quality using weighting functions optimized for the human eye.
4. *Encoding*: The coefficients of each sub-band of every transformed block are arranged into rectangular blocks, also know as code blocks, which are coded individually, one bit plane at a time. Starting from the most significant bit plane with a nonzero element, each bit plane is processed in three passes. Each bit in a bit plane is coded in three passes only, which are called *significant propagation*, *magnitude refinement*, and *cleanup*. The output is arithmetically coded and grouped with similar passes from the other code blocks to form *layers*. These layers are partitioned into packets which are the fundamental units of the encoded stream.

Fig. 3 shows the block diagram of the JPEG2000 encoder. The JPEG2000 decoder simply applies the inverse operations of the encoder in reverse order. The JPEG2000 decoder simply inverts the operations of the encoder.

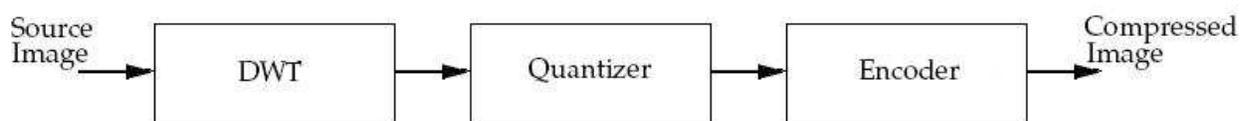


Fig. 3. JPEG2000 encoder.

## 2.2 Channel coding of images

When digital images are transmitted over a wireless channel, channel coding is generally used to combat errors due to noise in wireless channels. This section provides an overview of channel encoding for still image communication over the channel. Recent developments in the wireless communication and mobile computing have sparked interest in the area of multimedia communication over wireless channels (Modestino & Daut, 1979), (Modestino et al., 1981), (Sabir et al., 2006). The state-of-the-art encoding and decoding methods rely on joint source channel coding for optimal communication performance.

For example, Cai et al. (Cai & Chen, 2000) presented a scheme of transmitting images over a wireless channel. Fig. 4 shows the proposed encoding and decoding framework.

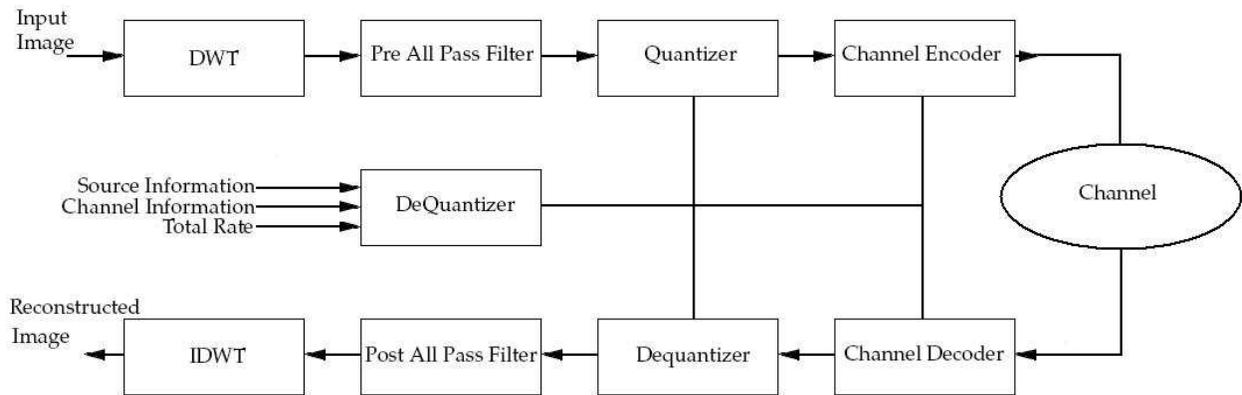


Fig. 4. Encoder-decoder framework presented in (Cai & Chen, 2000).

Similarly, Wu et al. (Wu et al., 2005) have developed a theoretical framework for image transmission over a wireless channel. This work exploits the error resilience property of JPEG2000 and achieves optimal distortion performance. Fig. 5 shows the semantic description of the proposed method. The problem is considered as a constraint optimization problem, i.e., minimizing the distortion by keeping the constraints on rate (or length of codewords), which can be expressed as,

$$\min \sum_b E[D_b(V_b)] = \min \sum_b D_{b,0} - E[\Delta D_b(V_b)] \quad \text{s.t.} \quad \sum_b L_b(V_b) \leq L_T \quad (2)$$

where  $E[D_b(V_b)]$  is the expected value of distortion  $D_b$  where  $V_b$  denotes the code. The term  $D_{b,0}$  is distortion when channel noise is zero, the term  $E[\Delta D_b(V_b)]$  is the expected reduction in distortion when a code  $V_b$  is adopted.  $L_b$  is the length of bit stream of code block  $b$ . For noisy channel Equation 2 reduces to

$$\min \sum_b -E[\Delta D_b(V_b)] \quad \text{s.t.} \quad \sum_b L_b(V_b) \leq L_T \quad (3)$$

which can be minimized through constrained optimization as,

$$\min \left\{ \sum_b -E[\Delta D_b(V_b)] + \lambda \sum_b L_b(V_b) \leq L_T \right\} \quad (4)$$

If we assume that the JPEG2000 decoder can efficiently correct the data then the constrained optimization can be expressed as,

$$\begin{aligned} f(V_b) = & - \sum_{j=1}^{N'_c-1} \left( \sum_{k=1}^j d_k \right) \left( \prod_{m=1}^j \left[ 1 - P\left(r_m, \frac{l_m}{r_m}\right) \right] \right) P\left(r_{j+1}, \frac{l_{j+1}}{r_{j+1}}\right) \\ & - \left( \sum_{k=1}^{N'_c} d_k \right) \left( \prod_{m=1}^{N'_c} \left[ 1 - P\left(r_m, \frac{l_m}{r_m}\right) \right] \right) + \lambda \sum_{n=1}^{N'_c} \frac{l_n}{r_n} \end{aligned} \quad (5)$$

where  $r_i$  is the rate of  $i$ th code pass,  $d_i$  is the associated distortion reduction of  $l_i$  bytes,  $P(r_i, \frac{l_i}{r_i})$  is the probability that error has occurred using  $i$ th code pass and  $N'_c$  is the number of code in a certain  $V_b$ .

Exhaustive search method can be used to obtain an optimal  $V_b$ . This method searches all possible  $V_b$  code-blocks which is equivalent to searching overall search space, that is,  $\sum_{i=1}^{N'_c} R^i$ . It can be observed that exhaustive search method is memory wise inefficient. To address this limitation, constrained search technique can be used. To this end, Equation 5 is reformulated as,

$$f(V_b) = \sum_{i=1}^{N'_c} \left[ \left\{ \prod_{j=1}^i \left( 1 - P(r_j, \frac{l_j}{r_j}) \right) \right\} (-d_j) + \lambda \frac{l_i}{r_i} \right] \quad (6)$$

Solving Equation 6 results in optimum rate which can be expressed as,

$$r_p^* = \begin{cases} \operatorname{argmin}_{r_{k+2} \in \mathfrak{R}(k+2)} J_{k+1}(r_{k+1}), & \text{when } p = k + 2 \\ \operatorname{argmin}_{r_p \in \mathfrak{R}(p)} J_{p-1}(r_{p-1}^*), & \text{when } p = k + 2 \end{cases} \quad (7)$$

where  $J_k$  is the cost function at stage  $k$ . Further details on this method can be found in (Wu et al., 2005).

### 3. Video

Video contains spatial redundancy as well as temporal redundancy due to moving images. Similar to image data, for efficient transmission or storage, all redundancy should be removed. However, to make video transmissions immune to transmission errors due to noisy channel, channel coding is used. This section provides a brief overview of video compression methods and transmission of video over a wireless channel.

#### 3.1 Source coding of videos

Video is a sequence of frames/pictures being played at certain rate, for example, 15 or 30 frames per second. Video compression methods exploit spatial and temporal redundancy in the video data. Most video compression techniques are lossy and operate on the premise that much of the data present before compression is not necessary for achieving good perceptual quality. For example, DVDs use MPEG-2 video coding standard that can achieve compression ratio of 15 to 30, while still producing a picture quality which is generally considered high-quality for standard-definition video. Video compression is a trade off between disk space, video quality, and the cost of hardware required to decompress the video in a reasonable time. However, if the video is over compressed in a lossy manner, visible (and sometimes distracting) artifacts can appear. Fig. 5 shows a general video encoder.

#### 3.2 MPEG video basics

The Moving Picture Expert Group (MPEG), worked to generate the specifications under the International Organization for Standardization (ISO). The MPEG video actually consists of three finalized standards, MPEG-1, MPEG-2 and the most recent MPEG-4. The MPEG-1 & -2 standards are similar in basic concepts. They both are based on motion compensated

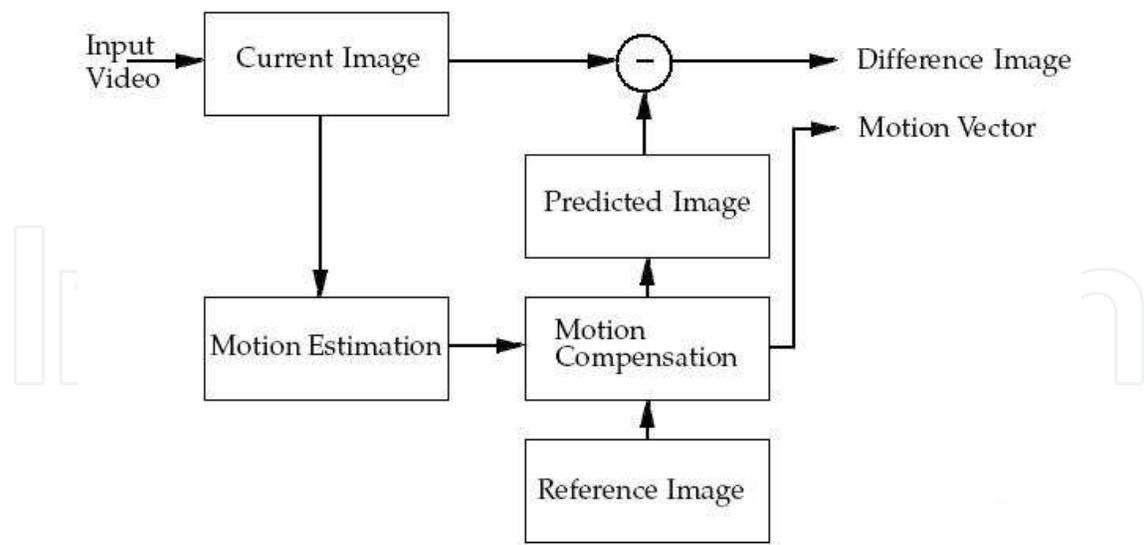


Fig. 5. General video encoder Drozdek (2002).

block-based transform coding techniques, while MPEG-4 deviates from these more traditional approaches in its usage of software image construct descriptors, for target bit-rates in the very low range (< 64Kb/sec).

### 3.2.1 MPEG video layers

MPEG video is broken up into a hierarchy of layers to combat errors, random search and editing, and synchronization. From the top level, the first layer is known as the video sequence layer, and it is any self-contained bit stream, for example a coded movie or an advertisement. The second layer is the Group of Pictures (GOP), which is composed of 1 or more groups (GOPs) of intra (*I*) frames and non-intra (*P* or *B*) frames. Of course, the third layer is the picture layer itself, and the next layer beneath it is called the slice layer. Each slice is a contiguous sequence of raster ordered *macroblocks*, most often on a row basis in typical video applications, but not limited to this by the specification. Each slice consists of macroblocks, which are 16x16 arrays of luminance pixels, or picture data elements, with *8times8* arrays of associated chrominance pixels. The macroblocks can be further divided into distinct *8times8* blocks, for further processing such as transform coding. Each of these layers has its own unique 32 bit start code defined in the syntax to consist of 23 zero bits followed by a 1, then followed by 8 bits for the actual start code. These start codes have no limitation on number of zero bits preceding them.

### 3.2.2 Intra frame coding techniques

The term *intra frame coding* refers to the fact that various lossless and lossy compression techniques are performed relative to information that is contained only within the current frame, and not relative to any other frame in the video sequence. In other words, no temporal processing is performed outside of the current frame. The basic processing blocks used for I-frame coding are the video filter, DCT, DCT coefficient quantizer, and run-length amplitude/variable length coder.

### 3.2.3 P/B frame coding techniques

The I-frame coding techniques are limited to processing the video signal on a spatial basis, relative only to information within only the I-frame only. Considerable amount of compression efficiency is obtained by exploiting inherent temporal redundancies. Temporal processing to exploit this redundancy uses a technique known as motion compensation. Details of prediction (P) and bidirectional (B) frame coding techniques are discussed next.

1. **P Frame Encoding:** Starting with I-frame, the encoder can forward predict a future frame. This is commonly referred to as a P frame, and it may also be predicted from other P frames, although only in a forward time manner. Consider, for example, a GOPs consisting of 5 frames. In this case, the frame ordering can be expressed as I, P, P, P, I, P, P, P, I, .... Each P frame in this sequence is predicted from the frame immediately preceding it, whether it is an I frame or a P frame.
2. **B Frame Encoding:** B frames are commonly referred to as bi-directional interpolated prediction frames. These frames can use forward/backward interpolated prediction. As an example of the usage of I, P, and B frames, consider a GOPs consisting of 8 frames, for example, I, B, P, B, P, B, P, I, B, P, B, P, B, P, B, I, .... As I frames are coded spatially only and the P frames are encoded using forward prediction based on previous I and P frames. The B frames, on the other hand, are coded based on a forward prediction from a previous I or P frame, as well as a backward prediction from a succeeding I or P frame. Consider, for example, sequence of GOPs consisting of 8 frames, the first B frame is predicted from the first I frame and first P frame, the second B frame is predicted from the second and third P frames, and the third B frame is predicted from the third P frame and the first I frame of the next group of pictures, and so on. It can be observed from the above example that backward prediction requires that the future frames that are to be used for backward prediction be encoded and transmitted first, out of order. The main advantage of the usage of B frames is coding efficiency. In most cases, B frames results in less bits being coded overall. Quality can also be improved in the case of moving objects that reveal hidden areas within a video sequence. Backward prediction in this case allows the encoder to make more intelligent decisions on how to encode the video within these areas. Since B frames are not used to predict future frames, errors generated will not be propagated further within the sequence.

### 3.2.4 Motion estimation

The temporal prediction technique used in MPEG video is known as *motion estimation*. The basic premise of motion estimation is that in most cases, consecutive video frames will be similar except for changes induced by objects moving within the frames. In the trivial case of zero motion between frames, it is easy for the encoder to efficiently predict the current frame as a duplicate of the previous frame. When this is done, the only information necessary to transmit to the decoder becomes the syntactic overhead necessary to reconstruct the picture from the original reference frame. When there is motion in the images, the situation is not as simple. To solve motion estimation problem, a comprehensive 2D spatial search is performed for each luminance macroblock. Motion estimation is not applied directly to chrominance in MPEG video, as it is assumed that the color motion can be adequately represented with the

same motion information as the luminance. The MPEG standard does not define how this search should be performed.

### 3.2.5 Coding of residual errors

The residual error frame is generated by subtracting predicted frame from its reference frame. The residual error is coded using processing steps similar to I frame encoding. However, DCT coefficient quantization method is different. A constant matrix (of value 16) for each of the 64 locations is set as default quantization matrix for non-intra frames. The non-intra quantization step function contain dead zone around zeros to eliminate any single DCT coefficient to maximize the run length amplitude efficiency. The variable length code and the differential value are calculated for the residual block information to measure the motion vector based on their statistical likelihood of occurrence.

### 3.3 Channel coding of videos

The channel coding is an essential component in reliable transmission of video data over a wireless channel. This section provides a brief overview of commonly used channel coding methods used for video transmission over a wireless channel. Since video contains huge amount of data, for optimum use of resources, joint source-channel coding strategy is generally used.

Dyck and Miller (Dyck & Miller, 1999) have proposed rate allocation schemes for source and channel codes, designing of channel codes for specific source codes, and power allocation for modulated symbols, etc. Fig. 6 shows joint source-channel coding framework as proposed in (Dyck & Miller, 1999).

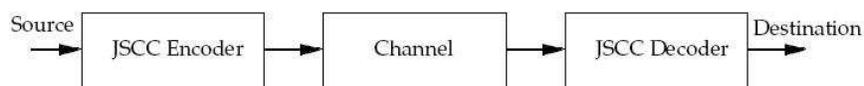


Fig. 6. Encoder decoder model presented in (Dyck & Miller, 1999).

In the JSCC block, the source encoder operates on an input vector  $X$  which belongs to  $n$ -dimensional space  $\mathbb{R}^n$  and is used to generate quantization index  $I \in \mathfrak{S}$ , where  $\mathfrak{S}$  is the index set such that  $\mathfrak{S} = \{1, 2, \dots, N\}$ . The channel maps  $I$  to  $J$  such that  $J \in \mathfrak{S}$ .

The JSCC decoder estimates  $\hat{X}$  from  $J$ . The optimal mapping of  $X$  to  $J$  that minimize the average distortion between source and destination is obtained by using Vector Quantization (VQ). The expected distortion for transmission of videos over wireless link is given as

$$E(D) = \sum_{i=1}^N \sum_{x \in \chi_r : x \in \mathbb{R}^n} \sum_{j=1}^N P(j|i) d(x, y_j) \quad (8)$$

where  $P(\cdot|\cdot)$  is the channel transition probability,  $d(\cdot, \cdot)$  is the dissimilarity metric between two vectors (in case of Minimum Mean Square Estimator (MMSE) the  $d(x, y) = \|x - y\|^2$ ),  $N$  is the number of decoder indices,  $\chi$  and  $R$  are the source and encoder realization respectively. The

optimal encoding rule to minimize Equation 8 is given as (Kumazawa et al., 1984),

$$y_i = \frac{\sum_{i=1}^N \sum_{x \in \mathcal{X}_r : x \in \mathfrak{R}_i} P(j|i) E[X|X \in \mathfrak{R}_i]}{\sum_{i=1}^N \sum_{x \in \mathcal{X}_r : x \in \mathfrak{R}_i} P(j|i)} \quad (9)$$

$$= E[X|j]$$

where

$$R_i = \sum_{j=1}^N P(j|i) d(x, y_i) \leq \sum_{j=1}^N P(j|k) d(x, y_i), \quad \forall k. \quad (10)$$

In the decoding process after assignment of appropriate index to the source symbol is decoded using either MAP or conditional Mean Square Estimation (MSE) method.

In case of MAP decoding, decoder estimates the transmitted sequence and use them to estimate the original transmitted symbols. The basic method for MAP decoder is to select the symbol that maximizes the a posteriori probability, i.e,  $P[i|j]$ . Since it is assumed that the channel is memory less, it can be written as  $\hat{i} = \underset{i}{\operatorname{argmax}} P[i|j]$ .

In case of conditional MSE decoding, the decoder decodes source symbols using conditional MSE. The joint source channel coding for conditional MSE is considered by (Miller & Park, 1998). The expected distortion can be written in the form

$$E[D(X, y)|J] = \frac{\sum_i \left( \sum_{t=1}^T E[d(X_t, y_t)|I = i] \right) P[J = j|I = i] P[I = i]}{\sum_i P[J = j|I = i] P[I = i]}. \quad (11)$$

To find  $y$  by minimizing distortion, many algorithms are available (Baum et al., 1970), for example

$$\hat{D} = \sum_{t=1}^T \sum_{l=1}^T \|y(l) - y_t\|^2 P[I_t = i|J = j]. \quad (12)$$

Based on the expected distortion given in Equation 12, the optimal decoder with minimum expected distortion, is given as

$$y_t = \sum_{k=1}^N y(k) P[I_t|J = j], \quad \forall t \quad (13)$$

where  $y_t$  is the decoder estimate which minimizes the distortion such that it becomes the optimal channel vector quantization decoder. In Equation 13, the a posteriori probabilities  $P[I_t|J = j]$  are computed either by BCJR algorithm (Bahl et al., 1974) by forward and backward recursive equation, or by Hidden Markov Model (HMM) (Baum et al., 1970).

In mobile communications, the channel behaves as a Rayleigh fading channel due to mobility and variable path delays. Under Rayleigh fading conditions, the channel can be more noisy, therefore, robust methods are required for such applications. The methods for decoding are discussed by considering vector quantization only. However, for more effective transmission, we need to introduce channel coding along with source coding (Man et al., 1997).

Kwon et al., (Kwon & kyoon Kim, 2002) have proposed an adaptive scheme for code rate selection in the video communication. This adaptive technique is useful for transmission over noisy channel in terms of its computational efficiency. The transmission side contains video encoder followed by a channel encoder and a rate controller, which adaptively learns the channel condition and allocates the rate accordingly. In this work, the distortion at any time  $t$  can be expressed as,

$$D(t) = D_s(R_s(t)) + D_c(R_c(t)) \quad (14)$$

where  $D_s(R_s(t))$  is the distortion caused by the source and  $D_c(R_c(t))$  is the distortion caused by the channel. These two distortions are assumed to be uncorrelated. The  $R_s(t)$  is the source coding rate, and  $R_c(t)$  is the channel coding rate. Generally, statistical characterization of both distortions are assumed to be known. To minimize the distortion via selecting a code rate, the channel is assumed to be independent and exhibits bursty nature. The adaptive scheme is based on average Peak Signal to Noise Ratio (PSNR). The adaptive scheme encodes at an optimal rate  $r^*$  which provides the maximum average PSNR, that is,

$$r^* = \arg \max_{r^*} \text{avgPSNR}(r^*, P_s, L_B) \quad (15)$$

where  $P_s$  is the symbol error probability and  $L_B$  is the average burst error length. The approach of finding the maximum of average PSNR for the given  $r$ ,  $P_s$  and  $L_B$  can be a time consuming process. However, for a given  $r$ ,  $P_s$ ,  $L_B$  and the average residual error probability  $P_R^{avg}$ , the optimum code rate can be found as,

$$r^* = \operatorname{argmin}_r [P_R(r, P_s, L_B) - P_R^{avg}] \quad (16)$$

Yingjun (Yingjun et al., 2001) have also proposed an adaptive scheme based on adaptive segmentation for video transmission over a wireless channel. It is known that the P frame has both spatial and temporal redundancy, hence, its error resistance is more than I frame, which only has spatial redundancy. The proposed scheme only focuses on error resistance of I frame. By protecting the I frame, this scheme provide good visual quality. This transmission scheme is shown in Fig. 7.

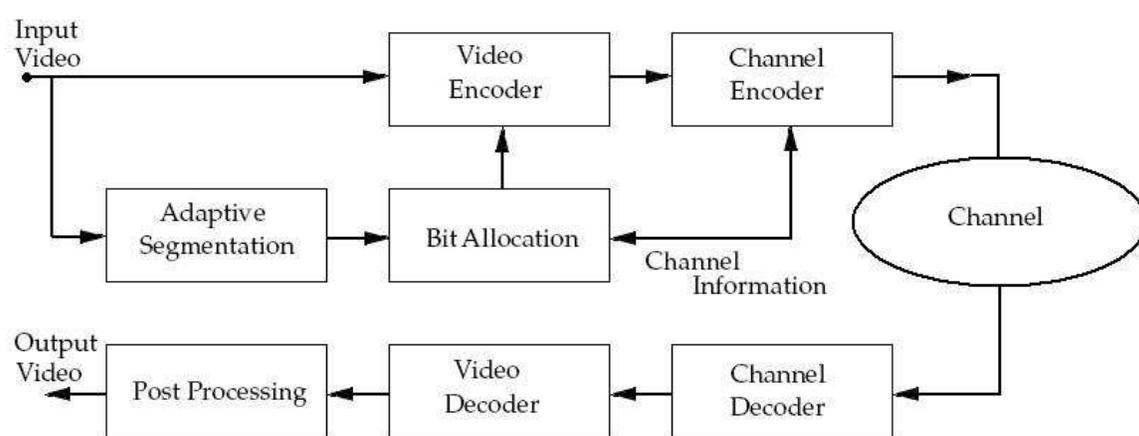


Fig. 7. Video transmission model (Yingjun et al., 2001).

In the transmission system, the video frames are divided into segments at macro level, the rate distortion analysis is performed which provides the basis of finding optimal bit allocation scheme (Lu et al., 1998). The bit allocation is a tradeoff between quantization level and the channel coding rate. In the next step, the resulted segmentation is coded and transmitted over a wireless channel and at the receiver, the inverse process is performed in reverse order to recover the transmitted video. In order to conceal the error, post processing block is added. The proposed joint source channel design based on Hotelling transformation and the rate distortion function is given as,

$$T_{rd} = C_V(V_{rd} - m_V) \quad (17)$$

where  $C_V$  is the variance and  $m_V$  is the mean of rate distortion vector  $V_{rd}$ . The component of transformed version of rate distortion vector  $T_{rd}$  is observed to have zero value. Hence the vector dimension is reduced significantly by this transformation. In addition, the transformed version helps to find an optimal bit allocation mechanism. The segmentation is performed in adaptive fashion. In the adaptive scheme, first the threshold  $T$  is decided for control segmented regions number. Then  $D_{rd}$  is rate distortion function which can be expressed as,

$$D_{rd}(i, j) = \|U_{rd,i} - U_{rd,j}\|. \quad (18)$$

In addition to adaptive learning algorithms, an iterative algorithm has also been proposed by Nasruminallah (Nasruminallah & Hanzo, 2011). The system model considered for iterative source channel coding is given in Fig. 8.

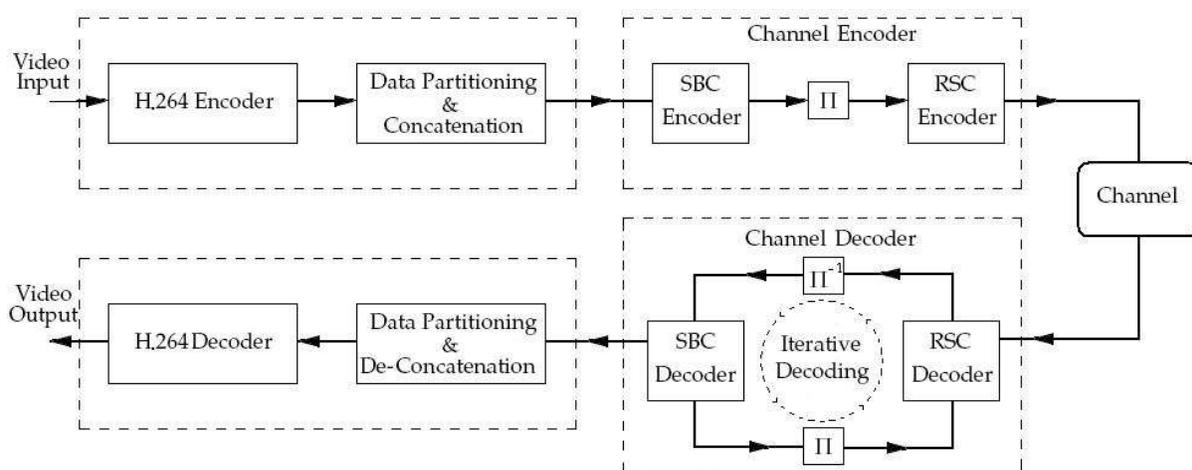


Fig. 8. System model in (Nasruminallah & Hanzo, 2011).

A more realistic formulation of H.264/AVC instead of conventional mathematical modeling for video source encoding is presented in (Nasruminallah & Hanzo, 2011). The video is encoded with H.264/AVC, followed by a data partitioning block, which de-multiplexes the data and concatenates it to generate a bit stream. This bit stream is then fed to the channel encoder which first passes it through the Short Block Codes (SBC). These encoded bits are then interleaved and fed to the Recursive Systematic Convolutional (RSC) encoder also known as serial concatenation, with inner scheme as RSC and outer scheme as SBC. The symbols at the output of RSC encoder are modulated (not shown in Fig. 8) and then transmitted over the channel. The channel corrupts the modulated signals with the assumption that the channel is

narrow band and temporally correlated with Rayleigh fading. The signal at the receiver end is demodulated and the corrupted version of the source symbols are fed to the RSC decoder, which shares the data with the SBC decoder to start the iterative process. The channel decoder uses standard iterative decoding process (Berrou & Glavieux, 1996), i.e., with the exchange of extrinsic information between the two decoders, for message decoding. The overall process can be summarized as the soft information is at the input of RSC decoder which processes it to extract the Log-Likelihood Ratios (LLR) estimates. These estimates are de-interleaved and fed to the SBC decoder. This process repeats until the decoding converges to some reasonable estimate. Kliewer (Kliewer et al., 2006) have proposed the condition to ensure convergence of the iterative decoding. The encoding of SBC encoder for rate  $\frac{K}{K+1}$  such as  $\frac{2}{3}$ ,  $\frac{3}{4}$ ,  $\frac{4}{5}$ , etc., can be achieved by placing the redundant bit at any position in a codeword by taking the Exclusive-OR (XOR) of the source bits as,

$$r = [b(1) \oplus b(2) \oplus b(3) \oplus \dots \oplus b(K)] \quad (19)$$

where  $\oplus$  denotes XOR operation.

This bit can be placed at any  $K + 1$  position in the codeword. Furthermore, to attain a more powerful code, the SBC encoder produces information rate  $\frac{K}{N}$  where  $N = m \times K$ . The first  $((m - 1) \times K)$ -tuples are produced by repeating the information block of size  $K$  and the last set is produced as,

$$r(p) = [b(1) \oplus b(2) \oplus b(3) \oplus \dots (b(p) = 0) \oplus \dots \oplus b(K)] \quad (20)$$

where  $p$  can take values between  $[1, 2, \dots, K]$ .

Nasruminallah has demonstrated the effectiveness of SBC and RSC methods based on different code rates for SBC and rate  $\frac{1}{4}$  for RSC with constituent code  $g_1 = [1011]$ ,  $g_2 = [1101]$ ,  $g_3 = [1101]$ ,  $g_4 = [1111]$ .

#### 4. References

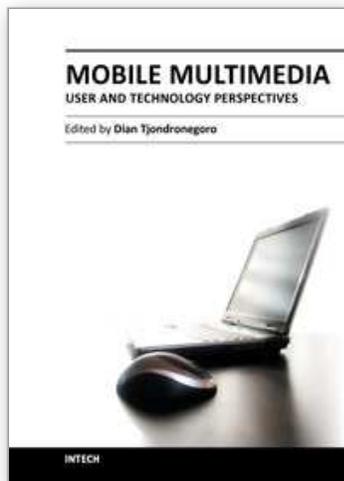
- Ayanoglu, E. & Gray, R. (1987). The design of joint source and channel trellis waveform coders, *IEEE Transactions on Information Theory* 33(6): 855–865.
- Bahl, L., Cocke, J., Jelinek, F., & Raviv, J. (1974). Optimal decoding of linear codes for minimizing symbol error rate, *IEEE Transactions on Information Theory* 20(2): 284–287.
- Baum, L. E., Petrie, T., Soules, G. & Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains, *Annals Mathematics Statistics* 41(1): 164–171.
- Berrou, C. & Glavieux, A. (1996). Near optimum error correcting coding and decoding: turbo-codes, *IEEE Transactions on Communication* 44(10): 1261–1271.
- Cai, J. & Chen, C. W. (2000). Robust joint source-channel coding for image transmission over wireless channels, *IEEE Transactions on Circuits and Systems for Video Technology* 10(6): 962–966.
- Cover, T. M. & Thomas, J. A. (1991). *Elements of Information Theory*, John Wiley and Sons Inc.
- Cover, T. M. & Thomas, J. A. (2006). *Elements of Information Theory*, John Wiley and Sons Inc.
- Drozdek, A. (2002). *Elements of Data Compression*, Bill Stenquist.

- Dyck, R. E. V. & Miller, D. J. (1999). Transport of wireless video using separate, concatenated, and joint source-channel coding, *Proceedings of IEEE* 87(10): 1734–1750.
- Farvardin, N. & Vaishampayan, V. (1987). Optimal quantizer design for noisy channels: An approach to combined source - channel coding, *IEEE Transactions on Information Theory* 33(6): 827–838.
- Goodman, D., Moulslley, T.J. & Holmdel, N. (1988). Using simulated annealing to design digital transmission codes for analogue sources, *IEEE Electronics Letters* 24(10): 617–619.
- Gray, R. M. & Ornstein, D. S. (1976). Sliding-block joint source/noisy-channel coding theorems, *IEEE Transactions on Information Theory* 22(6): 682–690.
- Kliewer, J., Goertz, N. & Mertins, A. (2006). Iterative source-channel decoding with Markov random field source models, *IEEE Transactions on Signal Processing* 54(1): 3688–3701.
- Kumazawa, H., Kasahara, M. & Namekawa, T. (1984). A construction of vector quantizers for noisy channels, *Electronics and Communications in Japan* 67(4): 39–47.
- Kwon, J. C. & Kim, J. K. (2002). Adaptive code rate decision of joint source-channel coding for wireless video, *IEEE Electronics Letters* 38(25): 1752–1754.
- Lu, J., Liou, M. L., Letaief, K. B. & Chuang, J. C. I (1998). Mobile image transmission using combined source and channel coding with low complexity concealment signal processing, *IEEE Signal Processing and Image Communication* 86(5): 974–997.
- Man, H., Kossentini, F. & Smith, M. J. (1997). Robust EZW image coding for noisy channels, *IEEE Signal Processing Letters* 4(1): 227–229.
- Miller, D. J. & Park, M. (1998). A sequence-based, approximate MMSE decoder for source coding over noisy channels using discrete hidden Markov models, *IEEE Transactions on Communication* 46(1): 222–231.
- Modestino, J., Daut, D. & Vickers, A. (1981). Combined source-channel coding of images using the block cosine transform, *IEEE Transactions on Communication* 29(9): 1261–1274.
- Modestino, J. W. & Daut, D. G. (1979). Combined source-channel coding of images, *IEEE Transactions on Communication* 27(11): 1644–1659.
- Nasruminallah & Hanzo, L. (2011). Joint source-channel iterative decoding of arithmetic codes, *IEEE Communications Surveys and Tutorials* 1(99): 1–27.
- Sabir, M. F., Sheikh, H. R., Heath, R. W. & Bovik, A. C. (2006). Joint source-channel distortion model for JPEG compressed images, *IEEE Transactions on Communication* 15(6): 1349–1364.
- Sayood, K., Gibson, J. D. & Liu, F. (1988). Implementation issues in map joint source/channel coding, *Twenty-Second Asilomar Conference on Signals, Systems and Computers* 1(1): 102–106.
- Shannon, C. E. (1948). A mathematical theory of communication, *Bell System Technical Journal* 27(1): 379–423, 623–656.
- Wu, Z., Bilgin, A. & Marcellin, M. W. (2005). Joint source/channel coding for image transmission with JPEG2000 over memoryless channels, *IEEE Transactions on Image Processing* 14(8): 389–400.
- Wyrwas, R. & Farrell, P. (1989). Joint source-channel coding for raster document transmission over mobile radio, *IEEE Proceedings Communications, Speech and Vision* 136(6): 375–380.

- Yingjun, S., Jianhua, L., Wang, J., Letaier, K. B. & Gu, J. (2001). Adaptive segmentation based joint source-channel coding for wireless video transmission, *IEEE VTS 53rd Vehicular Technology Conference VTC 3*(1): 1752–1754.
- Ziv, J. & Lempel, A. (1978). Compression of individual sequences by variable rate coding, *IEEE Transactions on Information Theory* 24(1): 530–536.

IntechOpen

IntechOpen



## **Mobile Multimedia - User and Technology Perspectives**

Edited by Dr. Dian Tjondronegoro

ISBN 978-953-307-908-0

Hard cover, 156 pages

**Publisher** InTech

**Published online** 20, January, 2012

**Published in print edition** January, 2012

As multimedia-enabled mobile devices such as smart phones and tablets are becoming the day-to-day computing device of choice for users of all ages, everyone expects that all mobile multimedia applications and services should be as smooth and as high-quality as the desktop experience. The grand challenge in delivering multimedia to mobile devices using the Internet is to ensure the quality of experience that meets the users' expectations, within reasonable costs, while supporting heterogeneous platforms and wireless network conditions. This book aims to provide a holistic overview of the current and future technologies used for delivering high-quality mobile multimedia applications, while focusing on user experience as the key requirement. The book opens with a section dealing with the challenges in mobile video delivery as one of the most bandwidth-intensive media that requires smooth streaming and a user-centric strategy to ensure quality of experience. The second section addresses this challenge by introducing some important concepts for future mobile multimedia coding and the network technologies to deliver quality services. The last section combines the user and technology perspectives by demonstrating how user experience can be measured using case studies on urban community interfaces and Internet telephones.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Hammad Dilpazir, Hasan Mahmood, Tariq Shah and Hafiz Malik (2012). Source Coding and Channel Coding for Mobile Multimedia Communication, Mobile Multimedia - User and Technology Perspectives, Dr. Dian Tjondronegoro (Ed.), ISBN: 978-953-307-908-0, InTech, Available from:

<http://www.intechopen.com/books/mobile-multimedia-user-and-technology-perspectives/source-coding-and-channel-coding-for-mobile-multimedia-communications>

**INTECH**  
open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen