MOTION ESTIMATION AND CONVOLUTIONAL CODING FOR VIDEO STREAMING OVER WIRELESS CHANNELS

SALAH A. ALIESAWI, SALAH S. MUSTAFA, SAEED A. GATHBAN
Department of Computer Science, College of Computer Science and Information Technology, University of Anbar, Iraq
E-mail: salaheng1996@gmail.com, dr salah_rawi@yahoo.com, saeedkareem1978@gmail.com

ABSTRACT

In this paper, a source and channel coding scheme is proposed for lossy video transmission over wireless channels. The source coding is used to reduce the redundancy so that the bit streams/video can be stored or sent efficiently over a network. The block-Matching (BM) with Quintet Search (QS) algorithm for motion estimation (ME) uses as a source coding or video compression technique. In the new QS algorithm, the number of searching points is reduced to five points instead of eight points in standard algorithm. The channel/convolutional coding (CC) adds useful redundancy to combat channel effects of the wireless environments and improve the system performance. The results of compression using different searching algorithms and different settings of the H.261 encoder parameters are computed for different videos with 176×144 dimension to select the proper settings that can be applied to the input video streams being transmitted over wireless links. The results show that the proposed system can produce a balance between the compression performance and preserving the video quality in video conference applications. Further, the performance of coded systems over additive white Gaussian Noise (AWGN) channels is also computed and compared with the system without channel coding.

Keywords: Block-Matching, Motion Estimation, Convolutional Coding, Video Transmission.

1. INTRODUCTION

The rapid developments in computing technology, data compression, and high bandwidth networks, have made it viable to design wireless communication systems for multimedia services, such as video streaming over the Internet, and multimedia messaging services, over wireless channels. The Internet does not offer satisfactory quality of service for multimedia transmission as the automatic repeat request (ARQ) in Internet transport protocol (TCP) may be shattering for multimedia streams due to their stringent delay requirements. On the other hands, the services with wireless mobile channels are even more inferior. Thus, it is a challenging task to design efficient transmission systems that can operate in such bandwidth-limited and unreliable noisy transmission environments [1-3]. To deal with these challenges, not only the video needs to be compressed efficiently, but also the transmission scheme needs to provide some error protection to deal with such wireless channels.

Multimedia data such as videos has also time constraints, where its transmission should be processed in real-time and must be played out continuously. Thus, compression is utilized to reduce the size of multimedia transmitted over the wireless channels to increase and improve the performance of such transmission systems. Compression discards the redundancies present in video using codecs that include the original video codec and make it suitable for the transmission and storage operations. The later step is the decoding process and return the video to the original state for playback and display. In [4], a new video coding scheme known as H.264 is proposed for video signal compression. The scheme is based on discrete cosine transform (DCT) and wavelet and allows best video compression rates without compromising in image quality. Video codec for video conferencing applications, is proposed in [5] and the algorithm is similar to the standard JPEG. In [6], a scalable video coding (SVC) is proposed for video transmission using orthogonal frequency division multiple Access (OFDM) systems. The system have demonstrated performance
improvement and reliable transmission in compared to a single carrier system. The optimal resource allocation scheme for H.264/SVC video transmission using OFDM is also proposed in [7] for cognitive radio (CR) networks. The scheme adopts a new probabilistic method to mitigate the imposed interference by cognitive users.

In [8], the researchers proposed uncoded video transmission system for multiuser wireless networks. The channel effects with different conditions are studied and the system has good resource allocation performance in compared with the traditional allocation algorithms. However, to reduce the error in the process of transmitted video and receiving data, there are several techniques, such as channel coding and modulation techniques, can be applied to reconstruct the original transmitted data and improve bit error rate (BER). Error correction code such as convolutional codes (CC), Reed-Solomon codes, Turbo codes have been applied such as [1,9,10], to avoid the multipath channel effects that causes extensive ghosting of the transmission. This ghosting causes inter symbol interference (ISI), blurring the time domain signal [9-13].

In this paper, a system for video transmission over noisy wireless channels is proposed by combining source coding and convolutional coding to deal effectively with errors due to channel effects and save transmission bandwidth. For source coding, the block matching (BM) algorithm for motion estimation (ME) is used as a powerful technique for high compression ratio (CR) [14,15]. The main approach of ME with new QS algorithm, is used to reduce computational complexity by reducing the searching points and redundancy present in the transmitted video. The rest of the paper is organized as follows: Section 2 presents the proposed video transmission system models with the source and channel coding. Section 3 presents the simulation environment setting, performance evaluation of source coding algorithm and transmission system using three different streaming videos and case studies (Video conferencing, Mobile networks). Finally, conclusions are drawn in Section 4.

2. VIDEO TRANSMISSION SYSTEM

The proposed transmission system models are shown in Fig.1. The video is inserted after being divided into a total frame size of (144 × 176). The frame size (144 × 176) is generally used in video conferencing or mobile video with model Quarter Common Intermediate Format (QCIF). It is commonly used in video compression format standard H.261/H.263 encoder/decoder. The source encoder carries out redundancy reduction coding for video streams and convert them into binary array of two dimensions to facilitate the transmission process. In figure 1 (b), the output stream is encoded by channel/convolutional coding, and modulated one by one, while in figure 1 (a), the output stream is only modulated by binary phase shift keying (BPSK) or QAM constellations. H.261/ H.263 acts as an inner code while channel coding acts as an outer code. Then, the stream can be transmitted over the wireless channels [15].

\[
r = hx + n; \quad (1)
\]
where \( n \) represents the AWGN samples with zero mean and unit variance and \( h \) is the channel fading coefficients of the physical propagation paths with zero mean and unit variance. For a simple AWGN channel without fading, the received signal is represented as \([12,15,16]\).

\[
r = x+n;
\]

(2)

In the receiver side, the inverse operation occurs. The received signal \( r \) is demodulated and converted from waveforms to digital signal. The BPSK demodulation restore the data and change the information from \((-1, 1)\) to \((0, 1)\). In figure 1 (b), Viterbi decoder is utilized to correct transmission errors using the protection bits inserted by the convolutional encoder and reconstruct the original form of bit stream. Then, bit streams are sent into H.261/H.263 decoder to reconstruct the transmitted video. A short description of the proposed video transmission system is listed below:

**Step 1:** Divide the uncompressed video into no of frames.

**Step 2:** Source encoding is used to transform the video streaming into its corresponding information (2-D) bits and reduce redundancy using H.261 with Quintet searching algorithm.

**Step 3:** Encode the compressed information 1-D bits using convolutional encoder with specified generator matrix and bit rate \((R)\).

**Step 4:** The encoded information bits are then modulated to convert the binary bits \((0, 1)\) into \((-1, 1)\).

**Step 5:** Introduce noise to simulate wireless channel errors using equation (2). The signals are assumed transmitted over an AWGN channel.

**Step 6:** The reverse operations (demodulation, channel and source decoding) are performed at the receiver side, to decode the received sequences of information bits and retrieve the transmitted video streaming.

**Step 7:** Count the number of bits errors by comparing the decoded bit sequences with the original one, and calculate the bit error rate (BER) as a function of signal to noise ratio (SNR).

### 2.1 Convolutional Encoding/ Decoding

The application of error correction codes is suitable where retransmissions are relatively costly or impossible. In general, channel coding is a model of introducing the controlled redundancy into transmitted binary data stream in order to improve the reliability of transmission. In convolution encoder, \( k \) data bits may be shifted into the register at once, and \( n \) code bits generated. The sliding window moved block through block along the input stream, and then can determine the output of the encoder.

**Figure 2: Code Structure of General Convolutional Codes.**

However, the blocks include \( k \) symbols, then output symbols are not only created from the current block but also from the previous \( v \) 1 blocks. The \( v \) which is known as the constraint length of this code is actually 1 plus the length of the longest shift register delay line in the encoder. The convolutional code \( R = k/n \) could be described as a \((n, k, v)\) code in figure 2, with the generator polynomial \( g \) that distinguishes the encoder connections. The used Viterbi decoding algorithm in receiver side, has the ability to detect and correct a limited number of errors without needing a reverse channel to request retransmission of data [17].

### 2.2 Source Video Coding

The input video has lots of redundancy i.e. spatial redundancy and temporal redundancy. Video compression is a practical implementation of source coding that uses coding techniques to reduce such redundancy. Most compression algorithms combine temporal motion compensation and spatial image compression, and the inverse process is called decompress/decoding. The ME scheme with
BM algorithm is typically used to find the better motion vector (MV) and reduce temporal redundancy. The approach is successfully adopted by video coding standards such as H.261, H.263, MPEG-1, MPEG-2 and MPEG-4 [9]. Fast BM algorithms have been proposed such as three-step search (TSS), new three-step search (NTSS), four-step search (4SS), and diamond Search (DS), to reduce the computational complexity [14-16]. The Quintet search (QS) algorithm proposed in [18] is used in this work. It is similar to the 3SS with five searching points in the beginning and different searching window. The video streaming converts into number of frames with RGB models, which is generally require more space than the other color models. Thus, these frames are converted into YCbCr color space, where Y is the luminance component and CbCr are the chrominance component. The algorithm flowchart is shown in figure 3, and it operates on groups of neighboring pixels or macroblocks with (16 × 16). These blocks of pixels are compared from one frame to the next, and the codec sends only the differences within those pixels and develops a motion vector (MV), which specifies the movement of a macroblock from one location to another. Figure 4 represents the searching process. In video streaming with more motion, the algorithm should encode more data to keep up with the larger number of pixels that are changing [16]. Then, the compression of various processes is done and gets the CR, and PSNR. In the absence of match, the search process stops. The steps of QS algorithm can be listed as:

**Step 1:** Input the original uncompressed video and divide it into no of frames.
**Step 2:** Arrange the video streaming in groups of frames and select reference frame and the next frame.
**Step 3:** Apply BM motion estimation algorithm with QS algorithm.
**Step 4:** Start the search location with center (0, 0), and select parameter p= 7, step size S = 4.
**Step 5:** Select the searching range (8 locations +/- S pixels.
**Step 6:** Choose a location from (5) locations depending on the min cost.
**Step 7:** Calculate the new step size, S = S/2.

**Step 8:** If (S=1) occur then stop search. If not points and continue the process until the desired result obtains.
**Step 9:** Find motion vector (MV) and apply coding MV.
**Step 10:** Finally, evaluate CR and PSNR.

![Figure 3: The Block diagram of the QS algorithm in ME Technique](image-url)
3. SIMULATION RESULT AND DISCUSSION

3.1 Simulation Environment & Encoder Parameters

The three uncompressed video streams of (144 × 176) frame resolution in figure 5 have been used in the work. Two videos are standards and one video is a camera recording within the room and with specifications dedicated to video conferencing. The video sequences include a wide range of low to high motion activities. The simulation is carried out by changing the source encoding parameters, which have greater impact on the encoded bit rate and video quality. In this work, 20 frames of each video sequence in YCbCr format are taken to perform simulations.

3.2 Performance Evaluation of H.261/Video Compression

In this section, different settings of the H.261 encoder parameters are applied and the performance of encoding algorithm using CR, PSNR and processing times is calculated to select the proper settings that can be applied to the input video streams being transmitted over wireless links.

The encoder parameters that have higher influence on video quality are selected as encoder controlling parameters. These are: the group of pictures (GOP) length and the quantization parameter (QP) of I, P and B pictures. The GOP contributes both in data compression and quality enhancement, while the quantization parameters manage the bit rate and the decoded video quality. Further, the QP specifies the number of quantization steps and has significant control over the CR. In this section, the GOP for various video characteristics are investigated in Tables 1-3, where few frames in a group are selected as I-picture to get higher coding efficiency. If the GOP length is high, selection of QP of I-pictures does not provide significant contribution on CR. However, the video quality depends on the appropriate selection of such parameters. Further, the transmitted data rate can be reduced by selecting appropriate video compression standards.

Table 1: Source Encoding Performance Of Forman Video Streaming.
3.3 Comparison with Other Algorithms

The results of BM for ME with QS algorithm are also compared with the technique using other different searching algorithms, such as seven search and Hexagonal search as in Table 4. The proposed algorithm was better than others in the average CR and PSNR values. Further, the proposed algorithm outperforms the others in terms of quality of the output frames.

Table 4: Performance Comparison With Other Algorithms Of Akiyo Video Streaming.

<table>
<thead>
<tr>
<th>Frames Number</th>
<th>Quintet search</th>
<th>Seven search</th>
<th>Hexagonal Search</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GOP = 1 Threshold 4</td>
<td>GOP = 1 Threshold 4</td>
<td>GOP = 1 Threshold 4</td>
</tr>
<tr>
<td></td>
<td>CR</td>
<td>PSNR</td>
<td>Coding</td>
</tr>
<tr>
<td>1</td>
<td>114.20</td>
<td>22.81</td>
<td>114.20</td>
</tr>
<tr>
<td>2</td>
<td>115.10</td>
<td>22.82</td>
<td>115.10</td>
</tr>
<tr>
<td>3</td>
<td>116.10</td>
<td>22.83</td>
<td>116.10</td>
</tr>
<tr>
<td>4</td>
<td>117.10</td>
<td>22.84</td>
<td>117.10</td>
</tr>
<tr>
<td>5</td>
<td>118.10</td>
<td>22.85</td>
<td>118.10</td>
</tr>
<tr>
<td>6</td>
<td>119.10</td>
<td>22.86</td>
<td>119.10</td>
</tr>
<tr>
<td>7</td>
<td>120.10</td>
<td>22.87</td>
<td>120.10</td>
</tr>
<tr>
<td>8</td>
<td>121.10</td>
<td>22.88</td>
<td>121.10</td>
</tr>
<tr>
<td>9</td>
<td>122.10</td>
<td>22.89</td>
<td>122.10</td>
</tr>
<tr>
<td>10</td>
<td>123.10</td>
<td>22.90</td>
<td>123.10</td>
</tr>
<tr>
<td>11</td>
<td>124.10</td>
<td>22.91</td>
<td>124.10</td>
</tr>
<tr>
<td>12</td>
<td>125.10</td>
<td>22.92</td>
<td>125.10</td>
</tr>
<tr>
<td>13</td>
<td>126.10</td>
<td>22.93</td>
<td>126.10</td>
</tr>
<tr>
<td>14</td>
<td>127.10</td>
<td>22.94</td>
<td>127.10</td>
</tr>
<tr>
<td>15</td>
<td>128.10</td>
<td>22.95</td>
<td>128.10</td>
</tr>
<tr>
<td>16</td>
<td>129.10</td>
<td>22.96</td>
<td>129.10</td>
</tr>
<tr>
<td>17</td>
<td>130.10</td>
<td>22.97</td>
<td>130.10</td>
</tr>
<tr>
<td>18</td>
<td>131.10</td>
<td>22.98</td>
<td>131.10</td>
</tr>
<tr>
<td>19</td>
<td>132.10</td>
<td>22.99</td>
<td>132.10</td>
</tr>
<tr>
<td>20</td>
<td>133.10</td>
<td>23.00</td>
<td>133.10</td>
</tr>
</tbody>
</table>

3.3 Performance Evaluation using AWGN Channel

For three sequences, the bit streams of encoded frames are transmitted assuming the ideal synchronization between the receiver and the transmitter. BPSK modulation with less power is used and the noise environment is considered AWGN channels. Simulations have been conducted for different signal to noise ratio (SNR), and BER characterizes the associated performance degradation of the transmission systems.
The performance of coded system is compared with the respective uncoded system under the same channel. Figures 6-8 show the performance of video transmission system of the three different (Forman, Akiyo, Mohand) videos. In these figures, it is seen that, the BER is inversely proportional to the SNR values. AWGN gives a desirable gain for higher values of SNR. High video quality at the receiving end can also be achieved by reducing the BER of the transmission channel.

To improve the performance of this system, CC with $R = 1/2$ and a constraint length of 7 is used to achieve desire BER and to transmit digital videos through wireless channels. In all simulation, it is observed that the coded system demonstrates better BER in contrast to others.

4. CONCLUSIONS

In this paper, video transmission systems are proposed using source and channel coding over wireless channels. The proposed systems were tested using three different streaming videos and two different case studies (video conferencing and mobile networks). The simulation results of ME with QS algorithm are computed and were better than other algorithms in terms of CR and PSNR. Further, the video quality of such source algorithm depends on the appropriate selection of the GOP length and QP. The results have also shown that the compressed video stream is more susceptible to channel errors, the video quality and BER value are inversely related to the SNR values. Thus, CC is used along with videos to minimize the effect of channel errors, and the performance is improved using CC in contrast to uncoded systems. However, future work combining source and channel coding with multiuser structures may help addressing the limitations of such techniques.
REFERENCES:


[18] Saeed A. Gathban, Salah A. Aliessiwi and Salah S. Mustafa , "Block Matching for Motion Estimation using a new Search Algorithm", accepted for publication in Journal of Engineering and Applied Sciences, 2017