CHAPTER 4

INTRAFRAME COMPRESSION

4.1 INTRODUCTION

In recent years, interest in multimedia has generated a lot of research in the area of video coding in academia and industry. Compression is essential for transmission of digital video over band-limited network and for easy storage. Video compression, a mere combination of image compression and motion compensation deals with elimination of spatial and temporal redundancy from the given video sequence, thereby reducing the storage and bandwidth constraints. The efficiency of compression techniques depends on the compression ratio, fidelity measures and its suitability for real time multimedia applications. Among the two salient intraframe and inter frame encoding scheme, this chapter discusses the performance evaluation of intraframe video compression using structural similarity index SSIM, a metric which exploits the human psycho visual system. SSIM is introduced for three different operations, namely shot detection and key frame extraction, spatial redundancy elimination and also as a fidelity evaluation measure. The proposed work operates by introducing a threshold to control the SSIM index value followed by entropy encoding, such as adaptive Huffman coding and Lempel Ziv Welch (LZW) coding.

The performance of the suggested scheme is evaluated with standard benchmark video sequences using compression ratio, peak signal to
noise ratio (PSNR) and structural similarity index measure (SSIM). The next section gives an overview of intraframe compression.

4.2 INTRAFRAME COMPRESSION OVERVIEW

A video signal basically comprises 3-dimensional array of color pixels with 2-dimension treated as spatial direction of moving pictures and one dimension considered as time domain. A video sequence can be assumed as a series of still images. It is the human eye that visualizes or interprets the frame as motion rather than as still information. Compression of such orderly arranged contiguous frames of video is essential as the amount of data to be stored or transmitted is enormous leading to memory/bandwidth constraints. Intraframe compression, also referred to as still image compression, is achieved by reducing the correlation among the pixels within a frame. One way of decorrelating the pixels within a frame is by means of transform coding. For many years the DCT has represented the state of the art in still image coding. JPEG is the standard that has incorporated this technology (Pennebaker and Mitchell 1992). JPEG 2000 standard uses the discrete wavelet transform (Taubman and Marcellin 2002; Ping Sing Tsai and Ricardo Suzuki 2008). The MPEG standard involves the application of DCT followed by quantization and run length encoding to eliminate the spatial redundancy. Further, from the existing literature, it is clear that wavelet and fractal based compressions are also used to eliminate the spatial redundancy (Jin Li and Jay Kuo 1999; Didier Le Gall 1991). In this dissertation, an alternative procedure that possesses the characteristics of human visual system is introduced with zero level quantization, so that the compression ratio can be improved.

4.3 THE PROPOSED COMPRESSION SCHEME

Structural similarity index is an objective quality metric applied for measuring and quantifying the effect of structural degradation between the
original and the reconstructed image signal (Wang 2004). The main functionality of the human visual system is to extract the structural details from the viewing field and SSIM exploits this by exhibiting the properties such as luminance, contrast and structure (Wang 2002). The efficiency of video compression can be improved if the properties of SSIM are applied to determine the spatial redundancy within the frame, apart from using it as a fidelity measure.

4.3.1 SSIM and LZW

Key frames extracted are considered as intraframes and these frames are divided into macro blocks of specific size which are merely square shaped groups of neighboring pixels. Then taking the first macro block as reference, SSIM index is calculated between the reference block and the remaining blocks in the frame. When the SSIM value is equal to one, it implies the two blocks compared are exactly similar or identical. Redundant blocks are identified based on a threshold selection and the index values are compared with the selected threshold. Blocks with SSIM above threshold are treated as similar blocks and instead of transmitting all similar blocks, only

![SSIM based Intraframe compression diagram](image-url)
one block is transmitted or stored after entropy encoding. For the similar blocks that are identified, only the block indices are transmitted. This results in bit saving as only few bits are needed to represent the block indices compared with the bit requirement for the block. Then the frame with similar blocks discarded is considered and once again SSIM is calculated between the blocks. Finally, the reduced frame with only dissimilar blocks are transmitted or stored along with block index. In this way, spatial redundancy is exploited to achieve compression. A suitable entropy encoding scheme like Adaptive Huffman coding (AHC) and Lempel-Ziv-Welch (Horspool 1991) encoding is applied to further encode the block for obtaining compression efficiency (Ziv and Lempel 1977; Ziv and Lempel 1978).

4.3.2 SSIM without entropy coding

In this section the key frames extracted are compressed using the SSIM concept alone without applying any entropy coding technique. The key frame is divided into blocks and SSIM index is calculated between the first block and the remaining blocks in the frame. The calculated SSIM values are compared with a threshold value and the similar blocks are identified. The index position of the similar blocks is determined and only these index position are retained and the similar blocks are discarded. The I frame with similar blocks discarded are considered and once again SSIM is calculated between the blocks and the same procedure is repeated. The reduced frame with only dissimilar blocks are transmitted along with block index.

4.3.3 Compression without SSIM and with Entropy Coding

In this section the different methods to achieve intraframe compression without using SSIM and using entropy coding techniques like LZW and adaptive Huffman coding are explained.
• Key frames are extracted from the video clipping.

• Case 1: Key frame is converted into data file and it is compressed using the LZW Code.

• Case 2: Key frame is converted into data file and it is compressed using Adaptive Huffman code.

• Case 3: Key frame is divided into blocks. For each block discrete cosine transform is taken. Then the transform coefficients are quantized and entropy coded.

4.4 PERFORMANCE EVALUATION

Experiments are performed using standard video sequences which are used as benchmark data for video compression problems. To evaluate the performance of the proposed scheme, compression and fidelity measures are used. Compression Ratio (CR) and bit rate are considered as compression evaluation parameters and PSNR and SSIM are considered as fidelity measures. By varying the block size and threshold the compression fidelity measures are evaluated and tabulated in the Table 4.1.

\[
Compression \ Ratio = 1 - \frac{Compressed \ Image \ Size}{Original \ Image \ Size}
\]

Table 4.1 shows the intraframe compression performance for Claire sequence by varying the threshold and the block size. When the threshold is lowered, compression ratio is improved with slight degradation in the quality of the recovered key frame and this is clear from the values of CR, PSNR and SSIM. Moreover, an increase in block size results in the reduction of compression ratio and with an increase in fidelity measure in both the cases i.e. SSIM combined with AHC and SSIM combined with LZW coding.
TABLE 4.1 Intraframe compression – Performance measure for Claire sequence

<table>
<thead>
<tr>
<th>Video</th>
<th>Threshold</th>
<th>Block size</th>
<th>SSIM + AHC</th>
<th>SSIM + LZW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>CR</td>
<td>mse</td>
</tr>
<tr>
<td>Claire</td>
<td>0.95</td>
<td>8 × 8</td>
<td>83.3</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>8 × 8</td>
<td>85.5</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>0.85</td>
<td>8 × 8</td>
<td>86.5</td>
<td>61.63</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>8 × 8</td>
<td>88.89</td>
<td>67.6</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>16×16</td>
<td>81</td>
<td>41.8</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>16×16</td>
<td>83.3</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>0.85</td>
<td>6×16</td>
<td>83.35</td>
<td>50.3</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>6×16</td>
<td>83.75</td>
<td>51</td>
</tr>
</tbody>
</table>

Figure 4.2 Intraframe compression for Coastguard sequence
Quality of the reconstructed key frame is evaluated using two objective quality measures viz. SSIM and PSNR. Figures 4.2 to 4.4 show the variation of SSIM calculated between the original key frame and the decompressed key frame for different threshold values for the sequences carphone, coast guard and flower. At lower threshold, SSIM value is low. This is because, as the threshold is lowered more number of blocks become
similar blocks and all these similar blocks are replaced by a single block, thereby resulting in compression. When the block size is $8 \times 8$, SSIM value is lower for all the sequences and an increase in block size results in better quality and this is clear from the higher value of SSIM.

Figure 4.5 Rate distortion curve for different video sequences (Block Size 8×8)
Figure 4.5 (Continued)
Figure 4.5 (Continued)

Figure 4.5 shows the bit rate and PSNR obtained as a result of compressing the key frames of different video sequences like Coast guard, Flower, Claire, Mobile, Mother-daughter and Carphone. For the standard video clippings considered, PSNR value is higher when the bit rate is high. High bit rate implies that the compression ratio is low. When higher
compression is achieved, bit rate reduces with a reduction in quality. This is because higher compression is possible only when the threshold is reduced. At reduced threshold value, SSIM index of many blocks exceeds the threshold values and these blocks with higher SSIM values are treated as similar or redundant blocks. These similar blocks are replaced by a single block. Thus compression is achieved with reduced quality.

Figure 4.6 Rate distortion curve for different video sequences. (Block Size 16×16)
Figure 4.6 (Continued)
Figure 4.6 (Continued)

Figure 4.6 shows the bit rate and PSNR obtained as a result of compressing the key frames of different video sequences with block size $16 \times 16$. It is observed from Figure 4.6 that the PSNR value increases with increased bit rate. High bit rate implies that the compression ratio is low.
Figure 4.7  Threshold versus compression efficiency (Mother-daughter video sequence) using SSIM plus adaptive Huffman coding

Figure 4.7 shows the variation of compression efficiency for different threshold values for the video sequence Mother-daughter using the SSIM concept. Adaptive Huffman coding is used to encode the dissimilar blocks in the key frame. Compression efficiency is low when the threshold is high and when the threshold is reduced; compression is improved irrespective of the block size.
Figure 4.8  Threshold versus compression efficiency (Mother-daughter video sequence) using SSIM plus LZW coding

Figure 4.8 shows the variation of compression efficiency for different threshold values for the video sequence Mother-daughter using the SSIM along with LZW coding. Here also, the compression efficiency is low when the threshold is high and it is improved when the threshold is reduced for a block size of 8 and 16. Moreover, the compression provided by this concept is higher compared with SSIM plus adaptive Huffman coding for the selected values of the threshold.
Figure 4.9  Performance evaluation of key frame compression using SSIM and LZW coding
Figure 4.9 (Continued)

Figure 4.9 shows the compression efficiency achieved in compressing the key frame of different video sequences for various values of the threshold. It is evident from the results that a lower threshold results in better compression efficiency.

PSNR is used as a fidelity measure to evaluate the quality of the reconstructed key frame. Figure 4.10 shows the variation of PSNR for different threshold.
Figure 4.10 Intraframe compression result - PSNR versus Threshold
Figure 4.10 (Continued)
Figure 4.10 (Continued)

It is evident from Figure 4.10 that the PSNR value is high for a higher threshold and as threshold is reduced, compression efficiency is improved with loss in quality.
Figure 4.11 Compression efficiency for different video sequence using SSIM and adaptive Huffman coding

Figure 4.12 Compression efficiency for different video sequence using SSIM and LZW

From Figures 4.11 and 4.12 it is evident that compression efficiency is better for the proposed method SSIM along with LZW than
SSIM with adaptive Huffman coding for the different sequences considered. Moreover, from the results it is clear that the compression efficiency is improved as the threshold is reduced.

**Figure 4.13 Intraframe compression – PSNR for different video sequences**

**Figure 4.14 Intraframe compression – SSIM for different video sequences**
Figures 4.13 and 4.14 show that both fidelity measures PSNR and SSIM are high for higher value of threshold and as the threshold is reduced, some loss in quality occurs with improvement in compression efficiency.

![Original Image](image1)

![Reconstructed Image](image2)

Figure 4.15 First key frame of Claire video sequence (SSIM along with LZW)

Figure 4.15 shows the original key frame and the reconstructed key frame of the Claire video sequence using SSIM and LZW with block size 8x8 and threshold 0.8
Figure 4.16 First key frame of Claire video sequence (SSIM along with adaptive Huffman coding)

Figure 4.16 shows the original key frame and the reconstructed key frame of the Claire video sequence using SSIM and adaptive Huffman coding with block size of $8 \times 8$ and threshold value equal to 0.8
Figure 4.17 Original key frame of Coastguard sequence

Figure 4.18 Reconstructed key frame - Coastguard sequence
(T=0.85 Block size = 8)
Figure 4.17 shows the original key frame of the Coastguard sequence and it is observed from Figure 4.18 and 4.19 the reconstructed key frame of the coast guard sequence are better in quality.

4.5 CONCLUSION

This chapter proposes a method to compress the key frames extracted from the video sequence. Several standard video clippings like Claire, Carphone, Mobile, Flower are used to test the performance of the proposed concept. Evaluations are made by varying the threshold and block size. Performance measures like PSNR, SSIM, bit rate and compression efficiency are used to evaluate the performance of the proposed method. From the results obtained, it is clear that combining SSIM with LZW produces a better compression ratio with some loss or degradation in visual quality. Objective quality measures like PSNR and SSIM take reasonable values. Moreover, the results of SSIM along with LZW are better compared with the results of SSIM added with adaptive Huffman in terms of quality and compression ratio. Inter frame compression schemes are explained in the next chapter.