International Journal of Electronics and Communication Engineering (IJECE)
ISSN(P): 2278-9901; ISSN(E): 2278-991X
Vol. 5, Issue 3, Mar - Apr 2016, 1-10

International Academy of Science, Engineering and Technology
Connecting Researchers; Nurturing Innovations

INTER-INTRA FRAME CODING IN MOTION PICTURE COMPENSATION USING NEW WAVELET BI-ORTHOGONAL COEFFICIENTS

MANIMOZHI DEVARAJ

Assistant Professor, Department of Electronics and Communication Engineering,

Arunai Engineering College, Thiruvannamalai, Tamil Nadu, India

ABSTRACT

Video compression has become one of the basic technologies of the multimedia age. In many applications, such as the design of multimedia workstations and high quality transmission and storage, the goal is to achieve transparent coding of Image and video at the lowest possible data rates. In other words, bandwidth cost money, therefore, the transmission and storage of information becomes costly. However, if we can use less data, both transmission and storage become cheaper. In this paper two techniques are used together to achieve high compression rate. In video frames, seam curving technique is used as Intra frame coding and DWT is used as Inter frame coding. Both Inter-Intra frame coding are used to achieve desired result.

KEYWORDS: Video Compression, Inter Frame Coding, Intra Frame Coding, Seam Curving, DWT, SPHI

INTRODUCTION

Video compression plays an important role in modern multimedia applications. The idea behind compression is to save time and the number of bits sent between images by taking the difference between them instead of sending each frame again. With video streaming and storage becoming so popular this is a very useful tool to have. Compression can be lossy or lossless. Lossless compression means that that when the data is decompressed, the result is a bit-for-bit perfect match with the original. While lossless compression of video is possible, it is rarely used, as lossy compression results in far higher compression ratios at an acceptable level of quality.

One common method for video or image compression is discrete Wavelet transform. DWT is a lossy compression algorithm that samples video frames at regular intervals, analyzes the frequency components present in the sample, and discards those frequencies which do not affect the image as the human eye perceives it. However, if the video is over compressed in a lossy manner, visible (and sometimes distracting) artifacts can appear. In this paper we propose an approach for video compression based on the new technique using DWT for inter coding and Seam curving for intra coding.

INTRA CODING USING SEAMLESS CARVING TECHNIQUE

Image compression is important for many applications that involve huge data storage, transmission and retrieval such as for multimedia, documents, videoconferencing, and medical imaging. Uncompressed images require considerable storage capacity and transmission bandwidth. The objective of image compression technique is to reduce redundancy of the image data in order to be able to store or transmit data in an efficient form. This results in the reduction of file size and allows more images to be stored in a given amount of disk or memory space.

The diversity and versatility of display devices today imposes new demands on digital media. For instance, designers must create different alternatives for web-content and design different layouts for different devices. Moreover, HTML, as well as other standards, can support dynamic changes of page layout and text. Nevertheless, up to date, images, although being one of the key elements in digital media, typically remain rigid in size and cannot deform to fit different layouts automatically. Other cases in which the size, or aspect ratio of an image must change, are to fit into different displays such as cell phones or PDAs, or to print on a given paper size or resolution.

Standard image scaling is not sufficient since it is oblivious to the image content and typically can be applied only uniformly. Cropping is limited since it can only remove pixels from the image periphery. More effective resizing can only be achieved by considering the image content and not only geometric constraints.

Seam-carving, can change the size of an image by gracefully carving-out or inserting pixels in different parts of the image. Seam carving uses an energy function defining the importance of pixels. A seam is a connected path of low energy pixels crossing the image from top to bottom, or from left to right. By successively removing or inserting seams we can reduce, as well as enlarge, the size of an image in both directions. For image reduction, seam selection ensures that while preserving the image structure, we remove more of the low energy pixels and fewer of the high energy ones. For image enlarging, the order of seam insertion ensures a balance between the original image content and the artificially inserted pixels. These operators produce, in effect, a content-aware resizing of images.

Seam carving can support several types of energy functions such as gradient magnitude, entropy, visual saliency, eye-gaze movement, and more. The removal or insertion processes are parameter free; however, to allow interactive control, we also provide a scribble based user interface for adding weights to the energy of an image and guide the desired results. This tool can also be used for authoring multi-size images.

The key insight is the realization that most texture artifacts can be eliminated through local image-space translations. The result is one texture for the whole object which minimizes visual artifacts. This simple strategy proves remarkably effective.

The process allows the user to resize an image by removing a continuous path of pixels (a seam) vertically or horizontally from a given image. A seam is defined as a continuous path of pixels running from the top to the bottom of an image in the case of a vertical seam, while a horizontal seam is a continuous line of pixels spanning from left to right in an image.

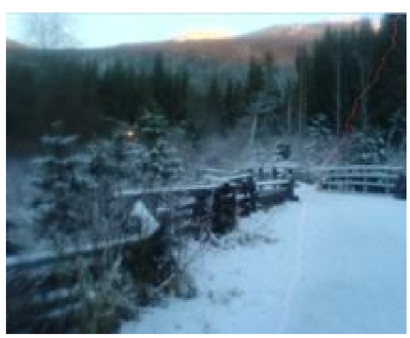

Figure 1: Image with Vertical Seam

Algorithm Implementation

The first step in calculating a seam for removal or insertion involves calculating the gradient image for the original image. The gradient image is a common image that is used in both horizontal and vertical seam calculation, and can be calculated either from the luminance channel of a HSV image, or calculated for each of the R, G, and B channels, then averaging the three gradient images. Figure 2 is included as an example gradient image. The Sobel operator was chosen for calculation of the gradient image in this project, but other gradient operators may be used.

Once the gradient image is calculated, the next step is to calculate the energy map image. The energy map image needs to be calculated separately for either vertical (Figure 3) or horizontal (Figure 4) seams, and also needs to be recalculated after every seam removal. It is calculated by the following process for the vertical seam case (a horizontal energy image can be calculated using the same function, where the input image is transposed): for each pixel (i,j) in the gradient image (see Table 1), the value at (i,j) in the energy map is the sum of the current value at (i,j) from the gradient image and the minimum of the three neighboring pixels in the previous row, i.e. min((i-1,j-1),(i-1,j),(i-1,j+1)), from the energy map. For i=1 (the initial row), the values in the energy map image are set to those in the gradient image, and for when the pixel (i,j) is along the edge of an image, only (i-1,j) and either (i-1,j-1) or (i-1,j+1) are used depending on if (i,j) is on the right or left edges, respectively.

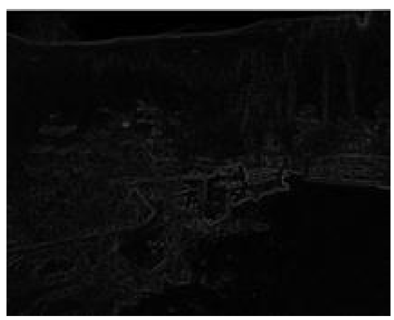

Figure 2: Gradient Image

Table 1: Pixel Indices

(i-1,j-1)	(i-1,j)	(i-1,j+1)
(i,j-1)	(i, j)	(i,j+1)
(i+1,j-1)	(i+1,j)	(i+1,j+1)

Figure 3: Vertical Seam Energy Map

Figure 4: Horizontal Seam Energy Map

Once the energy map is calculated, the method to find the optimal seam is to first find the minimum value in the last row (which becomes the (i,j)'th pixel), saving the pixel location for use in removal, then working backwards by finding the minimum of the 3 neighboring pixels of (i,j) in the (i-1)'th row and saving that pixel to the seam path. This process is repeated until the first row is reached, and results in the optimal seam, an example of which is shown in Figure 6.

Figure 5: Energy Map with Vertical Seam

After the optimal seam is found, the path of pixels that make up the seam are removed from both the gradient image and the original RGB image, and the remaining pixels are shifted right or up to form a continuous image.

The process can be repeated to remove a set of seams, horizontally or vertically and will result in an image with reduced dimensions, but with the overall scene content intact. An example of this is included as Figure 6, where the image was resized to 320x240 pixels, from 640x480 pixels and as can be seen, the resulting image will have artifacts if a large number of seams are removed.

Figure 6: Resized Image

For the case of seam insertion (increasing the image size), a seam can be calculated along a given direction, and the average of the two neighboring pixels along the seam can be inserted. If the desired image size is to be increased by N pixels in a given direction, the computation of the first N seems to be removed along that direction must first be completed, and then averaged pixels are inserted along each successive seam, hence the limitation on the maximum increase in image size in my implementation noted earlier in the features and functionality section. This method of calculating N seams is used to avoid inserting pixels along the same seam repeatedly.

INTER FRAME CODING USING SEAMLESS CARVING TECHNIQUE

To circumvent this problem, a series of techniques – called picture and video compression techniques – have been derived to reduce this high bit-rate. Their ability to perform this task is quantified by the compression ratio. The higher the compression ratio is, the smaller is the bandwidth consumption. However, there is a price to pay for this compression: increasing compression causes an increasing degradation of the video.

Since last two decade the discrete wavelet transform (DWT) has witnessed great success for Video compression. All those DWT based method whether they are conventional and directional, use wavelet 9/7 filter or wavelet 5/3 for better compression. This paper introduces new wavelet based bi-orthogonal filter coefficient that can give better result in case PSNR and MSE comparison to wavelet 9/7 filter and wavelet 5/3 filter.

Digital Representation of Video Signals

Colors are synthesized by combining the three primary colors red, blue, and green (RGB). The RGB color system is one means of representing color images. Even, the luminance (brightness) and chrominance (color) information can be represented separately. We can obtain the luminance signal Y which represents the “brightness” of the color, By calculating a weighted sum of the three colors R, G, and B. All three above mentioned systems use three components: luminance Y, blue color difference U (equivalent to Cb above) and red color difference V (equivalent to Cr above) to represent a color. This is called as YUV system.

The YUV representation system has certain advantages over the RGB system. As the human visual system (HVS) is less sensitive to chrominance than to brightness, the chrominance signals can therefore be represented with a lower resolution than the luminance without significantly affecting the visual quality. This by itself achieves some degree of data compression.

Motion Compensation

An MPEG video is a sequence of frames. As two successive frames of a video sequence often have small differences (except in scene changes), the MPEG-standard offers a way of reducing this temporal redundancy. There are three types of frames

I-frames (intra)

P-frames (predicted) and

B-frames (bidirectional)

The I-frames are the “key-frames”, which have no reference to other frames and their compression is not that high. P-frames can be predicted from an earlier I-frame or P-frame. Thus P-frames cannot be reconstructed without their referencing frame, but they need less space than the I-frames, because only the differences are stored. The B-frames are a two directional version of the P-frame, referring to both directions (one forward frame and one backward frame). B-frames cannot be referenced by other P- or B frames, because they are interpolated from forward and backward frames. P-frames and B-frames are called inter coded frames, whereas I-frames are known as intra coded frames.

Motion estimation or Motion compensation is a technique to realize the references between the different types of frames and the correlation between two frames in terms of motion is represented by a motion vector. The resulting frame correlation, and therefore the pixel arithmetic difference, strongly depends on how good the motion estimation algorithm is implemented.

Wavelet Video Coding

The wavelet transform decomposes a video frame into a set of sub-frames with different resolutions corresponding to different frequency bands. These multi resolution frames also provide a representation of a global motion structure in the scene at different scales. Wavelet transforms involve representing a general function in terms of simple, fixed building blocks at different scales and positions. The discrete wavelet transform (DWT) has gained wide popularity due to its excellent de correlation property, many modern images and video compression systems embody the DWT as the transform stage .After DWT was introduced, several codec algorithms were proposed to compress the transform coefficients as much as possible. Among them, stationary Wavelet Transform (SWT) and Set Partitioning in Hierarchical Trees (SPIHT) are the most famous ones.

The DWT and IDWT are the most computationally intensive and time critical portions of the algorithm. The DWT uses 7-tap and 9-tap FIR filters. Motion estimation and compensation on spatial domain is used in wavelet video coding in order to exploit the spatial correlation present in the video sequences. A discrete wavelet transform (DWT) is applied to generate a set of wavelet coefficients for each sub band which is generally coded separately.

The relationship between DWT and sub bands and focus on low frequency sub bands:

$$DWT(Fsource(i, j)) = LL(Fsource(i, j)) + LH(Fsource(i, j)) + HL(Fsource(i, j)) + HH(Fsource(i,j))$$

Where L denotes a low pass filter function, H denotes a high pass filter function, Fsource represents an original input frame, and (i, j) is the block location on a frame.

Table 2: New Filter Coefficients

9/7 FILTER COEFFICIENT		PROPOSED FILTER COEFFICIENT	
LPF	HPF	LPF	HPF
0	0	-0.0015	0.0015
0.0378	-0.0645	0.0027	0.0027
-0.0238	0.0407	0.0049	-0.0049
-0.1106	0.4181	-0.0128	-0.0128
0.3774	-0.7885	-0.0025	0.0025
0.8527	0.4181	0.0264	0.0264
0.3774	0.0407	-0.0050	0.0050
-0.1106	-0.0645	-0.0455	-0.0455
-0.0238	0	0.0211	-0.0211
0.0378	0	0.0756	0.0756
		-0.0568	0.0568
		-0.1404	-0.1404
		0.1817	-0.1817
		0.6594	0.6594
		0.6594	-0.6594
		0.1817	0.1817
		-0.1404	0.1404
		-0.0568	-0.0568
		0.0756	-0.0756
		0.0211	0.0211
		-0.0455	0.0455
		-0.0050	-0.0050
		0.0264	-0.0264
		-0.0025	-0.0025
		-0.0128	-0.0128
		0.0049	0.0049
		0.0027	-0.0027
		-0.0015	-0.0015

Still images are considered as 2-D signals. Applying the sub band/wavelet transform to such signals is most commonly done by using the 1-D transform version and applying it to the still image in both row-order and column-order. This is because implementation of the single dimension transform is more efficient than an equivalent 2-D transform, and was shown [10] to be an effective solution. Video images are considered as a 3-D signal, the three dimensions being the horizontal, the vertical, and the temporal dimension. In this section, we will summarize the implementation of the 1-D wavelet transform which is also representative of the sub band transform in this context.

The wavelet transform, as a data decor relating tool, has won acceptance because of its multi resolution analysis capabilities in which the signal being transformed is analyzed at many different scales to give a transformation whose coefficients can efficiently describe fine details as well as global details in a systematic way. In addition to this the locality of the wavelet basis functions as opposed to the Fourier transform. Wavelets also unify the many other techniques that are of local type, such as the Gabor transform and the short time Fourier transform. The wavelet/sub band transform is implemented using a pair of filters: a high pass filter and a low pass filter, which split a signal's bandwidth in two halves. The frequency responses of and are mirror images. To reconstruct the original signal an inverse transform is implemented, using the inverse transform filters, which are also mirror images.

The bi-orthogonal 9/7 filter coefficient and the new proposed filter coefficient are shown in the above table.

Set Partitioning in Hierarchical Trees (SPIHT)

This compression schemes is based on wavelet coding technique. The image is transformed using a DWT. In the beginning, the image is decomposed into four sub-bands by cascading horizontal and vertical two-channel critically sampled filter-banks. This process of decomposition continues until some final scale is reached. In each scale there are three sub-bands and one lowest frequency sub-band. Then successive-approximation quantization (SAQ) is used toper form embedding coding. This particular configuration is also called QMF pyramid. The SPIHT algorithm is used to the multi-resolution pyramid after the sub band/wavelet transformation is performed.

These are the steps involved in compressing the video frames. The decoder does exactly opposite, that is it performs arithmetic decoding on the input bit stream. Initially the coded video is readed, and then it is decoded using SPHIT encoder. After that the decoded video is passed through the inverse DWT with the proposed filter coefficients. Now convert the video from YCbCr to RGB format. Measure the MSE and PSNR values. Compare them with the classic approach parameters. After reconstruction of Video parameters are measured as follows:-

PEAK SIGNAL TO NOISE RATIO (PSNR)

The phrase peak signal-to-noise ratio, often abbreviated PSNR, is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Because many signals have a very wide dynamic range, PSNR is usually expressed in terms of the logarithmic decibel scale.

$\sum$

And root mean square error is given by:-

RMSE= $\sqrt{MSE}$

Here

Pi=Original data

Qi= Reconstructed Data

K= Size of video

The peak signal to noise ratio for reconstructed image is given by

PSNR=20 $\log_{10}$ (max (p_i)/RMSE)

CONCLUSIONS AND FUTURE WORKS

We presented an approach for content-aware resizing of image frame Seams and compressing the frames of the video using SPHIT algorithm. By doing these experiments we conclude that both techniques have its' own advantage and disadvantage. The PSNR of the decompressed video tends to be 4dB.

Future research efforts focus on better PSNR and MSE value. Our future work includes applying this schema with low computational complexity. The future direction of this research is to implement a compression technique using neural network.

REFERENCES

1. Vimal Kumar, Surendra Kumar Agarwal, "Image Compression by using New Wavelet Bi- Orthogonal Filter Coefficients", 978-1-4673- 1318-6/12/$31.00 ©2012 IEEE
2. R. L. Claypoole, G. M. Davis, W. Sweldens, and R. G. Baraniuk, "Nonlinear wavelet transforms for image coding

via lifting," IEEE Trans. Image Process., vol. 12, no. 12, pp. 1449–1459, Dec. 2003

3. W. Ding, F. Wu, X. Wu, S. Li, and H. Li, "Adaptive directional lifting-based wavelet transform for image coding," IEEE Trans. Image Process., vol. 16, no. 2, pp. 416–427, Feb. 2007.
4. C.-L. Chang, A. Maleki, and B. Girod, "Adaptive wavelet transform for image compression via directional quincunx lifting," in Proc. IEEE Workshop Multimedia Signal Processing, Shanghai, China, Oct. 2005
5. Cohen, A., Daubechies, I., And Feauveau, .I.: 'Biorthogonal bases of compactly supported wavelets', Comm. Pure Appl. Math., 1992, 45, pp. 485-560
6. Daubechies, I., And Sweldens, w.: 'Factoring wavelet transforms into lifting steps', J. Fourier Anal. Appl., 1998, 4, (3), pp. 247-269.
7. Mark J. Shensa, The Discrete Wavelet Transform: Wedding the A Trous and Mallat Algorithms, IEEE Transaction on Signal Processing, Vol 40, No 10, Oct. 1992.
8. Olivier Rioul and Martin Vetterli, "Wavelets and Signal Processing", IEEE Trans. on Signal Processing, Vol. 8, Issue 4, pp. 14 - 38 October 1991
9. Boykov, Y., and Jolly, M.-P. 2001." Interactive graph cuts for optimal boundary & region segmentation of objects in n-d images"
10. K. Mallikarjuna, S. Ifthukar" Efficient Lossy Image Compression Technique Based on Seam Identification and SPIHT Coding", International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064.
11. D. Manimozhi and Dr. T. K. Shanthi, "Video Compression by Using New Wavelet Bi-Orthogonal Filter Coefficients", In IJERT, Volume. 3, Issue. 03, March – 2014