Drift Effect Reduction in Frequency Scalable Coding

Jong-Nam Kim, Yo-Sung Ho, Taehwan Shin and Tae-Sun Choi

Kwangju Institute of Science and Technology
Ssangam-Dong Kwangsan-Gu Kwangju, 506-712, Korea
jnkim@sipl.kjist.ac.kr

Abstract: This paper describes two different approaches for drift effect reduction in frequency scalable coding; one is based on frequency sampling, and the other is based on motion vector correction. The frequency sampling method preserves the same average pixel energy as the original image, while the motion vector correction method provides more accurate motion vectors. Their performances are compared with those of the conventional ones.

1. Introduction

For scalable image coding, we usually employ either pyramid or subband coding structures [1,9]. The frequency pyramid coding scheme may have one motion prediction loop in the higher layer, or two motion prediction loops both in the lower and the higher layers [8]. The former is simpler, but has drift artifacts in the lower layer; the latter has no drift effects, but it has more complex structure [6]. Here we define drift effects as visual artifacts caused by inaccurate processing of pixel values, such as coefficient scaling and motion compensation, in scalable layered coding. In the frequency pyramid coding with one motion prediction loop, drift effects occur partly due to DCT coefficient scaling and partly due to motion vector scaling in the lower layer [2,3,4].

In this paper, we propose two schemes for drift effect reduction and compare their performances with those of conventional frequency-domain pyramid coding schemes. After we analyze drift effects in a two-layer frequency pyramid coding structure, details of the two proposals and simulation results are described.

2. Drift Effect in Frequency Pyramid Coding

In scalable pyramid image coding schemes, we subsample the original image in the spatial domain or in the frequency domain to get subimages of smaller resolution [5,7]. In the spatial pyramid, subsampling is achieved by selecting a few representatives out of a group of pixels. In the frequency pyramid, however, subsampling is performed by taking the upper-left 4 x 4 block out of 8 x 8 DCT coefficients. Fig. 1 illustrates the subsampling operations in the spatial domain and in the DCT domain.

![Subsampling Operations](image)

Fig. 1: Subsampling Operations

Fig. 2 shows a typical frequency pyramid encoder. Because the predicted image in the lower layer is obtained by subsampling from the original image in the higher layer and motion vectors in the lower layer are scaled by half of those of the higher layer, this system may demonstrate drift effects. In other words, the condition for motion compensation in the lower layer is not the same as one in the higher layer. Therefore, in the continuous inter-frame coding, inaccurate motion errors can be accumulated. Reconstructed images with these errors in the lower layer may result in severe artifacts to human viewers.
3. New Methods for Drift Reduction

3.1 Frequency Sampling Method

Since signal energy is generally concentrated on low frequency coefficients in the DCT domain, direct sampling cannot be used without any scaling because of too much energy for the subsampled 4 x 4 low frequency block. Therefore, an appropriate scaling factor must be multiplied to the chosen frequency samples. In conventional coding schemes, the selected DCT coefficients are scaled down by half regardless of signal energy distribution in the DCT domain. In general, image complexity indicates how much signal energy is spread out in the frequency domain.

In this section, we propose a new method for DCT coefficient scaling according to signal energy distribution in the DCT domain. Because signal energy in the 8 x 8 pixel block is not compacted into the top-left 4 x 4 low frequency coefficients in the DCT domain, discarded signal energy by the frequency subsampling operation must be compensated by use of the scaling factor for selected coefficients. If we compensate for the discarded signal energy appropriately, the frequency sampled image should contain one quarter of the total signal energy in the original image regardless of the image complexity. In the conventional scheme, the scaling factor 0.5 for the selected DCT coefficients cannot provide one quarter of the signal energy during the subsampling process.

In order to state the proposed method precisely, let’s call the energy of the 8 x 8 block as the total energy (TE), and the energy of the 4 x 4 energy as the partial energy (PE). The scaling factor is determined as follows to preserve exactly a quarter of the total signal energy in the original 8 x 8 block:

\[ 0.5 \times \sqrt{(TE/PE)} \]

With this scaling factor, the lower layer image has exactly one quarter energy of the higher layer image. This means that the average energy per pixel in the frequency sampled image is the same as that of the original image. The average energy per pixel is affected by image complexity in the conventional schemes, but it is independent of image complexity in the proposed scaling.

In the DCT block, the DC component represents the average luminance value of the image block. This average luminance should not be changed after block sampling. The DC energy of the 4 x 4 block without any scaling would be 4 times that of the 4 x 4 block image obtained by lowpass filtering and decimation. Therefore, the DC component in the frequency sampled block must be scaled down by half, while the other coefficients must be scaled down by \( 0.5 \times \sqrt{(TE/PE)} \) to preserve exactly one quarter of the total signal energy in the sampled 4 x 4 subimage block, as illustrated in Fig. 3.

![Fig. 3: Adaptive DCT Coefficient Scaling](image)

We note that spectra of the conventional scaling method and of the low pass filtering method are slightly different. The two methods are similar in discarding high frequency coefficients for lowpass filtering or frequency block sampling. Fig. 4 illustrates the compensated frequency spectrum by the proposed scaling.

![Fig. 4: Spectrum of Proposed Coefficient Scaling](image)
3.2 Motion Vector Scaling

As described earlier, another source of drift effects is inaccurate motion vectors for the lower layer decoder due to improper scaling of motion vectors from the higher layer.

Let the accurate motion vector of the drift-free coding scheme be \( m_v \), and let the scaled motion vector of the coding scheme with drift effects be \( m_v' \). Then the error of the motion vector is \( m_v - m_v' \).

Fig. 5 shows a new coding scheme to reduce drift effects by proper motion vector scaling. In Fig. 5, 8 x 8 DCT coefficients are subsampled to form a 4 x 4 subblock. The subsampled coefficients are scaled before quantization. Then, the quantized coefficients are inversely processed to reconstruct an image in the lower layer by the inverse quantization and the IDCT operations. Subsequently, motion compensation is performed with the reconstructed image in the lower layer and the scaled motion vector from the higher layer. Among several scaling factors ranging from 0.3 to 0.6, one factor that maximizes the SNR value between the decimated image of the input and the motion compensated image is selected and transmitted. The inversely quantized data is padded with zeros to match the image resolution. Finally, the inverse scaling is performed.

![Fig. 5: Drift-Free Encoder with Motion Vector Scaling](image)

In the conventional coding scheme, motion vectors of the higher layer are obtained from the lower layer with a scaling factor of 0.5. However, for an accurate motion compensation, motion vector scaling is not proportional to the image resolution ratio. To solve this problem, many approaches have been attempted. The proposed scheme reduces the difference between the scaled motion vector and the accurate motion vector by using additional blocks. This scheme is simpler and more effective than the drift-free frequency pyramid coder, since the selected scaling factor for the given range is transmitted in this scheme, instead of the motion vector for the lower layer.

4. Simulation Results

The following three figures are simulation results. The images used for simulation are Claire images. First frame was coded in intra coding mode and the other frames were coded in inter coding mode.

![Fig. 6: PSNR Values (Conventional Coding Scheme)](image)

![Fig. 7: PSNR Values (Coefficient Scaling)](image)

![Fig. 8: PSNR Values (Motion Vector Scaling)](image)
The SNR performance in low resolution by simulation was marginally better for the conventional scaling method than one for the proposed coefficient scaling method.

We can easily find the cause of the above results in the frequency spectrum of the low pass filtering method, the conventional scaling method and the newly proposed methods. Referenced small resolution image to measure PSNR was made by removing high frequency components without which is general low pass filtering method. Therefore, general low pass filtering method is similar to the conventional scaling method.

To check the validity of the proposed scaling method, the simulation was done with following operation: DCT, frequency decimation with coefficient scaling of the respective method, IDCT and spatial interpolation using bilinear interpolation. Then, the SNR is calculated for the original image and for the interpolated images from the decimated images by the above scaling methods. In this simulation, the SNR of the proposed scaling method is higher than one of the conventional scaling method.

In the proposed motion vector correction method, performance of PSNR was better than the conventional scheme. Because of the additional motion compensation and the PSNR routine of the lower layer encoder, the complexity is increased slightly compared to the drift scheme. Also, there is a data increase due to the transmission of the scaling factor. However, the complexity and the data increase are much lower than the drift-free scheme [1,8].

5. Conclusions

In this paper, we proposed two methods for drift effect reduction in scalable frequency pyramid coding structures. One is a proposed scaling method for subsampled DCT coefficients, and the other is a modified scheme from the conventional drift scheme which has only one motion estimation, motion compensation loop in the higher layer encoder. In simulation results, the performance of the proposed methods was better than the conventional scheme. But the inevitable increase of complexity is resulted in. This amount of complexity is little compared with the free-drift scheme which uses two motion prediction loops.

References