

# Video Coding With Linear Compensation (VCLC)

Arif Mahmood

Zartash Afzal Uzmi

Sohaib Khan

School of Science and Engineering  
Lahore University of Management Sciences, Lahore, Pakistan  
{arifm, zartash, sohaib}@lums.edu.pk

**Abstract**—Block based motion compensation techniques are commonly used in video encoding to reduce the temporal redundancy of the signal. In these techniques, each block in a video frame is matched with another block in a previous frame. The match criteria normally used is the minimization of the sum of absolute differences (SAD). Traditional encoders take the difference of the current block and its best matching block, and this differential signal is used for further processing.

Rather than directly encoding the difference between the two blocks, we propose that the difference between the current block and its first order linear estimate from the best matching block should be used. This choice of using linear compensated differential signal is motivated by observing frequent brightness and contrast changes in real videos. We show two important theoretical results: (1) The variance of the linear compensated differential signal is always less than or equal to the variance of differential signal in traditional encoders. (2) The optimal criteria for finding the best matching block, in our proposed scheme, is the maximization of the magnitude of correlation coefficient.

The theoretical results are verified through experimentation on a large dataset taken from several commercial videos. For the same number of bits per pixel, our proposed scheme exhibits an improvement in peak signal to noise ratio (PSNR) of up to 5 dB when compared to the traditional encoding scheme.

## I. INTRODUCTION

A digital video signal consists of a sequence of frames and is usually characterized by strong temporal correlation between adjacent frames. This correlation is exploited in standard video codecs to achieve significant compression, resulting in storage and communication efficiency. For this purpose block based motion compensation techniques are used and have become an integral part of modern video codecs such as H.263 [1] and H.264/AVC [2].

Block based motion compensation involves dividing each frame into non overlapping rectangular blocks, matching each block with another suitable block in a previous frame and finally taking the difference of the two matched blocks. Most video encoders use the minimization of Sum of Absolute Differences (SAD) as a criterion for finding the best match for a block [3], a process commonly known as motion estimation. Current video codecs expect a high correlation between the two matching blocks such that the variance of the difference signal is smaller than the variance of the current block to be encoded. In video literature, this type of encoding is referred to as predictive coding [4]. However, it is important to realize that this procedure simply takes the differential of the two matched blocks and is equivalent to differential encoding used in audio signals. The notion of predictive encoding in audio signals is

different and involves linear estimation of a signal sample from previously observed samples. That is, the characteristics of a sample are predicted from previous sample values. Although existing video encoding techniques predict the motion of a block, they do not attempt to predict the relationship between two matched blocks. We argue that it is beneficial to predict a relationship between two matched blocks as compared to the current practice of simply taking their difference.

For motion estimation, SAD presents a computationally efficient solution [5] and is therefore used in existing video encoders. However, SAD is implicitly based on the ‘brightness constancy’ assumption, i.e. the intensity values of a block of video are not expected to change from one frame to another, although the block may undergo a spatial shift. However, such ideal conditions rarely exist: brightness and contrast changes are frequently observed between frames, especially in commercial videos. Even under simple linear changes, such as brightness variation, SAD does not guarantee a correct match.

As a simple illustration of this fact, consider an  $\sqrt{n} \times \sqrt{n}$  image block. Suppose that the subsequent frame is brighter by a constant factor  $\Delta$  at each pixel. The correct matching location for this block will thus have a SAD value of  $n\Delta$ , and the difference signal at this location will have a variance of zero. However, it is quite likely that there would be other locations in the search area that will have a lower SAD value. This is because an addition of  $\Delta$  causes the intensity levels at the other locations to become closer to the intensity levels of the original block to be encoded. Thus, a motion estimator based on SAD will result in a match at an incorrect location where the variance of the difference signal can potentially be much higher than zero. Hence, even under slight departures from the brightness-constancy assumption, SAD no longer remains an accurate motion estimator for video encoding.

We consider the use of a first order linear estimator to model the changes in intensity of a block from frame to frame. This choice is motivated by observing the brightness and contrast changes in real videos. Instead of taking the difference between two matching blocks, we estimate one from the other and take the difference between the actual and estimated values. That is, if  $b_j$  is the best matching block for  $b_i$ , we use  $b_j$  to compute  $\hat{b}_i$  as the Minimum Mean-Square Error (MMSE) linear estimate of  $b_i$ , and then consider  $b_i - \hat{b}_i$  for further processing. We show that the variance of  $b_i - \hat{b}_i$  is always smaller than or equal to the variance of  $b_i - b_j$ , leading to better compression and resulting in storage and

communication efficiency. We further show that, when  $b_i - \hat{b}_i$  is used instead of  $b_i - b_j$ , the optimal criterion for finding the best match  $b_j$  for  $b_i$  is the maximization of the magnitude of correlation coefficient. The proposed scheme, Video Coding with Linear Compensation (VCLC), captures all first order variations in video signals. We have observed with extensive experimentation on eight commercial videos that considering a non linear predictor instead of a linear predictor results in diminishing gains, indicating that the video signal changes from frame to frame are primarily linear.

The rest of the paper is organized as follows: we provide a formal definition of the problem in Section II. Related work is discussed in Section III, while in Section IV we establish the theoretical justification of correlation coefficient to be used for motion estimation and MMSE linear estimation to be used in motion compensation process. Section V provides the details of overall encoding and decoding system based on VCLC. Experiments and results are discussed in Section VI followed by conclusions in Section VII.

## II. PROBLEM DEFINITION

We consider a digital video signal as a sequence of frames  $F$  indexed at discrete time  $k$ . For the purpose of encoding, each frame  $F(k)$  is divided into non overlapping blocks  $b(k, x, y)$ , each of size  $\sqrt{n} \times \sqrt{n}$ , where  $n$  is the total number of pixels in the block and the parameters  $x, y$  represent the spatial position of block  $b(k, x, y)$  within frame  $F(k)$ . Two primary steps in the video encoding process are:

- 1) Motion prediction (or motion estimation) which is carried out on each block  $b(k, x, y)$  by finding its closest match  $b(k', x', y')$ , where  $k' = k + \delta_k$ ,  $x' = x + \delta_x$ , and  $y' = y + \delta_y$ , in a judiciously selected search area within a previous frame.
- 2) Motion compensation, which essentially means finding the motion compensated differential signal

$$\Delta = b(k, x, y) - \mathbf{h}\{b(k', x', y')\}, \quad (1)$$

where  $\mathbf{h}(\cdot)$  is an arbitrary function that has to be chosen such that the variance of  $\Delta$ ,  $\sigma_{\Delta}^2$  is minimized.  $\Delta$  is also known as the motion compensated residue and  $\sigma_{\Delta}^2$  is known as inter-frame variance [6]. In current practice,  $\mathbf{h}(\cdot)$  is taken to be the identity function.

The primary goal of video encoding is to maximize the compression for which a heuristic is to minimize the variance of motion compensated residue ( $\Delta$ ). Thus  $\mathbf{h}(\cdot)$  is used as an estimation function for  $b(k, x, y)$ , such that the estimation error variance ( $\sigma_{\Delta}^2$ ) is minimized.

We, therefore, intend to find the function  $\mathbf{h}(\cdot)$  in the motion compensation step as well as the criteria for finding the closest match in motion estimation step such that  $\sigma_{\Delta}^2$  is minimized. It is expected, as we will show, that the criteria for finding the closest match and the estimation function  $\mathbf{h}(\cdot)$  are closely related to each other.

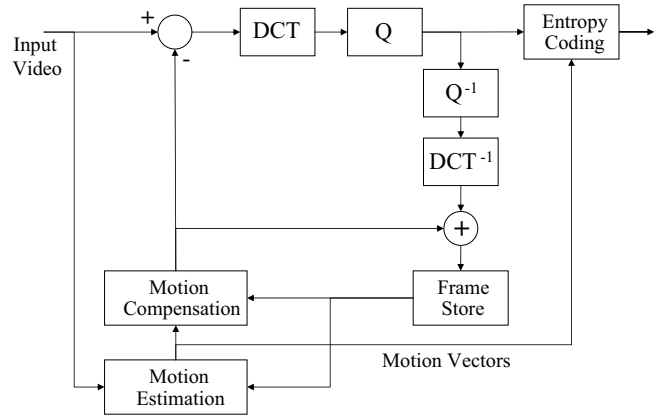


Fig. 1. A traditional generic interframe video codec.

## III. RELATED WORK

A number of schemes have been proposed and standardized for video encoding. All existing video encoders use minimization of SAD as the criteria for finding the closest match in the motion estimation step [3]. That is, the values  $k', x', y'$  are determined, such that the SAD value given by the following expression is minimized

$$SAD = \sum_{x=1}^{\sqrt{n}} \sum_{y=1}^{\sqrt{n}} |b(k, x, y) - b(k', x', y')|. \quad (2)$$

Furthermore, existing video encoders select  $\mathbf{h}(\cdot)$  in Equation (1) such that  $\mathbf{h}(\theta) = \theta$ . In this case, the resulting motion compensated differential signal ( $\Delta_d$ ) is given by

$$\Delta_d = b(k, x, y) - b(k', x', y'). \quad (3)$$

The variance of  $\Delta_d$  can be expressed in the following form:

$$\sigma_{\Delta_d}^2 = \sigma_b^2 + \sigma_{b'}^2 - 2\rho_{bb'}\sigma_b\sigma_{b'}. \quad (4)$$

where  $b \triangleq b(k, x, y)$  and  $b' \triangleq b(k', x', y')$  and  $\rho_{bb'}$  is the correlation coefficient between blocks  $b$  and  $b'$ . We define the gain of traditional video encoders as:  $\mathcal{G}_d = \sigma_b^2 / \sigma_{\Delta_d}^2$ . Using Equation (4), an expression for  $\mathcal{G}_d$  can be derived as

$$\mathcal{G}_d = \frac{\sigma_b^2}{\sigma_{\Delta_d}^2} = \frac{1}{1 + \frac{\sigma_{b'}}{\sigma_b} \left( \frac{\sigma_{b'}}{\sigma_b} - 2\rho_{bb'} \right)}. \quad (5)$$

If the video signal is assumed to be stationary, such that  $\sigma_b^2 = \sigma_{b'}^2$ , then Equation (5) reduces to

$$\mathcal{G}_{ds} = \frac{1}{2(1 - \rho_{bb'})}. \quad (6)$$

From Equation (6), we note that  $\mathcal{G}_{ds}$  is maximized when  $\rho_{bb'}$  is maximized:  $\rho_{bb'} \leq 1$ . However, minimization of SAD does not guarantee a maximization of  $\rho_{bb'}$ , thus SAD is not the optimal criteria for the maximization of  $\mathcal{G}_{ds}$ . Furthermore, in general the video signal is non-stationary and the true gain is given by (5), whose maxima cannot be guaranteed either by maximization of  $\rho_{bb'}$  or by minimization of SAD.

Nevertheless, from Equations (5) and (6), maximization of correlation coefficient appears to be a more attractive criteria

for motion estimation as compared to minimization of SAD. However, correlation coefficient has not been given serious consideration in the video encoding literature because of its high computational complexity [7]. In this paper, we restrict the scope of our discussion only to compression efficiency, and do not consider the computational efficiency aspects of match measures. It should be noted, however, that with the availability of powerful processors, it is practicable to implement complex algorithms to enhance the compression efficiency. For example, H.264/AVC uses much more complex algorithms than those employed by previous video encoding standards [8].

In our opinion, the use of correlation coefficient as a motion estimator has also been ignored in the video encoding community because of the comments made in the seminal paper by Jain and Jain [6], suggesting that the accuracy of the area correlation method is poor when the block size is small and the blocks are not undergoing pure translation. However, for block sizes commonly used in motion estimation algorithms, correlation coefficient actually outperforms SAD and other measures that are based upon the brightness constancy assumption. We have verified this by performing a large number of experiments on a number of scenes taken from ten commercial videos. Furthermore, in case of non-translational motion, all block based motion estimation algorithms suffer some degradation in performance. However, performance of correlation coefficient based estimators degrades much more gracefully [9].

We note that VCLC is a fundamental technique and other schemes and optimizations proposed in literature or included in standards may be used in addition to VCLC. Such additional schemes include the Overlapped Block Motion Compensation (OBMC) [10], [11] that was invented to handle the complex motion within a block. Similarly, sub-pixel motion estimation [12], that aims to increase the accuracy of motion compensation may also be used with VCLC.

#### IV. VIDEO CODING WITH LINEAR COMPENSATION (VCLC)

In traditional video encoding systems, the estimation function  $\mathbf{h}(\cdot)$  in Equation (1) is selected to be the identity function, such that  $\mathbf{h}(\theta) = \theta$ . That is, in these systems, the difference signal given in Equation (3) is used for further processing. Inherently, the use of Equation (3) is based on the brightness constancy assumption for pixel intensities. However, we have observed that brightness and contrast changes are so ubiquitous in natural videos, especially commercial videos like movies, that the assumption of constant pixel intensities breaks down frequently.

We have experimentally verified this observation by measuring the average MSE of the original block and the matched block while varying the estimation filter  $\mathbf{h}(\cdot)$ . Figure 2 shows the reduction in MSE as  $\mathbf{h}(\cdot)$  was changed from identity to first order linear and first order quadratic estimator. A significant decline in MSE was observed when a linear estimator was used compared to identity function. Increasing the complexity of

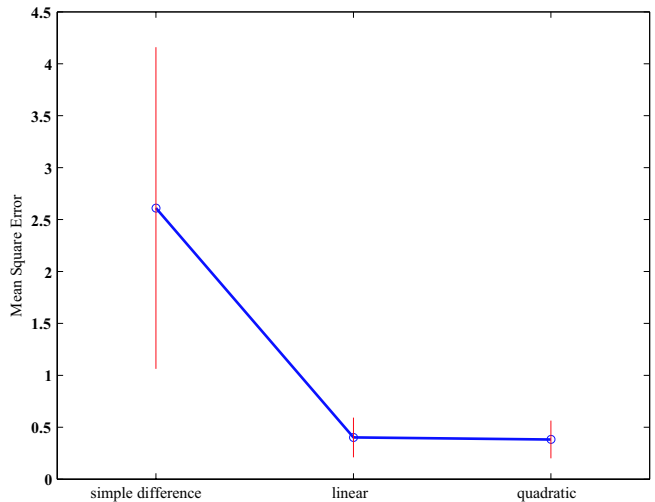


Fig. 2. The average and standard deviation of Mean Squared Error of different estimation filters  $\mathbf{h}(\cdot)$ . More than 400,000 8x8 blocks taken from eight commercial movies were used to compute these statistics.

the estimator from linear to quadratic resulted in diminishing returns, and subsequent improvements were not as significant.

Hence, in this paper, we propose that intensity changes between blocks in the nearby frames can be better modeled by a first order linear estimator. Therefore, we select  $\mathbf{h}(\cdot)$  for the estimation of block  $b$  as

$$h(b') = \alpha b' + \beta, \quad (7)$$

where  $\alpha$  and  $\beta$  are selected to minimize the mean square error between  $\mathbf{h}(b')$  and the block  $b$  that is being estimated. With each block of video, these two additional parameters are transmitted, but, as we will show, the corresponding reduction in the variance of the linear compensated difference signal justifies this overhead.

In the next two subsections, we first discuss the theoretical impact of choosing the first order linear model on the motion compensation strategy, and then discuss the optimal motion estimator under this model.

##### A. Motion Compensation using Linear Estimator

For the motion compensation step, current input block  $b(k, x, y)$  is linearly estimated from the best matching block  $b(k', x', y')$ , instead of computing the difference as in traditional methods. Thus we use

$$\hat{b}(k, x, y) = \alpha b(k', x', y') + \beta. \quad (8)$$

The parameters  $\alpha$  and  $\beta$  are selected such that the mean squared estimation error, given below is minimized:

$$\Lambda = \sum_{y=1}^{\sqrt{n}} \sum_{x=1}^{\sqrt{n}} (b(k, x, y) - \alpha b(k', x', y') - \beta)^2. \quad (9)$$

Minimizing  $\Lambda$  with respect to  $\alpha$  and  $\beta$  yields

$$\alpha = \rho_{bb'} \frac{\sigma_b}{\sigma_{b'}}, \quad (10)$$

$$\beta = \mu_b - \rho_{bb'} \frac{\sigma_b}{\sigma_{b'}} \mu_{b'}. \quad (11)$$

where  $\mu_b$  and  $\mu_{b'}$  are the means of the blocks  $b$  and  $b'$ . In the proposed VCLC scheme, we define motion compensated residue  $\Delta_p$ , similar to the traditional case, but using the MMSE linear estimate  $\hat{b}$  instead of  $b'$

$$\Delta_p = b(k, x, y) - \hat{b}(k, x, y). \quad (12)$$

It is straightforward to show that the mean of  $\Delta_p$  is always zero, regardless of the form of the original and the matched block. The variance of  $\Delta_p$  has a direct impact on compression efficiency: if  $\sigma_{\Delta_p}^2 < \sigma_{\Delta_d}^2$ , VCLC would lead to better compression compared to the traditional schemes. Since  $\Delta_p$  is zero mean, its variance is the minimum mean square error of estimation given by Equation (9), which can also be derived to the following form:

$$\sigma_{\Delta_p}^2 = (1 - \rho_{bb'}^2)\sigma_b^2. \quad (13)$$

The above relationship of  $\sigma_{\Delta_p}^2$ , i.e. the variance of linear compensated difference signal, should be compared to the expression of  $\sigma_{\Delta_d}^2$  in Equation (4), which is the variance of the simple difference. Using Equation (13), we can show that  $\sigma_{\Delta_p}^2$  is always less than or equal to  $\sigma_{\Delta_d}^2$ .

*Theorem 1:* For same motion estimator,  $\sigma_{\Delta_p}^2$  is upper bounded by  $\sigma_{\Delta_d}^2$ .

*Proof:* Since the square of any real number is non-negative, the following inequality holds

$$\left(\rho_{bb'} \frac{\sigma_b}{\sigma_{b'}} - 1\right)^2 \geq 0. \quad (14)$$

Rearranging we get

$$\sigma_b^2 - \rho_{bb'}^2 \sigma_b^2 \leq \sigma_b^2 - 2\rho_{bb'} \sigma_b \sigma_{b'} + \sigma_{b'}^2. \quad (15)$$

Comparing Equation (15) with Equations (4) and (13), it follows that

$$\sigma_{\Delta_p}^2 \leq \sigma_{\Delta_d}^2. \quad (16)$$

Note that this result holds true regardless of the form of the input signal. ■

Similar to the definition of  $\mathcal{G}_d$  in section III, we define motion compensation gain of the VCLC scheme as

$$\mathcal{G}_p = \frac{\sigma_b^2}{\sigma_{\Delta_p}^2} = \frac{1}{1 - \rho_{bb'}^2}. \quad (17)$$

Since  $\sigma_{\Delta_p}^2 \leq \sigma_{\Delta_d}^2$ , therefore  $\mathcal{G}_p \geq \mathcal{G}_d$ . Hence it can be concluded that the use of VCLC scheme will never result in a lower gain when compared with traditional encoding scheme.

### B. Motion Estimation with Correlation Coefficient

In previous discussion, advantage of VCLC scheme was shown over the traditional motion compensation techniques, independent of motion estimation process. This implies that if the VCLC scheme is used with traditional motion estimation, gain will still be improved. However we notice from Equation (17) that the gain of VCLC is maximized when  $|\rho_{bb'}|$  is maximized. This indicates that for VCLC, the optimal criteria for finding the closest match in the motion estimation step is not the minimization of SAD, rather it is the maximization of

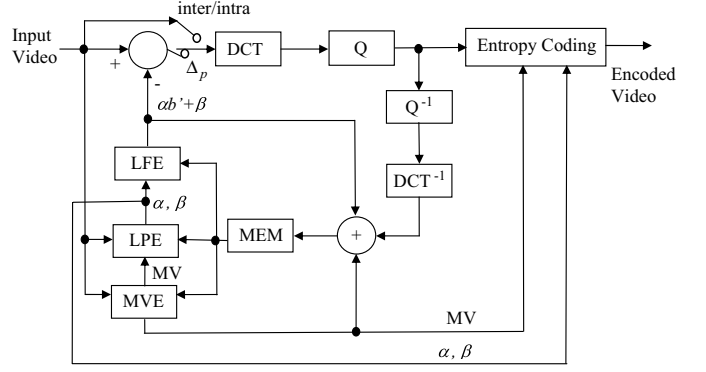


Fig. 3. Block diagram of VCLC video coder. MVE: Motion Vector Estimator, LPE: Linear Parameter Estimator, LFE: Linear Frame Estimator.

the magnitude of correlation coefficient in the search space. Thus the location of best matching block is given by

$$(\tilde{k}, \tilde{x}, \tilde{y}) = \arg \max_{k', x', y'} |\rho_{bb'}|, \quad (18)$$

where

$$\rho_{bb'} = \frac{\sum_{x=1}^{\sqrt{n}} \sum_{y=1}^{\sqrt{n}} (b(k, x, y) - \mu_b)(b(k', x', y') - \mu_{b'})}{\sigma_b \sigma_{b'}}. \quad (19)$$

Thus there is no other location where the linear compensated differential signal would have a lower variance or a higher gain than the one obtained by maximizing Equation (18) over  $(k', x', y')$ .

## V. SYSTEM OVERVIEW

Simplified block diagrams of VCLC encoder and decoder are shown in Figures 3 and 4 respectively. The input video frame to be encoded is sent to the Motion Vector Estimator (MVE) which also obtains a reference frame from the memory. The MVE finds the best matching blocks in the reference frame for each block in the input video frame, by maximizing the magnitude of correlation coefficient as given in Equation (18). For each block MVE provides the motion vector information to Linear Parameter Estimator (LPE) which computes  $\alpha$  and  $\beta$  for each block in accordance with Equations (10) and (11). LPE sends these parameters to Linear Frame Estimator (LFE) where the linear estimate of the complete frame is formed using the linear estimates of the individual blocks. The linear estimate of the complete frame is subtracted from the input video frame and the resulting residue error is further processed through transform coder (e.g., DCT), quantizer and entropy coder.

Traditional decoders require residue error information along with motion vectors in order to decode the current frame. VCLC decoder additionally requires transmission of  $\alpha, \beta$  parameters. We, however, observe that when using VCLC, the mean of motion compensated residue is zero, resulting in a zero DC value of the transform of each block

$$\mathcal{DC}_{\text{VCLC}} = 0, \quad (20)$$

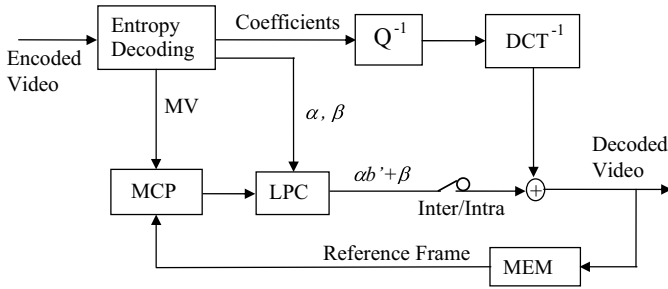


Fig. 4. A generic interframe predictive decoder with linear compensation. MCP: Motion Compensated Prediction, LPC: Linear Parameter Compensation.

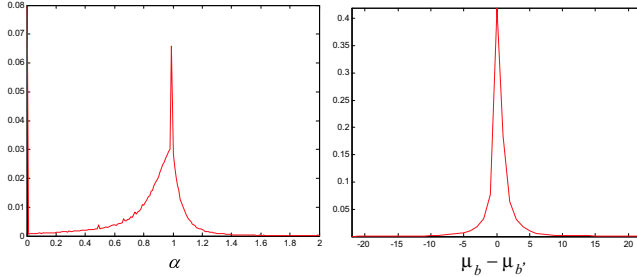


Fig. 5. (a) The histogram of  $\alpha$  always have a maxima at 1.00. (b) The histogram of  $\mu_b - \mu_{b'}$  always have a maxima at 0.00 and have shape similar to Laplacian distribution.

which essentially reduces transmission of one additional parameter as compared to traditional encoders. In traditional generic encoders (GE) [3], the DC value of a transformed block is the difference of means of input block  $b$  and its best matching block  $b'$

$$DC_{GE} = \mu_b - \mu_{b'}, \quad (21)$$

and it is generally non-zero. Note that in VCLC scheme, instead of transmitting  $\alpha, \beta$  parameters, we can transmit  $\alpha, (\mu_b - \mu_{b'})$  and reconstruct  $\alpha, \beta$  on decoder side using Equation (11). Therefore as compared to traditional systems, the actual overhead is only one parameter per block. Furthermore, for intermediate to larger block sizes, for example 8 by 8 or above, the cost of  $\alpha$  parameter, in terms of bits per pixel, turn out to be insignificant. For smaller block sizes, which are not very common in video encoders due to large number of motion vectors, the cost of sending an additional parameter becomes noticeable requiring the use of efficient quantization and coding for sending the same.

## VI. EXPERIMENTS AND RESULTS

We experimentally verified the theoretical VCLC results of previous sections by encoding scenes selected from numerous commercial videos. These videos often exhibit significantly larger changes in lighting compared to the standard test sequences often used in video codec research. On this dataset, the efficiency of VCLC was compared with that of traditional Generic Encoder (GE) [3] which used SAD for motion estimation and simple differences for motion compensation. In our

TABLE I  
COMPARISON OF TRADITIONAL GENERIC ENCODER (GE) AND VCLC  
MOTION COMPENSATION  $SNR$  (dB)

DataSet	$SNR_{VCLC}$			$SNR_{GE}$		
	4×4	8×8	16×16	4×4	8×8	16×16
Fast&Furious	36.11	36.51	29.03	29.09	33.04	25.32
BatmanBgins	41.83	39.33	32.97	34.05	34.66	28.82
KingKong	38.69	35.08	29.31	30.58	31.62	25.58
UnderWorld	42.70	39.58	35.25	34.95	34.34	29.15
Spiderman	35.01	31.56	26.44	29.64	29.46	24.41
PinkFloyd	40.59	37.66	35.14	37.18	35.83	33.29
Metallica	40.59	35.28	31.17	32.91	32.49	28.37
Blade	45.25	39.68	35.31	35.89	32.74	30.04
LordOfRings	39.69	35.84	33.60	34.15	31.91	30.13
MissionImps	36.70	31.87	27.66	29.42	26.60	23.75

experiments, the motion compensated residue of VCLC and GE was first transformed using DCT and then quantized by a uniform quantizer. The minimum number of bits needed to transmit the quantized residue was estimated by calculating its entropy. Note that VCLC improves the motion compensation efficiency while the other blocks of video encoder remain same.

The improvement in motion compensation is generally measured by the improvement in prediction  $SNR$  defined in [6]

$$SNR = 10 \log_{10} \frac{MI_{\max}^2}{\sum_{r=1}^M \sigma_r^2}, \quad (22)$$

where  $M$  is the total number of residue blocks,  $\sigma_r^2$  is the variance of a block of residue, and  $I_{\max}$  is the maximum pixel intensity. For traditional generic encoder:  $\sigma_r^2 = \sigma_{\Delta_d}^2$  and for VCLC:  $\sigma_r^2 = \sigma_{\Delta_p}^2$ . The  $SNR$  comparison as shown in table I was computed for three motion estimation block sizes:  $4 \times 4$ ,  $8 \times 8$  and  $16 \times 16$ . Maximum improvement in  $SNR$  was observed for  $4 \times 4$  block size, which was up to 11.3 dB, whereas, for  $8 \times 8$  and  $16 \times 16$  block sizes it was up to 6.8 dB and 6.1 dB respectively.

The  $SNR$  in equation (22) measures the performance of an encoding scheme based on the variance of the motion compensated residue without considering the effects of transform encoding and quantization. A better way to evaluate an encoding scheme is to characterize the end to end performance of the system by measuring the distortion in the decoded signal. Although in VCLC scheme only motion compensation and motion estimation steps are improved, by simple experiments we show that end to end performance is also improved. We computed peak signal to noise ratio ( $PSNR$ ) defined in [13] as

$$PSNR = 10 \log_{10} \frac{255^2}{MSE}, \quad (23)$$

where  $MSE$  is the mean squared error between original frame and the corresponding reconstructed frame.

In an end-to-end system, the additional parameters  $\alpha$  and  $(\mu_b - \mu_{b'})$  also have to be quantized before entropy coding. Typical histograms of both of these parameters are shown in figure 5. Individual parameter encoding was done by

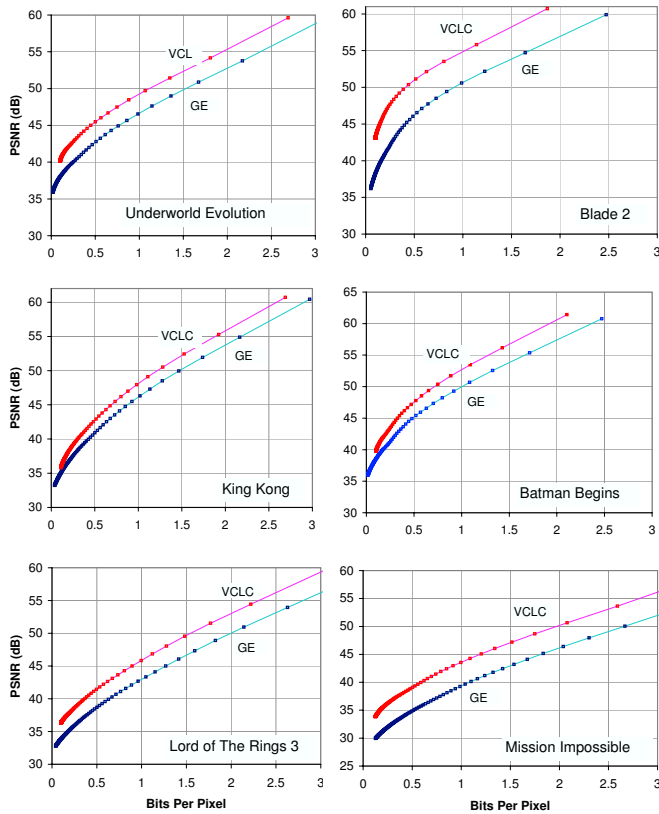


Fig. 6. Variation of  $PSNR$  with the variation of bits per pixel (bpp) for VCLC and GE.

quantizing these parameters, using the generalized Lloyd Max quantizer. The individual parameter encoding scheme required, on the average, 0.10 bpp overhead, for both parameters, for 8 by 8 block size. However, a simple block-based differential scheme is found to be more effective from the compression point of view as compared to the individual parameter encoding scheme. The block based differential scheme required, on the average, 0.0154 bpp overhead, for both parameters, for 8 by 8 block size. Figure 6 shows the average rate distortion curves for six videos encoding using  $8 \times 8$  block size. In figure 6, the slight rightward shift of the top curve, representing VCLC's performance, is due to the overhead of the two additional parameters. We note that the VCLC scheme exhibits an improvement of up to 5 dB in  $PSNR$  when compared with the traditional generic encoder.

## VII. CONCLUSION

In traditional encoders, motion compensation is done by taking the difference of the current block and its best matching block. The best matching block is selected by the minimization of the sum of absolute differences. In this paper it is

demonstrated that rather than directly encoding the difference between the two blocks, the difference between the current block and its first order linear estimate from the best matching block should be used. This choice of using linear compensated differential signal was based upon the fact that brightness and contrast changes frequently occur in real videos. Two important theoretical results were also demonstrated: (1) The variance of the linear compensated differential signal is always less than or equal to the variance of differential signal in traditional encoders, leading to better compression. (2) The optimal criteria for finding the best matching block, in the proposed scheme, is the maximization of the magnitude of correlation coefficient. We performed experiments on a large dataset taken from several commercial videos. For the same number of bits per pixel, the proposed scheme exhibited an improvement in  $PSNR$  of up to 5 dB when compared to the traditional encoding scheme.

## REFERENCES

- [1] ITU-T, "Video coding for low bit rate communication," *Recommendation H.263; version 1, Nov. 1995; version 2, Jan. 1998; version 3, Nov. 2000*.
- [2] ITU-T, ISO/IEC JTC1, "Advanced video coding for generic audiovisual services," *Int. Telecommun. Union-Telecommun. (ITU-T) and Int. Standards Org./Int. Electrotech. Comm. (ISO/IEC) JTC 1, Recommendation H.264 and ISO/IEC 14 496-10 (MPEG-4) AVC, 2003*.
- [3] M. Ghanbari, "Standard Codecs: Image compression to advanced video coding," *IEEE Telecom. Series 49, Institute of Electrical Engineers*, vol. 49, 2003.
- [4] R. Srinivasan and K. R. Rao, "Predictive coding based on efficient motion estimation," *IEEE Trans. Commun.*, vol. COM-33, no. 8, pp. 888–896, August 1985.
- [5] J. Vanne, E. Aho, T. D. Hamalainen, and K. Kuusilinna, "A high-performance sum of absolute difference implementation for motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 7, pp. 876–883, July 2006.
- [6] J. Jain and A. Jain, "Displacement measurement and its application in interframe image coding," *IEEE Trans. Commun.*, vol. 29, no. 12, pp. 1799–1808, December 1981.
- [7] D. Barnea and H. Silverman, "A class of algorithms for fast digital image registration," *IEEE Trans. Commun.*, vol. 21, no. 2, pp. 179–186, February 1972.
- [8] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, "Video coding with H.264/AVC: tools, performance, and complexity," *IEEE Circuits Syst. Mag.*, vol. 4, pp. 7–28, First Quarter 2005.
- [9] Q. X. Wu, "A correlation-relaxation-labeling framework for computing optical flow - template matching from a new perspective," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, no. 8, pp. 843–853, September 1995.
- [10] M. T. Orchard and G. J. Sullivan, "Overlapped block motion compensation: an estimation theoretic approach," *IEEE Trans. Image Processing*, vol. 3, no. 5, pp. 693–699, Sep 1994.
- [11] J. K. Su and R. M. Mersereau, "Motion estimation methods for overlapped block motion compensation," *IEEE Trans. Image Processing*, vol. 9, no. 9, pp. 1509–1521, September 2000.
- [12] B. Girod, "Motion compensating prediction with fractional pel accuracy," *IEEE Trans. Comput.*, vol. 41, no. 4, pp. 604–611, April 1993.
- [13] G. Sullivan and T. Wiegand, "Video compression— from concepts to the H.264/AVC standard," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 18–31, January 2005.