EARLY TERMINATION ALGORITHMS FOR CORRELATION COEFFICIENT BASED BLOCK MATCHING

Arif Mahmood, Sohaib Khan

Department of Computer Science, Lahore University of Management Sciences
{arifm,sohaib}@lums.edu.pk

ABSTRACT

Block based motion compensation techniques make frequent use of Early Termination Algorithms (ETA) to reduce the computational cost of block matching process. ETAs have been well studied in the context of Sum of Absolute Differences (SAD) match measure and are effective in eliminating a large percentage of computations. As compared to SAD, the correlation coefficient (ρ) is a more robust measure but has high computational cost because no ETAs for ρ have been reported in literature.

In this paper, we propose two types of ETAs for correlation coefficient: growth based and the bound based. In growth based ETA, ρ is computed as a monotonically decreasing measure. At a specific search location, when the partial value of ρ falls below the yet known maxima, remaining calculations are discarded. In bound based ETA, a new upper-bound on ρ is derived which is tighter than the currently used Cauchy-Schwartz inequality. The search locations where the proposed bound falls shorter than the yet known maxima are eliminated from the search space. Both types of algorithms are implemented in a cascade and tested on a commercial video dataset. In our experiments, up to 88% computations are found to be eliminated. In terms of execution time, our algorithm is up to 13.7 times faster than the FFTW based implementation and up to 4.6 times faster than the current best known spatial domain technique.

1. INTRODUCTION

Block-based Motion Compensation (BMC) techniques have become an essential component of all modern video encoders such as H.263 and H.264/AVC. In BMC techniques, each video frame is divided into non-overlapping rectangular blocks. Each rectangular block from the current frame is matched in a suitable area of a reference frame in order to find a best matching block. In traditional BMC techniques, minimization of Sum of Absolute Differences (SAD) is used as the criterion for the best match search and the process is known as *Motion Estimation*. Once the best match block is found, *Motion Compensation* is done by computing the pixel-by-pixel difference of the current block and the best match block.

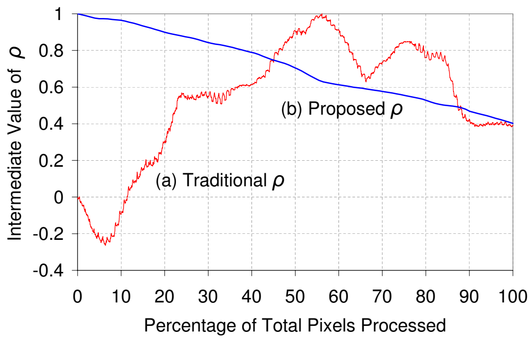

Fig. 1. Comparison of growth patterns of correlation coefficient ρ at a specific search block having $\rho = 0.40$. The traditional form (a) and the proposed monotonic form (b) follow different paths to reach the same final value.

The pixel-by-pixel difference, also known as *Motion Compensated Residue* (MCR), is further processed for encoding.

SAD based motion estimation is implicitly based upon brightness constancy assumption, whereas in real videos brightness and contrast changes between successive frames are frequently observed. In the presence of intensity variations, SAD may yield incorrect motion vectors yielding inefficient motion estimation. Furthermore, in the presence of intensity variations, traditional motion compensation technique may also become inefficient and the better approach may be to first compensate for the intensity variations and then compute the residue. It has been shown in [1] that the variance of linearly compensated MCR is upper bounded by the variance of traditional MCR. Also, it has been shown that minimization of variance of linearly compensated MCR is guaranteed by maximizing the correlation coefficient (ρ) over the search space. The minimization of the variance of the residue results in significant improvement in compression. Hence, ρ should be a preferred match measure as compared to SAD, especially for commercial videos which exhibit brightness and contrast changes more frequently.

Although block matching with maximization of ρ has advantage from the compression point of view [1], the high computational cost of ρ has often been cited as a criticism on its use in practice [2], [3], [4]; while SAD based BMC is computationally attractive, mainly because of the existence of Early Termination Algorithms (ETA). Two types of ETA are well studied for SAD: monotonic growth based ETA [4], [5] and the tight lower bound based ETA [6]. In monotonic growth based ETA, at a search location, as soon as the partial value of SAD becomes larger than the yet known minima, remaining calculations are skipped. In bound based ETA, a lower bound on SAD is evaluated at each search location. Locations exhibiting lower bound larger than the yet known minima are skipped. In contrast to SAD, to the best of our knowledge, no ETA has been developed for ρ based block matching. The growth pattern based ETA have not been applied to ρ based block matching, because of the non-monotonic growth pattern of ρ (Figure 1). Whereas, the bound based ETA have not been applied to ρ because of the absence of useful bounds on ρ. Currently, the best available bound on ρ is derived from the Cauchy-Schwartz inequality which is always fixed to +1, the maximum possible value of ρ, therefore it cannot be used for early termination.

Due to the non-monotonic growth pattern and the absence of tight bounds, the concept of early termination has been traditionally considered inapplicable to the ρ based block matching [2], [3], [4]. In contrast to this traditional understanding we, for the first time, not only argue the applicability of the early termination but also propose ETA of both types. For growth based ETA, we propose a monotonic form of ρ, such that during the block matching process, as the computations proceed on a specific search location, the value of ρ decreases monotonically (Figure 1). As soon as the current value of ρ falls below the yet known maxima, further calculations on that location become redundant and can be skipped. For bound based ETA, a new upper bound on ρ is derived, which is tighter than the currently used Cauchy-Schwartz inequality based bound. If at any search location, the proposed upper bound is found to be lower than the yet known best maxima, that specific search location can be eliminated from the search space. The bound based early ETA is used in cascade with the growth pattern based ETA to obtain significant speedup.

The proposed algorithm is tested on a dataset of commercial movies. On most of the datasets, more than 80% of the computations were found to be eliminated. On average, our technique is ten times faster than the best known implementation in Fourier domain (FFTW [7]) and three times faster than the best known spatial domain implementation (Bounded Partial Correlation [8]).

2. PROBLEM DEFINITION

A digital video is as a sequence of frames (F), indexed at discrete time k, each frame $F(k)$ of size $l \times m$ pixels. For encoding purpose, each frame $F(k)$ is divided into nonoverlapping blocks $b(k, i_o, j_o)$, each of size $n \times n$ pixels, where (i_o, j_o) represent the spatial location of first pixel of the block $b(k, i_o, j_o)$ within frame $F(k)$. Motion estimation is carried out for each block $b = b(k, i_o, j_o)$ by finding its closest match $b' = b(k', i'_o, j'_o)$, where $k' = k + \delta_k$, $i'_o = i_o + \delta_{i_o}$ and $j'_o = j_o + \delta_{j_o}$. The block b' is to be selected by the maximization of $\rho_{bb'}$ [1]

$$(\tilde{k}, \tilde{i}_o, \tilde{j}_o) = \arg \max_{k', i'_o, j'_o} |\rho_{bb'}|, \tag{1}$$

where

$$\rho_{bb'} = \frac{(\mathcal{CC}_{bb'} - \frac{1}{n^2}\mathcal{S}_b\mathcal{S}_{b'})}{\sigma_b \sigma_{b'}}. \tag{2}$$

$\mathcal{CC}_{bb'}$ is the cross correlation of blocks b and b':

$$\mathcal{CC}_{bb'} = \sum_{i=1}^{n}\sum_{j=1}^{n} b(k, i_o + i, j_o + j) b(k', i'_o + i, j'_o + j), \tag{3}$$

and σ_b is defined as:

$$\sigma_b = \sqrt{\mathcal{S}_b^2 - \frac{1}{n^2}(\mathcal{S}_b)^2}, \tag{4}$$

where $\mathcal{S}_b = \sum_{i=1}^{n}\sum_{j=1}^{n} b(i,j)$ and $\mathcal{S}_b^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} b^2(i,j)$.

The primary goal is to reduce the computational cost of correlation coefficient based motion estimation such that the original accuracy is preserved. Our approach to this problem is to obtain speedup by developing early termination algorithms. We therefore find a suitable form of ρ, so that the concept of early termination becomes applicable. We also find tight upper bounds on ρ and use them to verify the suitability of search locations for inclusion in calculations.

3. RELATED WORK

The efficient implementations of ρ have been based upon the form of ρ given by Equation (2) [9]. In this form, the cross-correlation term has the maximum complexity. The cross-correlation term has been computed in spatial domain with the computational complexity of the order of $\mathcal{O}\big((l-n+1)(m-n+1)n^2\big)$ and in frequency domain using FFT with the computational complexity of the order of $\mathcal{O}\big((l+n-1)(m+n-1)\log_2[(l+n-1)(m+n-1)]\big)$ [4]. The order of computational complexity of frequency domain implementation is lesser than the spatial domain implementation, however the basic operations in frequency domain implementation are significantly more complex. Therefore for small block sizes, typically used in video encoders, the spatial domain implementation turns out to be more efficient [9].

Recently a fast spatial domain technique has been proposed in [8] called as Bounded Partial Correlation (BPC) algorithm. In the BPC algorithm, cross correlation is calculated on some area of the blocks and upper bound on correlation is calculated on the remaining area, using Cauchy-Schwartz inequality. It has been shown that the sum of partial correlation and partial bound provides an upper-bound on the final cross-correlation. During block matching process, further calculations on a specific location are terminated if the calculated upper-bound falls shorter than the yet known maxima. Since the bound calculations are simpler than the correlation calculations, computations are reduced by selecting larger area for the bound; however it makes the bound more loose and it becomes more harder to obtained elimination. In the limiting case, if total block area is used for calculating bound, the Cauchy-Schwartz inequality yields an upper bound of +1, which will never become less than the known maxima and therefore no elimination will be obtained.

Many approximate schemes are also well studied in the context of block motion compensation. The more commonly used ones include two dimensional logarithmic search, three step search, conjugate direction search, cross search, orthogonal search. Note that the algorithms proposed in this paper are exact and will always yield the value and location of maxima same as that of brute force search. However the basic concept of our algorithms is also applicable to the approximate algorithms and may generate further speedup.

4. GROWTH BASED EARLY TERMINATION ALGORITHM FOR CORRELATION COEFFICIENT

In image processing literature, correlation coefficient between two blocks is commonly interpreted as the covariance of the two blocks ($\sigma^2_{bb'}$) normalized by the individual image standard deviations [3]:

$$\rho_{bb'} = \frac{\sigma^2_{bb'}}{\sigma_b \sigma_{b'}}. \quad (5)$$

Most of the current implementations of correlation coefficient are based upon this interpretation. Variance of an image block is always positive, therefore the variation in magnitude of $\rho_{bb'}$ while processing consecutive pixels are directly proportional to the variations in $\sigma_{bb'}$. Since $\sigma_{bb'}$ has a non-monotonic growth pattern, it results in a non-monotonic growth of $\rho_{bb'}$ as well. Figure 1 shows the non-monotonic growth pattern of $\rho_{bb'}$ when computed in traditional way. Due to non-monotonic growth of traditional form of ρ, early termination algorithms have long been considered inapplicable on ρ based block matching [2], [3].

We propose a monotonic form of $\rho_{bb'}$ suitable for early termination algorithms, while proceeding from the traditionally used form of ρ (5):

$$\rho_{bb'} = \sum_{i=1}^{n}\sum_{j=1}^{n} \widetilde{b}(k, i_o + i, j_o + j)\widetilde{b}(k', i'_o + i, j'_o + j), \quad (6)$$

where:

$$\widetilde{b}(k, i_o + i, j_o + j) = \frac{b(k, i_o + i, j_o + j) - \mu_b}{\sigma_b}, \quad (7)$$

μ_b is the mean of block b. Since $\widetilde{b}(k, i_o, j_o)$ has zero mean and unit variance therefore

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \{\widetilde{b}^2(k, i_o + i, j_o + j) + \widetilde{b}^2(k', i'_o + i, j'_o + j)\} = 2. \quad (8)$$

From Equations (6) and (8):

$$\rho_{bb'} = 1 - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \Delta^2_{bb'}(i, j), \quad (9)$$

where $\Delta_{bb'}(i, j)$ is given by:

$$\Delta_{bb'}(i, j) = \{\widetilde{b}(k, i_o + i, j_o + j) - \widetilde{b}(k', i'_o + i, j'_o + j)\}. \quad (10)$$

Since $\Delta_{bb'}(i, j)$ is not a complex number, its square is always positive. Therefore as the summation process proceeds, $\rho_{bb'}$ either decreases or remains same. Figure 1 shows such monotonically decreasing growth pattern for a specific search location as computations progress over a block.

The monotonic decreasing growth pattern of ρ, as given by (9), provides the opportunity of early termination. For a specific search location, the partial value of ρ calculated over $(p \leq n, q \leq n)$ pixels provides an upper bound on the final value of $\rho_{bb'}$ to be obtained at that location:

$$\rho_{bb'}(p, q) \geq \rho_{bb'} \; \forall p \leq n, q \leq n. \quad (11)$$

As a result, if at a specific search location, the current value of correlation coefficient $\rho_{bb'}(p, q)$ once falls below the yet known maxima ρ_{max}, it cannot improve after processing the remaining pixels. Therefore, further calculations at that search location becomes redundant and can be skipped. The percentage of pixels skipped at each search location depends upon the dissimilarity between the blocks b, b' and the magnitude of ρ_{max}.

Table 1. Comparison of Execution Time in Seconds

DataSet	FFTW	BPC	Proposed
Fast& Furious	444.44	133.76	47.58
BatmanBgins	803.40	205.92	72.31
KingKong	424.80	120	34.07
UnderWorld	687.48	149.6	50.01
Spiderman	197.12	66.304	14.42
PinkFloyd	105.35	40.635	17.63
Metallica	115.5	33	15.65
Blade	495.55	197.90	52.56
LordOfRings	531.84	128	66.95
MissionImps	185.76	43.2	26.88

5. BOUND BASED EARLY TERMINATION ALGORITHM FOR CORRELATION COEFFICIENT

In bound based ETA, an upper bound on ρ is derived as a function of the contents of the two blocks to be matched. Then that upper bound is used to filter out the search locations, exhibiting large dissimilarities with the block b. The proposed bound is based upon the following L_1 norm and L_2 norm inequalities [5]:

$$\big|||b||_1 - ||b'||_1\big| \leq ||b - b'||_1, \tag{12}$$

and

$$||b - b'||_1 \leq n||b - b'||_2. \tag{13}$$

From Equations (12) and (13):

$$\Big|||b||_1 - ||b'||_1\Big| \leq n||b - b'||_2, \tag{14}$$

which can be rearranged to the following form:

$$\sum_{i,j}(b(i,j) - b'(i,j))^2 \geq \frac{1}{n^2}\big(\sum_{i,j}|b(i,j)| - \sum_{i,j}|b'(i,j)|\big)^2. \tag{15}$$

From Equations (6) and (15), upper bound $\rho^u_{bb'}$ on $\rho_{bb'}$ is given by:

$$\rho^u_{bb'} = 1 - \frac{1}{2n^2}\big(\sum_{i,j}|\widetilde{b}(i,j)| - \sum_{i,j}|\widetilde{b'}(i,j)|\big)^2. \tag{16}$$

where $\widetilde{b}$ and $\widetilde{b'}$ are as defined by Equation (7). The bound $\rho^u_{bb'}$ may become significantly tighter than the previously used Cauchy-Schwartz inequality based bound [8], depending upon the contents of the blocks b and b'. The bound $\rho^u_{bb'}$ can be used for early termination of $\rho_{bb'}$ calculation as follows: at a specific search location $\rho^u_{bb'}$ is calculated before starting the correlation process and compared with the yet known best maxima, ρ_{max}. If $\rho^u_{bb'}$ is found to be lower than the ρ_{max}, correlation calculation at that specific search location becomes redundant and can be skipped. If the number of such locations turn out to be significant, overall block matching procedure becomes faster.

6. EXPERIMENTS AND RESULTS

The proposed algorithms are compared with FFTW based frequency domain implementation and with BPC based fast spatial domain implementation of ρ. The comparisons are done in full search mode, for 8×8 block size, on a dataset prepared from the selected frames of ten commercial movies [1]. The two proposed algorithms are implemented in a cascade form. At each search location, first the bound based elimination condition is checked. If the search location survives the elimination test, correlation calculations are started in the form given by (9). As the successive pixels are processed, the correlation value decreases. As soon as the current correlation is found to be lesser than a previous known maxima, remaining calculations are skipped. The amount of skipped calculations are recorded and a cumulative value for each dataset is shown in table 2. On most of the datasets, more than 80% calculations are found to be eliminated. Since actual calculations reduced significantly, large speedups are observed (table 1). The proposed algorithms are up to 13.7 times faster than the FFTW based correlation technique and up to 4.6 times faster than the BPC technique. Note that these speedups are not obtained by compromising the accuracy. The proposed algorithms have same accuracy as the brute force implementations of ρ. Thus, despite large speedups, the quality of motion estimation remains preserved.

Table 2. Comparison of Percent Calculation Elimination

DataSet	FFTW	BPC	Proposed
Fast& Furious	0	19.16	68.76
BatmanBgins	0	36.81	86.21
KingKong	0	21.03	81.14
UnderWorld	0	29.17	84.19
Spiderman	0	30.87	86.20
PinkFloyd	0	26.28	87.08
Metallica	0	25.14	81.89
Blade	0	32.18	88.75
LordOfRings	0	42.97	81.19
MissionImps	0	44.82	86.89

7. REFERENCES

[1] A. Mahmood, Z. A. Uzmi, and S. Khan, "Video coding with linear compensation (VCLC)," *IEEE ICC*, 2007.

[2] W. K. Pratt, *Digital Image Processing, 3rd Ed.*, 2001.

[3] B. Ziltova and J. Flusser, "Image registration methods: A survey," *Img. Vis. Comp.*, vol. 21, pp. 977–1000, 2003.

[4] D.I. Barnea and H.F. Silverman, "A class of algorithms for fast digital image registration," *IEEE Trans. Commun.*, vol. 21, no. 2, pp. 179–186, February 1972.

[5] H.S. Wang and R.M. Mersereau, "Fast algorithms for the estimation of motion vectors," *IEEE Trans. Image Processing*, vol. 8, no. 3, pp. 435–438, March 1999.

[6] W. Li and E. Salari, "Successive elimination algorithm for motion estimation," *IEEE Trans. Image Processing*, vol. 4, no. 1, pp. 105–107, January 1995.

[7] Matteo Frigo and Steven G. Johnson, "The design and implementation of FFTW3," *Proceedings of the IEEE*, vol. 93, no. 2, pp. 216–231, 2005.

[8] di Stefano L., S. Mattoccia, and F. Tombari, "ZNCC-based template matching using bounded partial correlation," *Pattr. Rec. Ltr.*, vol. 26, no. 14, 2005.

[9] J.P. Lewis, "Fast normalized cross-correlation," *Vis. Inter.*, pp. 120–123, 1995.