Journal of Engineering and Applied Sciences 4 (5-6): 358-368, 2009
ISSN: 1816-949X

Efficient Intra Mode Search Algorithms Based on Image Region Correlation for Inter-Frames in H.264/AVC Video

Byung-Gyu Kim
Department of Computer Science and Engineering, College of Engineering,
Sun Moon University, Republic of Korea

Abstract: Fast intra mode determination schemes based on an image region correlation and the Rate-Distortion (RD) cost are proposed for interframes in the H.264/AVC video standard. With the inter mode search procedure with 7 block types, an intra mode search causes a significant increase in the complexity and computational load for an inter-frame. At first, we adopt a Macro-Block (MB) tracking scheme to obtain the most correlated image region in the reference picture. To reduce the computational load of the intra mode search at the inter-frame, the Rate-Distortion (RD) costs of a tracked MB or image region for the current MB is used and we propose adaptive thresholding algorithms for skipping the intra mode search. For the IPPP sequence type, the overall encoding time can be reduced up to 52% based on comparative analysis of experimental results with JM reference software.

Key words: H.264/AVC video, inter mode, intra mode, macro-block tracking, adaptive thresholding

INTRODUCTION

Developments in video coding techniques have accelerated over the last several years. H.264/AVC video coding is the newest standard defined by the Joint Video Team (JVT) (Wiegand *et al*., 2003) for which various techniques have been adopted to obtain a high coding efficiency compared to previous standards. A variable block size for inter mode prediction maximizes the coding efficiency based on Rate-Distortion Optimization (RDO) in the H.264/AVC coding standard (Weigand *et al*., 2003). The block sizes are 16×16, 16×8, 8×16, 8×8, 8×4, 4×8 and 4×4. In addition, intra mode prediction (nine modes for a 4×4 luma block and four modes for 16×16 luma and 8×8 chroma blocks) follows inter mode prediction to determine the best residual image (Wiegand *et al*., 2003). Recently, nine intra mode predictions for an 8×8 luma block were added for high profile.

Generally, the intra mode is very suitable for a picture that has a scene change or a boundary between video segments. In this case, it is difficult to minimize the rate and distortion costs by using a motion estimation (inter mode search) procedure. Thus, an intra-mode search for all MBs of that picture may be more effective. A frame with all intra prediction Mbs is called an Intra frame (I frame). For inter-frames with high temporal correlations, most of MBs are inter modes as the best coding mode because of the efficiency of the inter mode. The intra mode is very rare for inter-frames. However, all intra prediction modes must be checked to determine a mode that has a better coding efficiency than the best inter mode. As a result, the complexity and computational load is dramatically increased for the inter-frame. Therefore, a fast intra mode decision scheme that can reduce the complexity of theH.264/AVC video encoder is needed.

There are many algorithms for fast intra mode searches (Lee and Jeon, 2004; Cheng *et al*., 2005; Kim and Kim, 2006) and we can categorize these intra mode decision methods into two groups. One is a group that determines, whether to skip the intra mode decision procedure (Lee and Jeon, 2004; Cheng *et al*., 2005; Kim and Kim, 2006). The other is a group that decreases the intra search time either by reducing the number of prediction directions for the current MB, or by early termination of each directional mode search using thresholding (La *et al*., 2007; Li *et al*., 2007; Wang *et al*., 2006a, b; Cheng and Chang, 2005; Fu and Xu, 2004; Pan *et al*., 2005; Rehman and Kumarayapa, 2007; Rui Su *et al*., 2006; Shengfa and Zhenping, 2006).

Lee and Joen (2004) proposed a selective intra mode search method based on transformed coefficients of the residual image and upper-row and left-column pixel values. For a comparison of the average rate for the best inter mode and the average boundary error for the current block, they defined AR (Average Rate) as the number of bits required for texture data without a header. They also defined ABE (Average Boundary Error) by employing upper-row and left-column pixel values of already encoded neighboring blocks. Then, they omitted all intra modes when AR<ABE. This method is effective for video sequences with low motion or large stationary regions.

Cheng *et al.* (2005) suggested an intra mode decision method for inter-frame coding. A threshold value is defined based on the minimum distortion values of 3 spatially adjacent blocks (left, top, top-right). If the cost values of the inter modes are less than the defined threshold value, the intra mode search is skipped in the inter-frame coding. In this method, the threshold value has been determined through experiment.

Kim and Kim (2006) suggested a method based on Rate-Distortion (RD) costs of the neighboring MBs of the current MB and they proposed an adaptive thresholding scheme. Under the assumption that the current MB is highly correlated with neighboring MBs, they used the Sum of the Absolute Differences (SAD) when the best inter mode was determined. The minimum RD cost of the defined neighboring MBs was then used as an adaptive threshold value to design as effective thresholding scheme. If the threshold is less than the SAD of the best inter mode for the given MB, motion consistency between the two frames is low. In this situation, it is more desirable to check the intra mode prediction to obtain a better mode with a smaller RD cost. Otherwise, the intra mode search is omitted.

La *et al.* (2007) proposed a fast intra mode decision algorithm for an H.264/AVC encoder based on the Dominant Edge Direction (DED). Their algorithm used an approximation of Discrete Cosine Transform (DCT) coefficient formula to determine the DED. By detecting the DED before the intra prediction, 3 modes instead of 9 modes are selected for the RDO calculation to determine the best mode in a 4×4 luma block. For 16×16 luma and 8×8 chroma blocks, only 2 modes are selected instead of 4 modes. This method is fast but produces a large loss of image quality and bit increment compared to other methods based on the direction of image texture (Pan *et al*., 2005; Wang *et al*., 2006b). A novel VLSI design for an edge based fast intra mode decision algorithm has been also presented for hardware implementation (Li *et al*., 2007). A histogram of the local edge based prediction mode selection method is used and is mainly focused on Intra frame coding.

A method based on the inter-block correlation in the intra mode domain was proposed by Wang *et al.* (2006a). They considered four modes of neighboring coded MBs/blocks as a good candidate for an intra mode search. After the candidate mode (minimum of four modes of neighboring coded MBs/blocks) is determined, it is checked by comparing its RD cost with a threshold. A value less than the threshold indicates that the candidate is good enough for the current MB/block. Otherwise, other modes are needed to determine the optimal prediction mode for the current MB/block.

For intra prediction of 4×4 blocks, Cheng and Chang (2005) have presented a three step intra prediction algorithm using a correlation between neighborhood directions. Fu and Xu (2004) suggested a fast intra prediction algorithm that reduces the redundancy between the chroma mode decision and the luma mode decision. They observed that the chroma mode decision procedure is highly correlated with the luma decision procedure. They modified the intra mode search routine so that each chroma prediction mode was only examined once and the I16 and I4 MB prediction routines were also performed once.

A directional field based approach has been reported by Pan *et al.* (2005), where several directions are selected by using an edge direction texture histogram according to block types. A similar approach based on Dominant Edge Strength (DES) was developed by Wang *et al.* (2006b), who partitioned four 8×8 sub-blocks for MBs and four 2×2 sub-blocks for 4×4 sub-blocks. The dominant edge direction was determined by a defined edge detection filter using the Sobel mask. With this dominant direction mode, candidate modes were selected for the dominant direction mode, the DC and two adjacent modes. This method speeds-up the intra-frame with a large bit increment.

Su Rui *et al.* (2006) proposed a fast mode decision algorithm based on an integer transform and an adaptive threshold. First, integer transform operations on the original image are executed to determine the directions of the local texture. According to this direction, only a few of the possible intra prediction modes are tested during the first step. If the summation of differences (SAD) of the reconstructed block corresponding to the best mode is smaller than an adaptive threshold, the calculation is terminated. Otherwise, other possible modes are needed. A speed-up gain of 45% was reported for the IIIII sequence type.

Shengfa and Zhenping (2006) presented an efficient intra prediction mode decision algorithm for an MPEG-2 to H.264 video transcoder. They reused the side information gathered in the video decoding stage to simplify the mode and direction decision process for the Intra prediction. Full Discrete Cosine Transform (DCT) coefficients are available for an Intra coded MPEG-2 block. Since the DC coefficient of every DCT block reflects the average energy to some extent, they were motivated to accelerate the estimation process of the intra-frame prediction of H.264 by using the DCT coefficients of the MPEG-2 8×8 DCT blocks. Rehman and Kumarayapa (2007) presented an approach for 4×4 sub-blocks based on a correlation of modes between neighboring 4×4 sub-blocks. They adaptively selected only five modes among nine prediction modes by using the position of a 4×4 sub block in the slice or frame. This approach reduced the complexity of the intra mode search of I4 MB by approximately 45%, but did not reduce the total encoding time. Most of researches has been based on an I-Slice, except for Lee and Jeon (2004)'s method, which is computationally complex because of transformation of the residual image, Kim and Kim (2006)'s method and Cheng and Chang (2005)'s method. For many fast mode decision schemes, the Rate-Distortion (RD) optimization has also been used as the following:

$$J_{RD} = SAD_{Mode} + \lambda\{R(Header) + R(Residual)\} \quad (1)$$

Where:

J_{RD} = A bitrate-distortion value used as a cost function

SAD_{Mode} = The sum of the absolute differences for the given mode

λ = Denotes the Lagrangian multiplier

R (x) = A bit amount for coding x

Header = Provides header information

Residual = The residual data for the given MB

We propose a fast intra mode SKIP determination method based on the basis of the Rate-Distortion (RD) cost of the tracked MB for the current MB in the reference frame. We introduce a simple MB tracking scheme for this purpose. The remainder of this study is organized on the fast intra mode search algorithms which are based on MB tracking and adaptive thresholding schemes.

MATERIALS AND METHODS

Proposed fast intra mode search algorithms: H.264/AVC encoder performs the intra mode prediction for the intra frame and the inter frames. In fact, there are a few intra mode MBs in one inter-frame (or slice). In a full intra mode search, all combinations of intra mode predictions are executed to determine the best intra mode prediction.

Figure 1 shows the occupation ratio of the intra mode MB as the final MB mode according to various QP values and the IPPP sequence type for inter-frames. The intra mode occupied <2%, except for the Foreman sequence, although the QP value becomes small. For QP variation, intra mode MBs increase for the Foreman and Carphone sequences. For the Foreman sequence, the occupation was within approximately 1~15% as QP values.

For a picture with high temporal correlation between successive frames, it would be better to relate the occurrence of the intra mode MB with a local property of the image region, but not a property of image sequence. There are two types of local image properties. The first type is an MB with a highly homogeneous texture. In this study, it is difficult to determine the best matching MB in the reference picture because all regions have similar distortion values. This is called the drift phenomenon. As a consequence, the optimal motion vector will be large and will be assigned a number of bits. Therefore, the overall rate-distortion cost becomes larger than the cost of the best intra mode. It is better for the best intra mode to be the final mode. The second type is an MB with fast motion relative to the reference picture. In this study, since the minimized distortion value is inclined to be large, the intra mode is better as the final MB mode in terms of the coding efficiency.

All MBs of the given frame or slice must be examined by the intra mode prediction process, causing a large computational load for the encoder, although this is only a small part of all intra mode MBs. Thus, a separate fast intra mode search algorithm is necessary to speed-up the encoding process.

Two consecutive frames in a video sequence are highly correlated. With slow object motion in the video, the mode information for the same MB in a previous frame may still affect the mode determination process of the current MB because of the high temporal correlation. We propose an MB tracking scheme to make use of this temporal correlation. An adaptive decision rule to determine, whether the current MB requires an intra mode search is also suggested based on the proposed MB tracking scheme.

A simple MB tracking strategy: To track an object is to locate the current object region in an adjacent frame. To do this, we need to define the desired object region in the image plane as reported by Byung-Gyu (2008). We need a search procedure for the desired object region with a predefined tracking criterion in the temporal domain.

To apply this tracking scheme to block-based video coding, we consider each MB as a desired object in a mode decision procedure. As shown in Fig. 2, a P16×16 block type for the current MB (at time t) is used to locate the region in the previous frame (at time t-1) with the highest correlation. This is an integer pel motion estimation procedure for the current MB based on a block matching technique.

Once the best motion vector and the most highly correlated region are obtained, we determine the most

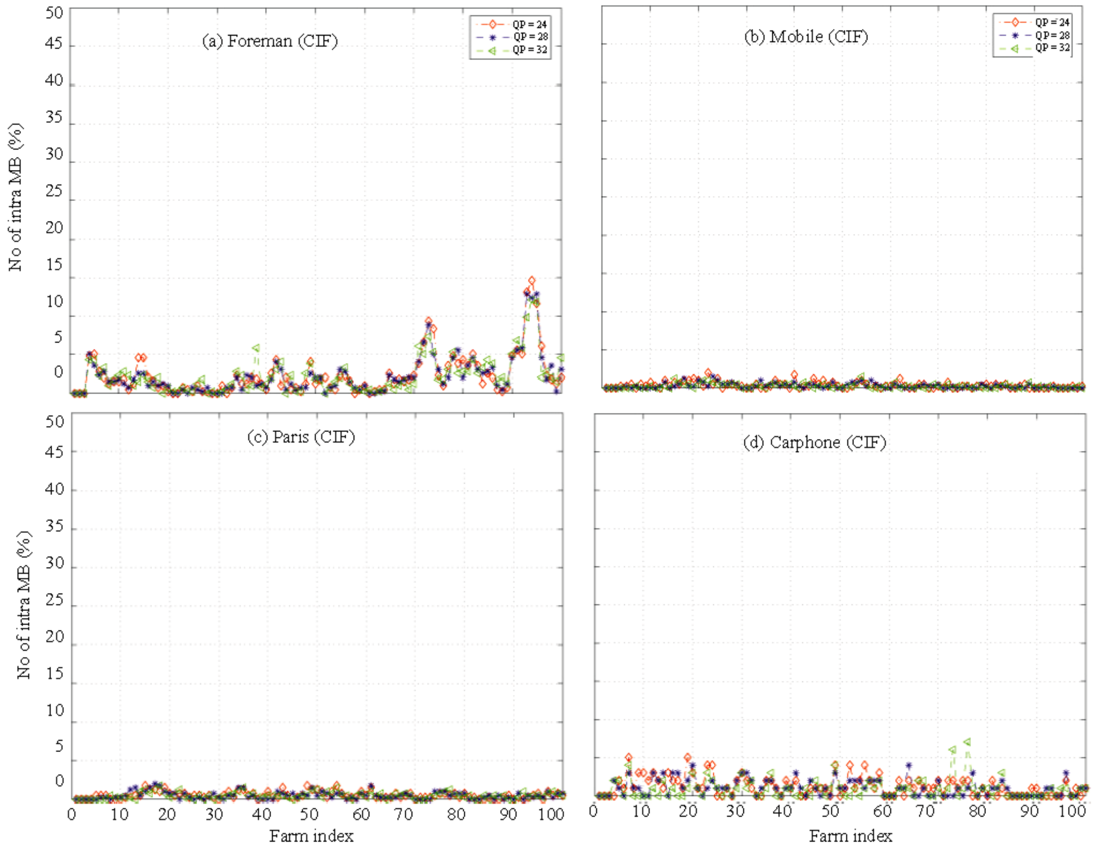

Fig. 1: The occupation of intra mode MBs with various QP values

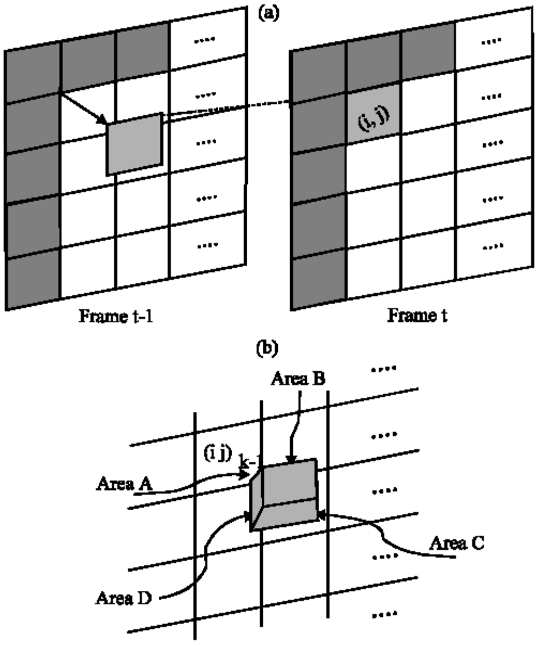

Fig. 2: An MB tracking scheme using P16×16 block motion estimation

highly correlated region of the current MB (MB (k, l)) in the previous frame as:

$$\text{Corr. Region} = \{\text{Area A, Area B, Area C, Area D}\} \quad (2)$$

Where (k, l) denotes an index of MBs that contain the overlapping (correlated) region of the current MB. A-D are the areas overlapping with the current MB (Fig. 2b). We use the above equation to determine the MB or image region that has the maximum correlation with the current MB. We can also use the overlapping area without determining the most correlated MB.

Adaptive thresholding schemes based on MB tracking: To achieve a fast intra mode detection scheme, we use the RD costs of the most correlated MB or image region with the current MB (MB_{kl}) using MB tracking.

Since, these tracked MB or image region in the reference frame is most correlated with the current MB, we may get information that will allow the intra mode prediction process to be skipped. Based on Eq. 1, the following relationship can be used for a given Mb_{kl}:

$$J_{RD_{kl}} = SAD^{kl}_{Mode} + \lambda\{R(Header) + R(MV) + R(Residual\ image)\} \tag{3}$$

Here, if R(Header) is usually negligible, then we can approximate Eq. 3 as:

$$\begin{aligned} J_{RD_{kl}} &\simeq SAD^{kl}_{Mode} + \lambda\{R(MV) + \lambda R(Residual\ image). \\ &= m\cos t_{kl} + \lambda R(Residual\ image) \end{aligned} \tag{4}$$

The right-most term is usually much smaller than the value of mcostkl. Thus, the RD cost of the current MB can be approximated as the value of $mcost_{kl}$. Also, the current MB (MB_{kl}) is highly correlated with the tracked MB or image region in the reference picture.

Using these facts, we propose two adaptive thresholding algorithms for skipping the intra mode search.

Algorithm based on tracked MB (Alg-1): The MB tracking scheme can provide approximate information for the current MB. We propose the following decision rule for extracting the intra mode SKIP:

$$\tilde{J}^{t-1}_{RD_{kl}} \geq m\cos t^{Inter\ best}_{kl} \tag{5}$$

Where:

$\tilde{J}^{t-1}_{RD_{kl}}$ = The RD cost of the tracked (most correlated) MB in the reference frame (t-1)

$m\cos t^{Inter\ best}_{kl}$ = Defined as the motion cost of Eq. 4

When the best inter mode is determined. Also, from Eq. 2 we define a correlation ratio, as follows:

$$\text{Corr. ratio} = \arg \text{Max}_{MB(kl)} \{\text{Area A, Area B, Area C, Area D}\} \tag{6}$$

If the correlation ratio is more than a predefined threshold τ, then we use the RD cost of the tracked MB to determine the most correlated MB. We use τ = 0.8~0.9 to guarantee enough correlation between two Mbs. If

$$\tilde{J}^{t-1}_{RD_{kl}}$$

is larger than the value of $m\ cost_{kl}^{Inter\ best}$ the selected inter-best mode is good enough for the current MB, so we can skip the intra mode search procedure. Otherwise, the intra mode search will be followed since it is possible to determine a more suitable mode among intra modes or an additional refinement stage.

Algorithm based on tracked image region (Alg-2): We do not use the most correlated MB of the current MB. Rather we use the most correlated image region. Unlike Alg-1, we compute a weighted RD cost of the tracked image region for the current MB (MB_{kl}) as:

$$\tilde{J}^{t-1}_{\omega RD_{kl}} = \sum_{ij} \omega_{ij} J^{t-1}_{RD11} \tag{7}$$

Where:

(i j) = The indices of MBs that overlap the current MB after MB tracking with a P16×16 mode

ω_{ij} = The correlation ratio of the tracked region (MB^{t-1}_{ij}) with the current MB as weights

$\tilde{J}^{t-1}_{RD_{ij}}$ = The RD cost values of the correlated MBs (as shown in Fig. 2). The defined weights in Fig. 2 are the normalized ratios of the overlapping areas Area A, Area B, Area C and Area D

We can design another form of adaptive thresholding scheme similar to Eq. 5 using this weighted RD cost:

$$\tilde{J}^{t-1}_{\omega RD_{kl}} \geq m\cos t^{inter\ best}_{kl} \tag{8}$$

If

$$\tilde{J}^{t-1}_{\omega RD_{kl}}$$

is larger than the value of $mcost_{kl}^{Inter\ best}$ the selected inter-best mode is good enough for the current MB, so we can skip the intra mode search procedure. Otherwise, we need the intra mode search to obtain better coding efficiency or to begin an additional refinement stage.

Refinement stage: We consider a simple optional refinement stage for extracting additional intra SKIP MBs. In the case of a P8×8 sub-block mode, the RD cost value is typically large. For the inter mode decision process in H.264/AVC coding, the P8×8 block type is chosen as the best mode when the current MB is composed of sub blocks (4×4) that have fast and different motion directions or complex boundaries with non-homogeneous motion components. The first reason why the RD cost of a P8×8 block is large is a large bit increment for the residual data because of increased distortion after motion estimation. The second reason is that motion vectors of 8×8 sub-blocks become large causing an increase in the number of bits for encoding motion information. Figure 3 shows, the average RD cost for the Foreman and Bus sequences.

The RD cost of the P8×8 MB type is large compared with the RD costs of other inter modes. It is desirable to select one among all intra modes as the best coding mode. However, when using only the averaged RD cost of

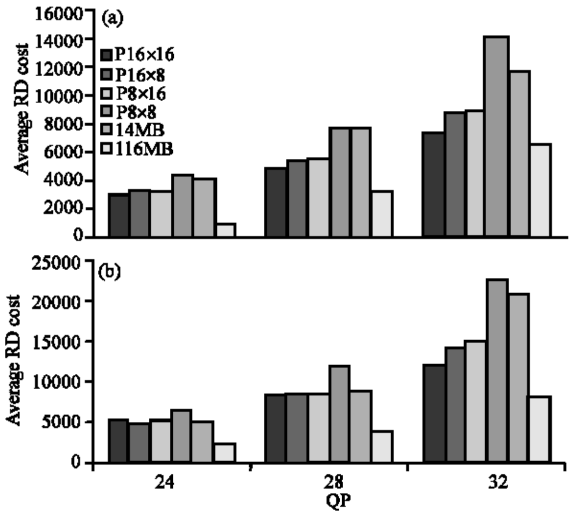

Fig. 3: The average RD cost values for (a) Foreman (CIF) and (b) Bus (CIF) sequences

P8×8 sub-block type MBs, this value is usually large. Thus, there may be an excessive number of MBs that are skipped if the motion cost value of the current MB is compared with this threshold value. This may cause an undesirable loss of quality and a bit increment. Therefore, we consider another term to eliminate this problem in the refinement stage. The I4 MB mode is usually dominant when intra mode is selected as the best coding mode because of stronger spatial correlation. Figure 4 shows the number of intra mode MBs for several sequences. The I4 MB mode is 2.5~10 times more common than the I16 MB mode for all tested sequences.

We use the average RD cost value of I4 MB, MBs that were pre-encoded for the algorithm. By using this additional term, we can handle image blocks that have fast motion or a complex texture. We can prevent excessive MB skipping of the intra mode search by considering the averaged RD cost of preencoded I4 MB, MBs. We can adaptively select an average threshold value as follows:

$$J^{TH}_{RD} = \frac{1}{n+1}(J'_{RD} + nJ^{TH}_{RD}) \qquad (9)$$

Where:

J'_{RD} = The RD cost value of an MB that is encoded as a P8×8 sub-block or the I4 MB mode

n = The number of Mbs encoded as a P8×8 sub block or the I4 MB mode up to the current MB. From the above equation, we can easily update the average RD value of J^{TH}_{RD}

If the value of J^{TH}_{RD} is larger than the value of most$_{kl}^{\text{Inter Best}}$, the selected inter-best mode is good enough for the current MB, so we can skip the intra mode search procedure. Otherwise, we need the intra mode search to obtain a better coding efficiency.

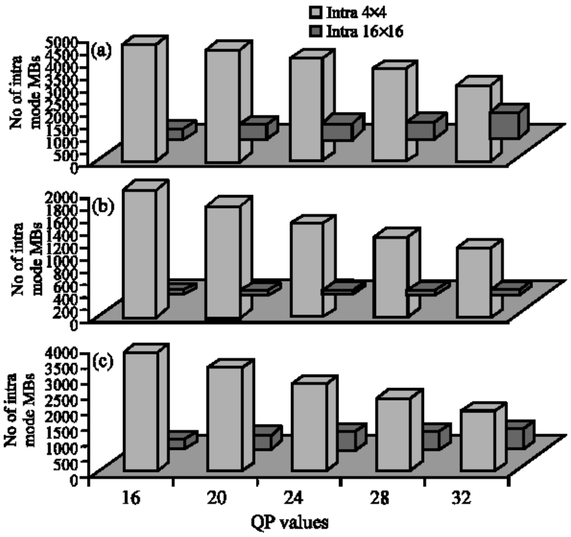

Fig. 4: The number of intra mode MBs for (a) Foreman (CIF), (b) Mobile (CIF) and (c) Carphone (QCIF) sequences

Both proposed algorithms include the refinement stage in this research. Although, the proposed algorithms use a thresholding scheme similar to the scheme of Kim and Kim (2006), there are differences.

First, we use the motion cost of the best inter mode for the current MB while, Kim use the SAD value of the best inter mode. Second, the proposed algorithms use the temporally most correlated region in the reference picture while Kim's method does not. Third, the defined adaptive threshold values are different from Kim's algorithm in terms of the most correlated MB.

We assumed that most of the correlated MBs were inter-predicted. However, this may not be true, we now consider when the tracked MB is an intra mode. Equation 4 is still valid because of approximating the RD cost of the best inter mode for the current MB. From Eq. 5:

$$\tilde{J}^{t-1}_{RD_{kl}}$$

is the RD cost of the tracked intra mode MB. This case is very clear. If

$$\tilde{J}^{t-1}_{RD_{kl}} > m\cos t^{\text{Inter Best}}_{kl}$$

the intra mode search can be skipped since

$$\tilde{J}^{t-1}_{RD_{kl}}$$

is the RD cost of the tracked intra mode MB. Otherwise, we must perform the intra mode search to determine a better coding mode.

RESULTS AND DISCUSSION

Various MPEG standard sequences were used with CIF and QCIF sizes to verify the performance of the proposed algorithms. Analyses were performed with encoding frames = 200, RD optimization enabled, QP = 20, 24, 28 and 32 sequence types of IPPP in the main profile, using CABAC, with a search range of MV = ±16 and the number of reference frames = 1 or 3. The Hadamard transform option was enabled. With this recording configuration, we can aim at PVR for mobile phone and digital video camcoder as target application systems.

JM 11.0 reference software of the JVT (joint video team) was used as a reference code for evaluation of the encoding performance. We compiled this JM software on Visual C++ 6.0 with Win 32 release mode. All algorithms for comparison were run on an HW platform of a Pentium4 PC with a 3.0GHz quad-core CPU and 1 GByte of RAM.

We defined three measures for evaluating the encoding performance, including average ΔPSNR, average ΔBits and an encoding-time saving factor, ΔT. The average ΔPSNR is the difference in (dB) between the average PSNR of the proposed method and the corresponding value of another method. The average PSNR is defined as:

$$\overline{\mathrm{PSNR}} = \frac{4\mathrm{PSNR}_{Y} + \mathrm{PSNR}_{Cb} + \mathrm{PSNR}_{Cr}}{6} \quad (10)$$

where, PSNR_{Y}, PSNR_{Cb} and PSNR_{Cr} are peak-to-noise ratios of the luminance and two chroma components, respectively. As performance improves, this criterion becomes larger. The average ΔBits is the bit-rate difference as a percentage (%) between compared methods. Lastly, the encoding-time saving factor ΔT is defined for a complexity comparison as:

$$\Delta T = \frac{T_{ref} - T_{proposed}}{T_{ref}} \times 100(\%) \quad (11)$$

under the condition that the Full Mode Search (FMS) is optimum (reference). As this value increases, the performance speed is increased. The required encoding time is measured by JM reference software and was defined as the total encoding time including IDR slices, motion vector searches and transforms. It must be noted that positive values for ΔPSNR and ΔBits indicate increments and negative values indicate decrements. We used two methods for an objective comparison of the encoding performance. These were Lee and Jeon (2004) and Kim and Kim (2006) methods, which are well-known as efficient and fast algorithms.

Figure 5 shows, the Rate-Distortion (RD) curves for four sequences. The proposed decision algorithms exhibited an RDO performance similar to the JM 11.0 original encoder with the full intra mode search. When intra mode prediction was disabled for inter-frames, a large loss of quality of approximately 0.1~0.3 (dB) for the Foreman, Football and Stefan sequences occurred. This is undesirable from the viewpoints of image quality and network bandwidth.

However, there was a negligible loss of quality in the Foreman sequence for the proposed algorithm 1. Also, algorithm 2 produced almost the same performance in overall bit rate. Although, we can observe that our algorithms produced similar performance for the Mobile sequence, an improvement of up to 52% was obtained for the speed of the encoding system. The algorithms yielded better quality at a higher bit rate for the Football sequence.

Table 1 shows the results for all algorithms using only one reference frame for the IPPP sequence type for QP = 20, 24, 28 and 32 due to a limitation of space. The proposed algorithm achieved a better speed-up factor with a minimal loss of image quality and a minimal bit increment compared with results for other fast methods. In sequences with stationary or slow motion, Lee and Jeon (2004)'s method yielded a better speed-up factor, by almost 5%, than Kim and Kim (2006)'s and the proposed algorithms. With no intra mode search, there was a negligible loss of quality and bit increment for the Flower and Paris sequences. Thus, the speed of the encoder can be improved just by removing the intra mode search procedure for these sequences. However, a significant degradation of image quality and a bit increment result for sequences with fast and global motion.

Kim and Kim (2006)'s algorithm is superior to Lee and Jeon (2004)'s method for sequences with fast object's motion or a large global motion (Stefan, Mobile, Football) as they mentioned. The proposed algorithms exhibited similar improvements in the encoding speed for the Paris and Foreman sequences. For sequences with fast object motion or a large global motion (Stefan, Mobile, Football), the algorithm-1 was faster than Kim's and Lee's methods by a factor of up to 25% for the average encoding time.

The algorithm-2 also improved the speed for the Stefan, Mobile and Football sequences, but was slower than the proposed Algorithm 1.

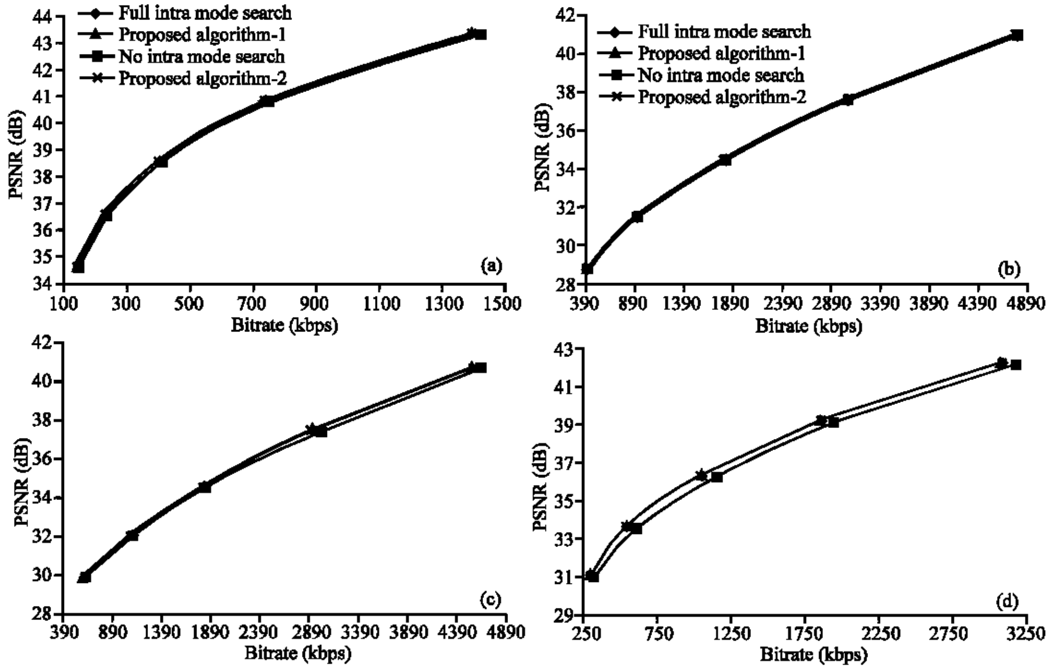

Fig. 5: RD curves for (a) Foreman (CIF), (b) Mobile (CIF), (c) Football (SIF) and (d) Stefan (QCIF) sequences

For QP = 24 with the Mobile sequence, the encoding speed was improved up to 52.17% with a bit decrement of 0.150%. Judging from total average values, the proposed algorithms produced a speed-up factor of 10% or a similar speed-up factor with a minimal loss of image quality and more bit-saving.

The results for all algorithms are shown for IPPP sequence types with multiple reference pictures (3 reference frames) in Table 2. Because of a multiple referencing structure for motion vector estimation, similar fashion is obtained although the speed-up gain is smaller than with just one reference frame. With no intra mode search a speed-up gain of almost 40% is achieved, while 30% gain is achieved using algorithm 1 (Table 2). Lee and Jeon (2004)'s method yields almost a 5% better speed-up factor than Kim and Kim (2006)'s in sequences with stationary or slow motion.

The proposed algorithms are better for the Football and Mobile sequences. Proposed algorithm-1 shows performance similar to Lee's method for sequences with stationary or slow motion, such as the Paris, Flower and Foreman sequences. For the Football and Stefan sequences, proposed Algorithm 1 is faster than Lee's method by >8%. Proposed Algorithm 1 achieves a total average value speed-up gain of 5% compared with Lee's method and 8% compared with Kim's method with a minimal loss of image quality and more bit saving. We now consider the influence of the suggested methods on the distribution of MB modes. As the QP value changes, the distribution of MB modes also gradually changes. The use of SKIP and 16×16 modes is decreased and the number of intra mode MBs are increased as QP value becomes smaller and vice-versa. For intra mode MBs, I4 MB is less common and I16 MB is more common as the QP value increases but I4 MB is dominant in most cases.

Results in Table 1 and 2 show that the speed-up gain of the proposed algorithms become smaller as the QP value increases due to the increase in use of the SKIP mode that omits both the inter mode and the intra mode search. As the use of the SKIP mode increases, the chance for elimination of the intra mode search is also increased in the total encoding time. Therefore, the speed-up gain of the proposed algorithms becomes smaller as the QP value increases.

If the RDO is off, the encoding system selects the best mode based only on the distortion (SAD) value. In this case, the suggested algorithms will operate to determine the mode that has a minimum SAD value based on the distortion. That is to say, the RD costs of Eq. 1 and 3-8 become SAD values. If we replace JRD_{kl} with SAD_{kl} in Eq. 4, we need not approximate this equation because of $JRD_{kl} = SAD_{kl}$. Therefore, Eq. 4 and 5 are still valid. For the weighted RD cost of the tracked image (Eq. 7), the

Table 1: Performance comparison of the proposed algorithms on the JM 11.0 reference encoder for IPPP sequences (Ref. frame = 1)

	QP = 20			QP = 24			QP = 28			QP = 32			Average values		
Contents	ΔPSNR	ΔBits	ΔT	ΔPSNR	ΔBits	ΔT	ΔPSNR	ΔBits	ΔT	ΔPSNR	ΔBits	ΔT	ΔPSNR	ΔBits	ΔT
Stefan (QCIF)															
No intra	-0.011	0.296	65.91	-0.065	4.462	62.37	-0.086	9.835	59.78	-0.112	12.459	57.00	-0.068	7.383	59.71
Lee's	0.001	-0.045	45.27	-0.006	-0.034	41.54	-0.002	0.100	33.77	-0.009	-0.207	33.13	-0.004	-0.046	35.52
Kim's	0.006	-0.070	42.56	-0.014	-0.099	38.31	-0.017	-0.166	36.48	0.001	-0.439	35.72	-0.005	-0.239	36.81
Alg-1	0.001	-0.070	47.63	0.001	-0.469	48.40	-0.014	-0.169	47.19	-0.016	0.188	44.96	-0.012	-0.259	46.54
Alg-2	0.003	-0.033	30.68	0.000	-0.194	32.67	-0.014	-0.193	33.56	-0.008	0.114	32.63	-0.005	-0.154	32.64
Flower (CIF)															
No intra	-0.022	-0.012	66.88	-0.014	0.040	64.62	-0.014	0.024	62.31	-0.003	0.107	59.60	-0.010	0.057	62.17
Lee's	-0.014	0.044	45.91	-0.009	0.029	45.62	-0.006	-0.006	32.89	-0.006	-0.014	30.78	-0.007	-0.001	38.80
Kim's	0.002	-0.200	29.52	0.002	0.026	31.90	-0.003	0.003	32.77	-0.004	-0.005	32.78	-0.003	-0.081	31.49
Alg-1	0.006	-0.306	45.80	0.006	-1.89	45.19	-0.010	-0.557	47.72	-0.027	-0.533	45.72	0.000	-0.918	46.91
Alg-2	0.006	-0.022	20.09	0.007	-0.016	24.89	-0.002	0.013	26.22	-0.002	0.013	25.79	0.000	-0.050	25.75
Mobile (CIF)															
No intra	-0.001	0.020	68.05	-0.006	0.051	66.15	-0.011	0.925	63.44	-0.013	0.138	60.35	-0.010	0.371	63.31
Lee's	-0.002	0.013	37.75	-0.006	-0.061	42.42	0.000	-0.026	30.92	-0.007	0.030	30.05	-0.006	0.009	33.53
Kim's	0.007	-0.412	38.75	0.002	-0.080	42.86	0.001	-0.108	42.38	-0.002	-0.104	40.57	-0.001	-0.148	40.83
Alg-1	0.002	-0.721	50.09	0.000	-0.150	52.17	0.001	-0.045	50.47	-0.005	0.050	47.97	-0.001	-0.202	50.18
Alg-2	0.004	-0.493	44.66	0.003	-0.103	49.4.	0.002	-0.110	48.44	0.008	-0.022	46.21	0.003	-0.252	47.18
Paris (CIF)															
No intra	-0.017	0.106	67.46	0.003	0.184	64.46	-0.042	0.122	61.45	-0.007	0.648	58.18	-0.015	0.330	61.36
Lee's	-0.012	-0.037	57.59	0.001	0.042	55.84	-0.030	-0.176	51.04	-0.021	0.128	49.22	-0.017	0.000	53.92
Kim's	0.004	-0.165	38.92	0.006	-0.257	37.79	-0.019	-0.537	37.61	-0.002	-0.350	36.89	-0.008	-0.395	36.43
Alg-1	-0.002	-0.311	52.79	0.003	-0.453	51.72	-0.016	-0.729	49.67	0.010	-0.679	47.29	-0.010	-0.555	50.12
Alg-2	-0.001	-0.091	44.50	0.015	-0.098	44.13	-0.019	-0.326	43.79	-0.001	-0.263	41.32	-0.009	-0.334	42.44
Foreman (CIF)															
No intra	-0.043	2.074	60.12	-0.045	2.364	57.08	-0.070	3.591	54.39	-0.108	3.802	51.99	-0.074	3.252	54.48
Lee's	-0.025	0.409	45.43	-0.015	0.297	41.14	-0.037	0.246	39.05	-0.077	0.510	39.76	-0.046	0.289	40.35
Kim's	0.002	0.199	27.29	-0.011	0.370	30.57	-0.027	0.345	31.58	-0.066	-0.792	31.49	-0.040	-0.437	30.24
Alg-1	-0.015	0.623	43.01	-0.021	0.907	43.26	-0.035	0.563	41.30	-0.078	0.984	39.65	-0.051	0.836	40.80
Alg-2	0.003	0.074	27.28	0.000	0.079	28.30	-0.054	-0.084	28.30	-0.034	-0.352	26.84	-0.032	-0.138	27.18
Football (SIF)															
No intra	-0.035	1.261	65.64	-0.043	2.450	62.27	-0.029	1.42	59.61	-0.022	2.22	55.61	-0.031	2.030	59.61
Lee's	-0.004	0.091	39.07	0.001	-0.069	25.54	-0.001	-0.011	18.76	0.001	0.102	19.87	-0.004	0.016	21.39
Kim's	-0.002	-0.958	29.41	-0.005	-1.180	31.93	-0.009	-0.800	33.04	-0.014	-1.005	31.15	-0.014	-1.083	32.04
Alg-1	-0.010	-1.228	41.43	-0.015	-1.286	47.93	-0.020	-0.864	47.02	-0.033	-1.015	43.77	-0.021	-1.055	46.24
Alg-2	0.001	-0.670	36.50	0.003	-0.693	38.41	-0.008	-0.283	37.83	-0.002	-0.175	34.90	-0.003	-0.463	36.04
Total average values															
No intra													-0.034	2.311	60.10
Lee's													-0.014	0.045	37.25
Kim's													-0.011	-0.397	34.64
Alg-1													-0.015	-0.358	46.79
Alg-2													-0.007	-0.231	35.37

Table 2: Performance comparison of the proposed algorithms on the JM 11.0 reference encoder for IPPP sequences (Ref. frames = 3)

	QP = 20			QP = 24			QP = 28			QP = 32			Average values		
Contents	ΔPSNR	ΔBits	ΔT	ΔPSNR	ΔBits	ΔT	ΔPSNR	ΔBits	ΔT	ΔPSNR	ΔBits	ΔT	ΔPSNR	ΔBits	ΔT
Stefan (QCIF)															
No intra	-0.048	2.224	42.16	-0.050	2.373	38.50	-0.046	5.073	35.84	-0.048	4.937	33.29	-0.064	3.622	37.45
Lee's	0.005	0.027	32.34	-0.005	-0.053	27.92	-0.006	-0.117	23.32	-0.017	-0.354	21.95	-0.007	-0.124	22.51
Kim's	0.008	-0.086	23.83	0.009	-0.110	25.78	-0.003	-0.336	24.37	-0.001	-0.750	27.44	-0.004	-0.263	24.51
Alg-1	-0.006	-1.008	32.98	-0.001	-1.068	32.52	-0.007	-0.945	30.41	-0.021	-0.750	27.44	-0.008	-1.239	30.84
Alg-2	0.004	-0.136	17.58	0.003	-0.129	19.79	0.005	-0.505	19.65	0.004	-0.103	17.85	-0.008	-0.299	18.72
Flower (CIF)															
No intra	-0.018	0.049	45.11	-0.015	0.026	42.31	-0.018	-0.001	39.90	-0.002	0.151	36.35	-0.013	0.056	40.91
Lee's	-0.016	-0.009	36.11	-0.012	0.006	31.19	-0.012	0.051	23.07	-0.006	0.126	20.58	-0.008	0.046	27.74
Kim's	-0.005	-0.388	18.97	0.003	-0.023	20.67	-0.003	-0.048	20.79	-0.007	-0.055	19.81	-0.002	-0.139	20.06
Alg-1	-0.01	-1.522	30.62	-0.002	-2.374	30.43	-0.026	-1.043	30.06	-0.045	-1.064	28.29	-0.037	-1.678	29.85
Alg-2	0.000	-0.080	11.13	0.006	0.040	15.37	0.000	0.053	16.03	-0.005	0.020	15.55	0.000	0.036	14.52
Mobile (CIF)															
No intra	-0.003	-0.012	48.31	-0.002	0.000	45.35	-0.006	0.020	41.53	-0.004	-0.098	37.57	-0.038	-0.022	43.19
Lee's	-0.003	-0.032	37.95	-0.002	0.064	31.43	-0.005	0.018	23.87	-0.001	-0.144	22.81	-0.001	-0.050	29.02
Kim's	-0.001	-0.579	26.68	0.004	-0.140	29.32	0.000	0.050	27.43	-0.003	-0.040	29.90	-0.002	-0.189	27.01
Alg-1	-0.012	-1.919	37.47	-0.001	-1.957	36.19	0.004	-0.420	33.51	-0.004	-0.489	29.90	-0.007	-1.328	34.27
Alg-2	0.000	-0.692	32.51	0.005	-0.218	34.20	-0.002	-0.148	31.67	0.001	-0.168	28.55	0.000	-0.415	31.73

Table 2: Continued

	QP = 20			QP = 24			QP = 28			QP = 32			Average values		
Contents	ΔPSNR	ΔBits	ΔT	ΔPSNR	ΔBits	ΔT	ΔPSNR	ΔBits	ΔT	ΔPSNR	ΔBits	ΔT	ΔPSNR	ΔBits	ΔT
Paris (CIF)															
No intra	-0.016	0.044	45.20	-0.008	0.092	41.38	-0.015	0.210	38.14	-0.005	0.352	35.10	-0.011	0.168	39.95
Lee's	-0.018	-0.008	41.02	-0.003	-0.085	36.08	0.003	0.019	32.38	0.001	0.034	29.94	-0.015	-0.003	34.86
Kim's	0.011	-0.105	22.87	-0.004	-0.422	24.46	0.015	-0.489	23.55	0.002	-0.740	21.86	-0.003	-0.591	23.19
Alg-1	0.010	-0.577	36.09	0.001	-0.991	33.76	0.005	-0.822	31.31	0.010	-1.044	28.55	-0.009	-0.859	32.43
Alg-2	0.016	0.064	27.08	-0.001	-0.196	28.18	0.002	-0.178	27.15	0.011	-0.441	24.85	-0.005	-0.411	26.82
Foreman (CIF)															
No intra	-0.048	1.730	36.44	-0.037	2.332	33.29	-0.052	2.503	30.81	-0.095	3.162	28.53	-0.058	2.432	32.27
Lee's	-0.025	0.470	27.37	-0.012	0.320	23.87	-0.034	0.142	22.22	-0.065	0.027	21.65	-0.041	0.274	23.78
Kim's	-0.002	0.032	16.85	0.005	0.424	18.20	-0.032	-0.019	17.92	-0.083	-1.418	17.53	-0.042	-0.810	17.58
Alg-1	-0.006	0.174	26.59	-0.004	-0.400	25.28	-0.096	-1.606	23.12	-0.142	-3.354	21.77	-0.080	-1.337	24.19
Alg-2	0.003	-0.046	15.69	0.003	0.012	15.80	-0.018	0.025	14.70	-0.040	-0.507	13.85	-0.026	-0.290	15.01
Football (SIF)															
No intra	-0.030	0.823	45.35	-0.029	0.928	40.87	-0.012	1.152	37.19	-0.009	1.514	32.69	-0.020	1.104	39.03
Lee's	-0.004	0.123	14.08	-0.003	0.017	17.45	0.000	-0.085	12.75	-0.004	-0.032	12.46	-0.003	0.013	14.18
Kim's	-0.003	-1.173	16.69	-0.006	-1.553	20.95	-0.005	-1.162	20.45	-0.018	-1.499	18.34	-0.015	-1.347	19.11
Alg-1	-0.019	-2.506	30.94	-0.018	-2.447	32.19	-0.007	-1.444	29.65	-0.029	-2.269	25.93	-0.040	-2.167	29.68
Alg-2	-0.001	-0.725	22.76	-0.005	-0.893	24.77	-0.014	-0.450	23.09	0.001	-0.431	20.14	-0.008	-0.669	22.69
Total Average values															
No intra													-0.034	1.226	38.80
Lee's													-0.012	0.026	25.34
Kim's													-0.011	-0.556	21.91
Alg-1													-0.030	-1.434	30.21
Alg-2													-0.011	-0.341	21.58

threshold value must be a weighted sum of the SAD values. In this study, Eq. 8 is still useful to deciding to skip the intra mode search. Thus, the proposed algorithms can speed-up the encoding time without considering the encoded bit rate.

CONCLUSION

We have proposed efficient intra mode SKIP detection algorithms based on the Macro-Block (MB) tracking scheme and the Rate-Distortion (RD) cost of the tracked MB for inter-frames in H.264/AVC video coding.

The proposed algorithms yield good performance because of adaptive RD thresholding schemes using the RD costs of the most correlated MB or image region that was tracked in the reference frame for the current MB. Based on comparative analysis, a speed-up factor of 45~52% was verified with a negligible bit increment and a minimal loss of image quality.

ACKNOWLEDGEMENTS

This research was supported by the MKE (The Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) Support program supervised by the NIPA (National IT industry Promotion Agency) (NIPA-2009-C1090-0902-0020).

REFERENCES

Byung-Gyu, K., 2008. Novel inter-mode decision algorithm based on macroblock (MB) tracking for the P-slice in H.264/AVC video coding. IEEE Trans. Circuits Syst. Video Technol., 18: 273-279.

Cheng, C.C. and T.S. Chang, 2005. Fast three step intra prediction algorithm for 4×4 blocks in H.264. Proceedings of the IEEE International Symposium on Circuits and Systems, May 23-26, Kobe, Japan, pp: 1509-1512.

Cheng, Y., Z.J.G. Wmg and K. Dai, 2005. Research on intra modes for inter-frame coding in H.264. Proceedings of the IEEE 9-th International Conference on Computer Supported Cooperative Work in Desi, May 24-26, Guilin China, pp: 740-740.

Fu, F.X.L. and L. Xu, 2004. Fast intra prediction algorithm in H.264-AVC. Proceedings of the 7th International Conference on Signal Processing, Aug. 31-Sept. 4, Beijing, China, pp: 1191-1194.

Kim, B.G. and J.H. Kim, 2006. Efficient intramode skip detection algorithm for H.264/AVC video encoder. Optical Eng., 45: 097008-097008.

La, B., M. Eom and Y. Choe, 2007. Fast mode decision for intra prediction in H.264/AVC encoder. Proc. IEEE Int. Conf. Image Process., 5: 321-324.

Lee, J. and B. Jeon, 2004. Fast mode decision for H.264. Proceedings of the IEEE International Conference on Multimedia and Expo, June 30, Taipei, pp: 1131-1134.

Li, S., X. Wei, T. Ikenaga and S. Goto, 2007. Architecture and memory: A VLSI architecture design of an edge based fast intra prediction mode decision algorithm for H.264/AVC. Proceedings of the 17th Great Lakes Symposium on Great Lakes Symposium on VLSI, March 11-13, Stresa-Lago Maggiore, Italy, pp: 20-24.

Pan, F., X. Lin, S. Rahardja, K.P. Lim, Z.G. Li, D. Wu and S. Wu, 2005. Fast mode decision algorithm for intraprediction in H.264/AVC video coding. IEEE Trans. Circuits Syst. Video Technol., 15: 813-822.

Rehman, J.U.Z.Y. and A. Kumarayapa, 2007. Fast intra prediction mode decision by adaptively selecting fewer number of modes. Proceedings of the 6th International Conference on Machine Learning and Cybernetics, July 12-15, Hong Kong, pp: 2385-2389.

Rui Su, G.Liu and T. Zhang, 2006. Fast mode decision algorithm for intra prediction in H.264/AVC. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, May 14-19, Toulouse France, pp: 921-924.

Shengfa, Y. and Z.Z.C. Zhenping, 2006. Efficient intra mode prediction decision for MPEG-2 to H.264/avc transcoding. Proceedings of the 8th International Conference on Signal Processing, Nov. 16-20, Guilin China, pp: 16-20.

Wang, C.C., T.S. Chen and C.W. Tung, 2006a. Fast intra-mode decision in H.264 using interblock correlation. Proceedings of the IEEE International Conference on Image Processing, Oct. 8-11, Atlanta, USA., pp: 1345-1348.

Wang, J.F., J.C. Wang, J.T. Chen, A.C. Tsai and A. Paul, 2006b. A novel fast algorithm for intra mode decision in H.264/AVC encoders. Proceedings of the IEEE International Symposium on Circuits and Systems, May 21-24, Island of Kos, pp: 3498-3501.

Wiegand, T., G.J. Sullivan, G. Bjontegard and A. Luthra, 2003. Overview of the H.264/AVC video coding standard. IEEE. Trans. Circuits Syst. Video Technol., 13: 560-576.