Edge-Enhancement in Intra Prediction Algorithm for H.264

Maha Sharkas, Mohamed E. Khedr, Abdusalam Z. Adham

Abstract— In an H.264 Advanced Video Coding system, an intra-prediction tool that uses neighbor picture elements within the current picture has been adopted for enhancing compression efficiency. In this paper, a new efficient H.264 intra prediction scheme is proposed. The new prediction scheme is called Modified Prediction Matrix Mode (MPMM). The main idea behind the proposed prediction algorithm is to differentiate between the best suitable intra predictions pixels, {vertical - horizontal}. The performance of our proposed modification is evaluated using Mat Lab code. The results show that our MPMM enhances the Peak Signal to Noise Ratio (PSNR) in addition to giving a clue for the object's edge pattern.

Index Terms— H.264, intra-prediction, prediction mode, Peak Signal to Noise Ratio.

1 INTRODUCTION

For communication of video information, in order to adapt to network environment of limited bitrate, redundant components reduction of video information is important part for all streaming media encoding technology. Predictive coding is common technique for most streaming media systems, because the image information has high similarity in time and spatial, predictive results close to true value can be achieved by predicting current data according to former and/or adjacent data. Then the complete image can be reconstructed by means of transmitting fewer samples values, the predicted difference value and predicting method, the information size will be much smaller than original image file size.

The video coding standard H.264/AVC, which was finalized by Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG [1], achieved great success in low bitrate coding of intra frame, thanks to the precise prediction carried out by the intra frame prediction algorithms. Intra prediction is employed to reduce the spatial correlation in a single frame of video sequence.

H.264 achieves better coding efficiency than the previous video codecs standards. However, this coding gain comes with a significant increase in computational complexity. Moreover the video compression efficiency achieved in H.264 standard is not a result of any single feature but rather a combination of a number of encoding tools, hence requires much processing power.

Maha sharks, Arab Academy for Science, Technology, and Maritime Transport, Electronics and Communications Engineering Department, P.O.B.1029, Alexandria, Egypt, msharkas@aast.edu
M. khedr, Arab Academy for Science, Technology, and Maritime Transport, Electronics and Communications Engineering Department, P.O.B.1029, Alexandria, Egypt, khedr@aast.edu
Abdusalam z. adham, Arab Academy for Science, Technology, and Maritime Transport, Electronics and Communications Engineering Department, P.O.B.1029, Alexandria, Egypt, absi87@gmail.com

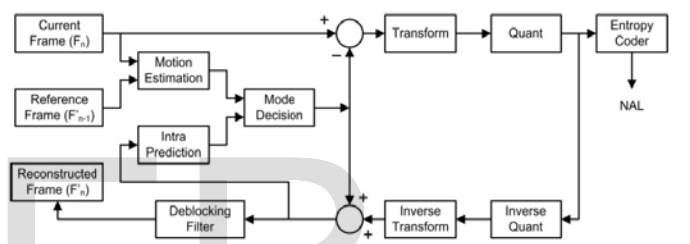

Fig.1 top-level block diagram of an H.264 encoder.

As it is shown in the top-level block diagram of an H.264 encoder in fig.1, one of the main tools is the intra prediction algorithm used in the baseline profile of H.264 standard [2].
Intra prediction algorithm generates a prediction for a Macro block (MB) based on spatial redundancy. At present, H.264 CODECs provide intra prediction for 4x4 up to 32x32 block size. In this paper, however, we will only focus on 4x4 block intra prediction [3].

2 H.264 INTRA PREDICTION

2.1 H.264 4X4 INTRA PREDICTION ALGORITHM

Intra prediction can greatly reduce redundancy of image information in space. Intra prediction algorithm predicts the pixels in a MB using the pixels in the available neighboring blocks. For the luminance (luma) component of a MB, a 16x16 predicted luma block is formed by performing intra predictions for each 4x4 luma block in the MB and by performing intra prediction for the 16x16 MB. There are nine prediction modes for each 4x4 luma block and four prediction modes for a 16x16 luma block.

A mode decision algorithm is then used to compare the 4x4 and16x16 predictions and select the best luma prediction mode for the MB.

ISSN 2229-5518

As shown in fig.2 (a), there are nine 4x4 luma prediction modes designed in a directional manner. Each prediction mode generates 16 predicted pixel values using some or all of the neighboring pixels A to M fig.2 (b). The arrows indicate the direction of prediction in each mode. The pixels A to M belong to the neighboring blocks and are assumed to be already encoded and reconstructed and are therefore available in the encoder and decoder to generate a prediction for the current MB [4].

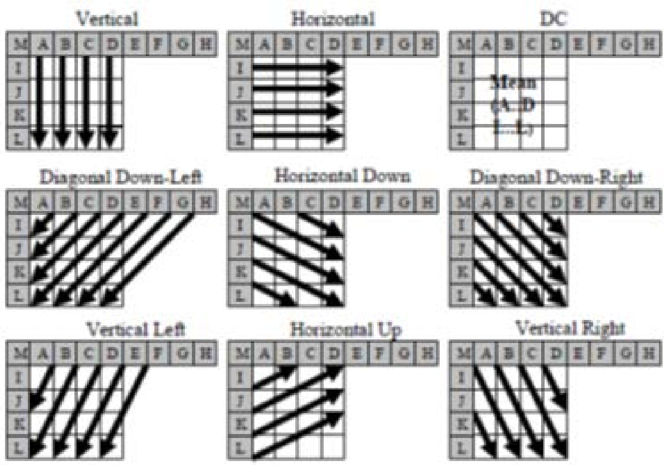

Fig.2 (a) nine 4x4 prediction modes

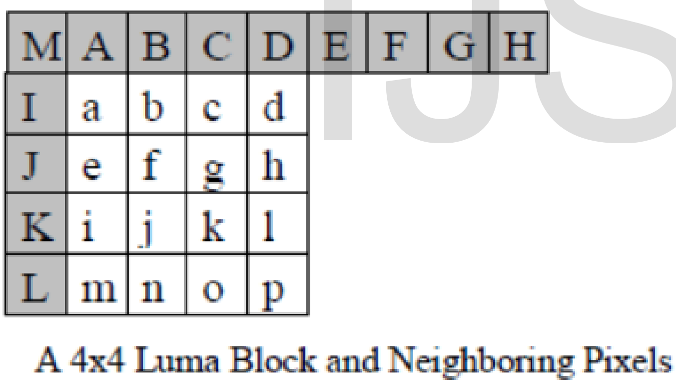

Fig.2 (b) A-M neighboring pixels for 4x4 block

The predicted pixels are calculated by a weighted average of the neighboring pixels A-M for each mode except Vertical and Horizontal modes where the same values of pixels are copied.

For gray level image, each macro-block only have brightness information, there are two ways of predictions:
16x16 pixel blocks and 4x4 pixel blocks, where 16x16 has 4 prediction mode and 4x4 has 9 prediction modes.
For 4x4, the 9 prediction modes are from mode0 to mode8, except mode2 which needs DC prediction (Mean value of reconstructed pixel gray level is adopted as gray level of all 16 pixels), the rest 8 prediction mode corresponds to a specific prediction direction. For current block, each pixel on this direction is evaluated according to reconstructed pixel on edge by a certain weighted formula, and then the model with minimum deviation is adopted as prediction model of the current 4x4 block.
For 16x16, the 4 prediction modes are from mode0 to mode3, which are vertical prediction, the horizontal prediction, the DC prediction and the Plane prediction. The previous 3 models have same prediction way with the 4x4 block. Plane prediction will receive the prediction value of gradual change form upper left corner to lower right corner, and it will be used for area with flat change of gray level, also the model with minimum deviation is adopted.
The deviation of all pixels by various prediction for current macro-block are compared, then which of 16x16 and 4x4 prediction will be chosen on one of the comparison techniques.

The basic process of intra prediction is as follows:
• Determine available information circumjacent macro block around the current macro-block, including the availability of macro-block on top, left and upper right of the current macro-block, (see Table 1).
• Dividing brightness macro-block into 16 blocks of 4x4, using 9 different prediction models, and calculating (SAD) for each 4x4 block in turn.
SAD (Sum of Absolute Difference) represents the summation of the absolute difference between predicted and actual value for each pixel of 4x4 sub-block, this measure the weight of correlation in the block.

After the calculation of prediction pixels, a number of cost functions that determine the quality of a prediction can be used. The most simple and common cost function is Sum of Absolute Differences (SAD) formulated by:

$$SAD(Orig, Pred_m) = \sum_{y=1}^{4}\sum_{x=1}^{4} |Orig(x,y) - Pred_m(x,y)| \qquad (1)$$

Where Orig and Predm are the original block and the predicted block, respectively. The prediction mode with the minimal SAD cost value is selected as the best mode [5].

TABLE I.
AVAILABILITY OF 4X4 LUMA PREDICTION MODES

Availability of Neighboring 4x4 Luma Blocks	Available 4x4 Luma Prediction Modes
None available	DC
Left available, Top not available	HORZ, DC, HORZ_UP
Top available, Left not available	VERT_RIGHT, DC, VERT_LEFT, DDL
Both available	ALL MODES

ISSN 2229-5518

2.2 IMAGE EDGE DETECTION TECHNOLOGY

Edge detection technology is adopted to detect local linear structure of the image, usually as a preprocessing step of image segmentation.
The edge of the image represents boundary between two regions with different average gray level. This boundary area contains a high frequency components, by the meaning of visualization, it represents a sudden change in pictures' texture or occurring of an object.

3 THE FEASIBILITY OF EDGE DETECTION FOR INTRA PREDICTION DATA

H.264 intra prediction process concern on evaluating the residual deviation after predicting the un-coded pixels according to the degree of similarity between pixels in the image. The residual data presents the deviation degree between current pixel and edge pixels on prediction direction. The larger the residual data is, the lower the degree of similarity between the current pixel and edge pixels is, which means the poor continuity. The goal of edge detection is to find the pixels with discontiguous gray level. After intra prediction of the image, the discontiguous pixels will cause larger prediction deviation in residual data.

For the 4x4 intra prediction mode of H.264, except Model 2, prediction value of each pixel for the current block is received by linear predictions on basis of gray level on edge of current block according to the specific direction. The prediction direction has the smallest prediction deviation should correspond to the direction of image texture (edge), because only on the direction of the texture, correlation between pixels is most obvious, the prediction value is the closest to actual value. The following can be deduced:

- If the prediction deviation of each pixel for current block are very small, it can present that the edge for actual pixels of the current block does not exist or the edge exists along the prediction direction of the block.
- If the prediction deviation of some pixels for current block are large, it can present that the edge deviate from prediction direction exist for actual pixels of the current block.

For 16x16 intra prediction mode, its larger prediction range determine that its prediction accuracy lower than 4x4 prediction mode, so 16x16 prediction mode commonly used in the image with smoother (more continuous) area.
The two deductions are also tunable for 16x16 prediction model.
Considering that there will be no complex boundary for pixel blocks of size 4x4, we might assume that in each 4x4 block, there is only one edge (in fact most of the cases should be in that way), so that the edge can be determined by simple threshold analyze for each pixels of the 4x4 block.
Even for the 4x4 block with more complex edge, the detected deviation on edge will not have large effect on overall result [6].

4 PROPOSED SYSTEM FOR EDGE-BASED ADAPTIVE INTRA PREDICTION

H.264/AVC employs intra prediction to reduce spatial redundancy between neighboring blocks. Different directional prediction modes are used to cater diversified video content. Basically, even after intra prediction, the residue still contains a lot of edge or texture information. Unfortunately, these high frequency components consume a large quantity of bits and the distortion is usually quite high. Based on this drawback, an Edge-based Adaptive Intra Prediction is proposed to reduce the residue energy especially for the edge region.

This paper proposes an enhancement procedure to improve the prediction at the high frequency parts of the picture – edge prediction enhancement. The proposed algorithm starts with edge detection procedure, accordingly, block prediction is performed taking edges into consideration.

As analyzed following, H.264 exploits the spatial correlation between adjacent blocks to predicate the pixels in the current block. When the current block is at the edge of moving objects or on the boundary between two objects, the correlation between the current block C and blocks A B becoming smaller than usual fig.3. So that the prediction accuracy decreased and the residual energy and bit rate increased. In fact, the correlation among the pixels in current block is bigger than that with adjacent blocks. If we take other pixels as the reference pixels to predict the rest pixels in the current block, this residue can be reduced and the prediction accuracy improved, thus the PSNR improved.

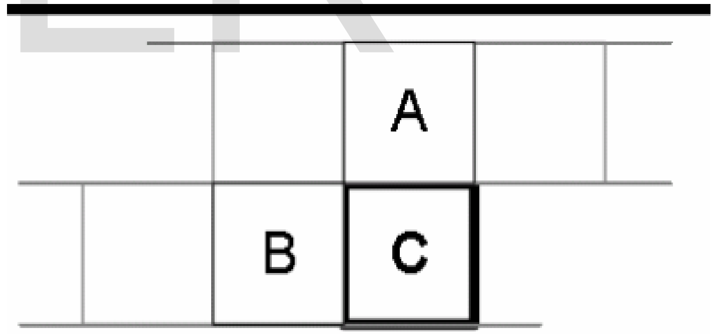

Fig.3 current block and neighbor blocks

4.1 NEW INTRA PREDICTION MODE

In many familiar video sequences, most of the texture directions are vertical and horizontal. So the Vertical mode and Horizontal mode are set to mode 0 and mode 1 respectively, in H.264.
Take both the prediction directions and correlations among the adjacent pixels in the same block into consideration, the new intra predictions are list as follows:

Mode 0: Vertical

ISSN 2229-5518

a=e=i=m =ref[x+2, y]
b=f=j=n =ref[x+2, y]
c=g=k=o =ref[x+2, y]
d=h=l=p =ref[x+2, y]

Mode 1: Horizontal
a=b=c=d =ref[x, y+2]
e=f=g=h =ref[x, y+2]
i=j=k=l =ref[x, y+2]
m=n=o=p =ref[x, y+2]

The proposed work found that; best value used as a base for predicting rest of the block is the median value of the entire block; therefore, instead of just measuring the SAD of each prediction mode, we have first to choose the best basis pixels for prediction.

To illustrate these words by means of values:
Let's work on an individual vector of the block, and examine two scenarios of block fields first a smooth field and secondly a progressive change field.
First scenario a vector of [1,2,3,4,5] fig.4 (a).

Conventional prediction mode will choose the first pixel as the base value for predicting the rest of the vector, such that:
Predicted vector is [1,1,1,1,1] fig.4 (b).

Upon this decision the sum of absolute error would be:
Error vector is [0,1,2,3,4] fig.4 (c). ==> Error value 10
On the other hand if an algorithm is concerning of finding the median value of this vector, it will choose [3] as a base sample rather than [1]. Then the predicted vector would be [3,3,3,3,3] fig.4 (d).
In this case the error vector as sum of absolute difference would be: Error vector is [2,1,0,1,2] fig.4 (e). ==> Error value 6
The second scenario is when there is a rapid change in values; this case illustrates an existence of object boundary within the block field.
So a block vector such: [1,2,7,8,10] fig.5 (a).
If a base sample chosen such conventional algorithm = [1]
Predicted vector is [1,1,1,1,1] fig.5 (b). Resultant error is [0,1,6,7,9] fig.5 (c). ==> Error value 23
But for proposed algorithm, the base –median- value would be [7]
Predicted vector is [7,7,7,7,7] fig.5 (d). Resultant error is [6,5,0,2,3] fig.5 (e). ==> Error value 16
Therefore, the proposed algorithm is concerning of locating the better base values for predicting the rest of the block with least sum of absolute difference.

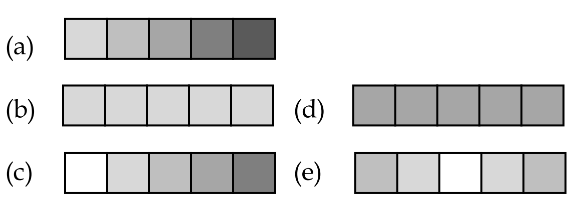

Fig.4 smooth texture prediction

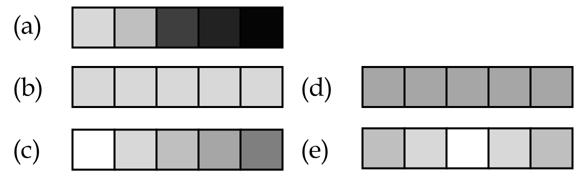

Fig.5 smooth texture prediction

5 EXPERIMENTAL RESULTS

Matlab7.0 is adopted to conduct simple fast edge detection and simulate intra prediction scheme, and the result is compared with traditional technique.
Test target image presented in fig.6 (a)Light House, (b)Tulips, (c)Koala, (d)Penguins, (e)Lena and (f)Foreman to facilitate the processing and examine variety of object shapes, intra prediction mode is forced to be the 4x4 model.

It is obviously the improvement of PSNR for all tested samples, the enhancement in average exceeds 10% better than traditional prediction mode. Moreover the results introduces another benefit from this technique, where the object boundaries could be clearly predicted, and this clue may employed to enhance other properties for H.264.
Wang et al. [7] proposed a simple dominant edge strength based algorithm to reduce the number of probable modes for complexity reduction. [7], each image block was divided into four sub-blocks and the edge strengths of each direction were derived from corresponding filtering operation.
Although algorithms [7][8] achieved much better results, the drawback of these two algorithms is that each sub-block of current block is represented by a single pixel value, so the results cannot exactly exhibit the edge direction of the current block. La et al. [9] used dominant edge direction to make a fast mode decision. Lin et al. [10] developed a simple algorithm based on direction detection of the edge inside the block. The above algorithms have limitations in coding performances and time consumption.

As represented in table.2, the proposed scheme always win the comparison when compared to the normal prediction mode. For instance in the Tulips image (figure a) where it consists of obvious vertical objects, we find the modified prediction mode increased the PSNR from 64.97 to 71.95. In addition, if results produced from horizontal mode (55.18) compared with results of vertical mode (64.97), its higher PSNR indicates that image contains vertically aligned objects.
Another image of Koala (fig.6 (c)) contains random spots in its texture rather than obvious boundaries, also the modified mode exhibited a better reconstructed image power of (54.69) (50.44) for modified and normal horizontal mode, respectively. And (50.26) (54.66) for vertical mode. The comparison also represented in graph fig.7. Where the bars represents the PSNR of each picture from the experiment set with different prediction and selection modes.

ISSN 2229-5518

Fig.6 (a) Light House

Fig.6 (b) Tulips

Fig.6 (c) Koala

Fig.6 (d) Penguins

Fig.6 (e) Lena

Fig.6 (f) Foreman

Table.2 IntraFrame Prediction Enhanced Edge prediction – PSNR for Edge prediction

Image	Mode H	Mode V	Mode Hi	Mode Vi	Object boundary
LightHouse	53.93	51.75	59.00	55.92	Horizontal Edges
Tulips	55.18	64.97	62.04	71.95	Vertical Edges
Koala	50.44	50.26	54.69	54.66	Random spots
Penguins	47.69	49.97	52.22	54.29	Spots + Vertical
Lena	51.63	58.79	51.07	65.17	Curved Edges
Foreman	52.45	53.78	59.41	57.70	Curved Edges

ISSN 2229-5518

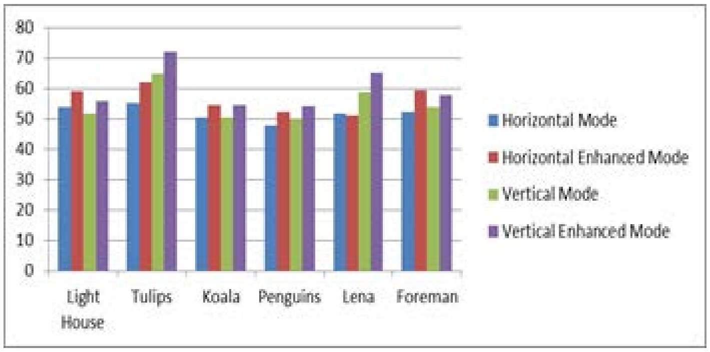

Fig.7 bars represent the PSNR of each picturefor proposed sheme against conventional prediction sheme

6 CONCLUSION

Traditional intra prediction scheme achieves quite high coding efficiency. However, due to the limitation of the traditional intra prediction methods, there is still large residue around edges or texture region, which degrades the RD performance.

In order to further exploit the spatial redundancy, an Edge-based Adaptive Intra Prediction is proposed to reduce the residue energy for edge regions. Compared with H.264/AVC, PSNR improvement can be achieved.

However, the proposed model enhanced the PSNR, its coding efficiency of the overhead data is still an open issue. Sometimes, the large overhead information eat up the improvement made by enhancing the prediction accuracy.

REFERENCES

[1] Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, "Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 ISO/IEC 14496-10 AVC)", JVT-G050, March 2003.

[2] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, “Overview of the H.264/AVC Video Coding Standard”, IEEE Trans. on CAS for Video Technology, vol. 13, no. 7, pp. 560–576, July 2003.

[3] Miaohui Wang, King Ngi Ngan, and Long Xu, " Efficient H.264/AVC Video Coding with Adaptive Transforms", IEEE Transactions on Multimedia,

[4] Ref: **Pixel Similarity Based Computation and Power Reduction Technique**

[5] Ref: Intra Prediction Architecture with Macro block/Block Reordering Milica Orlandi´c

[6] Ref: A Improved Image Edge Detection Algorithm Based on H.264 Intra Prediction

[7] J. F. Wang, J. C. Wang, J. T. Chen, A. C. Tsai, A. Paul, "A Novel Fast Algorithm for Intra Mode Decision in H.264/AVC Encoders," in Proc. IEEE Int. Symp. Circuit Syst., May 2006.

[8] An-Chao Tsai, Jhing-Fa Wang, Wei-Guang Lin, and Jar-Ferr Yang, “A Simple and Robust Direction Detection Algorithm for Fast H.264 Intra Prediction” IEEE International Conference on Multimedia and Expo, pp.1587-1590, July 2007.

[9] B. La, M. Eom, and Y. Choe, “Fast mode decision for intra prediction in H.264/AVC encoder,” International Conference of Image Processing, pp.321-324, Sept. 2007.

[10] W.G. Lin, A.C. Tsai, J.F. Wang, and J.F. Yang, “A simple direction detection algorithm for fast H.264 intra prediction,” Proc. IEEE TENCON, pp.1-4, Oct. 2007.