Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG
(ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6)
7th Meeting: Pattaya II, Thailand, 7-14 March, 2003

Document: JVT-G016
Filename: JVT-G016.doc

Title:	Fast Motion Estimation for JVT
Status:	Input Document to JVT
Purpose:	Proposal

Author(s) or Contact (s): Zhibo Chen, Peng Zhou, Yun He
Dept. of Electronic Engineering, Tsinghua University, Beijing, China
GuoZhong Wang, SVA, Shanghai
Tel: 86-10-62781413
Email: chenzhibo@tsinghua.org.cn
hey@tsinghua.edu.cn

Source: Tsinghua University, China and SVA Group, China

1. Introduction

In Awaji meeting, we proposed a hybrid Unsymmetrical-cross Multi-Hexagon-grid Search (UMHexagonS) algorithm for integer pel motion estimation together with a Center-Biased Fractional Pel Search (CBFPS) algorithm for JVT [1].
The proposed algorithm proposed in [1] shows very good capability in keeping the rate distortion performance (maximum distortion less than 0.1 dB) for different sequences from QCIF format to HD (High Definition) format with different motion degree, as well as a great computation reduction up to 90% compare to Fast Full Search used in reference software can be achieved.

And in this proposal we refined two parts in UMHexagonS algorithm proposed in [1]: initial search point prediction and early termination. By refinement on these two parts, averagely more than 90% (up to 95.54%) of computation reduction and averagely 0.04dB (maximum less than 0.096dB) PSNR drop compared with that of Fast Full search algorithm adopted in JM reference software. Another advantage is that a tradeoff between search speed and reconstructed quality can be achieved by purposefully change a modulation factor in out algorithm.
We will describe the refined two parts in the following sections: initial search point prediction and early termination. The main body of the UMHexagonS algorithm is the same that described in [1][2].

2. Initial search point prediction

Initial search point prediction is an important technique introduced by many fast motion estimation algorithms [3,4] setting the search area around the MBD (Minimum Block Distortion) point of the whole search window in order to improve the performance of motion estimation. Median prediction as described in [6] is frequently used in many algorithms and standard. Motion vectors of the collocated block in the previous frame and of the spatially adjacent blocks are also used in [3] as initial search point predictors. According to the multiple reference frames and multiple block modes adopted in JVT, we proposed four kinds of prediction modes in this proposal:

a) Median Prediction

As Fig.1 shows, median predictor is used in median prediction of motion vectors, the median value of the adjacent blocks on the left, top, and top-right (or top-left) of the current block is used to predict the motion vector of the current block (as Fig.1 shows):

$$M\vec{V} = median(Mv\_\vec{A}, Mv\_\vec{B}, Mv\_\vec{C})$$

The predicted motion vector is $(MV_{pred\_MP}(x), MV_{pred\_MP}(y))$. Some rules specify the predicted motion vector value has been defined in [6]: when block A lies outside the picture or GOB(Group of Blocks) boundary, it is replaced by (0,0), when block C lies outside the picture or GOB boundary, it is replaced by motion vector of block D, when two blocks B and C lie outside, however, they are replaced by the motion vector of the third block.

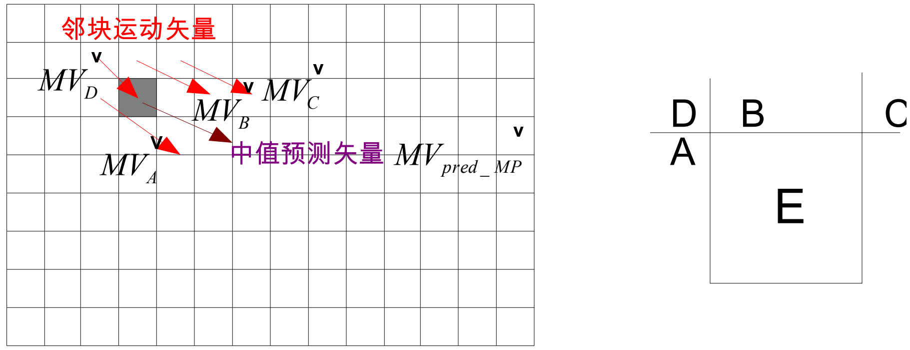

Fig.1 Spatial Median Prediction of motion vectors

b) UpLayer Prediction

As Fig.2 shows, there are seven inter prediction block modes defined in H.264. 8x8 modes (mode 4, 5, 6, 7) are first searched followed by 16x16 modes (mode 3, 2, 1) in current reference software[5]. Such a strategy is not beneficial in utilizing the motion relationship between different modes, therefore the search order of the modes here is changed according to the size of the block mode, a hierarchically search order from mode 1 to 7 is chosen as our mode search order and the motion vector of the up layer block (for example, mode 5 or 6 is the up layer of mode 7, and mode 4 is the up layer of mode 5 or 6, etc.) is used as one of the prediction candidates of lower layer, just as Fig.3 demonstrates.

$$\vec{MV}_{pred\_UP} = \vec{MV}_{UpLayer}$$

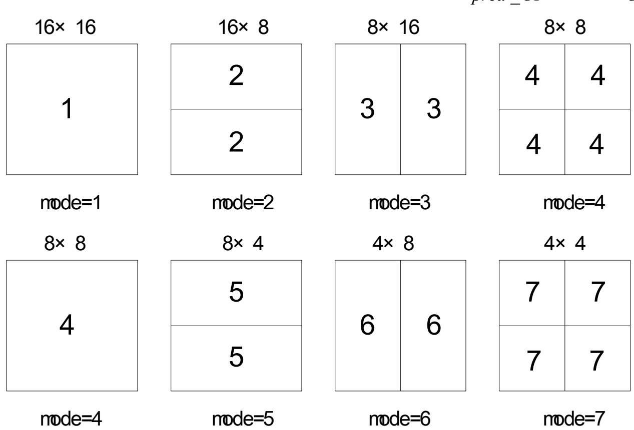

Fig.2 Seven Prediction Blocks Modes in H.264

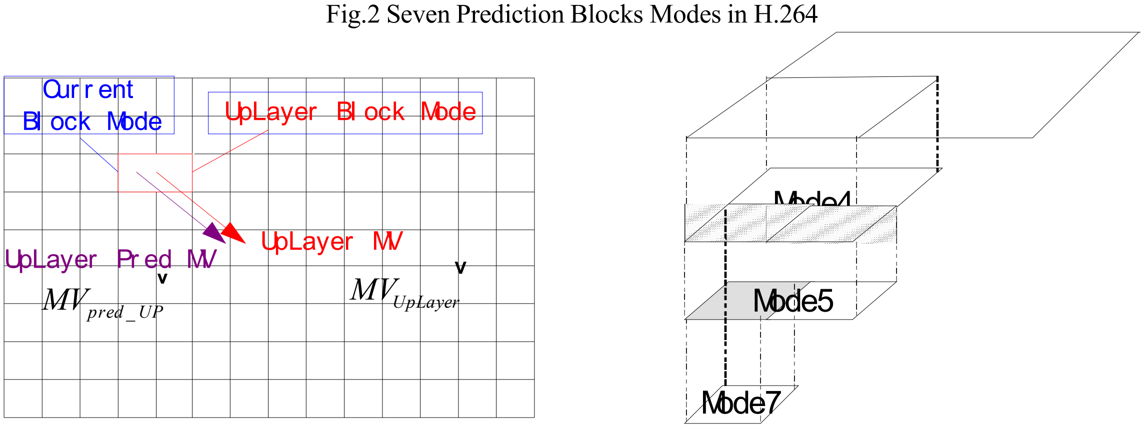

Fig.3 Spatial UpLayer Prediction of motion vector

c) Corresponding-block Prediction

For natural video sequence, the motion track of a moving object is continuous except scene change occur, therefore there is strong correlation of motion vector in the temporal domain, and then we utilize this property to give an accurate starting search position.

In this prediction mode, the motion vector of the corresponding block in the last frame is used as one motion vector candidate, as Fig.4 shows:

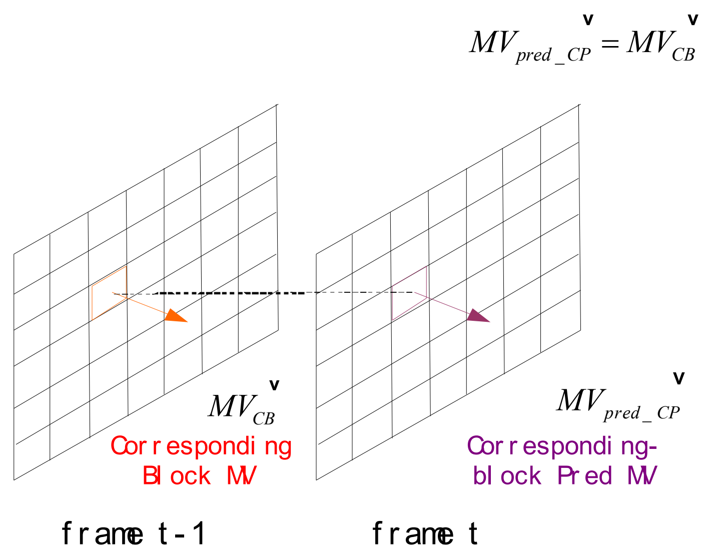

Fig.4 Temporal Corresponding-block Prediction of motion vector

d) Neighboring Ref-frame Prediction

Multi-reference frames motion compensation is adopted in JVT to increase prediction accuracy and coding efficiency. For the same current block, motion vectors in different reference frames exhibit a strong correlation in our experiment. Therefore current block's motion vector in reference frame t' can be predicted by scaling of current block's motion vector in reference frame $t'+1$, as Fig.5 shows:

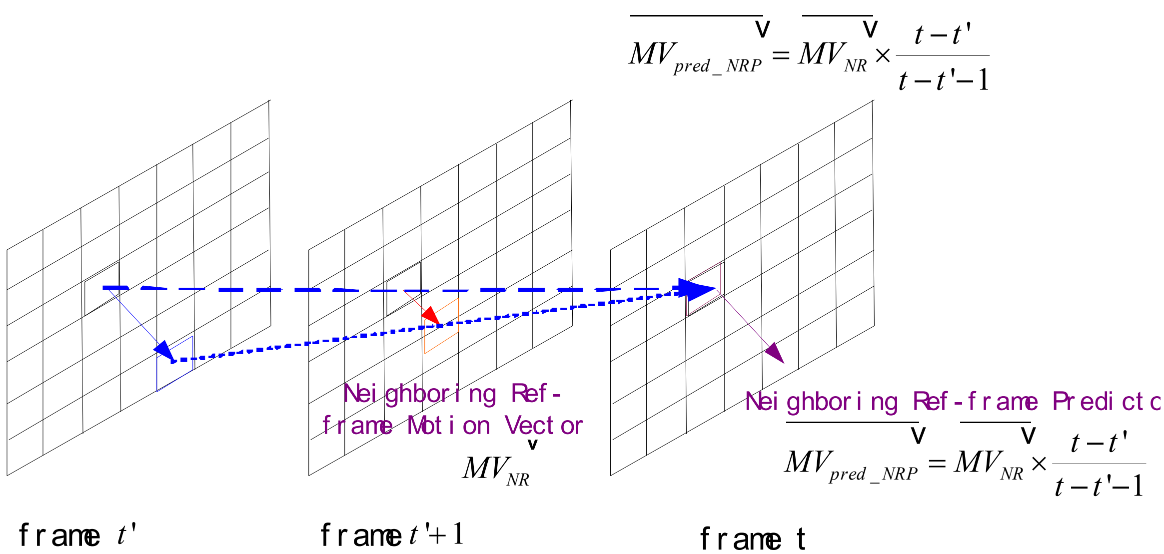

Fig.5 Temporal Neighboring Ref-frame Prediction of motion vector

It's clear that MV_{pred_UP} can not be used as prediction candidate for Mode 1(16x16 mode), MV_{pred_NRP} can not be used if the current reference frame is zero, and MV_{pred_CP} is not able to be used if the current frame is the firs P frame. Therefore in our algorithm: for mode 1, the median predictor, the (0, 0) vector and the motion vectors of adjacent blocks on the left, top, and top-right are chosen as the prediction candidates; and for other modes, the median prediction, the (0, 0) vector and motion vector of the up layer are chosen as the vector prediction candidates. MV_{pred_NRP} and MV_{pred_CP} are used as predictor when they are available.

The prediction with the minimum cost among these prediction candidates will be chosen as the initial search position of next search step. And the detail of the search strategy can be found in [1].

3. Threshold of Early Termination

3.1 Prediction of SAD

Similar to motion vectors, SAD of spatial adjacent or temporal adjacent blocks should also have a high correlation. Three spatially adjacent blocks (left, top, top-right) and the collocated block in the previous frame as predictors are used in [3]. And similar to initial search point prediction, we propose four kinds of prediction mode for SAD prediction in this proposal, according to the multiple reference frames and multiple block modes adopted in JVT.

These four modes are Median Prediction (MP) mode, Uplayer Prediction (UP) mode, Corresponding-block Prediction (CP) mode and Neighboring Ref-frame Prediction (NRP) mode.

a) Median Prediction

Similar to motion vector median prediction, as Fig.1 shows: A, B, C are the neighboring blocks around current block and corresponding MV are MV_A , MV_B , MV_C, and corresponding SAD are SAD_A , SAD_B , and SAD_C. Therefore we can predict the SAD by:

$$SAD_{pred\_MD} = \min(SAD_{Vx\_median}, SAD_{Vy\_media})$$

$$SAD_{Vx\_median} = f_{SAD}(Vx\_median)$$

$$Vx\_median = Median(MV_A(x), MV_B(x), MV_C(x))$$

where

$$SAD_{Vy\_median} = f_{SAD}(Vy\_median)$$

$$Vy\_median = Median(MV_A(y), MV_B(y), MV_C(y))$$

and the definition of $f_{SAD}()$ is:

$$\begin{cases} f_{SAD}(MV_A(x)) = f_{SAD}(MV_A(y)) = SAD_A \\ f_{SAD}(MV_B(x)) = f_{SAD}(MV_B(y)) = SAD_B \\ f_{SAD}(MV_C(x)) = f_{SAD}(MV_C(y)) = SAD_C \end{cases}$$

b) Uplayer Prediction

Similar as that described in motion vector Uplayer Prediction, we use the SAD of uplayer to predict SAD of current block mode:

$$SAD_{pred\_UP} = SAD_{UpLayer} / 2$$

c) Corresponding-block Prediction

Assuming current block is in frame t, and the reference frame is t'. Then the corresponding-block in last frame do motion estimation in frame $t'-1$, the SAD obtained is SAD_{Co}. Therefore SAD_{Co} can be used as current SAD predictor:

$$SAD_{pred\_CP} = SAD_{Co}$$

d) Neighboring Ref-frame Prediction

Assuming current block is in frame t, and the reference frame is t' ($t' < t-1$). And the SAD of current block searched in frame $t'+1$ is SAD_{NR}, which can be used SAD predictor:

$$SAD_{pred\_NRP} = SAD_{NR}$$

3.2 Introduction of modulated factor β

Two goals in introducing early termination are: a) the earlier the termination can be achieved, the better; b) be sure that the best matching position will not be missed or the termination motion vector is very close to the best one. First a new variable named *NSD* (*Normative SAD Difference*) is defined to describe the property of SAD difference of as:

$$NSD = (SAD_{pred} - SAD_{best}) / SAD_{pred}$$

Assuming the minimum SAD in the search window is SAD_{best}, and the predicted SAD according to spatial and temporal correlation of SAD is SAD_{pred}, and we define a threshold SAD_{thrh} for early termination with: $SAD_{thrh} = SAD_{pred}(1+\beta)$, where β is a modulated factor and will play a key role of keeping a tradeoff between reconstructed quality and search speed. The search strategy will stop if SAD of current MBD SAD_{exit} is smaller than SAD_{thrh}. Commonly there is a relation revealed like this: $SAD_{best} \le SAD_{exit} \le SAD_{thrh}$, and there should be no sense if $SAD_{thrh} \le SAD_{best}$.

To be sure that there will not be much quality degradation, we define a conservative threshold ΔTH, and we assume that the motion vector is accurate enough and the reconstructed quality will not degrade if:

$$SAD_{exit} - SAD_{best} < \Delta TH$$

and because $SAD_{exit} \le SAD_{thrh}$, therefore if inequation：

$$SAD_{thrh} - SAD_{best} < \Delta TH$$

is satisfied, then inequation (1.6) is also satisfied.

Residual pixel values at the input of the transform may be approximated by a Laplacian distribution with zero mean [7] and a separable covariance $r(m,n)=\sigma_{f^2}\rho^{|m|}\rho^{|n|}$, where m and n are the horizontal and vertical distances between two pixels, respectively, σ_{f^2} is the variance of the residual pixel values, and ρ ($|\rho|<1$) is the correlation coefficient. Then the transform can be described as:

$$F = AfA^T$$

where A is the transform matrix，f is the residual signal matrix，as well as the input of the transform. Then the variance of $(u,v)th$ transform coefficient $\sigma_F{}^2[u,v]$ can be written as [6]:

$\sigma_{F^2}[u,v]=\sigma_{f^2}[ARA^T]_{u,u}[ARA^T]_{v,v}$ ，where：

$$R = \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}$$

Without much influence on the end results, we can choose $\rho = 0.6$ from the experiment. Then we can deduce the relation between variance of spatial domain and frequency domain:

$$\sigma_F{}^2[u,v] = \sigma_f{}^2 * M[u,v], u = 0 \sim N; v = 0 \sim N$$

The integer transform adopted in JVT is based on 4x4 DCT transform but with some differences. Without loss of accuracy, we can deduce the M matrix based on 4x4 DCT:

$$M = \begin{bmatrix} 5.607424 & 2.125210 & 1.060864 & 0.678503 \\ 2.125210 & 0.805453 & 0.402067 & 0.257152 \\ 1.060864 & 0.402067 & 0.200704 & 0.128365 \\ 0.678503 & 0.257152 & 0.128365 & 0.082099 \end{bmatrix}$$

Assuming the quantization step is Q_{step}, then for two values of $(u,v)th$ transform coefficient, if the difference is satisfied with inequation (1.12), then this difference will not make difference in the reconstructed quality:

$$\sigma_{F1}{}^2[u,v] - \sigma_{F2}{}^2[u,v] < Q_{step}$$

from (1.10) and (1.12), we can get:

$$(\sigma_{f1}{}^2 - \sigma_{f2}{}^2) * M[u,v] < Q_{step}, u = 0 \sim 3; v = 0 \sim 3;$$

Based on the assumption that the residual signal is Laplacian distribution with zero mean[], we can approximate the variance in the spatial domain like this:

$$\sigma_f = \sqrt{2} * MMAE = \sqrt{2} * (SAD / B_{size})$$

where B_{size} is the current block size.

Here define σ_{f1} as the variance of spatial residual signal when SAD equals SAD_{thrh}, and σ_{f2} is the variance of the spatial residual signal when SAD equals SAD_{best}, then we can get:

$$\frac{2}{(B_{size})^2} * [SAD_{thrh}{}^2 - SAD_{best}{}^2] < \frac{Q_{step}}{M[u,v]}, u = 0 \sim 3; v = 0 \sim 3$$

$M[0,0]$ is the largest one, therefore if the inequation comes into existence when $[u,v] = [0,0]$, then it will come into existence for other $[u,v]$ values. Then we can assume $SAD_{best} \approx SAD_{thrh}$, :

$$\frac{4 * SAD_{thrh}}{(B_{size})^2} * [SAD_{thrh} - SAD_{best}] < \frac{Q_{step}}{M[0,0]}$$

then：

$$SAD_{thrh} - SAD_{best} < \frac{Q_{step}}{M[0,0]} * \frac{(B_{size})^2}{4 * SAD_{thrh}}$$

therefore the conservative threshold ΔTH we need is：

$$\Delta TH = \frac{Q_{step}}{M[0,0]} * \frac{(B_{size})^2}{4 * SAD_{thrh}}$$

Then we can study the property of modulated factor β in the following paragraph:
Because $SAD_{thrh} = SAD_{pred} * (1 + \beta)$, put it into inequation(1.16), we can get:

$$SAD_{pred}(1 + \beta) - SAD_{best} < \frac{Q_{step}}{M[0,0]} * \frac{(B_{size})^2}{4 * SAD_{thrh}}$$

both sides are divided by SAD_{pred}, and assume $SAD_{thrh} \approx SAD_{pred}$, then:

$$\beta < \frac{Q_{step}}{M[0,0] * 4} * (\frac{B_{size}}{SAD_{pred}})^2 - NSD$$

For JVT, the quantization step of DC value is:

$$Q_{step} = (2^{qbits} - f) / QE[q_{rem}][0][0]$$

Where $qbits = 15 + QP/6$, and $q_{rem} = QP\%6$, $f = (1 << qbits)/6$, QE is the defined quantization coefficient table and QP is the input quantization parameter [].
Following conclusions can be drawn from inequation(1.20)：

- The larger Q_{step} is，the larger β is. When quantization step is larger, the quantization noise is also larger and most of the details of reconstructed image are lost. In this case the difference between the best matching block and the sub-optimal matching block becomes blurry and it's possible to increase the value of β for speed up without much loss in reconstructed quality;
- The larger $\frac{SAD_{pred}}{B_{size}}$ is, the smaller β is. It's clear that $\frac{SAD_{pred}}{B_{size}}$ represents the energy of the residual signals. The detail of the image will increase if $\frac{SAD_{pred}}{B_{size}}$ increase and the matching error surface inside a search window will become more complex and it's very easy for motion search dropping into a local minimum. Therefore β is decreased in this case to keep the reconstructed image quality with the cost of speed.

- Statistical property of NSD can be analysed from experiments, and β can be chosen according to the distribution of NSD:

$$\beta_1 < \frac{Q_{step}}{M[0,0]*4}*(\frac{B_{size}}{SAD_{pred}})^2 - \alpha_1$$

$$\beta_2 < \frac{Q_{step}}{M[0,0]*4}*(\frac{B_{size}}{SAD_{pred}})^2 - \alpha_2$$

where: $P(NSD < \alpha_1) = 0.8$ and $P(NSD < \alpha_2) = 0.9$. Then inequation (1.22) and (1.23) separately have the probability of 80% and 90% to come into existence. Whereafter we are able to choose two SAD threshold:

$$SAD_{thrh\_1} = SAD_{pred}(1+\beta_1)$$

$$SAD_{thrh\_2} = SAD_{pred}(1+\beta_2)$$

Different skip policy will be used for termination by different threshold. Where SAD_{thrh_1} is larger and has a lower probability to ensure that the vector is accurate enough, therefore a more complex vector refinement algorithm (e.g. EHS [1]) should be used after the termination. And for SAD_{thrh_2}, which is smaller and has a higher probability to ensure the accuracy of the motion vector, a simple vector refinement algorithm (e.g. sub-steps with diamond search pattern inside EHS [1]) can be adopted to make the motion vector more accurate.

4. Simulation Results:

The experiment is taken on JM5.0c and the latest version JM6.1a separately. Typical sequences defined in common condition[10], some high motion sequences such as Stefan & Bus, and some HD sequence are tested. Our simulation results show a better performance than that of [1].
Average results computation defined in [9] are also included in our comparison. And these results were generated using an excel equivalent in our results due to its convenience and the result is very similar to that of [9]. Time-used statistics generated from the encoder is used, and an accurate timer function QueryPerformanceCounter() is used for calculating part running time of integer pel motion estimation and fractional pel motion estimation. Our experiments show that this calculation is stable and accurate enough for analysis the speed of each part.
Main parameters used in our experiments are: Hadamard transform is used, Search Range is set as 32 and OutFile Mode is 0, RD Optimization is used.

Our experiments results are all in JVT-D016.xls and a short description is given in the following tables. Averagely more than 90% (up to 95.54%) of computation reduction and averagely 0.04dB (maximum less than 0.096dB) PSNR drop compared with that of Fast Full search algorithm adopted in JM reference software.

In these tables: “Total” means the time saving in total time of the encoder, “Integer” means time saving for the integer pel motion estimation part, and “Fractional ” means time saving in the fractional pel motion estimation part. “Avg.bits” represents the average equivalent bit increasing under the same PSNR and “Avg.PSNR” represents the average equivalent PSNR loss under the same bit rate.

4.1 Experiment results on sequences defined in common condition:
For each test sequences, reference frame number equals 1 or 5 are tested separately and entropy coding equals CAVLC or CABAC are also tested separately.

Tabel 1. Results of sequences defined in common condition

		Total	Integer	Fractional	Avg. bits	Avg. PSNR(dB)
Ref =1	CAVLC	53.54%	93.11%	34.18%	+0.93%	-0.046
	CABAC	51.92%	93.21%	34.16%	+0.75%	-0.040
Ref =5	CAVLC	74.02%	92.99%	33.80%	+0.69%	-0.036
	CABAC	74.29%	93.24%	33.62%	+0.69%	-0.036

4.2 Experiment results on sequences with high motion:

We choose Stefan (CIF format, 30f/s, 300f) and Bus sequence (SIF format, 30f/s, 298f) as the typical high motion sequences in this test. And only P frame is used in this experiment. For each test sequences, reference frame number equals 1 or 5 are tested separately and entropy coding equals CAVLC or CABAC are also tested separately.

Tabel 2. Results of sequences with high motion

		Total	Integer	Fractional	Avg. bits	Avg. PSNR(dB)
Ref =1	CAVLC	51.07%	91.91%	33.38%	+1.28%	-0.056
	CABAC	50.59%	91.93%	33.10%	+0.91%	-0.045
Ref =5	CAVLC	72.69%	91.57%	31.93%	+0.84%	-0.039
	CABAC	72.42%	91.43%	31.67%	+0.81%	-0.036

4.3 Experiment results on B frames:

We choose Stefan (CIF format, 30f/s, 300f) and Bus sequence (SIF format, 30f/s, 298f) as the test sequence and 2 B frames are inserted between two P frames or I frame and P frame. Here in our experiment only the first frame is I frame.

Tabel 3. Results on B frame

		Total	Integer	Fractional	Avg. bits	Avg. PSNR(dB)
Ref =1	CAVLC	58.73%	91.15%	31.76%	-1.18%	+0.064
	CABAC	58.65%	91.02%	31.24%	+0.01%	-0.002
Ref =5	CAVLC	71.54%	89.78%	31.14%	-0.64%	+0.034
	CABAC	71.66%	89.86%	30.64%	+0.86%	-0.041

4.3 Experiment results on HD sequences:

Two HD sequences defined in [8] are test in our experiment: “Night” sequence is a sequence with Night scene, everything from no to fast motion, high dynamic. “Spincalendar” sequence is a sequence with studio scene of a spinning calendar, with various levels of detail. Number of reference frame is 5, the entropy coding method is CABAC, search range is 32 for “Night” sequence and 64 for “Spincalendar”.

Tabel 4. Results of HD Sequence

	Total	Integer	Fractional	Avg. PSNR(dB)
Night	66.48%	89.80%	30.01%	-0.007
SpinCalendar	85.67%	93.80%	30.36%	+0.011

5. Conclusion

Based on our former proposal [1], in this proposal two refined parts of UMHexagonS algorithm [1] are introduced: initial search point prediction and early termination. By refinement on these two parts, Averagely more than 90% (up to 95%) of computation reduction and averagely 0.04dB (maximum less than 0.096dB) PSNR drop compared with that of Fast Full search algorithm adopted in JM reference software. Comparing with the UMHexagonS algorithm proposed in [1], an averagely computation reduction ratio up to 40% and an average PSNR loss less than 0.01dB can be achieved. Another advantage is that a tradeoff between search speed and reconstructed quality can be achieved by purposefully change a modulation factor β in out algorithm, e.g. if α_1 and α_2 are chosen according to $P(NSD < \alpha_1) = 0.6$ and $P(NSD < \alpha_2) = 0.7$, then the threshold for early termination will increase and the search speed will increase with the cost of degradation of reconstructed image quality.

6. Reference

[1] Zhibo Chen, Peng Zhou, Yun He, “Fast Integer Pel and Fractional Pel Motion estimation in for JVT”, JVT-F017r1.doc, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, 6th meeting, Awaji, Island, JP, 5-13 December, 2002

[2] Zhibo Chen, etc. “Hybrid Unsymmetrical-cross Multi-Hexagon-grid Search Strategy for Integer Pel Motion Estimation in H.264”, Picture Coding Symposium, 2003, April, Saint_Malo, France

[3] Hye-Yeon Cheong Tourapis, Alexis Michael Tourapis, Pankaj Topiwala, “Fast Motion Estimation within the JVT codec”, JVT-E023, 5th Meeting: Geneva, Switzerland, 09-17 October, 2002

[4] Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, “Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG”, JVT-F100d2.doc, 6th meeting, Awaji, Island, JP, 5-13 December, 2002

[5] JM6.1, Reference Software of JVT

[6]A. K. Jain, “Fundamentals of Digital Image Processing”. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[7] I-Ming Pao; Ming-Ting Sun, “Modeling DCT coefficients for fast video encoding” Circuits and Systems for Video

Technology, IEEE Transactions on , Volume: 9 Issue: 4 , Jun 1999 , Page(s): 608 -616
[8] G. Sullivan, "Recommended Simulation Common Conditions for H.26L Coding Efficiency Experiments on Low-Resolution Progressive-Scan Source Material", document VCEG-N81, Sep. 2001.

[9] G. Bjontegaard, "Calculation of average PSNR differences between RD-Curves", document VCEG-M33, Mar. 2001

[10]"Description of new High Definition Video Test Sequences", ISO/IEC JTC1/SC29/WG11 MPEG02/ W5125 October 2002, Shanghai

(Append for Proposal Documents)

JVT Patent Disclosure Form

International Telecommunication Union Telecommunication Standardization Sector | International Organization for Standardization | International Electrotechnical Commission

Joint Video Coding Experts Group - *Patent Disclosure Form*

(Typically one per contribution and one per Standard | Recommendation)

Please send to:

JVT Rapporteur Gary Sullivan, Microsoft Corp., One Microsoft Way, Bldg. 9, Redmond WA 98052-6399, USA
Email (preferred): Gary.Sullivan@itu.int Fax: +1 425 706 7329 (+1 425 70MSFAX)

This form provides the ITU-T | ISO/IEC Joint Video Coding Experts Group (JVT) with information about the patent status of techniques used in or proposed for incorporation in a Recommendation | Standard. JVT requires that all technical contributions be accompanied with this form. *Anyone* with knowledge of any patent affecting the use of JVT work, of their own or of any other entity ("third parties"), is strongly encouraged to submit this form as well.

This information will be maintained in a "living list" by JVT during the progress of their work, on a best effort basis. If a given technical proposal is not incorporated in a Recommendation | Standard, the relevant patent information will be removed from the "living list". The intent is that the JVT experts should know in advance of any patent issues with particular proposals or techniques, so that these may be addressed well before final approval.

This is not a binding legal document; it is provided to JVT for information only, on a best effort, good faith basis. Please submit corrected or updated forms if your knowledge or situation changes.

This form is *not* a substitute for the *ITU ISO IEC Patent Statement and Licensing Declaration*, which should be submitted by Patent Holders to the ITU TSB Director and ISO Secretary General before final approval.

Submitting Organization or Person:	
Organization name	**Tsinghua University, China and SVA Group, China**
Mailing address	**100084 425# East mail building, Tsinghua University, Beijing, China**
Country	**China**
Contact person	**Zhibo Chen, Yun He**
Telephone	+86-10-62781413
Fax	+86-10-62770317
Email	chenzhibo@tsinghua.org.cn, hey@tsinghua.edu.cn
Place and date of submission	March/03/2003
Relevant Recommendation | Standard and, if applicable, Contribution:	
Name (ex: "JVT")	**JVT**
Title	Fast Motion Estimation for JVT
Contribution number	**JVT-G016**

(Form continues on next page)

Disclosure information – Submitting Organization/Person (choose one box)

☐ 2.0 The submitter is not aware of having any granted, pending, or planned patents associated with the technical content of the Recommendation | Standard or Contribution.

or,

The submitter (Patent Holder) has granted, pending, or planned patents associated with the technical content of the Recommendation | Standard or Contribution. In which case,

☐ 2.1 The Patent Holder is prepared to grant – on the basis of reciprocity for the above Recommendation | Standard – a free license to an unrestricted number of applicants on a worldwide, non-discriminatory basis to manufacture, use and/or sell implementations of the above Recommendation | Standard.

☐ 2.2 The Patent Holder is prepared to grant – on the basis of reciprocity for the above Recommendation | Standard – a license to an unrestricted number of applicants on a worldwide, non-discriminatory basis and on reasonable terms and conditions to manufacture, use and/ or sell implementations of the above Recommendation | Standard.

Such negotiations are left to the parties concerned and are performed outside the ITU | ISO/IEC.

☒ 2.2.1 The same as box 2.2 above, but in addition the Patent Holder is prepared to grant a "royalty-free" license to anyone on condition that all other patent holders do the same.

☐ 2.3 The Patent Holder is unwilling to grant licenses according to the provisions of either 2.1, 2.2, or 2.2.1 above. In this case, the following information must be provided as part of this declaration:

- patent registration/application number;
- an indication of which portions of the Recommendation | Standard are affected.
- a description of the patent claims covering the Recommendation | Standard;

In the case of any box ***other than 2.0*** *above, please provide the following:*

Patent number(s)/status ______________________________

Inventor(s)/Assignee(s) ______________________________

Relevance to JVT ______________________________

Any other remarks: ______________________________

(please provide attachments if more space is needed)

(form continues on next page)

Third party patent information – fill in based on your best knowledge of relevant patents granted, pending, or planned by other people or by organizations other than your own.

Disclosure information – Third Party Patents (choose one box)

☐ 3.1 The submitter is not aware of any granted, pending, or planned patents *held by third parties* associated with the technical content of the Recommendation | Standard or Contribution.

☐ 3.2 The submitter believes third parties may have granted, pending, or planned patents associated with the technical content of the Recommendation | Standard or Contribution.

For box 3.2, please provide as much information as is known (provide attachments if more space needed) - JVT will attempt to contact third parties to obtain more information:

3rd party name(s) ______________________

Mailing address ______________________
Country ______________________
Contact person ______________________
Telephone ______________________
Fax ______________________
Email ______________________
Patent number/status ______________________
Inventor/Assignee ______________________
Relevance to JVT ______________________

Any other comments or remarks: