

文章编号:1001-9081(2009)05-1362-03

基于 H. 264 的多参考帧快速运动估计算法

王莉莉¹, 黄晓革¹, 张明²

(1. 电子科技大学 电子工程学院, 成都 610054; 2. 空军工程大学 工程学院, 西安 710038)
(surpass0327@163.com)

摘要:在 H. 264 标准中,为提高图像质量和压缩效率,编码器支持 7 种块类型在多个参考帧中搜索最佳运动矢量。参考代码采用对参考帧逐一搜索的方式,极大地增加了计算复杂度。根据相邻帧间运动矢量的相关性和连续性,使用合成运动矢量来预测最佳的匹配位置。同时,结合分块模式的终止准则可以避免不必要的参考帧搜索。实验表明,较 JVT 参考模型该算法编码速度能提高 6 倍以上,同时保证了图像的高质量 and 低比特率。

关键词:H. 264; 运动估计; 多参考帧; 终止准则

中图分类号:TP391 **文献标志码:**A

Fast motion estimation algorithm for H. 264 with multi-references

WANG Li-li¹, HUANG Xiao-ge¹, ZHANG Ming²

(1. School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu Sichuan 610054, China;
2. School of Engineering, Air Force University of Engineering, Xi'an Shaanxi 710038, China)

Abstract: In H. 264 standard, the codec allows seven block types to search the best motion vector in several reference frames to achieve lower bit rate and higher quality. Since the reference code uses the full search scheme to obtain the best performance, it will increase the computational complexity significantly. Synthesis motion vectors were adopted to predict the optimal matching location according to correlation and continuity of motion vectors among adjacent frames. Meanwhile, adaptive criteria related to selected MB's mode were used to determine whether it is necessary to search more reference frames. The simulation results show that the speed of the proposed algorithm is over six times faster than that of the original scheme adopted in JVT reference software with similar video quality and low bit rate.

Key words: H. 264; motion estimation; multiple reference frames; half-stop criteria

0 引言

作为新一代的国际视频编码标准,H. 264^[1] 采纳了众多先进技术用以进一步提升压缩效率,其中之一就是多参考帧的运动估计。对于拍摄过程中镜头抖动、场景快速切换、物体边缘遮盖与重现等情况下,将运动搜索范围从原来的一个参考帧扩展到多个解码后的参考帧,通常能找到更精确的匹配块,从而有助于获得更高的编码效率。

在 JVT 参考模型中采用对每个参考帧逐一搜索的方法,计算复杂度随参考帧数目呈线性增加,又由于编码块共有 7 种模式(16×16、16×8、8×16、8×8、8×4、4×8 和 4×4),每个模式都要进行运动估计,计算量非常大。文献[2]首先用一个粗尺度机制搜索运动矢量,并根据结果选择最佳参考帧,然后在最佳帧中做细尺度的估计;文献[3-4]利用帧间运动矢量的连续性,采用合成运动矢量作为搜索起点,极大地降低了计算复杂度。文献[5]讨论了终止准则在多参考帧搜索的应用,文献[6]则进一步考虑了量化系数对阈值的影响,当搜索得到的残差小于给定阈值,则表明变换量化之后的残差系数全为零,可以跳过后续参考帧的估计。文献[7]重点研究了针对硬件实现时的多参考帧优化方案。

H. 264 支持多参考帧和多种分块模式,而各个参考帧和各种模式对编码效率的提升作用是不一样的,上述算法大都还没有考虑编码器的这一特性,对所有参考帧和所有块模式都

使用相同的搜索策略。因此,有必要根据多参考帧的特点来寻找更高效的运动估计算法。下面所有内容都是基于 P 帧来讨论的,相关结论可扩展到 B 帧中。

1 多参考帧特征分析

在运动估计时,运动矢量和参考帧选择为匹配因子 J 最小时对应的情形:

$$J = SAD(MV, Ref) + \lambda(R(MV) + R(Ref)) \quad (1)$$

其中, MV 为运动矢量, Ref 为参考帧, λ 为拉格朗日乘子, $R(MV)$ 、 $R(Ref)$ 分别为编码运动矢量和参考帧标号所用的比特数, SAD 表示编码块与参考帧中匹配块的绝对误差和 (Sum of Absolute Differences):

$$SAD(MV, Ref) = \sum_{m=1}^M \sum_{n=1}^N |f_k(m, n) - f_{Ref}(m + dx, n + dy)| \quad (2)$$

式中, f_k 和 f_{Ref} 分别表示当前帧和参考帧, $M \times N$ 为块尺寸, $MV = (dx, dy)$ 为位移矢量。

H. 264 标准中可以使用 1~16 个参考帧,图 1 显示了 foreman 序列在不同数量参考帧下的编码情况, QP 为 28, IPPP 结构。可见,不同位置的参考帧对提高编码效率的作用是不同的。随着距离的增加,较远参考帧对提高编码效率的作用逐渐降低。因此通常情况下,距离当前帧越近的参考帧相关性越强,被选择运动补偿的可能性也就越高。

收稿日期:2008-11-18;修回日期:2009-02-03。

作者简介:王莉莉(1983-),女,山东莱芜人,硕士研究生,主要研究方向:图像处理、视频压缩;黄晓革(1967-),男,四川成都人,副教授,博士研究生,主要研究方向:信号处理、计算机视觉;张明(1981-),男,四川阆中人,博士研究生,主要研究方向:图像处理、视频压缩。

考虑一般的 P 帧情形,令当前编码帧为帧 n , 帧 $n - i$ ($1 \leq i \leq N$) 为参考帧队列, N 为最大参考帧数。帧 $n - 1$ 选中的概率远大于其余参考帧,因此可以采用传统的单参考帧运动估计快速算法。而对于其他参考帧,可以按与当前帧距离由近到远的顺序依次搜索,这样做有两个优点:1) 运动矢量场在时空域上是高度相关的,可以用前面参考帧的搜索结果用于后续参考帧,预测搜索起点,加快运算速度;2) 一般而言,越靠后的参考帧选中的概率越小,当前面参考帧的估计结果足够准确时,可以终止后续帧的运动估计过程。

本文就 $n - i$ ($2 \leq i \leq N$) 帧的运动估计算法进行探讨,主要包含两个方面:运动矢量复用和终止条件。

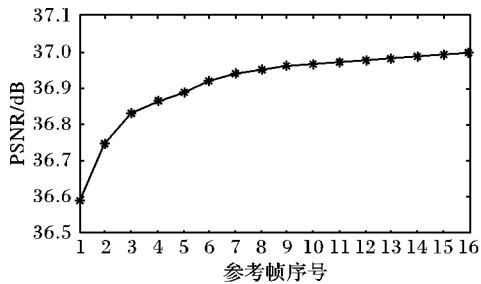


图 1 不同参考帧下 foreman 的编码结果

2 多参考帧快速运动估计算法

2.1 运动矢量复用

图像序列相邻帧之间有着很强的相关性,正是利用这种相关性,人们相继提出了许多快速搜索算法,如三步法、菱形法(DS)等。但是在多参考帧预测中,如果参考帧和当前帧距离较远,相关性就比较弱,直接运用这些快速算法就难以得到理想的效果。

在常见的快速运动估计算法中,都用到了邻块运动矢量作为搜索的起始位置,但在多参考帧情形下,由于终止条件的运用(见后),有可能邻块在该参考帧并没有进行搜索,因而必须寻找新的预测运动矢量。

通常运动物体在连续多个帧中都保持相似的运动状态,其内部宏块在多个参考帧中也必然有着很强的相关性。如图 2 所示,当前帧 n 的宏块 O 在参考帧 $n - k_1$ 的匹配块为 B ,运动矢量 $MV_{n-k_1}^{-k_1}$, B 在帧 $n - k$ 的运动矢量为 $MV_{n-k_1}^{-(k-k_1)}$, 则 O 在帧 $n - k$ 的运动矢量 MV_n^{-k} :

$$MV_n^{-k} \approx MV_{n-k_1}^{-k_1} + MV_{n-k_1}^{-(k-k_1)}; 1 \leq k_1 < k \quad (3)$$

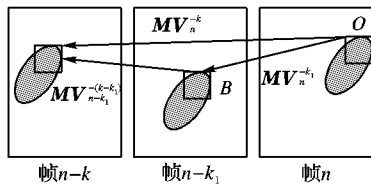


图 2 复合运动矢量

通常情况下 $MV_{n-k_1}^{-(k-k_1)}$ 不能预先得到。如图 3 所示,匹配块 B 往往覆盖 4 个宏块区域,只要编码帧 $n - k_1$ 时保存了其宏块在帧 $n - k$ 的运动矢量,就可以用 4 个覆盖宏块的运动矢量与 $MV_{n-k_1}^{-k_1}$ 相加,从而得到多个合成运动矢量用来预测搜索起点,即:

$$MV_n^{-k} \approx MV_{n-k_1}^{-k_1} + MV_{n-k_1}^{-(k-k_1)}(i); 1 \leq k_1 < k, i \in [0,3] \quad (4)$$

本文算法选择 $k_1 = k - 1$, 则式(4)变为:

$$MV_n^{-k} \approx MV_n^{-(k-1)} + MV_{n-(k-1)}^{-1}(i); 1 \leq k_1 < k, i \in [0,3] \quad (5)$$

这样编码过程中只需要保存宏块在第一个参考帧的运动矢量 $MV_n^{-1}(x,y)$ 即可, x,y 为宏块在高度和宽度上的标号。

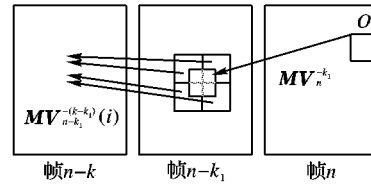


图 3 运动矢量复用

按照式(1)计算所有合成运动矢量及原点处的 J 值,选择使 J 值最小的运动矢量作为搜索起点。搜索起点在绝大多数情形下都非常靠近最佳匹配点,此时使用图 4 所示的小菱形模板反复搜索,直到中心点最佳匹配为止。

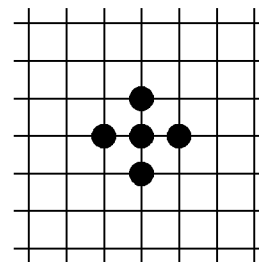


图 4 小菱形模板

2.2 终止准则

每搜索一个参考帧后,都要检测终止条件,如果满足终止条件,就结束后续的参考帧搜索过程。本文提出了三种检测条件。

2.2.1 阈值检测

如果能够找到一个合适的阈值 T , 一旦某个参考帧搜索完毕后 J 小于该阈值,则直接终止后续帧的搜索过程,可以大幅度减少算法的计算时间。

阈值的选择直接关系到算法的搜索速度和运动估计的精度,在很大程度上依赖于视频编码的具体环境(如纹理特性、码率等)。为了适应各种编码情况,利用宏块的空域相关性自适应地确定阈值。如图 5 所示,编码过程中宏块按扫描顺序编码,同一帧图像中与当前块相邻的已编码块共有 4 个,分别位于当前块的左方、上方、右上方和左上方,匹配度分别记为 J_A, J_B, J_C, J_D 。本文选取 T :

$$T = \text{median}(0, J_A, J_B, J_C, J_D) \quad (6)$$

$\text{median}()$ 表示中值。

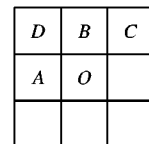


图 5 宏块的已编码邻块

2.2.2 不同模式间多参考帧选择的相关性判断

本文在编码过程中,改变了模式计算的顺序。首先计算 $16 \times 16, 8 \times 8, 4 \times 4$ 三种模式,后续的 $16 \times 8, 8 \times 16, 8 \times 4, 4 \times 8$ 等模式计算就可以根据前三种模式的结果做最大参考帧数限制。

考虑 16×8 和 8×16 情形,每个块既是 16×16 块的 $1/2$, 又是两个 8×8 子块的父块。因此可推断, 16×8 和 8×16 模式下各块的最佳参考帧最有可能分布在 16×16 模式和其两个 8×8 模式块的最大参考帧之内,其分布区间可表示如下:

$$\begin{cases} Ref_{16 \times 8}(i) \in [Ref_0, \max(Ref_{16 \times 16}, Ref_{8 \times 8}(2i), \\ Ref_{8 \times 8}(2i + 1))] \\ Ref_{8 \times 16}(i) \in [Ref_0, \max(Ref_{16 \times 16}, Ref_{8 \times 8}(i), \\ Ref_{8 \times 8}(i + 2))] \end{cases}; i = 0, 1 \quad (7)$$

Ref_0 是第一个参考帧。同理,对于 8×4 和 4×8 情形,可以将其最佳参考帧限制在 Ref_0 和 $8 \times 8, 4 \times 4$ 的最大参考帧之间。

2.2.3 运动矢量精度检测

假设同一物体在连续三帧 $n-2, n-1, n$ 的位置分别位于整像素、 $1/2$ 像素、整像素位置。编码帧 $n-1$ 时,运动补偿由 $n-2$ 帧的参考像素六抽头插值得到,插值过程不可避免地会使物体边缘模糊。则编码帧 n 时,帧 $n-2$ 可以提供更好的补偿效果。因此,编码块和参考块的相对位移精度也影响着最佳参考帧的选择。

H. 264 标准中运动矢量按其精度可分为整像素、 $1/2$ 像素和 $1/4$ 像素运动矢量。因此,算法中记录整像素、 $1/2$ 和 $1/4$ 像素精度运动矢量出现的标志,当三种精度的运动矢量全都出现时,终止后续参考帧的搜索过程。这里隐含的条件是:在相同精度的运动矢量下,较近参考帧的预测效果更好一些。

2.3 结合编码模式的终止准则

上面提出了 3 种判断条件,由于 H. 264 支持 16×16 到 4×4 等 7 种块模式,实际编码中大块模式占有较大的比例,因而对不同的块模式使用了不同的终止条件。对大块模式的终止条件较为苛刻的条件,从而搜索更多的参考帧,运动矢量更为精确;而小块模式的终止条件较为宽松,容易提前结束搜索过程。从而降低计算复杂度的同时,保证了编码图像的质量。具体如下:

1) 对于 16×16 和 8×8 模式,当同时满足阈值检测条件和矢量精度条件时,结束搜索过程;

2) 对于 4×4 模式,只要满足阈值检测条件和矢量精度条件之一时,结束搜索过程;

3) 对于 16×8 和 8×16 ,满足阈值检测条件、矢量精度条件、模式相关性条件中两项条件时,结束搜索过程;

4) 对于 8×4 和 4×8 模式,满足阈值检测条件、矢量精度条件、模式相关性条件中之一时,结束搜索过程。

2.4 算法流程

本文算法步骤如下:

1) 对第一个参考帧 ($n-1$) 进行运动估计,得到运动矢量 MV_n^{-1} ;

2) i 从 2 到 k , 对于参考帧 ($n-i$):

① 使用 2.3 节所述的终止准则(跟块模式有关),如果满足终止条件,则结束搜索过程;

② 根据式(5)得到四个合成运动矢量,根据式(1)计算这四个矢量及零运动矢量处的代价值 J , 选择使 J 最小的运动矢量作为搜索起点;

③ 以搜索起点为中心,反复使用图 4 所示的小菱形模板搜索,当匹配点位于模板中心时,完成该帧的运动估计过程。

步骤 1) 中,对于第一帧的搜索可以使用常规的运动估计算法,得到的运动矢量 MV_n^{-1} 需要保存,用于后续帧的编码。由于 H. 264 支持 7 种块模式,因此编码器中要保存 7 种块模式的 MV_n^{-1} 。

3 实验结果

为了验证新算法的性能,在相同条件下对参考模型 JM11.0 和本文算法进行了计算机仿真实验。选择 CIF 标准序列 akiyo、container、coastguard、foreman 和 mobiloe 的前 100 帧测试,实验中分别选用 5 和 10 个参考帧,IPPP 结构,量化系数 QP 为 28,宏块的搜索范围为 16×16 。采用了两种对比算法:全搜索法(Full Search, FS)和非对称十字型多层次六边形格点搜索算法(UMHexagonS)^[8]。本文算法第一个参考帧使用 UMHexagonS 算法。

实验结果如表 1、2 所示。其中 NCC 为标准化的计算复杂度(Normalized Computational Cost)。

$$NCC = \frac{\sum_{mode} N_{pts}(mode) \times W(mode)}{N_{mode} \times N_{ref}} \quad (8)$$

$N_{pts}(mode)$ 表示 $mode$ 模式下总的搜索点数, N_{mode} 为总的编码模式数, $N_{mode} = 7$, N_{ref} 为参考帧数量(不包含第一个参考帧), $W(mode)$ 为不同模式下对搜索点数的权重值,当编码为 16×16 模式时, $W(mode) = 1$; 16×8 或 8×16 模式时, $W(mode) = 1/2$; 8×8 模式时, $W(mode) = 1/4$; 8×4 或 4×8 模式时, $W(mode) = 1/8$; 4×4 模式时, $W(mode) = 1/16$ 。

从表 1、2 可以看出,对于所有的序列,本文算法的 PSNR 和平均码率跟 FS、UMHexagonS 算法都非常接近,损失几乎可以忽略。而本文算法每种模式平均每帧只需要搜索 3~7 个点即可,计算复杂度约为 FS 的 $1/200$,约为 UMHexagonS 算法的 $1/6$,适合硬件实现,有较大的实用价值。

表 1 三种算法在 5 个参考帧下的对比实验结果

序列	FS 算法			UMHexagonS 算法			本文算法		
	标准化的 计算复杂度	峰值信噪 比/dB	平均 码率/Kbps	标准化的 计算复杂度	峰值信噪 比/dB	平均 码率/Kbps	标准化的 计算复杂度	峰值信噪 比/dB	平均 码率/Kbps
akiyo	1089	39.75	93.85	10.62	39.76	93.78	4.08	39.74	93.35
container	1089	35.95	171.66	14.92	35.95	172.68	4.15	35.94	172.66
coastguard	1089	34.44	1330.45	57.87	34.43	1327.32	6.63	34.42	1331.85
foreman	1089	36.88	380.58	35.00	36.86	380.80	6.70	36.82	382.76
mobile	1089	34.23	1668.39	40.73	34.22	1669.99	6.49	34.19	1683.50

4 结语

本文基于运动物体的性质,充分考虑图像序列的时空相关性,将运动矢量复用和多种终止条件结合在一起,形成了新

的高效多参考帧搜索算法。实验证明,在低码率序列图像传输中,本文的多参考帧搜索算法在精度和计算复杂度方面都具有一定的优越性。下一步将结合模式选择、并行处理思想进行系统研究,努力探索帧间预测编码的最佳方案。

4 结语

针对传统 Hough 变换检测椭圆时存在的计算量大、无效采样、相互影响等问题,本文考虑了以下问题:如何利用边缘曲线段统计分布信息找到感兴趣初始参数,克服 RHT 的无效累积;如何约束高维参数空间搜索范围降低运算难度;如何减少曲线间相互影响,减少错误累积。基于此本文提出了基于局部 PCA 感兴趣参数约束的 Hough 多椭圆分层检测算法。

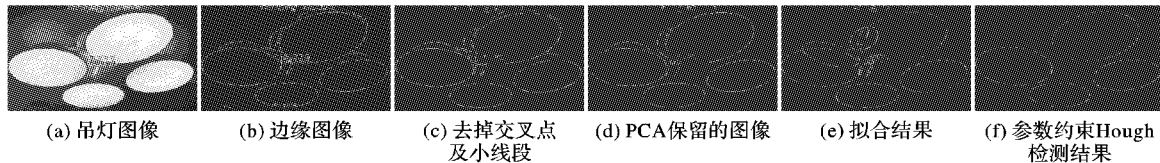


图 4 吊灯彩色图像椭圆检测结果

参考文献:

- [1] L XU, E OJA, P KULTANEN. A new curve detection method: Randomized hough transform [J]. *Pattern Recognition Letters*, 1990, 11(5): 331 - 338.
- [2] LIU ZHI-YONG, QIAO HONG, XU LEI. Multisets mixture learning-based ellipse detection [J]. *Pattern Recognition*, 2006, 39(4): 731 - 735
- [3] 于海滨, 刘济林. 基于中心提取的 RHT 在椭圆检测中的应用 [J]. *计算机辅助设计与图形学学报*, 2008, 19(9): 1107 - 1113.
- [4] KASEMIR K-U, BETZLER K. Detecting ellipses of limited eccentricity in images with high noise levels [J]. *Image and Vision Computing*, 2003, 21(2): 221 - 227.
- [5] ZHANG SI-CHENG, LIU ZHI-QIANG. A robust real-time ellipse detector [J]. *Pattern Recognition*, 2005, 38(2): 273 - 287.
- [6] FITZGIBBON A, PILU M, FISHER R B. Direct least square fitting of ellipses [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1999, 21(5): 477 - 480.

将标记函数、PCA 方向分析、曲线拟合和 Hough 变换四者有机结合起来,使椭圆检测复杂度大大降低,存储空间占用减少,精度有所提高,且稳定性较好,对圆的检测同样适用。该算法对边缘检测的连续性要求较高,否则将影响线段 PCA 及参数拟合的可靠性。实验结果表明,本文提出的方法计算消耗小、检测精度高,是一种简单有效的椭圆检测算法,为基于圆及椭圆检测的中高级计算机视觉问题提供了较好的预处理手段。

- [7] LU WEI, TAN JING-LU. Detection of incomplete ellipse in images with strong noise by iterative randomized Hough transform (IRHT) [J]. *Pattern Recognition*, 2008, 41(2): 1268 - 1279.
- [8] 唐珉, 胡占义. 参数空间分解法 [J]. *计算机学报*, 1999, 22(9): 911 - 917.
- [9] MAYBANK S J. Application of the fisher-rao metric to ellipse detection [J]. *International Journal of Computer Vision*, 2007, 72(3): 287 - 307.
- [10] GURU D S, SHEKAR B H, NAGABHUSHAN P. A simple and robust line detection algorithm based on small eigenvalue analysis [J]. *Pattern Recognition*, 2004, 25(1): 1 - 13.
- [11] FITZGIBBON A W, FISHER R B. A buyer's guide to conic fitting [C]// *Proceedings of the 6th British Conference on Machine Vision*. Surrey, UK, UK: BMVA Press, 1995, 2: 513 - 522.
- [12] BOOKSTEIN F L. Fitting conic sections to scattered data [J]. *Computer Graphics and Image Processing*, 1979, 9(2): 56 - 71.

(上接第 1364 页)

表 2 三种算法在 10 个参考帧下的对比实验结果

序列	FS 算法			UMHexagonS 算法			本文算法		
	标准化的 计算复杂度	峰值信噪 比/dB	平均 码率/Kbps	标准化的 计算复杂度	峰值信噪 比/dB	平均 码率/Kbps	标准化的 计算复杂度	峰值信噪 比/dB	平均 码率/Kbps
akiyo	1089	39.81	91.80	10.37	39.81	91.96	3.11	39.78	92.56
container	1089	35.98	168.81	15.07	35.98	169.22	3.12	35.94	169.72
coastguard	1089	34.45	1325.03	53.69	34.43	1324.01	4.95	34.43	1330.11
foreman	1089	36.97	373.57	34.44	36.95	374.33	4.94	36.87	376.00
mobile	1089	34.31	1624.79	41.30	34.30	1626.49	4.93	34.26	1654.93

参考文献:

- [1] Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H. 264/ISO/IEC 14496-10AVC) [Z]. Joint Video Team (JVT), 2003.
- [2] TING C-W, PO L-M, CHEUNG C-H. Center-biased frame selection algorithms for fast multi-frame motion estimation in H. 264 [C]// *Proceedings of 2003 International Conference Neural Networks and Signal Processing*. Washington, DC: IEEE Press, 2003, 2: 1258 - 1261.
- [3] CHEN M-J, CHIANG Y-Y, LI H-J, *et al.* Efficient multiframe motion estimation algorithms for MPEG-4 AVC/JVT/ H. 264 [C]// *Proceedings of the 2004 International Symposium on Circuits and Systems: ISCAS 04*. Washington, DC: IEEE Press, 2004, 3: 737 - 740.
- [4] SU YE-PING, SUN MING-TING. Fast multiple reference frame motion estimation for H. 264/AVC [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2006, 16(3): 447 - 452.

- [5] HSIEH B-Y, HUANG Y-W, WANG T-T, *et al.* Fast motion estimation algorithm for H. 264/MPEG-4 AVC by using multiple reference frame skipping criteria [C]// *Proceedings of Visual Communications and Image Processing*, SPIE 5150: VCIP 2003. Bellingham, WA: Society of Photo-Optical Instrumentation Engineers, 2003: 1551 - 1560.
- [6] HUANG Y-W, HSIEH B-Y, CHIEN S-Y. Analysis and complexity reduction of multiple reference frames motion estimation in H. 264/AVC [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2006, 16(4): 507 - 522.
- [7] CHEN T-C, TSAI C-Y, HUANG Y-W, *et al.* Single reference frame multiple current macroblocks scheme for multiple reference frame motion estimation in H. 264/AVC [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2007, 17(2): 242 - 247.
- [8] CHEN ZHI-BO, ZHOU PING, HE YUN. Fast integer and fractional pel motion estimation for JVT [C]// *6th Meeting of JVT-F017*. Awaji, Japan: [s. n.], 2002.