

中图法分类号: TN919.81 文献标志码: A 文章编号: 1006-8961(2011)06-0927-06

论文索引信息: 张新安, 官彦军, 陈爱武. AVS-M 亚像素运动估计快速算法 [J]. 中国图象图形学报, 2011, 16(6): 927-932

AVS-M 亚像素运动估计快速算法

张新安, 官彦军, 陈爱武

(湖南科技学院电子工程系, 永州 425100)

摘要: 针对 AVS-M 亚像素运动矢量搜索的复杂性, 提出一种基于预测的亚像素运动估计快速搜索算法。算法通过分析亚像素候选点的匹配误差关系排除可能性小的点, 保证了亚像素搜索的准确性。通过有效的搜索策略, 对于不同运动特征的视频序列, 在平均 PSNR 下降不超过 0.01 dB 的情况下, 与亚像素全搜索算法相比, 搜索点数减少 65.25% ~ 87.69%, 与中心偏置亚像素搜索算法相比, 搜索点数减少 43.67% ~ 67.38%, 有效减少了亚像素运动估计的计算量。

关键词: AVS-M; 视频编码; 亚像素运动估计; 预测矢量

Fast algorithm of sub-pixel motion estimation for AVS-M

Zhang Xin'an, Gong Yanjun, Chen Aiwu

(Department of Electronic Engineering of Hunan University of Science and Engineering, Yongzhou 425100 China)

Abstract: To reduce the complexity of AVS-M sub-pixel motion vector search, this paper proposes a sub-pixel motion estimation fast searching algorithm based on prediction. The algorithm expels the less possibility points by doing matching error relation of sub-pixel points selected analysis, which assures the accuracy of sub-pixel search. Through effective search strategy, compared with the half-pixel full search algorithm and CBFPS, experiments show that the proposed algorithm can respectively reduce the sub-pixel search points by 65.25% ~ 87.69% and 43.67% ~ 67.38% on average PSNR lost less than 0.01 dB to the video sequences with different motion characteristics. The algorithm efficiently decreases the computational number of sub-pixel motion estimation.

Keywords: AVS-M; video coding; sub-pixel motion estimation; prediction vector

0 引言

运动估计 (ME) 是 AVS-M^[1] 视频编码的关键部分, 同时也是整个编码器中计算量最大、耗时最多的部分。运动估计模块由整像素运动估计和亚像素运动估计两部分组成, 随着整像素运动估计快速算法的发展, 整像素的搜索点数大幅度减少, 一般可降至 10 个点以下, 导致亚像素运动估计的计算量占整个运动估计过程计算量的比重增大。因此, 如何减少

亚像素运动估计的计算量就成为实现实时视频编码的一个研究热点。

目前许多学者对亚像素运动估计算法进行了大量的研究^[2-7], 参考文献[2]提出一种通过计算值的范围判断最优亚像素点位置的算法, 但其必须先进行 1/2 像素搜索, 在此基础上利用中间结果和代价函数分布规律推出更精确的亚像素位置, 这仍需要较多的搜索点数和较大的存储空间, 有待进一步改进。文献[3]提出的中心偏置亚像素搜索算法 CBFPS 直接检测所预测的亚像素位置, 并在其周围

收稿日期: 2010-02-05; 修回日期: 2010-07-22

基金项目: 湖南省科技厅科技发展计划资助项目 (2009FJ4062); 湖南省教育厅高等学校科学研究资助项目 (08C364)。

第一作者简介: 张新安 (1957—), 男, 副教授。华东师范大学学士学位, 主要研究方向为数字信号处理及 DSP 应用。

E-mail: hnzxa@163.com。

做精细搜索,与全搜索相比,算法可以减少 53% 的计算量。但如果预测的亚像素位置不够准确,算法易陷入局部最小点。

提出一种适用于 AVS-M 基于预测的亚像素运动估计快速搜索算法。算法通过分析亚像素候选点的匹配误差关系排除可能性小的点,保证亚像素搜索的准确性。而且算法不需要测试最优整像素运动矢量周围点的匹配误差,可以与任意整像素搜索算法配合使用。对于不同运动特征的视频序列,该算法在保证图像质量和编码效率基本不变的同时,与全搜索算法相比,搜索点数减少 65.25% ~ 87.69%,可有效减少亚像素运动估计的计算量。

1 亚像素全搜索算法

AVS-M 定义了式(1)对亚像素运动估计的匹配误差进行计算,并选取最小匹配误差点为最优匹配点:

$$J(m, \lambda_{\text{MOTION}}) = \text{SATD}(s, c(m)) + \lambda_{\text{MOTION}} \cdot R(m - p) \quad (1)$$

式中, $\mathbf{m} = (m_x, m_y)^T$ 为运动矢量; $\mathbf{P} = (P_x, P_y)^T$ 为预测的运动矢量; λ_{MOTION} 是拉格朗日(Langrange)乘数因子; $R(m - p)$ 是编码运动向量所需要的比特数,通过查关于运动信息的查找表得到, SATD 按下面方式计算。

计算原始值和预测值之间的差值:

$$d(x, y) = s(x, y) - c(x - m_x, y - m_y) \quad (2)$$

对 $d(x, y)$ 进行 2 维的 Hadamard 变换,得到 $dT(x, y)$:

$$\text{SATD}(s, c(m)) = \left(\sum_{x,y} |dT(x, y)| \right) / 2 \quad (3)$$

AVS-M 亚像素运动估计全搜索过程如图 1 所示,大写字母代表整像素,小写字母代表 1/2 像素,数字代表 1/4 像素。

设整像素搜索已经得到最优矢量点 E, 首先计算 E 周围 8 个 1/2 像素点 a, b, c, d, e, f, g, h 的匹配误差,比较各点的匹配误差,求出最小匹配误差点,即最优 1/2 像素匹配点。然后,根据最优 1/2 像素点搜索 1/4 像素点。假设 c 为最优 1/2 像素匹配点,则搜索计算 c 周围 8 个 1/4 像素点 1, 2, 3, 4, 5, 6, 7, 8 的匹配误差,进行比较求得最优 1/4 像素匹配点。搜索过程中 1/2 和 1/4 亚像素利用邻近已编码像素点进行内插得到。全搜索算法共搜索计算了 16 个

点(8 个 1/2 像素点, 8 个 1/4 像素点), 计算量很大。

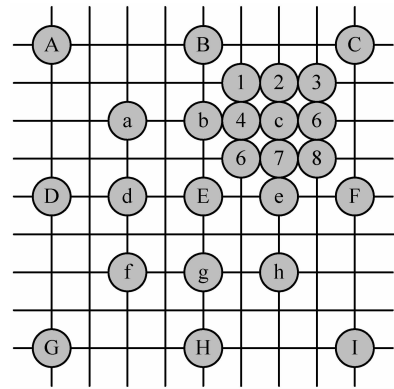


图 1 AVS-M 亚像素运动估计全搜索过程

Fig 1 AVS-M sub-pixel motion estimation total searching process

2 基于预测的亚像素运动估计快速算法

许多运动估计快速算法利用相邻块的运动矢量来预测当前块的搜索起点,以减少块匹配的搜索点数和防止陷入局部极小点。预测运动矢量包含整像素与亚像素两部分信息,如果预测运动矢量的整像素部分等于整像素运动搜索得到的最优矢量,那么预测矢量的亚像素部分很可能接近全局最优点,可以将预测的亚像素运动矢量作为搜索起点开始搜索。否则,不够准确的预测亚像素运动矢量会使匹配效果急剧下降,针对这种情况提出一种搜索策略,在保证图像质量的同时,可有效地减少搜索点数。

2.1 基于预测的亚像素运动矢量搜索

相邻块的运动矢量与当前块的运动矢量有很强的相关性, AVS-M 为了利用这种相关性定义了当前块的最佳接近矢量 MVP_{closest} , 选取 MVP_{closest} 作为当前块亚像素运动矢量的一个预测矢量,称之为最佳接近预测。

AVS-M 中有 7 种帧间预测模式,一个宏块根据其内部的纹理和运动特征可以分割成从 16×16 到 4×4 不同块尺寸的分割形式,较大尺寸的块与其内部较小尺寸块往往具有相似运动,因此选取上层模式的运动矢量作为当前模式亚像素运动矢量的另一个预测矢量 $MVP_{\text{up-layer}}$,称之为上层预测。

将预测运动矢量 pred_mv 和整像素搜索得到的

最优整像素运动矢量 mv 以 1/4 像素单位表示,按照式(4)可以提取预测亚像素运动矢量 $frac_pred_mv$ ^[3]。

$$frac_pred_mv = (pred_mv - mv) \% 4 \quad (4)$$

为了验证预测亚像素运动矢量的准确性,定义整像素搜索所得运动矢量 $MV_i = (MV_{ix}, MV_{iy})$ 与预测亚像素运动矢量的偏差的累积分布函数 CDF^[8]

$$P(\delta MV \leq d) = P\left[\bigcup_{i=1}^7 (|PMV_{ix} - MV_{ix}| + |PMV_{iy} - MV_{iy}| \leq d)\right] \quad (5)$$

式中 MV_{ix}, MV_{iy} 分别为预测矢量的水平和垂直分量, $i = 1, \dots, 7$ 为 7 种块尺寸。 δMV 和 d 以 1/4 像素精度为单位。表 1 列出了 6 个从运动较小到相对运动较大特征的不同序列的统计结果。

从表 1 可以看出,对于 Foreman 序列,最佳接近预测矢量和上层预测矢量与整像素搜索得到的最佳运动矢量相同的概率分别为 41.51% 和 47.33%,如果同时利用这两种预测矢量,相同的概率增加到 62.91%,本文算法选取两种预测矢量,比 CBFPS 只采用中值预测矢量的准确性要高很多。虽然将预测矢量作为最终运动矢量,此概率还不足够高,但如果

表 1 预测亚像素运动矢量偏差的统计结果

Tab. 1 To predict the statistical consequences of sub-pixel motion vector declination

序列	预测亚像素运动矢量	$d = 0$ /%	$d = 1$ /%	$d = 2$ /%	$d = 3$ /%
Akiyo	$MVP_{closest}$	73.41	92.60	97.41	99.22
	$MVP_{up-layer}$	73.07	92.52	97.32	99.21
	$MVP_{closest} + MVP_{up-layer}$	83.73	96.28	99.07	99.79
Silent	$MVP_{closest}$	72.18	92.37	96.36	99.17
	$MVP_{up-layer}$	72.67	91.74	96.47	98.68
	$MVP_{closest} + MVP_{up-layer}$	82.54	95.75	98.87	99.59
News	$MVP_{closest}$	69.80	87.67	94.50	97.71
	$MVP_{up-layer}$	72.12	88.31	95.54	98.06
	$MVP_{closest} + MVP_{up-layer}$	81.79	93.71	98.28	99.36
Carphone	$MVP_{closest}$	69.39	87.24	94.17	97.23
	$MVP_{up-layer}$	71.67	87.83	95.06	97.72
	$MVP_{closest} + MVP_{up-layer}$	81.35	93.22	97.76	99.21
Foreman	$MVP_{closest}$	41.51	68.72	83.14	91.82
	$MVP_{up-layer}$	47.33	69.69	83.83	91.09
	$MVP_{closest} + MVP_{up-layer}$	62.91	84.75	93.96	97.44
Mobile	$MVP_{closest}$	38.67	61.35	86.75	93.64
	$MVP_{up-layer}$	57.59	72.87	90.43	94.91
	$MVP_{closest} + MVP_{up-layer}$	64.37	82.73	94.63	97.82

将最佳接近预测矢量和上层预测矢量作为初始搜索点,做一步小菱形搜索(SDSP),就可以得到84.75%的最佳运动矢量,做三步小菱形搜索,准确率即达到97.44%。对于 Akiyo 等具有较小运动特征的序列,准确性可达到 99.79%。因此,取具有最小匹配误差的预测亚像素运动矢量作为初始搜索点,做三步小菱形搜索得到最终的运动矢量。

图 2 给出了一种典型的小菱形搜索过程。图 2 中 1A、1B 为第 1 步小菱形搜索得到的最优点和次优点,2、2A 为第 2 步小菱形搜索的点,2A、1A 为第 2 步小菱形搜索得到的最优点和次优点,3 为第 3 步小菱形搜索的点。先进行一步小菱形搜索,如果最优点不是小菱形的中心点,则在第 2 步小菱形搜索时,仅搜索与第 1 步小菱形搜索中的最优点和次优位置相同的点。依此类推在第 3 步小菱形搜索中仅搜索与第 2 步小菱形中的最优点和次优点位置相同的点。当最小匹配误差的点位于菱形的中心或者搜索步数已经达到 3 步时,算法终止。

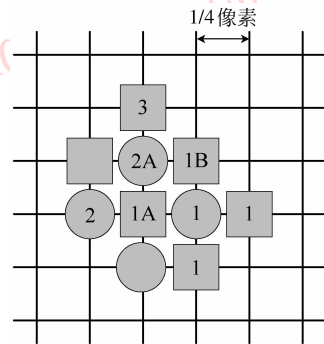


图 2 一种典型的小菱形搜索过程

Fig. 2 A kind of typical small rhombus searching process

从表 1 还可以看出,对于 Akiyo、News、Foreman 序列,所有待编码块中,最终的亚像素运动矢量分别有 83.73%、81.79%、和 62.91%,与预测矢量相同,如果能够预测出这些编码块,就可以跳过 SDSP 搜索过程,省去大量不必要计算。如果预测矢量与最终亚像素运动矢量相同,预测矢量将具有最小匹配误差。为此,对 7 种块尺寸设置自适应阈值 TH_i :

$$TH_i = \alpha \times Prev_J_{MOTION} \quad (6)$$

式中 $Prev_J_{MOTION}$ 为参考帧中相同块尺寸亚像素最小匹配误差的平均值。

算法的中止准则为:若 $J_{MOTION} < TH_i$,表明亚像素预测运动矢量已经足够精确,立即终止搜索过程。

2.2 亚像素精细搜索

如果预测矢量的整像素部分与整像素运动搜索得到的最优运动矢量不同,在亚像素搜索中就不能使用预测矢量的亚像素部分,因为不够准确的预测亚像素运动矢量会造成图像质量的较大下降。可以利用最优整像素运动矢量周围亚像素点匹配误差的关系,来排除一些可能性较小的点,从而加速搜索过程。

1/2 像素精细搜索如图 3 所示,以最优整像素运动矢量为搜索中心,1/2 像素为单位做一步小菱形搜索,保存最优点和次优点的位置和匹配误差,根据最优点和次优点的位置决定其他需要测试的点。如图 3(a)所示,如果最优点和次优点 A、B 不在一条直线上,那么只需再增加一个测试点 C。在(b)中,最优点和次优点 A、B 在一条直线上,那么需要再增加测试最优点 A 相邻的两个点 C、D。以上策略搜索完毕 1/2 像素运动矢量后,以最小匹配误差点作为 1/4 像素精度的搜索中心,1/4 像素为单位,做一步小菱形搜索,采用相同的策略对 1/4 像素点作精细搜索。

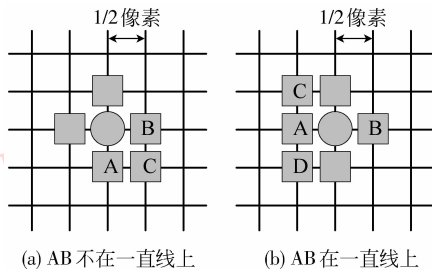


图 3 1/2 像素精细搜索示意

Fig 3 1/2 pixels fine searching schematic drawing

2.3 算法描述

根据上述研究和分析,本文提出基于预测的亚像素快速运动估计算法(PFPS)。为描述方便,定义最优整像素运动矢量与最佳接近预测矢量整数部分相同的块为最佳接近预测块(CPB),最优整像素运动矢量与上层预测矢量整数部分相同的块为上层预测块(UPB)。算法的主要步骤如下:

如果当前块是 CPB 或 UPB,计算搜索中心 $(0,0)$ 的 J_{MOTION} 。

如果当前块是 CPB,计算最佳接近预测矢量点的 J_{MOTION} ;如果当前块是 UPB,计算上层预测矢量点的 J_{MOTION} 。

选择具有最小匹配误差 $\min J_{\text{MOTION}}$ 的点。

如果 $\min J_{\text{MOTION}} < TH_i$,选择对应的运动矢量作为最终的亚像素运动矢量,结束搜索。

否则,以 $\min J_{\text{MOTION}}$ 点为中心做 SDSP 搜索,直至最小匹配误差点位于 SDSP 中心或搜索步数达到 3。否则,以 LDSP 计算最佳整像素点周围的 4 点,保存最优和次优点的位置和匹配误差。根据最优点和次优点的位置确定需要最佳计算的点。

如果最优或次优点中有搜索起点,增加计算非起点周围相邻的两点,更新最优点的位置。

如果最优点和次优点相邻且不是搜索起点,增加计算两点之间的点,更新最优点的位置。

否则,增加计算剩余所有点,更新最优点的位置。

如果当前精度是 1/4 像素精度,结束搜索。

否则在 1/4 像素精度点做精细搜索。

3 实验结果与分析

将本文算法置入 AVS-M 校验模型 WM3.3a,选取 6 个具有不同运动特征的 QCIF (174 × 144) 标准视频序列对算法性能进行测试。6 个序列中,“Aldyo、Silent”具有静止背景和较小运动,“News、Carphone”运动程度居中,“Foreman、Mobile”具有较强的运动和移动背景。这些视频序列基本涵盖了各种类型的视频序列特征,因此可认为测试结果具有普遍意义。编码器主要参数包括:搜索范围为 16;参考帧为 1;允许 RDO;Hadamard 变换开启。视频序列各包含 100 帧,帧率为 30。量化参数为 28,首帧编码为 I 帧,其余帧编码为 P 帧。测试平台:CPU 为 Intel Pentium4 3.0 GHz,内存为 DDR512 MB,操作系统为 Windows XP。与亚像素全搜索算法、CBFPS 算法和参考文献 [4] 的 ARFPS (adaptive range fractional pel search) 算法在相同条件下的测试结果如表 2 所示。

从表 2 可以看出,与亚像素全搜索算法相比,本文算法的搜索点数减少 65.25% ~ 87.68%,每帧平均亚像素运动估计时间减少 54.53% ~ 79.69%,而平均 PSNR 下降不超过 0.01 dB。与 CBFPS 算法相比,搜索点数减少 43.67% ~ 67.38%,每帧平均亚像素运动估计时间减少 35.59% ~ 67.51%,这是因

表 2 本文算法与其他算法的比较
Tab.2 The contrast between the algorithm in this paper and other algorithms

序列	算法	每帧平均编码时间/ms	每帧平均 ME 时间/ms	每帧平均亚像素 ME 时间/ms	搜索点数	PSNR/dB	平均码率/(kbit/s)
Akiyo	FS	416.66	289.57	227.12	16.00	40.86	105.89
	CBFPS	324.66	182.10	140.55	6.04	40.87	105.89
	ARFPS	281.12	157.68	120.71	5.23	40.85	105.95
	本文算法	225.56	89.10	45.66	1.97	40.87	106.05
Silent	FS	418.55	291.34	244.57	16.00	38.35	170.73
	CBFPS	335.09	207.12	170.45	6.34	38.34	171.03
	ARFPS	297.89	186.21	151.53	5.64	38.34	171.24
	本文算法	247.33	97.69	50.06	2.16	38.33	171.92
News	FS	418.44	303.55	247.01	16.00	39.04	223.45
	CBFPS	335.01	205.34	142.57	6.58	39.03	223.47
	ARFPS	301.37	184.72	128.26	5.92	39.03	223.68
	本文算法	237.77	106.10	57.10	2.28	39.02	227.12
Carphone	FS	428.77	321.88	256.12	16.00	39.09	312.35
	CBFPS	362.87	234.32	173.22	7.32	39.07	311.42
	ARFPS	330.65	213.51	157.84	6.67	39.07	312.83
	本文算法	284.76	163.54	107.76	3.34	39.06	313.37
Foreman	FS	432.32	330.12	262.12	16.00	37.96	318.06
	CBFPS	371.57	243.10	177.32	9.11	37.95	319.29
	ARFPS	339.76	222.29	162.14	8.33	37.95	319.91
	本文算法	302.01	180.43	113.10	5.04	37.97	321.98
Mobile	FS	449.87	333.89	262.65	16.00	35.36	914.86
	CBFPS	395.72	265.87	185.43	9.87	35.35	915.61
	ARFPS	362.44	243.53	167.51	9.04	35.35	916.45
	本文算法	315.86	186.64	119.42	5.56	35.35	916.92

为 CBFPS 算法只对 8×8 以下块尺寸的编码块采用快速算法,对较大块 (16×16 、 16×8 、 8×16) 仍使用全搜索。本文算法对全部编码块使用快速算法,这是因为本文算法采用了更多的预测矢量,预测矢量的准确性比 CBFPS 算法高,可以应用在较大的编码块。与 ARFPS 算法相比,搜索点数减少 38.50%~62.33%,每帧平均亚像素运动估计 (ME) 时间减少 28.71%~62.17%,平均 PSNR 基本相当,同时本文算法编码码率的变化很微小。

从表 2 中还可以看出,对于 Akiyo 和 Foreman 序列,本文算法和 CBFPS 算法的平均 PSNR 比全搜索算法稍高,这是因为亚像素全搜索算法只是搜索最优 $1/2$ 像素点周围的 8 个 $1/4$ 像素运动矢量,而本文算法和 CBFPS 算法通过预测 $1/4$ 像素搜索起点,并做小菱形循环搜索,可以减少陷入局部最小点的可能,所以提高了编码信号的峰值信噪比。

4 结 论

分析亚像素运动估计全搜索算法,在现有亚像素运动估计快速算法的基础上,提出一种适用于 AVS-M 基于预测的亚像素快速搜索算法。为提高预测矢量的准确性,算法采用更多的预测矢量,最佳接近矢量 $MVP_{closest}$ 和上层预测矢量 $MVP_{up-layer}$,同时采用早停止技术跳过不必要的搜索过程,当预测矢量与整像素搜索到的最优运动矢量存在较大差异时,算法通过分析亚像素候选点的匹配误差关系排除可能性小的点,保证了亚像素搜索的准确性。而且算法不需要测试最优整像素运动矢量周围点的匹配误差,可以与任意整像素搜索算法配合使用。对于不同运动特征的视频序列,该算法在保证图像质量和编码效率基本不变的同时,有效减少了亚像素运动估计的计算量。

参考文献 (References)

- [1] Beijing: AVS Workgroup. Information Technology—Audio Video Coding Standard, Part 7: Mobile Video [S]. Beijing: Chinese Standard Press, 2006: 6-8. [北京: AVS 工作组. 信息技术—先进音视频编码, 第 7 部分: 移动视频 [S]. 北京: 中国标准出版社, 2006: 6-8.]
- [2] Wang Weidong, Yao Qingdong. Fast algorithm of fractional pixel accuracy motion estimation [J]. Journal of China Institute of Communications, 2003, 24 (4): 128-132. [王维东, 姚庆栋. 小数像素运动估计快速算法 [J]. 通信学报, 2003, 24 (4): 128-132.]
- [3] Zhibo Ch, Peng Zh, Yun H. Fast integer and fractional Pel motion estimation [R]. Joint Video Team (JVT) of ISO/IEC MPEG and ITU-TVCEG, F017r1, 2002, 4: 8-12.
- [4] Zhou Cheng, Tan Yihua, Tian Jinwen, et al. A fast algorithm of sub-pixel motion estimation based on AVS [J]. Journal of Image and Graphics, 2007, 12 (9): 1520-1524. [周城, 谭毅华, 田金文, 等. 一种基于 AVS 的亚像素快速运动估计算法 [J]. 中国图象图形学报, 2007, 12 (9): 1520-1524.]
- [5] Chen Yonghua, Wang Yi, Liu Donghua, et al. Fast Three step sub-pixels search algorithm for AVS video coding [J]. Computer Engineering, 2008, 34 (9): 230-231. [陈永华, 王毅, 刘东华, 等. 用于 AVS 视频编码的三步快速子像素搜索算法 [J]. 计算机工程, 2008, 34 (9): 230-231.]
- [6] Zhou Bo, Wang Libo, Liu Donghua, et al. Rapid sub-pixels movement search algorithm for AVS video coding [J]. Application Research of Computers, 2008, 25 (7): 2238-2240. [周波, 王力波, 刘东华, 等. 用于 AVS 视频编码的快速子像素运动搜索算法 [J]. 计算机应用研究, 2008, 25 (7): 2238-2240.]
- [7] Dong Haiyan, Zhang Qishan. A fast haif-pixel motion vector search algorithm [J]. Journal of Image and Graphics, 2005, 10(7): 884-887. [董海燕, 张其善. 一种快速半像素运动矢量搜索算法 [J]. 中国图象图形学报, 2005, 10 (7): 884-887.]
- [8] Zhon Z, Xin J, Sun M T. Fast Motion Estimation and Inter Mode Decision for H.264/MPEG-4 AVC Encoding [M]. Amsterdam, Netherlands: ELSEVIER, 2005:156-158.