

基于多帧运动估计的帧率提升算法*

杨爱萍, 董翠翠, 侯正信, 李娜

(天津大学电子信息工程学院, 天津 300072)

摘要: 结合最大后验概率(MAP)估计的准确性和分层块匹配算法的快速性,研究了一种多参考帧运动估计算法,并提出了一种基于多帧运动估计的帧率提升(FRUC)系统方案。实验结果表明,基于该算法的内插帧无论从客观指标(时间复杂度、信噪比 PSNR)还是主观质量(视觉效果)均优于现有常用方法,且算法复杂度较低,便于硬件实现。

关键词: 帧率提升; 运动估计; 多参考帧; 块匹配; MAP 估计

中图分类号: TP391 **文献标志码:** A **文章编号:** 1001-3695(2012)10-3952-04

doi:10.3969/j.issn.1001-3695.2012.10.094

New algorithm for FRUC based on multi-frame motion estimation

YANG Ai-ping, DONG Cui-cui, HOU Zheng-xin, LI Na

(School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China)

Abstract: By combining the accuracy of MAP estimation and the high speed of hierarchical block matching, this paper studied the multi-reference frames motion estimation algorithm and proposed a system scheme, which could improve the rate of the frame based on it. Experimental results show that the interpolated frames of the proposed algorithm can perform better than what the current algorithm either in subjective (visual effect) and objective (time complexity, PSNR) indicators. Besides due to its low computation complexity, this algorithm can be easily realized in hardware.

Key words: frame rate up conversion (FRUC); motion estimation; multi-reference frames; block-matching; MAP estimation

0 引言

基于运动补偿内插的帧率提升算法中,运动估计的精度决定了运动补偿内插帧的质量^[1]。为了得到真实的运动估计,产生了一系列运动估计算法。现有的运动估计算法可以分为两大类^[2]:像素递归法,运动估计的精度高,但计算复杂;块匹配算法(BMA)^[3],简单和易于硬件实现。

基于块匹配的运动估计方法有很多,全搜索法(FS)最简单,可到达全局最优,但计算过于复杂;三步法^[4](TSS)快速高效,但容易陷入局部最优;由 Li 等人^[4]提出新三步法(NTSS)减小了局部最优的可能性;Shan 等人^[5]提出的钻石搜索算法(DS)适用于有运动矢量中心偏移特性的视频序列。后又出现了各种改进的搜索方法,双向运动估计^[6,7]可以解决传统单向运动估计容易出现的遮挡等问题;3DRS^[7]算法可以解决较多候选矢量和硬件代价之间的矛盾;EPZS 算法^[8]基于预测的思想可以从更多的预测矢量和搜索路径中选择最佳的运动矢量。但以上这些方法都是以匹配块大小不变为前提的,而实际中无法达到匹配块内各像素位移相等,采用块大小自适应的多分辨率运动估计可以解决上述问题。Bierling 等人^[9,10]提出了分层块匹配法;Burt 等人^[11]提出了利用图像金字塔结构进行运动估计。在此基础上,本文结合 MAP 估计的准确性和分层块匹配算法的快速性,研究了一种多帧运动估计算法^[12],并提出了一种基于多帧运动估计的帧率提升系统方案。

1 多参考帧运动估计算法

运动矢量的准确性是运动补偿帧率提升算法中首要考虑的问题。基于 Bayesian 模型的最大后验概率估计(MAP)算法可以得到更为精确的运动矢量^[13-17],该模型以吉布斯模型表征运动矢量场,且引入平滑约束。MAP 估计通常用优化的估计方法实现,常用的是迭代优化。文献[12]提出连续三帧或者三帧以上更多的帧来估计运动,能够较好地反映运动矢量场时间的相关性^[18,19],并且能有效地处理被遮挡区域^[20]。且可以利用 BMA 的结果来加速多帧 MAP 估计以提高计算速度。本文研究的运动估计算法采用三个连续帧 $F(x, t)$ 、 $F(x, t+1)$ 、 $F(x, t+2)$ 和一个先前已估计好的运动场 $M_b(x, t)$ 作为输入,如图 1 所示。其中 $x = (x, t)^T$ 和 t 分别表示像素或运动矢量的空间和时间坐标。

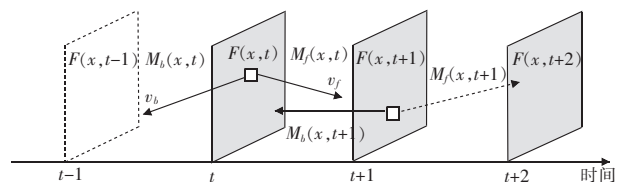


图1 算法采用的连续三帧和一个先前运动场估计示意图

1.1 多参考帧运动估计算法

本文研究的多参考帧运动估计算法流程如图 2 所示。具

收稿日期: 2012-03-28; 修回日期: 2012-04-30 基金项目: 国家自然科学基金资助项目(61002027)

作者简介: 杨爱萍(1977-),女,山东阳谷人,副教授,硕导,主要研究方向为图像及视频处理、压缩感知理论及应用(luckyflower0506@sina.com);董翠翠(1988-),女,河北衡水人,硕士研究生,主要研究方向为图像及视频处理;侯正信(1945-),男,教授,博导,主要研究方向为数字图像及视频、信号处理理论等;李娜(1986-),女,硕士,主要研究方向为图像及视频处理。

体步骤如下:

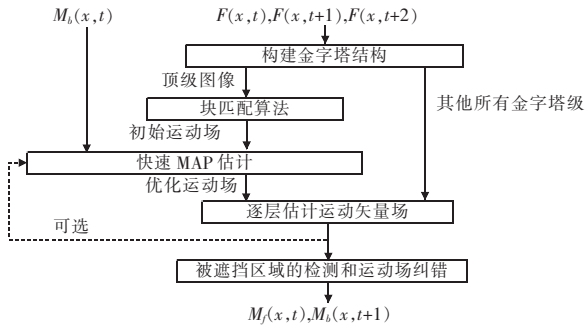


图 2 多参考帧运动估计算法流程

a) 将输入的三帧用分层运动估计构建一个金字塔结构。分层运动估计较全搜索块匹配算法可以提供更准确的运动轨迹和比较少的计算量^[9,10]。本文实验时采用高斯线性滤波器来构建金字塔^[11]。

b) 顶级图像的块匹配运动估计。金字塔结构构建完成之后,将最顶级的图像(具有最低分辨率的图像)进行块匹配运动估计,得到如图 1 所示的三个粗略的运动矢量场,即 $M_f(x, t)$, $M_b(x, t+1)$, $M_f(x, t+1)$ 。与传统的 BMA 算法不同的是,这里对于每一个块, BMA 算法将所有的搜索点的匹配误差记录在一个二维表格 E 中。BMA 算法得到的运动场作为 MAP 估计的初始值,然后用 MAP 估计算法迭代优化出更精准的运动场,存储的匹配误差用来加速该优化进程。

c) 逐层运动估计。MAP 估计后,产生的运动矢量的密度可用分层算法增加。处于不同塔级别的所有图像被分成具有相同尺寸的块,这意味着在 l 级的一个块对应 $l-1$ 级的四个块,且运动矢量密度从一级到另一级增加了四倍。在 $l-1$ 级的一个块将 l 级的三个运动矢量作为初始运动矢量。第一个运动矢量来自它的父块,另外两个来自其父块的水平和垂直相邻块,如图 3 所示。为了找到与当前块具有最小匹配误差值的运动矢量,在上述三个初始的运动矢量对应块周围的 ± 1 像素的搜索窗内搜索相应块,选择与当前块匹配误差最小的块对应的运动矢量。另外,如果有额外的计算存储空间,得到的运动矢量结果也可以用 MAP 估计进一步优化,如图 2 虚线所示。上述过程从一级到另一级不断重复进行,直到达到最高分辨率图像(原始图像)结束。

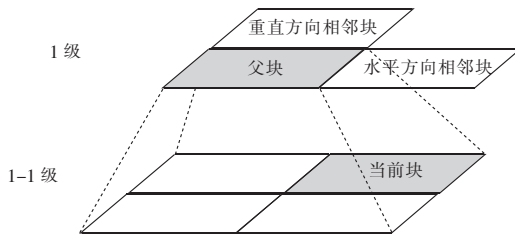


图 3 金字塔两级之间的块搜索

d) 被遮挡区域的检测。图 2 所示的最后一步,是用简单的方法检测被遮挡块并且纠正这些块中的运动矢量。为说明该方法,以 $F(x, t)$ 帧的一个块为例。从图 1 可看到,该块有两个运动矢量, $M_b(x, t)$ 的 v_b 和 $M_f(x, t)$ 的 v_f , v_b 向后指向 $F(x, t-1)$, v_f 向前指向 $F(x, t+1)$ 。如果该块在非遮挡区域,则在 $F(x, t-1)$ 和 $F(x, t+1)$ 间必然存在真实的相关性,这种情况下, v_f 的匹配误差 e_f 和 v_b 的匹配误差 e_b 将由观测噪声决定,则 e_f 和 e_b 的差值应当很小。另一方面,如果该块在被遮挡区域, $F(x, t-1)$ 和 $F(x, t+1)$ 便没有相关性,其中的一个匹配误差

可能远远大于另一个。如果 v_f 的匹配误差大于 v_b , 则将 v_f 用 $-v_b$ 替代,而且标记该块在被遮挡区域,反之,一样。被遮挡区域的信息需要高质量的内插帧来反映,此内插帧在帧率提升方案中有简略表述。

1.2 多帧 MAP 估计

1.2.1 MAP 运动估计

如 1.1 节所提,多帧 MAP 估计首先在三个帧 $F(x, t)$ 、 $F(x, t+1)$ 、 $F(x, t+2)$ 的金字塔结构的最高级图像中进行。为了简便起见,用 f_0, f_1, f_2 分别代表 $F(x, t)$ 、 $F(x, t+1)$ 、 $F(x, t+2)$ 三帧金字塔结构的最高级图像,如图 4 所示。

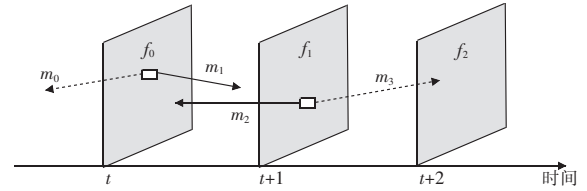


图 4 用已知的 f_0, f_1, f_2, m_0 进行 MAP 估计 m_1, m_2

用 m_0, m_1, m_2, m_3 分别表示与 $M_b(x, t)$ 、 $M_f(x, t)$ 、 $M_b(x, t+1)$ 、 $M_f(x, t+1)$ 相关的基于块的运动矢量场。以 f_0, f_1, f_2, m_0 为参考条件,分别对 m_0, m_1, m_2 作最大后验概率估计 $P(m_1, m_2, m_3 | m_0, f_0, f_1, f_2)$ 。由贝叶斯法则,该后验概率等同于式(1):

$$P(m_1, m_2, m_3 | m_0, f_0, f_1, f_2) = \frac{P(f_0 | m_0, m_1, m_2, m_3, f_1, f_2)}{P(m_0, f_0, f_1, f_2)} \times P(f_1 | m_0, m_1, m_2, m_3, f_2) \times P(m_0, m_1, m_2, m_3 | f_2) \quad (1)$$

当分别计算 m_1, m_2, m_3 概率的最大值时,由于 $P(m_0, f_0, f_1 | f_2)$ 不随 m_1, m_2, m_3 变化,式(1)中的分母可以忽略,得

$$(\hat{m}_1, \hat{m}_2, \hat{m}_3) = \arg \max_{(d_1, d_2, d_3)} [P(f_0 | m_0, m_1, m_2, m_3, f_1, f_2) \times P(f_1 | m_0, m_1, m_2, m_3, f_2) \times P(m_0, m_1, m_2, m_3 | f_2)] \quad (2)$$

首先,由于没有参考图像, m_0 对 f_0 提供的信息很少,且 f_2, m_3 与 f_0 没有直接的联系,所以 m_0, f_2 和 m_3 可以被忽略。又因噪声通常采用加性高斯白噪声作为模型^[12,21],所以,可得式(3)中的前两项,第一项条件概率 $P(f_0 | m_0, m_1, m_2, m_3, f_1, f_2)$ 处理完毕。

再者,如果没有 f_0, m_0, m_1, m_2 将对 f_1 提供很少的信息,因此,可以被忽略。所以,在同样的噪声模型下,可得式(3)中的第三项,第二项条件概率 $P(f_1 | m_0, m_1, m_2, m_3, f_2)$ 的处理完毕。

最后,由于仅仅一个单帧 f_2 对这些运动矢量场贡献很少的信息,因此条件 f_2 可以被忽略。运动矢量场常常假定用吉布斯随机场的实现,根据文献[6,12]中对矢量场的处理和文献[14,21]中时空邻近系统的子团表示,可得式(3)中的第四项,第三项条件概率 $P(m_0, m_1, m_2, m_3 | f_2)$ 的处理完毕。

最终得到简化的 MAP 运动估计算法

$$(\hat{m}_1, \hat{m}_2, \hat{m}_3) = \arg \min_{(m_1, m_2, m_3)} \left\{ \sum_{x \in \Lambda} [f_0(x) - f_1(x + m_1)]^2 + \sum_{x \in \Lambda} [f_0(x + m_2) - f_1(x)]^2 + \sum_{x \in \Lambda} [f_1(x) - f_2(x + m_3)]^2 + 2\sigma^2 \sum_{c \in C} V_c(m_0, m_1, m_2, m_3) \right\} \quad (3)$$

这样式(3)的最小化过程可由一个运动矢量场接着另一个运动矢量场^[21]进行,公式中的每一部分都可以通过查表得到,而最初的运动矢量值来自图 2 的块匹配算法。如上文所提, BMA 用来得到最初的运动矢量场,在块匹配过程中计算得到的匹配误差被存储下来,用来降低 MAP 估计的运算量。

综上,具体的多帧 MAP 估计流程如图 5 所示。

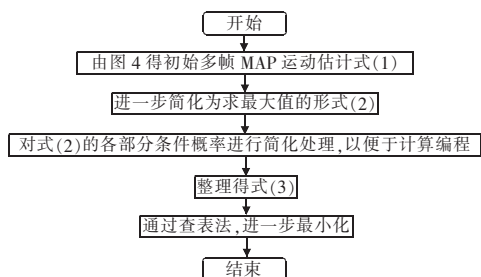


图5 多帧 MAP 运动估计流程

1.2.2 多帧 MAP 运动估计算法仿真及性能分析

本节通过实验仿真,说明多帧 MAP 运动估计算法的性能。实验中仅在金字塔的最高级图像中用了多帧 MAP 估计,如图 2 的虚线部分。其中,块尺寸定为 16×16 。对 CIF 格式的 foreman 序列进行实验,图 6 显示了 20 帧的前向和后向运动矢量场的均方误差。由图可看出,不同噪声方差的图像均在两次迭代后收敛到全局最小点,这是由于最小化过程结合了全搜索算法,得到了快速收敛性。

为了评价估计精度,将本文算法与分层 BMA 算法和传统全搜索 BMA 算法进行比较。图 7 显示了三种算法的 20 个运动矢量场的 MSE 的平均值。由图中可以看出,本文算法的 MSE 远远小于全搜索 BMA,且比分层 BMA 算法小 40%。

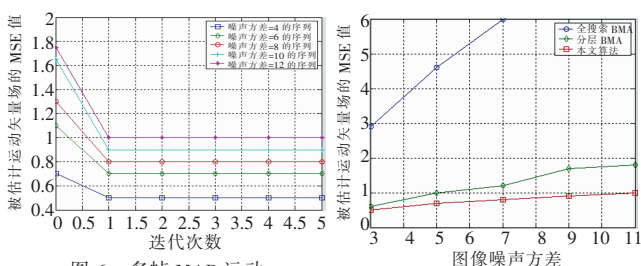


图6 多帧 MAP 运动估计算法的收敛性

图7 均方误差值

表 1 显示了 CIF 格式的 garden 序列的执行时间。本文算法相比分层 BMA 算法慢 0.2 s 左右。MAP 最优化只占了整个算法的 1% 的计算量。与全搜索 BMA 相比,本文算法快 0.3 s 左右。

2 基于多参考帧运动估计的帧率提升系统

本文提出的帧率提升系统算法框图如图 8 所示。运动估计部分利用多参考帧运动估计估计出精确的运动矢量。然后采用矢量中值滤波平滑矢量场^[22],运算简单,效果良好;再利用分级块腐蚀方法^[12]去除“块斑”,在没有引起运动矢量场模糊的前提下消除了“块斑”现象,并且抑制了一些非相关的运动矢量的产生;最后采用中值滤波运动补偿自适应方法^[23]进行插帧,得到高质量的内插帧,达到帧率提升的目的。

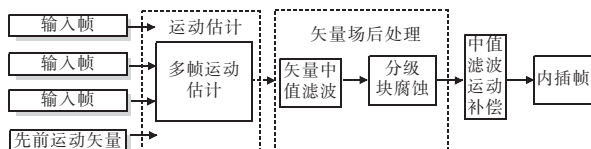


图8 基于多参考帧运动估计的帧率提升算法

3 实验仿真与结果分析

实验在 VC6.0 平台下,首先将帧速率降为原始帧率的一

半,通过在时间轴上进行下 2 采样得到,再依照帧率提升算法还原帧率。实验视频序列选用各具特点的标准 cif 测试序列 mobile、bus、coastguard、foreman。这几个视频序列包含各自的特点,可以全面反映帧率提升的优劣,如 mobile 序列包含有慢速移动和丰富细节;bus 序列包含中速运动且纹理复杂;coastguard 包含水平运动;foreman 前景运动差异明显,前景运动相对复杂,背景有微小晃动。

将本文提出的基于多参考帧运动估计的帧率提升算法,与帧平均算法、FS 和 EPZS 进行比较,它们均采用相同的矢量场后处理技术和运动补偿插帧算法。下面分别从客观质量、主观质量两个方面将本文算法与其他算法进行比较。

图 9 给出了 mobile 视频下采样序列利用不同的帧率提升算法得到的前 30 内插帧与原始视频帧序列比较所得的结果。从图中可以看出,基于本文算法得到的内插帧 Y 分量的 PSNR 值基本均在 30 dB 以上,高于另外三种算法,因此本文算法具有良好的客观质量。

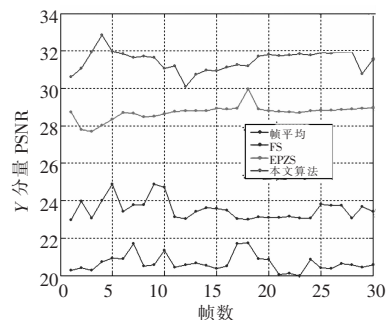


图9 mobile 视频序列

另外,对 mobile、bus、coastguard、foreman 视频序列进行帧率提升处理。为了充分说明上述各算法的客观性能,将计算出各自的时间复杂度(均取 300 帧),如表 2 所示。易见,本文算法的时间复杂度最低,效率最高。

表2 各种帧率提升算法的时间复杂度

原始序列	FS	DS	EPZS	本文
mobile	1.748 96	0.073 33	0.061 73	0.044 69
bus	1.250 10	0.604 20	0.043 33	0.036 40
coastguard	1.468 43	0.072 50	0.058 54	0.051 56
foreman	1.391 73	0.075 31	0.056 35	0.041 68

图 10 是由 mobile 视频序列中第 46 帧和第 48 帧采用不同算法内插出的第 47 帧图像, mobile 序列包含有丰富细节和慢速移动;(b)采用帧平均算法,明显地看出整幅图像出现了较大的模糊;(c)采用全搜索运动估计算法,虽然背景部分较清晰,但局部区域出现了明显的失真现象;(d)采用本文算法,可以看出内插图像清晰可见,视觉效果明显好于(b)(c)。

Foreman 图像中包含了丰富的前景运动,为观察这种特点图像的内插效果,实验对 foreman 视频序列进行处理。图 11 显示了 foreman 由第 99 帧和第 101 帧采用不同帧率提升算法内插出的第 100 帧图像;(b)(c)分别采用帧平均算法和全搜索算法,可以看到眼部出现了明显的失真和块效应;(d)采用本文算法得到的结果,可以看出,其效果接近于原始图像,具有良好的内插质量。

4 结束语

本文研究了一种快速多帧运动估计算法,该算法具有较快的收敛性和较高的估计精度,采用查表法大大降低了运算时

间,进而提出了基于多帧运动估计的帧率提升系统方案。实验结果表明,与现有常用的提升算法相比,无论是客观指标,还是主观质量,本文算法都具有显著优势,且能保证较好的实时性。

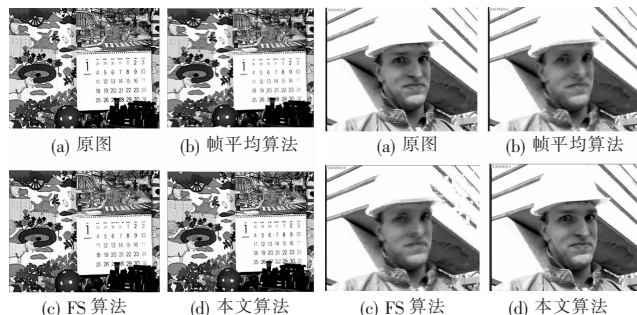


图 10 不同算法内插出 mobile 序列的第 47 帧

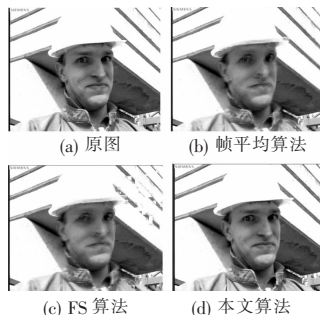


图 11 不同算法内插 foreman 的第 100 帧

参考文献:

- [1] BLANCHÉ P, VINCENT A, WANG D, *et al.* Advanced frame rate conversion and performance evaluation[J]. *The SMPTE Motion Imaging Journal*, 2006, 115(4): 153-159.
- [2] TEKALP A M. Digital video processing [M]. 2nd ed. [S. l.]: Prentice Hall, 1995.
- [3] 吴通. 基于 H. 264 块匹配运动估计的研究[D]. 武汉: 武汉理工大学, 2008.
- [4] LI Ren-xiang, ZENG Bing, LIU Ming. A new three-step search algorithm for block motion estimation[J]. *IEEE Trans on Circuits and Systems for Video Technology*, 1994, 4(4): 438-442.
- [5] SHAN Zhu, MA Kai-kuang. A new diamond search algorithm for fast block-matching motion estimation [J]. *IEEE Trans on Image Processing*, 2000, 9(2): 287-290.
- [6] CHOI B T, LEE S H, KO S J. New frame rate up-conversion using bi-directional motion estimation[J]. *IEEE Trans on Consumer Electronics*, 2000, 46(3): 603-609.
- [7] CHEN Tao. Adaptive temporal interpolation using bi-directional motion estimation and compensation[J]. *IEEE Trans on Image Processing*, 2002, 2(9): 313-317.
- [8] TOURAPIS H Y C, TOURAPIS A M. Fast motion estimation within the H. 264 codec[J]. *Proc of International Conference on Multimedia and Expo*, 2003, 3(7): 517-520.
- [9] BIERLING M. Displacement estimation by hierarchical block matching [C]//Proc of SPIE Conference on Visual Communication and Image Processing. 1988: 942-951.
- [10] KIM J W, LEE S U. On the hierarchical variable block size motion estimation technique for motion sequence coding [J]. *Optical Engineering*, 1993, 33(8): 2553-2561.
- [11] BURT P J, ADELSON E H. The Laplacian pyramid as a compact image code [J]. *IEEE Trans on Communication*, 1983, 31(4): 532-540.
- [12] WANG De-min, ZHANG Liang, VINCENT A. Motion-compensated frame rate up-conversion-part 1: fast multi-frame motion estimation [J]. *IEEE Trans on Broadcasting*, 2010, 56(2): 133-139.
- [13] KONRAD J, DUBOIS E. Bayesian estimation of motion vector fields [J]. *IEEE Trans on Image Processing*, 1992, 14(9): 910-927.
- [14] CHANG M M, TEKALP A M, SEZAN M I. Simultaneous motion estimation and segmentation [J]. *IEEE Trans on Image Processing*, 1997, 6(9): 1326-1333.
- [15] MCNERNEY P J, KONRAD J. Block-based MAP disparity estimation under alpha-channel constraints [J]. *IEEE Trans on Circuits System Video Technology*, 2007, 17(6): 785-789.
- [16] SHEN Huan-feng, ZHANG Liang-pei, HUANG Bo, *et al.* A MAP approach for joint motion estimation, segmentation, and super resolution [J]. *IEEE Trans on Image Processing*, 2007, 16(2): 479-490.
- [17] WU Jue, CHUNG A C S. A segmentation model using compound Markov random fields based on a boundary model [J]. *IEEE Trans on Image Processing*, 2007, 16(1): 241-252.
- [18] FEGHALI R. Multi-frame simultaneous motion estimation and segmentation [J]. *IEEE Trans on Consumer Electronics*, 2005, 51(1): 245-248.
- [19] GIBSON D, SPAN M. Robust optical flow estimation based on a sparse motion trajectory set [J]. *IEEE Trans on Image Processing*, 2003, 12(4): 431-445.
- [20] INCE S, KONRAD J. Occlusion-aware optical flow estimation [J]. *IEEE Trans on Image Processing*, 2008, 17(8): 1433-1451.
- [21] TEKALP M. Digital video processing [M]. Englewood Cliffs, NJ: Prentice Hall, 1995.
- [22] 宋锦. 帧速率上转换算法的研究与实现[D]. 天津: 天津大学, 2008.
- [23] 文俊, 刘重庆, 刘晓松. 视频标准帧速率上变换的自适应运动补偿方案[J]. *数据采集与处理*, 2003, 18(2): 204-208.