中图法分类号: TP391.41 **文献标志码**: A **文章编号**: 1006-8961(2011)08-1346-07
论文索引信息:郑嘉利, 覃团发, 倪光南. 结合率失真优化的自适应全局运动估计方法[J]. 中国图象图形学报,2011,16(8):1346-1352

结合率失真优化的自适应全局运动估计方法

郑嘉利[1),2)],覃团发[1)],倪光南[2)]
[1)](广西大学计算机与电子信息学院,南宁 530004) [2)](中国科学院计算技术研究所,北京 100190)

摘 要: 全局运动估计的关键在于全局运动模型的选择。结合率失真优化理论,提出一种自适应全局运动估计方法来达到编码优化的目的。该方法的主要思路是:对同一帧图像,分别使用平移运动模型、六参数运动模型和十二参数运动模型进行编码,用率失真优化算法计算 3 种运动模型下的拉格朗日代价函数值,拉格朗日代价函数值最小的运动模型被选为最佳的当前帧的运动模型。实验证明,该方法具有较好的鲁棒性,对不同分辨率的视频序列均有不同程度的编码增益。

关键词: 全局运动模型;全局运动估计;率失真优化

Adaptive global motion estimation method based on rate distortion optimization

Zheng Jiali[1),2)], Qin Tuanfa[1)], Ni Guangnan[2)]
[1)](*School of Computer and Electronic Information, Guangxi University, Nanning* 530004 *China*)
[2)](*Institute of Computing Technology, Chinese Academy of Sciences, Beijing* 100190 *China*)

Abstract: The key point of the global motion estimation is how to select a global motion model. In this paper, integrating with a rate distortion optimization algorithm, an adaptive global motion estimation method to optimize the coding is proposed. The scheme of the algorithm is as follows: for the same image, using the translated motion model, affine motion model and quadratic motion model to estimate the inter-frame motion and calculating their Lagrangian costs respectively. The motion model which yields least Lagrangian cost is selected as the best motion mode for current frame coding. Simulated results show that the proposed technique is robust and have better performance for various resolutions of video sequences.

Keywords: global motion model;global motion estimation;rate distortion optimization

1 全局运动估计的数学模型

在视频图像中,运动通常由摄像机运动和场景中的物体运动产生,由摄像机运动产生的运动将影响整个图像,称为全局运动。物体的运动称为局部运动。如果视频图像可以用一个全局运动建模,则用于视频编码时就可以节省大量的运动信息。全局运动估计(GME)[1-2]就是根据一定的镜头运动模型,利用视频序列中背景的运动信息得到该模型的具体参数。

要得到全局运动参数,首先要为背景图像运动确定一个合适的镜头数学模型,以描述全局运动的所有状况,包括平移、旋转、缩放和景深变化[3]等。人们提出各种 2 维运动模型[4]描述各种复杂的背景运动。在现今的 H.264 和 MPEG-4 标准里,运动估

收稿日期:2010-06-09;**修回日期**:2010-10-13
基金项目:广西自然科学基金项目(2011GXNSFA018169);广西教育厅科研项目(2010102LX016)。
第一作者简介:郑嘉利(1979—),男,副教授。2008 年于中国科学院计算技术研究所获计算机应用专业博士学位,主要研究方向为多媒体通信、视频编码技术。E-mail:lemontree312@live.cn。

计都是基于平移运动模型,如下式所示

$$\begin{aligned}\Delta x &= a_1\\ \Delta y &= a_2\end{aligned} \tag{1}$$

式中,$(\Delta x,\Delta y)$指的是当前点的运动矢量,a_1 代表物体水平方向移动的矢量,a_2 代表物体垂直方向移动的矢量。

平移运动模型是假设物体只在 2 维平面上做水平和垂直方向的平移运动,忽略旋转、缩放、错切等摄像机运动。然而,在实际应用中,摄像机运动,即全局运动往往占据整幅图像内运动集合的 50% 以上,如体育节目、车辆运动检测等。关于全局运动估计数学模型的研究成为全局运动估计算法的研究热点,不少学者提出各种不同的多项式全局运动估计数学模型,如 Rath 等人[5]提出一种四参数运动模型来模拟背景运动的平移、缩放和旋转效果,如下式所示

$$\begin{aligned}\Delta x &= a_3x - a_4y + a_1\\ \Delta y &= a_4x + a_3y + a_2\end{aligned} \tag{2}$$

式中,a_3 是摄像机镜头的缩放因子(也就是摄像机的焦距),a_4 是摄像机镜头的旋转因子。

Wiegand 等人[6]使用六参数的仿射变换运动模型来模拟更为灵活的平移、缩放和旋转背景运动效果,如下式所示

$$\begin{aligned}\Delta x &= a_1x - a_2y + a_3\\ \Delta y &= a_4x + a_5y + a_6\end{aligned} \tag{3}$$

式(3)可以变换为

$$\begin{pmatrix}\Delta x\\ \Delta y\end{pmatrix} = \lambda\begin{pmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix} + \begin{pmatrix}t_x\\ t_y\end{pmatrix} \tag{4}$$

式中,λ 是放大系数,φ 表示图像块的旋转角度,(t_x,t_y)表示图像块沿 x 和 y 方向的平移距离。对比式(3)和式(4),$a_1=a_5=\lambda\cos\varphi$,$a_2=-\lambda\sin\varphi$,$a_4=\lambda\sin\varphi$,$a_3=t_x$,$a_6=t_y$。

之后,Kunter 等人[7]用一个八参数的投影变换模型来描述任意的 3 维刚体运动

$$\begin{aligned}\Delta x &= \frac{a_1x + a_2y + a_3}{1 + a_7x + a_8y}\\ \Delta y &= \frac{a_4x + a_5y + a_6}{1 + a_7x + a_8y}\end{aligned} \tag{5}$$

Amri 等人[8]使用更复杂的十二参数二次方程变换模型来表征复杂的背景运动

$$\begin{aligned}\Delta x &= a_1x^2 + a_2x + a_3xy + a_4y^2 + a_5y + a_6\\ \Delta y &= a_7x_2 + a_8x + a_9xy + a_{10}y^2 + a_{11}y + a_{12}\end{aligned} \tag{6}$$

在目前的视频编码全局运动估计中,使用较普遍的是平移运动模型、六参数运动模型和十二参数运动模型。以往的研究表明,一般来说,越复杂的运动模型,可模拟的运动样式越多,运动估计越精确,编码压缩率越高。特别是在一些比较复杂的运动背景场景下,多参数运动模型(如上面列举的十二参数运动模型)往往比稀疏参数运动模型取得更好的运动预测效果。但是,多参数运动模型必须要付出更多的计算复杂度。由于全局运动估计是基于求解误差函数最小值的原理,在现有的算法框架下,如 Gauss-Newton 迭代法[9]和 Levenberg-Marquardt 迭代法[9],要估计的运动参数越多,迭代算法的复杂性越高(呈指数增长)。

考虑到视频背景内容的多样性,综合衡量计算复杂度和编码压缩效率,提出一种结合率失真优化的自适应全局运动估计方法。该方法的主要思路是:对同一帧图像,分别使用平移运动模型、六参数运动模型和十二参数运动模型进行编码,用率失真优化算法计算 3 种运动模型下的拉格朗日值,拉格朗日值最小的运动模型被选为最佳的当前帧的运动模型。具体步骤为:把图像划分为互不重叠的若干个图像块,把图像块的中心区域作为全局运动参数估计的初始化输入值;使用块匹配运动估计算法对全局运动中的大平移运动做初始化运动估计;用 Gauss-Newton 迭代法精细所得到的初始化运动估计结果,得到各个全局运动参数的值。以上步骤分别重复使用在平移运动模型、六参数运动模型和十二参数运动模型编码中。其中,在进行平移运动模型编码时,去掉 Gauss-Newton 迭代法步骤,因为平移运动模型公式里只有常数项。

下面以六参数运动模型为例,具体讨论以运动参数模型为基础的全局运动估计的迭代计算过程,同时描述本文所提出的结合率失真优化的自适应全局运动估计方法的技术细节。

2　全局运动参数估计

2.1　全局运动参数的初始估计

传统上的算法是把整幅图像作为全局运动估计的初始输入,这也是造成全局运动估计计算复杂度增大的一个原因。在我们提出的算法里,舍弃了把整幅图像作为全局运动估计的初始输入,而是把一幅图像划分为若干个子图像块,具体说来,就是把

CIF 图像划分了 3×3 个子图像块,把 QCIF 图像划分了 2×2 个子图像块。选取每个子图像块的中心区域作为全局运动估计的初始化输入。中心区域的大小是 64×64 像素。考虑到运动物体通常集中在图像中部,因此,位于图像中心的子图像块不作为初始化输入的候选值。

使用块匹配运动估计算法估算背景的平移运动,将得到的运动矢量作为下一步迭代式全局运动参数估计的初始化输入值。考虑到图像的静止区域(即运动矢量为零的块区域)往往映射到参考帧里也是同一个位置的区域,因此,把这些位于静止区域的块纳入全局运动估计中,不但干扰全局运动估计的精确性,而且增加计算复杂度。本文使用一种简单而有效的零运动矢量判别方法来剔除静止区域块。在做块匹配运动估计之前,首先,计算当前块和参考帧对应块之间的残差,如果亮度均方差 MSE(lum)和色度均方差 MSE(chr)同时满足以下条件,则当前块的模式是零矢量运动模式,如下式所示

$$\begin{aligned}\mathrm{MSE}(lum) &< \max\left(5.0, \frac{QP^2}{6}\right)\\ \mathrm{MSE}(chr) &< \max\left(5.0, \frac{QP^2}{12}\right)\end{aligned} \tag{7}$$

如果当前块的亮度均方差 MSE(lum)和色度均方差 MSE(chr)不同时满足以上条件,则当前块为有效的全局运动估计块,进行块匹配运动估计。

2.2 全局运动参数的最终估计

如式(2)所示,解六参数仿射运动方程的关键是解出 a_1, a_2, a_3, a_4, a_5 和 a_6 这 6 个运动参数。由于运动模型的运动参数的解具有不唯一性,因此需要使用层次性的迭代算法来估算运动参数的解。在本文中所使用的迭代算法是 Gauss-Newton 迭代法。在运动对象区域里,用均方预测误差极值法来估算仿射运动的 6 个参数,均方预测误差函数为

$$\sum_{x\in m, y\in m} DPD^2(x,y,a) = \sum_{x\in m, y\in m} E_n^2(x,y,a) = \sum_{x\in m, y\in m} (I_n(x,y)) - R_n[x+\Delta x(x,y,\alpha), y+\Delta y(x,y,\alpha)]^2 \tag{8}$$

式中,$DPD(x, y, a)$是当前帧第 i 个像素点的亮度值与参考帧中用全局运动参数位移后的对应插值像素点的亮度值之间的差。使用 Gauss-Newton 迭代法,式(8)可转化为

$$G\delta = \boldsymbol{g} \tag{9}$$

式中,δ 是每一次迭代过程中,参数初始值 a^{old} 和更新后的参数值 a^{new} 的差,即 $\delta = a^{\mathrm{new}} - a^{\mathrm{old}}$;$\boldsymbol{G}$ 是一个 6×6 的半正定对称矩阵,矩阵各个系数由 $DPD^2(x, y, a)$对运动参数 a_1, a_2, a_3, a_4, a_5 和 a_6 分别求二次偏导计算得出。$\boldsymbol{G}$ 矩阵系数为

$$\boldsymbol{G}(a) = \frac{DPD^2(x,y,a)}{\partial a_i \partial a_j} = \frac{DPD(x,y,a)}{\partial a_i} \cdot \frac{DPD(x,y,a)}{\partial a_i} \tag{10}$$

$\boldsymbol{g}$ 是一个六系数向量,有如下形式

$$\boldsymbol{g}(a) = DPD(x,y,a) \cdot \frac{\partial DPD(x,y,a)}{\partial a_i} \quad i = 1,\cdots,6 \tag{11}$$

把式(10)、式(11)代入式(9),可得出向量 δ 里各个运动参数的残差值。则更新后的运动参数值 $a^{\mathrm{new}} = \delta + a^{\mathrm{old}}$。将 a^{new}代入式(9)得出当前块内各像素点的运动矢量,通过运动补偿得到像素点预测值,进而算出当前块和预测块之间的预测均方差 $\mathrm{MSE}_{\mathrm{temp}}$。如果 $\mathrm{MSE}_{\mathrm{temp}}$满足以下条件,则停止迭代,产生最终仿射运动方程的 6 个运动参数;否则,将 a^{new}当成 a^{old}作为下一轮 Gauss-Newton 迭代的初始输入值,继续执行以上 Gauss-Newton 迭代步骤。

$$\mathrm{MSE}_{\mathrm{temp}} < \mathrm{MSE}_{\mathrm{min}} \tag{12}$$

式中,$\mathrm{MSE}_{\mathrm{min}}$是预先设置的 MSE 阈值,在本文中,$\mathrm{MSE}_{\mathrm{min}}$的值被设为 0.1。

3 率失真优化选择最佳运动模型

十二参数运动模型的全局运动参数计算方法与六参数运动模型的类似。把以上全局运动参数估计步骤重复使用在十二参数运动模型编码上,可相应得到当前帧的 12 个全局运动参数值,进而可计算出当前帧的预测残差。如前文所述,已经讨论了运动模型类型与编码性能以及编码比特数之间的关系。一般来说,运动模型越复杂,编码性能越好,同时也需消耗更多的比特数编码运动参数。我们的目的就是要选择一种运动模型,解决编码运动矢量所用的比特数与重构图像失真度之间的折衷问题,使得压缩码流大小与重构图像失真度都尽可能小。这一问题很自然可转化为率失真优化模式选择问题[10-11]。

H.264 中通过拉格朗日乘子法进行模式选择。拉格朗日乘子法是 R-D 优化的常用方法,它引入拉格朗日参数 λ,将每种模式的失真和编码比特以代价函数的形式表达出来,并根据编码结果在各个模

式中判别,使得代价函数最小的作为最优模式。代价函数定义为

$$J_{\mathrm{MODE}}(S_k, I_k \mid Q\lambda_{\mathrm{MODE}}) = D_{\mathrm{REC}}(S_k, I_k \mid Q + \lambda_{\mathrm{MODE}} \times R_{\mathrm{REC}}(S_k, I_k \mid Q)) \quad (13)$$

式中,I_k 表示某种编码模式;$D_{\mathrm{REC}}(S_k, I_k \mid Q)$表示重构图像与原始图像之间的失真度;$R_{\mathrm{REC}}(S_k, I_k \mid Q)$表示对宏块编码后数据及相关参数在码流中所占用的比特数,包括宏块头、运动矢量、运动补偿残差等所有信息编码所用的比特数;Q 表示量化参数;λ_{MODE} 表示拉格朗日参数。

率失真优化问题可归结为从编码参数集合中选择合适的参数,使得压缩码流大小与重构图像的失真度尽可能小的问题。这些参数包括模式信息、运动信息以及量化信息等。在本文中,主要的参数选择是何种类型的运动参数模型被选为最佳编码模式(假设在相同量化步长,相同块大小,使用相同数量的参考帧的前提下),同样也可使用式(13)所示的拉格朗日乘子法来进行模式选择。

依次用平移运动模型法、六参数运动模型法和十二参数运动模型法对测试序列进行全局运动估计补偿,在全局运动估计补偿、量化编码再解码重构过程中,使用拉格朗日乘子法把原图像和预测图像的失真度 D_{REC} 以及编码所使用的比特数 D_{REC} 代入式(13)可分别计算出各运动模型的代价函数值。设 $J_{\mathrm{translation}}$ 代表当前块使用平移运动模型的代价函数值,J_{affine} 代表当前块使用六参数运动模型的代价函数值,$J_{\mathrm{quadratic}}$ 代表当前块使用十二参数运动模型的代价函数值,则 3 者中代价函数值最小的,被选为最优模式,即 $\min\{J_{\mathrm{translation}}, J_{\mathrm{affine}}, J_{\mathrm{quadratic}}\}$。

4　仿真实验

实验采用 Immersive Media 公司制作的全景视频序列[12] Bridge 和 Garden 来进行测试。序列的分辨率是 2 048 × 768,属于柱面投影全景视频。将率失真优化选择最优运动模型方法集成到 H. 264/AVC 编码器中的可变块大小运动估计中,使得可变块大小运动估计编码的鲁棒性进一步得到拓展。

表 1 统计的是在本文提出的自适应全局运动估计方法中,各运动模型在不同块大小下被率失真优化选中的比例。由表中的数据不难看出,块尺寸越大,复杂运动模型被选中的可能性越高;反之,块尺寸越小,简单运动模型被选中的可能性越高。这是由于当块尺寸划分较大时,块内所包含的运动细节更多,更适合使用复杂运动模型;当块尺寸划分较小时,块的运动性质更容易被精细为某种单一的运动类型(一个极端的例子,当块大小是一个像素点时,运动类型可以规划为水平和垂直方向的平移运动)。因此,一般来说,简单运动模型用于编码运动矢量的比特数要比复杂运动模型多。

表 1　不同块大小模式下各运动模型被率失真优化选中的比例

Tab. 1　Percentage of motion model selected as the best model at different block sizes

块大小	运动模型	Bridge/%	Garden/%
16 × 16	平移运动模型	10.2	9.4
	六参数运动模型	68.3	12.6
	十二参数运动模型	21.5	78
8 × 8	平移运动模型	21.1	14.6
	六参数运动模型	53.3	30.6
	十二参数运动模型	25.6	54.8
4 × 4	平移运动模型	29.2	24.1
	六参数运动模型	48.3	26.3
	十二参数运动模型	22.5	49.6

本文所使用的自适应全局运动估计方法是基于拉格朗日乘子法,基本的规则是在编码运动矢量和预测残差的比特数,以及原图像和重构图像的失真度之间实现最佳折中。如表 2 所示,平移运动模型花费在运动矢量编码上的比特数最少,然而花费在预测残差上的比特数最多。平移运动模型、六参数运动模型与十二参数运动模型编码运动矢量的比特数依次递增;编码预测残差的比特数依次递减。这是由于随着运动模型的复杂度增加,运动参数个数增加,每一帧需要编码传送的运动矢量比特数也更多;由于对更多的摄像机运动效果模型化,运动估计效果更好,预测图像更接近原始图像,因此编码预测残差的比特数更少。从表 2 中可以看出,自适应全局运动估计方法编码运动矢量的比特数介于六参数运动模型方法与十二参数运动模型方法之间,但编码预测残差的比特数在测试的所有方法中最少,预测图像的峰值信噪比最高。

为了测试本文提出的自适应全局运动估计方法的性能,将该方法与分别单独使用平移运动模型、六参数运动模型和十二参数运动模型做编码性能比较。其中,如上文所述,平移运动模型是迄今

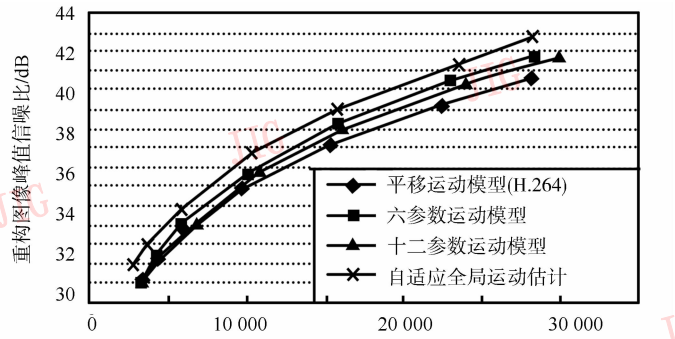
表 2 使用不同运动模型编码的运动矢量比特数和预测残差比特数以及峰值信噪比

Tab. 2 Bits spent on coding motion vectors and prediction error and PSNR by using difference motion models

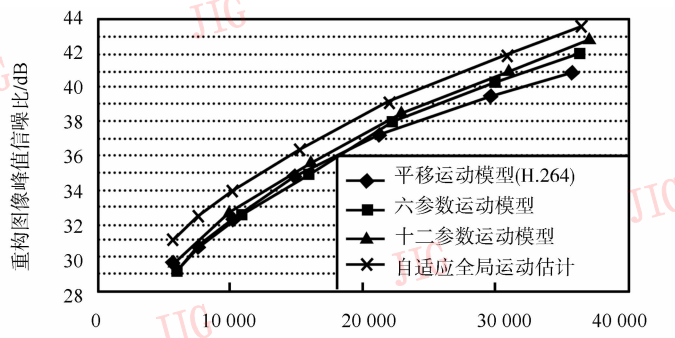
视频序列	运动模型	运动矢量/kbit	预测残差/kbit	预测图像噪比值信噪比/dB
Bridge	平移运动模型	26.16	212.31	20.25
	六参数运动模型	82.61	179.23	23.41
	十二参数运动模型	102.23	192.11	21.83
	本文所提出的方法	91.12	166.19	24.62
Garden	平移运动模型	22.95	203.42	21.22
	六参数运动模型	74.8	171.31	23.83
	十二参数运动模型	121.36	154.33	25.02
	本文所提出的方法	116.76	136.16	26.43

H.264/AVC 编码器所使用的运动估计模型。由图 1 的率失真优化曲线可以清楚看到,自适应全局运动估计方法的性能优于其他 3 种运动模型方法。相对传统的 H.264 所使用的平移运动模型方法,平均信噪比增益达到 2 dB。

如图 2 所示,预测图像中的灰色块是运动补偿后的残差块。残差块是原始图像与预测图像之间的残差值,预测图像中的残差块越多,表明运动估计的准确度越差。由图 2 可以很清楚地看到,使用自适应全局运动估计方法所生成的预测图像比使用传统的平移运动模型运动估计方法生成的预测图像所包含的残差块更少,因此,使用自适应全局运动估计方法所需要编码残差的比特数更少。尽管由于多参数运动模型需要编码传输更多的运动矢量个数来描述运动特征,编码运动矢量的比特数比平移运动模型要多,但是运动估计精确所带来的残差的下降和重构图像质量的上升,使得自适应全局运动估计方法总体编码性能比传统平移运动模型要好。



(a)视频序列 Bridge (2 048×76 825频/s)的RD性能图



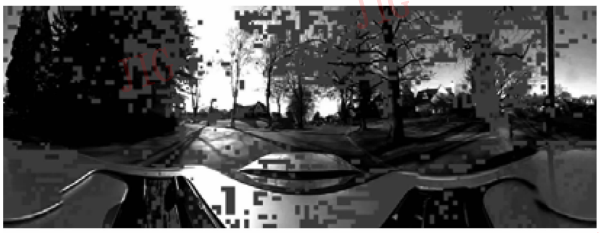
(b)视频序列 Garden (2 048×76 825频/s)的RD性能图

图 1 使用不同全局运动模型编码的 RD 性能对比图

Fig. 1 Rate-distortion curves by using different global motion model

(a) 视频序列 Bridge 使用 H.264/AVC 运动估计方法运动估计/补偿后的预测图像

(b) 视频序列 Garden 使用 H.264/AVC 运动估计方法运动估计/补偿后的预测图像



(c) 视频序列 Bridge 使用自适应全局运动估计方法运动估计/补偿后的预测图像

(d) 视频序列 Garden 使用自适应全局运动估计方法运动估计/补偿后的预测图像

(e) 视频序列 Bridge 原图像

(f) 视频序列 Garden 原图像

图 2　分别使用 H.264/AVC 运动估计方法与自适应全局运动估计方法运动补偿后的预测图像对比

Fig. 2　Prediction drawings produced by proposed method vs H.264/AVC

5　结　论

由于在连续的视频序列里,摄像机的运动并不是一成不变的,帧与帧之间摄像机的运动类型通常都不一样;帧内既有全局运动也有局部运动;而且最复杂的运动参数模型虽然能兼顾摄像机运动的各种类型,但是,由于需要编码传输的全局运动参数比特数较多,不一定是当前块最理想的运动参数模型模式。本文所提出的结合率失真优化的自适应全局运动估计方法对处理每一个块的 2 维运动具有较好的鲁棒性,兼顾全局运动与局部运动。具有局部运动的块的运动矢量单独编码传输,类似于 H.264/AVC 里的运动矢量编码方法;具有全局运动的块(即具有同一运动趋势的块)只需在宏块头插入运动参数模型类型标记,就可在解码端构造与编码端一致的运动参数模型,同时根据编码传输过来的摄像机运动参数,可恢复出每个块的运动矢量,进而对每个块进行运动补偿,生成解码图像。经仿真实验,本方法对各种分辨率的视频序列(QCIF、CIF、4CIF)均有不同程度的编码增益,对高分辨率的无缝拼接全景视频,平均信噪比增益可达到 2 dB。

参考文献(References)

[1] Haller M, Krutz A, Sikora T. Evaluation of pixel-and motion vector-based global motion estimation for camera motion characterization [C]//Proceedings of 10th International Workshop on Image Analysis for Multimedia Interactive Services. Piscataway, NJ, USA: IEEE Computer Society, 2009: 49-52.

[2] Zheng Jiali, Zhang Yongdong, Ni Guangnan. A fast global motion estimation method for panoramic video coding [C]// Proceedings of Pacific-Rim Conference on Multimedia.

Heidelberg:Springer-Verlag,2007: 152-155.

[3] Keller Y, Averbuch A. Fast gradient methods based on global motion estimation for video compression[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(4): 300-309.

[4] Guo Baolong, Ni Wei, Yan Yunyi. Video Signal Processing Techniques in Communications[M]. Beijing: Publishing House of Electronic Industry, 2007:87-118. [郭宝龙,倪伟,闫允一. 通信中的视频信号处理[M]. 北京:电子工业出版社,2007: 87-118.]

[5] Rath G B, Makur A. Iterative least squares and compression based estimations for a four-parameter linear global motion model and global motion compensation[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(7): 1075-1099.

[6] Wiegand T, Steinbach E, Girod B. Affine multipicture motion-compensated prediction[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2005, 15(2): 197-209.

[7] Kunter M, Krutz A, Mandal M, et al. Optimal multiple sprite generation based on physical camera parameter estimation[C]// Proceedings of Visual Communications and Image Processing. Bellingham WA,USA SPIE, 2007, 6508 (2): 0B.1-0B.10.

[8] Amri S, Zagrouba E, Barhoumi W. Background construction for video sequences with complex motions[C]//Proceedings of the international Group of e-Systems Research and Applications. Taiwan: IEEE Press, 2008: 11-26.

[9] Press W H. Numerical Recipes in C: the Art of Scientific Computing[M]. Cambridge: Cambridge University Press, 2003: 47-58.

[10] Sarwer M G, Po L M, Guo Kai, et al. Transform-domain rate-distortion optimization accelerator for H.264/AVC video encoding [J]. International Journal of Signal Processing, 2009, 5(3): 238-248.

[11] Tu Yukuang, Yang Jarferr, Sun Mingting. Efficient rate-distortion estimation for H.264/AVC Coders[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16: 600-611.

[12] Immersive Media Company. Panoramic Video Sequences [EB/OL]. (2008-10-12)[2010-05-26]. ftp://ftp.tnt.uni-hannover.de/pub/3dav.