

# 一种快速鲁棒的全局运动估计算法

秦 荣, 马志强, 张晓燕, 陈广居

(空军工程大学信息与导航学院, 陕西西安, 710077)

**摘要** 局部运动矢量会严重影响全局运动估计的结果, 不仅降低全局运动估计的准确度, 而且消耗大量的运算时间。针对上述问题提出了一种快速鲁棒的全局运动估计算法。首先采用快速的块匹配算法得到运动矢量场, 然后通过改进的梯度均值残差法去除估计的不准确运动矢量, 提高估计出的运动矢量场的准确性并使得全局运动估计所处理的宏块尽可能是全局运动块; 其次, 基于六参数仿射模型估计出初始全局运动参数, 为了消除错误运动矢量产生的误差, 使用匹配加权函数的方法对估计结果进行进一步修正, 最终得到最优的全局运动参数。实验结果表明该算法兼顾了全局运动估计的准确性和鲁棒性, 对大多数视频序列的全局运动估计具有算法复杂度低、运动参数估计准确的效果。

**关键词** 全局运动估计; 块匹配算法; 梯度均值残差; 仿射模型

**DOI** 10.3969/j.issn.1009-3516.2012.06.012

**中图分类号** TP391.4 **文献标志码** A **文章编号** 1009-3516(2012)06-0055-05

通常的视频序列中, 摄像机自身的运动会导致整个视频图像发生变化, 因此需要根据摄像机运动模型进行全局运动估计, 消除摄像机运动的影响。全局运动估计是计算机视觉、视频处理的基本手段之一, 被广泛的应用于视频分割<sup>[1]</sup>、视频压缩、电子稳像等领域, 有着重要的理论意义和高度的实用价值。

近年来, 国内外众多学者对此展开了深入研究, Hussein Alzoubi 和 W. David Pan 在 LMA 基础上, 利用降采样模板仅选取部分像素点参与迭代计算, 将算法的运算速度提高了十几倍。但同时一定程度上牺牲了估计的准确度<sup>[2]</sup>。Sorwar 假设前景运动对象一般位于图像的中心位置, 因而只采用边缘图像块的运动矢量估计运动参数, 减少了计算量, 但仍受假设条件的限制<sup>[3]</sup>。Bin Q 采用图像金字塔模型, 顶层通过结合三步搜索法和运动矢量预测的方法进行快速初始估计, 并考虑邻域信息剔除外点, 取得较好的估计效果<sup>[4]</sup>。陈正华为了提高估计的准确度, 提出了一种基于运动矢量可靠性分析的全局运动估计算法, 通过提取图像块的匹配特征并用支持向量机的方法将运动矢量分类, 提高了全局运动估计的性能<sup>[5]</sup>。郑嘉利针对全局运动估计提出了一种结合亮度残差阈值法和运动矢量残差阈值法的渐进精细阈值方法, 分两步逐步细分前景物体对象和背景区域, 提高了估计的准确性和鲁棒性<sup>[6]</sup>。虽然全局运动估计算法得到了一些改进, 但仍面临着估计精度和运算复杂度之间的矛盾, 存在估计的准确性不高, 计算量大和实时性差的问题。

针对上述问题, 本文提出了一种匹配加权的全局运动估计算法, 该算法的总体流程见图 1。

首先采用快速的块匹配算法得到运动矢量场, 然后通过改进的梯度均值残差法去除估计的不准确运动

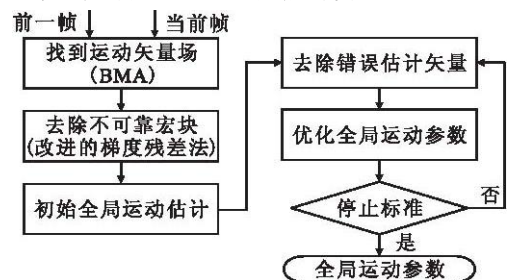


图 1 算法框图

Fig. 1 Block diagram of the proposed algorithm

收稿日期: 2012-06-26

基金项目: 陕西省自然科学基金资助项目(2010JM8014); 中国博士后科学基金资助项目(20100471838); 中国博士后特别基金资助项目(201104787)

作者简介: 秦 荣(1982-), 女, 陕西兴平人, 硕士生, 主要从事视频图像处理研究. E-mail: qinrong82@163.com

矢量;在基于六参数仿射模型估计的初始全局运动参数基础上,使用匹配加权迭代进一步优化,最终得到最优的全局运动参数。

## 1 块匹配算法得到运动矢量场

基于块的运动估计对于运动矢量场的表达清晰,算法简单,因而得到了广泛的关注,研究人员提出了很多快速算法<sup>[2-5]</sup>。块匹配算法基本思路是:首先,将视频图像序列中当前帧图像分割成大小相同的不重叠区域,称之为宏块,并在参考帧图像内预先设定一个搜索范围,按照一定的匹配准则搜索与之具有最小匹配误差的块,匹配块与当前块之间的坐标位移就是运动矢量 MV。

本文用绝对差值和(SAD)来衡量当前块与候选块的相似度,宏块的大小为  $8 \times 8$  像素。为了进一步降低运动估计的复杂度,采用4步搜索快速运动估计得到运动矢量场。图2给出了对 Silent 序列(CIF 格式)的第26、27帧进行4步搜索快速运动估计得到的运动矢量场。



图2 Silent 序列用块匹配算法得到的运动矢量场

Fig. 2 Silent estimated MV field using BMA

## 2 改进的梯度均值残差法去除外点

大多数视频运动场景既包含有全局运动又有局部运动,如果让获取的所有运动矢量都参与全局运动估计,那么处于局部运动区域的运动矢量(对全局运动参数估计造成干扰的外点),会严重影响全局运动估计的结果,不仅降低全局运动估计的准确度,而且消耗大量的运算时间。本文解决这一问题的思路是在全局运动估计之前,用梯度残差法将不属于全局运动内点集合的外点剔除出去,使得全局运动估计所处理的宏块尽可能都是全局运动块,从而提高估计的准确度减小运算量。

通过对视频图像的分析可以发现,全局运动和局部区域的纹理分布不相同。如果采用绝对差值和(SAD)最小作为块匹配准则进行块匹配运算,相对平坦的区域或者细节很小的区域,运算结果往往得不到正确的运动矢量。针对该问题,本文首先对大小为  $8 \times 8$  像素的各个宏块进行分析,判断其属于全局运动块或是局部运动块,然后去除局部运动块,使得参加全局运动估计所处理的宏块尽可能都是全局运动块,从而提高全局运动估计的准确度。

改进的梯度均值残差法的基本思想是:首先用 sobel 算子分别计算第  $k$  帧当前块内的所有像素的梯度均值  $G(i, j)$ ,如式(1)所示,  $m \times n$  为块大小,然后根据得到的运动矢量通过运动补偿后第  $k+1$  帧中对应块内所有像素的梯度均值  $G'(i, j)$ ,见式(2):

$$G(i, j) = \sum_{i=1}^m \sum_{j=1}^n \frac{G_{i,j}}{m \times n} \quad (1)$$

$$G'(i, j) = \sum_{i=1}^m \sum_{j=1}^n \frac{G'_{i,j}}{m \times n} \quad (2)$$

最后计算  $G(i, j)$  和  $G'(i, j)$  的残差绝对值大小,得到均值残差  $G_B$ ,见式(3):

$$G_B = \sum_{i=1}^m \sum_{j=1}^n \frac{|G_{i,j} - G'_{i,j}|}{m \times n} \quad (3)$$

考虑到通过运动补偿后的预测帧与当前帧之间亮度残差较大的集合是外点集合,本文将通过式(3)得到的第  $i$  个宏块的亮度残差  $G_i$  进行大小排序,按照一定的比例(本文到 10%)去除最高亮度残差块,这样可以快速去除部分外点。

对剩余的宏块通过比较  $G_B$  与阈值  $C_g$  的大小,把梯度均值残差小于阈值  $C_g$  的块判定为局部运动区域

块,从全局运动宏块集合中剔除出去。对阈值算法而言,阈值选取是关键所在,如果阈值  $C_g$  选择过大,那么全局运动估计的准确度就不高;如果阈值  $C_g$  选择过小,则很容易收敛于局部区域。本文的阈值  $C_g$  由下式确定:

$$C_g = \alpha \frac{1}{N} \sum_{i=1}^N G_i \quad (4)$$

式中: $N$  是去除最大亮度残差后的块的个数; $\alpha$  是根据不同视频图像类型统计得到的,本文中  $\alpha$  取值为 1.3。

### 3 匹配加权的全局运动估计

#### 3.1 初始全局运动参数估计

图像中各点的全局运动矢量都是在同一摄像机模型下的运动,可用同一个参数模型来表示。本文综合考虑运动估计效果和运算度,采用六参数仿射模型如下:

$$\begin{cases} x'_m = a x_m + \omega y_m + \alpha \\ y'_m = \omega x_m + \alpha y_m + \alpha \end{cases} \quad (5)$$

式中:参数  $a, \omega, \alpha, \alpha$  用于控制旋转与缩放; $\omega, \alpha$  决定平移幅度; $(x_m, y_m)$  表示第  $k$  帧第  $m$  块的位置; $(x'_m, y'_m)$  表示第  $k+1$  帧第  $m$  块的位置。

六参数仿射模型只需 3 个背景像素点的运动矢量即可根据式(5)求得 6 个运动参数,但若仅用很少几个像素的运动矢量来估计全局运动模型参数,得到的结果精度非常低,因此需要更多运动矢量参与全局运动模型参数估计来提高准确度,因此参数估计是一个解矛盾方程组的问题,用最小二乘法来迭代求解方程组,直到最后全局运动模型参数值收敛于一个较为稳定的值。设经过匹配算法估计出第  $k$  帧中第  $m$  块  $(x_m, y_m)$  在第  $k+1$  帧中的位置为  $[\tilde{x}_m, \tilde{y}_m]$ ,而经过六参数仿射模型变换到第  $k+1$  帧的位置为  $[x'_m, y'_m]$ ,视频图像梯度均值残差法去除外点后的宏块数为  $N_m$ ,定义误差函数如下:

$$E(\mathbf{P}) = \sum_{m=1}^{N_m} [(x_m - x'_m)^2 + (y_m - y'_m)^2] \quad (6)$$

将式(5)代入式(6)可以得到:

$$E(\mathbf{P}) = \sum_{m=1}^{N_m} [(\tilde{x}_m - a x_m - \omega y_m - \alpha)^2 + (\tilde{y}_m - \omega x_m - \alpha y_m - \alpha)^2] \quad (7)$$

那么最优全局运动参数  $\mathbf{P} = [a, \omega, \alpha, \alpha]^T$  就是使式(7)达到最小值的矢量  $\mathbf{P}$ :

$$\mathbf{P}_{\text{opt}} = \arg \mathbf{P} \min E(\mathbf{P}) \quad (8)$$

最优全局运动参数可通过最小二乘法对式(8)求最小值而得到。

#### 3.2 匹配加权

为了消除错误运动矢量带来的误差,本文提出了一种采用匹配加权函数的方法对估计结果进行进一步修正。令  $D'_k$  表示式(5)计算出的所有内点在第  $k+1$  帧中的位置集合, $\tilde{D}_k$  表示由块匹配得到的所有内点在第  $k+1$  帧中的位置集合。我们定义误差为:

$$E_D = D'_k - \tilde{D}_k \quad (9)$$

$E_D$  表示 2 个集合中对应元素相减形成的集合。分别求出  $E_D$  的均值  $\mu_D$  与方差  $\sigma_D$ ,由于前景运动目标和背景的全局运动偏差较大,表现为一些奇异点,其统计特性离均值较远,因此为每个宏块分配一个权值  $\omega_m$ :

$$\omega_m = \begin{cases} 1, & \text{if } |e_m - \mu_D| \leq 3\sigma_D \\ 0, & \text{if } |e_m - \mu_D| > 3\sigma_D \end{cases}, \quad e_m \in E_D \quad (10)$$

其中与背景统计特性相差较小的点被保留,与背景统计特性相差较大的点被判定为异常点,予以剔除。结合以上 2 个权值,将式(7)修正如下:

$$E(\mathbf{P}) = \sum_{m=1}^{N_m} \{ \omega_m [(\tilde{x}_m - a x_m - \omega y_m - \alpha)^2 + (\tilde{y}_m - \omega x_m - \alpha y_m - \alpha)^2] \} \quad (11)$$

进行迭代估计出全局运动参数,通过估计的最优全局运动参数对第  $k$  帧进行全局运动补偿后,再和第  $k+1$  帧进行差分运算,这样得到的差分图像可消除由于运动所引起的背景变化的影响。

## 4 实验结果及分析

本文实验是在 Internet Pentium Dual-Core E5200 2.5 GHz 处理器,1 GB 内存的硬件环境下进行的,仿真实验环境是 Windows XP 下的 MATLAB 7.10。实验分别对 Foreman(QCIF)、Coastguard(CIF)、Ferrari(CIF)、Stefan(CIF)、Tennis(PAL)、Bridge(PAL)等序列进行了全局运动估计实验,采用残差图对运动估计的准确性进行评价。

图 3(a)、(b)、(c)分别显示了 Foreman 序列的第 26、27 帧及预测帧图像,图 4(a)、(b)、(c)分别显示了 Stefan 序列的第 2、3 帧及预测帧图像,图 3(d)、图 4(d)是用直接帧差得到的图像,图 3(e)、(f)和图 4(e)、(f)分别是采用文献[4]提出的方法和本文提出的方法进行全局运动估计补偿后得到的残差图。从实验结果可以看出:经过全局运动补偿后的帧差图与直接帧差图比较,较好地去除了背景区域的影响,突出了感兴趣的具有局部运动的前景区域。

对视频图像进行分析可以发现,外点主要是由多匹配或前景运动造成的错误匹配构成的,当背景区域较为复杂时往往会导致较大的估计误差,比如体育运动的场地情况(见图 4)。通过比较图 4(e)、(f),可以看出本文提出的方法在代表全局运动的背景区域误差很小,得到了准确的估计结果。



图 3 Foreman 序列的全局运动估计结果

Fig. 3 Foreman global motion estimation result

将图像的平均峰值信噪比作为估计算法的评价指标,分别采用本文提出的方法和文献[4]的方法对 Stefan 序列前 60 帧图像和 Foreman 序列前 30 帧图像的平均峰值信噪比进行统计比较,从图 5(a)、(b)可以看出,本文提出的方法平均峰值信噪比明显高于文献[4]的方法,也就是说明了本文提出的全局运动估计算法准确度更高。



图 4 Stefan 序列的全局运动估计结果

Fig. 4 Stefan global motion estimation result

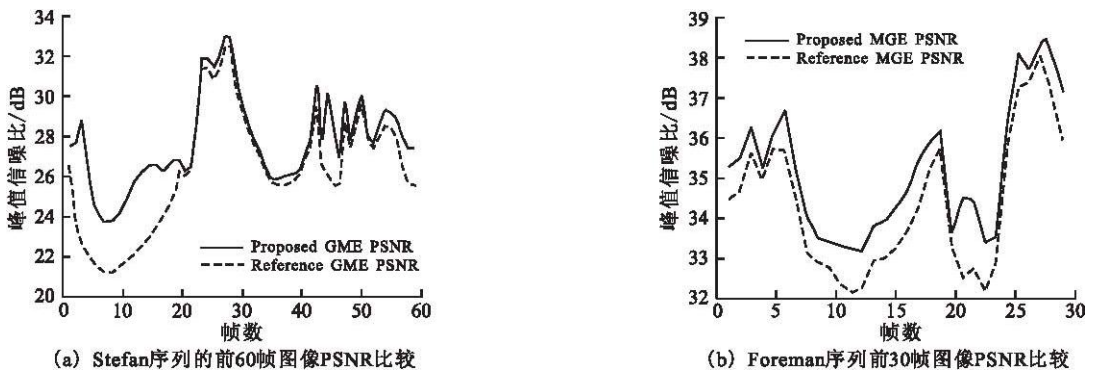


图 5 视频序列的 PSNR 比较

Fig. 5 PSNR of the video sequence

实验还分别对 5 个不同的视频序列采用本文提出的方法与其他文献的方法对全局运动估计时间进行了统计,对比数据见图 6,可以看出针对不同格式的视频序列,本文算法都有效减少了运算时间,提高了运动估计的实时性。

## 5 结语

本文提出了一种快速鲁棒的全局运动估计算法。首先采用快速的块匹配算法得到运动矢量场,然后通过改进的梯度均值残差法去除估计的不准确运动矢量;在基

于六参数仿射模型估计的初始全局运动参数基础上,使用匹配加权迭代进一步优化,最终得到最优的全局运动参数。通过实验将该算法与文献的方法进行比较,结果表明,该算法兼顾了全局运动估计的准确性和鲁棒性,对大多数视频序列的全局运动估计具有算法复杂度低、运动参数估计准确的效果。但由于块匹配估计运动矢量时的孔径问题和噪声问题,全局运动区域中还存在一些前景运动物体像素。我们下一步的研究工作是通过全局运动区域的一些辅助处理,得到更为精确的全局运动估计结果。

## 参考文献(References):

- [1] Liu Z, Zhang Z Y, Shen L Q. Moving object segmentation in the H.264 compressed domain[J]. Optical engineering, 2007, 46(1):017003.
- [2] Hussein Alzoubi, Pan W David. Fast and accurate global motion estimation algorithm using pixel subsampling[J]. Information sciences, 2008, 178(17):3415-3425.
- [3] Sorwar G, Murshed M, Dooley L. Fast global motion estimation using iterative least 2 square estimation technique [C]//IEEE 4th int conf inf comm signal process. Singapore:IEEE press, 2003:282-286.
- [4] Bin Q, Ghazal M, Amer A. Robust global motion estimation oriented to video object segmentation [J]. IEEE transactions on image processing, 2008, 17(6):958-967.
- [5] 陈正华, 章毓晋. 基于运动矢量可靠性分析的视频全局运动估计[J]. 清华大学学报:自然科学版, 2010, 50(4):623-627.  
CHEN Zhenghua, ZHANG Yujin. Global motion estimation based on the reliability analysis of motion vectors [J]. Journal of tsinghua university: science and technology, 2010, 50(4):623-627. (in Chinese)
- [6] 郑嘉利, 覃团发. 一种面向快速全局运动估计的渐进精细阈值方法[J]. 计算机科学, 2011, 38(6):293-297.  
ZHENG Jiali, QIN Tuanfa. Progressive fine granular threshold method oriented to fast global motion estimation [J]. Computer science, 2011, 38(6):293-297. (in Chinese)
- [7] 吴淑英, 陈卫东. 一种基于预测的改进的三步搜索算法[J]. 计算机应用与软件, 2008, 25(1):232-234.  
WU Shuying, CHEN Weidong. A prediction-based improved three step search algorithm[J]. Computer applications and software, 2008, 25(1):232-234. (in Chinese)
- [8] Li R X, Zeng B. A new three-step search algorithm for block motion estimation[J]. IEEE trans CASVT, 1994, 4(8):438-442.
- [9] 蒋晓悦, 赵荣椿. 几种块匹配运动估计算法的比较[J]. 计算机应用研究, 2004, 21(7):1-4.  
JIANG Xiaoyue, ZHAO Rongchun. Compare of several block matching algorithms[J]. Application research of computers, 2004, 21(7):1-4. (in Chinese)
- [10] Turagad D, Alkanhal M. Search algorithms for block matching in motion estimation [EB/OL]. (2010-12-01)[2012-06-26]. <http://www.ece.cmu.edu/~ee899/project/deepak-mid.htm>.

(编辑:徐楠楠)

(下转第 65 页)

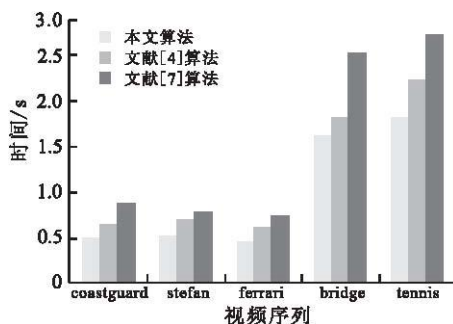


图 6 不同视频序列的估计时间比较

Fig. 6 Comparison of the estimated time