

文章编号:1007-2780(2012)02-0263-04

基于仿射变换的快速全局运动估计算法

王思珺^{1,2}, 赵 建^{1*}, 韩希珍¹

(1. 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033, E-mail: wangsijun0616@163.com;

2. 中国科学院 研究生院, 北京 100039)

摘 要: 全局运动估计是计算机视觉、视频处理等领域广泛采用的手段之一。为了降低全局运动估计计算复杂度, 提出一种基于像素梯度的全局运动估计方法。就运动模型而言, 采用了复杂性与准确性较好均衡的六参数仿射模型。为了提高计算速度, 采用了两层金字塔进行多分辨率计算, 而且在每层迭代计算中都抽取梯度较大的特征点进行计算, 同时结合了高斯-牛顿优化计算方法。实验验证了提出的全局运动估计算法的计算效率和精度。

关 键 词: 全局运动估计; 参数模型; 仿射变换; 迭代最小二乘法; 金字塔计算

中图分类号: TP391

文献标识码: A

DOI: 10.3788/YJYXS20122702.0263

Fast Global Motion Estimation Algorithm Based on Affine Transformation

WANG Si-jun^{1,2}, ZHAO Jian^{1*}, HAN Xi-zhen¹

(1. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences,

Changchun 130033, China, E-mail: wangsijun0616@163.com;

2. Graduate University of Chinese Academy of Sciences, Beijing 100039, China)

Abstract: Global Motion Estimation is an important tool widely used in computer vision, video processing, and other fields. To reduce the computational complexity of global motion estimation, a global motion estimation method based on the pixel gradient strategy was proposed. With regard to global motion models, six-parameter affine model is adopted because of its reasonable tradeoff between complexity and accuracy. In order to accelerate the algorithm, two-level pyramid is utilized in the calculation. It uses those gradient bigger pixels in the calculation of every pyramid level, and combines Gauss-Newton algorithm. The extensive experiments show that the proposed method is efficient and robust in terms of both computational complexity and accuracy.

Key words: global motion estimation; parameter model; affine transformation; iterative least squares method; pyramid calculation

1 引 言

图像中像素点的位置可能会随着时间而改

变,称之为像素运动。像素运动是由场景中实体运动或者摄像机运动产生的,并可用不同的运动模型来描述。通常,视频序列图像由前景和背景

收稿日期: 2011-05-20; 修订日期: 2011-06-15

基金项目: 吉林省科技厅重点项目(No. 20100310)

作者简介: 王思珺(1983—),女,吉林省吉林市人,硕士,主要研究方向是数字图像处理、视频压缩。

* 通讯联系人, E-mail: zhaojian6789@yahoo.com

构成,背景的运动是由摄像机运动造成的,称为全局运动;而前景物体所表现出来的运动是前景物体相对于摄像机的运动,称为局部运动。全局运动估计(Global Motion Estimation, GME)就是根据一定的摄像机运动模型,利用视频序列中背景的运动信息得到该模型的参数。全局运动估计广泛应用于视频压缩^[1-2]、视频编码^[3]、视频监控、视频检索^[4]等领域,因此,研究全局运动估计是非常有意义的。

全局运动估计的研究方法可划分为两类:两步法^[5]和基于像素的梯度下降法^[3,6-7]。前者通过块匹配法获得局部运动矢量,然后用迭代最小二乘法对局部运动矢量进行处理估计全局运动参数。后者采用梯度下降优化算法使当前帧和参考帧之间的残差平方和最小。

由于图像背景像素数量很多,而且有前景物体的局部运动干扰,因此在进行全局运动估计时有两个难点:首先是计算量大,难以达到实时应用的要求。其次是由于噪声点多,使估计的准确性难以得到保证。Rath^[5]提出利用迭代最小二乘的方法,在每次迭代过程中排除与前一次估计误差较大的运动矢量,经过有限次迭代能够消除误差较大的局部运动,但是无法剔除误差较小的局部运动,而且计算量大。Wang Demin^[8]通过计算运动场的偏导数,提出了一种快速的估计算法。Li^[9]提出了基于背景提取来提高全局运动估计精度和速度的方法,Moscheni^[10]也采用了背景提取的思路。G Sorwar^[11]基于运动前景一般分布在图像中间的假设,采用位于图像边缘的宏块矢量来估计运动参数,从而提高计算速度和排除前景运动对象的干扰。本文对全局运动估计进行了深入的研究,提出了一种基于像素梯度的改进的全局运动估计算法,并通过实验证明了本文算法的快速性、有效性。

2 全局运动仿射模型

针对不同的应用场景和摄像机运动特点,全局运动估计可以建立不同的运动模型。如四参数^[5]、六参数仿射^[12]模型,八参数透视模型^[3]以及八参数双线性模型^[4]等。一般,描述全局运动的模型参数越多,精度越高,计算复杂度也越高。六参数仿射模型可描述图像的平移、缩放、旋转、仿射等线性变换,而且其在模型复杂度和建模能

力之间取得了较好的折衷,本文利用该参数模型来进行全局运动估计,其表达式如下:

$$\begin{cases} x' = ax + by + c \\ y' = dx + ey + f \end{cases} \quad (1)$$

其中, (x, y) 代表参考帧中像素的坐标, (x', y') 代表当前帧中像素的坐标。定义参数矢量 $\mathbf{P} = (a, b, c, d, e, f)^T$, 其中分量 a, b, c, d 与缩放、旋转运动有关, 分量 e, f 与平移运动有关。只需要 3 个像素点的运动矢量即可求得 6 个运动参数的值。但是如果用很少像素的运动矢量来估计全局运动参数, 得到的结果精度很低, 因此需要更多像素的运动矢量参与全局运动估计, 这样, 参数估计就成了一个解矛盾方程组的问题, 可用最小二乘法来计算, 如高斯-牛顿迭代法^[13]。则对于当前帧中的 N 个宏块, 全局运动参数在最小二乘准则下的最优估计为:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^N x_k^2 & \sum_{k=1}^N x_k y_k & \sum_{k=1}^N x_k \\ \sum_{k=1}^N x_k y_k & \sum_{k=1}^N y_k^2 & \sum_{k=1}^N y_k \\ \sum_{k=1}^N x_k & \sum_{k=1}^N y_k & N \end{pmatrix}^{-1} \begin{pmatrix} \sum_{k=1}^N x_k x'_k \\ \sum_{k=1}^N y_k x'_k \\ \sum_{k=1}^N x_k x'_k \end{pmatrix} \quad (2)$$

$$\begin{pmatrix} d \\ e \\ f \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^N x_k^2 & \sum_{k=1}^N x_k y_k & \sum_{k=1}^N x_k \\ \sum_{k=1}^N x_k y_k & \sum_{k=1}^N y_k^2 & \sum_{k=1}^N y_k \\ \sum_{k=1}^N x_k & \sum_{k=1}^N y_k & N \end{pmatrix}^{-1} \begin{pmatrix} \sum_{k=1}^N x_k y'_k \\ \sum_{k=1}^N y_k y'_k \\ \sum_{k=1}^N x_k y'_k \end{pmatrix} \quad (3)$$

由于外点的存在, 最小二乘计算需要迭代进行, 使全局运动模型参数值收敛于一个较为稳定的值, 迭代估计的方法可以获得较精确的参数估计结果, 但是迭代计算方式一般需要耗费较大的计算资源。

3 改进的迭代求解算法

在运用递归最小二乘法估计全局运动参数时, 初始化输入值是图像中全部宏块经过块匹配

运动估计后得到的运动矢量,计算点越多一次迭代的计算量越大,但是如果计算点数太少就可能不能正确反映全局运动信息,抗干扰能力下降。如何有效选取可靠的计算点,减少冗余点数,是加速全局运动估计计算的一个关键。

在全局运动估计中都涉及到梯度运算,由于梯度运算会受到噪声点的影响,而且梯度值越小的话,该点越容易受到噪声点的影响,这些点梯度值就不可靠,如果让这些点参与计算,不仅会使计算量增大,而且会使计算结果不准确。将参与计算的像素集合缩小为梯度较大的像素子集能明显提高计算速度,而不会影响算法的鲁棒性。本文提出如下的特征点选取公式:

$$G = \sqrt{\nabla_x^2 + \nabla_y^2}$$

(4)

其中, ∇_x 和 ∇_y 分别代表该点的亮度值 $I(x,y)$ 在 x 方向和 y 方向的梯度,通过设定相应的阈值,可以将 G 大于阈值的点作为特征点。经多次实验计算,本文选取梯度值在前 10% 的大梯度像素点参与计算。

由于块匹配计算量大,本文采用金字塔分层的块匹配技术进行帧间运动估计,以提高运算速度,算法具体描述如下:

(1)用高斯图像构造法构造图像金字塔。将图像分为两层,由粗到精逐步求解运动参数矢量。底层为原始图像,顶层为原始图像经过 2:1 的采样,大小变为底层图像 1/4(长、宽都减少一半)。

(2)对金字塔顶层图像进行全局运动估计计算。

①用高斯滤波器进行滤波,滤除图像的随机噪声。

②把整幅图像分成互不重叠的块,块大小是 20×20 ,用块匹配算法对每一个块作运动估计,得到每一个块的运动矢量 $(\Delta x, \Delta y)$ 。理论上,块的大小越小,运动估计结果越精确,但同时复杂度大幅提高,本算法采用 20×20 块,兼顾考虑计算复

杂度与运动估计精确行。同时对块匹配运动估计,本文采用的是三步搜索快速运动估计。

③用 Sobel 算子计算当前块内所有像素的梯度值 $S(i,y)$,取梯度值较大的前 10% 像素点作为特征点参与迭代计算,求得相应的参数集。

(3)将金字塔顶层求得的参数映射到底层,作为底层运动参数估计的初始值,重复顶层相同的运动参数计算步骤,计算出底层图像的运动参数,即最终求得的参数集。

4 实验及结果分析

本文算法的硬件实验平台为 PC 计算机: Intel E5200 Pentium(R) Dual-Core CPU 2.5GHz, 4GB 内存。编程语言: C++, 软件环境: Visual C++ 6.0。

实验采用的测试序列为: RAW 格式的 $1\,004 \times 1\,004$ 大小的 Google 序列。分别取该序列的第 01 帧和第 02 帧进行背景的全局运动参数计算,图像序列如图 1 所示,计算结果如表 1 所示。



图 1 Google 序列图像. ①第 01 帧; ②第 02 帧.
Fig. 1 Image of Google sequence

表 1 Google 序列全局运动估计计算结果
Table 1 GME result of Google sequence

计算方法	参数 a	参数 b	参数 c	参数 d	参数 e	参数 f	运算时间/ms
算法改进前	1.0	0.0	1.7	0.0	1.0	1.9	156
算法改进后	1.0	0.0	1.5	0.0	1.0	1.7	47

从表 1 可看出,本文算法在保证计算精度的前提下,其运算时间下降为传统算法的 30.1%,大幅提高了运算效率。

5 结 论

针对传统全局运动估计算法中迭代方法计算量大、耗时的问题,本文提出了一种基于梯度的全

局运动估计快速算法,并将这种全局运动估计算法和已有的算法进行了比较。实验结果显示了该方法的有效性和优越性。此外,本文的方法并不仅仅适用于六参数仿射模型,它也可以类似地用来估计二参数、四参数模型,以及八参数双线性模型等,具有较强的实际意义。

参 考 文 献:

- [1] Mahmoud H, Bayoumi M. An efficient low-bit rate adaptive mesh-based motion compensation technique [C]// *Proc. 2nd Int. Workshop on Digital and Computational Video*. Tampa, Florida, USA; IEEE Computer Society, 2001:164-172.
- [2] Yosi K, Amir A. Fast gradient methods based on global motion estimation for video compression [J]. *IEEE Trans. Circuit Syst. Video Technol.*, 2003, 13(4):300-309.
- [3] Dufaux F, Konrad J. Efficient, robust, and fast global motion estimation for video coding [J]. *IEEE Transactions on Image Processing*, 2000, 9(3):497-501.
- [4] 俞天力,章毓晋. 基于全局运动信息的视频检索技术 [J]. *电子学报*, 2001, 29(12A):1794-1798.
- [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] Bin Qi, Ghazal M, Amer A. Robust global motion estimation oriented to video object segmentation [J]. *IEEE Transactions on Image Processing*, 2008, 17(6):958-967.
- [7] 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.
- [8] Wang D, Wang L. Fast and robust algorithm for global motion estimation [C]// *Proc SPIE Conf on Visual Communications and Image Processing (VCIP)*, San Jose, CA, USA; SPIE 1997, 3024:1144-1151.
- [9] Li H, Liu G, Li Y, et al. A robust, efficient, and fast global motion estimation method from MPEG compressed video [C]// *IEEE Proc. PCM2002, LNCS 2532*, Taiwan; SPIE, 2002: 151-158.
- [10] Moscheni F, Dufaux F, Kunt M. A new two-stage global/local motion estimation based on a background/foreground segmentation [C]// *IEEE Proc. ICASSP95*, Detroit, MI; IEEE, 1995:2261-2264.
- [11] Sorwar G, Murshed M, Dooley L. Fast global motion estimation using iterative least-square estimation technique [C]// *IEEE Proc. ICICSPCM 2003*, Singapore, IEEE, 2003:282-286.
- [12] 贺玉文,赵黎,钟玉琢,等. 快速鲁棒的全局运动估计算法 [J]. *软件学报*, 2001, 12(8):1220-1228.
- [13] Press W H, Flannery B P, Teukolsky S A, et al. *Numerical recipes in C: The Art of Scientific Computing* [M]. Cambridge: Cambridge University Press, 2003: 32-102.