

文章编号:1001-9081(2006)05-1024-03

## 基于低频子图的运动估计算法

余志雄, 孙洪, 贺军  
(武汉大学电子信息学院, 湖北 武汉 430079)  
(zhixiong\_yu@sohu.com)

**摘要:**为了减少运动估计的计算复杂度并提高其搜索性能,提出了一种基于低频子图的运动估计算法。该算法首先将当前帧和参考帧通过低通滤波器得到数据量减为1/4的低频子图,然后用全搜索得到最佳匹配的低频子块,最后对低频子块所覆盖的原参考帧区域进行精细搜索得到最优点。实验结果表明,与传统运动估计算法相比,该算法提高了运动估计的精确度和降低了计算复杂度,并且对各类视频都有很高的鲁棒性。

**关键词:**低频子图;运动估计;Haar 小波;块匹配算法;视频编码  
**中图分类号:**TP391.41    **文献标识码:**A

### Motion estimation based on low frequency sub-image

YU Zhi-xiong, SUN Hong, HE Jun  
(School of Electronic Information, Wuhan University, Wuhan Hubei 430079, China)

**Abstract:** In order to reduce the computational complexity of motion estimation and improve the accuracy, a fast motion estimation algorithm based on low frequency sub-image was proposed. The algorithm firstly puts Current Frame and Reference Frame through a low-pass filter to get low frequency sub-images with quarter pixels, then attains the best matched sub-block by Full Search, at last the area of Reference Frame covered by the sub-block are searched to get the best matched point. Compared with traditional algorithm, experimental results show that this algorithm gives a significant improvement in accuracy for motion estimation and reduces the computational complexity, and possesses strong robustness in different kinds of video sequences.

**Key words:** low frequency sub-image; motion estimation; Haar wavelet; block matching algorithm; video coding

### 0 引言

为了减少视频信号的数据量,ISO/IEC MPEG 工作组和ITU-T VCEG 工作组制定了一系列的视频压缩编码标准<sup>[1,2]</sup>,包括 H. 261、H. 263、H. 264、MPEG1、MPEG2、MPEG4 等。考虑到视频图像序列的相邻帧之间存在着很大的时间相关性,即时间冗余度,这些标准普遍采用了一种基于块匹配的运动估计算法来减小帧间时间冗余度,从而降低码率。

在运动估计算法中,全搜索算法(FS)对搜索窗内的所有点进行匹配搜索从而得到全局最优点,显然其搜索精度最高,但因为极大的计算复杂度而难以在实际中得到应用。近年来,国内外先后提出一系列的运动估计算法来降低运动估计模块的计算复杂度,这些算法一般都是利用运动估计误差曲面函数的极值特性,采用不同的搜索规则,通过逐步缩小搜索的范围从而实现快速搜索。二维对数法(TDL)、三步搜索算法(TSS)<sup>[3]</sup>、四步搜索算法(FSS)<sup>[4]</sup>等算法通过设定不同大小的模板,按照一定的规则逐步减小搜索范围从而减少搜索的点数。菱形搜索(DS)<sup>[5]</sup>、线性搜索<sup>[6]</sup>等算法利用误差曲面函数的极值特性,采用特定的模板进行逐步逼近最优点。

但是,这些算法往往是建立在以下假设的基础上的:  
1) 误差曲面函数存在唯一的全局最优点;2) 误差曲面呈单峰分布;3) 误差曲面不存在零梯度区域。然而,实际的误差曲面不一定满足上述特点,对于复杂视频图像尤其如此。所以常规运动估计算法往往会因为陷入局部最小点而降低搜索的性能。

利用低频子图全局覆盖、人眼敏感性和数据量低等性能,

本文提出了一种基于低频子图的运动估计算法。根据人眼视觉系统和图像的能量分布特性,人眼比较敏感和能量比较集中的部分都在图像的低频部分。并且低频子图覆盖了整个搜索区域的全部低频信息,但是低频子图大小却只有原来图像的1/4。因此,本文先采用 Haar 小波基进行图像分解得到低频子图,然后对低频子图进行全搜索得到初步运动矢量,最后在原参考帧中对初步运动矢量进行精确定位。实验结果表明,与传统的运动估计算法相比,该算法在降低了计算复杂度的同时提高了运动搜索的精确度,在仅用约 16.625 个搜索点数的情况下,无论是对局部运动、旋转运动还是全局运动的视频序列,其峰值信噪比(PSNR)与全搜索相差不到0.52dB。

### 1 运动估计算法的性能分析

常规运动估计算法通常都是基于块匹配的运动估计算法。在运动估计中,块匹配算法的思想描述如下:

将当前帧 C 分割成大小为  $M \times N$  的块,假设块中所有的像素都做相同的平移运动,块在当前帧 C 中相对于参考帧 P 的位置偏移  $d = [dx, dy]$  称为该块的运动矢量 MV(Motion Vector),  $dx$  和  $dy$  分别为运动矢量的水平分量和垂直分量。运动估计就是在参考帧 P 的给定窗口  $\Omega$  中搜索当前块的最佳匹配块并得到其运动矢量的过程。匹配过程可以用下式表述:

$$(\hat{d}_x, \hat{d}_y) = \arg \min_{(i,j) \in \Omega} MCF(i,j) \quad (1)$$

其中  $MCF$  为匹配代价函数(Matching Cost Function)。在传统的视频编码算法中,普遍使用  $SAD$ (Sum of Absolute Difference) 值作为匹配代价函数。 $SAD$  的表达式为:

收稿日期:2005-11-29;修订日期:2006-02-20    基金项目:国家自然科学基金资助项目(60372057)

作者简介:余志雄(1981-),男,湖北洪湖人,硕士研究生,主要研究方向:视频压缩编码;孙洪(1954-),女,湖北武汉人,博士生导师,主要研究方向:统计信号处理、图像分析、通信信号处理;贺军(1978-),男,湖北宜昌人,博士研究生,主要研究方向:图像和视频处理。

$$SAD(i,j) = \sum_{m=1}^M \sum_{n=1}^N |f_2(m,n) - f_1(m+i, n+j)| \quad (2)$$

$f_1(\cdot)$  和  $f_2(\cdot)$  分别为参考帧和当前帧中对应位置的像素值。

对于大小为  $M \times N$  的块,计算一次 SAD 值需要进行  $M \times N$  次比较运算和  $2 \times M \times N - 1$  次加法运算。若实现一个块的运动估计需要的搜索点数为  $T$ ,则完成一个块运动估计共需要  $T \times M \times N$  次比较运算和  $T \times (2 \times M \times N - 1)$  次加法运算。

可以看出,采用 SAD 值作为代价函数的块匹配算法中,计算复杂度与搜索的点数以及匹配块的大小近似成正比。

## 2 低频子图

### 2.1 子带分解

将一幅图像分解成一系列限带分量的集合,称为子带,它们可以重组在一起无失真的重建原始图像<sup>[7]</sup>。每个子带通过对输入进行带通滤波器得到。由于只含有原始图像的部分频率信息,每个子带图像的数据量都会极大减少。其中低频子带涵盖了原始图像的低频信息,其重建图像为原始图像的低分辨率图像。

利用小波的快速变换原理,对图像先进行行处理,再进行列处理,则可以得到原始图像的 4 个子带图像。图像  $I(m,n)$  的 4 频段子带分解的过程如图 1 所示。其中  $h_0(n)$  和  $h_1(n)$  组成一对正交滤波器,  $h_0(n)$  代表低通滤波器,  $h_1(n)$  代表高通滤波器;  $LL(m,n)$ 、 $LH(m,n)$ 、 $HL(m,n)$ 、 $HH(m,n)$  为分解后的 4 个子带图像。

### 2.2 Haar 小波基提取低频子图

Haar 小波基是最古老也是最简单的正交小波。对于  $N$  维的 Haar 小波分解,其子带的每个点只与原图的  $N \times N$  个点相关。对于 2 维的 Haar 小波变换,其变换矩阵  $H_2$ <sup>[7]</sup> 为:

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

由于在块匹配算法中常数因子并不改变算法的判定结果,为了减少算法的计算量,可以对编码帧和参考帧的变换矩阵同时乘以  $\sqrt{2}$  实现整数变换。

此时的正交滤波器为:  $h_0 = [1 \ 1]$ ,  $h_1 = [1 \ -1]$ 。则低频子图的提取只需对原始图像的行和列两次进行低通滤波即可。此时,低频子图为原始图像对应 4 个位置像素点值的和。

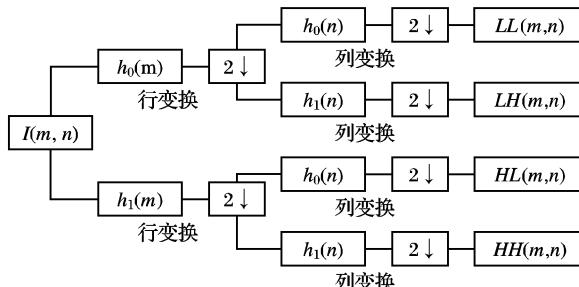


图 1 子带分解流程

$I(m,n)$  为原始图像,  $LL(m,n)$  就是所求的低频子图。

### 2.3 低频子图的性能

根据子带分解的特性,通过使用以上算法进行一阶子带分解,得到的低频子图具有如下特性:

- (1) 是原始图像的缩略图,与原始图像形状近似;
- (2) 包含原始图像的绝大部分能量;
- (3) 图像的水平大小和垂直大小均缩小为原始图像的  $1/2$ ,整个面积降为原始图像的  $1/4$ ;

(4) 每个像素仅与原始图像的 4 个对应像素有关。此时,每个子块需要  $3/4 \times M \times N$  个加法的小波变换计算量,近似为 0.375 个搜索点。

根据以上特性,对原图进行运动估计可以转化为:先对低频子图进行运动搜索得到低频子图的运动矢量  $d' = [d'_x, d'_y]$ , 再利用运动矢量  $d'$  对原图的 4 个对应位置进行精确定位。低频子图运动矢量  $d'$  的计算公式为:

$$(\hat{d}'_x, \hat{d}'_y) = \arg \min_{(i', j') \in \Omega'} SAD(i', j') \quad (3)$$

其中  $\Omega'$  代表低频子图搜索窗里面的所有位置。

## 3 基于低频子图的运动估计算法

### 3.1 低频子图运动估计的算法描述

为了介绍本算法的具体实现过程,这里采用搜索步长为 7 的搜索窗、 $16 \times 16$  的图像块进行运动搜索。则低频子图运动估计算法步骤如下:

(1) 用 Haar 小波基对当前帧  $C$  和参考帧  $P$  同时进行低频分解得到低频子图  $C'$  和  $P'$ 。

(2) 将低频子图  $C'$  划分为  $8 \times 8$  大小的子块,令低频子图子块的顶点为  $(i', j')$ ;对每个子块在低频参考帧  $P'$  内进行搜索窗为  $\pm 3$  的全搜索得到低频子图的运动矢量  $[d'_x, d'_y]$ 。

(3) 考虑到每个子块位置包含了原图 4 个匹配点的信息,以点  $(2i' + 2d'_x, 2j' + 2d'_y)$  为顶点,搜索参考帧  $P$  内覆盖低频信息的 4 个位置并得到最佳匹配点。此时的运动矢量即为算法的运动矢量  $[d_x, d_y]$ 。

### 3.2 性能分析

为了与传统算法的计算复杂度相统一,这里将低频子图的一次搜索等价于原图的  $1/4$  次搜索。则:

(1) 因为低频子图采用的是全搜索方法,所以总能够找到最匹配的低频子块,保证了搜索块中残差的低频分量最少。由于视频图像的能量总是集中在低频部分,所以低频子图的最佳匹配点往往覆盖了原图的最佳匹配点。

(2) 低频子图的搜索步长为原图的一半,使得搜索面积仅为原始面积的  $1/4$ ,即搜索点数降为原图的  $1/4$ 。

(3) 对于步长为 7 的搜索窗,本算法共包含 49 个子图搜索点和 4 个原图搜索点。根据上面的分析,加上小波变换引入的计算量后,整体运算量相当于 16.625 个原块的搜索,相对传统全搜索算法,计算量大幅下降。

(4) 由于不同于一般的快速运动估计算法,低频子图的搜索范围覆盖了原始图像的整个搜索窗,局部运动、旋转运动和全局运动对该算法的性能影响不大。

## 4 实验结果分析比较

为了验证本文算法的性能,在相同条件下对以下几种算法做了比较。实验采用的图像均为 CIF (Common Intermediate Format) 标准的灰度格式,匹配的准则为 SAD 准则,宏块大小为  $16 \times 16$ ,搜索窗大小为  $15 \times 15$ 。对三种典型的序列采用了 50 帧的搜索检测其平均性能。其中 claire 为典型的视频会议小序列运动; foreman 为前景运动加背景小幅度旋转运动; stefan 为大范围的剧烈运动。

从表 1 和表 2 中的结果可以得出以下结论:

(1) 对于小运动视频序列 (claire),菱形搜索采用大小两种模板的匹配,能快速精准的获取最优点,在平均 13.38 个搜索点的情况下,其平均 PSNR 值与全搜索相差仅 0.01dB。不过此时其他快速算法也均能获得较优的搜索结果,低频子图搜索和三步搜索与全搜索的差异分别为 0.07dB 和 0.09dB。

(2) 对于包含旋转、缩放、斜切等几何变换的视频序列 (foreman),即使无场景变化,由于基于块匹配算法的搜索原理,也不能得到很好的搜索结果,搜索的性能一般取决于搜索的范围。低频子图搜索范围覆盖了整个搜索窗并得到了最匹配的低频部分,因此平均 PSNR 值与全搜索相差不大,仅

0.52dB;而三步搜索和菱形搜索由于减少了搜索范围,平均PSNR值分别降低了1.89dB和1.50dB。

表1 各种搜索算法的PSNR(峰值信噪比/dB)

|        | claire | foreman | stefan |
|--------|--------|---------|--------|
| 全搜索    | 39.95  | 33.12   | 24.58  |
| 三步搜索   | 39.86  | 31.23   | 21.28  |
| 菱形搜索   | 39.94  | 31.62   | 22.06  |
| 低频子图搜索 | 39.88  | 32.60   | 24.24  |

表2 各种搜索算法的平均搜索点数

|        | claire | Foreman | stefan |
|--------|--------|---------|--------|
| 全搜索    | 255    | 255     | 255    |
| 三步搜索   | 25     | 25      | 25     |
| 菱形搜索   | 13.38  | 17.23   | 19.70  |
| 低频子图搜索 | 16.625 | 16.625  | 16.625 |

(3)对于大运动量的视频序列(stefan),由于减少了搜索范围,常规运动搜索算法都容易陷入局部最小值点,所以搜索的效果会大幅下降。与全搜索的PSNR值相比,三步搜索算法下降3.30dB,菱形搜索算法下降2.52dB。但是由于低频子图搜索算法覆盖了全部搜索范围,其PSNR值仅下降0.24dB。另外,由于偏离最佳搜索点,菱形搜索算法的迭代次数也会相应增多,此时菱形搜索算法的19.70个搜索点数也略高于低频子图搜索算法的16.625个搜索点。

(4)对于不同情况的视频序列,低频子图搜索算法都有较少的搜索点数,而且PSNR值与全搜索算法相差都很小,说明视频图像运动剧烈与否对低频子图搜索算法的精确度影响不大。而三步搜索算法和菱形搜索算法关于小运动序列claire的效果很好,但是对于旋转序列foreman和大运动序列stefan效果则偏差。这主要是因为此时误差曲面存在多个极

值点,导致搜索范围容易脱离全局最优点。

总之,由于通常视频图像能量主要集中在低频部分,本文提出的算法是一种运算速度快、适应范围广、搜索结果鲁棒性高的快速运动估计算法。

## 5 结语

本文根据低频子图的特性,先对低频子图进行全搜索得到低频子图的运动向量,再以该运动向量的两倍为中心搜索低频信息覆盖的4个匹配位置,实现精度为一个像素的运动估计算法,该算法在搜索方式和性能上类似于全搜索,计算量上接近常规快速搜索算法,而且对大运动和小运动视频序列均有比较高的峰值信噪比。

## 参考文献:

- [1] ITU-T Rec. H. 263, Video Coding for Low Bitrate Communication [S].
- [2] ITU-T Rec. H. 264/ISO/IEC11496-10, Advanced Video Coding [S].
- [3] LI R, ZENG B, LIOU ML. A new three-step search algorithm for block motion estimation[J]. IEEE Trans Circuits Systems for Video Technology, 1994, 4 (4) : 438 - 442.
- [4] PO ML, MA CW. A novel four-step search algorithm for fast block motion estimation[J]. IEEE Trans Circuits Systems for Video Technology, 1996, 6(6) : 313 - 317.
- [5] JAIN J, JAIN A. Displacement measurement and its application in interframe image coding[J]. IEEE Trans Communication, 1981, 29 (12) : 1799 - 1808.
- [6] DING GG, GUO BL. New Fast Motion Estimation Algorithm Based on Line Search[J]. Journal of xi'an JiaoTong University, 2004, 38 (2) : 36 - 39.
- [7] GONZALEZ RC, WOODS RE. 冈萨雷斯·数字图像处理[M]. 第2版. 阮秋琦, 阮宇智,译. 北京:电子工业出版社,2003.

(上接第1023页)

种方法都能去掉该场景下的强阴影,而图2(d)中左边运动目标的一部分也被当作阴影去掉,从图2(e)看出本文算法效果比图2(d)好,能避免误检,且能检测出场景上部的运动小目标。

图3显示了室内的阴影抑制,图3(a)为原始视频,图3(b)为提取的背景,图3(c)是未经过阴影抑制的运动前景检测结果,图3(d)为根据文献[3]方法进行阴影抑制的结果,图3(e)为本文算法进行阴影抑制的结果。由图3(d)与图3(e)可知,两种算法相比都能消除场景中运动目标脚下的阴影,但明显文献[3]中方法会在运动目标的胸部产生更多误检,本文的阴影抑制算法有更好的效果。

## 4 结语

本文基于混合高斯模型提出一种改进的阴影抑制算法,该算法首先应用改进的运动前景检测算法,完成了运动前景与静态背景的分离,然后利用疑似阴影模型完成了疑似阴影检测,最后应用改进的阴影抑制算法完成了运动目标阴影的去除。实验表明,在光照条件复杂的室内、室外环境下,与已有的典型算法相比,本文算法处理速度与效果方面都有明显的优势。

## 参考文献:

- [1] YONEYAMA A, YEH CH, KUO CCJ. Moving cast shadow elimination for robust vehicle extraction based on 2D joint vehicle/shadow models[A]. Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS'03)[C]. Miami, FL, USA. IEEE Computer Society, 2003. 229 - 236.
- [2] SALVADOR E, CAVALLARO A, EBRAHIMI T. Cast shadow seg-

mentation using invariant color features[J]. Computer Vision and Image Understanding, 2004, 95(2) : 238 - 259.

- [3] BAISHENG C, YUNQI L. Indoor and outdoor people detection and shadow suppression by exploiting HSV color information[A]. Proceedings of the Fourth International Conference on Computer and Information Technology[C]. Wuhan, China. IEEE Computer Society, 2004. 137 - 142.
- [4] MARTEL-BRISSON N, ZACCARIN A. Moving cast shadow detection from a gaussian mixture shadow model[A]. CVPR 2005[C]. IEEE Computer Society Conference, 2005.
- [5] PRATI A, MIKIC I, TRIVEDI MM, et al. Detecting moving shadows: algorithms and evaluation[J]. IEEE transactions on pattern analysis and machine intelligence, 2003, 25(7) : 918 - 923.
- [6] POWER PW, SCHOONEES JA. Understanding background mixture models for foreground segmentation[A]. Proceedings of Image and Vision Computing New Zealand[C]. Auckland, New Zealand, 2002. 267 - 271.
- [7] CHEUNG SCS, KAMATH C. Robust techniques for background subtraction in urban traffic video[A]. Proceedings of Electronic Imaging: Visual Communications and Image Processing 2004 (Part One)[C]. San Jose, California. Bellingham, WA: SPIE, 2004, (5308) : 881 - 892.
- [8] NADIMI S, BHANU B. Physical models for moving shadow and object detection in video[J]. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 2004, 26(8) : 1079 - 1087.
- [9] WANG H, SUTER D. A re-evaluation of mixture-of-gaussian background modeling[A]. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)[C]. Pennsylvania, USA, 2005. 1017 - 1020.