



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2013
08
VOL.18

ISSN1006-8961
CN11-3758/TB



北魏



隋



初唐

数字文化遗产研究 P968

综述

尺度不变特征变换算子综述

刘立, 詹茵茵, 罗扬, 刘朝晖, 彭复员 885

图像处理和编码

改进的H.264帧间模式选择算法

任克强, 张旭光, 罗会兰 893

均质雾天下摄像机动态标定的车速估计

宋洪军, 陈阳舟, 郜园园 901

基于人脸五官特征的空域差错掩盖算法

张江鑫, 谢晋, 邝万坤 913

采用圆谱-傅里叶矩的图像区域复制粘贴篡改检测

秦娟, 李峰, 向凌云, 殷蓓蓓 919

图像分析和识别

基于视频感知哈希的视频篡改检测与多粒度定位

朱映映, 文振焜, 杜以华, 邓良太 924

视觉注意模型的道路监控视频关键帧提取

刘云鹏, 张三元, 王仁芳, 张引 933

格拉斯曼流形上的半监督判别分析

姜伟, 陆瑶, 杨炳儒 944

精神疲劳实时监测中多面部特征时序分类模型

陈云华, 张灵, 丁伍洋, 严明玉 953

图像分类中的概率乘积核函数

杨赛, 赵春霞 961

轮廓整体结构约束的壁画图像相似性度量

唐大伟, 鲁东明, 杨冰, 许端清 968

应用色彩空间聚类方法实现道路建模

向宸薇, 王拓, 于舰 976

图像理解和计算机视觉

杂乱背景和摄像机移动下的时空兴趣点检测

刘长红, 陈勇, 王明文 982

形状和颜色混合不变量在图像检索中的应用

公明, 曹伟国, 李华 990

计算机图形学

结合统计分类和边缘特征的最优视图提取

潘翔, 陈敖, 章国栋 1004

医学图像处理

局部熵驱动下的脑MR图像分割与偏移场恢复耦合模型

张建伟, 杨红, 陈允杰, 方林, 詹天明 1011

改进的模糊聚类实质肺结节3维分割

李翠芳, 聂生东, 王远军, 孙希文, 郑斌 1019

遥感图像处理

高光谱海岸带区域分割的活动轮廓模型

王相海, 金弋博 1031

改进的基于非局部均值的极化SAR相干斑抑制

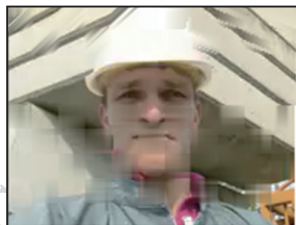
赵忠民, 赵拥军, 牛朝阳 1038

小波变换和稀疏表示相结合的遥感图像融合

刘婷, 程建 1045

《中国图象图形学报》论文摘要写作要求

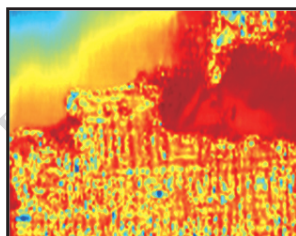
1054



基于人脸五官特征的空域差错掩盖算法 (第913页)



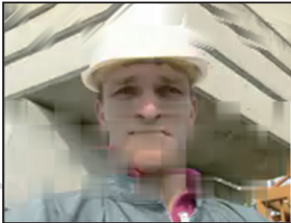
应用色彩空间聚类方法实现道路建模 (第976页)



改进的基于非局部均值的极化SAR相干斑抑制 (第1038页)

CONTENTS

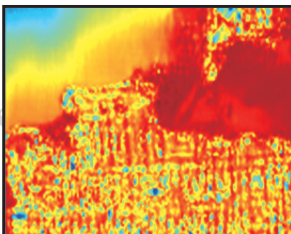
JOURNAL OF IMAGE AND GRAPHICS



Spatial error concealment algorithm based on the facial features (P913)



Road modeling using color spatial clustering (P976)



Improved polarimetric SAR speckle filter based on non-local means (P1038)

Review

- Summarization of the scale invariant feature transform
Liu Li , Zhan Yinyin, Luo Yang, Liu Chaohui, Peng Fuyuan 885

Image Processing and Coding

- Improved inter frame mode selection algorithm for H.264
Ren Keqiang, Zhang Xuguang, Luo Huilan 893
- Vehicle velocity estimation based on camera calibration under homogenous fog
Song Hongjun, Chen Yangzhou, Gao Yuanyuan 901
- Spatial error concealment algorithm based on the facial features
Zhang Jiangxin, Xie Jin, Kuang Wankun 913
- Detection of image region copy-move forgery using radial harmonic Fourier moments
Qin Juan, Li Feng, Xiang Lingyun, Yin Changming 919

Image Analysis and Recognition

- Video forgery detection and multi-granularity location based on video perceptual hashing
Zhu Yingying, Wen Zhenkun, Du Yihua, Deng Liangtai 924
- Key frame extraction based on the visual attention model for lane surveillance video
Liu Yunpeng, Zhang Sanyuan, Wang Renfang, Zhang Yin 933
- Semi-supervised discriminant analysis on Grassmannian manifold
Jiang Wei, Lu Yao, Yang Bingru 944
- Time-series classification model based on multiple facial feature for real-time mental fatigue monitoring
Chen Yunhua, Zhang Ling, Ding Wuyang, Yan Mingyu 953
- Probability product kernels in image classification
Yang Sai, Zhao Chunxia 961
- Similarity metrics between mural images with constraints of the overall structure of contours
Tang Dawei, Lu Dongming, Yang Bing, Xu Duanqing 968
- Road modeling using color spatial clustering
Xiang Chenwei, Wang Tuo, Yu Jian 976

Image Understanding and Computer Vision

- Spatio-temporal interest point detection in cluttered backgrounds with camera movements
Liu Changhong, Chen Yong, Wang Mingwen 982
- Application of combined shape and color invariants to image retrieval
Gong Ming, Cao Weiguo, Li Hua 990

Computer Graphics

- Best view selection based on statistical classification and edge features
Pan Xiang, Chen Ao, Zhang Guodong 1004

Medical Image Processing

- Brain MR image segmentation and bias correction model based on local entropy
Zhang Jianwei, Yang Hong, Chen Yunjie, Fang Ling, Zhan Tianming 1011
- Segmentation of sub-solid pulmonary nodules based on improved fuzzy C-means clustering
Li Cuifang, Nie Shengdong, Wang Yuanjun, Sun Xiwen, Zheng Bin 1019

Remote Sensing Image Processing

- The active contour model for segmentation of coastal hyperspectral remote sensing image
Wang Xianghai, Yin Yibo 1031
- Improved polarimetric SAR speckle filter based on non-local means
Zhao Zhongmin, Zhao Yongjun, Niu Chaoyang 1038
- Remote sensing image fusion with wavelet transform and sparse representation
Liu Ting, Cheng Jian 1045

中图法分类号: TP919.8 文献标识码: A 文章编号: 1006-8961(2013)08-0893-08

论文引用格式: 任克强, 张旭光, 罗会兰. 改进的 H. 264 帧间模式选择算法[J]. 中国图象图形学报, 2013, 18(8): 893-900. [DOI:10.11843/jig.20130802]

改进的 H. 264 帧间模式选择算法

任克强, 张旭光, 罗会兰

江西理工大学信息工程学院, 赣州 341000

摘要: 针对 H. 264 帧间模式选择算法存在的问题, 提出一种帧间模式选择快速算法。算法采用 8×8 块运动矢量同质性分析对宏块模式进行选择, 缩小候选模式范围, 降低 H. 264 帧间模式选择算法的计算复杂度; 使用模式相关和运动矢量合并对 UMHExagonS 算法进行优化, 减少大量的运动估计搜索点数, 取得了较高的编码效率。实验结果表明, 与使用 UMHExagonS 的帧间模式选择算法相比, 该算法在峰值信噪比及输出码流码率基本不变的前提下, 可以平均减少大约 50% 的运动估计时间。

关键词: H. 264/AVC; 模式选择; UMHExagonS; 模式相关; 运动同质性

Improved inter frame mode selection algorithm for H. 264

Ren Keqiang, Zhang Xuguang, Luo Huilan

Faculty of Information Engineering, Jiangxi University of Science and Technology, Ganzhou 341000, China

Abstract: Aiming at the problem of H. 264 inter frame mode selection algorithm, a fast inter frame mode selection algorithm is proposed in this paper. The algorithm selects the macroblock mode by using homogenous analysis of an 8×8 block motion vector, to narrow the scope of the candidate modes and decrease the computational complexity of the H. 264 inter frame mode selection algorithm. Meanwhile, the algorithm uses mode correlation and motion vector merging to optimize the UMHExagonS algorithm, which could reduce a large number of motion estimation search points, and obtain higher coding efficiency. The experimental results show that compared with inter frame mode selection algorithm using UMHExagonS, the proposed algorithm can reduce about 50% motion estimation time with only negligible change of PSNR and output stream bit rate.

Key words: H. 264/AVC; mode selection; UMHExagonS; mode correlation; motion homogeneity

0 引言

H. 264/AVC 视频压缩标准于 2003 年由联合视频专家组(JVT)提出^[1], 由于其优越的编码性能, H. 264/AVC 已被广泛的应用, 并成为学术界研究和讨论的热点。H. 264/AVC 的较高编码效率得益于

引入的许多先进技术, 其中, 多参考帧多种小块的帧间预测技术在很大程度上提高了视频的压缩率。然而, 对每个宏块的分割及预测方式遍历地进行搜索和比较, 将带来庞大的运算量。据统计, 遍历多模式的模式选择占据了 JM 参考软件整个编码时间的 90%^[2]。因此, 针对快速模式选择算法的研究就显得日益迫切和必要。

收稿日期: 2012-09-14; 修回日期: 2013-01-23

基金项目: 国家自然科学基金项目(61105042)

第一作者简介: 任克强(1959—), 男, 教授, 1993 年于华中理工大学获计算机组织与系统结构专业硕士学位, 主要研究方向为视频信号处理、信息隐藏。E-mail: renkeqiang@21cn.com

目前,有关快速模式选择算法的研究大多围绕“减少候选模式个数”展开,实现方法大致可以分为阈值法和预测法两类。阈值法在进行模式遍历的同时加入提前终止阈值,如果当前模式满足其对应的提前终止阈值,则中止模式选择,并令当前模式为最优模式。文献[2]提出了基于宏块运动强度的提前终止模式选择算法;文献[3]中使用宏块残差一致性和空间一致性条件作为提前终止方法;文献[4]综合使用了运动估计矢量、宏块纹理复杂度以及宏块率失真代价作为模式选择的提前终止策略。预测法在模式选择之前,首先对宏块中与模式选择相关的信息进行分析和处理,从而缩小候选模式范围。文献[5]使用当前块和参考块之间相位信息,设计了一种直接模式选择算法;文献[6]和文献[7]使用宏块的时空相关性对宏块模式进行预测,缩小候选模式范围。阈值法和预测法均可以较好地提升编码效率。然而,阈值法中的阈值通常是经验值,并不一定适合千变万化的视频序列,自适应的阈值则往往需要对已编码码流进行较复杂的分析,在一定程度上增加了编码负荷;预测法通常需要对宏块信息进行变换和处理,提高了编码器对硬件的要求。文献[8]和文献[9]提出了运动同质性的概念,并使用 4×4 块运动矢量同质性分析判断宏块所处区域的运动连续性;文献[8]认为,在小块运动矢量连续的区域,小块空间运动同质性较高,适合选择较大宏块模式编码;文献[9]对标准化了的运动矢量场进行3个方向的运动同质性分析来决定候选模式集。但是,大多数视频序列中小块模式所占比例通常很小^[10],对每个宏块都进行帧间 4×4 模式运动搜索无疑会浪费大量时间。此外,H.264标准中的UMHexagonS运动估计算法采用全局搜索策略,搜索运算量较大,调用该算法进行帧间模式选择将导致较高的编码复杂度。

针对上述问题,对H.264帧间模式选择算法进行改进,提出一种帧间模式选择快速算法,并对UMHexagonS算法进行优化,以实现较高的编码效率。

1 H.264 帧间模式选择算法

H.264采用“遍历所有模式并择优”的方法进行帧间模式选择。对一个宏块而言,需要对每种帧间模式分别进行运动估计。

1.1 帧间模式选择

为了提高块匹配精度,H.264为帧间亮度宏块提供了7种块分割方式和两级结构,1个宏块可分割为1个 16×16 块、2个 16×8 或 8×16 块、4个 8×8 块,其中每个 8×8 块(记作 $P8 \times 8$)可继续分割为1个 8×8 块、2个 8×4 或 4×8 块、4个 4×4 块。这7种模式依次编号为模式1至模式7,如图1所示。此外,为了提高压缩率,H.264还提供了SKIP/DIRECT直接预测模式以及Intra 16×16 和Intra 4×4 两种帧内模式。也就是说,H.264共为帧间宏块提供了10种候选模式,模式选择就是要在所有候选模式中选择一种最优模式作为编码模式。

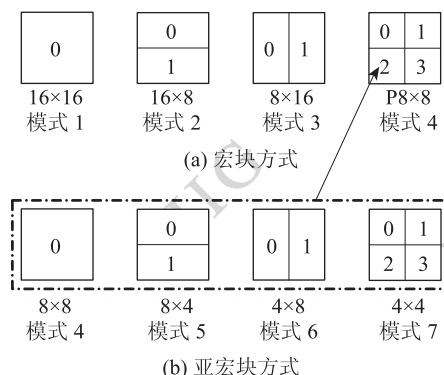


图1 宏块和亚宏块

Fig. 1 Macroblock types and subtypes

为了在候选模式中择优,H.264推荐使用拉格朗日乘数计算每种模式的代价,拉格朗日代价为

$$J = Dis + \lambda_{MODE} \cdot R \quad (1)$$

式中, Dis 表示重建块与原始块之间的量化差, λ_{MODE} 为当前模式的拉格朗日乘数, R 代表当前模式下编码所需比特数。

1.2 UMHexagonS 算法

UMHexagonS算法是一种混合模板运动估计搜索算法,由Chen等人提出并被JVT正式采纳^[11]。该算法与全搜索算法相比较可节省高达90%的运动估计时间^[12]。UMHexagonS算法可简单描述为以下4个步骤:

UMH_s1:起始搜索点预测(包括中值预测和上层预测);

UMH_s2:非对称十字形搜索;

UMH_s3:非均匀多六边形格点搜索;

UMH_s4:扩展六边形搜索。

此外,为了提高匹配精度,H.264允许在调用

UMHexagonS 算法进行整像素运动估计之后, 再进行 1/2 像素以及 1/4 像素精度搜索。

图 2 描述了 UMHexagonS 算法第 2 步至第 4 步的搜索过程, 其中 UMH_s3 和 UMH_s4 还可分别进一步细分为两个子步骤。

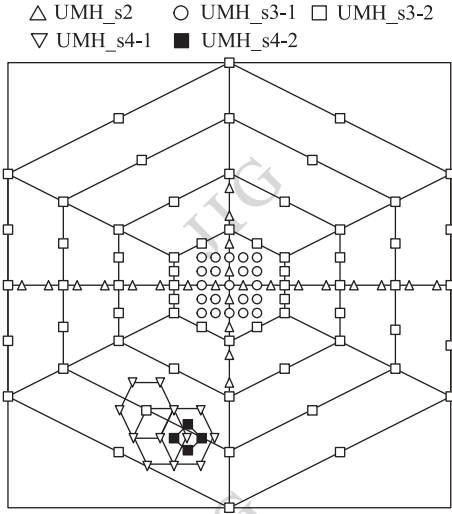


图 2 UMHexagonS 算法

Fig. 2 UMHexagonS algorithm

2 算法改进

2.1 模式选择的改进

H. 264 采用遍历的算法进行模式选择, 即计算每种候选模式的代价并选择代价最小的一种模式作为编码模式。遍历算法并不考虑宏块的自身特点, 仅追求最准确的模式选择及最高的压缩率。然而, 通过分析视频序列中宏块的自身特点往往可以对其编码模式做出准确预测, 减少模式遍历所需的大量编码时间。使用运动矢量的同质性分析便是一种有效的模式预测方法^[8,9], 并具有运算简单的特点。

运动同质性是描述两个或多个块运动相似程度的一种度量。使用小块的运动同质性分析可以作为上层模式选择的一种依据。在图 3 中, 如果宏块的 4 个 8×8 子块都具有较高的运动同质性(图 3(a)), 则宏块模式更倾向于选择模式 1; 如果子块 0 和子块 1、子块 2 和子块 3 分别具有较高的运动同质性(图 3(b)), 则宏块模式更倾向于选择模式 2; 如果子块 0 和子块 2、子块 1 和子块 3 分别具有较高的运动同质性(图 3(c)), 则宏块模式更倾向于选择模式 3。此外, 如果当前模式代价小于比其尺寸大的模式代价, 则可以跳过尺寸更大模式的选择而

不影响编码性能; 同理, 如果当前模式代价小于比其尺寸小的模式的代价, 则可跳过更小尺寸模式选择而不影响编码性能^[13]。图 3 中 $mv(k)$, ($k=0, 1, 2, 3$) 是编号为 k 的 8×8 子块运动矢量。

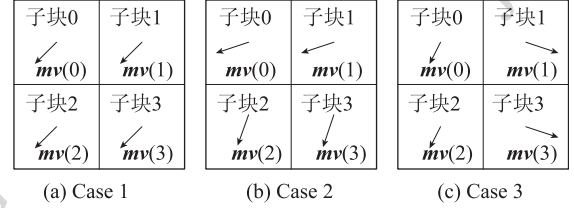


图 3 运动同质性

Fig. 3 Motion homogeneity

由于实际视频序列中亚宏块模式所占比例通常很小, H. 264 为帧间模式设计了两级结构(图 1), 只在模式 4 被选择时才开启亚宏块模式, 并且亚宏块模式通常处于运动复杂区域^[8], 其运动同质性较弱。因此, 对 H. 264 帧间模式选择算法进行改进, 基本思想是: 对于宏块模式, 使用 8×8 块运动矢量的同质性分析进行宏块模式选择; 对于亚宏块模式, 则保留其遍历的模式选择。具体改进方法如下:

1) 引入同质性判决因子 Ψ , 对 8×8 块运动矢量进行运动同质性分析从而选择宏块模式。

(1) 令 Ψ 为 3 种模式下运动矢量的绝对距离和, 即

$$\begin{cases} \Psi_1 = \sum_{i=0}^2 \sum_{j=i+1}^3 (|x_i - x_j| + |y_i - y_j|) \\ \Psi_2 = 3 \sum_{i=0,2} (|x_i - x_{i+1}| + |y_i - y_{i+1}|) \\ \Psi_3 = 3 \sum_{i=0,1} (|x_i - x_{i+2}| + |y_i - y_{i+2}|) \end{cases} \quad (2)$$

式中, Ψ_1 、 Ψ_2 和 Ψ_3 分别为模式 1、模式 2 和模式 3 的同质性判决因子; (x_i, y_i) 为编号为 i 的 8×8 子块的运动矢量坐标。

(2) 求 3 种模式下 Ψ 的最小值 $\Psi_{\min}$, 即

$$\Psi_{\min} = \min(\Psi_1, \Psi_2, \Psi_3) \quad (3)$$

设定阈值 TH , 如果 $\Psi_{\min} > TH$, 说明宏块处于复杂运动区域, 使用遍历算法进行模式选择; 如果 $\Psi_{\min} \leq TH$, 则依据同质性判决因子 Ψ 进行模式选择, 且当 $\Psi_i = \Psi_{\min}$ 时, 宏块模式更倾向于选择模式 i 。经多次实验验证, TH 取值为 12 时, 可以在基本不影响输出码流码率及峰值信噪比的前提下最大限度地节约编码时间。

2) 利用运动同质性分析结果和已搜索模式的

代价进一步缩小候选模式范围。

令 mc_i 为模式 i 的代价, 且 $\Psi_{\min} \leq TH$, 则:

(1) 如果 $mc_1 < mc_4$ 且 $\Psi_{\min} = \Psi_1$, 直接选取模式 1 为最佳模式。

(2) 如果 $mc_1 \geq mc_4$ 且 $\Psi_{\min} = \Psi_1$, 则进行亚宏块模式选择, 不进行模式 2 及模式 3 的搜索。

(3) 如果 $\Psi_{\min} = \Psi_2$, 则进行模式 2 搜索, 并选择 mc_1 、 mc_2 和 mc_4 中最小的一种模式为最佳模式; 如果模式 4 被选择, 则还需进行亚宏块模式搜索, 否则不进行。

(4) 如果 $\Psi_{\min} = \Psi_3$, 则进行模式 3 搜索, 并选择 mc_1 、 mc_3 和 mc_4 中最小的一种模式为最佳模式; 如果模式 4 被选择, 则还需进行亚宏块模式搜索, 否则不进行。

2.2 UMHExagonS 算法的优化

使用 UMHExagonS 运动估计算法的帧间模式选择, 需要对 7 种帧间模式 (图 1) 分别调用 UMHExagonS 模版进行运动估计, 计算相应模式的代价并选取代价最小的一种模式进行编码。由于 UMHExagonS 算法采用了多六边形格点搜索模板做全局搜索, 搜索点数相对较多, 使用该算法做帧间模式的选择会导致较高的编码复杂度。

H. 264 参考软件 JM 在进行帧间模式选择时, 每种模式都会调用完全的 UMHExagonS 搜索模板相对独立地进行运动估计, 小块模式除了参考上层模式的运动矢量作为其上层预测矢量外, 与上层模式的运动估计过程并无太多联系。然而, 以模式 1 为例, 在其进行最优匹配块的搜索时, 实际上也同时对其中 4 个 8×8 块做了匹配, 如图 4 所示。

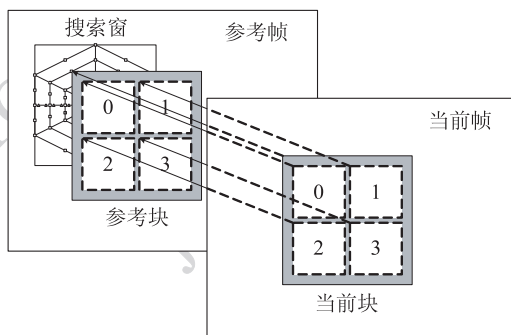


图4 模式间的相关性

Fig. 4 Modes correlativity

在同一个参考帧内, 由于模式 1 与模式 4 在起始搜索点的预测时使用的参考块不同, 二者的搜索中心并不相同, 搜索窗也不能完全重合, 但二者搜索

中心的距离往往相差不大, 搜索窗也有相当大的一部分相互重叠, 如图 5 所示。这一点在中、小运动视频序列中体现尤为显著。

因此, 对 UMHExagonS 算法进行了优化, 将其划分为“粗定位”(算法步骤 UMH_s1 至 UMH_s3) 和“细搜索”(算法步骤 UMH_s4 及后续分精度搜索) 两个等级: “粗定位”在搜索框内粗略地进行搜索以避免陷入局部最优, “细搜索”则在“粗定位”的基础上精确地找到全局最优的位置。

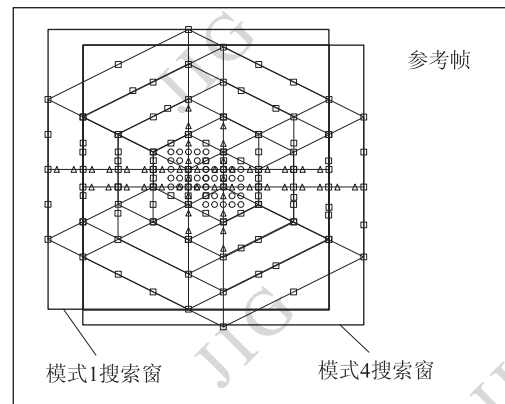


图5 模式间搜索窗的重叠

Fig. 5 Overlapped search windows between different modes

为了节约编码时间, 在进行模式 1 整像素运动估计的同时, 记录模式 4 中 4 个 8×8 块的最优运动矢量, 完成模式 4 的运动矢量“粗定位”, 相当于对模式 4 的“粗定位”使用了模式 1 的搜索窗。这种做法是基于 UMHExagonS 算法模板匹配点在搜索窗内分布相对均匀的特点, 可以节约 UMHExagonS 算法对模式 4 重复调用“粗定位”模板的搜索时间, 只要这种方法和 UMHExagonS 算法得到的最优匹配点距离相差不大, 二者在“粗定位”的效果上便是相同的, 并且, 这种方法与 UMHExagonS 算法在解决局部最优问题的思想上并不相悖。然后, 再对 4 个 8×8 块相互独立地进行“细搜索”, 便可以准确地找到最优匹配点所在的位置。

此外, 小块模式运动矢量对选定模式运动矢量具有一定的预测作用, 采用式(4)和式(5)对运动矢量进行合并, 作为选定模式的预测矢量, 同样可以达到“粗定位”的效果, 进一步减少选定模式调用 UMHExagonS 算法“粗定位”模板的搜索时间。

$$\begin{cases} pmv_{2,0} = (mv(0) + mv(1))/2 \\ pmv_{2,1} = (mv(2) + mv(3))/2 \end{cases} \quad (4)$$

$$\begin{cases} pmv_{3,0} = (mv(0) + mv(2))/2 \\ pmv_{3,1} = (mv(1) + mv(3))/2 \end{cases} \quad (5)$$

式中, $pmv_{i,j}$ 为在模式 i 下编号为 j 的宏块运动矢量预测值。

2.3 改进算法描述

采用运动矢量同质性分析与模式代价相结合的方法进行模式选择,并在模式间对 UMHexagonS 算法进行了优化,算法步骤如下:

1) 模式 1 运动估计,计算 mc_1 ,并对模式 4 运动矢量“粗定位”。

2) 模式 4 运动矢量“细搜索”,计算 mc_4 。

3) 计算 Ψ_1 、 Ψ_2 、 Ψ_3 及 $\Psi_{\min}$,若 $\Psi_{\min} > TH$,使用传统的遍历算法进行模式选择,并转步骤 10);否则转步骤 4)。

4) 若 $\Psi_{\min} = \Psi_1$ 且 $mc_1 < mc_4$,则转步骤 10)。

5) 若 $\Psi_{\min} = \Psi_1$ 且 $mc_1 \geq mc_4$,则转步骤 9)。

6) 若 $\Psi_{\min} = \Psi_2$,使用式(4)对模式 2 进行运动矢量预测,并以预测矢量为中心进行模式 2“细搜索”,计算 mc_2 ,并转步骤 8);否则转步骤 7)。

7) 若 $\Psi_{\min} = \Psi_3$,使用式(5)对模式 3 进行运动矢量预测,并以预测矢量为中心进行模式 3“细搜索”,计算 mc_3 。

8) 比较已搜索模式代价,若 mc_4 为最小,则转步骤 9);否则转步骤 10)。

9) 亚宏块模式全搜索。

10) 在已搜索模式中选出一代代价最小的模式作为编码模式,结束模式选择。

3 实验结果及分析

实验环境: Intel (R) Xeon (R) CPU E5504 @ 2.0 GHz; 2 GB 内存; Windows Server 2003 + JM18.0。JM 配置参数见表 1。

表 1 JM 软件配置参数

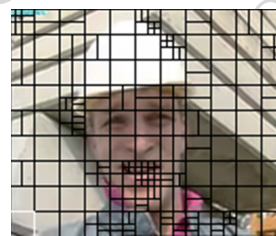
Table 1 JM software configuration parameters

配置参数	参数值
编码方式	帧编码
参考帧数目	1
序列类型	IPPP
编码帧数	300
运动搜索方式	UMHexagonS
RDO 优化	ON
其他配置	默认配置

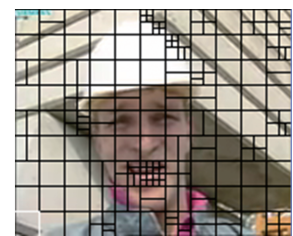
3.1 模式选择准确性

模式选择准确性比较实验以 QCIF (176 × 144) 测试序列“foreman”的第 42 帧和第 171 帧为例,将本文改进模式选择算法与遍历模式选择算法进行了比较,实验结果如图 6 所示,图中黑色实线表示最佳宏块分割方式。

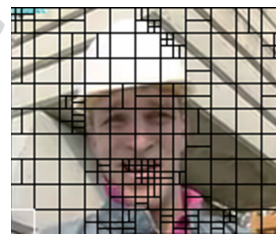
由图 6 可以看出,本文算法可以较为准确地对宏块进行划分,并且在码率影响不大的前提下,本文算法更倾向于选择大块模式来减少编码时间。



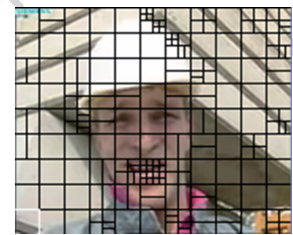
(a) 遍历算法第42帧



(b) 本文算法第42帧



(c) 遍历算法第171帧



(d) 本文算法第171帧

图 6 模式选择准确性比较

Fig. 6 Comparison of mode selection accuracy

3.2 UMHexagonS 算法优化正确性

为了验证对 UMHexagonS 算法优化的正确性,对 QCIF 分辨率测试序列“foreman”第 2 帧中所有 99 个宏块分别使用优化的 UMHexagonS 算法(IUMH)和原 UMHexagonS 算法(UMH)以模式 4 进行运动估计并记录最优运动矢量,将得到的 99 × 4 共计 396 组运动矢量按 (Macroblock 0 (block 0, block 1, block 2, block 3), Macroblock 1 (block 0, block 1, block 2, block 3), ..., Macroblock 98 (block 0, block 1, block 2, block 3)) 的顺序排列,以运动矢量的模来反映运动强度,以两种算法对应编号运动矢量的欧氏距离来衡量两种算法运动估计的差异,即

$$MS(n) = |\mathbf{mv}(n)| \quad (6)$$

$$MV_{\text{dist}}(n) = |\mathbf{mv}_{\text{IUMH}}(n) - \mathbf{mv}_{\text{UMH}}(n)| \quad (7)$$

式中, MS 为运动强度, MV_{dist} 为欧氏距离, $\mathbf{mv}(n)$ 为编号 n 的 8 × 8 块运动矢量,下标 UMH 表示使用 UMHexagonS 算法,下标 IUMH 表示使用本文优化的

UMHexagonS 算法。

实验结果如图 7 所示。图 7 y 轴以 1/4 像素为单位。从图 7 可以看出,本文对 UMHexagonS 算法的优化与原 UMHexagonS 算法在搜索效果上相差不大,尤其在中、小程度运动矢量的搜索上,本文对

UMHexagonS 的优化算法几乎可以得到与 UMHexagonS 算法相同的运动矢量。但是,当宏块运动程度相当剧烈时,优化算法的搜索结果相对 UMHexagonS 算法有一定偏差,其差异大致维持在 3 个整像素范围之内。

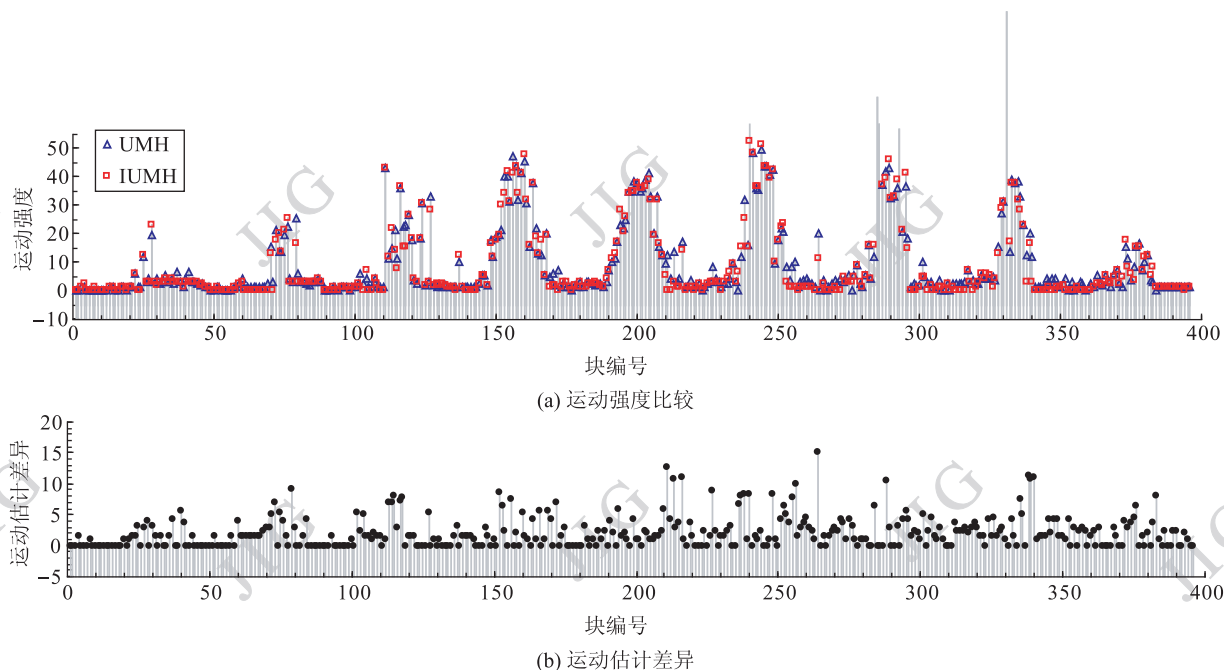


图 7 UMHexagonS 算法优化结果比较

Fig. 7 Comparison of UMHexagonS algorithm optimizing results

3.3 算法性能测试

为了测试本文在两个方面改进的帧间模式选择算法(本文算法)的性能,选取 foreman、coastguard、container、hall、mobile 和 news 共 6 组测试序列在 QCIF 和 CIF(352×288)分辨率下分别与 JM 中使用 UMHexagonS 运动估计的模式选择算法(JM 算法)进行了比较实验,采用视频序列亮度分量峰值信噪比 $PSNR(Y)$ 、输出码流码率 Br 以及运动估计时间 MET 3 个指标比较本文算法和 JM 算法的性能。

$$\Delta PSNR(Y) = PSNR(Y)_{opt} - PSNR(Y)_{ref} \quad (8)$$

$$\Delta Br = \frac{Br_{opt} - Br_{ref}}{Br_{ref}} \times 100\% \quad (9)$$

$$\Delta MET = \frac{MET_{opt} - MET_{ref}}{MET_{ref}} \times 100\% \quad (10)$$

式中, $\Delta PSNR(Y)$ 为 Y 分量峰值信噪比的变化量, ΔBr 和 ΔMET 分别表示输出码流码率和运动估计时间的变化百分比。opt 下标表示使用本文算法的测试值, ref 下标表示使用 JM 算法测试值。

表 2 为 QCIF 序列的实验结果,表 3 为 CIF 序列的实验结果。表 2 和表 3 的实验结果表明,对于不同运动复杂度和纹理复杂度的视频序列,本文算法的峰值信噪比及输出码流码率与 JM 算法相差不大,亮度分量峰值信噪比平均下降 0.022 dB,码率平均增加 1.13%。但编码速率比 JM 算法均有较大提升,运动估计时间平均减少 52.1% (QCIF) 和 49.04% (CIF);其中,对具有大量静止背景以及小运动的视频序列 container 和 hall 效果显著,运动估计时间分别减少 62.08% (QCIF)、58.34% (CIF) 和 59.41% (QCIF)、56.52% (CIF);对具有较剧烈全局运动的视频序列 foreman 效果则相对差一些,运动估计时间减少 33.18% (QCIF)、37.96% (CIF)。导致这种结果的一个主要原因是:对大运动宏块而言,模式间搜索窗的重叠区域较小,容易导致模式 4 中 4 个 8×8 块运动估计“粗定位”不够准确,使得其“细搜索”过程收敛速度较慢。

图 8 和图 9 以 CIF 分辨率 mobile 序列的前 50

表 2 QCIF 分辨率序列实验结果

Table 2 Experimental results of QCIF resolution sequences

序列	PSNR(Y)/dB		$\Delta PSNR(Y)$ /dB	Br/(kbit/s)		$\Delta Br/\%$	MET/s		$\Delta MET/\%$
	JM 算法	本文算法		JM 算法	本文算法		JM 算法	本文算法	
foreman	36.027	35.999	-0.028	124.87	126.70	+1.47	3390.48	2265.53	-33.18
coastguard	34.686	34.668	-0.018	202.68	208.78	+3.01	3846.47	1860.29	-51.64
container	36.257	36.234	-0.023	38.52	38.67	+0.39	2246.79	851.88	-62.08
hall	37.543	37.502	-0.041	51.18	51.93	+1.47	2083.79	845.75	-59.41
mobile	33.665	33.632	-0.033	439.47	444.79	+1.21	3990.34	1878.43	-52.93
news	36.952	36.940	-0.012	73.30	74.26	+1.31	2427.52	1131.59	-53.38

表 3 CIF 分辨率序列实验结果

Table 3 Experimental results of CIF resolution sequences

序列	PSNR(Y)/dB		$\Delta PSNR(Y)$ /dB	Br/(kbit/s)		$\Delta Br/\%$	MET/s		$\Delta MET/\%$
	JM 算法	本文算法		JM 算法	本文算法		JM 算法	本文算法	
foreman	36.462	36.458	-0.004	445.89	450.17	+0.96	14612.72	9065.70	-37.96
coastguard	35.314	35.302	-0.012	1024.70	1031.71	+0.68	16456.64	9633.58	-41.46
container	36.252	36.232	-0.020	184.40	185.73	+0.72	9165.22	3818.60	-58.34
hall	37.912	37.869	-0.043	277.88	279.08	+0.43	8805.80	3828.66	-56.52
mobile	34.499	34.482	-0.017	1794.03	1817.19	+1.29	15048.89	7964.85	-47.07
news	38.326	38.309	-0.017	203.70	205.07	+0.67	8704.87	4101.08	-52.89

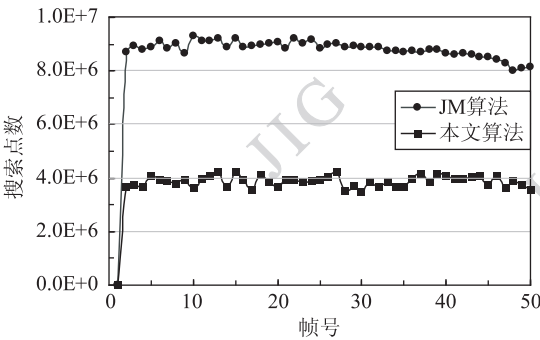


图 8 搜索点数的比较

Fig. 8 Comparison of search points

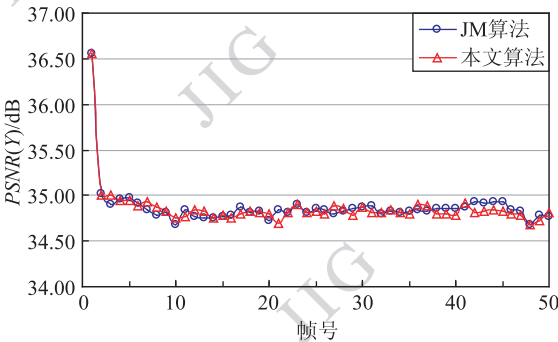


图 9 Y 分量峰值信噪比的比较

Fig. 9 Comparison of PSNR(Y)

帧为例,对本文算法和 JM 算法在搜索点数和亮度信号峰值信噪比两个方面分别进行了逐帧编码效果的比较实验。图 8 和图 9 的实验结果表明,本文算法在基本不影响每一帧峰值信噪比的基础上,与 JM 算法相比,可减少一半以上的运动估计搜索点数。

4 结 论

提出了一种基于模式间相关和运动同质性分析的模式选择快速算法。充分利用 UMHexagonS 算法非均匀多六边形格点搜索模板在搜索窗内分布均匀的特点,取得了较高的编码效率。一方面,利用模式间搜索窗的重叠,使用运动矢量并行定位的方法,在搜索效果基本一致的前提下,减少了传统算法对模式 4 重复使用 UMHexagonS“粗定位”搜索模板的运算量;另一方面,采用运动同质性分析和模式代价的比较,在缩小候选模式范围的基础上,使用运动矢量合并的方法,对选定模式再次避免了使用繁杂的 UMHexagonS“粗定位”搜索模板。如何进一步提高本文算法在大运动视频序列下的编码效率,将是后续研究的重点。

参考文献 (References)

- [1] Wiegand T, Sullivan G, Luthra A. Draft ISO/IEC 14496-10: 2002 (E). Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification [S]. Geneva: JVT, 2003.
- [2] Zeng H Q, Cai C H, Ma K K. Fast mode decision for H. 264/AVC based on macroblock motion activity [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2009, 19(4): 1-11.
- [3] Gu M H, Yu N M, Kou L K, et al. Fast mode decision algorithm for H. 264 based on early termination method [J]. Journal of Image and Graphics, 2011, 16(3): 305-309. [顾梅花, 余宁梅, 寇立康, 等. H. 264 快速模式选择算法中的提前终止策略 [J]. 中国图象图形学报, 2011, 16(3): 305-309.]
- [4] Zhang G B, Li L. A fast coding algorithm for H. 264/AVC based on best intermode decision and early-terminated strategies [J]. Journal of Image and Graphics, 2009, 14(1): 59-64. [章国宝, 李亮. 帧间编码模式选择及其择优早期终止的 H. 264/AVC 快速算法 [J]. 中国图象图形学报, 2009, 14(1): 59-64.]
- [5] Paul M, Lin W S, Lau C T, et al. Direct intermode selection for H. 264 video coding using phase correlation [J]. IEEE Transactions on Image Processing, 2011, 20(2): 461-473.
- [6] He Y B, Bi D Y, Ma S P, et al. An inter-mode decision algorithm based on spatiotemporal correlation in H. 264 [J]. Journal of Image and Graphics, 2009, 14(12): 2456-2462. [何宜宝, 毕笃彦, 马时平, 等. H. 264 中基于时空相关性的帧间模式选择算法 [J]. 中国图象图形学报, 2009, 14(12): 2456-2462.]
- [7] Han D, Kulkarni A, Rao K R. Fast inter-prediction mode decision algorithm for H. 264 Video encoder [C]//Proceedings of IEEE International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology. Shanghai, China: IEEE, 2012: 1-4.
- [8] Shen L Q, Liu Z, Zhang Z Y, et al. Fast inter mode decision using spatial property of motion field [J]. IEEE Transactions on Multimedia, 2008, 10(6): 1208-1214.
- [9] Liu Z, Shen L Q, Zhang Z Y. An efficient intermode decision algorithm based on motion homogeneity for H. 264/AVC [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2009, 19(1): 128-132.
- [10] Lee B, Kim M. An efficient inter-prediction mode decision method for temporal scalability coding with hierarchical b-picture structure [J]. IEEE Transactions on Broadcasting, 2012, 58(2): 285-290.
- [11] Chen Z B, Zhou P, He Y. Fast integer pel and fractional pel motion estimation for JVT [J]. JVT-F017, 2002, 12: 5-13.
- [12] Xu X Z, He Y. Improvements on fast motion estimation strategy for H. 264/AVC [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18(3): 285-293.
- [13] Zhu J X, Cao N, Chen Y S, et al. A fast macroblock inter mode decision algorithm for H. 264 [C]//Proceedings of IEEE WRI World Congress on Software Engineering. Xiamen, China: IEEE, 2009, 2: 189-193.