

二值 α 平面辅助的视频对象快速运动估计算法^{*}

宋传鸣^{1,2}, 王相海^{2,3+}, 张福炎^{1,3}

¹(南京大学 计算机科学与技术系,江苏 南京 210093)

²(辽宁师范大学 计算机与信息技术学院,辽宁 大连 116029)

³(南京大学 计算机软件新技术国家重点实验室,江苏 南京 210093)

Binary Alpha-Plane Assisted Fast Motion Estimation Scheme of Video Object

SONG Chuan-Ming^{1,2}, WANG Xiang-Hai^{2,3+}, ZHANG Fu-Yan^{1,3}

¹(Department of Computer Science and Technology, Nanjing University, Nanjing 210093, China)

²(College of Computer and Information Technology, Liaoning Normal University, Dalian 116029, China)

³(State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, China)

+ Corresponding author: Phn: +86-411-82154768, Fax: +86-411-84202626, E-mail: xhwang@graphics.nju.edu.cn

Song CM, Wang XH, Zhang FY. Binary alpha-plane assisted fast motion estimation scheme of video object. *Journal of Software*, 2008,19(4):829–841. <http://www.jos.org.cn/1000-9825/19/829.htm>

Abstract: This paper proposes a fast motion estimation scheme of arbitrarily shaped video object. The instructive role of the alpha-plane taking part in the motion estimation of video object is discussed, and a weighted binary alpha-plane matching criterion (WBAMC) is proposed by using boundary extension and boundary mask techniques. Based on priority search strategy, this paper proposes a fast binary alpha-plane assisted motion estimation (BAAME) scheme of video object. First, the BAAME uses alpha-plane and the WBAMC criterion to limit the search of boundary macro-blocks (MBs) into small unimodal area of two starting points so that a conventional fast motion estimation algorithm can be employed to search the motion vectors (MVs) of boundary MBs. Second, the BAAME predicts the starting points of opaque MBs by using MVs of neighboring boundary MBs and then employs a fast motion estimation algorithm to compute the MVs of opaque MBs. The proposed scheme can be combined with many spatial domain based and frequency domain based different motion estimation algorithms, and be effectively applied to object-based video codecs. The experimental results show that the BAAME scheme can always reach high motion estimation accuracy and better subjective quality for standard test video sequences which have different characteristics respectively. The proposed scheme can achieve 0.1~0.8dB higher prediction quality on average than DS (diamond search) and PSA (priority search algorithm) (BAAS (binary alpha-plane assisted search)+DS), and a

* Supported by the National Natural Science Foundation of China under Grant Nos.60372071, 60703084, 60723003 (国家自然科学基金); the Natural Science Foundation of Liaoning Province of China under Grant No.20032105 (辽宁省自然科学基金); the Natural Science Foundation of Jiangsu Province of China under Grant No.BK2007571 (江苏省自然科学基金); the Project Innovation of Graduate Students of Jiangsu Province of China under Grant No.CX07B-121Z (江苏省普通高校研究生科研创新计划); the Plan for Supporting the Colleges' Outstanding Talented Person of Liaoning Province of China under Grant No.RC-04-11 (辽宁省高等学校优秀人才支持计划)

Received 2006-11-08; Accepted 2007-01-04

little lower PSNR (peak signal-to-noise ratio) than FS (full search). Moreover, the BAAME scheme can speed up the motion estimation about 20 times in terms of computational complexity when compared with FS.

Key words: video coding; motion estimation; binary alpha-plane; video object; block matching criterion

摘要: 提出了一种任意形状视频对象的快速运动估计方法.详细分析了 alpha 平面在视频对象的快速运动估计过程中起到的指导性作用,采用边界扩展和边界掩码技术,提出了一种新的二值 alpha 平面匹配衡量准则 WBAMC (weighted binary alpha-plane matching criterion).结合优先搜索策略,提出了二值 alpha 平面辅助的视频对象快速运动估计算法 BAAME(binary alpha-plane assisted motion estimation).首先,利用 alpha 平面和 WBAMC 准则,将边界宏块的搜索范围缩小至两个搜索起点的单调区域,再采用传统的快速运动估计算法确定其运动向量;然后,用边界宏块的运动向量预测内部宏块的搜索起点;最后,采用快速运动估计算法搜索内部宏块的运动向量.这种方法可与多种空间和频率域运动估计算法相结合,有效地应用于基于对象的视频编码器中.实验结果表明,对于多种类型的标准测试视频流,BAAME 算法始终能够保持较高的估计精度和主观质量,运动补偿的平均 PSNR(peak signal-to-noise ratio)较 DS(diamond search)和 PSA(priority search algorithm)(BAAS(binary alpha-plane assisted search)+DS)高出 0.1dB~0.8dB,略低于 FS(full search),但是其计算复杂度与 FS 相比降低了 20 倍.

关键词: 视频编码;运动估计;二值 alpha 平面;视频对象;块匹配准则

中图法分类号: TP391 **文献标识码:** A

MPEG-4 中视频对象(video object,简称VO)^[1]概念及其编码机制的提出,有效地提高了视频编码器的交互性和适应能力.每个视频对象由一系列的视频对象平面(video object plane,简称VOP)组成;每个视频对象平面由对象包围盒(bounding box)来定义,并分为纹理信息和与其对应的alpha平面(alpha-plane)信息两部分.根据alpha平面,对象包围盒内部的宏块可被划分为外部宏块(transparent macroblock)、内部宏块(opaque macroblock)和边界宏块(boundary macroblock).为了消除视频对象存在的时间冗余,MPEG-4 采用了扩展的块匹配运动估计技术对上述 3 类宏块进行处理.首先,需要将参考VOP中处于包围盒内部的外部像素(transparent pixel)进行重复填充^[2],然后,利用多边形匹配(polygon matching)运动估计方法进行预测,并用SAD(sum of absolute differences)准则衡量预测误差,其方法如下:

$$SAD(u, v) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |I_c(i, j) - I_r(i + u, j + v)| \times \alpha(i, j) \quad (1)$$

其中,宏块大小为 $N \times N$, $I_c(i, j)$ 和 $I_r(i, j)$ 分别表示当前帧和参考帧中坐标为 (i, j) 的像素值, (u, v) 为候选运动向量, $\alpha(i, j)$ 为当前帧所对应的alpha平面中 (i, j) 点的值.选择SAD值最小的候选向量为当前宏块的运动向量,即

$$MV = \arg_{(u, v)} \min SAD(u, v), -D \leq u, v \leq D \quad (2)$$

其中, D 为运动向量的最大搜索范围.

基于块匹配的运动估计技术在准确度和复杂度之间达到了较好的折衷,因此,它在 MPEG-1/2/4, H.263, H.264/AVC 等多个视频编码标准中获得了成功应用.然而,基于块匹配的运动估计技术尚存在一些不足:

首先,尽管 3 步搜索法(three step search,简称TSS)、4 步搜索法(four step search,简称FSS)^[3]、基于块的梯度下降搜索法(block-based gradient descent search,简称BBGDS)^[4]、菱形搜索法(DS)^[5]、六边形搜索法(hexagon-based search,简称HEXBS)和十字-菱形-六边形搜索法(cross-diamond-hexagonal search,简称CDHS)^[6]等多种快速运动估计算法能够减小搜索点集,提高搜索速度,但是,这些算法均在一定程度上牺牲了运动补偿的主、客观质量.

其次,以宏块为单位进行运动估计,如果同一个宏块中包含两个或两个以上具有不同运动形式的物体,或者在一个边界宏块中,如果对对象的前景被较强的背景所影响,那么经过运动补偿后,就会不可避免地产生不连续的物体边界,以致影响运动补偿的视觉效果^[7].

再次,虽然 H.264/AVC 标准提出了先进的运动估计技术,提高了编码效率,但是在运动估计过程中,它的匹

配误差准则仍是基于灰度信息的,也没有将物体边界的连续性作为考察的主要因素之一。

针对上述边界估计不准确的问题,文献[7-9]提出了用二值边缘信息进行运动估计的快速算法,尽管其运算复杂度有所提高,却获得了较之传统快速运动估计算法更好的主、客观质量。文献[10]在上述思想的基础上,提出了一种应用于MPEG-4的任意形状视频对象的快速运动估计算法。它采用优先搜索(priority search,简称PS)策略,利用视频对象的alpha平面(下文中的“alpha平面”均指“二值alpha平面”)指导边界宏块的运动估计,并用边界宏块的运动向量预测内部宏块的向量,不仅加快了运动估计的速度,而且获得了平滑的运动场和较为连续的VO边界,提高了运动补偿的视觉质量。然而,在该算法中,alpha平面的匹配误差准则只能粗略地定位运动估计的起点(starting point,简称SP),并且所采用的均匀分布的起点模式(uniform starting point pattern)与运动向量的中心偏置特性^[11]不一致,这将增加该算法的计算量,甚至引入不准确的SP,以致影响运动补偿的质量。

本文首先论述 alpha 平面信息在 MPEG-4 视频对象的运动估计中将起到的指导作用,然后提出一种用于 alpha 平面信息辅助运动估计的匹配衡量准则 WBAMC(weighted binary alpha-plane matching criterion)。在此基础上,结合优先搜索策略,提出一种新的二值 alpha 平面辅助的视频对象快速运动估计算法(binary alpha-plane assisted motion estimation,简称 BAAME)。最后给出所提出算法的多项实验结果,并与相应的算法进行比较。

1 alpha 平面对运动估计的指导作用

多年来,众多专家以及VQEG(Video Quality Experts Group)组织^[12]根据人类视觉系统特性提出了多种评价图像和视频序列主、客观质量的模型。文献[13,14]将边界误差(edge errors)作为评价模型的重要参数之一。可见,物体边界的编码质量将影响最终的视觉质量,将其作为运动估计过程中需考虑的因素之一是完全必要的。

1.1 基于亮度信息的块匹配技术的不足

基于块匹配的运动估计技术的有效性建立在两个假设基础上,即:

- 1) 视频序列中的运动物体只存在平移运动,并且同一宏块内的所有像素具有一致的运动;
- 2) 在目标宏块的搜索窗口内,匹配误差曲面(error surface)是单峰分布的。

于是,目标宏块及其匹配宏块的匹配误差(如 SAD)应最小;利用快速运动估计算法总能找到最优的运动向量。然而事实上,上面的假设只是对自然序列运动特性的近似,这表现为以下两个方面:

- (1) 匹配误差最小的两个宏块,在视觉方面,不一定是最佳匹配。

图 1 为 Stefan 序列第 2 帧的部分画面,其中,图 1(b)为利用整数像素精度全搜索运动补偿得到的帧,将其与原始帧图 1(a)比较能够明显看出,球拍左侧的边框发生大幅断裂,图 1(c)标识了断裂所在的宏块。在解码端,由于预测误差被量化,重构帧很容易在此处出现不连续,而这恰恰是人眼最敏感的边界,以致发生可视的降质。

- (2) 误差曲面往往是多峰的。图 2 是图 1(c)中用白色方框标志的宏块在其搜索窗口内所得到的误差曲面,搜索精度为整数像素,匹配准则是 SAD。显然,误差分布曲面存在多个局部最优点,这样,传统的快速运动估计算法则易于陷入局部最优,进而获得不准确的运动向量。

综合(1)、(2)所述,基于亮度信息的块匹配技术主要有两个方面的不足:

- 其一,块匹配准则不能保证获得良好的主观效果;
- 其二,基于该准则的快速运动估计算法也不能保证搜索到全局最优的运动向量。

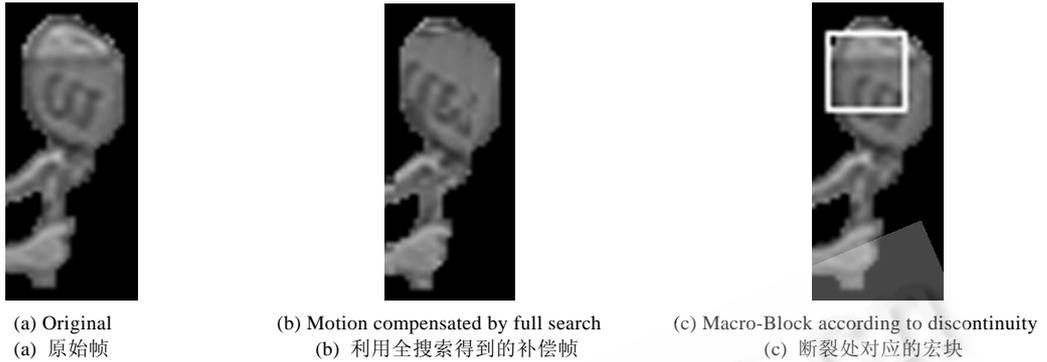


Fig.1 The macro-block with minimum displaced difference will not certainly achieve the best subjective motion compensated quality

图 1 匹配误差最小的宏块不一定能获得最优的主观补偿质量

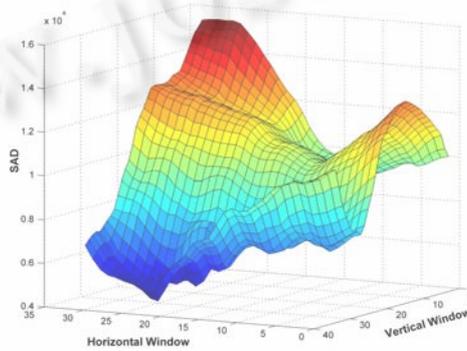


Fig.2 The error surface of the marco-block depicted in Fig.1(c)

图 2 图 1(c)中所示宏块的匹配误差曲面

1.2 alpha平面信息的指导作用

MPEG-4 用纹理信息和 alpha 平面共同描述一个 VOP,其中,纹理信息刻画 VOP 的亮度变化,而 alpha 平面则刻画了 VOP 的边界.在运动估计过程中,目标宏块与其匹配宏块的亮度匹配误差不仅应很小,而且其 alpha 平面的匹配误差也应很小.如果在对图 1(a)的搜索过程中兼顾其 alpha 平面,那么,搜索则不会收敛于一个边界完全不符的宏块,导致边界缺损.当然,简单地引入 alpha 平面也是不科学的,必须考察:

- 1) alpha 平面是否适合进行运动估计;
- 2) 匹配误差准则的选取;
- 3) alpha 平面所带来的性能提高;
- 4) 所需的额外计算量.

首先,文献[10]利用异或(exclusive-OR)运算定义了匹配误差准则:

$$BAMS(u, v) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} BA_c(i, j) \otimes BA_r(i+u, j+v) \quad (3)$$

其中, $BA_c(\cdot, \cdot)$ 和 $BA_r(\cdot, \cdot)$ 分别表示当前宏块和参考宏块的alpha平面, \otimes 表示异或运算.依据该准则,可以得到图 1(c)所示宏块在搜索窗口内的误差曲面(如图 3 所示).

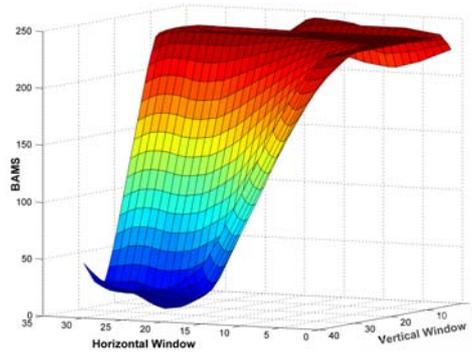


Fig.3 The error surface of the alpha-plane of the macro-block depicted in Fig.1(c) based on BAMS criterion

图 3 图 1(c)所示宏块的 alpha 平面基于 BAMS 准则的匹配误差曲面

对比图 2 与图 3 可以发现, alpha 平面的匹配误差曲面与亮度信息的匹配误差曲面,其总体变化趋势相近,并且前者的单调性更明显(图 3 只出现了两个波谷).文献[10]通过对其他测试序列的研究,也获得了同样的结论.这说明,在 alpha 平面匹配误差的局部最小点周围,往往分布着亮度信息匹配误差的局部最小点.若匹配误差准则选择合适,则在 alpha 平面匹配误差的全局最小点或者少数几个次最小点周围,分布着亮度信息匹配误差的全局最小点的可能性也很大(第 3.1 节的实验验证了此推论).因此,利用 alpha 平面匹配误差的分布来定位运动向量,可以将基于亮度信息匹配准则的运动估计限制在一个较小的搜索点集,有利于提高运动估计速度.

其次,将 alpha 平面用于边界宏块的运动估计和补偿,有利于得到连续的对象边界.而边界宏块的运动向量还可以作为内部宏块运动向量的搜索起点,这不仅有利于得到更加平滑的运动场,有利于获得更高的主观质量,也可以进一步加快运动估计的速度.另外,将 alpha 平面匹配误差的较小点作为快速运动估计的起点,还将有利于搜索得到全局最优的运动向量,进而获得更高的客观质量和编码效率.

再次,由于 alpha 平面是二值的,其运算均为位运算,因此, alpha 平面匹配的运算量至多仅相当于亮度信息运算量的 $1/16^{[10]}$.这样,将 alpha 平面引入运动估计,并不会明显增加运动估计所需的计算量.

最后, alpha 平面信息很简单,它只是反映了对象的边界特征,即使某两个宏块的 alpha 平面匹配误差最小,也并不能保证这两个宏块是最优匹配(第 2.1 节的实验将证明这一点).尤其当宏块内的边界像素较少时,这种匹配则有可能是“孔径效应(aperture problem)^[15]”的结果.而亮度信息反映了对象更丰富的特征,如亮度变化、纹理强度等,它是某个对象区别于其他对象的关键依据.当某两个宏块的众多特征均很好匹配时,才能说明这种匹配是有效的.因此在运动估计过程中, alpha 平面信息只能起到指导作用,不能起决定作用.

综上所述,运动估计应兼顾 VOP 的亮度信息和 alpha 平面的匹配误差,以 alpha 平面的匹配为指导,确定亮度信息运动估计的起始点,在保证较低运算量的前提下,最大限度地搜索到全局最优的运动向量.

2 二值 alpha 平面辅助快速运动估计算法

2.1 PSA(BAAS+DS)快速运动估计算法的不足

Hui 等人研究了 MPEG-4 的外部宏块、内部宏块和边界宏块的不同特性,提出了一种基于优先搜索策略的快速运动估计算法 PSA(priority search algorithm)(BAAS(binary alpha-plane assisted search)+DS(diamond search))^[10].算法的基本思想是:内部宏块的运动与边界宏块的运动是高度相关的:首先,对每个边界宏块,采用均匀分布的搜索起点模板(starting point pattern,简称 SPP)(如图 4(a)所示)和 alpha 平面信息,计算出该宏块的 BAMS(binary alpha-plane matching score)值小于某阈值时对应的 SP(updated SPP,简称 USPP),再在 USPP 中搜索 SAD 值小于另一阈值的 SP(final SPP,简称 FSPP),最后以 FSPP 集合中的各点作为搜索起点,采用传统的快速运动估计算法(如 DS)计算该边界宏块的运动向量;当包围盒内所有边界宏块的运动向量全部计算完毕后,对于某内

部宏块,将与该宏块相邻的边界宏块或内部宏块的运动向量作为起点,采用快速运动估计算法即可搜索得到目标宏块的运动向量.

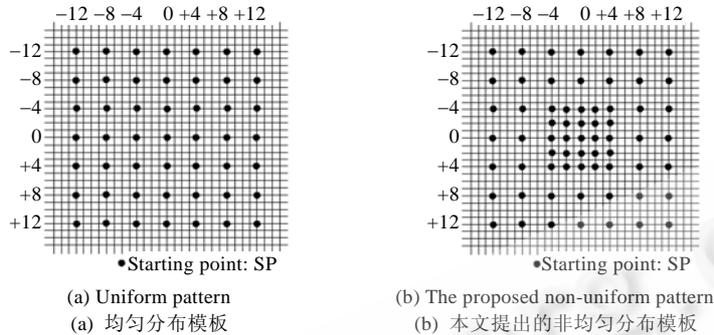


Fig.4 Starting point pattern

图 4 搜索起点分布模板

BAAS(binary alpha-plane assisted search)算法在计算量较低的情况下,为基于亮度信息的运动估计选取了较接近全局最优点的 SP 作为搜索起点,有利于运动估计得到全局最优的运动向量.然而,由于未能充分利用 alpha 平面以及建立更加合理的 SPP,使得该算法尚存在一些不足.这主要表现在两个方面:

1) alpha 平面的匹配误差准则 BAMS 尽管简单,但仍不足以有效区别两个具有不同 alpha 平面的宏块.图 5(b)和图 5(c)是两个不同参考宏块的 alpha 平面,并且形状迥异,可是这两个宏块与目标宏块的 BAMS 误差却相同,即:两个参考宏块成为全局最优匹配宏块的概率是相同的;而在视觉上,参考宏块 1 则更有可能成为匹配宏块.此时,两个参考宏块对应的 SP 要么全部被舍弃,要么全部加入 USPP.若是后者,仍需计算目标宏块与参考宏块的 SAD,而参考宏块 2 的 SAD 是完全不需要计算的.因此,BAMS 匹配误差准则只能粗略定位 SP;它引入的错误 SP 将增加运动估计的计算量.究其原因,在于没有利用 alpha 平面的边界特征.

2) 文献[11]通过大量的实验统计发现,自然视频序列的运动向量具有中心偏置特性,而 BAAS 采用均匀分布的 SPP:一方面,远离中心处的 SP 将增加搜索的计算量;另一方面,靠近中心处的 SP 则可能因为分布稀疏而不利于定位到全局最优,影响估计的精度.



Fig.5 Two reference macro-blocks of different shape but with the same BAMS

图 5 形状迥异的参考宏块具有相同的 BAMS 值示意图

2.2 二值alpha平面的匹配衡量准则

基于上述分析,本文提出了一种新的二值 alpha 平面匹配衡量准则 WBAMC,其衡量指标包括:1) 边界匹配程度:若目标宏块与参考宏块能够很好地匹配,那么,其边界也应相互重叠;否则,这两个宏块是误匹配的可能性就很大;2) 参考宏块与目标宏块的距离:根据中心偏置特性,距离中心越远的候选向量成为运动向量的可能性越小.据此,本文提出用边界掩码技术衡量边界的匹配程度、用距离加权方法提高零向量及其附近向量被选中的概率.

2.2.1 alpha 平面的边界掩码

在 VO 的运动过程中,其 VOP 的边界往往随时间发生或多或少的变化,这就可能导致本应重叠的两个宏块

边界产生一定的偏离.为了提高匹配衡量准则对上述边界时变的鲁棒性,本文首先根据 alpha 平面确定边界像素,然后将宽度为 1 个像素的边界扩展成宽度为若干个像素的区域,在该区域内的每个像素,按照其与原始边界像素的距离,分配给不同的权值.本文约定,由这些权值组成的信息称为“边界掩码”,而具有边界掩码的 alpha 平面称为“MA(masked alpha)平面”.MA 平面不仅强化了 alpha 平面的边界特征,还有利于较细致地度量两个边界的偏离程度.图 6 为 Stefan 序列第 1 帧的边界掩码示意图,其中,图 6(a)为原始边界,图 6(b)为扩展后的边界,其宽度为 4 个像素.为直观起见,图中用白色代表较小的权值,用灰色代表较大的权值.

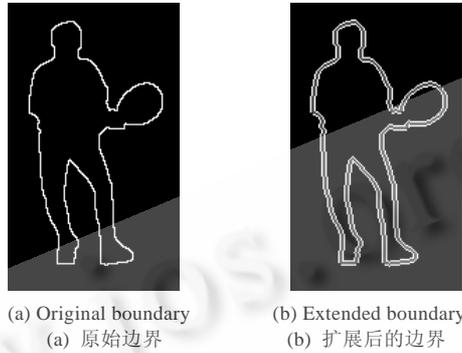


Fig.6 The boundary mask of alpha-plane
图 6 alpha 平面的边界掩码技术示意图

计算边界掩码的具体算法如下:

- Step 1. 在对象包围盒中,找到 alpha 平面的第 1 行信息;
- Step 2. 自左向右扫描该行信息,如果某个像素是边界像素,则将该像素的左侧作为对称轴(如图 7(a)中箭头所指),将对称轴左侧和右侧的各 μ 个像素置为边界像素(图 7(b)为当 $\mu=2$ 时的情况).设对称轴两侧,距离对称轴 1 个像素的边界掩码为 α ,距离 2 个像素的边界掩码为 β ,以此类推,且满足 $\alpha > \beta > \dots$ (图 7(c)所示为当 $\mu=2$ 时的情况).除边界区域外,alpha 平面上其余像素的信息保持不变;
- Step 3. 若包围盒中的行均已扫描完毕,则转入 Step 4;否则,扫描得到下一行像素,并转入 Step 2;
- Step 4. 在对象包围盒中,找到 alpha 平面的第 1 列信息;
- Step 5. 自上而下扫描该列信息,利用与 Step 2 相同的方法计算边界掩码.若某个像素在行扫描时已经计算出了掩码,则在行掩码和列掩码中取较大者作为该像素的边界掩码;
- Step 6. 若包围盒中的列均已扫描完毕,则算法结束;否则,扫描得到下一列像素,并转入 Step 5.

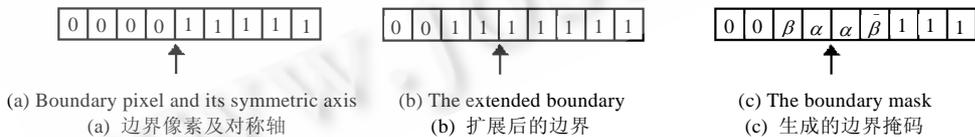


Fig.7 The computation of boundary mask for $\mu=2$ in one-dimensional case, where 0 denotes transparent pixels and 1 denotes opaque pixels

图 7 一维情况下,计算 $\mu=2$ 时边界掩码的示意图,其中,0 表示外部像素,1 表示内部像素

2.2.2 WBAMC 匹配衡量准则

在 MA 平面的基础上,本文提出了一种新的匹配衡量准则 WBAMC,分为以下两个步骤:

- 1) 计算无加权 BAMC,定义如下:

$$BAMC(u, v) = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} MA_c(i, j) \& MA_r(i + u, j + v) \tag{4}$$

其中, (u, v) 为候选向量, N 和 M 分别表示宏块的高度和宽度, $MA_c(\cdot, \cdot)$ 和 $MA_r(\cdot, \cdot)$ 分别表示当前边界宏块和参考边界

宏块的MA平面值,“&”表示“与(and)”运算;

2) 计算 WBAMC,定义如下:

$$WBAMC = \begin{cases} BAMC(u, v) - \delta \times (|u| + |v|), & (u, v) \neq (0, 0) \\ BAMC(u, v) + \sigma \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} BA_c(i, j), & (u, v) = (0, 0) \end{cases} \quad (5)$$

其中,|·|表示取绝对值运算, δ 和 σ 为比例系数, $BA_c(\cdot, \cdot)$ 表示当前宏块的alpha平面.

2.3 alpha平面辅助的边界宏块运动估计算法

基于 WBAMC 准则,本文提出了 alpha 平面辅助的边界宏块快速运动估计算法.根据运动向量的分布特性,本文首先确定了一个非均匀分布的搜索起点模板(non-uniform SPP,简称 NUSPP)(如图 4(b)所示).由于边界掩码的扩展半径为 2,为了使边界特征的匹配能够覆盖整个搜索窗口,故而将模板上各起点间的基本步长设置为 4.在此基础上,中心区域的步长缩小 1 倍,以增加中心区域的搜索精度.

算法的基本思想是:根据 NUSPP(non-uniform starting point pattern),计算目标边界宏块与每个 SP 对应的参考宏块之间的匹配程度,其 WBAMC 值越大,说明该 SP 接近全局最优点的可能性就越大.从所有 SP 中,选择 WBAMC 最大值点 SP_1 ,以及另外一点 SP_2 ,且 SP_2 满足:

$$WBAMC_{SP_2} = \max\{WBAMC_{SP} | |X_{SP_1} - X_{SP}| + |Y_{SP_1} - Y_{SP}| > 4\} \quad (6)$$

其中, X 和 Y 分别表示 SP 点的横坐标和纵坐标.以 SP_1 和 SP_2 为搜索起点,采用某种快速运动估计算法搜索运动向量.具体算法如下:

- Step 1. 按照第 2.2.1 节中的算法,利用 alpha 平面,分别计算目标 VOP 和参考 VOP 的边界掩码;
- Step 2. 在目标 VOP 的包围盒中,查找一个未进行运动估计的边界宏块.若找到,转入 Step 3;否则,算法结束;
- Step 3. 按照第 2.2.2 节中的匹配衡量准则,计算出目标宏块的 NUSPP 中每个 SP 相应的 WBAMC 值;
- Step 4. 从所有 SP 的 WBAMC 值中查找最大点 SP_1 和满足式(6)的 SP_2 ;
- Step 5. 分别以 SP_1 和 SP_2 为运动估计的起始搜索点,利用 DS 算法搜索 SAD 值最小的点,并将其对应的向量作为目标宏块的运动向量.转入 Step 2.

2.4 alpha平面辅助快速运动估计算法

对于自然视频的同一个VOP,其运动的空间相关性是比较明显的,这不仅体现为内部宏块的运动与相邻的边界宏块的运动高度相关,还体现在相邻的内部宏块之间也存在着很强的运动相关性^[16].因此,优先搜索策略可以在保证补偿质量的前提下,提高运动估计的速度.

在上面两节的基础上,同时结合优先搜索策略,本文提出了一种 alpha 平面辅助的任意形状视频对象快速运动估计算法 BAAME.给定一个目标 VOP,算法的具体过程如下:

- Step 1. 根据第 2.3 节中的算法,首先计算出 VOP 内所有边界宏块的运动向量;
- Step 2. 在 VOP 的包围盒内,查找一个未进行运动估计的内部宏块.若找到,则转入 Step 3;否则,算法结束;
- Step 3. 求出与目标宏块相邻的、并经过运动估计的所有宏块的运动向量及零向量(0,0)的平均值,作为目标宏块的搜索起点,即

$$\vec{V}_{init} = avg(u, v) \quad (7)$$

其中, $(u, v) \in \{\text{相邻宏块的运动向量}, (0, 0)\}$, avg 表示求平均值的运算.这样可以充分利用运动的空间相关性和中心偏置特性,并能节省 PSA(BAAS+DS)算法中计算多个候选 SP 的 SAD 的运算量;

- Step 4. 以 \vec{V}_{init} 为起点,采用 DS 算法搜索目标宏块的运动向量.转入 Step 2.

3 实验结果与分析

为了验证本文所提出算法 BAAME 的运动估计性能,分别以 CIF(common intermediate format)格式的标准视频流 Foreman, Coastguard 和 Stefan 的 1-90 帧为例进行了大量的实验,并将实验结果与 FS, DS, PSA(BAAS+DS)

算法比较.对于每一帧图像,将其前一帧作为参考帧,用第 2.4 节中的方法进行运动估计,并用参考帧中的最佳匹配块直接替代目标帧中的每个块,从而得到目标帧的补偿.

实验过程中,PSA(BAAS+DS)算法的参数设置与文献[10]相同;BAAME 的参数设置通过实验获得,分别为 $\mu=2, \alpha=7, \beta=3, \delta=8, \sigma=1/16$.需要特别指出的是,将 δ 和 σ 设置为 8 和 1/16 的目的还在于,与 δ 和 σ 的乘法运算可以通过移位运算来实现,进而提高运算速度.另外,FS 和 DS 的搜索窗口均为 33×33 ;在 PSA(BAAS+DS)和 BAAME 算法中,USPP 和 NUSPP 的搜索窗口均为 33×33 ,而随后进行的 DS,其搜索范围缩小至 8×8 .所有算法的宏块大小设置为 16×16 .

3.1 WBAMC 匹配衡量准则的分布曲面

本文提出的 WBAMC 准则,兼顾了 alpha 平面的边界特征及自然视频的运动向量分布特点,能够更加准确地评价两个边界宏块的匹配程度.图 8 所示为利用 WBAMC 准则对图 1(c)中所示宏块在搜索窗口内的匹配程度进行衡量得到的分布曲面(特别需要指出的是,由于 WBAMC 准则用来衡量匹配程度,值越大,表明匹配可能性越大,因此,它的分布与 SAD, BAMS 准则的分布所代表的意义完全相反.为了方便比较,本文对原始的分布作如下改动:用全局最大值减去每个 SP 处的 WBAMC 值,用所得的差代替该 SP 处的值,即相当于把原始的 WBAMC 分布曲面翻转了 180°).

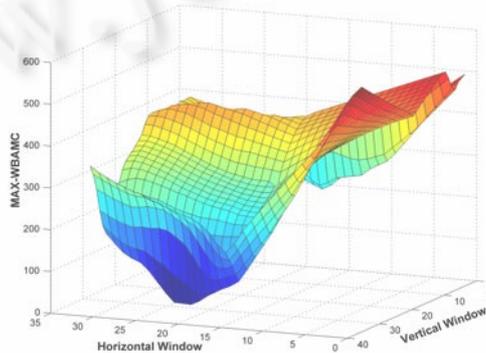


Fig.8 The distribution surface of the WBAMC matching criterion (having rotated by 180 degree)

图 8 WBAMC 匹配衡量准则的分布曲面(翻转 180°)

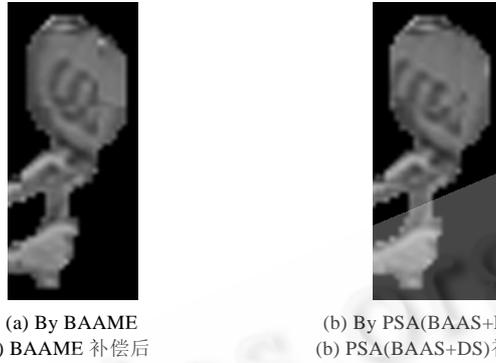
对比图 2、图 3 和图 8 可以发现, WBAMC 准则的分布趋势较之 BAMS 准则更加接近于 SAD 的分布趋势,这主要是因为 WBAMC 考察了更多特征缘故.另外, WBAMC 分布的单调性也比较明显,尽管不能避免出现多个峰值的现象,可是每个峰的内部,其单调性是很强的,这有利于将峰值点集中于少数几个 SP,而这几个 SP 有很大的可能性将会处于 SAD 全局最优点的单调区域内,进而提高了整个运动估计算法搜索到全局最优向量的准确度.实验表明,只要利用两个峰值点即可达到与 FS 相当的估计质量,而 BAMS 则需要更多的 SP.因此,本文提出的 WBAMC 匹配衡量准则是有效的.

3.2 运动估计/补偿质量的比较

为了验证算法的主观补偿质量,分别用 BAAME 算法和 PSA(BAAS+DS)算法对图 1(a)所示的原始帧进行运动估计和补偿,得到的补偿帧如图 9 所示.对比发现,本文提出的算法能够补偿得到更加连续的边界,获得更高的主观质量,进一步说明了 WBAMC 准则的有效性.

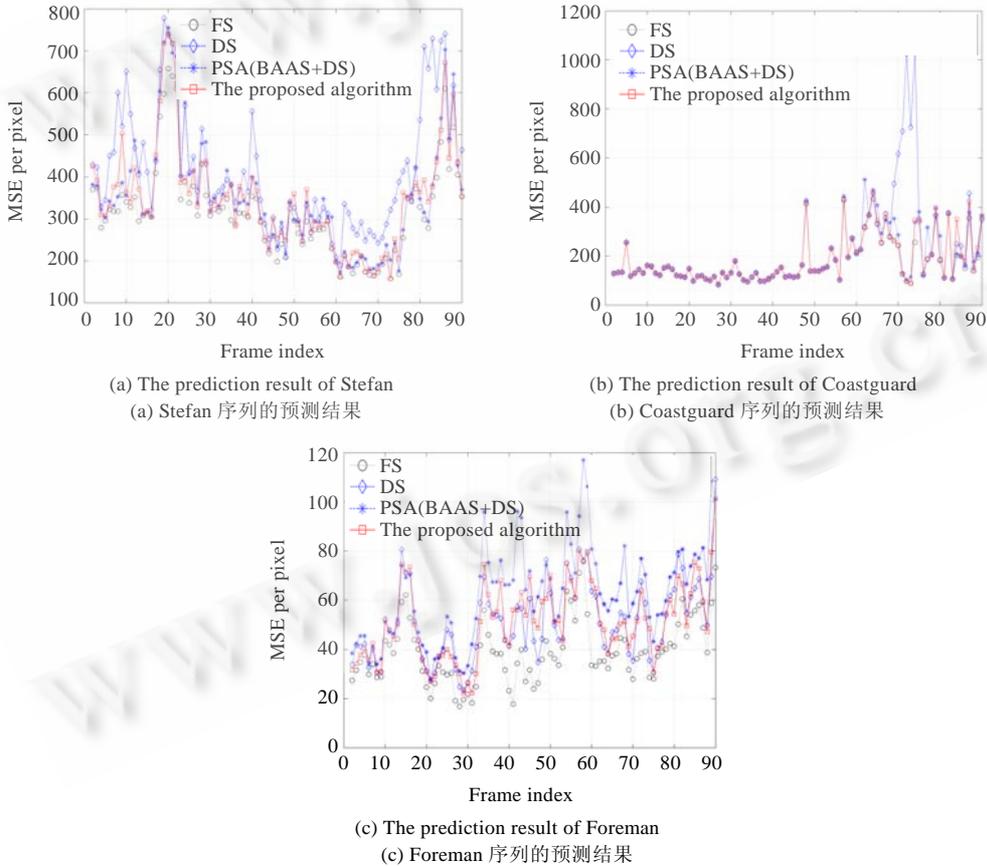
另外,为了验证算法的客观补偿质量进行了大量的统计实验,并分别用 MSE(mean square error)和峰值信噪比衡量重构图像帧的质量.图 10(a)~图 10(c)所示分别为 Stefan, Coastguard 和 Foreman 序列使用上述 4 种运动估计方法得到的补偿 MSE 的比较结果.从中可以发现,对于空间细节复杂而运动复杂性中等的一类视频(如 Stefan),以及空间细节复杂度中等而运动复杂性较低的一类视频(如 Coastguard),DS 算法往往陷入局部最优,其

预测 MSE 总是很大;而 PSA(BAAS+DS)和 BAAME 由于利用了 alpha 平面和边界特征,能够较为准确地预测边界宏块,并指导内部宏块的估计,从而获得了优于 DS 的补偿质量;对于空间细节较简单而运动复杂度较低的一类视频(如 Foreman),DS 的补偿质量有了一定程度的提高,甚至 PSA(BAAS+DS)的补偿质量逊于 DS.



(a) By BAAME (a) BAAME 补偿后 (b) By PSA(BAAS+DS) (b) PSA(BAAS+DS)补偿后

Fig.9 The motion compensated Fig.1(a) by BAAME and PSA(BAAS+DS) respectively
图 9 图 1(a)经过 BAAME 和 PSA(BAAS+DS)的运动补偿帧



(a) The prediction result of Stefan (a) Stefan 序列的预测结果 (b) The prediction result of Coastguard (b) Coastguard 序列的预测结果 (c) The prediction result of Foreman (c) Foreman 序列的预测结果

Fig.10 Comparison of the motion compensation quality of the proposed motion estimation scheme with three traditional schemes

图 10 本文提出的运动估计方法与其他 3 种方法补偿结果的比较

表 1 列出了利用 4 种运动估计方法对 3 个视频序列进行运动补偿后所得到的平均峰值信噪比.从中能够看出,对于不同特性的视频序列,本文算法的补偿质量总是优于 DS 和 PSA(BAAS+DS)算法,且仅次于 FS.这表明本文提出的运动估计算法是稳定而有效的,并且 WBAMC 准则能够比 PSA(BAAS+DS)的 BAMS 准则更高效地辅助基于亮度信息的运动估计.

Table 1 The average PSNR motion-compensated by four motion estimation schemes

表 1 4 种运动估计方法所得到的平均 PSNR

Sequence	FS	DS	PSA(BAAS+DS)	BAAME
Stefan	23.4	22.2	22.8	23.0
Coastguard	26.0	25.5	25.8	25.9
Foreman	32.5	31.2	30.4	31.2

3.3 计算复杂度的比较

计算时间复杂度是衡量运动估计算法性能的另一项重要指标,为了方便比较,本文对 4 种运动估计方法的计算量进行了统计,并将每个 VOP 的平均计算量作为算法的平均计算复杂度.根据文献[10],本文约定一次 8 位定点整数加法,其复杂度相当于一次位运算复杂度的 16 倍.

FS 算法的计算量取决于 VOP 内部像素的数量.PSA(BAAS+DS)算法的计算量主要集中在:1) alpha 平面的匹配;2) FSPP 的确定;3) DS 搜索.本文提出算法的计算量主要集中在:1) alpha 平面的匹配;2) DS 搜索.本文统计了 4 种算法对各个测试序列进行运动估计时所需要的加法、取绝对值、位运算和乘法运算的数量,并计算出每个 VOP 所需各种运算的平均数量,结果见表 2.

Table 2 Comparison of average computational complexity of four motion estimation schemes

表 2 4 种运动估计方法的平均计算复杂度比较

Sequence	Operation	FS	DS	PSA(BAAS+DS)	BAAME
Stefan	Addition	9 755 690	193 253	874 921	766 772
	Absolute	4 908 527	97 132	250 553	171 866
	Bit-Wise OP	0	0	23 071 (369 132/16)	26 940 (431 046/16)
	Multiplication	0	0	59	6
Coastguard	Addition	10 244 693	168 009	763 444	1 057 056
	Absolute	5 152 384	84 433	175 929	295 602
	Bit-Wise OP	0	0	25 354 (405 656/16)	29 660(474 556/16)
	Multiplication	0	0	67	3
Foreman	Addition	73 989 198	1 389 933	2 558 885	2 275 582
	Absolute	37 157 324	697 495	958 263	773 632
	Bit-Wise OP	0	0	39704 (635 260/16)	46 440(743 039/16)
	Multiplication	0	0	110	226

由表 2 的实验数据可知:

- 1) 本文所提出算法的计算复杂度高于 DS,这是由于对每个边界宏块的运动估计都需要执行 2 次 DS;
- 2) 所提出算法的计算复杂度总体上低于 PSA(BAAS+DS),即 FSPP(final starting point pattern)的确定浪费了后者的时间;
- 3) 从各个序列的平均情况来比较,本文所提出算法的计算复杂度较之 FS 降低了 20 倍,而较之 PSA(BAAS+DS),则降低了 4%.

最后,本文提出的方法可以作为一种运动估计技术的框架,广泛应用于基于对象的各种视频编码器中.它不仅可以与现有的各种空间域快速运动估计方法^[3-6,17-23]相结合,提高传统方法运动补偿的主、客观质量,还可以与各种频域运动估计方法^[24-27]相结合,进一步提高其估计速度.

4 结 论

本文详细分析了 alpha 平面在基于对象的运动估计中所发挥的作用,并提出了一种新的用于 alpha 平面的匹配衡量准则 WBAMC.基于该准则,提出了用于边界宏块的快速运动估计算法.最后,结合优先搜索策略提出了

一种 α 平面辅助的视频对象快速运动估计方法.实验结果表明,BAAME 方法性能稳定,运动补偿的平均 PSNR 较 DS 和 PSA(BAAS+DS)高出 0.1~0.8dB,略低于 FS,但是其计算复杂度较 FS 降低了 20 倍.

α 平面的引入,对于基于对象的视频编码器中运动估计技术的研究具有一定的理论意义.WBAMC 准则是衡量两个边界宏块匹配程度的简单而快速的方法.

致谢 感谢辽宁师范大学计算机与信息技术学院多媒体信息处理讨论班上老师和同学们的热烈讨论,以及在算法实验过程中肖羽同学所提供的帮助.

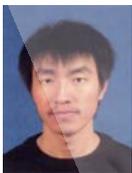
References:

- [1] ISO/IEC JTC1/SC29/WG11. Information technology—coding of audio-visual objects, part 2: Visual. ISO/IEC 14496-2, Int'l Standard (2nd ed.), 2001.
- [2] ISO/IEC JTC1/SC29/WG11. MPEG-4 video verification model 18.0. N3908, 2001.
- [3] Po LM, Ma WC. A novel four-step search algorithm for fast block motion estimation. *IEEE Trans. on Circuits and Systems for Video Technology*, 1996,6(3):313–317.
- [4] Liu LK, Feig E. A block-based gradient descent search algorithm for block motion estimation in video coding. *IEEE Trans. on Circuits and Systems for Video Technology*, 1996,6(4):419–423.
- [5] Zhu S, Ma KK. A new diamond search algorithm for fast block-matching motion estimation. *IEEE Trans. on Image Processing*, 2000,9(2):287–290.
- [6] Cheung CH, Po LM. Novel cross-diamond-hexagonal search algorithms for fast block motion estimation. *IEEE Trans. on Multimedia*, 2005,7(1):16–22.
- [7] Chan YL, Siu WC. Edge oriented block motion estimation for video coding. *IEEE Proc.: Vision, Image and Signal Processing*, 1997,144(3):136–144.
- [8] Chan YL, Siu WC. An efficient search strategy for block motion estimation using image features. *IEEE Trans. on Image Processing*, 2001,10(8):1223–1238.
- [9] Lee JH, Ra JB. Efficient motion estimation using edge-based binary block-matching and refinement based on motion vector correlation. In: *Proc. of the IEEE Int'l Conf. on Image Processing*. IEEE Press, 2001. 957–960. <http://ieeexplore.ieee.org/iel5/7594/20718/00958654.pdf?tp=&arnumber=958654&isnumber=20718>
- [10] Hui KC, Siu WC, Chan YL. Fast motion estimation of arbitrarily shaped video objects in MPEG-4. *Signal Processing: Image Communication*, 2003,18(1):33–50.
- [11] Li R, Zeng B, Liou ML. A new three-step search algorithm for block motion estimation. *IEEE Trans. on Circuits and Systems for Video Technology*, 1994,4(4):438–442.
- [12] ITU-R Document 6Q/14. Final report from the video quality experts group on the validation of objective models of video quality assessment. Phase II (FR-TV2), 2003. <http://www.vqeg.org>
- [13] Miyahara M, Kotani K, Algazi VR. Objective picture quality scale (PQS) for image coding. *IEEE Trans. on Communication*, 1998, 46(9):1215–1226.
- [14] Ginesu G, Massidda F, Giusto DD. A multi-factors approach for image quality assessment based on a human visual system model. *Signal Processing: Image Communication*, 2006,21(4):316–333.
- [15] Lee J, Ro YM. Hierarchical block matching algorithm in MRME. In: *Proc. of the Int'l Conf. on Image Processing (ICIP 2000)*. IEEE Press, 2000. 836–839. <http://ieeexplore.ieee.org/iel5/7221/19490/00901089.pdf?tp=&arnumber=901089&isnumber=19490>
- [16] Chung KL, Chang LC. A new predictive search area approach for fast block motion estimation. *IEEE Trans. on Image Processing*, 2003,12(6):648–652.
- [17] Li B, Li W, Tu YM. A fast block matching algorithm using smooth motion vector field adaptive search technique. *Journal of Computer Science and Technology*, 2003,18(1):168–173.
- [18] Wang Z, Liu ZG, Wang GX. An adaptive search algorithm for block motion estimation. *Journal of Computer Research and Development*, 2003,40(7):1036–1041 (in Chinese with English abstract).

- [19] Lu ZH, Li H, Chu JH, Yu SL. A fast motion estimation search algorithm with adjustable predictive quality. Acta Electronica Sinica, 2004,32(6):980–982 (in Chinese with English abstract).
- [20] Zheng QH, He JM, Fan ZW, Wang JM. A fast motion-estimation algorithm based on adaptive block search range for H. 264. Journal of Image and Graphics, 2005,10(11):1350–1354 (in Chinese with English abstract).
- [21] Zhang JY, Li B, Yao CL. A fast sub-pixel motion estimation algorithm based on best position calculation. Chinese Journal of Computers, 2005,28(9):1541–1547 (in Chinese with English abstract).
- [22] Xia JX, Huang SJ. Algorithm of adaptive rood pattern search using characteristic of different kinds of block in VOP. Journal of Communications, 2005,26(8):117–121 (in Chinese with English abstract).
- [23] Zou XC, Feng Y, Zhao XB. A fast block-matching motion estimation algorithm. Journal of Image and Graphics, 2006,11(7):938–942 (in Chinese with English abstract).
- [24] Wang XH. Research on spatial scalable video coding based on wavelet. Computer Science, 2002,29(7):115–117 (in Chinese with English abstract).
- [25] Argyriou V, Vlachos T. Quad-Tree motion estimation in the frequency domain. In: Proc. of the IEEE Int'l Conf. on Multimedia and Expo. (ICME 2005). IEEE Press, 2005. 1432–1435. <http://ieeexplore.ieee.org/iel5/10203/32544/01521700.pdf?tp=&arnumber=1521700&isnumber=32544>
- [26] Song CM, Wang XH. Research progress in video motion estimation techniques in the wavelet domain. Chinese Journal of Computers, 2005,28(10):1716–1727 (in Chinese with English abstract).
- [27] Song CM, Wang XH. A new scalable video motion estimation scheme in the wavelet domain. Chinese Journal of Computers, 2006, 29(12):2112–2118 (in Chinese with English abstract).

附中文参考文献:

- [18] 王贇,刘治国,王光兴.自适应块匹配运动估计搜索算法.计算机研究与发展,2003,40(7):1036–1041.
- [19] 鲁照华,李华,褚晶辉,俞斯乐.一种预测质量可控的快速运动估计搜索算法.电子学报,2004,32(6):980–982.
- [20] 郑庆辉,何加铭,范志蔚,王军明.一种 H.264 自适应块搜索范围快速运动估计算法.中国图像图形学报,2005,10(11):1350–1354.
- [21] 张苒尹,李波,姚春莲.基于最优位置计算的快速亚像素运动估计.计算机学报,2005,28(9):1541–1547.
- [22] 夏金祥,黄顺吉.基于 VOP 的块特性的自适应十字搜索模式运动估计法.通信学报,2005,26(8):117–121.
- [23] 邹晓春,冯燕,赵歆波.一种块匹配的快速运动估计算法.中国图像图形学报,2006,11(7):938–942.
- [24] 王相海.基于小波的视频空间可分级编码研究.计算机科学,2002,29(7):115–117.
- [26] 宋传鸣,王相海.小波域视频运动估计研究进展.计算机学报,2005,28(10):1716–1727.
- [27] 宋传鸣,王相海.一种新的小波域视频可分级运动估计方案.计算机学报,2006,29(12):2112–2118.



宋传鸣(1980—),男,辽宁沈阳人,博士生,CCF 学生会员,主要研究领域为图像,视频编码,多媒体水印技术.



张福炎(1939—),男,教授,博士生导师,CCF 高级会员,主要研究领域为计算机图形学,多媒体技术.



王相海(1964—),男,博士,教授,CCF 高级会员,主要研究领域为计算机图形学,多媒体信息处理.