

面向运动估计器硬件实现的加权块匹配检测算法

张 霞，郑南宁，张光烈，吴 勇，王少瑞，徐维朴

(西安交通大学人工智能与机器人研究所，西安 710049)

[摘要] 在各种视频处理算法中，运动补偿型算法大大提高了各种视频处理效果。运动估计器的硬件实现是各种运动补偿视频处理算法在实际系统中运用的关键。由于块匹配运动估计算法较低的运算复杂度和硬件实现难度以及块匹配检测标准函数很高的调用频率，已经广泛应用到各种实际系统中。文中提出了加权最小最大误差的匹配检测标准，能够降低运动估计器的运算复杂度，减少估计器的硬件面积，提高硬件速度，而且能够降低递归搜索块匹配运动估计算法固有误差传递带来的负面影响。

[关键词] 视频处理；运动补偿；运动估计；块匹配检测标准

[中图分类号] V475; V476 **[文献标识码]** A **[文章编号]** 1009-1742(2002)01-0047-07

1 引言

为提高电视图像质量，通常进行以下处理：去隔行，帧频提升，降噪和图像边缘增强等。在不同的算法类型中，运动补偿型算法是最高级的算法，如运动补偿去隔行、运动补偿帧频提升和降噪等。数字视频与数字图像的主要区别在于数字视频信号具有时间维度，其中包含运动信息。在运动信息能够准确得到的情况下，运动补偿型算法会比非运动补偿型算法的效果好。在运动补偿算法中，运动信息的准确度以及获得的难易度是各种运动补偿型算法的核心。运动估计模块通常称为运动估计器。运动估计器的硬件实现及其价格成为各种运动补偿型数字视频处理算法在实际系统中应用的关键。

在数字视频信号降噪和视频格式转换（去隔行、帧频提升等）算法中，基于块运动估计和运动补偿算法得到了广泛应用^[1]。降低运算复杂度是实现运动估计器的关键。块匹配运动估计器的改进通常从下面三个方面着手^[2]：块与像素的二次采样，采用智能搜索策略，寻找简单高效的块匹配检

测标准。对一帧图像，在匹配块检测过程中不必利用块中的所有像素，对块中像素进行采样率高达4:1的二次采样，实验证明不降低匹配块的效果，而且还可以对噪声产生抑制作用。当用运动估计器设计视频格式转换时，要得到的运动场信息是平滑的，其次考虑它的正确性。因此所采用的算法中应该包含运动场的平滑限制。三维递归搜索策略^[3]是考虑了这些方面而提出的一种快速搜索策略。

在块匹配运动估计过程中，块匹配检测标准的调用是非常频繁。降低块匹配检测标准的运算复杂度，将会使整个估计器的运算速度大大降低。因此，块匹配检测标准必须是计算简单、硬件实现容易的。上面所提的三维递归搜索策略有一个固有的缺点——误差传递性。为了保证运动场的平滑性，三维递归搜索策略将当前块的邻块运动矢量为其预测值，并对该值稍加修改形成当前块的待测矢量集合。这样，在保证运动矢量场一致性的同时，会导致估计的运动矢量的误差传递。在递归方向上进行该误差传递过程，使运动估计器不能够快速地对突然出现的运动物体做出反应。特别是对于大面积亮

[收稿日期] 2001-06-03；修回日期 2001-08-31

[基金项目] 国家“九五”攻关计划资助项目(96-120-28)

[作者简介] 张 霞(1978-)，女，河南长垣县人，西安交通大学硕士研究生

度变化缓慢的图像区域，由于这个区域内各个块之间的亮度差异较小，噪声的影响变得很大，会使在这个区域内检测到的运动矢量场在大小和方向上呈随机分布，这时如果出现一个运动物体，运动估计器对运动速度的变化不能做出快速的正确反应。为了减少递归搜索运动估计器的这个缺点的影响，修改块匹配检测标准，为匹配检测标准增加一个惩罚系数，使在大面积亮度变化缓慢的图像区域内检测到的运动矢量均在零矢量附近。这样，递归搜索运动估计器可对速度的突然变化做出较快的反应。

提出一种改进的加权最小最大误差标准。这种块匹配检测标准计算简单，硬件实现容易，运算速度快，并且所增加的权值能够有助于在递归搜索策略中降低误差传递对运动估计结果的影响。

2 块匹配运动估计算法原理及块匹配检测标准

在块匹配运动估计算法中，当前帧 n 中一个以 \mathbf{X} 为中心的块 $B(\mathbf{X})$ 的位移矢量 $\mathbf{D}(\mathbf{X}, n)$ 通过在一个搜索区域 $S_A(\mathbf{X})$ 中搜索一个相似块得到。这个搜索区域也以 \mathbf{X} 为中心，但处于相邻帧中，如帧 $n-1$ 。搜索到的相似块的中心相对于 \mathbf{X} 有一个位移矢量 $\mathbf{D}(\mathbf{X}, n)$ 。为了寻找 $\mathbf{D}(\mathbf{X}, n)$ ，一些待测位移矢量 \mathbf{C} 被一个匹配误差检测函数 $\epsilon(\mathbf{C}, \mathbf{X}, n)$ 所评估，这个函数量化了两个块之间的相似性。块匹配运动估计算法见图 1^[4]。

定义待测位移矢量 \mathbf{C} 的集合为 C_S^{\max} ，这个集合包括了前一帧中搜索区域 $S_A(\mathbf{X})$ 内相对于 \mathbf{X} 的所有可能位移（均处于图像的像素网格上）：

$$C_S^{\max} = \{ \mathbf{C} \mid -N \leq C_x \leq +N, \\ -M \leq C_y \leq +M \}. \quad (1)$$

式中 N 和 M 是限制搜索区域 $S_A(\mathbf{X})$ 的常数。当前帧中一个以 $\mathbf{X} = (X_x, X_y)$ 为中心，大小为 $X \times Y$ 的块 $B(\mathbf{X})$ 所包括像素 $\mathbf{x} = (x, y)$ 的位置为：

$$B(\mathbf{X}) = \{ \mathbf{x} \mid (X_x - 0.5X) \leq x \leq \\ X_x + 0.5X, (X_y - 0.5Y) \leq y \leq X_y + 0.5Y \}. \quad (2)$$

全搜索块匹配运动估计算法得到的位移矢量 $\mathbf{D}(\mathbf{X}, n)$ 是一个具有最小匹配误差检测函数 $\epsilon(\mathbf{C}, \mathbf{X}, n)$ 值的待测位移矢量 \mathbf{C} ：

$$\mathbf{D}(\mathbf{X}, n) \in \{ \mathbf{C} \in C_S^{\max} \mid \epsilon(\mathbf{C}, \mathbf{X}, n) \leq \\ \epsilon(\mathbf{F}, \mathbf{X}, n), \forall \mathbf{F} \in C_S^{\max} \}. \quad (3)$$

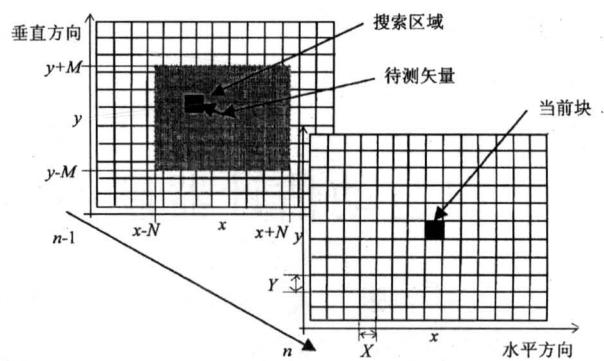


图 1 块匹配运动估计

Fig. 1 Block-matching motion estimation

一般来说，具有最小匹配误差的运动矢量 $\mathbf{D}(\mathbf{X}, n)$ 被分配给块 $B(\mathbf{X})$ 中的所有像素 \mathbf{x} 。一个待测运动矢量 \mathbf{C} 的块匹配误差值是当前块和前一帧中移位块的各像素亮度值的函数（称为成本函数），并在块 $B(\mathbf{X})$ 上求和。

$$\epsilon(\mathbf{C}, \mathbf{X}, n) = \sum_{\mathbf{x} \in B(\mathbf{X})} C_{\text{cost}}(F(\mathbf{x}, n), \\ F(\mathbf{x} - \mathbf{C}, n - p)). \quad (4)$$

视频信号为隔行扫描， $p=2$ ；若为逐行扫描， $p=1$ 。这个成本函数非常直接和易于计算，但由于这个函数很高的重复计算率将给处理系统带来很大的计算负担。为减少块匹配运动估计算法的计算复杂度以便于硬件实现，通常从几个方面考虑：

- 1) 采用一个简单的成本函数作为匹配准则；
- 2) 对匹配块进行像素二次采样；
- 3) 采用智能搜索策略，避免对所有的矢量进行检测。

对匹配块进行像素二次采样，公认采用 2:1 或者 4:1 的采样率对运动估计结果的负面影响不大。快速搜索算法包括三步法、四步法、交叉搜索法等等。这些快速算法大大降低了运动估计算法的运算速度，但是以找到局部最优来代替全搜索算法的全局最优。一种适合于运动补偿数字视频格式转换的运动估计算法是三维递归搜索运动估计算法。在这个算法设计中考虑了硬件实现和运动矢量场的一致性，运用了递归搜索策略。递归搜索策略在降低运算复杂度的同时，也带来了递归算法的一个固有缺点——误差传递。

鉴于匹配检测标准函数的高重复计算率，为提高块匹配运动估计算法的运算速度，需要简单的匹配检测标准。通常所用的匹配检测标准有^[5]：

- 互相关函数 (CCF)^[6]

$$F_{CC}(C, X, n) = \frac{\sum_{x \in B(X)} F(x, n) F(x - C, n - p)}{\left[\sum_{x \in B(X)} F^2(x, n) \right]^{1/2} \left[\sum_{x \in B(X)} F^2(x - C, n - p) \right]^{1/2}},$$

$$D(X, n) \in \{ C \in C_S^{\max} | F_{CC}(C, X, n) \geq F_{CC}(F, X, n), \forall F \in C_S^{\max} \}。 \quad (5)$$

- 平均平方误差 (MSE)^[7]

$$E_{MS}(C, X, n) =$$

$$\frac{1}{XY} \sum_{x \in B(X)} [F(x, n) - F(x - C, n - p)]^2,$$

$$D(X, n) \in \{ C \in C_S^{\max} | E_{MS}(C, X, n) \leq E_{MS}(F, X, n), \forall F \in C_S^{\max} \}。 \quad (6)$$

- 平均绝对误差 (MAE)^[8]

$$E_{MA}(C, X, n) =$$

$$\frac{1}{XY} \sum_{x \in B(X)} |F(x, n) - F(x - C, n - p)|,$$

$$D(X, n) \in \{ C \in C_S^{\max} | E_{MA}(C, X, n) \leq E_{MA}(F, X, n), \forall F \in C_S^{\max} \}。 \quad (7)$$

- 像素差值分类 (PDC)^[9]

$$T(C, x, n) = \begin{cases} 1 & |F(x, n) - F(x - C, n - p)| \leq Th, \\ 0 & \text{其他} \end{cases}$$

$$C_{PD}(C, X, n) = \sum_{x \in B(X)} T(C, x, n),$$

$$D(X, n) \in \{ C \in C_S^{\max} | C_{PD}(C, X, n) \geq C_{PD}(F, X, n), \forall F \in C_S^{\max} \}。 \quad (8)$$

当在频域内计算匹配误差时, CCF 是非常方便的。MSE 由于其数学上的连续性, 常用于理论分析。在实际中, 由于 CCF 和 MSE 有乘法运算, 在硬件上很难实现。文献 [8] 中的仿真结果表明, MAE 和 MSE 的效果类似, 由于硬件复杂度原因, MAE 得到了广泛应用。

PDC 是一个很简单的准则, 但是它的效果损失严重, 而且准则中要用到一个经验阈值, 这个阈值对于不同的序列可能有不同的最佳值, 因此现在已很少采用。有些文献提出对上述几个准则的修改方案以降低运算复杂度, 如 Reduced bit mean absolute error (RBMAE)^[10], Reduced bit pixel difference classification (RBPDC)^[11] 等等。文献 [12] 中提出一种加权平均绝对误差 (WMAE) 标准, 认为在大面积亮度变化缓慢的区域内, 由于各个匹配块之间的亮度差异比较小, 噪声对匹配结果的影响变得很大。这样, 在这个区域内得到的运动场中运动矢量的方向和大小呈随机分布。WMAE 标准降低了这种情况的出现概率。设待测位移矢量 $C = (C_x, C_y)$,

C_y , WMAE 的公式为

$$E_{WMA}(C, X, n) =$$

$$(1 + K(C_x^2 + C_y^2)) E_{MA}(C, X, n),$$

$$D(X, n) \in \{ C \in C_S^{\max} | E_{WMA}(C, X, n) \leq E_{WMA}(F, X, n), \forall F \in C_S^{\max} \}。 \quad (9)$$

式中权值 $K = 0.01 \sim 0.2$ 是一个经验常数。当运动信息比较明显的时候, 正确位移矢量的 E_{MA} 远小于其他不正确位移矢量所对应的 E_{MA} , 权值 K 不影响正确位移矢量的选择。当运动信息不明显, 即在大面积亮度变化缓慢的区域, 各个待测运动矢量的 E_{MA} 比较相当时, 权值 K 可使靠近的待测位移矢量被选中的可能性更大。这个标准可以使大面积亮度变化缓慢区域的运动矢量场更加有规则。为了进一步降低运算复杂度, 文献 [13] 提出了一种新的有效的块匹配检测标准最小最大误差 (MME):

$$E_{MM}(C, X, n) =$$

$$\max(|F(x, n) - F(x - C, n - p)|),$$

$$D(X, n) \in \{ C \in C_S^{\max} | E_{MM}(C, X, n) \leq E_{MM}(F, X, n), \forall F \in C_S^{\max} \}。 \quad (10)$$

MME 与传统的 MAE 标准相比, 主要优点在于 MME 对硬件的要求大大降低。例如, 对于一个块大小为 8×8 、每一个像素有 8 b 亮度信息的视频信号, MAE 和 MME 匹配的硬件结构见图 2。

MAE 需要 8×8 个 8 b 的绝对差值相加, 需要一个 16 b 的加法器。然而, MME 硬件仅需一个对 8 b 绝对差值进行比较的 8 b 比较器, 并使运动矢量选择的比较器从 16 b 降为 8 b。这样, MME 降低了运动估计器的硬件面积和复杂度。同时, 由于 8 b 比较器比 16 b 比较器和 16 b 加法器速度快, 提高了整个运动估计器硬件处理速度。

由于 MME 检测标准对每一对匹配块只取其最大的像素绝对差值, 所以, 它对噪声更加敏感。文献 [13] 建议选取仅次于最大的像素绝对差值作为匹配误差来降低噪声的影响, 但在噪声较大的场合仍无法令人满意, 而且在大面积亮度变化缓慢区域会预测出更加不规则的运动场。笔者提出一种加权的

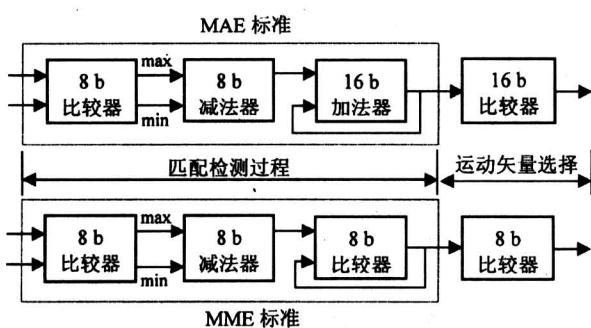


图2 块匹配检测标准 MAE 与 MME 的硬件结构图

Fig.2 The hardware architectures of MAE and MME

MME (WMME) 标准。WMME 与 WMAE 类似, 根据待测运动矢量 C 的大小给 MME 增加权值:

$$\begin{aligned} E_{w2MM}(C, X, n) = \\ (1 + K(C_x^2 + C_y^2)) E_{MM}(C, X, n), \\ D(X, n) \in \{C \in C_S^{\max} | E_{w2MM}(C, X, n) \leqslant \\ E_{w2MM}(F, X, n), \forall F \in C_S^{\max}\}. \quad (11) \end{aligned}$$

考虑到平方运算的硬件难实现性, 用绝对值代替平方运算得到:

$$\begin{aligned} E_{WMM}(C, X, n) = \\ (1 + K(|C_x| + |C_y|)) E_{MM}(C, X, n), \\ D(X, n) \in \{C \in C_S^{\max} | E_{WMM}(C, X, n) \leqslant \\ E_{WMM}(F, X, n), \forall F \in C_S^{\max}\}. \quad (12) \end{aligned}$$

模拟结果表明, WMME 式 (12) 中的常数 K 应该适当增大才能达到与 w2MME 相类似的结果。WMME 与 WMAE 和 w2MME 相比降低了运算复杂度, 它与 MME 相比增加了抗噪声的能力。

3 试验结果

试验结果通常用主观评价标准评定, 也有采用客观评价的。

3.1 试验结果主观评价

主观评价主要根据视觉效果对图像作出判断。

图3是MSE和MAE标准的效果对比。图3a和图3b是原始序列中的两帧相邻逐行图像(大小为 352×240); 图3c和图3d是用全搜索块匹配方法采用不同的匹配标准得到的运动场, 全搜索块匹配算法采用了 8×8 的匹配块, 运动检测窗口为 14×14 ; 图3e和图3f是由第一帧原始图像通过检测到的运动场进行块移位得到的第二场图像。由运动

场看出, 两者所检测到运动基本上是一致的, 从图3e和图3f的主观效果图同样可以得出相同的结论。由图3a和图3b可见, 在背景蓝天的白云部分运动场中, 运动矢量的方向和大小呈随机分布。

图4是用WMAE检测得到的运动结果, 从图3d和图4a对比可以看出, 图中白云处的错误运动矢量几乎已经消失。

图5是MME, w2MME和WMME的试验结果主观质量比较。图5a是由MME得到的矢量场, 对比图3d可以看出, MME受噪声的影响大, 图像中有许多较大的方向各异的运动矢量是不正确的。采用w2MME后, 运动场图5c中的错误运动矢量明显减少。但是与图4a比较, 采取相同的 $K=0.01$ 值, 仍有部分错误的运动矢量, 这与理论分析相符合。采用相同的 K 值, WMME的效果比w2MME的效果差, 但WMME中适当提高 $K=0.2$ 值后, 得到的运动场图5e要比图5c的效果好。对比图5e和图4a可以看出, 两个矢量场很相似, 由此可以得出结论, WMME比WMAE降低了运算复杂度, 而其效果是类似的。所以, 在硬件实现时选择了WMME匹配检测标准。

3.2 试验结果客观评价

以上给出的实验结果都是用主观评价标准评价的。通常在文献中, 也给出一些客观的评价标准, 如峰值信号噪声比(PSNR), 熵(entropy)等。PSNR的定义为^[14]:

$$R_{PSN} = 10 \lg \left\{ \frac{255^2 MN}{\sum_x [F(x, n) - F_o(x, n)]^2} \right\}$$

这里, MN 是帧的大小, $F(x, n)$ 是生成的帧, $F_o(x, n)$ 是用作参考标准的原始帧。图像中的边界在视觉上是非常重要的, 但占图像的很小部分, 其重要性并未在PSNR中反映出来。由于没有更可靠的视觉质量评价标准, PSNR对于评价图像的视觉质量虽然不是一个很好的选择, 但还是得到了很广泛的应用。

全搜索块匹配运动估计方法按六种匹配准则得到的运动场信息生成第二帧图像的客观评价, 即图3e、图3f、图4b、图5b、图5d、图5f中各图片的PSNR见表1。由表1可以看出, PSNR可以部分地反映图像的主观效果, 但也有偏差。MSE具有最大的 R_{PSN} 。MAE和MSE的 R_{PSN} 类似。MME和MSE、MAE相比具有最小的 R_{PSN} 。这些与主观效果相一致, 是在算法效果和运算的复杂度

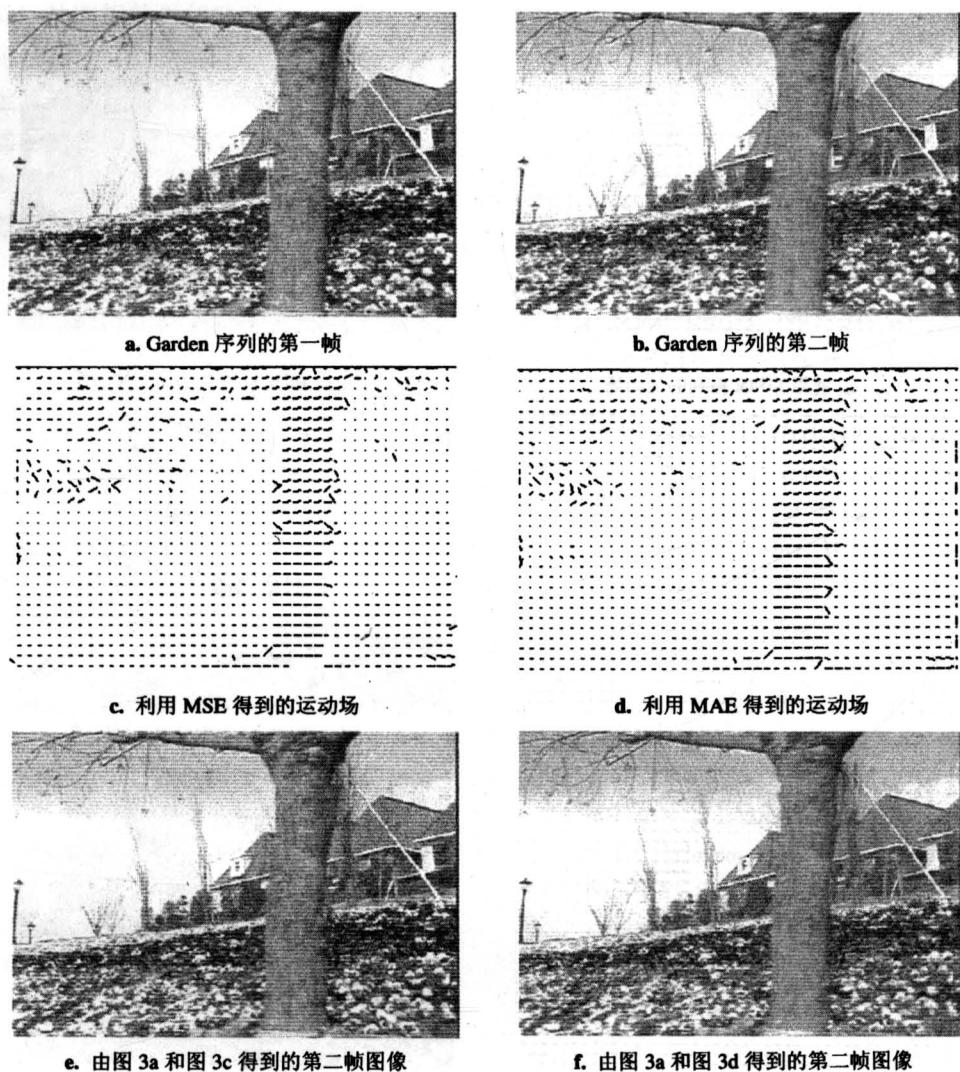


图 3 利用 MSE 和 MAE 的运动估计结果比较

Fig.3 The performance of motion estimation with MSE or MAE

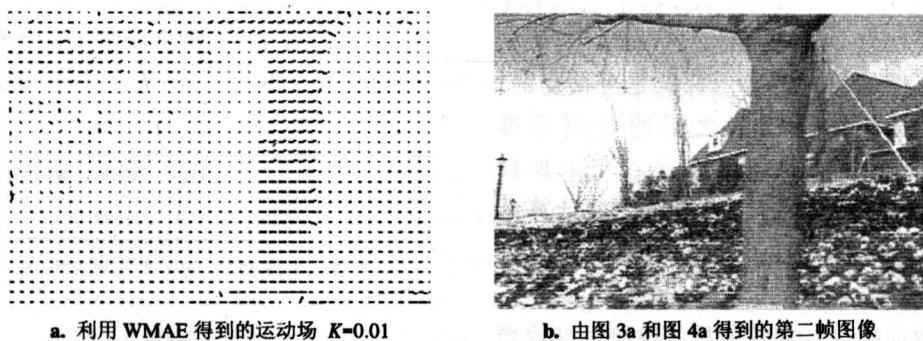


图 4 利用 WMAE 检测到的运动估计结果

Fig.4 The performance of motion estimation with WMAE

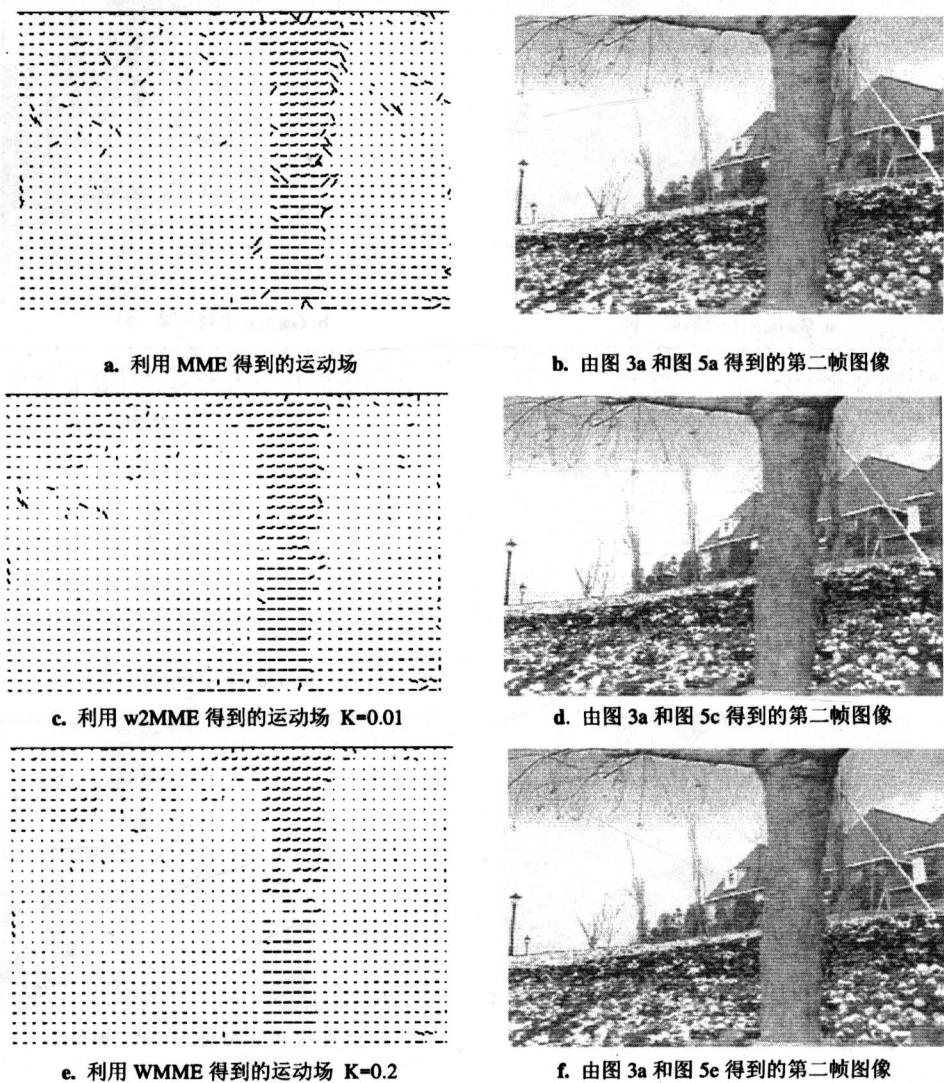


图 5 利用 MME、w2MME 和 WMME 得到的运动估计结果比较

Fig.5 The performance of motion estimation with MME, w2MME or WMME

之间折衷，是算法复杂度降低的必然结果。为降低算法的复杂度，图像质量的这点降低是可接受的。从表 1 可以看出，加了权值之后的检测标准 (WMAE, w2MME, WMME) 的 R_{PSN} 与未加权值时相比都有所降低，这是因为 PSNR 的计算结果是和 MSE 相一致的。完全按照 MSE 检测出来的最佳运动矢量能导致最大的 R_{PSN} 值，但大 R_{PSN} 值不能说明块效应减少。加权值后的检测标准检测的最佳运动矢量会导致很好的视觉效果，但不一定能够产生大的 R_{PSN} 值。另外，从表 1 中还可以看出，降低检测标准运算的复杂度对检测结果的负面影响是可接受的。

表 1 六种匹配准则的比较*

Table 1 The comparison of six block-matching criteria

	MSE	MAE	MME	WMAE	w2MME	WMME
R_{PSN}/dB	24.616	24.458	22.591	24.198	23.474	22.077

* 图像大小为 352×240 ；匹配块为 8×8 ；搜索窗为 14×14

4 结论

从匹配检测的效果和硬件实现的难易考虑，针对三维递归搜索块匹配运动估计算法的固有缺点，提出了一种新的块匹配检测标准 WMME——加权最小最大误差。这个匹配标准与其他匹配标准比

较, 能够在类似的匹配效果的情况下比较容易地用硬件实现, 而且能够降低运动估计器芯片的面积, 提高芯片的运算速度。由实验结果可见, 这个标准对运动估计器的输出结果影响不大, 而且能降低噪声影响。在要求进行运动估计的硬件系统中, 加权最小最大误差标准适合与三维递归搜索策略一起应用。

参考文献

- [1] Buzug T M. Foreword to the special issue on motion detection and compensation [J]. Philips Journal of Research, 1998, 51(2): 197~201
- [2] de Haan G. Progress in motion estimation for consumer video format conversion [J]. IEEE Transactions on Consumer Electronics, 2000, 46(3): 449~459
- [3] Haan G D, Biezen P W A C, Huijgen H, et al. True-motion estimation with 3-D recursive search block matching [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1993, 3(5): 368~379
- [4] de Haan G. Progress in motion estimation for consumer video format conversion [J]. IEEE Transactions on Consumer Electronics, 2000, 46(3): 449~459
- [5] Chen M J, Chen L G, Chiueh T D, et al. A new block-matching criterion for motion estimation and its implementation [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1995, 5(3): 231~236
- [6] Tekalp A M. Digital video processing [M]. Prentice Hall, 1996
- [7] Kappagantula S, Rao K T. Motion compensated predictive coding [M]. SPIE 27th Proc 432 1985. G4~70
- [8] Dufaux F. Multigrain blocks matching motion estimation for generic video coding [D]. Ph D dissertation, Swiss Fed Institute of Tech, Lausanne, Switzerland, 1994
- [9] Gharavi H, Mills M. Block-matching motion estimation algorithms: new results [J]. IEEE Transactions on Circuits and Systems, 1990, 37: 649~651
- [10] Baek Y, Oh H S, Lee H K. Block-matching criterion for efficient VLSI implementation of motion estimation [J]. Electronic Letters, 1996, 32(13): 1184~1185
- [11] Rath G B, Makur A. A fast matching criterion for VLSI implementation of block-based motion estimation [J]. Signal Processing, 1999, 73: 297~301
- [12] Castagno R, Haavisto P, Ramponi G, et al. A method for motion adaptive frame rate up-conversion [J]. IEEE Trans Circuits and Systems for Video Technology, 1996, 6(5): 436~446
- [13] Chen M J, Chen L G, Chiueh T D, et al. A new block-matching criterion for motion estimation and its implementation [J]. IEEE Trans Circuits and Systems for Video Technology, 1995, 5(3): 231~236
- [14] Nam J Y, Seo J S, Kwak J S, et al. New fast-search algorithm for block matching motion estimation using temporal and spatial correlation of motion vector [J]. IEEE Trans Consumer Electronics, 2000, 46(4): 934~942

A Weighted Block-matching Criterion for the Hardware Implementation of Motion Estimators

Zhang Xia, Zheng Nanning, Zhang Guanglie, Wu Yong, Wang Shaorui, Xu Weipu
(Xi'an Jiaotong University, Xi'an 710049, China)

[Abstract] In all kinds of digital video processing algorithms, motion compensated algorithm can acquire perfect performance because the motion information in the video signal has been considered. The hardware implementation of the motion estimator is the core of the various motion compensated digital video processings, which will be applied in real systems. Block-matching motion estimating algorithm is widely used in real system because of its low computing complication, easy realization and high call frequency of the block-matching criterion in the hardware systems. A new criterion called weighted minimized maximum error is proposed in this paper. This criterion can reduce the complexity of the motion estimator, decrease the area of the hardware and increase the speed of the hardware. On the other hand, the criterion is suitable to be applied to the recursive searching strategy which has an inherent weakness called error propagation.

[Key words] video processing; motion compensation; motion estimation; block-matching criterion