

一种运动自适应的帧内插算法¹

骆立俊 邹采荣 何振亚 陆 阳*

(东南大学无线电系 南京 210096)

*(合肥工业大学计算机与信息科学系 合肥 230009)

摘要 本文提出了一种运动自适应的帧内插算法,可以用于逐行扫描的数字视频系统中做帧速率转换。该算法首先采用了线性预测块匹配搜索算法来估计内插帧中每一个象素的运动矢量;然后对得到的运动场进行中值滤波和判别,剔除运动不一致或错误的运动矢量;最后沿着运动轨迹线性内插得到内插帧的象素值。仿真分析表明本文提出的算法,在一个大的搜索区域进行快速搜索和判别,可以有效地解决内插帧中快速运动目标的运动模糊和有切换帧时的图象重叠,提高了帧速率转换中内插帧的主观视觉质量。

关键词 运动估计, 中值滤波, 数字视频系统

中图分类号 TN941, TP391

1 引 言

随着数字电视和多媒体信息系统的发展,国际上出现了大量有不同帧速率的视频信号节目源,不同帧速率的节目交流需要做帧速率转换。与此同时,帧速率变换还可以使运动图象的显示格式与人眼视觉特性相匹配,从而有效地改善图象的主观质量,消除爬行、行抖动和亮度闪烁等现象^[1]。帧速率转换分帧速率上转换和帧速率下转换两种,本文主要讨论逐行扫描图象系列的帧速率上转换。帧速率上转换需要做帧的运动补偿和内插,内插帧的运动信息可以从图象系列中得到。最简单的帧内插方法是重复上一帧,这种方法不需要运动信息,但会产生运动目标的跳动运动,使运动不连续^[2]。简单地利用前后帧的线性内插来得到内插帧,会产生运动目标的运动模糊^[3]。

文献[4]提出了一种运动自适应的帧速率上转换(FRU)算法,它的基本思想是以当前象素的左邻近点的运动矢量为中心点,在 3×3 的区域内进行全搜索得到当前象素的运动矢量,然后对得到的运动场进行中值滤波,剔除运动不一致的运动矢量,最后沿着运动轨迹线性内插得到内插帧的象素值。分析表明,这种小范围的块匹配搜索算法对于只有低速运动目标的运动图象系列来说是简单有效的,对于有快速运动目标和有场景切换的运动图象系列,在内插帧中会产生快速运动目标的运动模糊和有切换帧时的图象重叠,两种情况分别如图1(a),图1(c)和图1(e),以及图1(b),图1(d)和图1(f)所示。本文将象素运动矢量的搜索区域扩大为 32×32 的区域,并对有无切换帧进行判别分析,提出了一种运动自适应的帧内插算法。该算法采用线性预测块匹配搜索算法来得到内插帧中每一个象素的运动矢量,然后对得到的运动场进行中值滤波和判别,剔除运动不一致或错误的运动矢量,最后沿着运动轨迹线性内插得到内插帧的象素值。本文详细地给出了线性预测块匹配搜索算法,运动自适应的帧内插算法。仿真分析中对运动自适应的帧内插算法的性能进行了讨论分析。

¹ 1997-01-15 收到, 1997-09-11 定稿

国家攀登计划(神经网络)重大项目和东南大学与非立浦公司消费电子领域合作项目资助

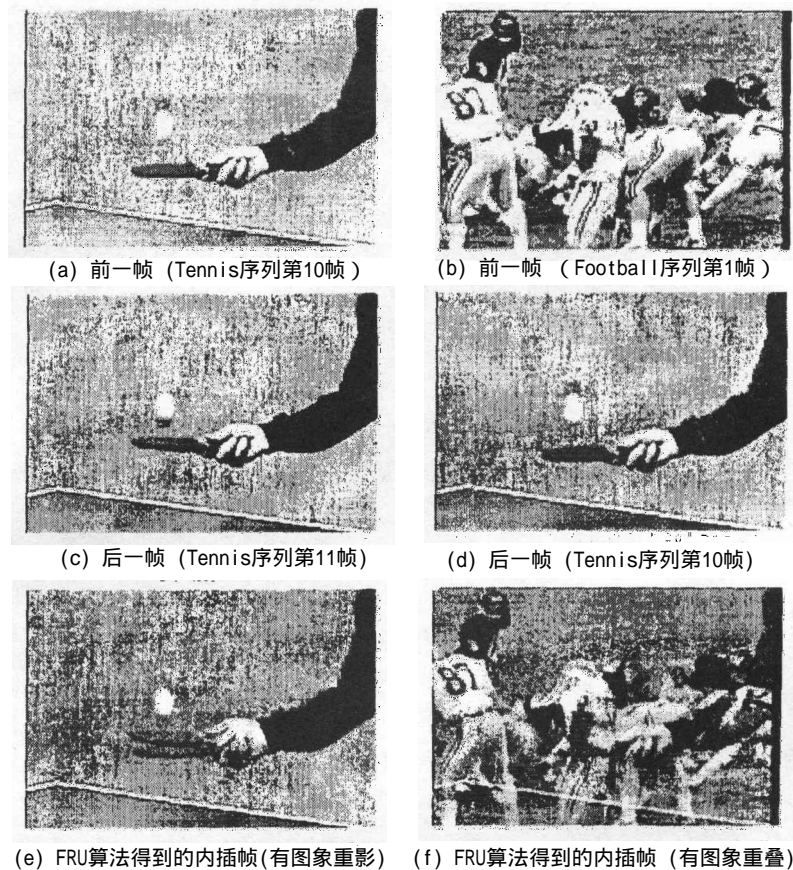


图1 FRU 算法在有快速运动目标和有场景切换时的内插帧

2 线性预测块匹配搜索算法

在帧速率变换中, 计算复杂度主要取决于运动估计和运动补偿。内插帧的运动信息来自它的前后两帧, 它的运动场可以用块匹配方法进行运动估计。目前运动估计算法有很多种, 全搜索运动估计 (FS) 算法^[5] 由于精度高, 硬件易实现, 是应用最广泛的运动估计算法, 但它的计算量大, 特别是在大范围搜索时, 加大了硬件实现难度。为了减少 FS 算法的计算量, 出现了许多改进的运动估计算法, 如三步搜索法 (3SS)^[6], 共轭方向搜索法 (CDS)^[7], 二维对数搜索法 (LOGS)^[8], 交叉搜索法 (CS)^[9], 动态搜索窗调整搜索法 (DSWDS)^[10], 改进的三步搜索法 (N3SS)^[11], 四步搜索法 (4SS)^[12] 等。

在帧速率变换中, 匹配块的大小和象素分布有四种形式^[4], 如图 2 所示。匹配块的尺寸太小或太大, 都不能准确地预测象素的运动矢量, 故本文分析采用图 2(d) 给出的匹配块。内插帧的运动估计以匹配块为单位, 在参考帧 (下一帧) 中以相对应的匹配块为中心的搜索区域内进行搜索, 寻找最佳匹配块, 得到相应的运动矢量。运动估计只对亮度信号进行, 色度信号借用亮度信号的估计结果。

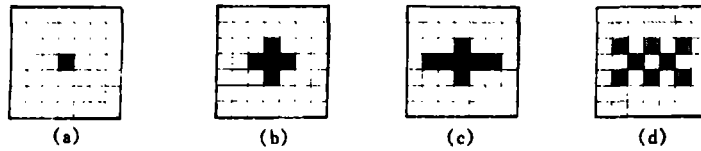


图 2 四种形式匹配块的大小和象素分布

衡量运动估计最佳匹配的准则有很多种,如均方误差 (MSE)、归一化互相关函数 (NCCF)、平均绝对误差 (MAD) 等,分析表明各种匹配准则的性能差别不显著,比较而言 MAD 的运算量最小,便于硬件实现,所以用得最多,在此定义为

$$\text{MAD}(i, j) = \frac{1}{8} \sum_m \sum_n |S_{k-1}(m-i, n-j) - S_{k+1}(m+i, n+j)|, \quad (1)$$

其中 i, j 分别为匹配块水平和垂直方向的偏移量,在搜索区域内取值; $S_k(m, n)$ 为第 k 帧位于 (m, n) 的象素值。

在视频序列图象中,物体运动是连续的,所以采用匹配块来描述的象素运动矢量在空间和时间上具有相关性。图 3 是两个图象序列采用 FS 算法做运动估计后运动位移分布直方图,其中搜索区域为 $(\pm 16, \pm 16)$ 。从图 3(a) 和图 3(c) 中可以看出,通常视频序列图象中物体运动范围和运动速度都不大,所以运动矢量在搜索区域内不是均匀分布的,图象序列 "Miss America" 和 "Table Tennis" 分别有 80% 和 77% 的运动位移分布在 5×5 的中心区域窗内,分别有 50% 和 59% 的运动位移分布在 3×3 的中心区域窗内。与此同时,内插帧的运动场在空间和时间上具有相关性,当前象素的运动矢量,在大多数情况下与内插帧中相邻象素的运动矢量有着相近的大小和方向。

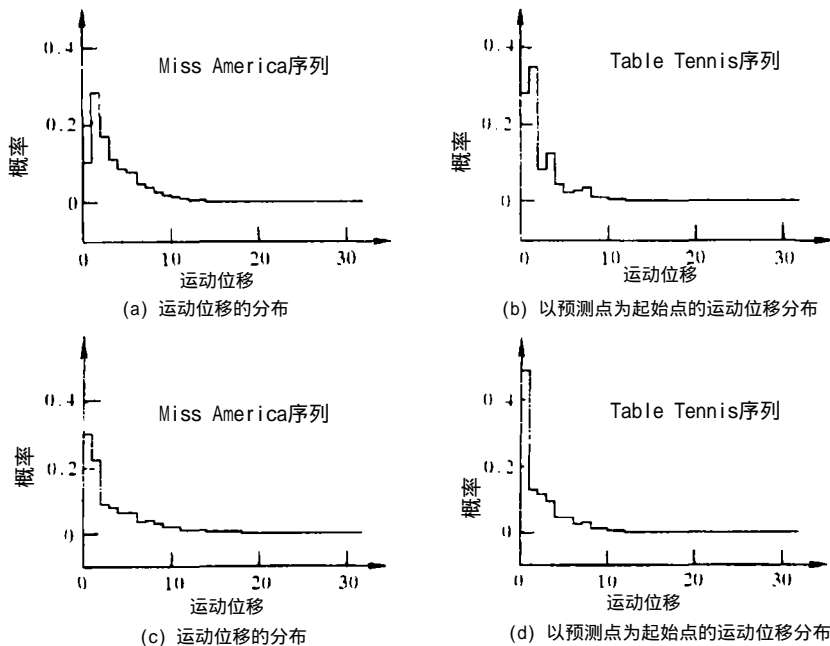


图 3 两个图象序列的运动位移分布直方图

图 4 是线性预测块匹配搜索 PSA 算法中用以预测的像素的分布图。我们用三个邻近像素 1, 2, 3 运动矢量的线性加权值来作为预测矢量, 可以得到图象序列 “Miss America” 和 “Table Tennis” 中像素的预测矢量和实际的运动矢量之间距离的分布直方图分别如图 3b 和图 3d 所示, 两个图象序列分别有 81% 和 80% 的运动位移分布在以预测点为中心的 3×3 区域窗内, 可以看出序列图象中像素的实际运动矢量与预测矢量之间的距离具有很强的空间分布特性——中心偏置分布特性。

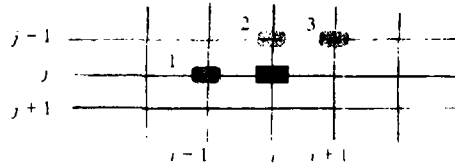


图 4 线性预测块匹配搜索算法中用以预测的像素的分布

PSA 算法首先用内插帧内已知的相邻像素运动矢量来预测当前像素运动矢量的搜索起始值, 然后以预测点为起始点, 采用 3×3 的搜索窗进行搜索步长为 1 的移动窗搜索, 直到搜索到达搜索域的边界或搜索的局部最小点位于搜索窗的中心时停止。令 $V_k = [V_{xk}, V_{yk}]^T$ 表示像素 $k=1, 2, 3$ 所对应的相邻像素的运动矢量, $P = [P_x, P_y]^T$ 表示当前像素运动矢量的预测值, $W = [W_1, W_2, W_3]^T$ 表示权值矢量, 当前像素位于 (i_0, j_0) , 预测点为 (i_1, j_1) , 其中

$$W = [W_1, W_2, W_3]^T = [1/3, 1/3, 1/3]^T,$$

$$P = [P_x, P_y]^T = \left[\sum_{k=1}^3 W_k V_{xk}, \sum_{k=1}^3 W_k V_{yk} \right]^T,$$

$$i_1 = i_0 + P_x, j_1 = j_0 + P_y. \quad (2)$$

PSA 预测搜索算法的具体步骤如下:

步骤 1 依据 (2) 式计算运动矢量搜索初始值 (i_1, j_1) , 取搜索步长为 1, 计算以预测的初始点 (i_1, j_1) 为中心的 3×3 搜索窗上的 9 个搜索点所各自对应的 MAD。以当前 MAD 最小的点作为下一次搜索窗的中心点, 如果当前搜索窗的中心点为 MAD 最小的点, 中止搜索;

步骤 2 取搜索步长为 1, 计算以新的搜索中心点为中心 3×3 搜索窗上的 9 个搜索点所各自对应的 MAD, 然后以 MAD 最小的点作为下一次搜索窗的中心, 如果当前搜索窗的中心点为 MAD 最小的点, 中止搜索。在搜索过程中如果当前 MAD 最小的点位于搜索窗的四个角, 下一次的搜索只需要增加 5 个点, 否则的话下一次的搜索只需要增加 3 个点。步骤 2 重复直到搜索窗的中心到达区域 $(\pm 16, \pm 16)$ 的边界为止。

3 运动自适应的帧内插算法

运动自适应的帧内插算法分为四部分: 场景切换的判别, 线性预测块匹配搜索得到内插帧中像素的运动矢量, 运动场的中值滤波, 沿着运动轨迹线性内插得到内插帧的像素值。

帧内插算法采用了线性预测块匹配搜索算法来得到内插帧中像素的运动矢量, 然后对得到的运动矢量场进行中值滤波, 剔除运动不一致或错误的运动矢量。令 $U_k = [u_{xk}, u_{yk}]^T$,

$k = 1, \dots, 9$, 分别表示以内插帧像素 (x, y) 为中心的 3×3 窗内总共 9 个像素的运动矢量。采用中值滤波可以得到像素 (x, y) 的最终的运动矢量 $V = [u_x, v_y]^T$ 为

$$\begin{aligned} V_x &= \text{median}[u_{x1}, \dots, u_{x9}], \\ V_y &= \text{median}[u_{y1}, \dots, u_{y9}]. \end{aligned} \quad (3)$$

帧内插算法在沿着运动轨迹线性内插计算内插帧的像素值时分以下两种情况:

(1) 定义图象序列前一帧 (即第 $k-1$ 帧) 与后一帧 (即第 $k+1$ 帧) 运动自适应的帧间均方差为

$$\text{MSE}_{k-1} = \frac{1}{N_1 N_2} \sum_{x=1}^{N_1} \sum_{y=1}^{N_2} \min_{i,j} \{ \text{MAE}(i, j, x, y), \quad -16 \leq i \leq 16, \quad -16 \leq j \leq 16 \}, \quad (4)$$

其中

$$\text{MAE}(i, j, x, y) = \sum_{m=x}^{x+15} \sum_{n=y}^{y+15} |S_{k-1}(m, n) - S_{k+1}(m+i, n+j)|, \quad (5)$$

如果 $\text{MSE}_{k-1} > 15 \cdot \text{MSE}_{k-2}$, 则判定前一帧与后一帧间存在场景切换, 取内插帧的像素值 $S_k(x, y)$ 为

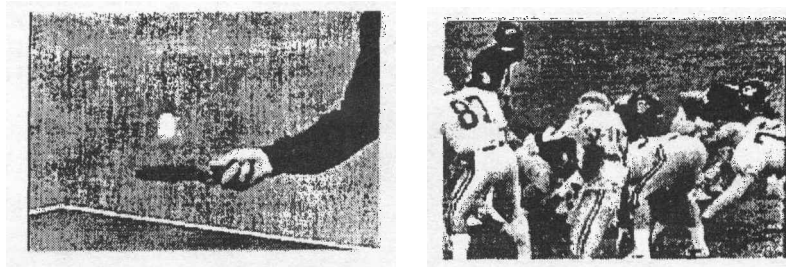
$$S_k(x, y) = S_{k-1}(x, y). \quad (6)$$

(2) 如果前一帧与后一帧之间无场景切换, 线性预测块匹配搜索得到内插帧中像素的运动矢量 $V = [i, j]^T$, 取内插帧的像素值 $S_k(x, y)$ 为

$$S_k(x, y) = \frac{1}{2} [S_{k-1}(x-i, y-j) + S_{k+1}(x+i, y+j)]. \quad (7)$$

4 仿真结果

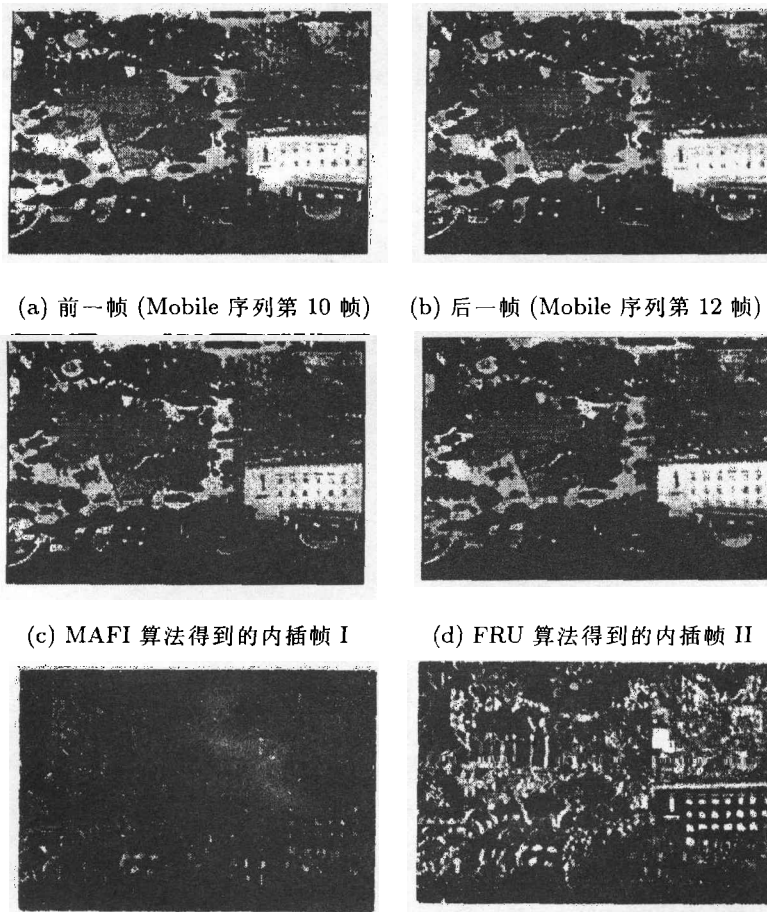
运动自适应的帧内插 (MAFI) 算法的计算机仿真实验采用了图 2(d) 所示的匹配块, 搜索域为 32×32 , 运动估计的匹配准则选取为平均绝对误差 MAD, 内插帧内像素的运动矢量估计采用线性预测块匹配搜索 PSA 算法, 场景切换则采用运动自适应的帧间均方差进行判别。MAFI 算法在图 1 给出的有快速运动目标和有场景切换两种情况下的内插帧结果分别如图 5(a) 和图 5(b) 所示。仿真分析表明本文提出的算法, 在一个大的搜索区域进行快速搜索和判别, 可以有效地解决内插帧中快速运动目标的运动模糊和有场景切换时的图象重叠, 提高了帧速率转换中内插帧的主观视觉质量。为了进一步对 MAFI 算法与文献 4 给出的 FRU 算法进行比较, 我们采用 Mobile 序列的第 10 帧和第 12 帧作为内插的前后帧 (分别见图 6(a) 和图 6(b)), MAFI 算法和 FRU 算法内插得到的内插帧 I 和内插帧 II 分别如图 6(c) 和图 6(d) 所示, 内插帧 I 和内插帧 II 与 Mobile 序列的第 11 帧的帧差图象分别见图 6(e) 和图 6(f), 在图 6(d) 中, 日历上的数字和旋转的彩球有比较明显的运动模糊。从图 5 和图 6 中可以看出, 本文提出的 MAFI 算法优于文献 4 提出的 FRU 算法, 但由于 MAFI 算法假定运动模型是线性常速度的, 没有考虑到运动目标的加速度, 所以 MAFI 算法对变速度的运动目标的运动场估计仍存在着较大的误差, 这可以从图 6(e) 的帧差图象中观察到。



(a) 内插帧 (无明显图像重影) (b) 内插帧 (无图像重叠)

图 5 MAFI 算法在有快速运动目标和有切换帧时的内插帧

本文提出的 MAFI 算法的计算复杂性最主要来自于线性预测块匹配搜索的搜索次数, 几种不同的运动估计搜索算法的搜索次数的比较如表 1 所示。从表 1 可以看出, MAFI 算法与文献 4 的 FRU 算法相比, 计算量有一些增加。



(e) 内插帧 I 与 Mobile 序列第 11 帧的帧差图象 (f) 内插帧 II 与 Mobile 序列第 11 帧的帧差图象

图 6 MAFI 算法与 FRU 算法的比较结果

表 1 几种不同的运动估计搜索算法的平均搜索次数的比较

序列图象	FS	3SS	4SS	FRU	MAFI
Miss America	1024	33	19.52	9	12.01
Claire	1024	33	18.30	9	11.07
Flower Garden	1024	33	21.32	9	14.57
Table Tennis	1024	33	19.45	9	11.98
Football	1024	33	18.56	9	11.36
Mobile & Calendar	1024	33	21.35	9	13.87
平均	1024	33	19.75	9	12.48

5 结 论

本文提出了一种 MAFI 算法, 算法首先采用了线性预测块匹配搜索算法来得到内插帧中象素的运动矢量, 然后对得到的运动场进行中值滤波和判别, 剔除运动不一致或错误的运动矢量, 最后沿着运动轨迹线性内插得到内插帧的象素值. 算法还对有无场景切换进行了准确的判别, 仿真分析表明本文提出的算法, 在一个大的搜索区域进行快速搜索, 可以有效地解决内插帧中快速运动目标的运动模糊和有场景切换时的图象重叠, 提高了帧速率转换中内插帧的主观视觉质量. MAFI 算法与文献 4 的 FRU 算法相比, 计算量有一些增加.

参 考 文 献

- [1] Hentchel C. Linear and nonlinear procedures for flicker reduction. *IEEE Trans. on Consumer Electronics*, 1992, CE-38(4): 151-156.
- [2] Csillag Peter, Boroczky Lilla. Motion-compensated frame rate conversion using an accelerated motion model. *Proceeding of ICASSP'96, Altalanda, USA: 1996, vol. 4, 2319-2322.*
- [3] Schamel G. Pre- and post-filtering of HDTV signals for sampling rate reduction and display up-conversion. *IEEE Trans. on CAS*, 1987, CAS-34(11): 1432-1439.
- [4] Castagno Roberto, Haavisto Petri, Giovanni Ramponi. A method for motion adaptive frame rate up-conversion. *IEEE Trans. on CAS VT*, 1996, 6(5): 436-446.
- [5] ISO/IEC 13818-2 Coding of moving picture and associated audio, 1995.
- [6] Koga T, Iinuma K, *et al.* Motion-compensated inter-frame coding for video conferencing. in *Proc. NTC81, New Orleans, LA: Nov. 1981, C9.6.1-C9.6.5.*
- [7] Kappagantula S, Rao K R. Motion compensated inter-frame image prediction. *IEEE Trans. on Commun.*, 1985, COM-33(9): 1011-1015.
- [8] Jain J R, Jain A K. Displacement measurement and its application in inter-frame image coding. *IEEE Trans. on Commun.*, 1981, COM-29: 1799-1808.
- [9] Ghanbari M. The cross-search algorithm for motion estimation. *IEEE Trans. on Commun.*, 1990, 38(7): 950-953.
- [10] Lee L W, Wang J F, Lee J Y, and Shie J D. Dynamic search-window adjustment and interlaced search for block-matching algorithm. *IEEE Trans. on CAS VT*, 1993, 3(2): 85-87.
- [11] Li R, Zeng B, Liou M. A new three-step search algorithm for block motion estimation. *IEEE Trans. on CAS VT*, 1994, 4(4): 438-442.

- [12] Po L M, Ma W C. A novel four-step algorithm for fast block motion estimation. IEEE Trans. on CAS VT, 1996, 6(3): 67-71.

AN ALGORITHM FOR MOTION ADAPTIVE FRAME INTERPOLATION

Luo Lijun Zou Cairong He Zhenya Lu Yang*

(*Department of Radio Engineering, Southeast University, Nanjing 210096*)

*(*Department of Computer and Information Science, Hefei University of Technology, Hefei 230009*)

Abstract A motion adaptive frame interpolation algorithm for progressive frame rate conversion of digital video systems is proposed. Firstly a motion vector is searched for each pixel in the interpolated frame using the block-matching search method of linear prediction. Then the resulting motion field is median filtered and differentiated to remove inconsistent or wrong vectors. Finally averaging along the motion trajectory is used to get the interpolated pixel values. Simulation results indicate the proposed method with fast search and differentiation in a large area can solve the motion blurring of fast moving object and the scene overlapping of a scene change in the interpolated frame, and the corresponding objective quality of the interpolated frame is also improved.

Key words Motion estimation, Median filtering, Digital video system

骆立俊: 男, 1970 年生, 博士生, 研究方向为视频信号处理和神经网络等.

邹采荣: 男, 1962 年生, 教授, 研究方向为视频信号处理, 多维信号处理, 以及神经网络等.

何振亚: 男, 1923 年生, 教授, 博士生导师, 研究方向为神经网络, 自适应信号处理, 以及数字图象处理等.

陆 阳: 男, 1966 年生, 工程师, 研究方向为数字图象处理, 计算机工业控制, 数字信号处理, 以及神经网络等.