

一种适用于遮挡问题的运动补偿帧插值方法¹

赵建伟 王朋 刘重庆

(上海交通大学图像处理与模式识别研究所 上海 200030)

摘要: 针对二维运动估计中常出现的遮挡问题, 提出一种基于网格模型的运动补偿自适应帧插值技术, 将图像帧中分为 3 种不同区域, 根据各自的特点分别进行运动估计和运动补偿插值. 在准确检测遮挡区域前提下, 单独对遮挡区域处理有效地减小了运动估计误差. 提出采用特征窗口运动补偿匹配的方法进行网格节点运动矢量估计, 消除了传统块匹配方法导致的块效应, 得到了亚像素的匹配精度. 实验证明, 该方法简单易行, 有效地解决了遮挡区域的运动估计问题, 插值图像具有较好的视觉效果.

关键词: 遮挡问题, 网格模型, 运动估计, 运动补偿帧插值

中图分类号: TN919.8 **文献标识码:** A **文章编号:** 1009-5896(2004)05-0771-06

Motion Compensated Frame Interpolation Method Applied for Occlusion Issue

Zhao Jian-wei Wang Peng Liu Chong-qing

(Inst. of Image Proc. & Pattern Recog., Shanghai Jiaotong Univ., Shanghai 200030, China)

Abstract A motion compensated adaptive frame interpolation technique based on mesh model is proposed considering occlusion issue existed in two-dimensional motion estimation. The whole image is divided into three types of regions in order to perform motion estimation and motion compensated frame interpolation respectively. A more accurate occlusion detection method is presented to improve the precision of detection, which is the main premise of the adaptive interpolation. The motion estimation error introduced by occlusion regions is reduced effectively. Mesh node motion estimation based on feature window matching is used to achieve sub-pixel resolution and overcome block artifacts produced by block matching. Experimental results indicate that the proposed algorithm is easy to implement, and solves motion estimation problem existed in occlusion regions effectively. The interpolated image shows better visual performance.

Key words Occlusion issue, Mesh model, Motion estimation, Motion compensated frame interpolation

1 引言

运动补偿帧插值 (MCFI) 是沿运动轨迹方向进行插值的一项重要视频处理技术, 主要应用于甚低码率编码^[1]和帧率转换^[2]. 运动估计技术是运动补偿的核心部分, 目的是通过获取图像帧中各像素点沿运动轨迹的运动矢量得到整个二维运动矢量场. 块匹配法由于其容易实现, 计算简洁等优点广泛应用于序列图像运动估计中, 尤其是 H.263 以及 MPEG 压缩编码国际标准. 但是块匹配法仍存在一些问題, 如运动场不连续, 存在块效应, 不能处理旋转、缩放等复杂运动. 基于网格模型的运动估计方法^[3-8]较好地解决了上述问题, 与块匹配法相比, 其优点是能够更真实地表示单个物体的物理运动, 得到光滑的运动矢量场, 有效克服块匹配法中出现的块效应问题, 而且由于采用更复杂的空间变换, 网格模型可以有效地处理旋转、缩放等复杂运动. 但值得注意的是, 所有基于网格模型的方法都增强了帧内各部分的联接关系, 这相当于给

¹ 2002-08-26 收到, 2002-11-28 改回

二维运动矢量场施加了全局平滑约束,而这种约束关系尤其不适用于遮挡(Occlusion)区域进行准确的运动估计^[7]。遮挡区域的运动估计不准是导致整个运动矢量场估计不准的主要因素。

本文针对二维运动估计中常出现的遮挡问题,提出一种基于网格模型的运动补偿自适应帧插值技术。通过像素分类将图像帧中分为3种不同区域,根据各自的特点进行运动估计和运动补偿自适应插值,在一定程度上减小了遮挡区域带来的运动估计误差。同时,本文提出的准确检测遮挡区域的方法是准确对各区域进行自适应插值的重要前提,节点运动估计过程是基于网格模型的运动估计技术的关键,实验证明算法简单易行,在一定程度上解决了遮挡问题,能获得较好的视觉效果。

2 遮挡问题

遮挡问题是运动估计中较难解决的问题,即在前后两帧图像中,由于目标的运动造成背景的覆盖/暴露(Covered/Uncovered),而覆盖/暴露区域的像素点在前后两帧图像中不存在对应关系,导致无法估计到准确的运动。通常,遮挡区域(图1)可大致分为两种:即将被覆盖的背景区域(Background To Be Covered, BTBC)和即将暴露的背景区域(Uncovered Background, UB),主要是由于物体的全局运动或局部运动而产生在物体的边界或物体内部,又称为运动估计失效的区域。由于基于网格模型的运动表示方法增强了帧内各部分的联接性,所以不适用于遮挡区域进行准确的运动估计;运动估计不准确,就会很大程度上影响帧插值的视觉效果。所以,本文通过准确检测遮挡区域,并对其单独进行运动估计来完成遮挡区域的帧插值过程。

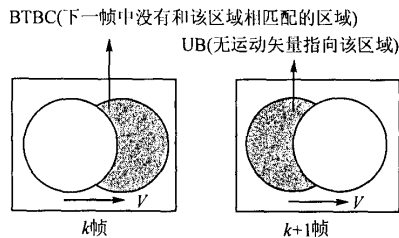


图 1 遮挡问题示意图

3 运动补偿帧插值

3.1 像素分类

通过像素分类可将整个图像帧分为静止区域,遮挡区域(BTBC和UB区域)和运动区域。

3.1.1 运动变化区域检测 通过计算 k 帧和 $k+1$ 帧的帧差(),得到变化检测模板(CDM),确定两帧运动剧烈的区域:若FD大于预先设定的阈值,则认为该像素点属于运动变化区域。利用中值滤波和形态学(开、闭运算)去除孤立的像素点,得到封闭的连通区域,最后设置阈值去除小面积的像素集合。由此可将图像帧分为运动变化区域和静止区域。需要指出的是,遮挡区域存在于两帧的运动变化区域,所以要在CDM内部检测遮挡区域。除了遮挡区域,运动变化区域内的其它像素点统称为运动区域。

3.1.2 BTBC区域检测 BTBC区域指的是即将在下一帧被覆盖的像素集合,以及将被其他运动物体遮盖的部分背景区域。本文检测BTBC区域的步骤如下:

(1) 利用传统的块匹配法进行粗匹配,插值后采用Horn&Schunck^[9]的算法进行前向稠密运动矢量场估计;利用估计出的稠密运动矢量场,进行后向运动补偿得到 k 帧的预测图像 \tilde{I}_k ;

(2) 计算 k 帧的运动补偿帧差 (DFD), 若 $DFD_k(x, y)$ 大于预先设定的阈值, 则认为该像素点为 BTBC 区域的点, 再进行形态学 (开、闭运算) 操作得到平滑的初始 BTBC 区域;

$$DFD_k(x, y) = |\tilde{I}_k(x, y) - I_k(x, y)| \tag{1}$$

(3) 将初始的 BTBC 区域利用 K -近邻法分割成亮度信息一致的区域, 空间不连续的区域则认为不同的分割区域; 用仿射模型描述每个区域, 通过估计出的前向稠密运动场得到仿射模型参数;

(4) 利用区域匹配搜索最佳匹配参数, 使得每个区域的 $DFD_k(x, y)$ 最小, 对于每个区域的 $DFD_k(x, y)$, 通过设定阈值得到精确的 BTBC 区域。

3.1.3 网格节点的运动估计 网格节点的运动估计是基于网格模型的运动估计方法中非常重要的环节, 节点的运动矢量估计应有很高的精度, 否则容易造成网格拓扑结构的改变, 并导致目标跟踪失败。本文采用特征窗口运动补偿匹配^[10]的方法进行网格节点运动矢量估计, 利用 Newton-Raphson 梯度下降法寻找局部最优解, 并结合双线性内插法获得很高的亚像素匹配精度, 有效地避免了传统块匹配产生的块效应。

节点的运动矢量估计过程取一局部窗口进行匹配, 考虑到各点的运动变形问题, 定义仿射运动矢量场:

$$\delta = DX + d \tag{2}$$

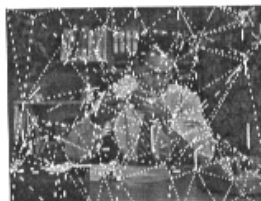
其中 δ 代表像素点 $X(x, y)$ 的运动矢量, D 是变形参数矩阵, d 是平移参数向量。令 $A = I + D$, 第 k 帧 I 的节点坐标 X 在第 $k + 1$ 帧 J 的位置为 $AX + d$, 即 $J(AX + d) = I(X)$ 。给定两幅图像 I 和 J , 以及 I 中的一个窗口, 运动估计就是要确定 A 和 d 的参数。由于图像中存在噪声, 而且仿射模型也只是对运动模式的近似逼近, 所以运动参数估计过程相当于一个最小化的优化计算问题, 为寻找最优的参数矩阵 A 和平移向量 d 使得运动补偿误差最小, 定义代价函数为:

$$\epsilon = \int \int_W [J(AX + d) - I(X)]^2 w(X) dX \tag{3}$$

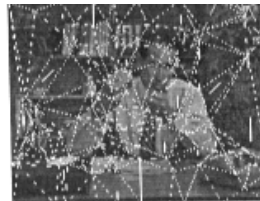
其中 W 代表特征窗口, $w(X)$ 是加权函数。

代价函数最小化过程: (1) 对式 (3) 求关于参数矩阵 A 和 d 的偏导数, 并令 $\partial\epsilon/\partial D, \partial\epsilon/\partial d$ 为零; (2) $J(AX + d)$ 通过 Taylor 级数展开, 取其一阶导数, 得到线性方程; (3) Newton-Raphson 最小化迭代求解 A 和 d 。

节点运动估计结果图如图 2 所示。



(a) 第 437 帧的初始网格



(b) 第 443 帧的网格

图 2 网格节点运动估计

网格节点的运动矢量建立了 k 和 $k+1$ 帧的对应关系, 可以通过对网格每个单元 (Patch) 的 3 个顶点求解式 (4), 即进行后向仿射变换得到每个单元的仿射参数 $a_1 \sim a_6$:

$$\begin{cases} x = a_1 x' + a_2 y' + a_3 \\ y = a_4 x' + a_5 y' + a_6 \end{cases} \quad (4)$$

其中 (x, y) 是 k 帧的像素点, (x', y') 对应 $k+1$ 帧的像素点. 再根据式 (4) 进行前向运动补偿得到 $k+1$ 帧各网格单元内部像素点的位置, 即可得到 $k+1$ 帧各像素点的运动矢量.

3.1.4 UB 区域检测 UB 区域指的是位于 $k+1$ 帧的, 与 k 帧没有对应像素的集合. 通过网格节点的运动估计, 并进行仿射变换运动补偿后, $k+1$ 帧中的模型失效区域就是 UB 区域. 根据式 (4) 进行前向运动补偿得到 $k+1$ 帧的预测图像 \tilde{I}_{k+1} , 其余处理与 3.1.2 节中式 (2)~(4) 类似, 得到精确的 UB 区域.

3.2 自适应插值

由于运动矢量场的非线性, 所以本文通过自适应插值完成帧插值过程. 通过像素分类以及遮挡区域检测将整个图像帧分为静止区域, 遮挡区域 (BTBC 和 UB) 和运动区域, 每个区域拥有自己的运动矢量. 为了使算法简单有效, 本文对于静止区域和运动区域, 采用仿射变换进行基于网格的运动估计, BTBC 区域采用后向运动估计, UB 区域采用前向运动估计. 得到了各区域的运动矢量, 也就得到了整个图像帧的运动矢量场, 就可以根据每个区域的特点分别进行运动补偿帧插值. 给定第 k 帧、 $k+1$ 帧, 通过两帧插值得到中间任一时刻 t ($k < t < k+1$) 对应的图像帧, 则插值帧 t 帧中像素点 $X(x, y)$ 的亮度的自适应插值公式为

$$I_t(X(x, y)) = \begin{cases} [I_k(X(x, y) - m\mathbf{MV}(x, y)) + I_{k+1}(X(x, y) + (1-m)\mathbf{MV}(x, y))]/2, & \text{(静止区域、运动区域)} \\ I_k(X(x, y) - m\mathbf{MV}(x, y)), & \text{(BTBC 区域)} \\ I_{k+1}(X(x, y) + (1-m)\mathbf{MV}(x, y)), & \text{(UB 区域)} \end{cases} \quad (5)$$

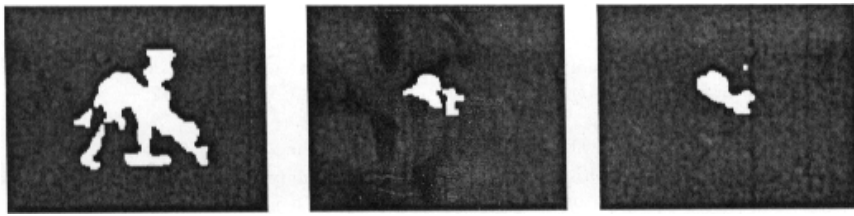
其中 $\mathbf{MV}(x, y)$ 表示像素点 $X(x, y)$ 的运动矢量, $0 < m < 1$.

4 实验结果

为了验证算法的有效性, 本文对不同的 QCIF(176×144) 视频序列进行测试. 图 3 是对 Salesman 序列第 437, 443 帧进行像素分类的结果, 图 3(a), 3(b), 3(c) 分别表示运动变化区域检测结果和遮挡区域检测结果. 图 4 是分别使用块匹配算法和本文算法得到的插值结果比较. 图 4(a), 4(b) 表示使用块匹配方法得到的插值图像和运动矢量场, 其中块大小取为 16×16; 图 4(c), 4(d) 表示使用本文算法得到的插值图像和运动矢量场. 可以看出本文算法得到的插值图像有效地消除了块匹配算法导致的块效应, 而且运动矢量场明显优于块匹配得到的矢量场, 平滑且可以很好地描述物体运动. 为了更清晰表示本文算法的优越性, 分别对标准测试序列 Salesman, Carphone 进行测试, 使用的方法包括块匹配算法、像素递归法和本文的算法, 插值结果的峰值信噪比 (PSNR) 比较结果如图 5 所示. 由图 5 可以看出, 本文的算法性能明显优于其他两种算法, 主要原因在于本文算法考虑到遮挡问题的存在, 并进行了相应的处理.

5 结论

遮挡问题是序列图像运动估计中一个很难解决的问题, 遮挡区域的运动估计不准是导致整个运动矢量场估计不准的主要因素. 本文利用基于网格模型的运动估计方法对该问题进行探讨, 并提出一种基于网格模型的运动补偿自适应帧插值技术. 通过像素分类将图像帧中分为 3 种不同区域, 根据各自的特点分别进行运动估计和运动补偿插值, 有效地减小了遮挡区域带来的运



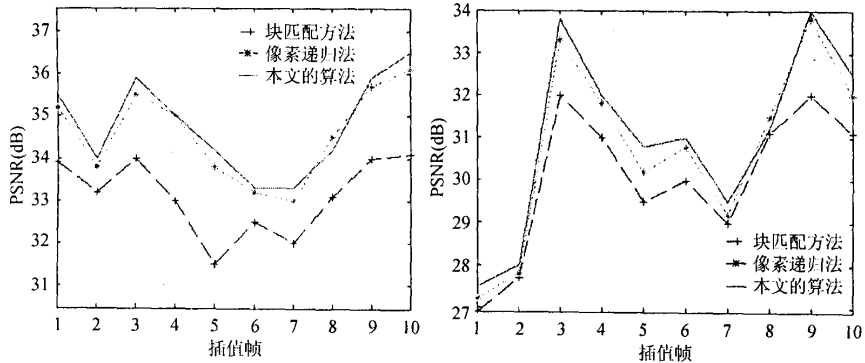
(a) 运动变化区域检测 (b) BTBC 区域检测 (437 帧) (c) UB 区域检测 (443 帧)

图 3 像素分类



(a) 使用块匹配方法得到的插值图像 (b) 使用块匹配方法得到的运动矢量场 (局部放大) (c) 使用本文算法得到的插值图像 (d) 使用本文算法得到的运动矢量场 (局部放大)

图 4 插值结果对比图



(a) Salesman 序列

(b) Carphone 序列

图 5 PSNR 比较图

动估计误差。本文提出的遮挡检测方法可以更加准确检测遮挡区域，是准确对各区域进行自适应插值的重要前提。同时，本文采用特征窗口运动补偿匹配的方法进行网格节点运动矢量估计，不仅消除了传统块匹配方法导致的块效应，而且得到了亚像素的匹配精度。

本文提出的运动补偿帧插值方法简单易行，对解决二维运动估计中的遮挡问题具有一定的实践意义。

参 考 文 献

- [1] Thoma R, Bierling M. Motion compensating interpolation considering covered and uncovered background. *Signal Processing: Image Communication*, 1989, 1(1): 191-212.
- [2] Sezan M, Lagendijk R. Frame rate conversion of motion picture films for progressive scan HDTV. *Signal Processing VI, Proc. EUSIPCO'92, Brussels, Belgium, 1992: 1279-1282.*
- [3] Sullivan G J, Baker R L. Motion compensation for video compression using control grid interpolation. in *Proc. ICASSP, Toronto, Ont., Canada, 1991: 2713-2716.*
- [4] Wang Y, Lee O. Use of two-dimensional deformable mesh structures for video coding, part I—the synthesis problem: Mesh-based function approximation and mapping. *IEEE Trans. on Circuits Syst. Video Technol.*, 1996, 6(6): 636-646.
- [5] Wang Y, Lee O, Vetro A. Use of two-dimensional deformable mesh structures for video coding, part II—the analysis problem and a region-based coder employing an active mesh representation. *IEEE Trans. on Circuits Syst. Video Technol.*, 1996, 6(6): 647-659.
- [6] Altunbasak Y, Tekalp A. Occlusion-adaptive content-based 2-D mesh design and tracking for object-based coding. *IEEE Trans. on Image Processing*, 1997, 1P-6(9): 1270-1280.
- [7] 吴枫, 高鹏, 高文. 基于网格模型的运动估计技术. *电子学报*, 2000(5): 47-51.
- [8] Horn B K P, Schunck B G. Determining optical flow. *Artif. Intell.*, 1981, 17(2): 185-203.
- [9] Shi J, Tomasi C. Good features to track. TR93-1399, Cornell University, 1993.

赵建伟: 女, 1976 年生, 博士生, 主要研究方向为序列图像运动估计、数字视频处理、计算机视觉。
王 朋: 男, 1974 年生, 博士生, 主要研究方向为运动目标的检测和跟踪、视频图像处理。
刘重庆: 男, 1939 年生, 博士生导师, 教授, 主要研究方向为图像处理与模式识别、计算机视觉。