

基于像素与子块的背景建模级联算法

解文华, 易本顺, 肖进胜, 甘良才
(武汉大学 电子信息学院, 湖北 武汉 430072)

摘要: 针对子块级背景建模方法无法保证所提取前景形状的精确性及像素级背景建模方法无法有效处理非平稳场景的问题, 提出了一种背景建模分层模型, 首先采用文中子块级建模算法得到较为粗糙的背景区域和前景区域, 然后利用混合高斯模型对特定图像区域执行像素级的前景提纯或背景模型更新操作, 2 种不同层次的算法通过非对称前向反馈机制进行级联。实验结果表明, 所提分层模型在能够有效处理非平稳场景的同时保证了所提取前景形状的精确性, 且对光照突变不敏感, 建模效果优于级联算法中任一独立算法, 而处理时间小于 2 种独立算法处理时间之和, 满足了实时处理要求。

关键词: 目标检测; 背景建模; 混合高斯模型; 非平稳场景

中图分类号: TP391.7

文献标识码: A

文章编号: 1000-436X(2013)04-0194-07

Cascaded algorithm for background modeling using pixel-based and block-based methods

XIE Wen-hua, YI Ben-shun, XIAO Jin-sheng, GAN Liang-cai

(School of Electronic Information, Wuhan University, Wuhan 430072, China)

Abstract: Block-based background modeling couldn't obtain the exact shape of foreground, while pixel-based approaches couldn't handle non-stationary backgrounds effectively. To solve the problem, a hierarchical scheme for background modeling was presented. The hierarchical model used block-based method proposed to obtain coarse background and foreground regions firstly, and then the operations of pixel-level foreground refining and model updating based on Gaussian mixture model were performed on special regions of the input image. These two algorithms in different levels were combined by adopting an asymmetric feed-forward strategy. Experimental results show that the hierarchical method proposed can obtain the exact shape of foreground and process non-stationary scenes well, in addition, it is insensitive to illumination change and can provide better results than any single approach in it, meanwhile, the integrated computation time is shorter than the sum of those of running the block and pixel-level methods, and satisfies real-time processing.

Key words: target detection; background modeling; Gaussian mixture model; non-stationary scenes

1 引言

运动目标检测作为计算机视觉的重要研究课题之一, 在智能视频监控、基于视觉的人机交互、视频压缩等领域有着广泛的应用。背景差分法作为一种重要的运动目标分割方法, 其主要思想是通过比较当前帧与背景模型间的差别来实现对运动目标的检测, 其性能依赖于所使用的背景建模技术^[1,2]。

常用的背景建模方法一般可分为基于像素的背景建模方法和基于子块的背景建模方法两大类^[3,4]。

基于像素的方法为每个像素点建立独立的模型, 可以得到较为精细的运动物体形状信息。例如, ZivKovic 等提出的自适应混合高斯背景建模方法可以使得每个像素的高斯分量数目自适应地变化, 从而使得所建立的背景模型更准确^[5]; Elgammal 等提出使用基于核密度的非参数 (KDE, kernel density estimation) 模型来对背景分布进行建模, 该方法不事先假设背景分布, 从而避免了分布模型的选择和分布的参数估计^[6]; Han 等提出一个序贯核密度近似 (SKDA, sequential kernel density approximation) 方

法, 是 KDE 方法的近似与 KDE 方法相比, SKDA 方法减小了计算量和所需要的内存^[7]; Kim 等提出了码本(CB, codebook)模型算法, 不用预先假设背景分布, 每个像素使用一个包含多个码字的码本表示, 每个像素的样本值基于颜色差异和亮度差异聚类为一系列的码字, 可以有效处理多背景情况^[8]。但上述方法仅利用了每个像素在时间域上的分布信息, 而没有利用像素在空间域上的分布信息, 受模型学习速度的影响, 场景的突然变化不能被立即选入背景模型, 从而使得此类方法对场景的非平稳变化依然比较敏感, 在处理非平稳场景时(例如突然一阵风引起的树木晃动或水面波动等)依然会产生大量的虚警。

在基于子块的方法中, 一帧图像被划分为若干相互重叠或者互不重叠的子块, 利用这些子块的颜色或纹理等特征进行背景建模。例如, Heikkila 等利用每个像素邻域内的 LPB(local binary pattern)直方图去描述背景模型^[9]; 文献[10]通过求取子块颜色梯度变化值来确定其背景, 然后根据各子块的计算结果重建背景模型。此类方法对背景的局部运动不敏感, 更适用于非平稳场景, 但提取出的运动目标较为粗糙, 不适于对运动目标分割要求较高的应用场合。

针对像素级和子块级背景建模方法分别存在自身难以克服的缺点, 本文提出了一种有效的背景建模分层模型, 首先采用文中所提基于子块的背景建模方法, 将整帧图像划分为互不重叠的子块, 对每一子块的颜色统计特征建立背景模型, 并根据子块颜色均值的变化情况自适应调节前/背景判定阈值, 从而保证子块级前景检测的顽健性, 然后根据子块级检测结果, 利用像素级的混合高斯(MOG, mixture of Gaussians)建模方法对特定图像区域的像素做更细致的前/背景分类或对其背景模型进行更新。实验结果表明, 文中算法具有对光照突变和背景局部运动不敏感的特性, 前景检测效果高于级联算法中任一独立算法, 而处理时间小于 2 种独立算法的处理时间之和。

2 子块级算法

在基于子块的背景建模方法中, 将一帧大小为 $P \times Q$ 的图像划分为多个互不重叠的 $M \times N$ 的子块(为处理方便, 一般选择 M 、 N 使其能分别整除 P 和 Q), 对每个子块单独进行处理。为了对每一子块建

立背景模型, 需提取出子块的颜色或纹理等特征, 并对该特征建模, 本文采用文献[11]中的方法来提取描述子块的颜色特征向量 $\mathbf{v} = \{ \mu_{ht}, \mu_{hb}, \mu_{lt}, \mu_{lb} \}$

$$\mu = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N x_{m,n} \quad (1)$$

$$\left\{ \begin{array}{l} \mu_h = \frac{\sum_{m=1}^M \sum_{n=1}^N (x_{m,n} | x_{m,n} \geq \mu)}{\sum_{m=1}^M \sum_{n=1}^N \begin{cases} 1, x_{m,n} \geq \mu \\ 0, \text{其他} \end{cases}} \\ \mu_l = \frac{\sum_{m=1}^M \sum_{n=1}^N (x_{m,n} | x_{m,n} < \mu)}{\sum_{m=1}^M \sum_{n=1}^N \begin{cases} 1, x_{m,n} < \mu \\ 0, \text{其他} \end{cases}} \end{array} \right. \quad (2)$$

$$\left\{ \begin{array}{l} \mu_{ht} = \frac{\sum_{m=1}^M \sum_{n=1}^N (x_{m,n} | x_{m,n} \geq \mu_h)}{\sum_{m=1}^M \sum_{n=1}^N \begin{cases} 1, x_{m,n} \geq \mu_h \\ 0, \text{其他} \end{cases}} \\ \mu_{hb} = \frac{\sum_{m=1}^M \sum_{n=1}^N (x_{m,n} | \mu \leq x_{m,n} < \mu_h)}{\sum_{m=1}^M \sum_{n=1}^N \begin{cases} 1, \mu \leq x_{m,n} < \mu_h \\ 0, \text{其他} \end{cases}} \end{array} \right. \quad (3)$$

$$\left\{ \begin{array}{l} \mu_{lt} = \frac{\sum_{m=1}^M \sum_{n=1}^N (x_{m,n} | \mu_l \leq x_{m,n} < \mu)}{\sum_{m=1}^M \sum_{n=1}^N \begin{cases} 1, \mu_l \leq x_{m,n} < \mu \\ 0, \text{其他} \end{cases}} \\ \mu_{lb} = \frac{\sum_{m=1}^M \sum_{n=1}^N (x_{m,n} | x_{m,n} < \mu_l)}{\sum_{m=1}^M \sum_{n=1}^N \begin{cases} 1, x_{m,n} < \mu_l \\ 0, \text{其他} \end{cases}} \end{array} \right. \quad (4)$$

其中, $x_{m,n}$ 为子块中 (m, n) 处的像素值。 μ 为子块内各像素点亮度的平均值, 将子块内的像素按亮度值是否大于 μ 分为高亮度部分和低亮度部分, 并分别求出两部分像素的均值 μ_h 和 μ_l , 然后将高亮度部分和低亮度部分中的像素点分别根据是否大于 μ_h 和 μ_l 做进一步划分, 总共划分为 4 个部分, 并分别求出相应的均值 μ_{ht} 、 μ_{hb} 和 μ_{lt} 、 μ_{lb} , 以此均值参数来描述子块的颜色特征。

利用上述特征进行背景建模可以快速有效地检测出前景区域。为了更好地应对类似于树枝摇动的具有多峰分布特性的场景, 文中借鉴混合高斯背景建模的思想, 使每个子块 i 的背景模型包含一组

K 个带有权值的颜色特征向量 $\{v_0^i, v_1^i, \dots, v_K^i\}$, 每个特征向量(模型)的权值为 ω_k^i , 并按权值大小降序排列, K 个权值相加和为 1, K 值越大反映的背景越详细, 但计算量也相应增加, 一般取 3~5, 依背景复杂程度而定。 K 个特征向量与对应权值的初始化可采用如下方式进行: $v_0^i = v^i$, $\omega_0^i = 1$, $v_1^i = v_2^i = \dots = v_K^i = 0$, $\omega_1^i = \omega_2^i = \dots = \omega_K^i = 0$ 。 v^i 为第一帧图像中子块 i 的颜色特征向量。

当新一帧图像到来后, 对每个子块 i , 首先利用公式 $B_i = \arg \min_b (\sum_{k=1}^b \omega_{k,t}^i > T)$ 做前景成分和背景成分的判别, $\omega_{k,t}^i$ 表示 t 时刻下子块 i 的第 k 个模型权值, T 为背景阈值。满足此条件的前 B_i 个模型为背景成分, 后 $K-B_i$ 个模型为前景成分。

为判定某子块 i 属于背景块还是前景块, 首先计算出该子块的颜色特征向量, 然后利用欧式距离做模型匹配, 若匹配上前 B_i 个模型则为背景块, 若匹配上后 $K-B_i$ 个模型或没有匹配上任何模型则属于前景块。 t 时刻子块 i 的颜色特征向量 v_t^i 与其第 k 个模型 $v_{k,t}^i$ 之间的欧式距离 $D(v_t^i, v_{k,t}^i)$ 计算式为

$$D(v_t^i, v_{k,t}^i) = \sqrt{\sum_{j=1}^4 (v_{tj}^i - v_{k,tj}^i)^2} \quad (5)$$

其中, v_{tj}^i 和 $v_{k,tj}^i$ 分别为 v_t^i 与 $v_{k,t}^i$ 中第 j 个向量元素。若满足 $D(v_t^i, v_{k,t}^i)$ 小于距离阈值 T_D , 则认为当前子块的颜色特征与第 k 个模型相匹配。如果有匹配, 利用当前子块的颜色特征 v_t^i 更新最佳匹配(距离值最小)模型的特征向量值 $v_{o,t}^i$

$$v_{o,t+1}^i = (1 - \alpha_b) v_{o,t}^i + \alpha_b v_t^i \quad (6)$$

其中, α_b 为模型学习因子。

而对所有模型的权重按下式进行更新

$$\omega_{k,t+1}^i = (1 - \alpha_\omega) \omega_{k,t}^i + \alpha_\omega M_k, \quad \omega_{k,t+1}^i = \omega_{k,t+1}^i / \sum_{m=1}^K \omega_{m,t+1}^i \quad (7)$$

其中, α_ω 为权重学习因子, 对于最佳匹配模型, M_k 为 1, 其余为 0。

如果不存在相匹配的模型, 则需要添加新模型或用新模型替换掉权重最小的模型, 新模型的初始化操作按下式进行:

$$\begin{cases} \omega_{n,t+1}^i = \alpha_\omega \\ v_{n,t+1}^i = v_t^i \end{cases} \quad (8)$$

即将新模型的权重赋值为 α_ω , 并将新模型的特征向量赋值为当前子块的颜色特征值。

为尽量减少子块中噪声点对匹配结果的影响, 所划分子块尺寸越小, T_D 取值应越大。由于直接使用上述颜色特征进行建模无法应对光照突变等情况, 文中加入强度方差这一参数来自适应调整 T_D 的取值, 从而提高块建模算法的顽健性。若当前子块的强度方差与重建出的背景图像中相同块的强度方差相差很大, 则意味着该区域为运动区域的可能性更大。两子块的强度方差相似度由下式进行衡量

$$S = \frac{|U_B \cdot U_C^T|}{\|U_B\| \cdot \|U_C\|} \quad (9)$$

式中: $U_B = \{|\mu_B - x_0^B|, |\mu_B - x_1^B|, \dots, |\mu_B - x_{MN-1}^B|\}$, $U_C = \{|\mu_C - x_0^C|, |\mu_C - x_1^C|, \dots, |\mu_C - x_{MN-1}^C|\}$ 均为 $1 \times MN$ 的矢量, μ_B, μ_C 分别为两子块的强度均值, 而 x_i^B, x_i^C 分别为两子块中第 i 个像素点的像素值。

由式(9)可知, S 的取值范围为 0~1, 其值越大意味着 2 个子块越相似。 T_D 的取值由 $T_D = T_D(\beta + S)$ 重新确定(其中, β 为经验常数, 一般取 0.8), 即在当前子块与背景越相似时, 认为该区域发生运动的可能性越小, T_D 取较大的值, 使当前子块更趋于被判为背景; 反之, T_D 取较小的值, 使当前子块更趋于被判为前景。

3 混合高斯建模方法

基于像素的背景建模方法主要是根据每个像素在时间域上的分布信息构造其背景模型。混合高斯背景建模方法将像素的亮度分布用几个高斯函数的加权和来近似, 能够描述像素的双峰或多峰状态, 在处理比较复杂、光照缓慢变化和存在小幅度重复运动的场景时具有一定的顽健性和较好的实时性^[1,3], 已成为当前目标检测中主流的背景建模方法, 因此文中选用混合高斯背景建模方法作为级联算法中的像素级建模方法。

设 x_t 为 t 时刻的某一像素值, 其概率密度函数可以写为 H 个高斯分布的线性组合^[1,2]

$$P(x_t) = \sum_{i=1}^H \frac{\omega_{i,t}}{(2\pi)^{d/2} |\mathbf{Cov}_{i,t}|^{1/2}} \cdot e^{-\frac{1}{2}(x_t - \mu_{i,t})^T \sum_{i,t}^{-1} (x_t - \mu_{i,t})} \quad (10)$$

其中, $\omega_{i,t}, \mu_{i,t}$ 和 $\mathbf{Cov}_{i,t}$ 分别为 t 时刻第 i 个高斯分布的权值、均值和协方差矩阵。

H 个高斯分布按权重大小降序排列, 仅取前面

的几个高斯分布表示背景分布。随着场景的变化, 每个像素的混合高斯模型都需要被不断地学习更新, 方法是将像素的当前值 x_t 与其混合高斯模型中的 H 个高斯成分逐一比较, 若 x_t 与第 i 个高斯成分均值 $\mu_{i,t}$ 之间的差小于 δ 倍的该高斯成分的标准差 (通常 δ 设为 2.5~3.5), 则该高斯成分被 x_t 更新, 其余高斯成分保持不变。

4 级联策略

基于子块的背景建模方法监视了场景中更为全局的变化, 具有对背景局部运动不敏感的特性, 因此更适用于处理非平稳背景的情况, 但由于其分块操作, 提取出的运动目标较为粗糙, 无法满足对运动目标分割精度要求较高的应用场合; 而基于像素的背景建模方法, 如文中采用的混合高斯背景建模方法, 虽然能够提取出较为精细的运动目标形状信息, 但是由于其忽略了各像素点的空间相关性, 在对非平稳背景进行建模时容易造成误判。

为了在对非平稳背景有效建模的同时能够得到准确的运动目标形状信息, 文中联合上述 2 种不同层次的背景建模方法, 使其优势互补, 构成一种分层背景建模模型。不同于参考文献[12,13]等将 2 种不同类型的建模方法作为 2 个相互独立的算法模块, 并根据 2 个模块的检测结果做融合判决, 文中在 2 个模型之间采用非对称前向反馈机制, 在利用上述基于子块的方法得到运动目标区域信息的基础上, 使用基于像素的混合高斯建模方法对特定区域做进一步检测或模型更新, 有效提高了级联算法的执行效率。具体集成策略如下。

1) 精确运动目标形状提取

基于子块的背景建模方法由于充分考虑了相邻像素的空间相关性, 因此可有效抑制树叶摇动、水纹波动等非平稳背景的影响, 但所得到的运动目标形状一般比较粗糙。如图 1 所示, 其中, 图 1(a) 为原始视频图像, 图 1(b) 为采用文中子块级背景建模方法提取出的运动目标, 可见块效应较为严重, 造成运动目标(跳舞的男子)变形, 无法应用于人体行为分析等对运动目标分割精度要求较高的应用场合。

文中进一步利用混合高斯模型来消除块效应, 以得到较为精确的目标检测结果。为降低算法复杂度, 仅在由子块级算法检测出的前景区域上进行混合高斯建模, 从而达到像素级的目标提纯目的, 但直接在上述区域执行像素级提纯操作, 并不能得到

理想的效果, 如图 1(c) 所示, 提纯后的运动目标依然存在一定的块效应。这是由于在基于子块的背景建模方法中, 一方面, 当背景建模所使用的颜色或纹理等特征无法有效区分运动目标与背景时, 会造成漏检, 即前景缺失; 另一方面, 运动目标边缘的一些细节往往被划分进不同的子块, 当这些细节所占区域远小于分块的尺寸时, 其对整个子块的统计特征影响较小, 同样会造成前景缺失, 这意味着子块级背景建模方法并不能得到完整的前景区域, 直接将混合高斯建模方法作用于由子块级建模方法检测出的前景区域时, 无法应对由于前景块缺失而引起的块效应。

因此在使用混合高斯背景建模方法做运动目标的提纯时, 其作用区域不但包括已检测出的前景块还应包括前景块周围的区域, 如图 1(a) 所示(左肩位置), 较小的矩形框内为文中子块级算法检测出的前景区域, 与较小矩形框中心相重合、四边相平行的较大矩形框内为前景提纯区域, 对该区域仅仅执行前/背景点判别操作, 而不执行背景模型的更新操作, 得到如图 1(d) 所示的最终运动目标检测结果。

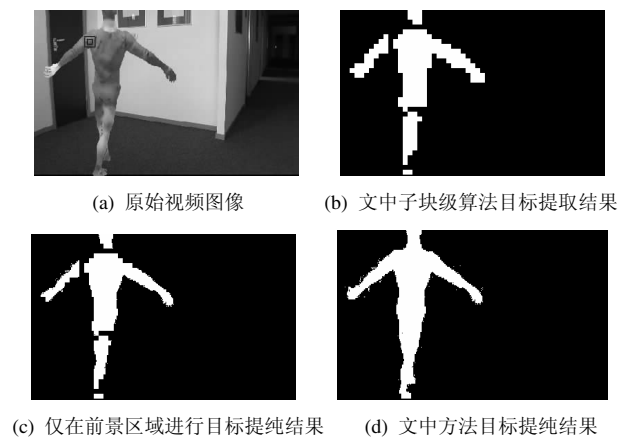


图 1 精确运动目标形状提取

2) 背景模型更新

为了使基于像素的混合高斯建模方法能适应背景的缓慢变化, 需要对各像素点的背景模型进行更新。为了降低算法复杂度, 文中级联算法并不是对整幅图像的所有像素点都进行背景模型的更新, 而是仅仅更新由子块级算法检测出的属于背景区域内的所有像素点的背景模型。这样做的原因是: 首先, 基于子块的方法采用一组像素点来做背景/前景的判决, 其背景判决结果较为可信, 因此在由子块级算法检测出的背景区域内无需再利用混合

高斯模型做进一步判决,从而提高算法的效率,但为了适应背景的缓慢变化,需要以当前像素的颜色信息来更新背景模型;其次,在子块级算法中被判决为前景的区域内通常同时包含真实前景点和不稳定背景点,因此在这些区域需要利用混合高斯模型做更精确的背景/前景点判决,并达到消除块效应的目的,但是由于这些区域所包含的信息更多为前景信息,所以不需要更新这些区域内各像素点处的背景模型。

综上所述,文中级联算法的流程图如图 2 所示。通过采取上述级联策略,整个算法具有对光照和背景局部运动不敏感的特性,既能有效处理扰动背景(如树叶晃动、水纹波动等),又能精确提取出运动目标的形状信息,前景检测效果高于级联算法中任一独立算法,而处理时间小于 2 种独立算法的处理时间之和。

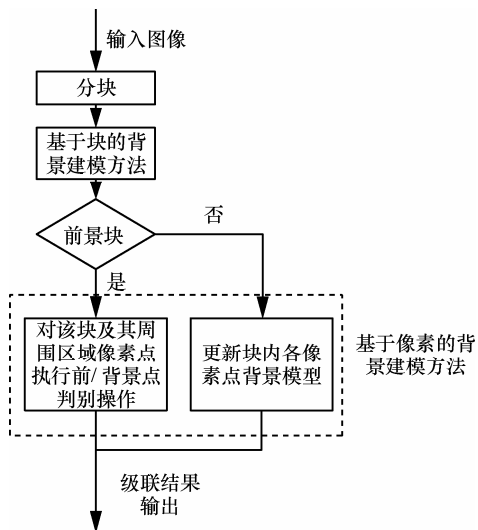


图 2 文中级联算法流程图

5 实验结果与分析

为验证文中背景建模分层模型的有效性,在多个视频序列上进行了测试,并与 CB 算法^[5]和混合高斯建模算法(MOG)^[2]进行了比较。在 5 个序列上,不同算法的测试效果如图 3 所示,图 3 中第 1、2 行分别为原始测试帧和对应的基准图;第 3、4、5 行分别为 CB 算法、MOG 算法和文中块算法的建模效果;第 6 行为文中级联算法的建模效果,级联算法采用第 4 节的级联策略对上述 MOG 算法和块算法进行级联,所采用的初始化参数与 MOG 算法和块算法相同,为较直观地对比 4 种算法的建模效果,

均未作阴影检测、形态学滤波等后续处理。图 3(a)、图 3(b)、图 3(c)所示分别为 duck 序列、WavingTrees 序列和 Campus 序列,这 3 个序列的主要特点是,其背景为存在水纹波动或树枝(叶)摇动的非平稳背景;图 3(d)所示为 video8_long 序列,其主要特点是,由于画面右上角走廊灯的开和关,造成背景出现局部的光照变化;图 3(e)所示为 cubicle 序列,其主要特点是当监控场景中出现较大的新目标时,由于摄像头的自动白平衡功能,导致画面出现大面积的亮度变化。上述 5 段视频的分辨率分别为 320×240、160×120、160×128、384×240、352×240。

由图 3 第 3、4 行可知,基于像素的 CB 算法和 MOG 算法虽然都具有一定的处理非平稳场景的能力,但是当背景出现较大幅度的扰动时,例如在图 3(a)、图 3(b)、图 3(c)中水纹波动和树枝(叶)摇动幅度较大,均导致了大量误检;同时由于 CB 算法和 MOG 算法对光照突变较为敏感,如图 3(d)、图 3(e)所示,当场景中出现人为开关灯或者监控摄像头自动白平衡等情况导致画面大面积亮度变化时,误检较为严重。从图 3 第 5 行可以看出,虽然文中块算法所提取的前景目标存在块效应,但由于其充分利用了各像素点的空间相关性,因此能有效地处理较大幅度的背景扰动,如在图 3(a)、图 3(b)、图 3(c)中,文中块算法较好地抑制了水纹波动和树枝(叶)摇动的干扰,准确提取出了运动目标;另外,如图 3(d)、图 3(e)所示,文中块算法通过引入强度方差相似度这一参数,在一定程度上克服了光照突变的影响。图 3 第 6 行所示为文中级联算法的建模效果,从图中可以看出,级联算法有效结合了 MOG 算法和块算法的优点,在有效处理非平稳场景(图 3(a)、图 3(b)、图 3(c)所示)和光照突变场景(图 3(d)、图 3(e)所示)的同时保证了所提前景形状的精细度,主观建模效果优于以上 3 种算法。

为进一步对级联算法的建模效果进行客观评价,采用下式的打分机制。

$$F = 2 \cdot \frac{pr}{p+r} \left(p = \frac{t}{t+f}, r = \frac{t}{t+n} \right) \quad (11)$$

其中, F 代表得分值,最高分为 1,最低分为 0,分值越高表明建模效果越好, t 、 f 、 n 分别为正确检测、错误检测和未检测出的前景点个数。

文中针对上述 5 个视频序列对 CB 算法、MOG 算法、块算法和级联算法的建模效果进行了打分,

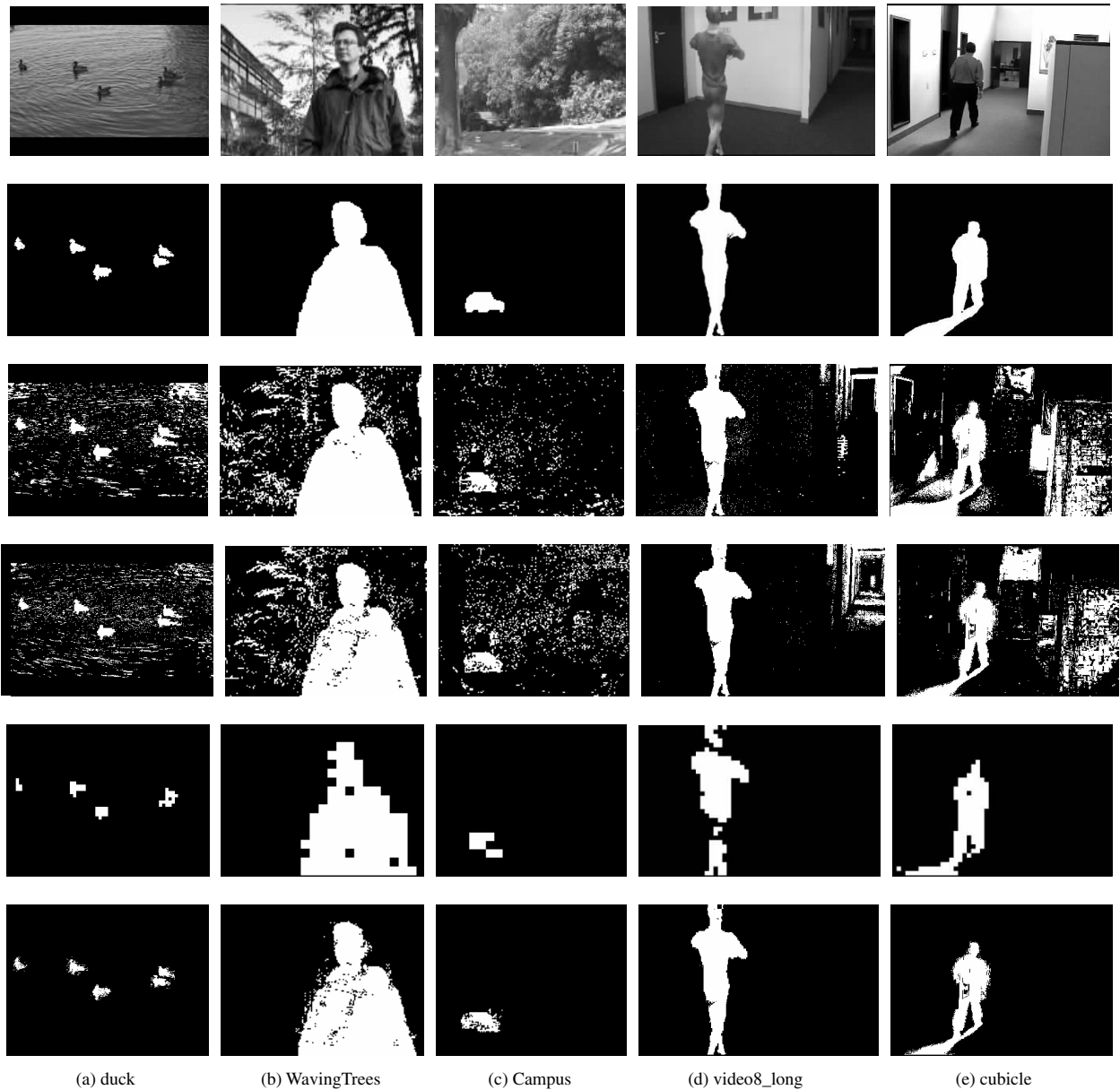


图 3 不同算法建模效果对比

打分结果如表 1 所示。

视频序列	打分结果			
	CB 算法	MOG 算法	块算法	级联算法
duck	0.300	0.331	0.675	0.881
WavingTrees	0.854	0.865	0.924	0.965
Campus	0.340	0.230	0.656	0.715
video8_long	0.823	0.781	0.824	0.975
cubicle	0.432	0.509	0.882	0.893

由表 1 可知，CB 算法与 MOG 算法在处理较为

复杂的非平稳场景与光照突变场景时，错检率较高，客观得分较低；块算法虽然能有效应对非平稳场景与光照突变场景，但由于其对图像进行分块处理，存在严重的块效应，因此客观评分结果也普遍偏低；而级联算法同时兼顾了像素级算法与子块级算法的优点，无论在处理非平稳场景上还是在处理光照突变场景上，其性能与客观得分均高于以上 3 种算法。

表 2 给出了针对上述 5 个视频序列，MOG 算法、块算法和级联算法处理时间的对比结果，为便于比较，所有的测试序列均被转换到 320×240 大小。实验硬件配置为 2.4GHz 酷睿 i3 双核处理器，2GB

DDR3 内存,操作系统为 Windows XP。由表 2 可知,由于文中级联算法并不是将子块级建模方法和像素级建模方法进行简单的级联,而是在 2 种建模方法之间采用了非对称前向反馈机制,利用子块级建模反馈信息,对不同区域执行不同的像素级处理操作,因此文中级联算法的处理时间小于 2 种独立算法处理时间之和,达到了实时处理要求。

表 2 不同算法处理时间对比

视频序列	处理时间/(毫秒/帧)		
	MOG 算法	块算法	级联算法
duck	21.4	5.3	24.2
WavingTrees	20.6	7.6	23.5
Campus	22.8	6.8	23.1
video8_long	19.2	6.4	22.8
cubicle	20.6	7.9	24.5

6 结束语

本文提出了一种有效的背景建模分层模型,首先利用文中所提块算法对输入图像进行子块级背景建模,得到较为粗糙的背景块和前景块,然后根据子块级建模结果,利用像素级的混合高斯建模算法对特定图像区域进行前景提纯或背景更新等操作。实验结果表明,文中级联算法能有效处理树叶摇动等非平稳场景,并具有对光照突变不敏感的特性,主观及客观建模效果均优于级联算法中的任一独立算法,与此同时,其处理时间小于 2 种独立算法处理时间之和,满足了实时处理要求。

参考文献:

[1] 朱齐丹,李科,张智等.改进的混合高斯自适应背景模型[J].哈尔滨工程大学学报,2010,31(10):1348-1353.
ZHU Q D, LI K, ZHANG Z, *et al.* An improved Gaussian mixture model for an adaptive background model[J]. Journal of Harbin Engineering University, 2010, 31(10):1348-1353.

[2] PAL A, SCHAEFER G, CELEBI M E. Robust codebook-based video background subtraction[A]. Proceedings of the 35th International Conference on Acoustics Speech and Signal Processing[C]. Dallas: IEEE, 2010. 1146-1149.

[3] LIN H H, CHUANG J H, LIU T L. Regularized background adaptation: a novel learning rate control scheme for Gaussian mixture modeling[J]. IEEE Transactions on Image Processing, 2011, 20(3):822-836.

[4] KIM T, PAIK D T, PAIK J. Block matching-based background generation and non-rigid shape tracking for video surveillance[A]. Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition[C]. Beijing: China, 2007. 415-420.

[5] ZIVKOVIC Z, HEIJDEN F V D. Efficient adaptive density estimation

per image pixel for the task of background subtraction[J]. Pattern Recognition Letters, 2006, 27(7):773-780.

[6] ELGAMMAL A, DURAISWAMI R, HARWOOD D, *et al.* Background and foreground modeling using nonparametric kernel density estimation for visual surveillance[J]. Proceedings of the IEEE, 2002, 90(7):1151-1163.

[7] HAN B, COMANICIU D, YING Z, *et al.* Sequential kernel density approximation and its application to real-time visual tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(7):1186-1197.

[8] KIM K, CHALIDABHONGSE T H, HARWOOD D, *et al.* Real-time foreground-background segmentation using codebook model[J]. Real-Time Imaging, 2005, 11(3): 172-185.

[9] HEIKKILA M, PIETIKAINEN M. A texture-based method for modeling the background and detecting moving objects[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(4):657-662.

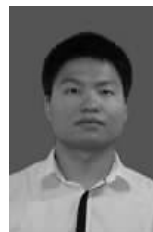
[10] 李刚,曾锐利,林凌等.基于帧间颜色梯度的背景建模[J].光学精密工程,2007,15(8):1257-1262.
LI G, ZENG R L, LIN L, *et al.* A novel background modeling method based on color grade of different frames[J]. Optics and Precision Engineering, 2007, 15(8):1257-1262.

[11] GUO J M, LIU Y F, HSIA C H, *et al.* Hierarchical method for foreground detection using codebook model[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011, 21(6):804-814.

[12] 李红波,曾德龙,吴渝.基于LBP与码本的视频分层背景建模和目标检测[J].数据采集与处理,2009,24(4):502-507.
LI H B, ZENG D L, WU Y. Layered video background model and object detection based on local binary pattern and codebook[J]. Journal of Data Acquisition & Processing, 2009, 24(4):502-507.

[13] 王永忠,梁彦,潘泉等.基于自适应混合高斯模型的时空背景建模[J].自动化学报,2009,35(4):371-378.
WANG Y Z, LIANG Y, PAN Q, *et al.* Spatiotemporal background modeling based on adaptive mixture of gauss[J]. Acta Automatica Sinica, 2009, 35(4):371-378

作者简介:



解文华(1987-),男,河南许昌人,武汉大学博士生,主要研究方向为多媒体处理和计算机视觉。

易本顺(1965-),男,湖北武汉人,武汉大学教授、博士生导师,主要研究方向为多媒体网络通信、视音频处理、无线通信和光纤通信等。

肖进胜(1975-),男,湖北武汉人,武汉大学副教授,主要研究方向为视音频信号处理和网络通信等。

甘良才(1942-),男,湖北咸宁人,武汉大学教授、博士生导师,主要研究方向为自适应数字通信、扩频通信、移动通信和软件无线电等。