

基于 2D 时空熵门限的运动目标检测

刘 洋 李玉山

(西安电子科技大学电路 CAD 研究所 西安 710071)

摘要: 该文给出一种基于二维(2D)时空熵门限进行运动目标检测的方法。研究了通用运动目标检测方法的特点,运用 2D 熵门限分割方法检测运动目标的结构区域,在 2D 熵门限的基础上,提出了一种快速熵门限求解算法。实验结果表明,这种方法可以很好地检测出运动目标,同时可大大提高运算速度。

关键词: 运动目标检测, 二维熵门限, 形态学方法

中图分类号: TP391.4 文献标识码: A 文章编号: 1009-5896(2005)01-0039-04

The Moving Object Detection Based on 2D Spatio-temporal Entropic Thresholding

Liu Yang Li Yu-shan

(Institute of Electronic CAD, Xidian University, Xi'an 710071, China)

Abstract A method of moving object detection based on two dimensional(2D) spatio-temporal entropic thresholding is proposed in this paper. The characteristics of several common approaches for moving object detection are analyzed, and 2D entropic thresholding method is used to detect the structural region of the moving object. The 2D thresholding method is first deduced. A fast algorithm for entropic thresholding is put forward. The experimental results show that this method can be used to detect the moving object, and reduce the computational time efficiently.

Key words Moving object detection, 2D entropic thresholding, Morphological method

1 引言

运动目标检测在计算机视觉,模式识别以及运动编码方面起着重要的作用。在运动目标检测的相关文献中研究人员提出了大量的运动目标检测和提取算法,例如贝叶斯方法^[1],形态学方法^[2],基于边缘的检测方法^[3]和基于门限的检测方法等^[4,5]。传统的运动目标检测方法通常是首先估算出稠密的运动区域,然后根据运动信息来分割图像。如果邻近的像素单元服从相同的霍夫或仿射变换,则将它们合并成一个区域。基于贝叶斯框架的运动目标检测方法采用马尔科夫随机场对图像建模,能够对图像进行运动估计和图像分割。但是贝叶斯方法的运算量很大,并且其中的很多算法需要大量的目标和区域模型作为输入参数。形态学方法,例如基于流域分割算法,也是运动目标检测和提取的一种有用工具。但由于梯度图像有可能与原始图像一样具有波动性,所以基于流域分割算法的有效性受到所使用的梯度图像的限制。基于边缘的方法被认为是更适合于语义视频图像分割的一种方法。因为人们对于运动目标的边缘是很敏感的,但从背景中提取运动目标边缘的效果很大程度上取决于预设的门限值。基于门限的方法,例如,灰度直方图熵门限法,用门限将图像划分成运动和静止区域。但是传统的基于门限的运动目标提取

算法存在两个问题:首先是如何选取有效的视觉特征来提取运动目标,这些视觉特征必须尽可能利用人类视觉系统原理和图像分量的主要性质。其次是如何根据图像分量的变化来自适应地确定全局特征门限值。由于不同的图像序列包含不同的图像分量,而且它们的变量特性也明显不同,所以对于运动目标检测,任何预设的固定门限值都是不合适的。一个有效的运动目标检测方法必须建立运动目标沿时间轴的对应关系,同时由于运动目标不仅具有运动区域的时间相关性,而且具有空间灰度值的相关性,所以应该运用图像序列的帧间对比度和局部方差来进行有意义的运动目标检测^[5]。此外,对不同的图像序列应该自适应地确定其全局特征门限值,我们将在传统 2D 熵门限方法的基础上通过推导得出一种改进的算法。

2 传统的 2D 熵门限

门限是将图像中的目标区域从背景中分离出来的一种通用方法。很多提出的关于图像分割的方法是根据一维灰度直方图进行的。例如,一维灰度直方图熵门限算法^[4]。一维熵门限方法仅仅利用图像的空间信息,对噪声比较敏感,随着信噪比(SNR)的降低,分割目标的性能也受到很大的影

响。Abutaleb^[6]提出的2D熵门限方法充分利用了图像中像素的灰度及其邻域灰度平均值,形成二维直方图,并依次为依据选取最佳阈值,该方法对信噪比较低图像也能取得比较好的效果。1992年,Brink^[7]通过最大化背景和目标的熵中较小的那一个,而不是求解两类熵和的最大值对Abutaleb的方法进行了改善。

2D熵门限方法中,二维参数对(像素灰度值、邻域灰度平均值)的分布概率形成一个二维直方图。2D熵门限方法与一维熵门限方法相类似,一维熵门限方法是根据求出的门限将图像划分为物体和背景两类,而2D熵门限方法是根据求出的门限矢量 (S, T) 将图像划分为4个区域,如图1所示。

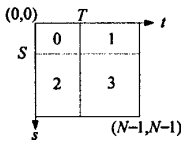


图1 传统2D熵门限划分的图像区域

图中的区域0和区域3主要用来表示背景和物体的分布,因为属于背景或物体区域的像素主要分布在对角线象限内,而区域1和区域2主要反映图像的边缘像素和噪声的分布。所以对应0和3区域上的二维参数对的分布概率在对应的区域应该有最大值,并对应于最大熵。

假设 $c_{i,j}$ 为在当前位置 (x,y) 的像素灰度值为 i ,在 (x,y) 邻域的像素灰度平均值为 j 的次数,它反映了相邻像素点灰度值的变化情况。 $c_{i,j}$ 的定义如下:

$$c_{i,j} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \delta_{i,j}, \quad \delta_{i,j} = \begin{cases} 1, & I(x,y)=i, \bar{I}(x,y)=j \\ 0, & \text{其它} \end{cases} \quad (1)$$

其中 $I(x,y)$ 表示 (x,y) 处的像素灰度值, $\bar{I}(x,y)$ 表示其邻域像素灰度平均值。由 $c_{i,j}$ 元素组成的矩阵我们称之为共生矩阵,并用 C 来表示。共生矩阵对角线上的峰值对应图像的区域部分,非对角线部分对应图像的边缘部分。共生矩阵可以在整幅图像或在窗口区域中进行计算,可以是对称的或非对称的。因此对应于静止背景和运动物体的分布概率可以分别定义为

$$P_{i,j}^0(S,T) = c_{i,j} / \left(\sum_{u=0}^S \sum_{v=0}^T c_{u,v} \right) \quad (2)$$

$$P_{i,j}^1(S,T) = c_{i,j} / \left(\sum_{u=S+1}^{N-1} \sum_{v=T+1}^{N-1} c_{u,v} \right) \quad (3)$$

根据Brink的最大熵阈值理论^[7],背景和物体的熵可分别定义为

$$H^0(S,T) = - \sum_{i=0}^S \sum_{j=0}^T P_{i,j}^0(S,T) \lg P_{i,j}^0(S,T) \quad (4)$$

$$H^1(S,T) = - \sum_{i=S+1}^{N-1} \sum_{j=T+1}^{N-1} P_{i,j}^1(S,T) \lg P_{i,j}^1(S,T) \quad (5)$$

门限矢量 (S,T) 应该满足方程

$$H(S,T) = \max \left[\min_{\substack{S=0, \dots, N-1 \\ T=0, \dots, N-1}} (H^0(S,T), H^1(S,T)) \right] \text{ 来使背景和物体的后}$$

验熵最大化。

为了找到全局最大值,上述的优化过程采用搜索算法来求得门限矢量。对于每一个 (S,T) 都要计算 $O(N^2)$ 次来求得目标和背景的熵,由于有 N^2 个 (S,T) ,所以求得全局门限矢量的运算次数是 $O(N^4)$ 。我们分析了共生矩阵和2D熵门限的性质,提出了一种快速2D熵门限求解算法。

3 快速2D熵门限求解算法

根据Brink背景和物体的熵的定义,我们可以看到,在计算每个 (S,T) 熵的时候很多的点都被重复进行运算,造成很多不必要的资源浪费。我们把 $\sum_{i=0}^S \sum_{j=0}^T c_{i,j}$ 和 $\sum_{i=S+1}^{N-1} \sum_{j=T+1}^{N-1} c_{i,j}$ 分

别用 $P_0(S,T)$ 和 $P_1(S,T)$ 来表示,即

$$P_0(S,T) = \sum_{i=0}^S \sum_{j=0}^T c_{i,j}, \quad P_1(S,T) = \sum_{i=S+1}^{N-1} \sum_{j=T+1}^{N-1} c_{i,j} \quad (6)$$

那么,在 $(S,T+1)$ 点对应的概率和熵可以通过下式求解:

$$\left. \begin{aligned} P_0(S,T+1) &= \sum_{i=0}^S \sum_{j=0}^{T+1} c_{i,j} = \sum_{i=0}^S \sum_{j=0}^T c_{i,j} + \sum_{i=0}^S c_{i,T+1} \\ &= P_0(S,T) + \sum_{i=0}^S c_{i,T+1} \\ P_1(S,T+1) &= P_1(S,T) - \sum_{i=S+1}^{N-1} c_{i,T+1} \end{aligned} \right\} \quad (7)$$

$$\begin{aligned} H^0(S,T+1) &= - \sum_{i=0}^S \sum_{j=0}^{T+1} \frac{c_{i,j}}{P_0(S,T+1)} \lg \left(\frac{c_{i,j}}{P_0(S,T+1)} \right) \\ &= - \sum_{i=0}^S \sum_{j=0}^T \frac{P_0(S,T)}{P_0(S,T+1) P_0(S,T)} \frac{c_{i,j}}{P_0(S,T)} \lg \left(\frac{P_0(S,T)}{P_0(S,T+1) P_0(S,T)} \frac{c_{i,j}}{P_0(S,T)} \right) \\ &= - \frac{P_0(S,T)}{P_0(S,T+1)} \sum_{i=0}^S \sum_{j=0}^T \frac{c_{i,j}}{P_0(S,T)} \lg \left(\frac{c_{i,j}}{P_0(S,T)} \right) \\ &\quad - \frac{P_0(S,T)}{P_0(S,T+1)} \sum_{i=0}^S \sum_{j=0}^T \frac{c_{i,j}}{P_0(S,T)} \lg \left(\frac{P_0(S,T)}{P_0(S,T+1)} \right) \\ &\quad - \sum_{i=0}^S \frac{c_{i,T+1}}{P_0(S,T+1)} \lg \left(\frac{c_{i,T+1}}{P_0(S,T+1)} \right) \\ &= \frac{P_0(S,T)}{P_0(S,T+1)} H^0(S,T) - \frac{P_0(S,T)}{P_0(S,T+1)} \lg \left(\frac{P_0(S,T)}{P_0(S,T+1)} \right) \\ &\quad - \sum_{i=0}^S \frac{c_{i,T+1}}{P_0(S,T+1)} \lg \left(\frac{c_{i,T+1}}{P_0(S,T+1)} \right), \end{aligned}$$

$$\begin{aligned} H^1(S,T+1) &= \frac{P_1(S,T)}{P_1(S,T+1)} H^1(S,T) - \frac{P_1(S,T)}{P_1(S,T+1)} \lg \left(\frac{P_1(S,T)}{P_1(S,T+1)} \right) \\ &\quad + \sum_{i=S+1}^{N-1} \frac{c_{i,T+1}}{P_1(S,T+1)} \lg \left(\frac{c_{i,T+1}}{P_1(S,T+1)} \right) \quad (8) \end{aligned}$$

同理,在 $(S+1, T)$, $(S+1, T+1)$ 点相应的概率和熵可以通过下式求解:

$$\left. \begin{aligned} P_0(S+1, T) &= P_0(S, T) + \sum_{j=0}^T c_{S+1, j} \\ P_1(S+1, T) &= P_1(S, T) - \sum_{j=T+1}^{N-1} c_{S+1, j} \end{aligned} \right\} \quad (9)$$

$$\left. \begin{aligned} H^0(S+1, T) &= \frac{P_0(S, T)}{P_0(S+1, T)} H^0(S, T) - \frac{P_0(S, T)}{P_0(S+1, T)} \lg \left(\frac{P_0(S, T)}{P_0(S+1, T)} \right) \\ &\quad - \sum_{j=0}^T \frac{c_{S+1, j}}{P_0(S+1, T)} \lg \left(\frac{c_{S+1, j}}{P_0(S+1, T)} \right) \\ H^1(S+1, T) &= \frac{P_1(S, T)}{P_1(S+1, T)} H^1(S, T) - \frac{P_1(S, T)}{P_1(S+1, T)} \lg \left(\frac{P_1(S, T)}{P_1(S+1, T)} \right) \\ &\quad + \sum_{j=T+1}^{N-1} \frac{c_{S+1, j}}{P_1(S+1, T)} \lg \left(\frac{c_{S+1, j}}{P_1(S+1, T)} \right) \end{aligned} \right\} \quad (10)$$

$$\left. \begin{aligned} P_0(S+1, T+1) &= P_0(S, T) + \sum_{i=0}^S c_{i, T+1} + \sum_{j=0}^T c_{S+1, j} + c_{S+1, T+1} \\ P_1(S+1, T+1) &= P_1(S, T) - \sum_{i=S+1}^{N-1} c_{i, T+1} - \sum_{j=T+1}^{N-1} c_{S+1, j} + c_{S+1, T+1} \end{aligned} \right\} \quad (11)$$

$$\left. \begin{aligned} H^0(S+1, T+1) &= \frac{P_0(S, T)}{P_0(S+1, T+1)} H^0(S, T) \\ &\quad - \frac{P_0(S, T)}{P_0(S+1, T+1)} \lg \left(\frac{P_0(S, T)}{P_0(S+1, T+1)} \right) \\ &\quad - \sum_{i=0}^{S+1} \frac{c_{i, T+1}}{P_0(S+1, T+1)} \lg \left(\frac{c_{i, T+1}}{P_0(S+1, T+1)} \right) \\ &\quad - \sum_{j=0}^{T+1} \frac{c_{S+1, j}}{P_0(S+1, T+1)} \lg \left(\frac{c_{S+1, j}}{P_0(S+1, T+1)} \right) \\ &\quad + \frac{c_{S+1, T+1}}{P_0(S+1, T+1)} \lg \left(\frac{c_{S+1, T+1}}{P_0(S+1, T+1)} \right) \\ H^1(S+1, T+1) &= \frac{P_1(S, T)}{P_1(S+1, T+1)} H^1(S, T) \\ &\quad - \frac{P_1(S, T)}{P_1(S+1, T+1)} \lg \left(\frac{P_1(S, T)}{P_1(S+1, T+1)} \right) \\ &\quad + \sum_{i=S+1}^{N-1} \frac{c_{i, T+1}}{P_1(S+1, T+1)} \lg \left(\frac{c_{i, T+1}}{P_1(S+1, T+1)} \right) \\ &\quad + \sum_{j=T+1}^{N-1} \frac{c_{S+1, j}}{P_1(S+1, T+1)} \lg \left(\frac{c_{S+1, j}}{P_1(S+1, T+1)} \right) \\ &\quad - \frac{c_{S+1, T+1}}{P_1(S+1, T+1)} \lg \left(\frac{c_{S+1, T+1}}{P_1(S+1, T+1)} \right) \end{aligned} \right\} \quad (12)$$

同时,根据共生矩阵的对称性特点,我们知道 $P_0(S, T) = P_0(T, S)$, $P_1(S, T) = P_1(T, S)$, $H_0(S, T) = H_0(T, S)$, $H_1(S, T) = H_1(T, S)$, 所以我们只需要计算共生矩阵对角线的上半部分就可以了。

我们知道小波分解能够有效地对二维信号作二次抽样,同时有效地保留原始信号的低频分量。利用 Wang 等人^[8]提出的通过小波分解进行求解 2D 熵门限的思想,我们采用两个步骤来求解全局特征门限矢量 (S, T) : 第 1 步,我们将计算得到的共生矩阵 C 进行 n 级小波分解,得到大小为 $N/2^n \times N/2^n$ 的共生矩阵 C 的低频分量 C_{low} , 然后将 C_{low} 看作新的共生矩阵利用上述迭代算法求解门限矢量 (S_{low}, T_{low}) 。第 2 步,根据第 1 步计算出来的门限矢量 (S_{low}, T_{low}) , 全局特征门限矢量 (S, T) 应位于以 $(2^n \times S_{low}, 2^n \times T_{low})$ 为中心,高度和宽度均为 2×2^n 的矩形区域内,运用快速迭代方法对共生矩阵 C 计算门限矢量 (S, T) 。

在实验中,我们对实验图像的共生矩阵 C 进行了三级小波分解,在奔腾 4 1.7G 微处理器上,用 VC++ 语言验证了快速 2D 熵门限算法的性能,与 Brink 算法相比,该快速算法可以节省大约 98% 的计算时间。比较结果(毫秒级)见表 1。

表 1 快速 2D 熵门限算法与 Brink 算法计算时间的比较(ms)

图像 (W×H)	使用 2D 熵门限快速求解 算法			Brink 算法总体 时间
	第 1 步 计算 时间	第 2 步 计算 时间	总体 时间	
Lena(256×288)	16	125	141	94982
Alex(352×288)	16	109	125	102031
Couple(256×256)	16	62	78	98266
Miss(256×256)	16	78	94	97640
Autumn(345×206)	16	94	109	81500

4 基于 2D 时空熵门限的运动目标检测

一个有效的运动目标检测方法必须建立运动目标沿时间轴对应关系,同时由于运动目标不仅具有运动区域的时间相关性,而且具有空间灰度值的相关性,所以扩展 2D 熵门限到时域进行有意义的运动目标检测。因为人眼对二阶差值比较敏感,因此选取图像序列的帧间对比度和局部方差作为图像序列运动目标检测的二维参数。在运动检测过程中,我们首先将图像序列分割成 3×3 的块,对于每一个块计算图像序列的帧间对比度(FCON)和局部方差(LCON)。

帧间对比度表示的是时间方向上不同帧同一位置上的像素灰度的连续性。定义如下:

$$FCON(m) = \sum_{x=0}^2 \sum_{y=0}^2 |I(x, y, t_n) - I(x, y, t_p)|^2 \quad (13)$$

其中 $I(x, y, t_n)$ 表示当前块 (x, y) 位置的灰度值, $I(x, y, t_p)$ 表示的是参考图像当前块 (x, y) 位置的灰度值。

局部方差表示的同一帧中相邻像素灰度值的连续性。定义如下:

$$LCON(m) = \sum_{x=0}^2 \sum_{y=0}^2 |I(x, y, t_n) - \text{MEAN}(m)|^2 \quad (14)$$

其中 $MEAN(m) = \frac{1}{9} \sum_{x=0}^2 \sum_{y=0}^2 I(x, y, t_n)$, 是当前块的平均像素灰度值。

由于 FCON 和 LCON 反映了运动图像序列时间和空间两个方向上像素灰度的连续性, 因此可以用于运动图像序列的分析和运动物体的检测。将式(2)~(5)中基于像素的参数换成基于 3×3 块的参数, 此时 $c_{i,j}$ 表示的是帧间对比度为 i , 局部方差为 j 的次数, 则可以得到基于块的最佳分割阈值。根据所确定的门限, 可以将当前帧分割成 4 个区域: 静止背景 (“0” 区域)、静止边缘 (“1” 区域)、运动边缘 (“2” 区域) 和运动区域 (“3” 区域), 见图 2。

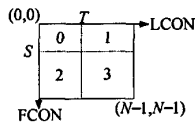


图 2 2D 时空熵门限划分的图像区域

利用上述的 2D 时空熵门限分割算法, 我们对公用图像序列和使用 USB 摄像板采集的图像序列进行运动目标的检测和提取。针对块分辨率造成的检测结果的轮廓不够平滑以及检测出的运动区域内部出现空洞的问题, 我们对运用 2D 时空熵门限方法分割后的结果, 采用形态学区域填充算法和闭运算, 以达到消除内部空洞和平滑轮廓的目的。实验结果如图 3 和图 4 所示。



(a) 公用图像 (b) 2D 时空熵门限检测结果 (c) 形态学填充和平滑结果

图 3 采用公用图像的实验结果



(a) 采集图像 (b) 2D 时空熵门限检测结果 (c) 形态学填充和平滑结果

图 4 采用 USB 摄像板采集图像的实验结果

比较图 3 (b), 图 4(b)与图 3 (c), 图 4(c)的结果, 可以看出, 使用 2D 时空熵门限可以很好地检测出运动目标, 但检测出的运动目标外部轮廓不够平滑, 内部有少许的空洞。采用形态学区域填充算法和闭运算可以很好地消除内部空洞和平滑外部轮廓, 从而提取有意义的运动目标。

5 结论

本文研究了运动目标检测的相关方法, 在研究传统 2D 熵门限结构方法的基础上, 提出了一种利用小波变换和迭代方法进行 2D 熵门限快速求解的算法。与 Brink 算法相比, 这种快速 2D 熵门限求解算法可以节省大约 98% 的计算时间。此外, 运用运动图像的时间空间相关性, 还能很好地检测出运动目标。这种算法可以广泛应用于运动视觉和视频压缩编码的许多场合。

参考文献

- [1] Forsyth D A, Ioffe S, Haddon J. Bayesian structure from motion. The Proc. of the Seventh IEEE International Conference on Computer Vision, Kerkyra, Greece, Sept 1999, vol.1: 660 - 665.
- [2] Wang D. Unsupervised video segmentation based on watersheds and temporal tracking. *IEEE Trans. on Circuits and Systems for Video Technology*, 1998, 8(5): 539 - 546.
- [3] Bilal Ahmad M, Tae-Sun Choi. Edge detection-based block motion estimation. *Electronics Letters*, 2001, 37(1): 136 - 144.
- [4] Fan Jianping, Zhu Xingquan, Wu Lide. Automatic model-based semantic object extraction algorithm. *IEEE Trans. on Circuits and Systems for Video Technology*, 2001, 11(10): 1073 - 1084.
- [5] Fan J, Elmagarmid A K. Statistical approaches to tracking-based moving object extraction. The Proc. of 1999 International Conference on Information Intelligence and Systems, Bethesda, MD, USA, 31 Oct. - 3 Nov. 1999: 375 - 381.
- [6] Abutaleb A S. Automatic thresholding of gray level pictures using two-dimensional entropy. *Computer Vision Graphics Image Process*, 1989, 47(1): 22 - 32.
- [7] Brink A D. Thresholding of digital images using two dimensional entropies. *Pattern Recognition*, 1992, 25 (8): 803 - 808.
- [8] Wang Qing, Wang Qiurang, Feng D D, et al.. A fast 2D entropic thresholding method by wavelet decomposition. The Proceedings of 2002 International Conference on Image Processing, Rochester, New York, June 2002, vol.3: 265 - 268.

刘洋: 男, 1980 年生, 博士生, 主要从事计算机运动视觉算法与研究等方面的研究。

李玉山: 男, 1945 年生, 教授, 博士生导师, 主要从事视觉芯片设计、机器视觉等方面的研究。