



(12) 发明专利

(10) 授权公告号 CN 102750712 B

(45) 授权公告日 2015.06.17

(21) 申请号 201210187511.X

TECHNOLOGY》. 2011, 第 21 卷 (第 4 期), 393–407.

(22) 申请日 2012.06.07

刘翠微等. 基于时空视频块的背景建

(73) 专利权人 中山大学

模. 《北京理工大学学报》. 2012, 第 32 卷 (第 4 期), 390–393.

地址 510275 广东省广州市海珠区新港西路
135 号

审查员 王亮

(72) 发明人 林倞 江波 徐元璐 梁小丹

(74) 专利代理机构 广州粤高专利商标代理有限公司 44102

代理人 林丽明

(51) Int. Cl.

G06T 7/20(2006.01)

(56) 对比文件

CN 102013022 A, 2011.04.13, 全文.

Xiaobai Liu et al. Integrating
Spatio-Temporal Context with Multiview
Representation for Object Recognition in
Visual Surveillance. 《IEEE TRANSACTIONS
ON CIRCUITS AND SYSTEMS FOR VIDEO

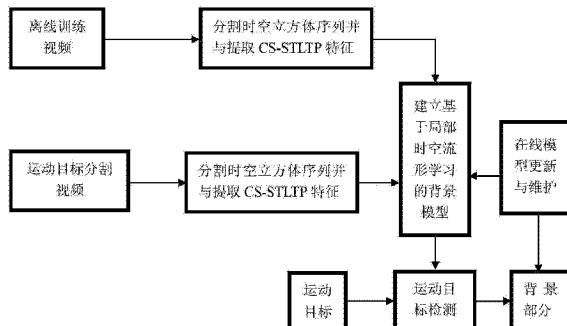
权利要求书2页 说明书7页 附图3页

(54) 发明名称

一种基于局部时空流形学习的运动目标分割方法

(57) 摘要

本发明公开了一种基于局部时空流形学习的运动目标分割方法，包括：1. 输入视频时空立方体的分割与光照不变性特征的提取；2. 基于局部时空流形学习的背景模型的建立；3. 基于局部时空流形的运动目标分割与模型的在线更新维护；利用本方法可以有效地描述局部的时空变化，处理输入视频中扩展图像中 SIFT 特征点集合的尺度适应性，解决在运动目标分割时不能有效排除动态背景，光照变化的问题，实现为智能监控平台提供可靠的，有效的运动目标。



1. 一种基于局部时空流形学习的运动目标分割方法, 其特征在于包括以下步骤 :
 - 1) 输入离线视频 ;
 - 2) 对离线视频进行时空立方体分割与对称时空局部纹理编码特征提取, 获得时空立方体特征的时空纹理变化描述序列 ;
 - 3) 根据时空立方体特征的时空纹理变化描述序列建立基于局部时空流形学习的背景模型 ;
 - 4) 输入运动视频 ;
 - 5) 对运动视频进行时空立方体分割与对称时空局部纹理编码特征提取 ;
 - 6) 通过判断运动视频时空立方体与通过背景模型预测的时空立方体间的距离, 得到目标运动和背景部分 ;
 - 7) 根据背景部分对背景模型进行更新和维护 ;
 - 8) 输入新的运动视频, 则跳至步骤 4), 否则结束 ;
- 步骤 2) 包括以下步骤 :
 - 21) 对离线视频进行时空立方体分割得到视频中每个单元位置的时空立方体序列 ;
 - 22) 对离线视频进行对称时空局部纹理编码特征提取, 获得时空立方体特征的时空纹理变化描述序列 ;
- 步骤 22) 对离线视频进行对称时空局部纹理编码特征提取包括以下步骤 :
 - 221) 取时空立方体任一像素点, 在像素点的 $3 \times 3 \times 3$ 的三维立方体临域中计算其特征向量, 把 $3 \times 3 \times 3$ 的三维立方体分割为 4 个时空平面提取像素点的时空局部纹理编码 ;
 - 222) 使用中心像素点对称的像素对, 通过设定尺度阈值对像素对进行比较编码, 计算像素对的对称时空局部纹理编码, 将对称时空局部纹理编码拉成一个大的特征向量作为该时空立方体的时空纹理变化描述。
2. 根据权利要求 1 所述基于局部时空流形学习的运动目标分割方法, 其特征在于步骤 3) 包括以下步骤 :
 - 31) 将步骤 21) 获得时空立方体序列延展成序列矩阵 ;
 - 32) 采用主成分分析对序列矩阵进行分解, 根据序列矩阵的主成分分析的特征值设定阈值, 大于设定阈值的 d 个特征值则作为局部时空流形的结构描述参数 ;
 - 33) 采用局部时空流形的结构描述参数重新表达序列矩阵, 得到局部时空流形的状态序列 ;
 - 34) 通过线性拟合计算局部时空流形的变化描述参数及对局部时空流形状态变化进行线性拟合后的误差矩阵 E ;
 - 35) 采用主成分分析对误差矩阵 E 进行分解, 根据误差矩阵 E 的主成分分析的特征值设定阈值, 大于设定阈值的 d 个特征值则作为产生式噪声描述参数。
3. 根据权利要求 1 所述基于局部时空流形学习的运动目标分割方法, 其特征在于所述步骤 6) 根据局部时空流形预测下一个时空立方体的局部时空流形状态和特征向量, 设定判断距离的阈值, 如果运动视频中某像素点对应的特征向量与预测出来的特征向量的欧式距离小于阈值则为背景部分, 否则为运动目标点。
4. 根据权利要求 3 所述基于局部时空流形学习的运动目标分割方法, 其特征在于所述步骤 7) 根据背景部分对背景模型进行更新是采用设定更新阈值与 CCIPCA 算法结合的增量

式更新。

一种基于局部时空流形学习的运动目标分割方法

技术领域

[0001] 本发明涉及视频监控领域,具体涉及运动目标分割技术,光照不变性特征的提取,局部时空流形离线学习,运动目标与背景分割,局部时空流形在线更新维护等领域。

技术背景

[0002] 目前,以视频监控为核心的公共场合安防系统正在大幅度的推广,不同于传统的人力式视频监控带来的巨大的人力花费和可能出现的大量误报漏报的问题,智能视频监控以其直观、准确、及时和信息内容丰富而越来越受到业界的关注和推广。近年来,随着计算机性能、网络的飞速发展,在原来智能监控中存在的计算瓶颈,传输瓶颈得以突破。复杂的、效果高的智能监控技术正在逐渐取代原有的简单的、注重运行效率的传统智能监控技术。运动目标分割,作为智能视频监控的一项基础型技术,一直都是大量的研究机构的关注的重点。

[0003] 运动目标分割技术的主要目的是排除监控视频中的背景对高层次的智能视频监控技术(运动目标跟踪,跨线检测,行为识别等)造成的干扰。其基本的方法是通过对监控视频中的背景提取特征,然后使用统计 / 概率模型进行背景建模。当视频中新来一帧后,通过判断图像能否被背景模型表达来判断是否出现了前景物体并分割出来。

[0004] 传统的运动目标分割方法,如多高斯(GMM),直接相减法,背景均值图像法等,均基于监控视频中的场景基本静止(且摄像头不会发生抖动),无突然的光照变化等等的基本假定,这也就决定了他们的适用范围固定在一些室内的简单场景下。而随着视频监控的不断推广,越来越多的室外复杂场景,传统的运动目标分割方法已经无法满足视频智能分析的需求。

[0005] 当下的监控视频中,常存在以下的几个问题:

[0006] 1、动态背景,如监控视频场景中出现的随风摇摆的树木,喷泉,水面,天气变化(雨,雪)等,此时简单的背景模型无法对动态背景中所有可能出现的情况进行充分的表示。

[0007] 2.、光照变化,如室内场景的开关灯,摄像头的自动光补偿,均会造成视频在全局层面上出现大幅度的变化。

[0008] 3、摄像头抖动,受客观条件影响,拍摄监控视频的摄像头难免会出现一定的抖动,这一部分对于目前的运动目标跟踪而言是一个巨大的难题,尚没有一种较好的方法能够简单且有效的解决该问题。

[0009] 为了解决上述的问题 1 和 2,许多改进的运动目标分割技术被提出。从背景模型的类型来分,大致可以分为以像素点为基础的分割,以图像块为基础的分割,以时空立方体为基础的分割三类。

[0010] 以像素点为基础的分割,其基本思想在于为每个像素点在整个监控视频中维护若干个简单的概率模型。但是不同于传统方法简单的使用像素值作为建模的样本,当下的以像素点为基础的运动目标分割技术一般使用局部的纹理编码(LBP)来描述一个像素点在其所处的局部空间内的特征。此类特征一般仅记录当前像素点与其邻域像素点的像素值的相

对大小,也就相当于记录了这一块区域的纹理变化规律,这样在发生光照变化的时候相对像素值大小基本保持不变,从而提高对视频中的动态场景和光照变化的描述能力,此类方法最近的研究发表在 2010 年 CVPR 上”Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes”。以像素点为基础的分割技术特点在于算法简单运行效率较高,但是因为局部的纹理特征仅能描述极为简单的变化,而在较为复杂的动态背景和光照变化的视频中,此类方法的效果较差。

[0011] 以图像块为基础的分割,不同于以像素点为基础的分割,通过使用固定大小的图像块作为背景模型的基础单元,使用主成分分析(Principle Component Analysis),判别式模型(Generalized 1-SVM)等来描述背景。这类方法在很多较为复杂的场景中效果较好,但是存在着两个局限性 :1. 对于时间较长的监控视频,很多模型会出现明显的退化现象。2. 图像块的基本假设将视频从时间和空间上分离,对时空上变化上的表达不够全面。此外,缺少一种以图像块为基础的局部纹理特征,故此类方法对光照变化较为敏感。

[0012] 以时空立方体为基础的分割,在图像块的基础上更进一步,通过从视频中分割出一个个的时空立方体,并假定每个时空立方体描述了背景在视频中一个简单的时空变化来对背景进行建模。常见的模型为稀疏字典表达 (Sparse Coding), 时空显著性 (Spatiotemporal Saliency)。以时空立方体为基础的分割能够很好描述动态背景在视频中的变化,当然,与之对应是模型复杂度和计算量的增加。与图像块为基础的分割类似的是,对于如何使用一种特征描述时空立方体中的纹理统计量,尚未有人提出方法。故此类方法如何能够提高在光照变化的效果有待研究。

[0013] 鉴于前面对于运动目标分割方法的描述,一种有效的分割技术应着力于解决两个问题 :1. 如何使用较为简单的模型描述动态背景的变化。2. 如何针对建模的基本单元提出一种光照不变性特征。

发明内容

[0014] 本发明为了解决运动目标分割中的复杂动态背景和突然光照变化的问题,实现为智能视频监控中运动目标跟踪,目标行为识别,跨线 / 跨区域检测提供稳定的,可靠的前景区域。

[0015] 本发明的技术方案为 :一种基于局部时空流形学习的运动目标分割方法,包括以下步骤 :

[0016] 1) 输入离线视频 ;

[0017] 2) 对离线视频进行时空立方体分割与对称时空局部纹理编码特征提取,获得时空立方体特征的时空纹理变化描述序列 ;

[0018] 3) 根据时空立方体特征的时空纹理变化描述序列建立基于局部时空流形学习的背景模型 ;

[0019] 4) 输入运动视频 ;

[0020] 5) 对运动视频进行时空立方体分割与对称时空局部纹理编码特征提取 ;

[0021] 6) 通过判断运动视频时空立方体与通过背景模型预测的时空立方体间的距离,得到目标运动和背景部分 ;

[0022] 7) 根据背景部分对背景模型进行更新和维护 ;

- [0023] 8) 输入新的运动视频,则跳至步骤 4),否则结束。
- [0024] 更进一步,步骤 2) 包括以下步骤:
- [0025] 21) 对离线视频进行时空立方体分割得到视频中每个单元位置的时空立方体序列;
- [0026] 22) 对离线视频进行对称时空局部纹理编码特征提取,获得时空立方体特征的时空纹理变化描述序列。
- [0027] 更进一步,步骤 22) 对离线视频进行对称时空局部纹理编码特征提取包括以下步骤:
- [0028] 221) 取时空立方体任一像素点,在像素点的 3*3*3 的三维立方体临域中计算其特征向量,把这个 3*3*3 的三维立方体分割为 4 个时空平面提取像素点的时空局部纹理编码;
- [0029] 222) 使用中心像素点对称的像素对,通过设定尺度阈值对像素对进行比较编码,计算像素对的对称时空局部纹理编码,将对称时空局部纹理编码拉成一个大的特征向量作为该时空立方体的时空纹理变化描述。
- [0030] 更进一步,步骤 3) 包括以下步骤:
- [0031] 31) 将步骤 21) 获得时空立方体序列延展成序列矩阵;
- [0032] 32) 采用主成分分析对序列矩阵进行分解,根据序列矩阵的主成分分析的特征值设定阈值,大于设定阈值的 d 个特征值则作为局部时空流形的结构描述参数;
- [0033] 33) 采用局部时空流形的结构描述参数重新表达序列矩阵,得到局部时空流形的状态序列;
- [0034] 34) 通过线性拟合计算局部时空流形的变化描述参数及对局部时空流形状态变化进行线性拟合后的误差矩阵 E ;
- [0035] 35) 采用主成分分析对误差矩阵 E 进行分解,根据误差矩阵 E 的主成分分析的特征值设定阈值,大于设定阈值的 d 个特征值则作为产生式噪声描述参数。
- [0036] 更进一步,步骤 6) 根据局部时空流形预测下一个时空立方体的局部时空流形状态和特征向量,设定判断距离的阈值,如果运动视频中某像素点对应的特征向量与预测出来的特征向量的欧式距离小于阈值则为背景部分,否则为运动目标点。
- [0037] 更进一步,步骤 7) 根据背景部分对背景模型进行更新是采用设定更新阈值与 CCIPCA 算法结合的增量式更新。
- [0038] 采用本技术方案为了保证背景模型的稳定可靠,使用离线训练视频的方式建立背景模型。可以有效地描述局部的时空变化,处理输入视频中扩展图像中 SIFT 特征点集合的尺度适应性,解决在运动目标分割时不能有效排除动态背景,光照变化的问题,实现为智能监控平台提供可靠的,有效的运动目标。

附图说明

- [0039] 图 1 为本发明的系统框图;
- [0040] 图 2 为实施例中时空光照不变性对称纹理编码示意图;
- [0041] 图 3 为运动目标分割结果。

具体实施方案

- [0042] 如图 1 所述,本发明包括以下步骤 :
- [0043] 1) 输入离线视频 ;
- [0044] 2) 对离线视频进行时空立方体分割与对称时空局部纹理编码特征提取,获得时空立方体特征的时空纹理变化描述序列 ;
- [0045] 3) 根据时空立方体特征的时空纹理变化描述序列建立基于局部时空流形学习的背景模型 ;
- [0046] 4) 输入运动视频 ;
- [0047] 5) 对运动视频进行时空立方体分割与对称时空局部纹理编码特征提取 ;
- [0048] 6) 通过判断运动视频时空立方体与通过背景模型预测的时空立方体间的距离,得到目标运动和背景部分 ;
- [0049] 7) 根据背景部分对背景模型进行更新和维护 ;
- [0050] 8) 输入新的运动视频,则跳至步骤 4),否则结束。

[0051] 其中各步骤的具体实现方式如下

- [0052] (1) 步骤 2) 对离线视频进行时空立方体分割与对称时空局部纹理编码特征提取
- [0053] 时空立方体指的是输入视频中分割提取出的融合时间与空间变化的基础分析单元。在系统中,对于连续的 t 帧,我们在每一帧的相同位置单元提取一个 $h*w$ 的图像块。然后将所有的 t 个图像块组合成一个 $h*w*t$ (在系统中统一使用 $4*4*5$) 的时空立方体。依据此原理,可以将输入视频分割为一个个的时空立方体序列 $V_i = \{v_{i,1}, v_{i,2}, \dots, v_{i,n}, \dots\}$, i 代表单元位置的索引。

[0054] 为了在时空立方体上提取能够对抗光照变化的特征,本发明提出一种改进的对称时空局部纹理编码(CS-STLTP),特征的提取过程见图 2。

[0055] 不同于其他的局部纹理编码只对一像素点的空间临域($3*3$)进行计算,为了加入时空的变化信息,我们对像素点在三维临域中进行计算编码。为了提高特征的计算提取效率,对一个时空立方体中任一像素点在该像素点的 $3*3*3$ 三维立方体临域内计算其特征。从立体几何的角度看,为了完整表达所有的时空变化,可以把这个 $3*3*3$ 的立方体分割成四个时空平面,然后分别计算局部纹理编码,最后拉成一个长特征向量作为这个像素点的时空变化描述。

[0056] 考虑到这样特征向量的长度变为了原来的 4 倍(原来只计算一个 $3*3$ 平面,现在计算 4 个 $3*3$ 时空平面),采取了对称编码的形式,将特征向量缩短一半,这样既能够保证时空变化信息能够被很好的描述,又保证了程序的执行效率。对称编码的方法是,原来使用中心的像素点与其 8 个相邻的像素点一一进行比较,然后计算得到一个纹理编码,现在忽略中心的像素点,将 8 个与中心像素点相邻的像素点分成 4 组关于中心像素点对称的像素对,然后对每个像素对分别计算,最后得到一个纹理编码。这样的比较计算次数由原来的 8 次变成了现在的 8 次,加快了特征计算速度。

[0057] 在实际的监控视频中,因为视频采集设备的种种问题,得到的视频一般都存在着一定的噪声,为了抑制噪声对特征造成的影响,不是单纯的比较两个像素点值之间的大小关系,而是设定一个尺度变化阈值,若与之比较的像素点值在其尺度变化范围内,假定这两个像素点是同一类并忽略掉他们之间的细微的差异。如果在其尺度变化范围之外,则根据

大小关系分成两种情况。在我们的系统中,像素值比较得到编码的规则设定如下:

$$[0058] \quad S_\tau(p_i, p_j) = \begin{cases} 1, & p_j \geq (1 + \tau)p_i \\ 0, & (1 + \tau)p_i > p_j > (1 - \tau)p_i \\ -1, & p_j \leq (1 - \tau)p_i \end{cases}$$

[0059] 式中 p_i, p_j 为不同像素点的像素, $S_\tau(p_i, p_j)$ 为纹理编码。

[0060] 按照上面的特征计算方法,一个像素点的特征长度为 $4*4=16$ 维(如果对每个颜色通道抽取特征则为 48 维),将时空立方体中每个像素点计算得到的对称时空局部纹理编码(CS-STLTP)拉成一个大的特征向量作为该时空立方体的时空纹理变化描述。

[0061] (2)步骤 3)根据时空立方体特征的时空纹理变化描述序列建立基于局部时空流形学习的背景模型;

[0062] 根据步骤 2)得到了时空立方体特征描述序列,在本步骤中,通过给予的离线训练的时空立方体特征描述序列,使用一种局部时空流形描述这个序列在其中的变化关系。

[0063] 为了简化时空立方体序列所在的局部时空流形结构与变化,首先假定时空立方体序列满足局部线性变化,因此采用一种线性动态系统(LDS)来描述。同样的假设在动态纹理的研究领域得到了广泛的认可和使用。许多复杂的纹理变化,如火焰,烟雾,水面等存在着复杂的难以描述的运动变化的纹理变化,都能够通过使用线性动态系统进行建模,合成,识别等工作。因此,假定在监控视频中存在的动态背景能够被线性动态系统较好的描述。在本系统中,采用自动回归平均移动模型(ARMA)进行建模。

[0064] 因此,将时空立方体序列的局部时空流形结构与变化描述为

$$[0065] \quad v_{i,n} = C_{i,n}z_{i,n} + \omega_{i,n}$$

$$[0066] \quad z_{i,n+1} = A_{i,n}z_{i,n} + B_{i,n}\varepsilon_{i,n}$$

$$[0067] \quad \omega_{i,n} \sim N(0, \Sigma_{\omega_i}), \quad \varepsilon_{i,n} \sim N(0, 1)$$

[0068] 公式中的 $C_{i,n}$ 描述了局部时空流形的结构(以线性代数的观点可以认为是局部时空流形的基底), $z_{i,n}$ 表示当前局部时空流形的状态, $A_{i,n}, B_{i,n}$ 描述了局部时空流形状态的变化, $\omega_{i,n}, \varepsilon_{i,n}$ 分别表示时空立方体中存在的噪声和局部时空流形状态变化中存在的噪声。在系统中,假定一个时空立方体序列的局部时空流形的维度为 d ,局部时空流形状态变化中产生式噪声的维度为 d_ε 。

[0069] 为了确定上述的模型中的参数,确定模型参数的具体步骤如下:

[0070] 将时空立方体序列 $V_i = \{v_{i,1}, v_{i,2}, \dots, v_{i,n}\}$ 延展成一个序列矩阵 $W_{i,n}$, $W_{i,n}$ 的每列为一个时空立方体的特征向量。这样把对序列的描述转换成线性代数中的对线性空间描述。

[0071] 基于局部时空流形的维度为 d 的基本假定, $C_{i,n}$ 可以认为是该局部时空流形的一组基底,因此选择主成分这组特殊的基底。使用主成分分析(PCA)对矩阵 $W_{i,n}$ 进行分解,矩阵 $W_{i,n}$ 的前 d 个主成分(PC)即为局部时空流形的结构描述 $C_{i,n}$ 。

[0072] 根据已经得到的局部时空流形的结构描述 $C_{i,n}$,使用 $C_{i,n}$ 重新表示时空立方体序列 $V_i = \{v_{i,1}, v_{i,2}, \dots, v_{i,n}\}$,得到每个时空立方体对应的局部时空流形状态 $Z_i = \{z_{i,1}, z_{i,2}, \dots, z_{i,n}\}$ 。

[0073] 假定时空立方体序列对应的局部时空流形状态满足线性变化,那么整个局部时空流形状态序列 $Z_i = \{z_{i,1}, z_{i,2}, \dots, z_{i,n}\}$,下一状态 $z_{i,n}$ 是通过上一状态 $z_{i,n-1}$ 乘以状态变化描述 $A_{i,n}$ 计算得到的。因此,求解状态变化描述 $A_{i,n}$ 可以通过线性拟合计算得到。 $A_{i,n}$ 表达为

[0074] $\min_{A_{i,n}} \left\| [z_{i,2} \ z_{i,3} \ \dots \ z_{i,n}] - A_{i,n} [z_{i,1} \ z_{i,2} \ \dots \ z_{i,n-1}] \right\|$

[0075] 在计算得到 $A_{i,n}$ 以后, 计算对局部时空流形状态变化进行线性拟合后的误差 E_i ,

$$E_i = [z_{i,2} \ z_{i,3} \ \dots \ z_{i,n}] - A_{i,n} [z_{i,1} \ z_{i,2} \ \dots \ z_{i,n-1}]$$

$$= B_{i,n} [\epsilon_{i,1} \ \epsilon_{i,2} \ \dots \ \epsilon_{i,n-1}]$$

[0078] 因此误差 E_i 可以认为是由产生式噪声 $\epsilon_i \sim N(0, 1)$ 生成的, 同时在前面已经假定局部时空流形状态变化中产生式噪声的维度为 d_ϵ , 那么, 误差 E_i 能够继续被主成分分析降维, 即选择误差 E 的前 d_ϵ 个主成分作为噪声的局部时空流形状态变化中产生式噪声 $B_{i,n}$ 。

[0079] 在实际应用中, 监控视频中的背景场景存在着多种情况。对于视频场景中静止的单元区域, 如天花板, 天空, 地面等, 它们的时空立方体序列所处的局部时空流形状态基本保持不变, 那么他们的局部时空流形的结构和变化都应该使用低维参数进行描述。对于视频场景中存在动态背景的单元区域, 如摇摆的树木, 波动的水面, 摆动的窗帘等, 它们的时空立方体序列所处的局部时空流形结构复杂, 状态变化也较为复杂, 此时应该使用高维的参数描述。

[0080] 因此, 在系统中, 根据时空立方体序列延展矩阵的特征值的大小自适应地确定该时空立方体序列的局部时空流形的维度和局部时空流形的状态变化中存在的产生式噪声的维度。即对序列矩阵 $W_{i,n}$, 误差矩阵 E_i 的主成分分析的特征值设定阈值, 然后选择所有特征值大于阈值的主成分作为 $W_{i,n}$, E 的基底。即

$$d^* = \arg \max \Sigma^d > T_d$$

$$d_\epsilon^* = \arg \max \Sigma_\epsilon^d > T_{d_\epsilon}$$

[0083] 其中 Σ^d , Σ_ϵ^d 分别对应于序列矩阵 $W_{i,n}$, 误差矩阵 E_i 的第 d 个特征值。采用自适应局部时空流形描述的好处在于简化了对于输入视频中的静止单元区域的模型, 减少了在运动目标分割与模型的在线更新维护阶段这些单元区域的计算量。

[0084] (3) 基于局部时空流形的运动目标分割与模型的在线更新维护

[0085] 根据已经训练得到的局部时空流形实现在线的运动目标分割。然后根据分割后得到的背景部分, 对局部时空流形进行更新维护。

[0086] 给出一个训练得到的局部时空流形, 可以根据局部时空流形的变化规律预测下一个时空立方体的状态, 再由预测的状态重建预测的下一个时空立方体的特征向量。即

$$\hat{z}_{i,n+1} = A_{i,n} z_{i,n} + B_{i,n} \epsilon_{i,n}$$

$$\hat{v}_{i,n+1} = C_{i,n} z_{i,n}$$

[0089] 在系统中, $z_{i,n}$ 表示离线训练计算得到的最近一个局部时空流形状态, $\epsilon_{i,n}$ 被假定为服从标准的正态分布。

[0090] 根据预测得到的时空立方体的特征向量, 通过比较预测的得到的特征向量与新来的一个时空立方体的特征向量。能够判别出新来的时空立方体中那些像素点为运动目标, 那些像素点为背景。在系统中, 为了保证分割算法的运行效率, 使用了简单的设定距离阈值的方式分割运动目标。也就是当新来一个像素点对应的特征向量与预测出来的该像素点的特征向量的欧式距离小于阈值分割 T_p 时, 认定该像素点为背景点, 否则认定为运动目标点。将背景点在二值的分割结果图像中赋值为 0, 将运动目标点在结果图像中赋值为 255, 就得到了新来的 t 帧图像(时间长度足够分割成时空立方体) 的运动目标分割结果。

[0091] 因为输入的视频的背景场景可能会随着时间的变化而发生变化,所以每个单元区域的局部时空流形也有可能发生改变。所以有必要在线的更新维护每个局部时空流形,以确保局部时空流形能够描述最近的结构和变化规律。

[0092] 但是如果直接把新来的时空立方体作为更新局部时空流形的样本的话,这些时空立方体中的运动目标部分就会被混入更新到背景模型当中。也就是说如果不对新来的时空立方体加以区分的话,运动前景就会污染背景模型。所以通过设定一个简单的更新阈值,来排除一部分对于背景模型的污染。当新来的一个像素点的特征向量与预测出来的该像素点的特征向量的欧式距离小于更新阈值 T_u 时,接受这个特征向量作为更新样本,反之使用预测的特征向量作为更新样本。

[0093] 因为局部时空流形的线性变化假设和输入视频中可能存在的各种噪声,无法保证所有的背景都能够处于局部时空流形中。所以更新阈值无法保证这些背景点的能够被更新到局部时空流形中。同样,更新阈值也无法保证那些与背景极为相似的运动目标能够被排除在外,所以在本系统中,使用增量式的更新维护方法以缓解运动目标错分,漏分带来的模型污染的问题。

[0094] 对于局部时空流形的更新分为两步:1、更新局部时空流形的结构;2、更新局部时空流形的变化。

[0095] 在系统中,将更新阈值与 CCIPCA 相结合的方式为在局部时空流形的更新维护中既保证了系统的运行效率,又保证了更新后的局部时空流形能够很好地描述变化后的局部时空流形结构。

[0096] CCIPCA 的更新算法将更新样本视为一个作用在原流形基底上的力,通过迭代式的计算更新样本在原流形的每个基底上的投影并根据该投影偏转对应的每个基底,实现增量式的流形基底更新。

[0097] 当更新完毕局部时空流形的结构以后(由 $C_{i,n}$ 更新为 $C_{i,n+1}$),可以由更新样本反向计算出其对应的更新后的局部时空流形的状态,即

$$[0098] z_{i,n+1} = C_{i,n+1}^T \hat{v}_{i,n+1}$$

[0099] 通过线性拟合得到局部时空流形的变化描述 $A_{i,n}$ 并在在线更新阶段,希望更新后的 $A_{i,n+1}$ 能够更好的描述最近的局部时空流形的变化规律。这是因为希望通过 $A_{i,n+1}$ 预测出来的局部时空流形的下一状态尽可能的准确。所以在算法中对最近的 L 个状态进行拟合得到更新后的局部时空流形的变化描述 $A_{i,n+1}$ 。

[0100] 最后,重新计算线性拟合误差 E_i ,然后再次对 E 进行主成分分析提取前 d_e 个主成分得到更新后的局部时空流形的产生式噪声 $B_{i,n+1}$ 。

[0101] 图 3 为运动目标分割结果:第一列为原视频中的一帧(包含人流密集的公共场合(场景 1),动态纹理(场景 2、3、5),突然光照变化(场景 4)),第二列手标得到的运动目标分割结果(groundtruth),第三列为我们的基于局部时空流形学习的运动目标分割方法的计算结果。

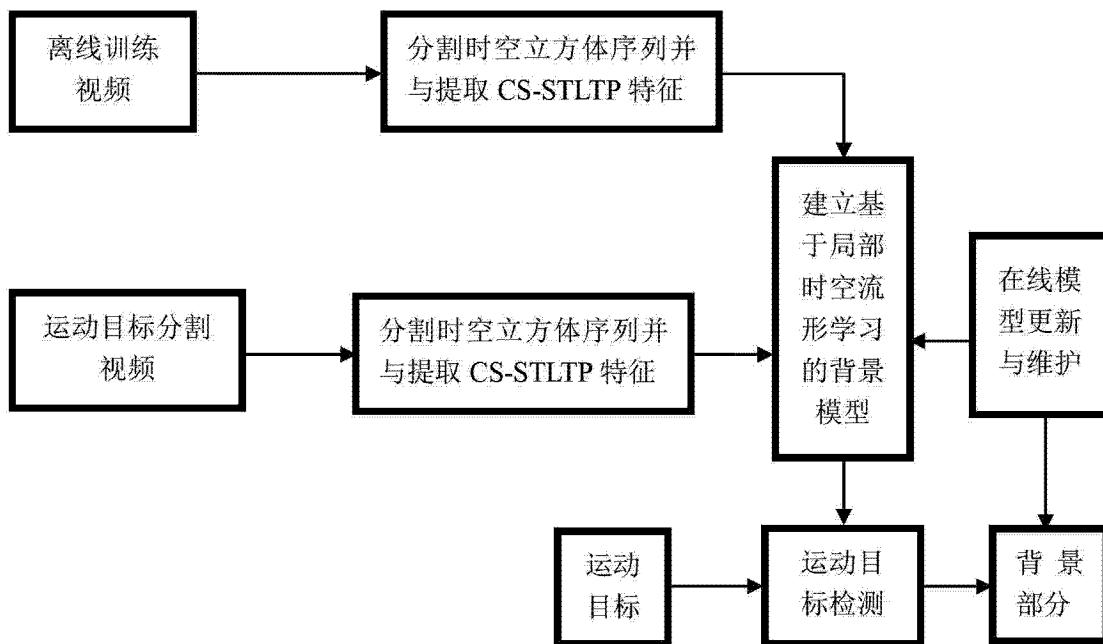


图 1

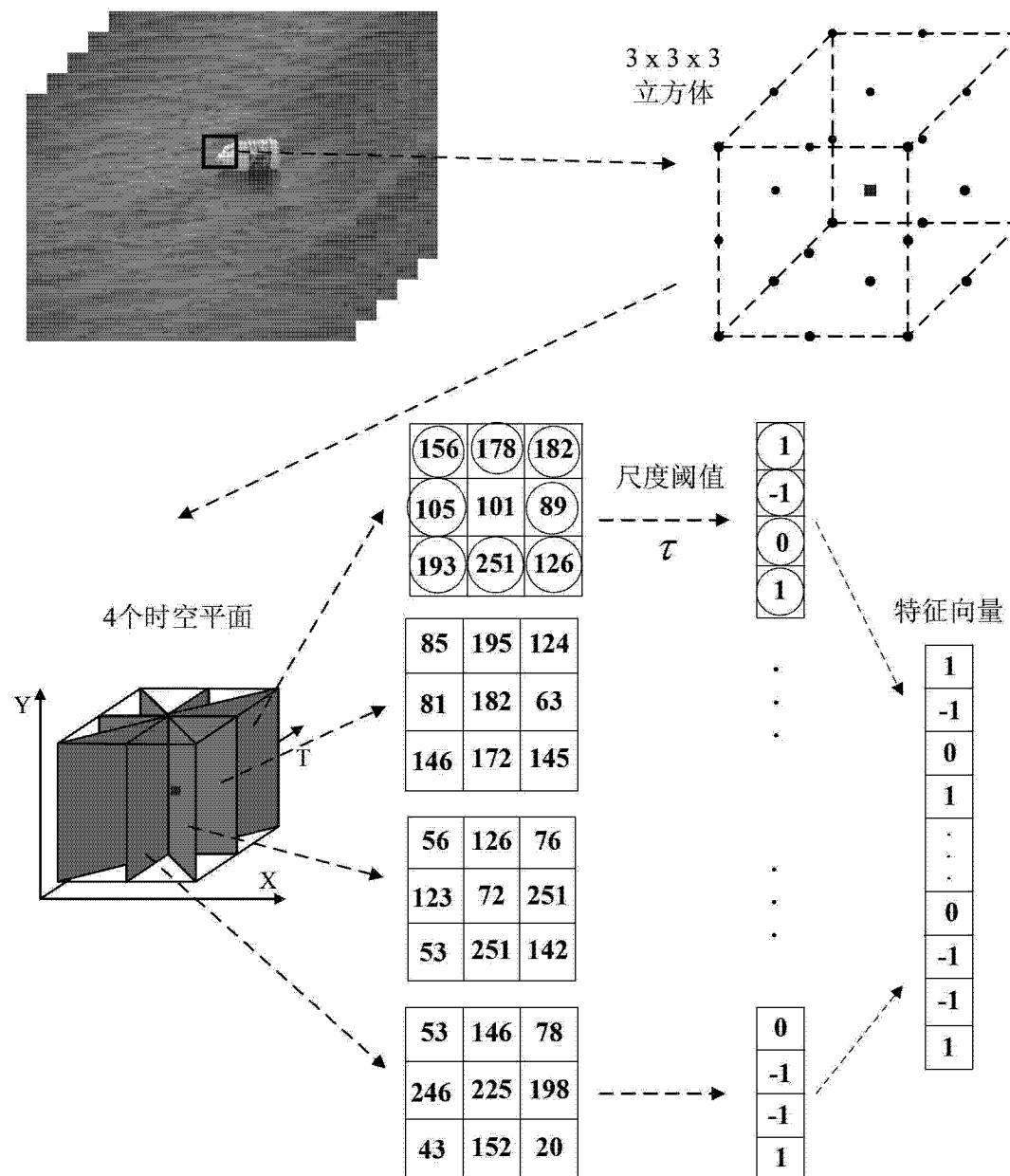


图 2

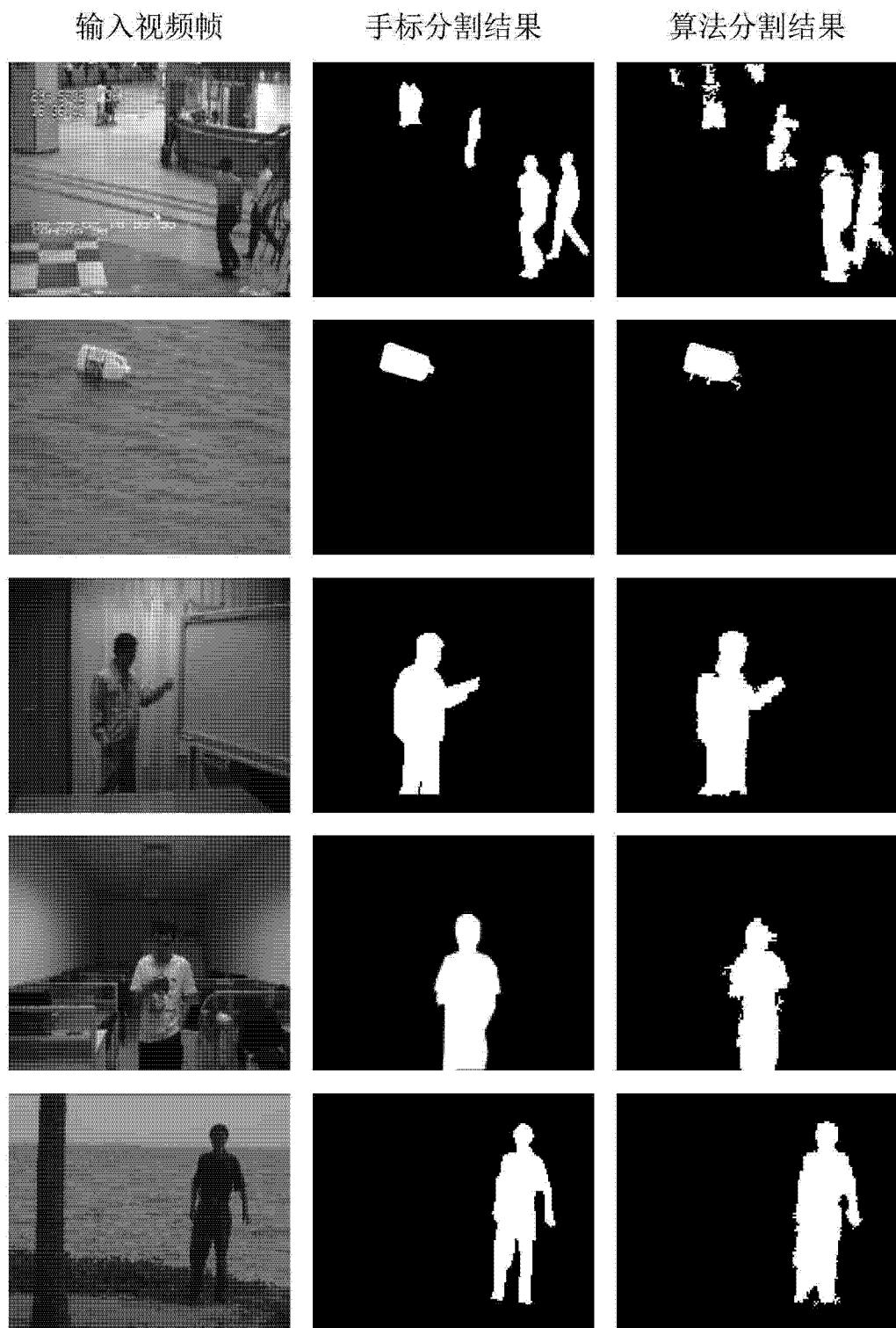


图 3