

文章编号: 2096-1618(2018)05-0509-08

野外早期火灾烟雾视频检测技术研究

高丰伟, 魏 维, 程 阳

(成都信息工程大学计算机学院, 四川 成都 610225)

摘要:针对视频中野外早期火灾烟雾形状模糊, 颜色相对背景对比不明显, 外观特性多样等特点, 提出一种并行的基于深度学习和动态纹理特征的烟雾识别方法。实验中, 针对视频烟雾检测中视频帧的噪声, 采用均值滤波的方法对视频帧进行过滤。并给出一种基于野外早期火灾烟雾视频的深度卷积神经网络结构, 并对该结构进行训练和测试。针对烟雾的动态纹理特征, 提出建立线性动态系统模型, 结合支撑向量机对烟雾视频帧进行分类识别。最后结合卷积神经网络的识别结果和动态纹理特征识别结果提出一种混合矩阵的判定方法。实验结果表明, 方法相对于传统烟雾识别方法在早期火灾烟雾的检测上有更高的准确率。

关键词:视频烟雾检测; 卷积神经网络(CNN); 线性动态系统(LDS); 动态纹理

中图分类号:TP391.41

文献标志码:A

doi:10.16836/j.cnki.jcui.2018.05.005

0 引言

火灾是当今世界上多发性灾害中发生频率较高的一种灾害, 也是时空跨度最大的一种灾害, 给人类的生活、生产乃至生命安全构成严重威胁^[1]。传统的点式光电感烟探测装置, 探测范围小, 在野外等较开阔的区域不仅成本大, 而且灵敏度低。近年来, 随着视频监控的普及, 视频火灾探测技术因其响应速度快、检测范围广以及能获取火灾周边的信息^[2]等特点备受国内外研究者的关注。

因为烟雾通常在火焰出现前就升起, 所以野外火灾烟雾检测对于早期火灾预警系统尤其重要。长距离视频早期野外火灾烟雾探测比短程或室内火灾烟雾探测更困难, 主要因素如下: (1) 早期野外火灾烟雾的主要特征是低速的散发传播速度, 形状模糊, 颜色相对背景对比不明显, 外观特性多样。(2) 野外火灾烟雾的流体旋涡和螺旋特性不可探测。(3) 雾气、云等都与火灾烟雾具有相似的颜色和纹理特征。

视频烟雾检测技术的研究方法主要是提取烟雾的颜色、运动及纹理等特征, 之后运用分类器对这些特征进行训练, 并对待检测视频作出识别, 判断其是否为烟雾区域。Ye Wei 等^[3]先对图像序列进行 Surfacelet 变换, 再对变换后的系数进行广义高斯建模, 获得与系数相对应的模型参数作为特征, 最后用 KL 距离做相似性度量, 对烟雾进行识别; Chunyu Yu 等^[4]提出了基于前景图像累积和光流法的检测方法。先用帧差法和块图

像处理法提取火焰特征, 并使用块图像处理法和光流法提取烟雾特征, 最后综合这两种特征, 使用反向神经网络进行火焰和烟雾的识别; 胡燕等^[5]在深入分析火灾烟雾图像特征的基础上, 通过 Harris 检测算法找到强度变化剧烈和图像边缘的特征点, 根据光流场与运动场的对应关系由成像平面中光流的变化估计烟雾的相对运动, 计算运动矢量信息, 实现多特征烟雾检测。ByoungChul Ko 等^[6]在火灾烟雾的空间域提取 HOG (histogram of oriented gradient) 特征, 并在时间域提取 HOOF (histogram of oriented optical flow) 特征, 运用这两种特征组成的码本, 再结合具体的加权方法生成 BOF (bag of features) 直方图, 最后使用随机森林进行分类。

传统的动态纹理基本采用光流法, 以及空间域和时空域的运动方向等对动态特征建模, 然而由于烟雾具有非刚性的特点, 形状、运动方向等并不确定, 因此难以提取出合适于烟雾的动态纹理特征。2003 年, Doretto 等^[7]基于系统识别的思想, 提出了一种对动态纹理建模的方法——线性动态纹理系统 (linear dynamic system, LDS), 并给出该方法的一种近似解, 通过大量的实验表明^[8-10], 该方法在烟雾等动态纹理上表现出极大的优势。Panagiotis Barmpoutis 等^[11]结合时空特征、运动特征、动态纹理特征分别对烟雾进行识别, 最后利用各个特征判断出的结果进行 DSC (dynamic score combination) 识别。之后, 该作者在文献[11]的基础上, 对其中的 LDS 作出了改进, 提出了一种高阶线性动态纹理系统^[12], 使烟雾识别率得到了进一步的提升。

近年来, 由于深度神经网络 (deep neural network, DNN) 在计算机视觉领域具有高效性、普适性等优势, 也开始被应用到视频烟雾检测中。DNN 使用原始图

像作为输入,其能自动学习到从底层像素级到高层表示级的特征,免去了人工提取特征的步骤,让机器自动从数据中获取特征。卷积神经网络(convolutional neural networks, CNN)是 DNN 的一种重要网络模型,其在视频烟雾识别中表现出了一定的优势。Tao 等^[13]使用深度学习框架 Caffe,实现了卷积神经网络,并对多种不同的卷积神经网络以及德国航天中心的一个比较成熟的野外火灾算法 fshell 在野外火灾烟雾视频集上的效果做了对比,研究结果显示 CNN 在野外火灾烟雾识别中也能同样表现的很好。陈俊周等^[14]将 CNN 引入烟雾静态纹理特征提取,并将光流序列作为烟雾的动态纹理特征,提出了一种级联的卷积神经网络烟雾识别框架,该方法有效的降低了非烟雾区域的误检率。本文引入了线性动态纹理系统和卷积神经网络,将烟雾视频帧作为原始图片输入卷积神经网络提取烟雾的特征,对视频中的烟雾区域进行分类识别,同时,将烟雾视频帧作为线性动态系统的输入量,提取烟雾视频帧的动态纹理特征,使用 SVM 进行分类识别,最后结合各自判定结果采用混合判定的决策方法对烟雾视频帧作出判断,最终实验结果显示,文中方法在视频烟雾识别准确率和误检率上相对于传统的方法有更好的表现。

1 系统框架流程

首先对烟雾视频集中的每一帧进行均值滤波处理,然后将视频分为 F 帧 32×24 的烟雾视频集,首先将烟雾视频集放入卷积神经网络进行分类识别,再将其输入线性动态系统,得到烟雾视频集的动态纹理特征,该动态纹理特征利用支撑向量机进行分类识别,最终根据卷积神经网络识别结果和动态纹理识别结果采用混合判定的决策方法确定该视频序列是否为烟雾。烟雾识别系统框架如图 1 所示。

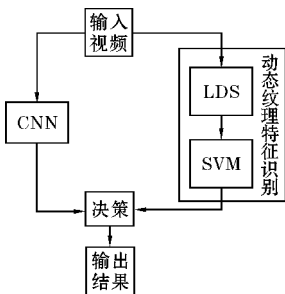


图1 烟雾识别系统框架图

2 视频预处理

如图 2,烟雾视频集的选取主要选择野外的视频

火灾烟雾,包括形状、稀薄程度、运动方向等特征不同的烟雾,以及雾天的、晴天的以及傍晚等不同场景下的早期火灾烟雾,这些视频共 10 个。这些视频主要是来源于网上,然后经过筛选,裁剪,归一化操作对视频进行处理,统一为对野外环境下的早期火灾烟雾具有一定的代表性的视频集。

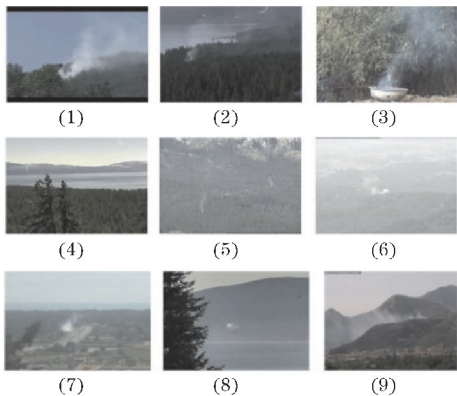


图2 烟雾视频集



图3 原始图片

如图 3,由于选择的视频集具有一定的噪声干扰,首先对烟雾视频集进行去噪处理,提取视频中的每一帧,先对整幅图像进行均值滤波,为了方便进行运算,采用卷积的方式对图像进行滤波,卷积核如公式(1)所示,大小为 5×5。

$$M_5 = \frac{1}{25} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

(1)

卷积的计算公式为:

$$I_m = I \times M_5 = \sum_{i=0}^h \sum_{j=0}^w (I(x+j,y+i)M_5(j,i))$$

(2)

其中, I_m 表示滤波后图像, I 表示原图像, h 和 w 分别表示高斯滤波模板的高和宽, x 和 y 分别表示当前滑动窗口的相对于原点的偏移量。

滤波后的图像如图 4 所示。图像看起来模糊化,这是因为图像的细节就是变化快的部分,即高频成分,滤波阻止了这些高频信号通过,则图像模糊,图像的模

糊平滑处理会减少图像噪声并降低图像细节层次。

依次对烟雾视频集中的所有视频进行均值滤波得到去噪后的烟雾视频集。由于烟雾具有非刚性的特点,形状,运动方向等并不确定,因此实验中,所用视频序列全部采用数字图形图像处理常用的宽为 32×24 的小块。非烟雾视频集的选取主要从这些烟雾视频集中的干扰因素中选取,如浮动的云,飘动的树以及夜晚的车灯等。

由于烟雾不仅有颜色等静态特征,还具有随时间变化的动态纹理特征,为了更好的对烟雾和非烟雾进行区别,提出了一种结合卷积神经网络和动态纹理特征对待检测视频进行识别的方法。



图 4 滤波后的图片

3 基于卷积神经网络的烟雾识别

近年来神经网络在图像处理,语音识别等领域都

有非常出色的表现,神经网络是一种模仿生物神经网络的结构和功能的数学模型或计算模型,用于对函数进行估计。卷积神经网络是在普通神经网络的基础上加入卷积层的特殊的神经网络,属于前馈神经网络的一种,对于语音、图像等二维数据有较好的效果。

对于图像而言,图像中单一的像素值并没有实际意义,当一个像素点和周围的像素点组合起来才能表达出图像中的信息。卷积神经网络通过卷积层可以将一块区域中的像素点结合起来,提取他们之间的关系,经过多层卷积层的特征提取,最终将提取出的特征输入普通神经网络进行分类,经过多次的训练,最终确定网络中各个神经元的最优参数,这样就训练好了一个卷积神经网络模型。同时卷积神经网络中局部连接和权值共享的特点也大大减少了神经网络的计算量。

为了让处理图像的卷积神经网络能够处理视频,抽取视频序列块中图像帧作为实验输入数据。实验方法中 CNN 的输入图像是从烟雾数据集中按每 10 帧提取一帧,有重叠的规则进行提取,图像大小为 32×24,共计提取 5000 张训练图像。

神经网络结构设计如图 5 所示。
具体的网络结构设计如下:

(1)输入层。输入层输入数据为 32×24 像素大小的图片,为了缩短计算时间,图片在输入到网络之前先进行灰度化、二值化处理,最终将只包含 0 和 1 的图像输入进网络进行训练。

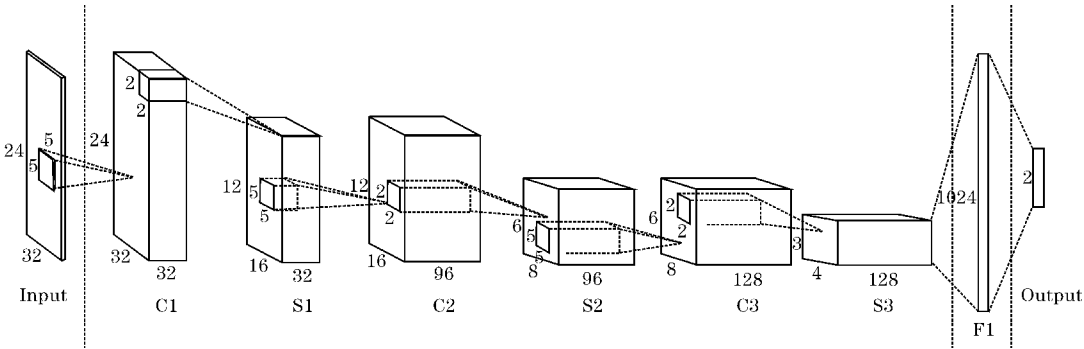


图 5 卷积神经网络结构

(2)卷积层 C1。取卷积核大小为 5×5,步长为 1,特征图个数为 32 对输入图像进行卷积,经过修正线性单元 (rectified linear unit, ReLU) 函数作为激活函数,其最终输出结果为 32×24×32 维数据。卷积核的数值使用方差为 0.1 的正态分布随机生成,偏置的值取 1,ReLU 激活函数的公式为

$$f(x) = \max(0, x) \tag{3}$$

(3)池化层 S1。使用 Max Pooling 方法对卷积层 C1 的输出结果进行重采样,输出结果为 16×12×32 维

数据。

(4)卷积层 C2 和 C3 的功能和 C1 相同,卷积核均为 5×5,步长为 1,特征图个数分别为 96 和 128,输出数据的维度分别为 16×12×96 和 8×6×96,池化层 S2 和 S3 的功能和 S1 相同,输出数据的维度分别为 8×6×96 和 4×3×128。

(5)F1 为全连接层,设置 1024 个神经元与 S2 相连,使用 ReLU 作为激活函数。

(6)F2 为输出层,设置 2 个神经元与 F1 相连,最

终对 1024 维向量进行 2 分类。

网络参数配置如表 1 所示。

表 1 网络参数配置

层号	类别	特征图数	核大小	步长
C1	卷积层	32	5×5	1
S1	池化层	32	5×5	2
C2	卷积层	96	5×5	1
S2	池化层	96	5×5	2
C3	卷积层	128	5×5	1
S3	池化层	128	5×5	2
F1	全连接层	神经元个数:1024		
F2	输出层	神经元个数:2		

在全连接层 F1 中,使用了 Dropout^[15] 方法避免过拟合情况的发生。Dropout 是 Hinton 提出的,通过阻止特征检测器的共通作用来提高神经网络的性能。在训练模型的过程中,Dropout 是指随机让网络中的某些隐含层节点的权重不工作,将这些节点在网络中屏蔽掉不参与运算,但是其权重保留,只更新工作的节点的权重,这样就能保证任意两个节点不能同时工作,从而避免了有依赖关系的两个隐含节点共同作用的情况,阻止了某些特征和其他特征相互依存的情况。网络中使用的优化方法为 Adam 优化器^[16]。卷积神经网络的训练图片集和测试图片集互不相同。在卷积神经网络的输出层,输出分类识别结果。

由于烟雾具有非刚性的特点,纹理特征不明显,使用传统的特征提取方法较难选取合适于烟雾的特征,卷积神经网络可以通过对大量图片的训练,自动学习最能代表烟雾的特征,免除了人工选择和提取特征的麻烦。对于视频数据来说,随时间变化的动态特征包含了更多的信息,因此文中加入了能表示烟雾随时间变化的动态纹理特征。

文中动态纹理通过将一组连续的视频帧输入线性动态系统模型 (linear dynamic system, LDS),进而求解出可以代表该视频的特征参数矩阵,作为烟雾的特征。

4 动态纹理特征识别

动态纹理是指描述某种动态景观的具有时间相关重复特征的图像序列^[17]。例如自然界中滚滚的麦浪,袅袅升起的烟雾,流淌的河水,随风摆动的树叶以及燃烧的火焰等都属于动态纹理特征。动态纹理与静态纹理的区别表现在:静态纹理特征表示的是单一的图像纹理特征,而动态纹理特征表示的是由一幅幅的单一图像所组成的,并在时间上具有一定的相关性的视频

纹理特征。

4.1 建立线性动态系统模型

将线性动态纹理系统模型 (linear dynamic system, LDS) 的参数解作为动态纹理特征。LDS 是 Derotto 等^[7] 在时空自回归模型的基础上提出的用来描述动态纹理变化特性的模型,将动态纹理视为含有任意协方差的高斯白噪声的二阶平稳随机过程。而含有任意协方差的高斯白噪声的二阶平稳随机过程可以建模为动态系统的输出量。自该系统模型被提出之日起,就有大量的国内外学者将其应用到烟雾识别实验中,实验表明:该模型足以表明烟雾的动态纹理特征。

LDS 模型给动态纹理一个最初的假设:一个动态纹理的每一帧图像,都是一个动态系统在某一个时刻的输出,而这一动态系统是由一个独立同分布 (independent and identically distributed, IID) 零均值高斯白噪声所驱动的随机过程。

对于一个具有 F 帧图像序列的视频和每帧含有 p 个像素点的图片,令 $\{Y(t) \in R^p\}_{t=1}^F$,则任一 t 时刻的图像 $Y(t)$,可以得到带噪声的图像为 $I(t) = Y(t) + w(t)$,其中 $w(t)$ 为噪声,且对于每一帧都符合独立同分布。在时间维度上,则使用低维度的隐马尔科夫过程 $z_t \in R^n$ 来表示观察序列 $\{I(t) \in R^p\}_{t=1}^F$ 过程的相关性。

由此该视频序列可建立如下线性动态系统模型:

z(t+1) = Az(t) + Bv(t) (4)

I(t) = Cz(t) + w(t) (5)

其中 $z(t) \in R^n$ 是时刻 t 的隐藏状态, $A \in R^{n \times n}$ 为隐藏状态的动态模型, $C \in R^{p \times n}$ 映射隐藏状态到系统的输出量, $w(t) \sim N(0, R)$ 和 $Bv(t) \sim N(0, Q)$ 分别是视频实际测量噪声和过程噪声。一般情况下, $Q = BB^T$, B 矩阵是噪声的协方差,用 v_t' 代替 $Bv(t)$ 。系统的指令 n 是隐藏状态的维度, p 是视频序列中每一帧的像素的数量。则 $M_{LDS} = (A, C)$ 表示线性动态系统的特征描述子,其主要优势在于 C 可以表示视频图片序列的显性动态特征,而 A 表示该视频序列的隐藏状态动态特征。显性动态特征一般是有规律,可预测,且容易被观测到的特征,比如烟雾在运动;而隐藏状态特征是指背景元素的运动,一般不易被观测到,且无规律性可言,从宏观上不易被发现,比如烟雾视频中属于背景点的像素的运动。

4.2 分类参数确定

给出一个视频序列,可以确定上述动态系统的参

数,例如 (C^0, A, C, Q, R) 。很多的优化方法例如 N4SID^[18] 和 EM^[19] 算法都能用于学习这些参数。但是由于帧序列的像素数代表输出量的维度,因此给计算带来很大的麻烦。

一般情况下,如果给包含水波,烟雾,旗帜等的动态纹理进行分类,则需要大量的 MLDS 特征描述子,每个 M 的维度又非常大,则计算量就会过大。如果计算量过大,就会影响系统检测的及时性,和敏感度。为了减少计算量,需要烟雾视频序列在时间和位置上都有一个很好的确定,因此文中线性动态系统输入的视频序列固定为 F 帧 32×24 的视频块。

为了构建一个零均值的视频序列,令当前视频序列像素均值为 C^0 。则动态系统模型可变为

$$z(t+1) = Az(t) + Bv(t) \quad (6)$$

$$I(t) = C^0 + Cz(t) + w(t) \quad (7)$$

其中,

$$C^0 = \frac{1}{N} \sum_{t=1}^F I(t) \quad (8)$$

得到上述改进后的动态系统模型后,则需要对模型参数进行 A, C , 等进行估计。传统的基于最大似然法和最小预测误差法的解法都存在着计算复杂或者无解析解的问题。采用 Doretto 等提出的基于 PCA (principal component analysis) 来计算系统参数的方法。

对于一个零均值的视频序列 $\{I(t)\}_{t=1}^F$, 压缩, 排序后输入系统, 即矩阵 Y , 之后对矩阵进行奇异值分解 (singular value decomposition, SVD), 如下:

$$Y = [I(1) - C^0, \dots, I(F) - C^0] = U \Sigma V^T \quad (9)$$

系统参数可以确定为如下:

$$C = U, Z = \Sigma V^T$$

其中, $Z = [z(1), z(2), \dots, z(F)]$ 是系统的估计状态。此时, Z 的计算完全不依靠式(6), 这正是子最优方法。给出状态序列 Z , 用最小平方法可以计算出矩阵 A :

$$A = [z(2), z(3), \dots, z(F)] [z(1), z(2), \dots, z(F-1)]^T \quad (10)$$

其中 Z^T 代表 Z 的伪逆, 同时:

$$Q = \frac{1}{N-1} \sum_{i=1}^{F-1} V_i (V_i)^T \quad (11)$$

其中, $V_i = Bv(t) = z(t+1) - Az(t)$ 。

至此, 整个动态系统的参数可以通过一个视频序列完全确定下来。这种动态特征和只有局部纹理模型的单一图像梯度特征是明显不一样的。其中 C 可以表示视频图片序列的显性特征, 同时, A 能够表示该视频序列的隐藏状态下的动态特征。鉴于上述采用 $M =$

(A, C) 来表示动态纹理事件的动态特征。

4.3 基于动态纹理特征的烟雾识别

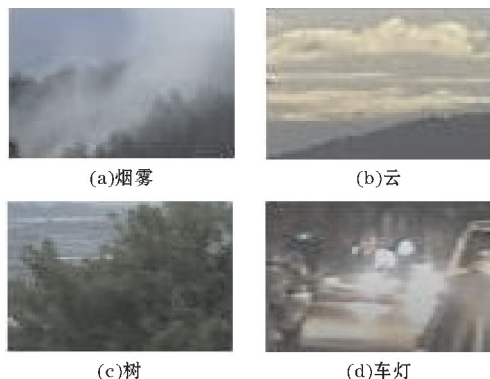


图6 动态纹理特征实验集

给出一个未知视频序列如何通过动态纹理特征识别出该序列是否为烟雾, 就要找出烟雾视频序列和非烟雾视频序列的动态纹理特征的区别。假如烟雾的动态纹理特征用 $M1 = (A1, C1)$ 表示, 非烟雾的动态纹理特征用 $M2 = (A2, C2)$ 表示, 未知视频序列的动态特征用 $M = (A, C)$ 表示, 分类识别就是判断 M 是属于 $M1$ 类的还是 $M2$ 类的。

对野外环境下的烟雾和非烟雾视频集进行分类识别, 非烟雾视频集的选取主要从烟雾视频集中的干扰因素中选取, 在野外环境下, 对烟雾的干扰最严重的主要是浮动的云, 飘动的树以及夜晚的车灯等干扰因素, 因此, 如图6, 选择浮动的云、飘动的树以及夜晚的车灯作为干扰因素分别用动态纹理特征做了实验进行验证。

图7和图8给出烟雾与云、树、车灯的隐藏状态下的动态纹理特征 A 与显性动态纹理特征 C 的 mesh 图。在 Matlab 环境下, 对一个 m 行 n 列的矩阵画 mesh 图是生成一张曲面, x 坐标的取值 $1 \sim m$, 间距为 1, y 坐标的取值是 $1 \sim n$ 构成的网络节点, 曲面的高度就是 $m \times n$ 矩阵里面的元素值, 如果 m 和 n 的值比较大, 则 mesh 图不仅能突出表现出矩阵元素的数值差异, 还能显示矩阵元素在数值分布上的变化状态。

从图8可以看到烟雾与浮动的云、飘动的树、夜晚的车灯的显性动态纹理特征矩阵都表现出一定的运动性, 相似度比较高, 但是仍然在数值上存在一定的差异。然而在图7中可以发现, 这些动态序列隐藏状态下的动态纹理特征却有着明显的不同, 烟雾的隐藏状态特征出现明显的峰值, 而其他3类非烟雾动态纹理特征虽也有峰值出现, 但变化状态却更趋于平稳。

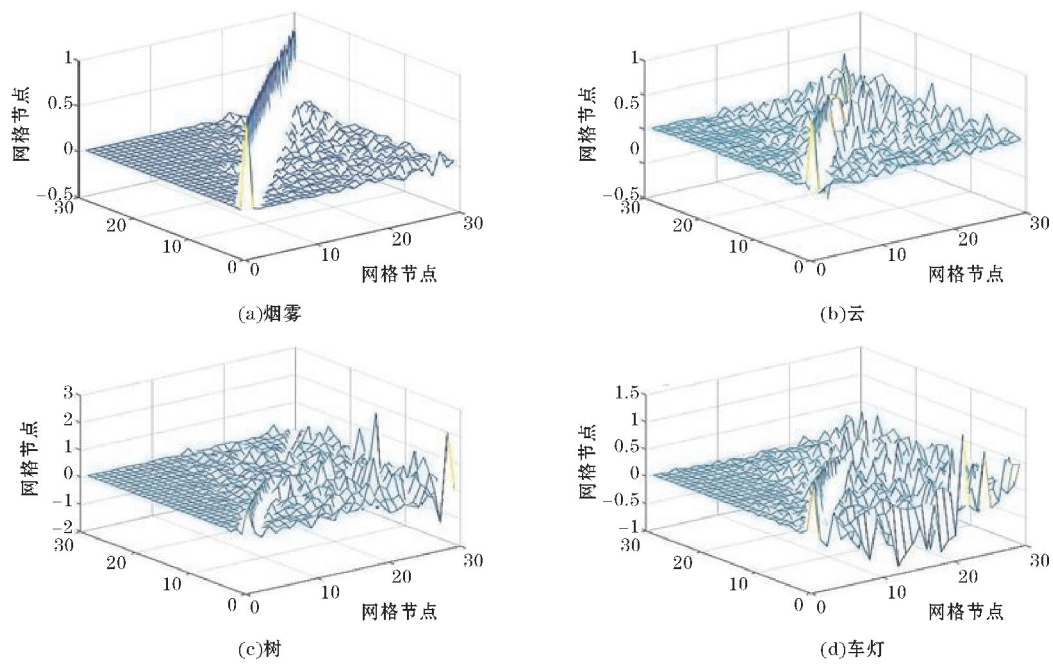


图 7 烟雾、云、树、车灯隐藏动态纹理特征 A mesh 图

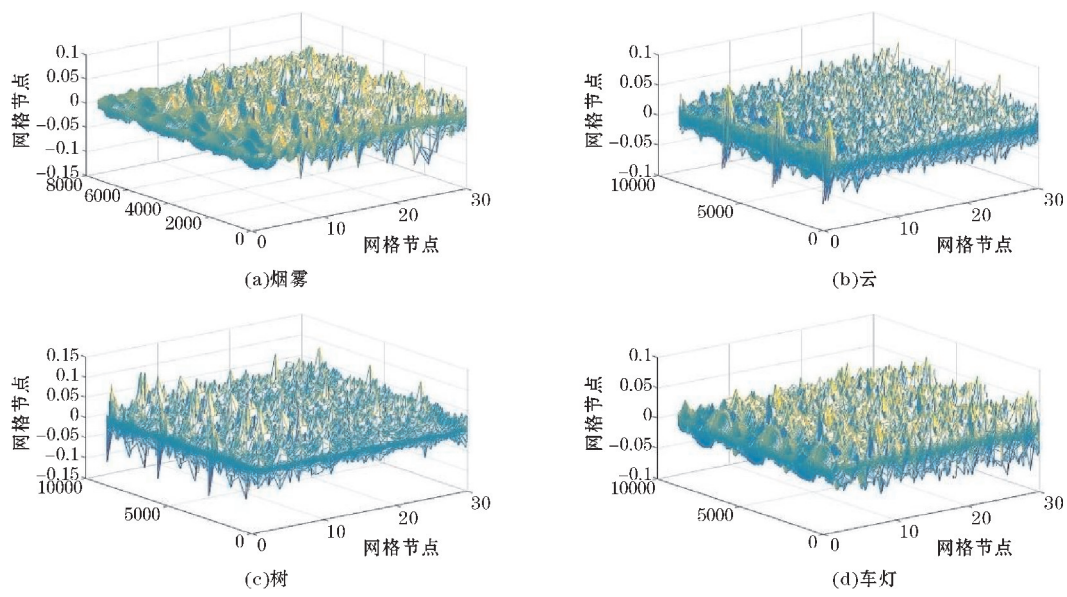


图 8 烟雾、云、树、车灯的显性动态纹理特征 C mesh 图

4.4 分类器的训练及视频帧 F 的选取

为了能够识别烟雾动态纹理特征,使用台湾大学林智仁(Lin Chih-Jen)教授等开发设计的一个简单、易于使用和快速有效的 SVM 模式识别与回归的软件包——LibSVM。LibSVM 的训练采用已经收集的烟雾集和非烟雾集进行训练,共包含视频序列块 1000 个,训练样本中烟雾和非烟雾样本分别包含已收集视频块的 80%,其余的 20%作为测试集,每个视频块的帧数为 F 。

为了确定能够更好的表达动态纹理特征的时间长度,分别选取帧数为 20,50,80,110,140 帧的烟雾视频块和非烟雾视频块进行实验验证。实验结果如表 2 所示。

从表中可以看出 50 帧的视频帧对烟雾动态纹理的描述准确率最高。帧数过少,不足以表达动态纹理的特性,帧数过多则有可能后续帧中没有烟雾的出现,过多过少都会导致烟雾准确率的识别。因此将帧数为 50 的视频集作为 LibSVM 的训练集对 LibSVM 进行训练。

表 2 视频帧数对烟雾准确率的影响

视频帧数	烟雾准确率/%
20	80.0
50	94.28
80	60.0
110	88.57
140	88.57

5 实验结果和分析

依据 CNN 对视频中早期火灾烟雾进行识别,并引入线性动态系统,进行动态纹理特征识别。CNN 的模型采用 Tensorflow 训练,网络的模型结构如图 5 所示,训练参数的选择如表 1。TensorFlow 是谷歌基于 Dist-Belief 进行研发的第二代人工智能学习系统,其被用于语音识别或图像识别等多项机器深度学习领域,另外 TensorFlow 完全开源,任何人都可以用广泛使用,因此近年来其应用越来越广。实验中 CNN 的输入图像是从烟雾数据集中按每 10 帧提取一帧的规则进行提取的图像大小为 32×24 的 5000 张图片。文中动态纹理特征的识别,使用的是线性动态系统模型,为了减少计算量,构建了一个零均值的视频序列,并引入了一种基于 PCA 线性动态系统解法。线性动态系统的输入是 50 帧 32×24 的视频片段。

5.1 评价指标

为评价实验结果的有效性。实验采用如下判定指标:

$$ACC = (TP + TN) / N$$
$$TPR = TP / (TP + FN)$$
$$TFR = FN / (FN + TN)$$

其中,TP 表示烟雾样本中被识别为烟雾的实例数;TN 表示烟雾样本中未被识别为烟雾的实例数;FP 表示非烟雾样本中被识别为烟雾的实例数;FN 表示非烟雾样本中被识别为非烟雾的实例数;ACC 表示准确率;TPR 表示真正率;TFR 表示负正率。

5.2 实验结果

单独使用 CNN 对数据集进行识别和单独使用基于 LDS 的动态纹理特征识别的结果如表 3 所示。

表 3 CNN 和 LDS 的烟雾识别结果

方法	ACC/%	TPR/%	TFR/%
CNN	93.57	94.2	92.95
LDS+SVM	94.28	92.85	93.05

在 CNN 烟雾识别部分,烟雾和非烟雾出现误检的原因主要是非烟雾帧中的干扰因素在颜色,和纹理特性上与烟雾的非常像,例如云等。而在动态纹理识别部分,出现误检的原因是很多干扰因素的动态纹理与烟雾的运动特征非常相似,例如流动的云,随风摆动的树等。

为提高提高烟雾识别的准确率,将动态纹理特征和基于 CNN 的烟雾识别方法结合起来采用混合判定的决策方法对视频帧进行烟雾的识别。如表 4,其中规定:如果基于 CNN 的烟雾识别结果和动态纹理特征识别结果都判定该视频帧为烟雾,则确定该视频帧为

烟雾序列;如果其中有某一种方法判定该视频帧为烟雾,则结合传统的 HOG 和 SVM 烟雾识别系统对该视频帧进行第二次识别,如果是,则视频帧为烟雾,否则为非烟雾;如果其都判定该视频帧为非烟雾,则确定该视频帧为非烟雾。

表 4 混合矩阵决策方法

CNN	LDS	HOG+SVM	T/F
T	T		T
F	F		F
T	F	T	T
T	F	F	F
F	T	T	T
F	T	F	F

最终结果如表 5 所示。

实验结果表明,结合基于 CNN 的烟雾识别和动态纹理特征识别混合判定决策方法,使烟雾视频帧的准确率和真正率以及负正率都有明显的提高。另外与文献[20]中准确率相比,解决了在光照弱和烟雾浓度稀薄的情况下,识别率低的问题。

表 5 实验结果与传统烟雾识别方法比较

方法	ACC/%	TPR/%	TFR/%
文中方法	95	95.65	94.36
HOG+SVM	87.5	88.4	89.63
文献[20](光照、烟雾正常)	96.57		
文献[20](光照、烟雾稀薄)	87.67		

6 结束语

提出一种结合 CNN 的烟雾识别和动态纹理特征识别的视频烟雾检测方法,创新点在于结合了传统的手动提取特征方法和 CNN 的自动提取烟雾图像特征的方法,与传统的方法相比,该方法在烟雾准确率和误检率上都有明显的提高,同时对于光照和烟雾浓度的敏感度都有极大的降低。但是为了减少计算量,动态纹理的提取舍弃了大部分的动态纹理特征,同时对 LDS 的参数解法对非线性的分类效果并不是很好,因而下一步将对 LDS 作更深入的研究,并考虑采用码本的形式加入更多的动态纹理特性。

参考文献:

[1] 刘嘉,汪则灵. 浅谈火灾危害与环境保护[J]. 经营者,2010,17:180-181.

[2] etin A E, Dimitropoulos K, Gouverneur B, et al. Video fire detection-Review[J]. Digital Signal Processing,2013,23(6):1827-1843.

[3] Ye Wei, Zhao J, Wang S, et al. Dynamic texture

- based smoke detection using Surfacelet transform and HMT model[J]. Fire Safety Journal. 2015, 73: 91–101.
- [4] Chunyu Yu, Zhibin Mei, Xi Zhang. A real-time video fire flame and smoke detection algorithm[J]. Procedia Engineering. 2013, 62: 891–898.
- [5] 胡燕, 王慧琴, 姚太伟, 等. 基于 Harris 特征点检测与跟踪的火灾烟雾识别[J]. 计算机工程与应用, 2014, 50(21): 180–183.
- [6] ByoungChul Ko, JunOh Park, Jae-Yeal Nam. Spatiotemporal bag of features for early wild fire smoke detection[J]. Image and Vision Computing, 2013, 31: 786–795.
- [7] Doretto G, Chiuso A, Wu Y N, et al. Dynamic Textures. [J]. International Journal of Computer Vision, 2003, 51(2): 91–109.
- [8] Coviello E, Mumtaz A, Chan A B, et al. Growing a bag of systems tree for fast and accurate classification[C]. Computer Vision and Pattern Recognition. IEEE, 2012: 1979–1986.
- [9] Ravichandran A, Chaudhry R, Vidal R. Categorizing Dynamic Textures Using a Bag of Dynamical Systems[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2013, 35(2): 342–353.
- [10] Mumtaz A, Coviello E, Lanckriet G R G, et al. A Scalable and Accurate Descriptor for Dynamic Textures Using Bag of System Trees[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 37(4): 697–712.
- [11] Barmoutis P, Dimitropoulos K, Grammalidis N. Smoke detection using spatiotemporal analysis, motion modeling and dynamic texture recognition[C]. Signal Processing Conference. IEEE, 2014: 1078–1082.
- [12] Dimitropoulos, Kosmas, Barmoutis, Panagiotis, et al. Nikos. Higher order linear dynamical systems for smoke detection in video surveillance applications[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 26(12): 1–13.
- [13] Chongyuan Tao, Jian Zhang, Pan Wang, Smoke detection based on deep convolutional neural networks[C]. IECON 2016–42nd Annual Conference of the IEEE Industrial Electronics Society, 2016: 150–153.
- [14] 陈俊周, 汪子杰, 陈洪瀚, 等. 基于级联卷积神经网络的视频动态烟雾检测[J]. 电子科技大学学报, 2016(6): 992–996.
- [15] Hinton G E, Srivastava N, Krizhevsky A, et al. Improving neural networks by preventing coadaptation of feature detectors[J]. Computer Science, 2012, 3(4): 212–223.
- [16] Kingma D P, Ba J. Adam: A Method for Stochastic Optimization[J]. Computer Science, 2014.
- [17] 朱旭阳, 李思昆. 动态纹理合成技术研究综述[J]. 系统仿真学报, 2001(S2): (23–25)+38.
- [18] P V Overschee, B D Moor. N4SID: Subspace Algorithms for the Identification of Combined Deterministic-Stochastic Systems[J]. Automatica, 1994, 30: 75–93.
- [19] R. Shumway and D. Stoffer. An Approach to Time Series Smoothing and Forecasting Using the EM Algorithm[J]. Time Series Analysis. 1982, 3(4): 253–264.
- [20] 唐杰, 周洋, 杨萌, 等. 采用颜色混合模型和特征组合的视频烟雾检测[J]. 光电子·激光, 2017, 28(7): 751–758.

Research for Early Wildfire Smoke Video Detection

GAO Feng-wei, WEI Wei, CHENG Yang

(College of Computer Science and Technology, Chengdu University of Information Technology, Chengdu 610225, China)

Abstract: A smoke recognition method of deep learning and dynamic texture features based on parallel for the early fire smoke video's blurry shape and the color contrast is not obvious compared with the background, and the appearance characteristics are varied. In the experiment, using the method of mean filter for filtering the video frame. And gives a convolutional neural network structure for early fire smoke based on video, then training and testing of the structure. For the dynamic texture features of smoke, making a linear dynamic system model and detecting the video frames based on support vector machine. The smoke is proposed based on a judgment method of the mixing matrix convolution neural network recognition results and dynamic texture recognition results. The experimental results show that this method compared with the traditional smoke recognition method has higher accuracy in the detection of early fire smoke.

Keywords: video smoke detection; Convolutional Neural Network(CNN); Linear Dynamic Systems(LDS); dynamic textures