

空-频域联合投票的交通视频阴影去除方法

宋传鸣 洪旭 王相海

辽宁师范大学计算机与信息技术学院 辽宁 大连 116029



摘要 交通场景中的静止或运动阴影往往会降低车辆目标跟踪的精度,因此有效地去除阴影是交通监控视频处理的重要环节之一。然而,目前尚无一种能够同时发掘阴影的空间域和频率域特性且抵抗静止和运动阴影干扰的阴影去除方法。为此,提出了一种基于空-频域联合投票策略的交通视频阴影去除方法。首先,将视频帧从 RGB 颜色空间转换到 HSV 颜色空间,再进行非下采样剪切波变换;其次,假设变换系数服从高斯分布,采用变换系数的均值和标准差计算每个尺度的加权掩码;然后,根据多尺度变换系数的零树分布特性,利用粗尺度的加权掩码校正细尺度的加权掩码,将各个尺度、各个颜色通道的加权掩码进行线性组合后得到一个公共掩码,再采用基于最小二乘法拟合的最大熵方法计算自适应分割阈值,对公共掩码进行二值化;最后,联合频率域加权掩码、S 通道和 V 通道的掩码进行投票,进而确定去除阴影后的运动车辆区域。实验结果表明,该算法可有效去除交通监控视频中的静态/运动阴影,抑制阴影的干扰,将传统 Meanshift 算法的输出车辆轨迹与真实轨迹间的平均欧氏距离缩小 95%,且未出现目标丢失的现象,增强了智能分析算法的鲁棒性。研究结果说明,该算法有效联合交通监控视频的空间域和频率域表示,充分发掘了运动车辆区域与阴影区域之间的纹理特性和颜色特性差异,有利于获得更精确的阴影去除结果,进而提高车辆目标跟踪的精度。

关键词: 交通监控视频;阴影去除;非下采样剪切波;联合投票;多尺度加权掩码

中图分类号 TP399

Shadow Removal of Traffic Surveillance Video by Joint Voting in Spatial-Frequency Domain

SONG Chuan-ming, HONG Xu and WANG Xiang-hai

School of Computer and Information Technology, Liaoning Normal University, Dalian, Liaoning 116029, China

Abstract The static or moving shadows in traffic scenes tend to reduce the accuracy of vehicle target tracking. Thus, it is an important step to effectively remove the shadows in the processing of traffic surveillance videos. However, there hardly is an efficient shadow removal method yet at present, which resists both static and moving shadows by simultaneously exploring the spatial and frequency characteristics of shadows. Under such a circumstance, this study proposed a shadow removal method for traffic video by using a joint voting strategy in spatial-frequency domain. The surveillance video is converted from RGB space into HSV space and then performed non-subsampled shearlet transform (NSST). Assuming that NSST coefficients follows the Gaussian distribution, the mean and standard deviation of coefficients is used to compute the weighted mask for each scale. Subsequently, the weighted mask at coarse scale is employed to adjust the mask at fine scale, according to the zerotree characteristics of multi-scale coefficients. The weighted masks of different scales and color channels are thus linearly combined to form a unified mask, which is then binarized by an adaptive threshold calculated by the maximum entropy segmentation based on the least square method. Finally, the moving vehicle area after shadow removal is determined by a joint voting strategy by using the weighted frequency-domain mask, the S-channel mask and the V-channel mask respectively. Experimental results show that the proposed algorithm can effectively remove the static and moving shadows in traffic surveillance video. It reduces the average Euclidean distance by 95% between the ideal trajectory and the output vehicle trajectory of traditional mean shift algorithm, suppressing the interference of shadows. Meanwhile, the proposed algorithm enhances the robustness of the intelligent analysis and avoids the phenomenon of losing the target. Our research indicates that it is conducive to obtaining more accurate shadow removal result to effectively combine the representation of traffic surveillance video in spatial and frequency domains, and to fully explore the differences of texture features and color features between moving vehicle areas and shadow areas. The accuracy of vehicle target tracking will be therefore improved.

Keywords Traffic surveillance video, Shadow removal, Non-subsampled Shearlet transform, Joint voting, Multiscale weighted mask

收稿日期:2019-04-08 返修日期:2019-08-31 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:大连市高层次人才创新支持计划项目(2015R069)

This work was supported by the Dalian Foundation for Youth Science and Technology Star(2015R069).

通信作者:宋传鸣(chmsong@lnnu.edu.cn)

1 引言

在光照稳定、前景单一、背景静止、摄像头位置与焦距固定等理想的采集条件下,基于监控视频的智能交通系统通过典型的检测和识别算法即可取得令人满意的处理和分析效果。然而,在真实的交通场景下,采集设备的拍摄环境往往较为复杂,所获取的视频质量难以达到理想水平,从而降低了基于监控视频的智能交通系统的可靠性。在这种情况下,Toyama等^[1]和Bouwmans等^[2]分别提出了视频监控领域的十大挑战及运动目标检测的13个难点^[3],其中均涉及阴影对交通监控视频处理效率的负面影响。因此,有效地去除阴影是交通监控视频处理的重要环节之一。尽管国内外学者已经提出了多种阴影检测与去除的算法,但是目前尚无一种能够同时抵抗静止或运动阴影干扰且无需人工交互的阴影去除方法。

为此,本文在分析阴影的空间域和频率域特性的基础上,提出了一种空-频域联合投票的交通视频阴影去除方法。该方法首先将视频帧进行颜色空间转换和非下采样剪切波变换,计算出不同尺度下各个方向子带的二值掩码,并根据多尺度变换系数的父子间传递性的特点,利用粗尺度掩码校正细尺度掩码,再将其线性组合得出多尺度加权掩码;其次,通过一种基于最小二乘拟合的最大熵分割算法来处理多尺度加权掩码,获得阴影区域和车辆区域的粗分割结果;最后,根据阴影的亮度、饱和度特点,对频率域的粗分割结果与饱和度、亮度通道的分割结果进行投票决策,求得最终的阴影去除效果。实验结果验证了本文算法的有效性。

2 相关工作

根据阴影的特点,研究人员从算法的决策过程、采用的模型和依据的特征等方面提出了3类典型的阴影检测和去除方法:基于彩色空间像素点的方法、基于边缘检测的方法、基于灰度空间轮廓线的方法。

基于彩色空间像素点的方法主要利用阴影的亮度和颜色特性,分析前景与背景像素的强度比,再在HSV颜色空间结合多个阈值来进行判断。该类方法对不同强度的阴影的处理能力较强。Cucchiara等^[4]利用多阈值对HSV颜色空间中前景和背景像素的强度比进行判断,所提出的DNM1方法在一定程度上抑制了阴影,但由于涉及到多个阈值,其难以做到对不同环境的自适应性;并且当目标区域与阴影区域具有相似的颜色和灰度值时,该方法无法区分与阴影具有相似灰度值的运动目标,容易造成误判。Deng等^[5]先将视频帧转换到YCbCr颜色空间,再利用相应的亮度和色度自适应阈值来去除阴影。但是,该算法所需阈值过多,自适应能力较差。Choi等^[6]提出将归一化的RGB与1阶梯度信息相结合来判断阴影,在一定程度上降低了误判率。Xiang等^[7]利用基于光照建模的局部强度比模型来改善跟踪算法对光照条件改变的鲁棒性。Ouvirach等^[8]利用高斯混合模型在HSV颜色空间提取运动前景,然后利用最大似然方法判断提取的运动前景像素属于目标还是阴影。该方法有效提升了检测效果,但是依然存在大量的误判,且计算量较大。与Ouvirach等的方法类似,Liu等^[9]也利用高斯混合模型对HSV空间的每个像素进行投影建模,为降低误判率,他们引进了基于马尔可夫随机场

(Markov Random Field, MRF)的预分类器来提取视频帧中阴影的颜色特征,并且对连续多帧阴影的特征进行统计,从而保证了预分类器能够有效地适应阴影变化,取得了不错的效果。但是,当训练样本无法匹配阴影的变化速度,即车辆与阴影的相对运动较快时,全局阴影统计特征将不再可信,误判率也随之上升。Salvador等^[10]证明了阴影区域与背景区域在颜色构成方面具有一致性特性,进而提出了一种基于颜色特征的阴影去除方法。Chen等^[11]根据颜色特征、归一化向量距离分布直方图来检测并消除阴影像素,并利用亮度比值置信区间还原误判的前景像素,提出了一种亮度和色度特征相融合的运动阴影去除算法。总体上看,该方法属于基于单个特征的方法,并未考虑阴影的纹理信息,其去除阴影的效果仍然不够理想。

由于阴影区域较平滑,而车辆目标往往包含一定纹理和边缘信息,因此基于边缘检测的方法和基于灰度空间轮廓线的方法从纹理和边缘信息的角度来分析和检测阴影,进而抑制阴影对目标跟踪过程的影响。Tian等^[12]借助基于互相关的纹理分析方法,通过对比当前帧和背景模型处于相同位置的像素与其邻域像素之间的纹理相似度,提出了一种基于纹理信息归一化互相关的方法来缓解阴影干扰,具有一定的鲁棒性。考虑到小波变换的边缘方向分析能力,Guan等^[13]选取各个小波子带的标准差作为阈值,在一定程度上去除了传统背景差分方法产生的阴影;Khare等^[14]进一步采用相对标准差作为阈值来处理小波子带。Wang等^[15]根据小波系数的尺度间关系,通过构造运动前景的零树小波掩码,并采用粗尺度掩码校正细尺度掩码,获得了较为准确的子带阈值和阴影检测结果。Xu等^[16]则将高斯模型和决策树分类器集成到条件随机场中,用决策树分类器识别阴影,再用高斯模型去除阴影,从而提出了一种基于学习的单幅单色图像阴影检测和去除算法,其优势在于使用的参数数量较少,适用于单色图像阴影的去除。与基于灰度或颜色的方法相比,上述方法利用纹理、边缘的特性来抵抗阴影的影响,取得了一定的效果。然而,这些挖掘阴影纹理特性的检测方式多数需要输入背景帧来获取图像的阴影与前景信息,因此去除静态阴影的效率有限;并且这些方法在提取纹理、边缘的过程中所采用的小波变换仅能捕获沿着水平、竖直和对角线方向的直线奇异对边缘和纹理的刻画尚有不足,进而制约并影响了阴影去除的准确率。

综合两方面的工作来看,目前尚无一种能够同时抵抗静止和运动阴影干扰的阴影去除方法。

3 阴影的性质与影响

3.1 阴影的性质

本节从形成阴影的物理机制出发,分析其空间域的亮度、颜色、纹理特性和频率域的频谱分布。

尽管阴影的表现有多种情形,但是大部分阴影均具有以下共同特性^[16-17]。

1)阴影的亮度低于前景的亮度。如图1所示,在整个平面上只有小球正前方的本影区域没有受到光源照射,并且其周围一定范围内呈灰暗颜色,因此阴影区域拥有相对较低的亮度。

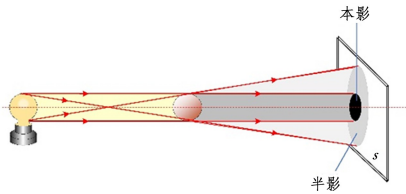


图1 阴影示意图

Fig.1 Diagram of shadow

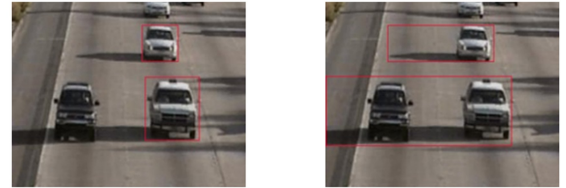
2)处于阴影区域的背景边缘、色调、纹理等特征一般不会被改变。一方面,无论是环境光还是光源,本身一般不具有结构性边缘和纹理。另一方面,阴影是由不透明物体阻碍光源沿直线传播所产生的,物体表面的反射光无法进入阴影区。因此,阴影不会受该物体表面反射光的影响,阴影的边缘、纹理、色调仅取决于其所在背景区的边缘、纹理和色调,而与不透明物体的表面属性无关。从这个意义上讲,经过彩色空间转换后,阴影区域的色彩特征一般不会受到影响^[11]。

3)从频率域的角度而言,阴影一般属于低频。由于反射光线被物体遮挡,阴影区域的亮度仅由周围环境的辐射光强度决定,因此阴影区域的亮度范围将被明显压缩,其亮度的梯度值亦相应减小。鉴于这一点,阴影区的能量在频率域中应主要集中在低频分量上^[15]。

3.2 阴影对交通监控视频处理的影响

在交通监控视频中,物体所产生的阴影一般会与车辆重叠或连在一起,使得阴影与车辆拥有相似的运动特征信息。

然而,运动特征是车辆跟踪算法所考查的主要因素,这就不可避免地导致跟踪算法难以区分车辆与阴影,提高了阴影被误检为目标车辆的概率。为证明这一点,图2给出了利用文献[18]的方法对运动车辆进行检测跟踪的结果。显然,该算法错误地将阴影判定为运动目标车辆的一部分,降低了车辆检测的精度。此外,当多个车辆并行行驶时,若车辆的阴影之间存在重叠或连接,运动目标检测和跟踪算法也可能将不同车辆误判为同一个目标。



(a)预期提取的目标区域

(b)实际提取的目标区域

图2 阴影对运动车辆跟踪算法的不利影响

Fig.2 Negative effects of shadows on vehicle tracking

可见,静止和运动阴影均会对基于监控视频的智能交通系统的处理精度造成不利影响。

4 空-频联合的交通视频阴影去除算法的总体框架

为兼顾阴影的亮度特性和边缘、纹理等空间域和频率域特性,本文在3.1节的基础上提出了一种空-频域联合投票的交通视频阴影去除方法,其总体流程如图3所示。

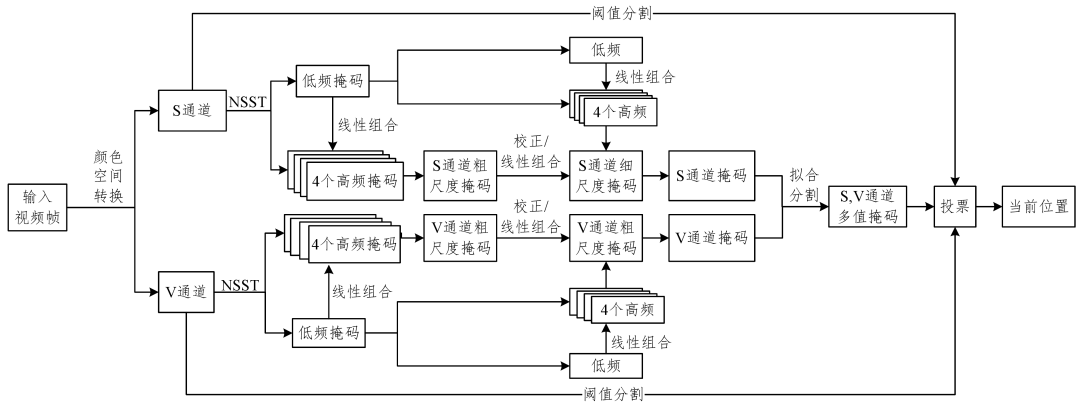


图3 空-频联合的交通视频阴影去除算法的总体流程图

Fig.3 General flowchart of shadow removal algorithm of traffic surveillance video by joint spatial-frequency domain

5 空-频联合的交通视频阴影去除方法

5.1 基于非下采样剪切波变换的零树结构掩码

现有算法大多根据阴影和运动前景的时空域分布特性,利用一定的阈值实现阴影区域的检测和去除。但是,单纯发掘颜色属性的方法易受到不同环境的干扰,尤其在去除监控视频的动态阴影时,自适应阈值的选取存在一定困难。由第3节的分析可知,阴影信息在频率域上分布在低频部分。若能充分考虑这一特点,则可在一定程度上抑制环境噪声的影响。鉴于此,本节讨论一种以低频信息为指导的、逐尺度求精的方法来构造运动前景的掩码,进而去除交通监控视频中的阴影。

鉴于小波变换的局限性,本文选用非下采样剪切波对视频帧F的S通道和V通道分别进行2级非下采样剪切波变换,以获得其多尺度的频域表示,进而得到F的1个低频分量

和8个高频方向分量。

利用式(1)计算最低频子带的二值掩码 $M_{LL}^c(x, y)$:

$$M_{LL}^c(x, y) = \begin{cases} 1, & X_{LL}^c(x, y) > \mu_{LL}^c \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

其中, $X_{LL}^c(x, y)$ 表示在颜色通道 c 的最低频子带中坐标 (x, y) 处的变换系数, μ_{LL}^c 为S通道和V通道的最低频子带系数的均值,上标 c 表示颜色通道且 $c \in \{S, V\}$ 。

对于各尺度下的所有高频方向子带,利用式(2)计算其二值掩码 $M_{i,j}^c(x, y)$:

$$M_{i,j}^c(x, y) = \begin{cases} 1, & |X_{i,j}^c(x, y)| > \sigma_{i,j}^c \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

其中, $X_{i,j}^c(x, y)$ 表示颜色通道 c 的第 i 个尺度下第 j 个方向子带中坐标 (x, y) 处的变换系数, $\sigma_{i,j}^c$ 为S通道与V通道的所有高频方向子带系数的标准差,下标 i 表示尺度且 $i \in \{1, 2\}$,

下标 j 表示方向且 $j \in \{1, 2, 3, 4\}$ 。

进一步,考虑到同一尺度下的各个子带具有相同的频带宽度,它们对同一场景往往表现出一致的幅频响应,本文将同一尺度下所有子带的二值掩码进行线性加权,具体过程如式(3)和式(4)所示:

$$\mathbf{M}_1^c(x, y) = 0.6 \times \mathbf{M}_{1L}^c(x, y) + 0.1 \times \sum_{j=1}^4 \mathbf{M}_{1,j}^c(x, y) \quad (3)$$

$$\mathbf{M}_2^c(x, y) = 0.6 \times \mathbf{M}_1^c(x, y) + 0.1 \times \sum_{j=1}^4 \mathbf{M}_{2,j}^c(x, y) \quad (4)$$

其中, $\mathbf{M}_{1,j}^c(x, y)$ 表示颜色通道 c 的第 1 个尺度、第 j 个子带中,坐标 (x, y) 处的加权掩码; $\mathbf{M}_{2,j}^c(x, y)$ 表示颜色通道 c 的第 2 个尺度、第 j 个子带中,坐标 (x, y) 处的加权掩码。因为粗尺度反映图像的低频信息,细尺度反映图像的高频信息,所以粗尺度的加权掩码 $\mathbf{M}_{1,j}^c(x, y)$ 可稀疏表示阴影区域和运动前景的主要轮廓,而细尺度的加权掩码 $\mathbf{M}_{2,j}^c(x, y)$ 有利于捕获阴影区域和运动前景的细腻边缘。

图 4 为一幅交通视频帧的 V 分量图像在不同尺度下的加权掩码。从中可见, $\mathbf{M}_{1,j}^c(x, y)$ 较为稳定,但相对粗糙;而 $\mathbf{M}_{2,j}^c(x, y)$ 较为精细,但易被噪声干扰。鉴于这两种掩码的特点,有必要将二者进行有机组合,从而获得关于阴影和运动前景更全面的判断。

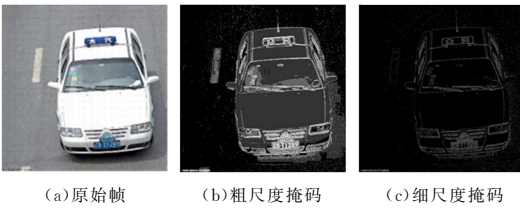


图 4 不同尺度下的加权掩码

Fig. 4 Weighted mask under different scales

为了融合不同尺度的加权掩码,本文引进了 NSST 变换系数的尺度间相关性,利用粗尺度掩码来校正与指导细尺度掩码,从而抑制噪声的干扰。对于某个非下采样剪切波变换系数 X ,文献[19]将其相邻较粗尺度下处于相同方向子带、相同位置的系数称为父系数,记为 PX (如图 5 所示),即称 PX 是 X 的父系数,而 X 是 PX 的子系数。

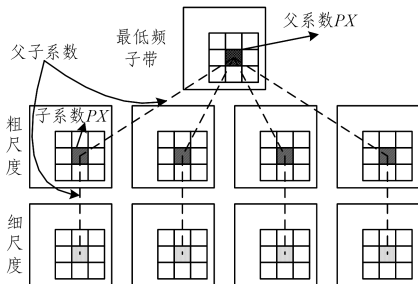


图 5 NSST 变换系数的父子关系示意图

Fig. 5 Diagram of parent-child relationship among NSST coefficients

同时,约定最低频子带的每个系数是与其同一尺度下各个高频方向子带中处在相同空间位置的系数的父系数。从理论上讲,父系数的支撑区间一般是子系数支撑区间的 4 倍。如果父系数的幅值较小,说明图像在较大范围内均未产生明显的像素值突变,所以在较小的子系数支撑区间内,自然也不会有较大梯度的变化。

为了证明父子系数的尺度间相关性,本文统计了 2 帧交

通监控视频受阴影污染退化的图像经过 2 级 NSST 变换后得到的父子系数条件概率 $p(X|PX)$,如图 6 所示。

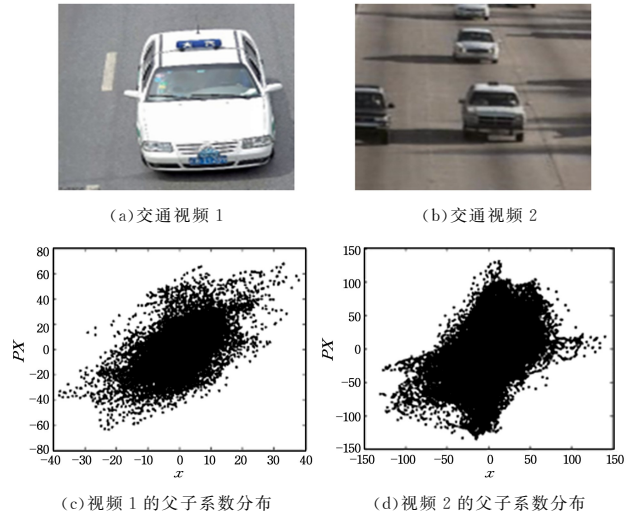


图 6 NSST 父子系数的条件概率分布图

Fig. 6 Conditional probability distribution of NSST's parent-child coefficients

不难发现,图中的条件概率总体上沿着 45° 方向分布,说明非下采样剪切波变换的父子系数之间具有较明显的线性相关性。综合理论和实验角度的分析,当父系数为“小”系数时,其子系数大概率属于“小”系数。NSST 的这种父子系数相关性被称为零树。

根据 NSST 变换系数的零树特性,若某个粗尺度系数的加权掩码 $\mathbf{M}_1^c(x, y) = 0$,则由式(1)、式(2)可知,该系数的幅值小于子带均值或标准差(即“小”系数),此时相应的细尺度系数的加权掩码 $\mathbf{M}_2^c(x, y)$ 也应该为 0。若该细尺度系数的加权掩码 $\mathbf{M}_2^c(x, y) = 1$,意味着该高频系数很可能受到了噪声的干扰,出现了伪边缘响应。据此思路,本文利用粗尺度的加权掩码 $\mathbf{M}_1^c(x, y)$ 来校正较细尺度的加权掩码 $\mathbf{M}_2^c(x, y)$:若粗尺度的加权掩码 $\mathbf{M}_1^c(x, y) = 0$,则直接将细尺度的加权掩码 $\mathbf{M}_2^c(x, y)$ 清零;若粗尺度的加权掩码 $\mathbf{M}_1^c(x, y) = 1$,则保持 $\mathbf{M}_2^c(x, y)$ 不变。

经过上述校正,粗、细尺度的加权掩码在不同频率范围内对运动前景做出了初步决策,并且零树结构保证了这两个决策的内在一致性。接下来,进一步将粗尺度和细尺度的加权掩码进行加权组合,得到两个尺度下的统一掩码 $\mathbf{M}^c(x, y)$,从而获得运动前景的一个更全面的、多频率的判断。该过程如式(5)所示:

$$\mathbf{M}^c(x, y) = w_1 \times \mathbf{M}_1^c(x, y) + w_2 \times \mathbf{M}_2^c(x, y) \quad (5)$$

其中,权重系数 w_1 和 w_2 可通过实验确定,本文将分别设置为 0.75 和 0.25。

图 7 给出了图 4(a) 的 V 分量的统一掩码。



图 7 图 4(a) 的统一掩码

Fig. 7 Nnified mask of Fig. 4(a)

类似地,还可以计算出S分量的统一掩码。为了联合使用两个通道的统一掩码去除阴影,本文将V通道和S通道的统一掩码也进行加权求和,进而得到两个颜色通道的公共掩码 $M^{SV}(x,y)$:

$$M^{SV}(x,y) = 0.5 \times M^S(x,y) + 0.5 \times M^V(x,y) \quad (6)$$

其中, $M^S(x,y)$ 和 $M^V(x,y)$ 分别表示S通道和V通道在坐标 (x,y) 处的统一掩码。

5.2 基于最小二乘拟合的自适应分割

视频帧的公共掩码 $M^{SV}(x,y)$ 在本质上是度量每个像素属于运动前景的概率。 $M^{SV}(x,y)$ 愈接近 1, 说明 (x,y) 处的像素愈可能属于运动前景; 反之, 说明该像素愈可能属于背景或阴影。不过, 与这种概率式的描述相比, 车辆跟踪等智能分析算法更加关心各个像素究竟属于运动前景还是背景和阴影。因此, 我们必须根据公共掩码 $M^{SV}(x,y)$ 做出一种二值化的判断, 将像素集合划分为运动前景和背景, 这等价于一个分割问题。与典型的图像分割不同, $M^{SV}(x,y)$ 的值域是 $[0, 1]$ 之间的实数集, 无法利用最大类间方差法(即大津法)、最大熵阈值分割法等传统阈值法实现分割。其根本原因在于, 这些方法均须在 256 个灰度级中以枚举的方式找出一个最优的灰度级, 使之取得最大的类间方差或信息熵。而由实数连续统计理论可知, $[0, 1]$ 区间存在无穷多个实数, 欲枚举出每个实数对应的类间方差或信息熵是完全不可能的。

为了确定一个实数阈值 $T \in [0, 1]$, 对于每个像素 (x,y) , 若其公共掩膜 $M^{SV}(x,y) < T$, 则令其二值公共掩码 $M_{\text{binary}}^{SV}(x,y) = 0$; 否则, 令 $M_{\text{binary}}^{SV}(x,y) = 1$ 。显然, 可将掩码 $M_{\text{binary}}^{SV}(x,y)$ 视为一个二元离散信源。根据信息熵的性质, 该离散信源的熵服从一条上凸的 2 次曲线 $H(T) = a_2 T^2 + a_1 T + a_0$, 且存在最大值。只要能够准确地拟合出这条曲线, 就可由 2 次函数的韦达定理计算出 $H(T)$ 取得最大值时的阈值 T 。根据该思路, 本文提出一种基于最小二乘拟合的自适应分割方法。

首先, 以 0.1 为一个 bin, 将公共掩码 $M^{SV}(x,y)$ 的值域等分成 10 个 bin, 即 $[0, 0.1)$, $[0.1, 0.2)$, $[0.2, 0.3)$, $[0.3, 0.4)$, $[0.4, 0.5)$, $[0.5, 0.6)$, $[0.6, 0.7)$, $[0.7, 0.8)$, $[0.8, 0.9)$, $[0.9, 1.0]$, 并统计所有像素的掩码 $M^{SV}(x,y)$ 处于各区间的频率 $P(m)$, 从而建立 $M^{SV}(x,y)$ 的直方图。

其次, 以 $t/10$ 作为全局阈值, 计算将 $M^{SV}(x,y)$ 的各像素划分为前景像素或背景像素的信息熵:

$$P_b = \sum_{m=1}^t P(m)$$

$$P_o = \sum_{m=t+1}^{10} P(m)$$

$$H\left(\frac{t}{10}\right) = - \sum_{m=1}^t \frac{P(m)}{P_b} \log_2 \frac{P(m)}{P_b} - \sum_{m=t+1}^{10} \frac{P(m)}{P_o} \log_2 \frac{P(m)}{P_o} \quad (7)$$

其中, $t \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ 。

然后, 将求得的 $H(0.1), H(0.2), \dots, H(1.0)$ 代入式(8), 利用最小二乘法对信息熵曲线 $H(T)$ 进行拟合, 得到该一元 2 次曲线的 3 个系数 a_2, a_1 和 a_0 :

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum_{m=1}^{10} 1 & \sum_{m=1}^{10} m & \sum_{m=1}^{10} m^2 \\ \sum_{m=1}^{10} m & \sum_{m=1}^{10} m^2 & \sum_{m=1}^{10} m^3 \\ \sum_{m=1}^{10} m^2 & \sum_{m=1}^{10} m^3 & \sum_{m=1}^{10} m^4 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{m=1}^{10} H(m) \\ \sum_{m=1}^{10} mH(m) \\ \sum_{m=1}^{10} m^2 H(m) \end{bmatrix} \quad (8)$$

最后, 令 $T = -a_1/20a_2$ 作为全局阈值, 将公共掩码 $M^{SV}(x,y)$ 进行阈值化, 得到二值公共掩码 $M_{\text{binary}}^{SV}(x,y)$:

$$M_{\text{binary}}^{SV}(x,y) = \begin{cases} 1, & M^{SV}(x,y) > T \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

我们以图 4 的交通监控视频帧为例, 采用上述方法对它的公共掩码(见图 7)进行自适应阈值分割, 结果如图 8 所示。



图 8 公共掩码的自适应阈值分割结果

Fig. 8 Adaptive thresholding result of unified mask

从图可见, 原始帧中的阴影已被有效去除。但是由于公共掩码是通过 NSST 变换的低频分量在零树结构的约束下所获得的, 在去除阴影的同时, 也容易损失车辆目标的部分细节特征, 进而对后续的视频分析和处理产生不利影响。

5.3 基于投票策略的多通道掩码融合

考虑到空间域的像素能够比低频子带更完整地保留视频帧的细节特征, 我们将视频帧的 S 通道及 V 通道进行阈值化, 如图 9 所示。



(a) S 通道的分割结果



(b) V 通道的分割结果

图 9 多通道阈值分割结果

Fig. 9 Results of multiple channel thresholding

可以发现, 一方面, S 通道和 V 通道的分割结果确实具有更多的细节信息, 如车牌号、车徽、车顶标牌等, 这表明空间域的分割结果可以作为频率域分割结果的有效补充; 另一方面, V 通道与 S 通道相比, 前者仍存在许多零散的噪声点(见图 9(b)), 后者则几乎不存在噪声, 但是个别特征出现损失(如车灯), 对车辆目标完整度的保持尚不够令人满意。鉴于这一点, 本文提出一种基于投票策略的多通道掩码融合方法来确定运动前景的区域, 以在去除阴影的同时, 通过多个颜色通道、空域和频域相互联合的方式, 最大程度地保留车辆目标的特征, 从而提高二值公共掩码 $M_{\text{binary}}^{SV}(x,y)$ 对运动前景的判别精度。

首先, 利用最大类间方差法计算视频帧的 S 通道和 V 通道的二值掩码 $M_{\text{binary}}^S(x,y)$ 和 $M_{\text{binary}}^V(x,y)$ 。

其次, 对于每个像素 (x,y) , 若二值公共掩码 $M_{\text{binary}}^{SV}(x,y) = 1$, 则判定该像素 (x,y) 为运动前景; 若二值公共掩码 $M_{\text{binary}}^{SV}(x,y) = 0$, 则只有当 S 通道和 V 通道的二值掩码均提示为运动前景时, 才将该像素 (x,y) 判定为运动前景, 从而避免引入 V 通道的噪声; 否则, 将该像素 (x,y) 判定为阴影或静止的背景。具体计算过程如式(10)所示:

$$M'_{\text{binary}}(x, y) = \begin{cases} 1, & M_{\text{binary}}^{\text{SV}}(x, y) = 1 \\ 1, & M_{\text{binary}}^{\text{SV}}(x, y) = 0 \text{ 并且 } M_{\text{binary}}^{\text{S}}(x, y) = 1 \\ M_{\text{binary}}^{\text{V}}(x, y) = 1 \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

最后,将所得的二值掩码 $M'_{\text{binary}}(x, y)$ 作为最终的运动前景分析结果。

5.4 算法的详细步骤

在上述分析的基础上,本文提出一种空-频域联合投票的交通监控视频阴影去除方法,其详细步骤如下。

步骤 1 输入一个含有阴影的交通监控视频 V_l 。

步骤 2 从 V_l 中读入一个尺寸为 $W \times H$ 像素的未处理视频帧 F , 并将其从 RGB 颜色空间转换到 HSV 颜色空间。

步骤 3 对视频帧 F 的 S 通道及 V 通道分别进行 2 级非下采样剪切波变换,各个尺度下的方向子带个数均为 4。

步骤 4 计算 S 通道和 V 通道的最低频子带系数的均值 μ_{LL}^c 。

步骤 5 计算 S 通道和 V 通道的不同尺度、不同方向子带中高频系数的标准差 $\sigma_{i,j}$ 。

步骤 6 根据式(1)计算最低频子带的二值掩码 $M_{\text{LL}}^c(x, y)$ 。

步骤 7 根据式(2)计算各个高频方向子带的二值掩码 $M_{i,j}^c(x, y)$ 。

步骤 8 根据式(3),为尺度 1 下的所有子带计算一个加权掩码 $M_{1,j}^c(x, y)$ 。

步骤 9 根据式(4),为尺度 2 下的所有子带计算一个加权掩码 $M_{2,j}^c(x, y)$ 。

步骤 10 根据多尺度变换系数的零树分布特性,利用粗尺度的加权掩码 $M_{1,j}^c(x, y)$ 校正较细尺度的加权掩码 $M_{2,j}^c(x, y)$ 。

步骤 11 根据式(5),将尺度 1 和尺度 2 下的加权掩码进行线性组合,得到两个尺度下的统一掩码 $M^c(x, y)$ 。

步骤 12 根据式(6),将 S 通道和 V 通道的统一掩码进行线性组合,得到两个颜色通道的公共掩码 $M^{\text{SV}}(x, y)$ 。

步骤 13 利用最小二乘拟合方法计算 S 和 V 两个颜色通道的公共掩码 $M^{\text{SV}}(x, y)$ 的自适应分割阈值。

步骤 14 以投票方式确定运动车辆区域的二值掩码 $M_{\text{binary}}(x, y)$ 。

步骤 15 利用结构元素

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

对 $M_{\text{binary}}(x, y)$ 进行形态学膨胀运算,得到二值掩码 $M_{\text{binary}}(x, y)$ 。

步骤 16 根据式(11),将二值掩码 $M_{\text{binary}}(x, y)$ 与视频帧 F 相乘,提取出运动车辆的候选区域:

$$O(x, y) = M_{\text{binary}}(x, y) \times F(x, y) \quad (11)$$

其中, $O(x, y)$ 表示输出视频帧 O 中在坐标 (x, y) 处的像素值, $F(x, y)$ 表示视频帧 F 中在坐标 (x, y) 处的像素值。

步骤 17 若 V_l 的全部视频帧均已处理完毕,则输出车辆目标在各视频帧中的位置,算法结束;否则,返回步骤 2。

6 实验结果与分析

鉴于目前尚没有一个公认的数据集和客观指标可用于评

价阴影去除算法的性能,我们选取了两种方法来验证本文算法的有效性:1)基于实验结果图的主观对比;2)利用现有的车辆跟踪算法来间接比较阴影去除前、去除后的跟踪效率,用以验证阴影去除的效果。

6.1 基于实验结果图的主观比较结果

本文选取两组视频流进行实验:Highway I 测试集和我们自行搜集的交通视频集。对于这两组视频流而言,阴影去除的难度主要在于:整个视频流较暗且阴影较强,致使阴影与车辆难以区分。

本文算法与目前较为经典的 4 种阴影检测与去除方法对两组测试视频流的处理结果如图 10 和图 11 所示,比较方法包括 DNM1 方法^[4]、零树小波方法^[15]、非全局 MRF 方法^[9]和全局 MRF 方法^[9]。可以看出, DNM1 方法将大量运动阴影误检为车辆区域(见图 10(b)和图 11(d));并且,就检测结果而言, DNM1 与零树小波方法在车辆区域、阴影区域和背景区域均会产生大量噪点,表明两种方法对运动前景和背景的判别精度尚有局限。而全局 MRF 和非全局 MRF 算法并没有有效去除阴影的外轮廓(见图 11(b)和图 11(c))。值得一提的是, DNM1 与零树小波方法均需要提前输入背景帧作为指导,再通过背景差分法获取运动车辆区域,而本文算法则无此特殊要求。

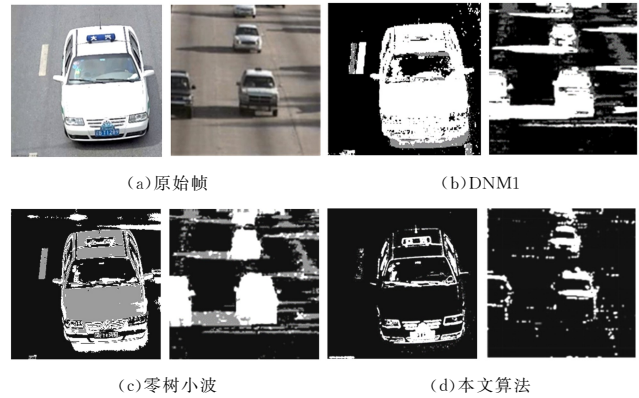


图 10 4 种不同算法的阴影去除结果

Fig. 10 Shadow removal results of four different methods

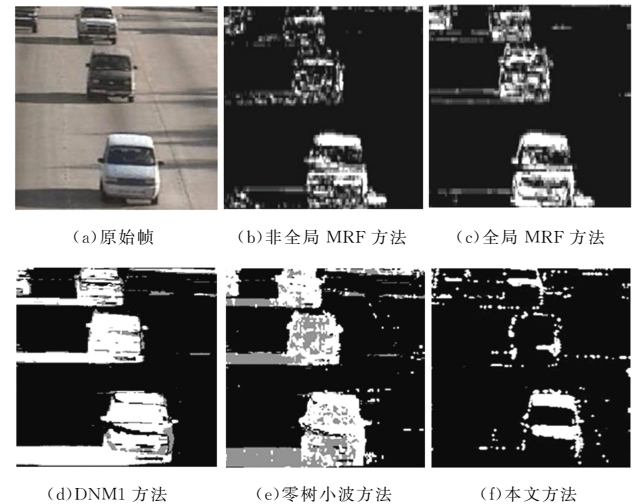


图 11 6 种不同算法的阴影去除结果

Fig. 11 Shadow removal results of six different methods

综合多方面因素,本文算法能够更加有效地去除由运动

车辆产生的运动阴影,且无需背景建模和人工交互。

6.2 基于车辆跟踪算法的间接比较结果

本文选择了两组拍摄手段不同的监控视频进行实验。这两组视频均是在强光照、复杂阴影的环境下拍摄的,同时含有静止阴影和运动阴影。

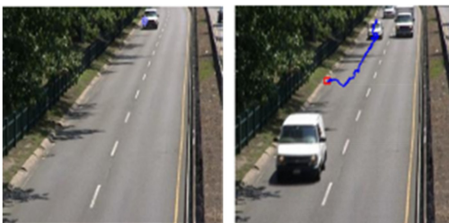
本文将两种经典的车辆跟踪方法作为对比方法,即均值漂移(Meanshift)法和基于轮廓波的改进均值漂移法^[20]。为了使结果更加直观,用蓝色线条表示跟踪得到的车辆运行轨迹。3组实验的跟踪结果如图12—图14所示。

在第一组和第二组实验中,由于路面阴影较重,传统均值漂移方法在跟踪开始后不久便丢失车辆目标;虽然基于轮廓波的均值漂移方法利用多尺度变换能够更准确地匹配车辆目标的特征,但是它在跟踪过程中存在跟踪目标的漂移现象(图12(b)和图13(b)的蓝色轨迹出现了明显拐点),并且在强阴影处也丢失了跟踪目标。第三组测试视频源自自行车记录仪,它与目标车辆之间系同步或近似同步的运动,因此蓝色轨迹偏离红色目标框的位移量越小,就表明跟踪结果越精准。从图14(a)和图14(b)可见,在去除阴影以前,传统Meanshift算法和改进Meanshift算法的输出轨迹均大幅偏离目标。此外,为了度量跟踪所得轨迹的偏离程度,本文统计了两个跟踪算法在每一帧上的输出轨迹与真实轨迹之间的平均欧氏距离,结果如表1所列。由于第三组视频在拍摄过程中存在抖动,我们未对其偏离程度进行分析。

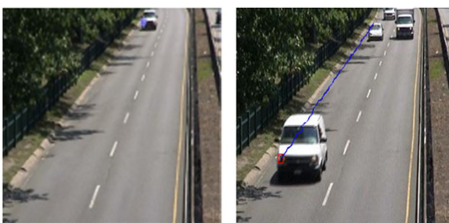
综合图12—图14和表1的结果,交通场景中的静态阴影和运动阴影均会对车辆跟踪算法造成显著影响,其会明显降低跟踪精度,甚至可能导致跟踪失败。



(a) 去除阴影前,Meanshift 算法的初始追踪位置和跟踪结果



(b) 去除阴影前,改进 Meanshift 算法的初始追踪位置和跟踪结果



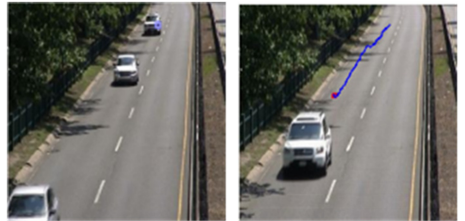
(c) 去除阴影后,Meanshift 算法的初始追踪位置和跟踪结果

图 12 第一组视频在去除阴影前后跟踪结果的比较
(电子版为彩色)

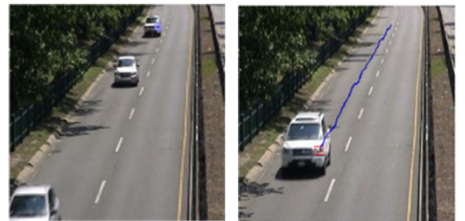
Fig. 12 Comparison between tracking results before and after shadow removal for first video



(a) 去除阴影前,Meanshift 算法的初始追踪位置和跟踪结果



(b) 去除阴影前,改进 Meanshift 算法的初始追踪位置和跟踪结果



(c) 去除阴影后,Meanshift 算法的初始追踪位置和跟踪结果

图 13 第二组视频在去除阴影前后跟踪结果的比较
(电子版为彩色)

Fig. 13 Comparison between tracking results before and after shadow removal for the second video



(a) 去除阴影前,Meanshift 算法的初始追踪位置和跟踪结果



(b) 去除阴影前,改进 Meanshift 算法的初始追踪位置和跟踪结果



(c) 去除阴影后,Meanshift 算法的初始追踪位置和跟踪结果

图 14 第三组视频在去除阴影前后跟踪结果的比较

Fig. 14 Comparison between tracking results before and after shadow removal for the third video

表1 车辆跟踪轨迹与实际运行轨迹的距离统计

Table 1 Distance statistics between tracked trajectory and actual moving trajectory

(单位:像素)

	Meanshift 跟踪 的平均距离	改进 Meanshift 跟踪 的平均距离	去除阴影后 Meanshift 跟踪的平均距离
第一组	89.21	63.75	3.72
第二组	73.34	47.65	4.04

相比之下,当应用本文算法去除了交通场景中的阴影后,传统 Meanshift 算法获得了平直的车辆轨迹和稳定的跟踪结果(见图 12(c)、图 13(c)和图 14(c)),未出现目标丢失的现象。具体地,跟踪轨迹与真实轨迹间的平均欧氏距离缩短了 95%。可见,本文算法可有效去除交通监控视频中的静止阴影和运动阴影,并抑制阴影的干扰,进而增强智能分析算法的鲁棒性。

结束语 本文提出了一种基于空-频域联合投票策略的交通视频阴影去除方法。与现有方法相比,本文方法利用非下采样剪切波变换充分发掘运动车辆区域与阴影区域之间的纹理特性差异,从而更加准确地对两者进行区分;利用变换系数的尺度间相关关系和零树分布特性,构造了目标车辆区域的统一掩码和二值掩码,获得了更有效的阴影区域去除结果。此外,考虑到本文算法需要利用 S 通道和 V 通道进行投票,当交通监控图像出现过曝光或曝光不足的现象时,两个通道的采样质量均会出现明显下降,以致影响本文算法的阴影去除准确率。我们将在今后的工作中进一步深入研究该问题的解决思路。

参 考 文 献

- [1] TOYAMA K, KRUMM J, BRUMITT B, et al. Wallflower: principles and practice of background maintenance[C]//Proceedings of the IEEE International Conference on Computer Vision. New York: IEEE Press, 1999: 255-261.
- [2] BOUWMANS T. Traditional and recent approaches in background modeling for foreground detection: an overview [J]. Computer Science Review, 2014, 11-12(1): 31-66.
- [3] GUO L L. Research on motion target detection technology for complex scenes [D]. Kunming: Yunnan University, 2017.
- [4] CUCCHIARA R, GRANA C, NERI G, et al. The sakbot system for moving object detection and tracking[C]//Proceedings of the European Workshop on Advanced Video-Based Surveillance Systems. Boston: Springer Press, 2001: 145-157.
- [5] DENG Y L, WU L F, LI Y T. An effective shadow removal approach [J]. Signal Processing, 2011, 27(11): 1724-172.
- [6] CHOI J M, YOO Y J, CHOI J Y. Adaptive shadow estimator for removing shadow of moving object [J]. Computer Vision and Image Understanding, 2010, 114(9): 1017-1029.
- [7] XIANG J H, FAN H, LIAO H H, et al. Moving object detection and shadow removing under changing illumination condition [J/OL]. <https://doi.org/10.1155/2014/827461>.
- [8] OUIVIRACH K, DAILEY M N. Extracting the object from the shadows: maximum likelihood object/shadow discrimination

[C]//Proceedings of the IEEE International Conference on Electrical Engineering/ Electronics, Computer, Telecommunications and Information Technology. New York: IEEE Press, 2013: 1-5.

- [9] LIU Z, HUANG K, TAN T. Cast shadow removal in a hierarchical manner using MRF [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(1): 56-66.
- [10] SALVADOR E, CAVALLARO A, EBRAHIMI T. Cast shadow segmentation using invariant color features [J]. Computer Vision and Image Understanding, 2004, 95(2): 238-259.
- [11] CHEN R, LI P, HUANG Y. Motion shadow removal algorithm based on multi-feature fusion [J]. Computer Science, 2012, 45(6): 291-295.
- [12] TIAN J, ZHU L, TANG Y. Outdoor shadow detection by combining tricolor attenuation and intensity [J]. EURASIP Journal on Advances in Signal Processing, 2012, 2012: 116-123.
- [13] GUAN Y. Unsupervised motion detection with background update and shadow suppression [J]. Journal of Control Theory & Applications, 2012, 10(4): 517-522.
- [14] KHARE M, SRIVASTAVA R K, KHARE A. Dual tree complex wavelet transform based shadow detection and removal from moving objects[C]//Proceedings of the SPIE 9029, Visual Information Processing and Communication V. Bellingham: SPIE Press, 2014.
- [15] WANG X H, WANG K, LIU M Y, et al. Zero-tree wavelet based vehicle motion shadow filtering method for traffic video [J]. Pattern Recognition And Artificial Intelligence, 2016, 29(12): 1104-1113.
- [16] XU M L, ZHU J J, LV P, et al. Learning-based shadow recognition and removal from monochromatic natural images [J]. IEEE Transactions on Image Processing, 2017, 26(12): 5811-5824.
- [17] NUR S A, IBRAHIM M M, ALI N M, et al. Vehicle detection based on underneath vehicle shadow using edge features[C]//Proceedings of the IEEE International Conference on Control System, Computing and Engineering (ICCSCE). New York: IEEE Press, 2016: 407-412.
- [18] LV J. Research and implementation of target detection and shadow suppression algorithm in video monitoring system [D]. Shenyang: Northeast University, 2010.
- [19] WANG X H, NI P G, SU X, et al. The nonsubsampling Contourlet HMT model [J]. SCIENCE CHINA Information Sciences, 2013, 43(11): 1431-1444.
- [20] WANG X H, LIU M Y, SU Y H, et al. Traffic video vehicle tracking algorithm based on Contourlet and MeanShift [J]. Journal of Liaoning Normal University (Natural Science Edition), 2017, 40(2): 192-198.



SONG Chuan-ming, born in 1980, associate professor, is a member of CCF. His main research interests include image and video coding, and traffic surveillance video processing.