

基于视觉的交通监控系统固定背景区段 视频传输的研究

姚孝明 钱清泉

(西南交通大学电气学院自动化研究所, 成都 610031)

摘要 多年来, 视频图像传输总是依赖于图像压缩比和网络带宽之间的权衡来实现。根据具有固定背景的基于视觉的交通监控系统的视频传输特点采用单参考图背景差提取对象的方法研究, 提出了一种基于特定对象的图像传输方式: (1) 固定背景分别在采集端与控制中心预存; (2) 采集端应用单参照图背景差方法提取目标对象, 生成最小矩形覆盖网格区域及其掩码矩阵; (3) 网络通道可以根据应用情况灵活选择; (4) 控制中心则根据目标对象编码或其小波变换结果以及网格区域和掩码在背景图上重构出指定背景中目标运动视频。其优势在于: (1) 大量节省网络带宽; (2) 提供系统高级处理的有效输入如空间位置关系图以及可以校正的目标对象速度数据; (3) 便于操作人员注意主要预警事件; (4) 为系统扩能提供空间等。实验结果表明了这种方法的可行性。

关键词 交通监控系统 背景差方法 图像传输 小波变换 图像融合

中图分类号: TP391 **文献标识码**: A **文章编号**: 1006-8961(2005)08-1042-05

Study on Section Video Transmission with Preset Background of Vision-based Traffic Surveillance Systems

YAO Xiao-ming, QIAN Qing-quan

(Institute of Automation, College of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031)

Abstract In this paper, an object-of-interest based video transmission approach is presented for vision-based traffic surveillance systems using the background subtraction with one reference image. This background is pre-stored at both the image-capturing side and the monitoring center; at the acquisition side, using the background subtraction with only one reference image the target objects are extracted, deriving a minimum rectangular grid-based region and a mask matrix for each of the extracted object. The monitoring center, after receiving the necessary codes, can reconstruct the video frames by fusing the objects or their wavelet-transformed versions into the preset background. The advantages are: (1) the networking resources are greatly saved; (2) some high-level processing functionality like collision avoidance, fault alarming can be improved; (3) operators' attention can focus on important events; (4) new services can be added or developed. Experimental results have proved its effectiveness.

Keywords traffic surveillance systems, background subtraction, video transmission, wavelet transform, image fusion

1 引言

基于视觉的交通监控系统是智能交通领域的重要研究课题。通过对路边高架安装的摄像机所拍的视频图像进行分析, 可以完成诸如目标提取、目标跟

踪以及目标活动模式分析等工作, 为交通流量控制、违规车辆的截获等实际应用开辟了新的发展前景。

多年来, 图像传输总是依赖于图像压缩比和网络带宽之间的权衡来实现^[1,2]。目前市场上所见的大多是通过选择合适的信道来实现一定传输率的图像传输。以基于对象/内容为出发点的 MPEG2、

基金项目: 国家自然科学基金项目(50407009); 四川省应用基础研究基金项目(02GY029-039)

收稿日期: 2004-10-09; **改回日期**: 2004-12-27

第一作者简介: 姚孝明(1964 ~), 男, 博士研究生。1985年于西南交通大学电气工程系获学士学位, 2003年于南昌大学获计算机软件与理论专业硕士学位。主要从事数据安全、视频监控、图像处理、计算机视觉及其在应用、网格计算等方面的研究。发表论文多篇, 有国家专利1项。E-mail: xiaomingyao@163.com

MPEG4 以及 H. 26L 等标准虽然考虑了对各帧对象变化的预测补偿,但其对象平面并非完全独立,基本算法仍然是基于宏块的。为此提出面向特定对象即车辆对象的处理方法。

一般地,由于车辆运行的背景变化不大,同时系统更为关心车道上的主体——车辆,因此,通过与背景求差可得到车辆图像及其运动数据,此即有名的“背景差方法”^[3,4]。背景差方法面临了巨大的挑战:噪声,光影以及背景部分对象的运动如树枝的摇摆、路面的局部损坏等,也因此获得巨大的进展和成功,许多改进方法被提出,如动态背景维护的各种统计模型;依据光照恒常理论或像素线性相关的原理设计出了消除光影的统计检验的方法;根据色彩、边缘梯度以及纹理特征的提取方法等等^[5-12]。其局限主要是:(1)需要多幅参照图;(2)需要实时维护背景图;(3)计算量大;(4)实时性不强;(5)误检率较高。

近年来,Ohta 提出了用单参照背景图消除光影统计检验的方法,计算简单,无需背景维护^[13]。由于单参照背景的实用前景,我们通过对光照的整体分布特征和局部分布特征的分析提出了统计均匀算子,得到了在不同光照条件下用同一背景提取目标的迭代算法,实验结果较为理想。

利用图像格论将图像像点视为某一 Voronoi cell 的中心点。这样,一幅图像的任意区域可以用包罗若干个格的矩形区域来表示^[11]。因此,可以预先设定图像平面上车辆运动范围,如果摄像机的安装位

置设计恰当,某些背景运动可以得到有效忽略。在这样的框架下,利用单背景参照进行目标提取和运动图像重构的技术思路是有价值的。也就是说,在单背景参照进行目标提取后,能够将目标对象按照运动测量结果在同一背景上重构生成新的运动图像,使得后者能够准确反映目标与背景的图像位置关系并为人类观察者所接受。其优势是多方面的:(1)可以大量节约带宽:例如,固定背景可以预先存储,无需传输,可复用目标提取后也不必传输,只需传输掩码和位置码;(2)提供基本高级分析数据如目标空间位置关系图输入数据,目标运动参数(格点速度);(3)便于操作人员注意主要预警事件,不为无细节分心;(4)为进一步智能化提供前期基础;(5)为开展其他业务提供带宽资源等。

本文研究了实现这一思路的困难与方法,讨论了“网格定位,小波变换,掩码嵌入”的初步方案,实验结果表明了方案的可行性。

2 网格定位编码方法及目标形体位置相关的多分辨率特征

2.1 网格定位编码方法与目标掩码矩阵游程编码

图像格论给出了图像平面内任意区域的基于 Voronoi cell 的同一描述,为将一个像点视为格点提供了理论依据^[11]。由此,可以建立基于格点的整数坐标系。一般地,为了方便应用,以图像平面左上角像点为原点(0,0),如图 1 所示。

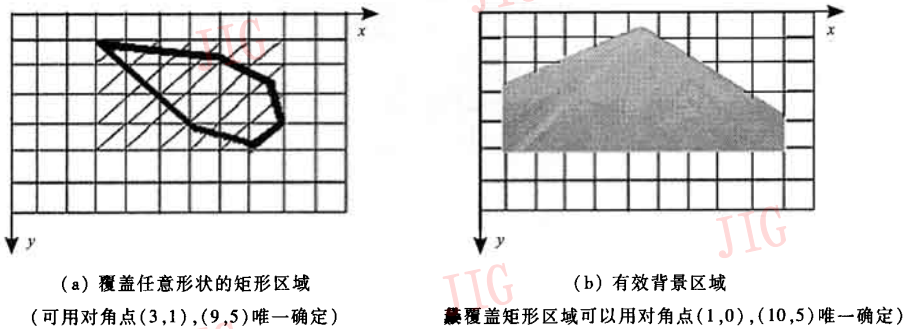


图 1 网格坐标系

Fig. 1 Grid Expression

可看出,(1)无关背景区域可以有效界定、忽略;(2)可以用像点网格确定最小覆盖目标的矩形区域,由此也就确定了目标的图像位置;(3)最小覆盖矩形区域可以包含非目标区域;(4)不同位置,目

标形体有所不同。

应该指出,由图像平面提取出来的运动参数显然并非目标的实际运动参数,需要通过透视变换的比例关系进行转换校正,这有赖于摄像机的安装及

其参数设置。

已有许多采用背景差法提取目标的方法。对于光影、背景局部变化、由于地面震动产生的镜头抖动以及目标之间的部分遮挡等问题已有相应的处理方法。其中,通过多视图的比照,一些部分遮挡的多目标可以有效分离^[10];采用 Kalman 滤波可以抑制抖动影响,对象匹配与分类也能有效完成^[14]。当目标距离镜头超过一定值时,分辨率可能过低导致提取过程失效。这时,恰当根据局部分辨率的分布情况来限定有效可视范围,可以力保目标的有效提取。可见,在本文,假定目标可以提取、分类,遮挡多目标可以有效分离是合理的。

一个目标可以由某个矩形网格区域完全覆盖。一个矩形区域可以由其对角坐标精确决定。一个简单的定位编码方法是对这两个点坐标进行编码。方式有两种:绝对坐标方式与相对坐标方式。所谓绝对坐标方式就是两个点的坐标均用绝对值,即

$$R = (x_1, y_1, x_2, y_2) \tag{1}$$

而相对坐标方式则是前一点的坐标用绝对值,后一点用相对值,即

$$\begin{cases} R = (x_1, y_1, \Delta x, \Delta y) \\ \Delta x = x_2 - x_1, \Delta y = y_2 - y_1 \end{cases} \tag{2}$$

一般地,相对方式按照编码法则由于数值较小,效率较高。特别地,若编码区域为正方形,则上述 4 元组的编码即转化为 3 元组的编码,效率更高。

网格定位码集 $\{R\}$ 即可完全确定区段内所有目标的当前位置。通过透视变换可以得到其相对空间位置关系图。这种关系图对分析目标的活动模式有着重要的应用价值。

很明显,如果只传输提取的目标区域,依照网格定位码集 $\{R\}$,在控制中心的固定背景上依次进行像素覆盖就可以完成图像的重构。如果不考虑传输误差,那么拍摄图像与重构图像两者是相同的。有时,技术上要求掩去矩形区域内的非目标部分。为此,采用表达目标轮廓形状的掩码矩阵 M ,即一个大小与网格矩形区域相同的逻辑矩阵,其元素在目标区域内取值为 1,否则为 0。这类矩阵可以转换为多个 0 与多个 1 交替的 1 维数组,最适合采用游程编码方法。设首元为 a ,其逆为 \bar{a} ,则编码形式为

$$an_1\bar{a}n_2an_3\cdots\bar{a}n_i an_{i+1}\cdots$$

其中, $\sum n_i = p \times q$ 为掩码矩阵 M 大小。

2.2 目标形体位置相关的多分辨率特征

观察分析图 1,可以看到:(1)道路形体表现为

一定的视图方向,不同的方向展示道路不同侧面,侧面差异与距离镜头远近有关;(2)差异不大的侧面及其分辨率大小取决于与镜头距离,远则小,近则大。这是一种位置相关的多分辨率特征。可以推知,道路上车辆具有同样的视图特征。

关于多分辨率图像的小波处理已有大量文献论述。任何一幅高分辨率图像可以分解为另一幅较低分辨率图像及其高频细节;反之,一幅较低分辨率图像配以适当高频细节也可重构满意的高分辨率图像。选择适当的小波函数可以进行连续缩放^[15]。对目标中心距离很小的运动变化,可以忽略其形体差异,对距离较大者进行小波变换产生其压缩/放大版本。因此,能够有效复用已经传输的目标数据。考虑到车辆运动总是不规则的,一旦改变到某个新位置使得车辆视图产生明显改变时,小波处理结果会产生明显的“假”像:在某位置摆放了形体与运动方向不合的车辆!这时,就要求传送完整的目标区域编码。而其最佳编码还有待进一步研究。这样,传输事件被分为 3 类:(1)无补偿法:只传送定位码与掩码;(2)有补偿法:另加高频系数码(放大)或压缩标志码;3)区域图像压缩编码。由于同一车辆运动范围以及其形体视图变化总是有限的,如果对车辆形体对象采用类似文字的 LZW 编码的标记方法^[2],通过建立一个动态目标数据库,可以获得更多的计算与传输收益。例如,对于系统内的中间区段,进入的车辆全部来自上一区段,因此,如果区段设置相似,则目标形体数据可以共享。

控制中心根据以上各类编码数据进行图像重构。为了消除网格区域中其他成分,引入掩码矩阵 M 的补阵 \bar{M} ,满足:

$$M + \bar{M} = U \tag{3}$$

其中, U 为大小与 M 相同,全部元素取值为 1 的全 1 阵。若矩阵的点乘运算用“ $\cdot *$ ”表示,目标在新区域 R' 的嵌入过程可以表示为

$$R_{new} = M \cdot * R + \bar{M} \cdot * R' \tag{4}$$

其中,“ $\cdot *$ ”为矩阵的点乘运算。

式(4)称之为“掩码嵌入”。其视觉效果一般采用主观评定(5 级)与客观评定(PSNR)相结合的方法^[1,2],通过在调试阶段得到的结果,决定传输数据类别细节。另外,由于光影的去除,重构图像对敏感观察者而言,也会得到“假”的人工印象,但这无碍于对交通状况的判断和分析。因此,本文方法更适合监控系统对控制目标的高级理解的应用而不是一般性的观赏。

3 区段交通图像的新传输方式及其实验结果

基于 MPEG4 等标准的视频压缩考虑了各帧空-时信息冗余而获得较高压缩比。其图像采集端功能较简单,即图像采集、压缩以及通信功能。由于信道传输中存在误码的可能,因此,一般还需考虑信道编码(即差错控制编码)。本文所讨论的新方式则有如下几点变化:

(1) 图像采集端功能扩展:显然,有背景图存储单元,图像分析单元(包括目标提取、分类与标识等)以及相应网格定位编码和掩码游程码等关键模块;

(2) 传输数据的变化:动态目标数据,根据目标检出的情况确定,亦即无目标检出时不传输,有目标时按目标空间位置关系依次传输目标区域及其定位码和掩码;

(3) 控制中心的变化:相应于采集端,需要有解码设备,并按式(4)进行图像融合,产生新图像帧。

过程差异大致可用图 2 来说明。

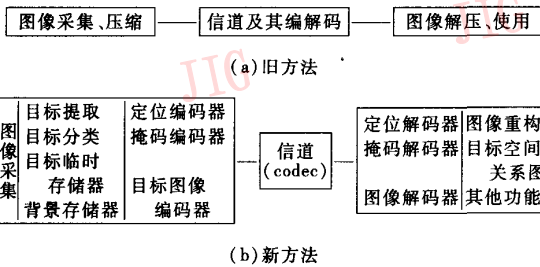


图 2 新旧传输方法的比较

Fig. 2 Comparison between the proposed approach and the conventional ones

考虑到交通实时信源以及实际信道的复杂性,表 1 给出了极端情况下(对象无复用、无专用压缩法)新方法压缩比、传输时间以及占用带宽比的估计数据。

表 1 极端情况下新方法相对旧方法的压缩比、传输时间以及信道增益的估计数据

Tab.1 Estimated compression ratio, transmission time and bandwidth gains of the proposed approach relative to old methods in worst case

图像帧格式 (CIF)	车辆对象大小	相对压缩比	传输时间	带宽比 (W_n/W_o)
288 × 352	63 × 80	17.7:1	$\leq 0.5t_o$	1:16

其中, t_o 为传统方法所需传输时间, W_n, W_o 分别为新、旧占用带宽。

将获得的基本背景图与提取的车辆矩形区域及其用小波压缩版本进行融合,结果如图 3、图 4 所示。定位码:前车 (154, 80, 217, 160), 后车 (120, 150, 152, 191), 由于未采用掩码,边缘痕迹较为明显。



图 3 提取的车辆对象矩形区域及其压缩版本
Fig.3 Rectangular region of an extracted vehicular object and its compression version with sym2 wavelet

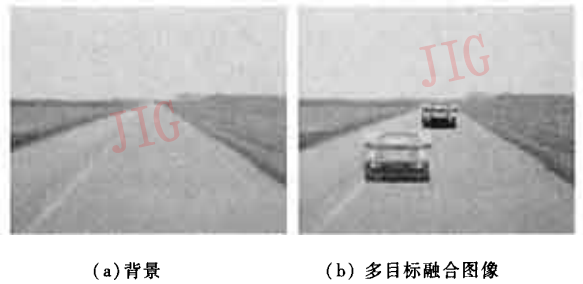


图 4 背景与融合图像

Fig. 4 Background and the fused frame; location code

由图可见,车辆形体位置相关的多分辨率特征十分明显,车辆及其压缩版本在定位编码的确定下,图像主观评价可以达到 2 级(4 分)甚至 1 级(5 分)。此外,掩码的引入可以产生好的边缘过渡,但是如前所述,光影特征将被消除。

因此固定背景与提取车辆及其压缩版本的融合所构造的新图像可以表达实际图像的主要信息。此外,在一定分辨率范围内,中心端通过对定位码的透视校正处理可以直接得到目标的空间关系图,为系统进行高级活动分析以及增加新的业务功能奠定技术基础。例如,通过图像平面与实际路面的透视校正,车辆中心格点距离关系可以转换为较为准确的实际距离关系,例如上图前后车辆中心格点距离为 70.714,而按模型地面中心计算约 71.823,两者接近。因此,可以用来速度估计,预测新的目标位置关系并由此给出潜在事故预警等等。可以肯定,作为基于视觉的系统,事先确定摄像机的安装高度与视

域范围可提高校正精度并得到较好的融合图像。另一方面,预存背景使得通道传输的内容为车辆及其实时掩码,从沿用传统方法以及对对象复用不可得的最坏情况估计,表 1 表明,对 CIF 格式图像帧,车辆平均大小不大于 63×80 ,掩码约每车辆 12 比特左右,可见,100 辆车辆的掩码也仅需 150 字节左右,因而可以不计,这时压缩比可提高约为 17.7:1,传输时间约为原来的一半,新旧占用带宽比为 1:16。由此可见,方法的有效性是明显的。

4 结 论

综上所述,通过在采集端增加目标对象提取、分类等智能处理模块和必要的存储单元与编码单元可以产生包括定位码、掩码以及对对象区域压缩码的合成数据流。控制中心可据此重构包含主要信息的实际视频。数据流大小取决于对象检出的情况:无检出时不发送,有检出时为覆盖检出对象的区域图像码与定位码以及掩码的总和。实验结果表明,通过对象的变换复用该数据流大小比 MPEG4 的标准结果还要小数十倍,从而节省大量带宽。此外,额外的计算与设备开销还带来了其他的好处:定位码与掩码提供位置信息,通过模型与透视校正可以得到目标的速度估计值,为系统高级分析如事故预测等提供了较好数据源。

人员监控有着许多不足,图像不必处处完美。可以料想,随着目标对象提取技术的进一步发展,图像边缘的界定更符合人类视觉特点,那么,融合的图片也会更加趋于完美。

致 谢 对西南交通大学博士生创新基金的资助表示感谢。

参考文献 (References)

- 1 Shen Lan-sun, Zhou hi, Tian Dong, *et al.* Video Coding and Low-bit-rate Transmission [M]. Beijing: Electronic Industry Publishing House, 2001. [沈兰荪, 卓力, 田栋等. 视频编码与低速率传输 [M]. 北京: 电子工业出版社, 2001.]
- 2 Wu Le-nan. Data Compression: Principles and Applications [M]. Beijing: Electronic Industry Publishing House, 1995. [吴乐南. 数据
- 3 Wren C, Azarbayejani A, Dorrell T, *et al.* Pfindex: Real-time tracking of the human body [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(7): 780 ~ 785.
- 4 Stauffer C, Grimson W E L. Adaptive background mixture models for real time tracking [A]. In: Proceeding of the Computer Vision and Pattern Recognition Conference [C], Fort Collins, Colorado, USA, 1999: 246 ~ 252.
- 5 Elgammal Ahmed, Harwood David, Davis Larry. Nonparametric background model for background subtraction [A]. In: Proceedings of the 6th European Conference on Computer Vision [C], Trinity College Dublin, Ireland, 2000, 2: 751 ~ 767.
- 6 Takashi Matsuyama, Takashi Ohya, Hitoshi Habe. Background subtraction for non-stationary scenes [A]. In: Proceeding of the 4th Asian Conference on Computer Vision [C], Taipei, Taiwan, 2000: 662 ~ 667.
- 7 Seki M, Fujiwara H, Sumi K. A robust background subtraction method for changing background [A]. In: Proceeding of the Fifth IEEE Workshop on Applications of Computer Vision [C], Palm Springs, California, 2000: 207 ~ 213.
- 8 He Yinghua, Wang Hong, Zhang Bo. Background updating in illumination-variant scenes [A]. In: Proceeding of the International Conference on Intelligent Transportation Systems [C], Shanghai, China, 2003, 1: 515 ~ 519.
- 9 Collins R T, Kanade T. Multi-camera tracking and visualization for surveillance and sports [A]. In: Proceedings of the Fourth International Workshop on Cooperative Distributed Vision [C], Kyoto Japan, 2001: 27 ~ 55.
- 10 Wang Yao, Ostermann Jorn, Zhang Ya-Qin. Video Processing and Communications [M]. Beijing: Tsinghua University Press, 2003.
- 11 Jiang H, Drew M S. A predictive contour inertia snake model for general video tracking [A]. In: Proceeding of the 2002 International Conference on Image Processing [C], Rochester, NY, USA, 2002, 3: 413 ~ 416.
- 12 Finlayson G D, Hordley S D, Drew M S. Removing shadows from images [A]. In: Proceeding of the 7th European Conference on Computer Vision [C], Copenhagen, Denmark, 2002, 4: 823 ~ 836.
- 13 Naoya Ohta. A statistical approach to background subtraction for surveillance systems [A]. In: Proceeding of the International Conference on Computer Vision [C], Vancouver, Canada, 2001, 2: 481 ~ 486.
- 14 Prabhakar B, Kadabar Damodar V. Automatic detection and matching of moving objects [J]. Central Research Laboratory Technical Journal, 2001, 3(3): 32 ~ 37.
- 15 Wu Fan. Scale-dependent representations of relief based on wavelet analysis [J]. Geo-spatial Information Science, 2003, 5(1): 66 ~ 69.