

# 视频字幕去除系统的设计及实现

宋 璧 刘政凯 邵肖伟

(中国科技大学信息处理中心,合肥 230027)

E-mail: gozkui@sohu.com

**摘 要** 老式视频中固化的字幕影响了不同语种间视频的交流和处理。论文基于视频字幕的特点,提出了一种视频字幕的定位及提取方法,并结合图像修复技术,实现了一个自动视频字幕去除系统。实验结果表明,该系统能够较好地去除字幕。

**关键词** 文字提取 图像修复 视频处理

文章编号 1002-8331-(2004)27-0107-02 文献标识码 A 中图分类号 TP391

## Design and Implementation of a Video Title Removal System

Song Bi Liu Zhengkai Shao Xiaowei

(Information Processing Center, University of Science and Technology of China, Hefei 230027)

**Abstract:** The titles embedded in old video have obstructed video intercommunion and processing among different languages. This paper proposes a title location and extraction approach based on the characteristics of the titles in video and removes them by image inpainting technology. It presents an automated video title removal system and experiment results show that this system performs fairly well.

**Keywords:** text extraction, image inpainting, video processing

### 1 引言

随着科学技术的飞速发展,视频已经成为人们日常生活中不可缺少的娱乐活动之一。老式的视频由于技术限制,字幕被固化于视频中,成为视频的一部分。但在实际应用中人们发现,固化的字幕严重阻碍了不同语种间的视频交流。因此,现有的视频技术都采用视频与字幕相分离的方式,使用软件来动态显示字幕。对于原有的老式视频,如果将字幕从视频中去除并修复,然后添加上多语种字幕文件,则也可以实现软件显示字幕。基于上述需求,设计并实现了一个视频字幕去除系统。

视频字幕去除系统中的关键技术是字幕提取和图像修复。现有的图像文字定位方法主要是利用文字区域所具有的纹理特征来处理的,如 Wu 等人使用多尺度高斯滤波计算水平和垂直纹理特征并聚类来定位文字<sup>[1]</sup>, Li 等人通过小波变换提取特征并采用神经网络分类器来检测文字<sup>[2]</sup>, Cai 等人利用文字边缘的空间分布特点作为文字检测的依据<sup>[3]</sup>等。但现有的文字定位算法的主要目的是文字识别,因此精度较低,只要定位至大致的矩形区域就可以满足要求。这种精度不能满足图像修复的要求,因此还需要进行文字提取,获得较精确的文字模板,以进行图像修复。

在图像修复领域,主要有基于 PDE 的修复方法和基于纹理的修复方法两大类。经典的基于 PDE 的修复方法有 Bertalmio 等人在 2000 年提出的基于投影迭代传播的 BSCB 算法<sup>[4]</sup>,以及 Chan 等人在 2001 年提出的基于 Total Variation 模

型的修复算法<sup>[5]</sup>。这些算法主要是利用边界信息向待修复区域内进行各向异性的迭代扩散,计算量大,在修复区域较小时有较好的效果。基于纹理的修复方法由纹理生成技术演变而来,如 Criminisi 等人提出的基于纹理生成的修复方法<sup>[6]</sup>,在待修复区域的边界通过块匹配的方式选择合适的纹理填充。这种方法计算量小,对大区域也有较好的修复效果。考虑到视频字幕区域较大以及运算量的因素,文中采用纹理修复算法来进行图像修复。

基于上述想法,论文实现了一个视频字幕去除系统,利用视频中字幕位置的帧间相关性和字幕的边缘信息进行精确定位,然后使用彩色量化分割的方法提取文字模板,最后用纹理修复算法填补原有字幕区域,使之达到或接近未加字幕前的视觉效果。

### 2 视频字幕模板提取

与经典的从扫描图象中提取文字相比,在数字视频中提取字幕面临新的挑战。首先,为了减轻低分辨率带来的块效应,视频帧中的字幕采用了抗锯齿插值算法,从而造成了文字边缘的模糊;其次,为了防止字幕被同色背景所掩盖,大部分字幕都存在反色的描边,如白字黑边,这导致了在抗锯齿插值时字幕像素的灰度变化很大,不易提取;最后,图像修复对字幕提取的精度要求较高,如果没有提取出完整的字幕区域则可能造成错误的修复结果,这进一步增加了字幕提取的难度。论文提出一种

字幕提取算法,通过对字幕区域彩色量化提取字幕颜色信息,然后利用文字的连通性采用种子点算法进行分割,最后使用形态学运算来修正文字区域,取得了较好的结果。

2.1 视频字幕定位

观察可知,视频中的字幕一般位于视频下部,呈水平排列,相同字幕的帧间位置基本不变,且整个视频中字幕的颜色是不变的。基于以上特点,针对视频下部 1/3 的区域利用帧间相关性来进行字幕定位,处理过程如图 1 所示。

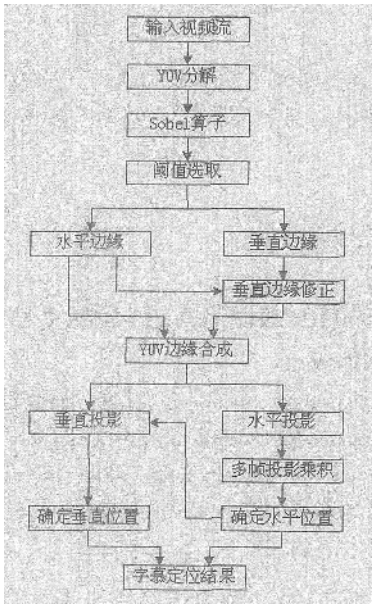


图 1 视频字幕定位流程图

这里采用 Fan 等人所提出的彩色边缘检测方法<sup>[7]</sup>来获取边缘信息。首先将原图进行 YUV 分解,分别用 Sobel 算子获得垂直边缘和水平边缘,然后采用基于判别式准则的直方图阈值选取算法<sup>[8]</sup>选取全局阈值,再将图像分解成子块,选取局部阈值进行二值化,在获得初步的水平边缘和垂直边缘后,考虑到文字部分的水平边缘和垂直边缘有一定的相关性,对水平边缘进行垂直膨胀,然后与原垂直边缘进行相与,得到修正后的垂直边缘,这样可以有效地去除由非文字产生的垂直边缘;最后将 YUV 分量合成得到垂直边缘和水平边缘。

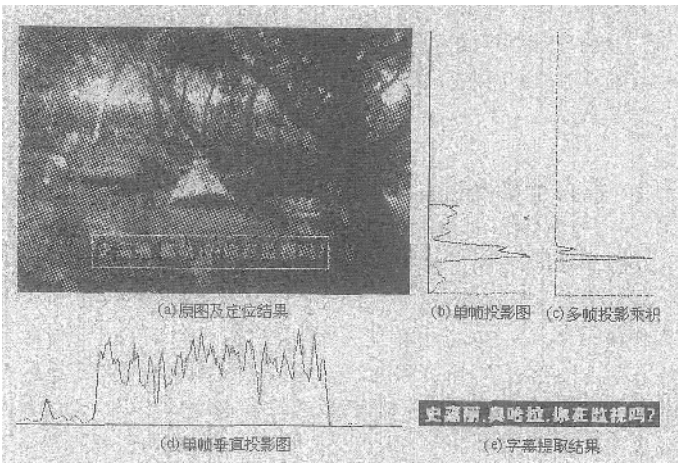


图 2 视频字幕定位及提取结果

幕的边缘丰富,水平排列的字幕在垂直边缘的水平投影上一般都形成一个峰值,如图 2(b)所示。但对于背景比较复杂的情况,字幕在水平投影图中形成的峰可能会不突出。基于相同字幕的帧间位置基本不变这个特点,采用了多帧水平投影图相乘的方法,有效地抑制了背景影响,突出了字幕形成的峰值,如图 2(c)所示。然后根据水平投影采用自动阈值法确定字幕的水平位置,再将该水平区域中的边缘图像进行垂直方向投影,如图 2(d)所示,以确定字幕框的左右边界,从而获得完整的字幕框,完成字幕定位。出于字幕提取的需要,字幕的上下边界向外扩展了 25%。最终的字幕定位结果如图 2(a)所示。

2.2 视频字幕提取

如前所述,视频中字幕的颜色变化较大,因此可以通过彩色量化的方法来降低颜色分布的复杂度,以便于后续处理。论文采用 Orchard 等人提出的等级树结构彩色量化方法<sup>[9]</sup>。首先确定彩色量化级数,文中取为 8,并把所有像素都归为一类;然后对所有的类分别进行 KL 变换,把特征值最大的类分裂成两类,直到分解至指定的类数为止;每类像素的颜色用该类的均值表示。

彩色量化后先对字幕的中间 4 行进行直方图分析,得到直方图 A;然后对字幕区域的上下边缘 4 行进行直方图分析,得到直方图 B。因为 A 中字幕出现的概率较高,B 中字幕出现的概率很小,所以取直方图 A-B 中最大的类作为字幕的种子点,进行区域生长,这样可以较好地提取出因抗锯齿插值而颜色有一定变化的点。考虑到描边的存在,因此对种子点生长结果进行形态学膨胀,以确保完整字幕的提取。最终得到一个二值化的字幕模板,也就是待修复区域模板,如图 2(e)所示。

3 视频修复

现有的视频修补算法主要利用相邻帧的信息来填充待修复帧,这对于持续多帧不变的字幕显然不合适。因此使用了图像修补算法逐帧进行修复。基于前述讨论,使用 Criminisi 等人提出的纹理修复算法<sup>[6]</sup>进行图像修复。

该算法先在待修复区域边界根据边缘方向和导数方向确定待修复块的优先级,此时待修复块中部分信息已知,只有部分区域需要修复;然后统计待修复块中不需要修复的像素个数来计算置信度,最后根据待修复块中的已知信息,在全图中搜索最相似的块并填充待修复部分;反复进行基于块匹配的修复操作直至全部被修复完为止,这样就得到了最终的修复结果。

4 实验结果

基于上述技术,在 Windows2000 环境下使用 Visual C++ 6.0 实现了一个视频字幕去除系统,能对输入的视频序列自动提取字幕并修复,输出修复后的视频序列。

以电影《飘》中的一帧为例,给出文字定位、字幕提取及修复结果。图 2(a)为标注了定位结果的原图,其中白框表示定位结果,可见字幕区域定位结果比较精确;图 2(b)为单帧的水平投影图,图 2(c)为多帧乘积的水平投影图,可见水平位置定位较准确,多帧投影图相乘的方法可以有效地抑制背景影响;图 2(d)为单帧的垂直投影图,与实际位置相当匹配;图 2(e)为视频字幕提取结果,可以满足图像修复的要求。图 3 为采用纹理修复算法得到的修复结果,可见与原图的一致性相当好,目视基本观察不出修复的痕迹,较好地完成了视频修复的目标。

预警预测输入数据进行预警预测。

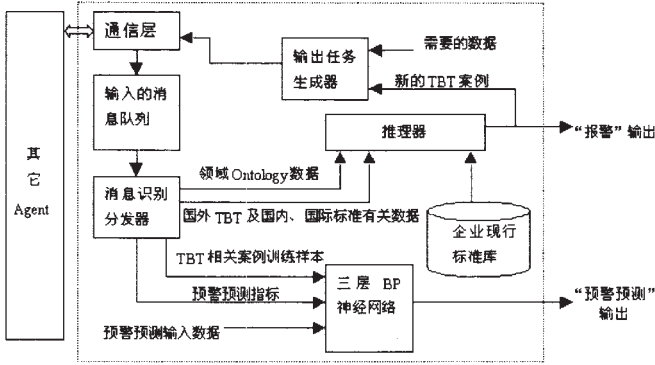


图 4 企业 TBT 预警预测 Agent 结构

### 5.3 国外 TBT 相关文档库 Agent 结构

国外 TBT 相关文档库 Agent 结构如图 5 所示。

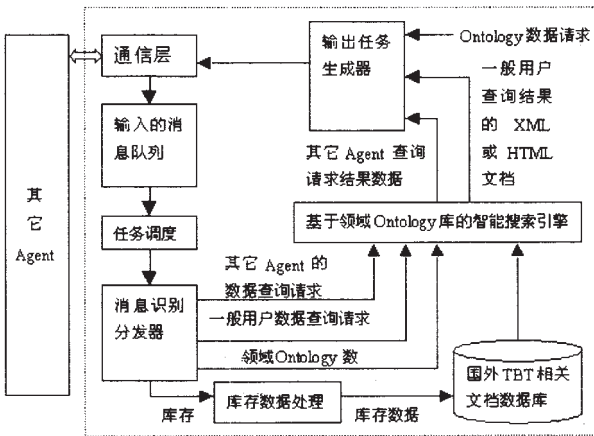


图 5 国外 TBT 相关文档库 Agent 结构

运行时,该 Agent 可能要处理大量的查询请求。其它 Agent 发来的数据,由通信层进入消息队列。由任务调度选出合适的任务,经消息识别分发器还原数据,供库存处理和搜索引擎使用。若是库存数据,经过库存数据处理,对存入的文档进行提取

关键词处理,然后将它们写入国外 TBT 相关文档数据库。

对于查询请求,假如查询引擎需要相应的领域 Ontology 数据,该 Agent 立即构造相应请求发给领域 Ontology 库 Agent;当该 Agent 需要的领域 Ontology 数据到达时,由基于领域 Ontology 库的智能搜索引擎处理查询,并产生查询结果;一般用户的查询结果为 Internet 浏览网页文档,其它 Agent 的查询请求结果为适于相应 Agent 处理的数据。所有这些输出经输出任务生成器打包由通信层发送出去。

### 6 结论

近年来,我国贸易出口受到 TBT 的影响呈现越来越严重趋势,形势十分严峻。结合信息技术、自然科学的方法,尽快建立我国实用、高效、及时、准确的出口产品 TBT 预警预测系统是我国亟待解决的重大问题。论文提出了使用 Multi-Agent 和 Ontology 技术的 TBT 预警预测系统解决方案,给出了出口产品 TBT 预警预测系统框架结构,为彻底解决这个问题提供了技术上的支持,为实现相关文档的自动收集、分发,数据的智能检索,以及 TBT 的智能预警预测,搭好了平台。同时群体决策支持系统的有机融入,为企业提供了一个规避、突破、跨越或应对 TBT 最优化的决策提供了信息化环境。该系统具有良好的可扩充性和自适应性。由于完善的 TBT 预警预测系统是一个庞大的系统工程,在当前系统平台的基础上,笔者正在紧张地进行着系统的进一步完善的大量研究工作。(收稿日期 2004 年 2 月)

### 参考文献

- 1.夏友富,俞雄飞,李丽.TBT 屏障——技术性贸易壁垒发展趋势及其对中国出口贸易的影响[J].国际贸易,2002:10
- 2.王晖.认清形势、采取措施、应对国外技术壁垒[J].技术性贸易措施,2003:1~2
- 3.胡舜耕等.面向自动文摘的多 Agent 系统中的协调算法研究[J].计算机研究与发展,2001;11:1302~1309
- 4.Sycara K P.Multiagent systems[J].AI magazine,1998:79~92
- 5.Jan Uhir.Creation of architectural ontology—user's experience[C].In: Proceedings of the 14th International Workshop on Database and Expert Systems Applications(DEXA 03)2003 IEEE 2003

(上接 108 页)



图 3 去除字幕后的修复结果

### 5 结论

论文针对视频字幕定位、提取以及图像修复等问题,设计并实现了一个视频字幕去除系统,实验结果表明该系统能较好地完成这一任务。(收稿日期 2004 年 5 月)

### 参考文献

- 1.Wu Y,Manmatha R,Riseman E.TextFinder: an automatic system to

- detect and recognize text in images[J].IEEE Trans Pattern Anal Machine Intell,1999,21(11):1224~1229
- 2.Li H,Doermann D,Kia OAutomatic Text Detection and Tracking in Digital Video[J].IEEE Trans Image Processing,2000,9(1):147~156
- 3.Cai M,Song J,Lyu M.A New Approach For Video Text Detection[C].In IEEE Proc ICIP 2002,1:117~120
- 4.Bertalmio M et al.Image Inpainting[C].In SIGGRAPH 2000,1:417~424
- 5.Chan T,Shen F.Mathematical Models for Local Nontexture Inpaintings[J].SIAM Journal on Applied Mathematics,2002,62:1019~1043
- 6.Criminisi A,Perez P,Toyama K.Object Removal by Exemplar-Based Inpainting[C].In IEEE Proc CVPR 2003,2:721~728
- 7.Fan J et alAutomatic Image Segmentation by Integrating Color-Edge Extraction and Seeded Region Growing[J].IEEE Trans Image Processing,2001,10(10):1454~1466
- 8.Otsu N.A Threshold Selection Method from Gray-level histograms[J].IEEE Trans System Man and Cybernetics,1979,9(1):62~66
- 9.Orchard M,Bouman C.Color Quantization of Images[J].IEEE Trans Signal Processing,1991,39(12):2677~2690