

**本科毕业设计（论文）说明书**

**题目： 智能视频摘要软件**

**学 院 软件学院**

**专 业 软件工程**

**学生姓名 刘 远 一**

**学生学号 200730553455**

**指导教师 黄 翰**

**提交日期 2011 年 6月1 日**

**摘 要**

近年来，随着社会的迅速进步，银行、电力、交通、安检以及军事设施等领域对安全监控系统的需求与日俱增。在现在的视频监控领域中，一般来说很难有这么多人力资源来监控或者查看这么庞大的视频资源。

本文针对上述视频数据量大的问题，开发出智能视频摘要软件，将一个监控视频中的所有运动物体的运动提取出来，放到同一个视频中去播放，这样就大大降低了监控视频所需要的人力资源消耗。这对提高监控效率以及安全防护等起着很大的作用。

本文先对视频摘要软件的主要功能进行简单的介绍，对适用范围以及存在问题进行说明；而后对视频摘要软件的处理流程进行介绍；然后对整个视频摘要软件的模块构成和算法原理进行详细分析。

**关键词**：视频摘要；运动物体检测；智能监控

**Abstract**

In recent years, with the rapid social progress, the requirements of Safety Monitoring System Requirements in banking, electricity, transportation, security and military facilities. In the present field of video surveillance, it is difficult to have so many human resources to monitor such a large video.

In this article, to solve the problem of large amount of video data, I developed a software named Summary of Intelligent Video Software. It can extract all of the movement events in a surveillance video and play all of the movement events in a single video. This greatly reduces the need for surveillance video of the consumption of human resource. This will improve monitoring efficiency and is good for security.

This article first introduces the main function, scope and exiting problems of the software. Then introduces the processing of the software. At last, the article detailed analysises the modules of the software and the algorithm theory.

**Keyword**: Video summary, Moving object detection, Intelligent monitoring

**第一章 绪 论**

**1.1 背景与意义**

近年来，随着社会的迅速进步和国力的不断增强，社会生产中各行各业对安全防范和现场监控记录需求与日俱增，并且要求越来越高。虽然监控系统己经广泛地被使用，但是现有的视频监控系统通常只是录制视频图像，提供的视频信息都是监控场景原始视频片段，不管是人工实时监控或者是非实时查看监控视频都会耗费很多的人力资源。

此软件名为《智能视频摘要软件》，主要功能是分析视频，运动物体跟踪和检测，运动事件的提取，单个运动事件播放，全部运动事件播放等。此软件主要应用于监控视频的分析与提取，比如说一个24小时的监控视频，在这24小时内进入这个场景的人或者其他运动物体不频繁或者很少，就可以用这个软件将这些运动事件提取出来，然后用户查看监控视频的时候就可以有选择性的一个个运动事件去播放，或者将这些事件整合到同一个视频里面一起播放。这样做的好处就是不用把原始的24小时的监控视频从头到尾播放一次，方便使用者监控视频，节省很多时间。

这个软件比较适用于场景不复杂的监控视频，比如说室内监控，一些偏僻的街道监控等。不适用于人来人往的闹市，这些运动物体太多太复杂，如果一个监控视频每时每刻都有运动物体，那用这个软件提取出运动事件也就没意义了，还不如从头到尾看完监控录像。

**1.2 国内外同类产品的技术情况**

现有的智能视频监控技术主要有：物体识别[7]，车牌识别[16]，车流统计，人脸检测和打架等反常行为识别等。以上这些智能视频监控技术都带有确定性，监控目标具有一定的颜色、形状等特征，其未能解决的问题是，如果你不知道某一物体的具体特征的时候，就要以人工的方式去将一个监控视频从头到尾查看一次，这样非常耗费人力资源和时间资源。

**1.3 本论文的主要工作与内容框架**

本文针对智能视频分析进行了一些研究，提出了一种提取视频摘要的方法，阐述了智能视频摘要软件的基本原理，以及对保存一个事件的数据结构进行描述，还有本地分析文件的数据格式等。然后介绍了怎么将各个技术环节以及系统模块逐个整合起来，最终得到一个完整的智能视频摘要软件。在论文最后，还对智能视频摘要软件进行了测试，并撰写测试报告。

**第二章 核心算法：帧间差分法**

**2.1 帧间差分法的基本原理[2]**

帧间差分法是一种通过对视频图像序列中2个或者3个相邻帧采用基于像素的时间作差分运算来获得运动目标轮廓的方法，它可以很好地适用于存在多个运动目标和摄像机移动的情况。当监控场景中出现物体运动时，相邻的帧之间会出现较为明显的差异，相邻的两帧相减，得到两帧图像亮度差的绝对值，判断它是否大于一个设定的阈值来判断视频或图像序列是否有运动特性，确定图像序列中有无物体运动。这里的帧差法主要是用于检测出视频中的运动物体最小矩形轮廓。

**2.2 帧间差分法进行运动物体检测的原理以及流程[14]**

首先收集某一段时间内的原始视频序列图像，，…，用于运动物体的检测。设，，是其中的3幅连续图像（1<=a<=n-2，b=a+1，c=b+1），分别对应图2-1中的图（a）、（b）、（c）。如图2-2则是，，经过灰度处理的灰度图像，。然后用先作为基础背景，进行 — 运算，再将这2者的差值进行二值化处理后得到图2-3（a）所示的二值图像；同理再用作为基础背景，进行 — 运算，再将这2者的差值进行二值化处理后得到图2-3（b）所示的二值图像。图2-3（a）（b）两幅图中的白色像素块就是运动物体的大致轮廓，然后我们再用opencv中给出的轮廓提取函数[13][15]将白色像素块的矩形轮廓大小和左上角的坐标提取出来，然后在原始图像上面用检测出来的矩形框圈出运动物体，就完成了运动物体检测的过程，如图2-4是被圈出的运动物体。



(a) (b) (c)

图2-1 视频原始序列图像



（a） （b） （c）

图2-2 视频灰度序列图像



（a） （b）

图2-3 视频相邻两帧图像灰度图之差



（a） （b）

图2-4 运动检测结果

**2.3 本章小结**

本章主要描述了智能视频摘要软件核心算法，即帧差法的简要介绍，帧差法的主要原理，并且给出了算法的主要流程，算法的截图结果等。

**第三章 软件需求**

**3.1 需求描述**

智能视屏摘要软件旨在开发一个软件系统，该系统能将视频中的运动事件作为视频的摘要提取出来并可以供用户查看摘要事件。智能视频摘要软件主要分为2个大功能：视频摘要提取（视频分析）和视频摘要的播放。

视频摘要提取功能，也可以说成视频分析功能，就是用运动物体检测与跟踪的方法对视频进行分析，将视频中所有运动事件提取出来并作为视频的摘要，然后将摘要信息保存。

视频摘要的播放，也就是将视频中的视频摘要重播再现。这里分为两种方式：一种是单个摘要事件的播放，用户可以在第一个功能中生成的事件列表选择某一个具体事件进行播放；另一种是所有事件的播放，这里将视频中的所有摘要事件集中在同一个视频中播放，并对立面的事件进行半透明处理，给每一个事件进行标号，以便识别。

如图3-1是智能视频摘要软件系统的用例图。



图3-1 系统用例图

**3.2 本章小结**

本章先是介绍了智能视频摘要软件的功能需求，并对它们的输入输出等进行了细化说明，然后给出了整个软件系统的用例图以及整体流程图，最后简单的说明了运行该软件所需要的硬件环境和软件环境。

**第四章 系统设计**

**4.1 系统用户界面设计**



图4-1 智能视频摘要软件主界面

**系统主界面**如图4-1所示。

①“选择视频文件”按钮：选择一个本地视频文件；②“选择文件夹”按钮：选择一个包含多个视频的本地文件夹；③显示所选择的文件路径；④“分析视频”按钮：点击此按钮进行视频分析；⑤显示分析视频的进度条；⑥显示分析结果；⑦“播放选中事件”按钮：播放选中的事件；⑧“播放所有事件”按钮：播放所有事件；⑨显示摘要事件的列表框。

**4.2 系统模块设计**

表4-1 系统模块设计表

|  |  |  |  |
| --- | --- | --- | --- |
|  | 系统模块名称 | 模块功能名称 | 在系统中的类结构 |
| 主程序 | UI模块 | 软件系统界面 | VideoMainForm.Designer.cs**、**VideoMainForm.cs  PlayForm.Designer.cs、PlayForm.cs  Program.cs |
| 界面事件响应 |
| 视频处理模块 | 分析视频 | VideoAnalyzeProcess.cs  EventNode.cs  Global.cs |
| 播放单个/所有事件 |
| 生成所有事件视频文件 |
| 本地文件操作模块 | 本地文件读入 | FileOperation.cs |
| 本地文件写入 |

如表4-1阐述了只能摘要软件系统的3大主模块：UI模块、视频处理模块和本地文件操作模块，其中每个模块又被分为数个小的功能模块。

**4.3 系统数据结构设计**

由于此智能视频摘要软件是一个单机版的视频分析软件，要是采用数据库的方式来存储视频摘要事件的信息就要求用户对数据库配置有一定的了解，考虑到数据库配置起来比较麻烦，为了让用户更方便地使用这个智能视频摘要软件，所以采用将视频摘要信息保存到本地的txt文件这种保存方式。

**4.3.1 EventNode类结构设计**

表4-7 EventNode类数据结构

|  |  |  |
| --- | --- | --- |
| **变量类型** | **变量名** | **变量的含义** |
| int | startFrame | 视频摘要事件的开始帧。 |
| int | endFrame | 视频摘要事件的结束帧。 |
| List<Rectangle> | trackList | 一个Rectangle类型的List列表，保存了整个视频摘要事件每一帧（或者间隔n帧）的矩形轮廓以及矩形轮廓左上角的坐标。 |
| Rectangle | rect | 此变量在运动物体检测与跟踪算法中，保存了上一帧中运动事件的矩形轮廓以及其左上角坐标。 |
| bool | mark | 此变量在运动物体检测与跟踪算法中，标记了该视频摘要事件是否已经完结。 |

如表4-7写出了EventNode的数据结构，一个EventNode变量记录了一整个视频摘要事件的所有信息。主要变量是startFrame、endFrame、trackList3个，前两者分别表示事件的开始帧和结束帧，然后trackList主要是保存从开始帧到结束帧这个帧段中摘要事件在每一帧（或者间隔n帧）中出现的位置以及其矩形轮廓大小。这种数据结构的好处是节省空间，与把摘要事件作为连续的图片提取出来相比，只是保存了事件出现的位置以及轮廓大小的EventNode节省的空间要大好多，一个是保存图片的List，一个是保存矩形轮廓以及轮廓左上角位置。

如表4-7中的Rectangle类型是C#中自带的类，它表示了一个矩形的长宽以及矩形左上角的坐标，这里使用这类变量来存储矩形轮廓的大小和位置。

**4.5 本章小结**

本章主要是介绍了整个软件系统的设计，包括软件界面设计，系统内部模块设计，并对系统的各个模块进行深入分析，然后还介绍了本软件系统所用到的数据结构，最后是列举了整个系统的代码清单文件。

**第五章 软件测试**

**5.1 测试环境**

CPU：AMD双核，2.30GHz

内存：DDR2代，3.25GB

显卡：ATI Radeon HD 3850，512MB

硬盘空间：10G以上的硬盘空间，用以存放视频结果

**5.2 测试结果分析**

**5.2.1 压力测试**

这里所谓的压力测试就是测试这个智能视频摘要软件的瓶颈所在。下表5-1的数据用来显示当视频长度不一样或者事件数不一样的时候软件的运行情况：

表5-1 压力测试表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **视频编号** | **视频长度/时：分：秒** | **分析时间/时：分：秒** | **事件个数/个** | **内存最大占用量/MB** | **CPU平均占用率** | **软件运行情况** |
| 1 | 00:05:09 | 00:00:30 | 25 | 148.4 | 48% | 正常运行 |
| 2 | 00:30:00 | 00:03:50 | 18 | 110.5 | 47% | 正常运行 |
| 3 | 01:01:00 | 00:7:30 | 100 | 820.6 | 51% | 正常运行 |
| 4 | 01:00:00 | 00:10:17 | 199 | 1557.9 | 52% | 内存溢出 |
| 5 | 04:47:00 | 00:15:44 | 20 | 147.1 | 48% | 正常运行 |
| 6 | 09:36:00 | 00:25:44 | 29 | 185.7 | 51% | 正常运行 |

从表5-1中可以看出（1）不管其他测试参数是多少，CPU的占用率都基本维持在50%左右；（2）同时从表中“视频长度”那列和软件“软件运行情况”那列的结果来看，软件可以处理很长的视频这是无可非议的；（3）从“事件个数”、“内存最大占用量”和“软件运行情况”这3列可以明显看出视频的摘要事件数量越多，所需要的内存就越大，4号视频有199个事件，内存占用了1557.9MB，直接导致内存溢出。

所以我得出的结论就是：智能视频摘要软件的瓶颈就是摘要事件数以及计算机的物理内存大小。后来我控制事件个数进行测试，发现在150个事件左右的时候内存会溢出，就是说此软件目前最多能处理150个事件以内的监控视频。

**5.2.2 速度测试**

速度是指软件分析视频的速度，这里主要研究智能视频摘要软件的分析速度跟什么因素有关。

速度测试一：从上面的表5-1可以知道软件分析一个视频所需要的时间随着视频长度增加而增加的，并且不难得出，当视频的其他参数条件（分辨率，单位时间内的事件数，视频帧率）都一样时，视频分析时间跟视频的总长度是成正比的。

速度测试二：

个

秒



图5-1 事件个数与分析时间关系图

如图5-1所示，横坐标表示事件个数，纵坐标表示分析所需时间。这个测试我用了23个视频，每个视频除了事件数的其他参数都相同：帧率=15fps，长度=30min，分辨率=320\*480。由图5-1可以看出，软件分析视频的时间是随视频摘要事件数的增加而增加的，因为测试环境是电脑，有很多不稳定因素，他们的关系曲线大致可以拟合成一条直线。

**5.3 本章小结**

本章主要介绍了软件的运行流程以及对应的运行效果，并对软件进行了性能测试报告，说出了软件的瓶颈所在和分析了软件的运行速度跟其他因素之间的关系等。

**参考文献**

[1]王建平,刘伟,王金玲.一种视频运动目标的检测与识别方法[J].计算机技术与自动化,2007,26(3):78-81

[2]石时需,郑启伦,黄翰.基于自适应混合差分的快速视频目标检测法[J].计算机科学.2008(7):224-226.

[3] Tao Xiang, Shaogang Gong. Video Behavior Profiling for Anomoly Detection. Pattern Analysis and Machine Intelligence, IEEE Transactions on. Volume 30, Issue 5, May 2008 Page(s):893-908.

[4]田莘.基于MeanShift算法的目标跟踪问题研究[D]，西安科技大学，2010年.

[5]王建宇,陈熙霖,高文,赵德斌.背景变化鲁棒的自适应视觉跟踪目标模型.软件学报.2006,17(5):1001-1008.

[6]吴思,林守勋,张勇东.基于动态背景构造的视频运动对象自动分割.计算机学报.2005,28(8):1386-1392.

[7]李海波.视频监控中基于人的检测与研究.北京邮电大学,2010.

[8]肖思兴.复杂场景下的运动目标检测与跟踪算法研究[D].厦门大学,2009.

[9]王长军.[基于视频的目标检测与跟踪技术研究](http://cdmd.cnki.com.cn/Article/CDMD-10335-2006073629.htm)[D].浙江大学,2006.

[10]曹志刚,李宇成.[运动目标的前景与背景实时分割](http://cpfd.cnki.com.cn/Article/CPFDTOTAL-YJZD201005001235.htm)[A].全国冶金自动化信息网2010年年会论文集[C],2010.

[11]张晓晖,刘丁.仿人视觉的机器人视觉[J],计算机应用于研究.2007(18):11-14.

[12]张玉芳,毛嘉莉,熊忠阳.一种改进的K-means算法[J].计算机应用,2003,23(8):31-33.

[13] Huang Maochi, Yen Shwuhuey. A Real-time and color-based computer vision for traffic monitoring system [J]. IEEE International Conference on Multimedia and Expo(ICME), 2004(3):2119-2122.

[14] Thongkamwitoon T, Aramvith S, Chalidabhongse T H. An adaptive real-time background subtraction and moving shadows detection [A]. In: IEEE International Conference on Multimedia and Expo (ICME) [C]. 2004:1459-1462.

[15] Stauder J, Mech R, Ostermann. Detection of moving cast shadows for object segmentation [J]. IEEE Transactions on Multimedia, 1999, 1(1):65-76.

[16]肖梅,韩崇昭,张雷.交通监控系统中基于多源信息融合的运动阴影检测[J].西安交通大学学报,2005,39(10):1077-1080.

[17]工亮,胡卫明,谭铁牛.人运动的视觉分析综述[J].计算机学报,2002,25(3):225-237.

[18] KALRA S, CHONG MN. Bidirectional motion estimation via vector Propagation [J]. IEEE Trans. Circuits Syst. Video Technol, 1998(8):976-987.

[19] KALRA S, CHONG MN, KRISHNAN D. A new auto-regressive(AR) model based algorithm for motion picture restoration[A]. IEEE ICASSP97[C], 1997(4):2557-2560.

[20]李跃发. 基于视频的车辆检测与跟踪技术研究[D].西安电子科技大学,2007.