1. 题目介绍，背景和意义

各位评审老师，各位同学们好，我的论文题目是《基于摘要的监控视频信息检索系统》。首先我介绍一下选题背景和意义。

随着信息技术的不断发展和社会各方面的安防需求，监控摄像头被安装在每一个角落，很多地方都换上了高清监控摄像头，使得监控视频包含了更多的信息。但是在一些场景下监控视频往往无人问津，存储了大量的无用数据，比如说华工南校区的正门这种很少有人经过的地方，一天的监控视频绝大部分时间都在拍着一个无人的背景。要事后取证查看这样的视频也要花费很多人力资源和时间资源。

为了解决这样的问题，本文研究并开发出了一个《基于摘要的监控视频信息检索系统》，用于相对不复杂的监控场景的实时监控或者本地监控视频分析，有视频摘要和事件检索功能，方便监控人员快速浏览和查找监控视频的所有事件。

1. 系统两大模块，视频摘要和事件检索

视频摘要是视频摘要是对视频内容的一个浓缩，通过对视频内容进行分析，仅提取视频中有意义的内容，然后形成视频摘要，达到短时间浏览长视频的效果。视频摘要的主要特点是可以将不同时间的运动事件合并在一起播放，监控人员可以快速直观地看到某个时间段内所有运动物体，而不用重新浏览整个监控视频。

视频摘要主要有3大步骤，运动物体检测，运动物体跟踪和视频摘要生成。

运动物体检测是非常重要的一步，其检测效果直接影响到运动物体跟踪效果，运行速度影响到是否满足实时监控的需求。本文实现并对比了六种运动前景提取方法，基于背景差分的有三种，分别是中值法背景建模、均值法背景建模和混合高斯背景建模；基于帧差法也有三种，分别是传统帧差法、三帧差分法和与运动历史图结合的帧差法。

这是六种方法获取的前景图，这里的图像操作都是基于灰度图的，从彩色RGB图转化为灰度图待会会讲到。在背景建模方法中，中值法和均值法都比较容易留下拖尾，容易跟其它运动物体连接到一起，而混合高斯背景建模方法获取到的前景图像最好。基于帧差法的方法中，帧差法会有图像空洞和重影问题，三帧差分法虽然解决了重影问题，但是使得检测到的前景像素非常少，有可能一个运动物体会分解为两个运动点团，与运动历史图结合的真差法能很好的填充图像中的空洞，但是没有解决重影的问题。

这个表是对六种前景提取算法的效率对比，可以看到基于帧差法的三种算法都能满足实时监控的需求，而效果最好的混合高斯背景建模方法处理速度太慢，和前景效果综合比较后我选择了与运动历史图结合的帧差法来做运动物体的前景检测。

图像灰度化操作一般都是用这里的公式1 从RGB彩色空间转换到灰度空间，但是这条公式中都是浮点数乘法操作，比较耗时，因此可以将其转化为公式2的形式，将系数全变为整数，但是公式2引入了除法操作，为了使运算效率更高，这里继续将除法操作转化为位移操作，得到公式3。这3个公式的转化结果是基本一样的，而公式3的转化速度比其它公式都要快。

先将检测到的运动物体前景二值图进行中值滤波去噪，消除大部分的椒盐噪音。

然后再将去噪后的二值图像进行形态学闭操作处理，形态学闭操作可以填充图像内部微小缝隙，而不改变图像大小，这样使得同一个运动物体不会因为微小缝隙而断裂开来。

连通区域检测是用线段编码的方式进行，将每一行中相邻的像素点标记为一条条线段，然后逐行扫描这些线段，将相连的线段标记为同一个连通区域，也就是一个运动点团。

运动点团跟踪采取矩形框跟踪的方式，用一个最小外接矩形表示该运动物体，然后根据运动的连续性，同一个物体连续两帧的矩形框会有很大的重叠面积，如图（1）所示，当重叠面积都大于前后两帧的矩形框A和B的60%时，就判定为同一个物体。

但是实际运动情况往往复杂得多，本文采用分情况处理的方式，对运动点团分裂、合并以及交叉运动的两个物体这几种情况进行处理。

每一个运动事件记录其开始帧，结束帧和运动物体在每一帧中对应的最小外接矩形框。

视频摘要生成过程就是把不同时间发生的运动事件拼接到同一时间上去，使得监控人员可以用最短时间浏览完所有运动事件。

1. 视频内容检索

视频内容检索是在视频摘要的基础上的，在记录了一个监控视频所有事件的信息后，就可以针对监控视频中的这些事件的片段进行检索，而不用对整个监控视频检索，大大缩小了检索时间。

视频内容检索的特征包括事件发生时间、入侵区域、运动方向、目标类型以及颜色特征。由于对每一个运动事件都记录了其开始帧，以及每一帧中矩形框的位置，因此可以很简单的得到其发生时间、运动方向以及入侵的地方。

由于手上的监控视频大部分是人物或者车辆的视频，因此这里的类别只分人物、车辆和其它。本文采用opencv中自带的基于Haar特征的级联分类器来训练人物和车辆的分类器，人物的样本和车辆的样本如ppt中所示。然后对人物分类器和车辆分类器分别进行准确率统计，人物分类器的准确度为78.1%，车辆分类器的准确度为93.7%。人物分类器准确度低是因为在监控视频中由于摄像头位置的差异导致拍摄到的行人差异很大，很难训练出一个适用于所有场景的人物分类器，但是在某一场景下训练出来的行人分类器检测同一场景时能达到很好的效果。车辆分类器准确率较高的原因是现在的交通监控摄像头拍摄车辆的角度都基本比较固定，使得车辆有明显的共同特征。

颜色检索是让用户先输入一个参考颜色，然后找出主颜色与该参考颜色相同的运动事件。本文采用文献[27]中杨振亚等人提出的RGB颜色空间的矢量-角度距离色差公式来计算两个RGB颜色的色差，如果一个运动前景图中有一定数量的像素点颜色与参考颜色相似，则判定此运动事件为目标事件。

现在演示一下系统的运行界面，首先这是实时监控的界面，虽然系统支持实时监控，但是结果查看还是要到本地模式。现在读入一个事先分析好的监控视频，在右侧是显示事件列表，每一个事件都有一个标号。这里是查看视频摘要，在查看视频摘要的时候，如果对某个运动物体感兴趣，可以根据其左上角的事件编号，在事件列表中回溯到原始的视频中去查看。最后是检索功能，在这个界面上选择需要进行检索的特征，比如说找出所有往这个方向运动的事件，然后检索即可在事件列表中得到相应的结果。

1. 系统测试

最后是对整个系统进行测试，这个是对系统在视频摘要模块的运行速度测试，可以看到当视频分辨率为1280\*720时处理速度达到40帧/秒，满足了该高清分辨率下实时监控的需求。

视频摘要最忌漏掉运动事件，所以本文也对视频运动事件检出率的一个测试，事件的平均检出率达到95.7%。