**发明专利申请公布说明书**

发明名称

基于时空融合的智能提取视频摘要方法

摘要

本发明提供一个基于时空融合的智能提取视频摘要方法，用户选择一个源视频，采用基于时空融合的智能提取视频摘要方法对视频进行视频摘要提取，然后将所有视频摘要都保存到一个包含所有视频摘要的新视频中，用户可以通过此新视频来观看源视频的所有视频摘要。

**权利要求书**

1. 一种基于时空融合的智能提取视频摘要方法，其特征在于，包括以下5个步骤：
2. 从本地磁盘读入一个avi格式的视频文件；
3. 使用帧差法对读入的视频进行运动物体检测，得到运动物体的矩形轮廓；
4. 使用矩形轮廓匹配来对步骤(b)中检测出的运动物体的矩形轮廓进行精确跟踪；
5. 将步骤(c)中跟踪的运动物体的整个运动过程作为一个视频摘要，采用特定数据格式记录每一个视频摘要的详细信息；
6. 在步骤(d)中对整个视频的视频摘要记录完毕后，将记录的每一个视频摘要从源视频提取出来，并将所有视频摘要合成到同一个新视频中，得到一个包含所有视频摘要的视频文件。
7. 权利要求1中基于时空融合的智能提取视频摘要方法，其特征在于，步骤(a)中的视频文件格式要求为avi编码格式的视频。
8. 权利要求1中基于时空融合的智能提取视频摘要方法，其特征在于，步骤(b)包括以下6个步骤：
9. 在视频中获取连续的两帧图像；
10. 将这两帧图像转化为灰度图像；
11. 将这两帧图像的灰度图进行做差操作，得出差图像；
12. 对差图像进行二值化操作；
13. 对二值化的差图像进行中值滤波操作，消除小的噪声；
14. 对图像进行矩形轮廓检测，所得到的矩形轮廓就是视频中运动物体的矩形轮廓。
15. 权利要求1中基于时空融合的智能提取视频摘要方法的步骤(c)中，其特征在于，在第i次使用权利要求1中步骤(b)检测出的运动物体的矩形轮廓集合和在第i+1次使用权利要求1中步骤(b)检测出的运动物体的矩形轮廓集合进行匹配，轮廓匹配度大于70%的识别为同一个运动物体。
16. 权利要求4中所述的轮廓匹配度，其特征在于，轮廓匹配度的计算方法为P = ，P为2个矩形轮廓的匹配度，为2个矩形轮廓重叠部分的面积，为第一个矩形轮廓的面积，为第二个矩形轮廓的面积。
17. 权利要求1中基于时空融合的智能提取视频摘要方法的步骤(d)中所述的视频摘要，其特征在于，如果视频摘要的长度太短，将会被认为是噪音而被过滤掉。
18. 权利要求1中基于时空融合的智能提取视频摘要方法的步骤(d)中所述的特定数据格式，其特征为包括以下4个内容：
19. 摘要事件的开始帧号；
20. 摘要事件的结束帧号；
21. 摘要事件在每一帧中出现的矩形轮廓左上角坐标；
22. 摘要事件在每一帧中出现的矩形轮廓的高和宽；
23. 权利要求1中基于时空融合的智能提取视频摘要方法中，其特征在于，步骤(e)包括以下5个步骤：
24. 读取权利要求1中基于时空融合的智能提取视频摘要方法的步骤(d)中保存的第n个视频摘要信息；
25. 精确定位到源视频中第n个视频摘要的起始帧；
26. 遍历此视频摘要的每一帧，复制视频摘要在每一帧中的矩形轮廓区域的图像，也就是运动物体，得到一个连续的图像序列；
27. 重复步骤(e-1)到步骤(e-3)，直到将视频中所有视频摘要处理完毕；
28. 将步骤(e-3)中得到的图像序列按照其矩形轮廓左上角的坐标以及序列号粘贴到另一个所有视频摘要的图像序列中，并将所有摘要事件的图像序列保存为一个avi格式的所有视频摘要文件。
29. 权利要求8中的步骤(e-5)中，其特征在于，粘贴的时候会对图像进行透明化处理，以免遮挡了其他视频摘要，并且对每一个视频摘要进行编号，以便识别。

**说明书**

基于时空融合的智能提取视频摘要方法

**技术领域**

本发明一般涉及计算机智能监控视频处理领域。更具体的说主要用于将视频中的视频摘要提取出来合并到一个视频里面同时播放，用以减少监控人员查看视频的时间的基于时空融合的提取智能视频摘要方法。

**背景技术**

近年来，随着社会的迅速进步和国力的不断增强，银行、电力、交通、安检以及军事设施等领域对安全防范和现场记录报警系统的需求与日俱增，要求越来越高，视频监控在生产生活各方面得到了非常广泛的应用。虽然监控系统己经广泛地存在于银行、商场、车站和交通路口等公共场所，但实际的监控任务仍需要较多的人工完成，而且现有的视频监控系统通常只是录制视频图像，提供的信息是没有经过解释的视频图像，只能用作事后取证，没有充分发挥监控的实时性和主动性。并且一个监控视频往往就有几十个小时，而且监控摄像头也数不胜数，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，将视频监控自动化，智能提取视频的感兴趣部分能大大减少人力资源的占用。这对提高监控效率以及安全防护等起着很大的作用。

现有的智能视频监控技术主要有：特定物体识别，车牌识别，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，如果你并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。

**发明内容**

针对当前智能视频监控技术的不足，提供了一种基于时空融合的智能提取视频摘要方法。本发明的目的在于快速将一个监控视频里面所有视频摘要都提取出来，并且合并到一个视频里面，在不知道目标的具体特征的时候让用户用最短的时间来观看所有视频摘要，从而快速找出想要找的目标。

本发明提供了一种运动物体检测的方法，采用帧差法将视频中的运动物体检测出来，这是视频摘要提取的最基本操作。

本发明提供了一种运动物体跟踪的方法，其通过轮廓匹配的方式对检测出的运动物体轮廓进行跟踪，并将运动物体的整个运动过程作为一个视频摘要。

本发明提供了一种视频摘要的存储方法，通过一定的数据格式将视频摘要的起始帧号，视频摘要的结束帧号，视频摘要中每一帧的轮廓左上角坐标和视频摘要中每一帧的轮廓的高和宽记录下来，达到记录整个视频摘要的目的。

本发明提供了一种将一个视频中所有视频摘要合成到同一个新视频中的方法，通过读取所有记录的视频摘要信息，然后将每一个视频摘要从源视频中提取出来，再合成到同一个新视频中，得出一个包含所有视频摘要的视频文件，此视频文件长度等于最长的视频摘要长度，用户可以用最短的时间浏览完所有视频摘要。

本发明提供一种基于时空融合的智能提取视频摘要方法，其特征在于，包括以下5个步骤：

1. 从本地磁盘读入一个avi格式的视频文件；
2. 使用帧差法对读入的视频进行运动物体检测，得到运动物体的矩形轮廓；
3. 使用矩形轮廓匹配来对步骤(b)中检测出的运动物体的矩形轮廓进行精确跟踪；
4. 将步骤(c)中跟踪的运动物体的整个运动过程作为一个视频摘要，采用特定数据格式记录每一个视频摘要的详细信息；
5. 在步骤(d)中对整个视频的视频摘要记录完毕后，将记录的每一个视频摘要从源视频提取出来，并将所有视频摘要合成到同一个新视频中，得到一个包含所有视频摘要的视频文件。

根据本方法的步骤(a)，本发明支持所有avi格式的视频文件，其他格式的视频文件需要先转换为avi格式。

根据本方法的步骤(b)，要对步骤(a)中读入的avi格式视频采用帧差法进行运动物体的检测，得出运动物体的矩形轮廓，这是提取视频摘要的最基本操作，同时也是对运动物体进行跟踪的前提。

根据本方法的步骤(c)，采用轮廓匹配法来对步骤(b)中检测出的运动物体的矩形轮廓进行跟踪，以得出运动物体的整个运动过程。

根据本方法的步骤(d)，将运动物体的整个运动过程作为一个视频摘要，并将视频摘要的起始帧号，视频摘要的结束帧号，视频摘要中每一帧的矩形轮廓左上角坐标，视频摘要中每一帧的矩形轮廓的高和宽这些信息记录下来，并由这些信息记录一个完整的视频摘要。

根据本方法的步骤(e)，读取步骤(d)中得到的所有视频摘要信息，利用图像复制和粘贴技术，将这些视频摘要从源视频中逐个提取出来，然后逐帧合成到同一个新视频中，最终得到一个包含所有视频摘要的视频文件。这样就完成了视频摘要的提取，用户可以通过这个视频文件用最短的时间浏览完源视频的所有视频摘要。

现有技术存在的一些问题和本发明的相对优势如下：

1. 特征性太强

现在的视频处理技术一般都是搜索技术，这些技术要求你要知道目标的颜色、形状等其他特征，如果给出一个视频，但是没有给出相应特征，那根本无法对此视频进行智能化处理。

1. 特定技术对场景要求严格

现有的这些技术所能处理的视频对场景也有特定要求，能处理的视频场景也比较傲单一。比如一个特定的视频处理技术只能处理车流场景，而不能处理人物场景；另一个视频处理技术只能处理人物场景，而不能处理车流场景等，不能应用于多种场景。

1. 本发明的相对优势

和本发明比，本发明并不需要知道目标的具体特征，如果你在不知道目标特征的情况下就可以利用次方法迅速的浏览所有视频摘要，然后用最短的时间找到自己感兴趣的视频摘要。

本发明的优势还体现在，本发明对视频场景也没有特定的要求，可以同时处理人物场景，车流场景等多种场景的视频，用户不用担心因为视频场景不同而要换一种不同的提取视频摘要技术。

**附图说明**

附图，包括其中以提供本发明的进一步的理解并构成该说明书的一部分，说明本发明的实施例并与说明书一起用于解释本发明的原理。

在附图中：

图1为说明基于时空融合的智能提取视频摘要方法的流程图。

**具体实施方式**

在此，本发明的说明书附图将被详细说明。

如图1所示，基于时空融合的智能视频摘要方法的主要流程包括如下5个步骤：

1. 读入一个avi格式的视频文件；
2. 进行运动物体轮廓检测；
3. 进行运动物体跟踪；
4. 记录视频摘要信息；
5. 提取视频中所有视频摘要并合并到同一个视频文件中。

步骤(a)允许用户通过文件选择窗口从本地硬盘选择一个视频文件作为视频摘要提取源，此视频文件要求是avi编码格式。

步骤(b)采用帧差法对视频进行运动物体的检测，帧差法是目前运动物体检测算法中最快的算法之一，通过两帧连续图像的灰度图做差，从而将图像中的相同背景区域中的像素点设为0，剩下不为0的像素区域即使运动物体的轮廓，然后再将这些轮廓用最小矩形框圈出，最终得到运动物体的矩形轮廓。

步骤(c)采用轮廓匹配法进行运动物体的跟踪，轮廓匹配法主要原理是，一个运动物体在连续几帧内的矩形轮廓重叠率总是很高，这样对在相邻帧中检测出的运动物体矩形轮廓进行重叠率匹配，重叠率超过70%的标记为同一个运动物体的轮廓。如果在上一帧中的某个矩形轮廓没有在下一帧找到与之相匹配的矩形轮廓，则判定前一帧中的矩形轮廓所对应的运动物体结束运动。

步骤(d)将步骤(c)所跟踪到的物体的完整运动过程作为一个视频摘要，并且用一定的数据格式将每一个视频摘要存储下来。这里所用到的数据格式包括视频摘要的起始帧号，视频摘要的结束帧号，视频摘要中每一帧的矩形轮廓左上角坐标，视频摘要中每一帧的矩形轮廓的高和宽。这里是通过记录视频摘要在源视频中出现的时间段（这里用视频帧位置表示），用以快速定位到源视频中对应的位置，然后再将这个时间段中视频摘要在每一帧中出现的位置（即矩形轮廓的左上角坐标和高宽）记录下来，以达到用最少数据量记录整个视频摘要的目的。

步骤(e)是在步骤(d)的基础上，将记录好的视频摘要通过图像的局部复制技术将视频摘要每一帧的最小矩形区域图像从源视频中提取出来，然后用图像粘贴技术将这些提取出来的图像逐帧合并到一个新的视频中，这样最终得到一个包含源视频所有视频摘要的视频文件，用户可以通过这个视频文件快速浏览源视频的所有视频摘要。

**说明书附图：**

读入avi视频文件

记录视频摘要信息

进行运动物体轮廓检测

进行运动物体跟踪

提取视频中所有视频摘要并且合并到同一个视频文件中

结束

图1