

基于图像处理和摄影测量的 空间视频对象的检索方法

肖 辉 编译

摘 要 为了支持视频扫描参照和双向搜索,必须检索每幅图像帧的空间对象。可是检索每秒30帧的图像的所有帧的空间对象要花费很多时间。因此我们采用内插的方式检索空间对象的其它帧,任意1帧的空间对象检索仅要约1秒。可是处理1秒间隔的图像帧,视频空间对象空间冗余较低,使检索空间对象变得困难。因为几乎不可能仅基于方法的图像进行对象的跟踪。本文为图像帧的空间对象设计了一个有效的索引方法,结合了基于该方法的图像摄影测量的方法。此方法可以有效的检索图像帧的空间对象以及能够恢复视频空间对象的内容。

关键字 GIS 视频索引 对象跟踪 摄影测量 图像处理

因为多媒体技术的发展,描绘空间数据的视频数据能够很容易的获取、保存和分发。视频是描绘空间信息的有效媒体,应用在很多地理信息系统(GIS)中。视频的引入是GIS的一个新的方向。关于视频空间数据综合有很多研究。视频非常庞大的信息的优点是能够提供比常规媒体如地图更加逼真而更符合人的感知的信息。

因为视频本身并没有空间对象的绝对信息,空间对象的视频检索必须能够处理空间信息的查询。因为GIS的功能主要是查询空间对象,不同于普通的基于恢复多媒体的内容的检索方式。为了进行检索,提取视频对象特有区域(如轮廓的表达)是很重要的问题。

人工方法生成轮廓是不可行的。本文提出对视频中空间对象的半自动对象跟踪方法,但对速度为每秒30帧的视频这样的方法就不够了,因为需要非常多的处理时间和用户交互。

为了解决上述方法存在的问题,我们要减少已经获得输出的相同对象的处理时间。本

文根据跟踪1秒间隔的分散帧中的空间对象而生成对象轮廓,其它帧内插生成对象轮廓。因为帧间隔1秒使空间冗余变小,使跟踪空间对象变得更困难,所以我们提出了结合摄影测量和图像处理的方法。跟踪的结果用于生成空间对象的轮廓和高效率的图像检索。

1 相关资料

1.1 针对空间数据的视频的恢复

对于空间数据的视频的恢复分为两种类型。图像的特殊对象的视频帧可以恢复(图到视频恢复)。这样的帧的对象的区域还可以显示出来。这种恢复可以生成的空间对象比图(纸质或数字形式)实际外观和更加逼真的信息。相反,视频帧的特地区域和对象的特征可以恢复了(视频到图恢复)。

为处理恢复,要描述的空间对象的信息有:像素的坐标,世界坐标和特征。在这些信息中像素的坐标是最重要的,也是两种类型恢复的本质。信息为描述对象的最小矩形或者连接对象的特征点(称为轮廓)。我们以前的

研究中提出支持恢复对象轮廓的处理的 MPEG-7 元数据。本文提出的对象跟踪方法的目的是为视频检索生成对象轮廓。

1.2 摄影测量

摄影测量通过相机参数、量测像片坐标和地面控制的严格准确的计算得到对象空间的点的坐标。共线条件为一直线上的物方空间点 (X_A, Y_A, Z_A) 、像片空间的对应点 (X_a, Y_a, Z_a) 、透视中心 (X_L, Y_L, Z_L) 的关系。空间后方交会是确定像片的 6 个外方位元素 $(\omega, \phi, \kappa, X_L, Y_L, Z_L)$ 的方法。通常参数包括透视中心的物方空间坐标 (X_L, Y_L, Z_L) 和三个角元素 (ω, ϕ, κ) 。角元素表达了物方空间坐标系与像空间坐标系之间的关系。这种解算方法至少需要三个控制点即已知 XYZ 物方空间坐标。通过空间后方交会确定立体像对的像片的外方位元素。重叠的立体区域的物方空间点就可以计算出来。进行空间前方交会即两张像片上的相同物方点所对应的光线将交会在同一个点上。空间后方交会和前方交会是建立在共线条件之上的。

2 空间对象索引

我们提出的方案的基本的方法是对 1 秒间隔的帧的空间对象进行跟踪, 其它帧通过内插方法生成空间轮廓。可是, 1 秒间隔的帧的空间冗余相对较低, 使得跟踪空间对象变得相对困难。因此为了高效率的跟踪帧对象, 我们提出了混合的对象跟踪方法, 即结合摄影测量和图像处理的方法。此方法的总的流程见图 1, 详细的介绍将在下面的部分讲述。

2.1 利用摄影测量方法计算特征点

上面部分论述的摄影测量方法可以应用到运动物体绘图系统采集的立体视频中。利用共线条件使特征点的象坐标转变为地面坐标。相反的, 任意帧的特征点的地面坐标也可

以转换为象坐标, 同样的也可以得到相应帧的相机方位和位置。两个坐标系统的转换可以通过摄影测量的方法得到, 理论定义为:

$$P(X, Y, Z) = \text{Intersection}(p_l, p_r, C_l, C_r, E_l, E_r) \quad (1)$$

$$p(x_a, y_a) = \text{Resection}(P, C_l, C_r, E_l, E_r) \quad (2)$$

其中 P 是地面坐标, p 是像坐标(左片和右片 p_l 和 p_r), C 是相机位置, E 是相机方位。通过两个坐标系统的转换和必要的参数就可以得到像对的空间对象的地面坐标, 反之亦然。跟踪视频中的空间对象的方法如下。首先, 利用两个函数得到特征点。帧中首先生成的对象, 人工输入其左片和右片的特征点。使用公式(1)和必要的参数将特征点转换为地面坐标。对于下一帧, 将这个地面坐标再利用公式(2)将其转换为像坐标。

显然, 由公式(1)(2)和其参数输入的数据是有误差的, 这是因为下一帧的特征点的寻找是有误差的。利用两个转换公式接近特征点的像素就可以找到。不能准确的确定, 但可以估计出任意其它帧的特征点位置。以前的实验表明地面坐标误差低于 50cm, 像坐标误差低于 30 个像素。下面部分将介绍如何改正特征点。改正了一个特征点, 那么其它帧的跟踪处理可以重复进行直到没有特征点为止。

2.2 改正特征点

下面部分将介绍空间对象的像坐标通过摄影测量的转换公式的计算。由于人工输入和转换公式存在的误差使得点的计算变得不准确, 所以必须对特征点进行改正。通过寻找计算得到的点周围的初始最小的特征点进行改正。

有很多方法可以改正特征点, 最常用的方法是用匹配的办法。匹配可以分为区域匹配

和灰度匹配。本文使用的 BMA 属于使用最为广泛的灰度匹配。与特征匹配不同,灰度匹配不需要提取特征图像信息,这意味着问题简单化了以及处理的花费降低了。

应用 BMA 存在下面两个问题。首先,如果人工输入了生成的第一帧一个特征点,那么特征点会很少,这可能是不准确的输入的原因。为了使输入造成的误差最小,将当生成较大的特征点时再输入。因此输入后面帧的特征点,相反的关系到对象跟踪时间。然后因为对象的大小和对象的帧与帧间的特征点的变化,BMA 变得相对困难。为了解决这个问题,我们改进了 BMA,使用可变大小的块。当特征点在下一帧找到时,减少特征点周围的初始块大小。利用这种方法,提高块匹配的精度和减少了处理的时间。关于 BMA 的改进如下。

搜索方向:为了减小人工输入的误差,倒转处理顺序,先进行对象跟踪。当特征点较多时再输入特征点。

搜索窗口的大小 搜索窗口的大小为 32×64 。这是因为误差主要是垂直方向的(车辆上的相机装备对垂直方向的震动比较敏感),而且像坐标的实验误差低于 30 个像素。搜索窗口的大小的确定与数据获取情况与车辆行驶速度和路面情况都有关系。

搜索方法:块搜索方法如 TSS 和 TDLS,它们的块大小都是 16×16 。本文的块匹配方法的块大小是变化的,所以本文的搜索采用另外的搜索窗口。

块大小:很多时候,块匹配方法的块大小为 16×16 或者 8×8 ,这取决于速度和精度之间的平衡。本文的块匹配方法的块的大小为 32×32 ,因为块的大小重采样过的原始块不同。

块大小的调整 为了提高匹配精度和减少处理时间,当在下一帧中寻找和计算特征点

时,减少重采样过的特征点周围的原始块的块大小。两帧间的块减少的比例取决于相机和物体的几何距离增加的比例。这通过相机和目标的几何位置可以得到,目标由运动物体绘图系统和摄影测量的公式得到。在实验中窗口大小一般减少为约为原始大小的 70%。

块匹配测度 灰度匹配有很多的测度如差绝对值和测度(SAD),零平均差绝对值和测度(ZSAD)和正交关系测度(NCC)。本文采用 ZSAD 为块匹配测度。ZSAD 定义为:

$$\sum_{(u,v) \in W} (I_1(u,v) - \bar{I}_1) - (I_2(x+u, y+v) - \bar{I}_2) \quad (3)$$

其中 I 是块的灰度。本文中利用到的灰度值是彩色图像的 RGB 值转换成 HIS(色调,饱和度,亮度)空间的灰度值。

2.3 目标分割

特征点跟踪后,进行目标分割(如图 1)。一个特征点计算和找到后,将此点作为种子点,通过对对象分割可以找到目标。利用区域生长,通过增加相似的相邻像素来增长由种子点形成的区域。相似性可以由亮度,颜色,纹理或者它们的结合来确定。本文的实验采用灰度级进行区域生长。在目标分割之前,为了得到更好的分割结果,依次进行收缩,腐蚀和收缩操作。注意到尽管本文利用了目标分割方法,但对于这些方法的比较超出了本文的研究范围。

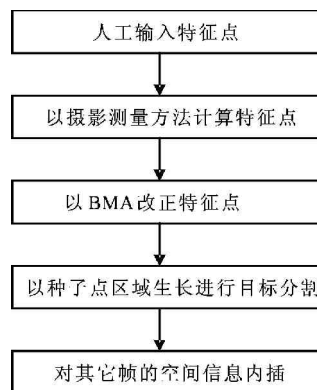


图 1 目标跟踪流程图

最终由处理帧间的其它帧生成目标的轮廓。这可以通过线性内插得到。由所有的帧生成一个轮廓 此轮廓可以记录为视频索引的元数据 此索引即为需要的结果。

3 实验

3.1 块匹配

块匹配的结果如图 2。窗口大小为 32×32。下一帧中实验的窗口分两种 固定大小窗口和可变大小窗口。图 2(a)是原始图像 ,图 2(b)是下一幅图像(1 秒之前)图 2(c)是用户输入的图 2(a)的特征点(输入包括左片和右片) ,图 2(d)计算的下一幅图像图 2(b)的特征点。此点标记为矩阵表示来通过摄影测量的办法计算的特征点 此点标记为十字丝表示 32×32 固定大小窗口的块匹配的结果 此点标记为圆表示可变大小窗口 22×22 的块匹配的结果。发现目标变得更小 ,可变大小窗口的块匹配找到更准确的目标点。固定大小窗口的块匹配的处理时间更短 ,且与窗口大小成比例。实验中相邻帧的窗口的尺寸减少约 70%。

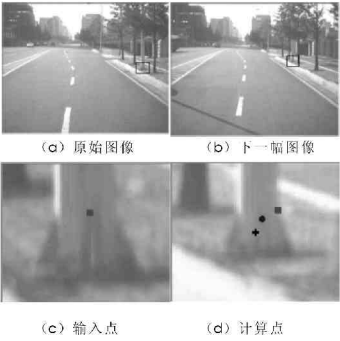


图 2 块匹配结果

图 2 显示自动生成的四帧目标跟踪的结果 ,速度为 1 帧/秒。图 3 中帧的标号对时间是颠倒的。对于其它帧继续进行初始特征点的计算通过相应的块匹配方法的改正。相应的特征点离直接输入的特征点仅仅远 1 到 2 个像素。图 3 中 ,目标跟踪和分割在第五帧失败了 ,因此没有显示。之所以失败是因为反复的转换导致的误差累计和对于块匹配来说目标变得太小。下面比较了采用可变尺寸块的块匹配结果的精度。块匹配的实验即用固定尺寸也采用可变尺寸 来观察结果。表 1 显示了每个块尺寸的块匹配的精度。比较人工计算的坐标和自动生成的坐标间的 RMS 误差值。RMS 较低的意味着精度更高。

表 1 块匹配精度比较

Block Sampling Scale (Block Size)		Frame 1 (User Input)	Frame 2	Frame 3	Frame 4
Fixed Size	1.0(32×32)	0	1.0656	1.8592	3.1223
	0.8(26×26)	0	1.08	1.7462	3.0314
	0.6(19×19)	0	0.05	0.0674	0.2386
	0.4(13×13)	0	0.907	3.8799	53.6092
Variable Size		0	0	0.0586	0.2021

3.2 目标分割

本文利用通过块匹配寻找到的特征点进行目标的分割,得到目标的形状。原始图像和特征点如图 4(a)和(b)显示。图 4(c)和(d)显示了矩形区域内的寻找到的单一的目标区域。此区域是本文方法的最终结果,记录为视频元数据的目标的轮廓。

4 结论

本文提出了一种新的目标跟踪的方法,可用于空间目标的视频索引。视频索引可以使 我们更加直接的浏览空间对象,提高 GIS 领域内的视频的有效性。

视频表示的信息是空间目标的真实的信息表达,可以建立双方向地图和视频间的搜

索。支持视频的空间目标搜索对视频元数据是必须的。视频元数据的大多信息是视频帧目标的轮廓,所以提出了生成目标轮廓的半自动方法。此方法通过块匹配和目标分割进行目标跟踪。因为此图像处理方法可能使得 1 帧/秒的视频的处理失败,提出了一种新的应用摄影测量的来弥补目标跟踪的方法。此方法通过摄影测量的公式得到象坐标的空间目标的预位置,改正块匹配方法,以得到精确的位置。调整块匹配窗口大小辅助执行。通过车辆视频检验此方法。

尽管此方法可以得到好的结果,但对摄影测量操作的精度和数据稳定性比较灵敏。如果计算出的特征点离真实的像素太远此方法会失败。今后必须用更好的方法解决此问题。

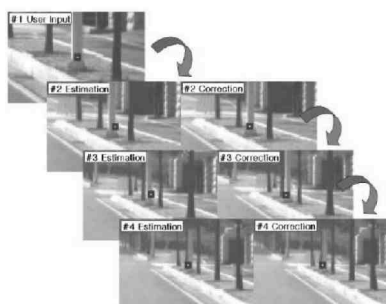


图 3 空间目标跟踪结果

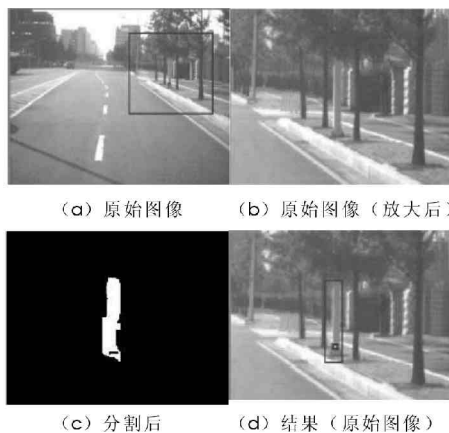


图 4 分割结果