

视频对象分割及跟踪方法研究

沈未名¹ 江 柳¹ 种衍文¹

(1 武汉大学多媒体网络通信工程湖北省重点实验室, 武汉市珞喻路 129 号, 430079)

摘 要: 讨论和分析了视频对象全自动提取及后续帧中的跟踪过程及采用的方法, 对其算法、分割性能和结果进行了比较和评述。并结合 Hausdorff 和 Snake 跟踪器, 对初始对象轮廓进行跟踪。结果证明, 这种跟踪方法能有效地解决对象被遮挡和形变以及背景移动的问题。

关键词: 视频对象; 对象分割; 对象跟踪; 光流; 活动轮廓; 聚类分析

中图分类号: TP391

1 视频对象分割技术

视频对象分割按照人工参与程度分为半自动分割和全自动分割。视频对象全自动分割根据其分割过程主要分为运动分割、时空联合法、变化检测三类。

1.1 运动分割法

运动分割主要通过研究光流场, 从序列图像中近似计算不能直接得到的运动场, 然后根据运动场的特征进行视频分割。运动场是一个纯粹的几何概念, 它描述了物体的三维运动在图像平面上的投影, 光流描述物体表面亮度变化速度的分布。光流与运动场不一定总是完全对应的, 如外部光照不变, 物体围绕自身旋转, 运动场不为零, 但是物体表面的亮度并没有发生变化, 光流处处为零; 相反, 物体没有运动, 运动场为零, 而外部光照发生变化, 导致物体表面的亮度发生变化, 从而光流不为零。但在理想情况下, 这种表现运动能够反映物体的运动情况。

利用图像时域和空域的一阶导数估算光流, 即图像上某一点灰度的一阶时间变化率等于场景亮度变化率与该点运动速度的乘积, 表征该定义的光流计算式称为约束方程。然而, 像素的运动速度是一个二维矢量, 仅有一个约束方程不足以惟一确定光流的两个分量。因此, 需要附加一个

约束条件。有关方法见文献[1~7]。

运动分割的优势是能够处理两个或多个物体运动的情况, 但其缺点是很难快速而准确地估算光流场, 而且仅有运动信息, 也无法准确定位物体的边缘。

1.2 时空联合法

在视频对象分割过程中, 完全采用运动信息而忽略空间信息将无法准确定位物体边缘。空间分割及区域跟踪方法能够在一定程度上解决这个问题, 从而保证分割结果的准确性和完整性。

时空联合法包含空间分割、区域跟踪和区域合并三个阶段。首先对图像进行降噪预处理; 然后将首帧图像根据空间特征, 如纹理、灰度分割成许多独立且灰度均匀的小区域, 并在后续帧中跟踪这些小区域; 最后根据相似性准则合并区域, 并进行后处理, 最终得到运动物体。

空间分割主要利用相邻像素间的灰度、纹理特征。目前应用较广泛的分割算法主要有动态聚类^[8,9]、基于数学形态学的分水岭算法^[3]、Snake 方法、分裂和合并以及统计方法等。

区域跟踪通过在后继帧中估算空间分割, 得到各个区域的运动矢量或者区域内各个像素的运动频率实现。文献[3]采用分层块匹配法构造运动场, 采用最小均方近似值计算区域中每个像素的运动参数, 区域的运动矢量由这些参数合成得到。由于参数是由多个像素结合在一起估算出来

的, 因而受噪声影响较小。

空间分割只是将视频图像分成了多个灰度均匀的区域, 并没有找到具有语义的运动物体。此时需要利用区域跟踪阶段得到的运动矢量或者运动频率, 根据相似性规则将相关区域合并成语义对象, 实现对象提取。进行区域合并的判定规则主要分为以下两种。

1) 运动频率。联合多个后续帧计算区域内每个像素的运动频率^[8], 当这个运动频率大于预设的阈值时, 此像素被认为是运动的; 然后计算每个区域内运动像素的个数, 以此确定区域是属于静止背景还是运动物体。此区域跟踪方法只适合静止背景的情况, 而且联合多帧进行跟踪必然会造成一定的延时。

2) 运动相似性。计算每个区域的运动参数, 根据均方运动补偿误差或其增量来衡量两个相邻区域的运动相似性。对每两个相邻区域 R_i 和 R_j 估计一组运动参数, 并计算用这组运动参数进行补偿时的均方运动补偿误差增量 $\Delta_{i,j}$; 对每两个相邻区域寻找 $\Delta_{i,j}$ 的最小值, 若该值小于预设的阈值, 则合并相邻区域。

Moscheni 等人^[9] 考虑到灰度相差不大的相邻区域属于前景物体的可能性极大, 并且同一区域使用相邻区域的运动参数集得到的残余运动能够可靠地反映出运动特征的相似程度, 因而提出一种基于相邻区域灰度对比度的空域相似性判定准则和基于补偿区域与参考区域的加权残差分布的时域相似性判定准则。该算法具有鲁棒性和有效性。

1.3 变化检测法

该方法也称为帧差技术, 通过在当前帧与参考帧的互帧差基础上检测运动物体, 然后将空间信息应用到检测结果中, 并对其进行修正。该方法与时空联合法相似, 均结合了空间和时间信息。所不同的是, 变化检测法通常将检测到的运动信息作为分割前景对象的主要线索。

视频对象的变化检测根据图像亮度或梯度信息在后续帧中的变化情况实现。变换区域检测最基本的方法就是直接对相邻两帧之差进行阈值化, 将当前帧分为相对于后一帧的变化和未变化区域。但由于光照变化、视频采集仪器产生的噪声、成像过程中的信息丢失等问题的不可避免性, 很难确定统一、准确的阈值。具体见文献[10~13], 本文采用彩色多值算子, 避免了梯度算子导致的过分割以及对噪声敏感等问题。同时, 该算法能够自动形成闭合的区域, 提取较精确的物体

轮廓。但是该算法过分地依赖标记过程中阈值的选取。

1.4 其他分割方法

近年来出现的网格模型可以很好地处理旋转、缩放以及变形运动。根据图像的帧差、梯度^[14] 等特征提取特征点, 并依据约束 Delaunay 三角形法将这些特征点构造成网格, 或者将网格的各个结点平均分布在图像内部^[15], 通过最小化目标函数值将网格向符合图像特征的位置进行变形, 最终得到运动物体, 但计算复杂度大大增加。

Snake 又称作活动轮廓, 首先由 Kass 等人在 1988 年提出。随后, 许多学者对其模型进行了发展。经典的 Snake 模型通常被定义为内部能量和外部能量之和, 其中内部能量反映了曲线的弯曲度与平滑度, 外部能量则反映了图像的梯度、纹理等特征。轮廓最初设置在物体真实轮廓的附近, 随着能量函数值的逐渐降低, 轮廓向符合图像某些特征的位置移动, 能量最小化时的轮廓即为最终结果。由于常规 Snake 模型需要接近物体真实边界的初始轮廓值, 因而限制了其在视频对象全自动分割方面的应用。Besson 等人^[16] 突破了这种限制, 提出了基于区域并联合运动信息的活动轮廓方法, 初始轮廓设置为当前图像的边界。基于 Snake 模型的活动轮廓是图像分割的一个强有力的工具, 其分割过程也涉及到目标函数最优解的计算, 其所带来的运算量也是巨大的。

2 视频对象在后续帧中的跟踪

通过对初始分割结果进行跟踪, 从而实现后续帧的对象分割。

2.1 区域跟踪

区域跟踪方法的主要思想是利用区域的时间相关性实现对象的后续帧跟踪。Gu^[17] 提出一种基于区域分类的对象跟踪方法, 即将下一帧图像分割成灰度均匀的区域, 然后将这些区域向后投影到已产生对象的那一帧, 并进行运动估计。若补偿区域属于运动对象的面积大于背景的面积, 则该区域属于运动物体, 否则属于背景。Vigus^[18] 使用 Kalman 滤波器作为跟踪器, 将跟踪区域的坐标作为 Kalman 滤波器的初始输入计算区域在下一帧的位置, 并对 Kalman 滤波器的输出结果从内向外进行一个螺旋式的搜索, 找到最佳匹配区域。

此外, 部分文献中提到光流跟踪法, 即采用光流运动参数模型估计对象的运动轨迹。但由于对象在运动过程中可能会发生被遮挡、变形的情况,

所以光流不能很好地处理运动对象的非刚体运动以及遮挡问题。

2.2 边界跟踪

前面介绍的 Snake 技术在边界跟踪方面有其独特的优势,它可以处理任意形状物体的任意形变。首先将分割得到的物体边界作为跟踪的初始模板,然后确定表征物体真实边界的目标函数,并通过降低目标函数值,使初始轮廓逐渐向物体的真实边界移动,最终达到对象的真实边界。跟踪过程实际上是解的寻优过程,其带来的计算量也比较庞大,而且由于 Snake 模型的盲目性,对于快速运动的物体或者形变较大的情况,其跟踪效果不理想。

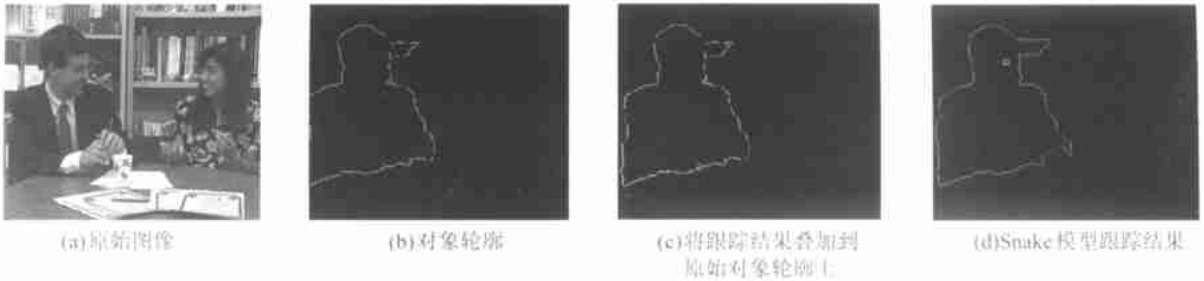


图1 跟踪结果

Fig. 1 Result of Tracking

3 结语

视频对象分割是综合人工智能和模式识别、图像理解、计算机视觉相关学科的一个研究方向,其目的是通过模拟人类视觉系统实现运动对象和背景的分离。这个过程需要利用对象的运动特征,联合对象的空间信息也是十分必要的。然而对象分割一直是机器视觉和图像处理领域中的一个难点,这是因为:①对象的定义不是惟一的,自然界的可视实体都能够作为分割对象;②对象提取实际上是一个分割的过程,众多学者认为其是一个最艰难的任务;③传统的同一性准则,如灰度、纹理等并不能导致语义物体的形成,对更复杂、更具有语义特点的同一性准则的开发需要进一步的研究;④依靠计算机检测和提取运动物体的技术还处在不成熟的发展阶段,利用现有的视频分割技术能够分离出单独的运动对象,但未能达到令人满意的效果;⑤视频分割算法的复杂度比较高。

因而,如何实现视频对象的语义分割尚待进一步研究,如对运动物体模型的精确估算、噪声影响的降低、轮廓的精确性提取等都是视频分割过程中应着重考虑的问题。同时,分割应该从更加综合的

Park 等人^[19]提出一种带有方向性的 Snake 模型,即根据边界曲线特征与图像特征,调节 Snake 曲线的发展方向,从而防止错误的边缘定位。

Hausdorff 跟踪器^[19]是一种模糊匹配的对象跟踪方法,即从最不相似的形状区域中找出最相似的形状区域。本文首先采用 Hausdorff 跟踪器找到对象轮廓在下一帧的最佳匹配位置,然后采用 Snake 模型拟合发生在该位置上的非刚性形变。这种跟踪方案不仅能够处理对象被遮挡和背景移动的情况,而且在一定程度上克服了 Snake 的盲目性。跟踪结果如图 1 所示。

角度,如实用化、系统化、工程化的角度考虑。

参 考 文 献

- 1 Nagel H H. Displacement Vectors Derived from Second-Order Intensity Variations in Image Sequences. *Computer Graphics Image Processing*, 1993, 21: 85~117
- 2 Castagno R. Video Segmentation Based on Multiple Features for Interactive and Automatic Multimedia Applications: [Ph. D Dissertation]. Italy: University of Trieste, 1998
- 3 黄波,杨勇.一种基于时空联合的视频分割算法. *电子学报*, 2001(1)
- 4 Zhu S, Ma K. A New Diamond Search Algorithm for Fast Block-Matching Motion Estimation. *IEEE Trans. on Image Processing*, 2000, 19(2)
- 5 Eckehard S, Peter E, Bernd G, et al. Motion-based Analysis and Segmentation of Image Sequences Using 3D Scene Models. *Signal Processing* 66 1998, 233~247
- 6 刘李杰,蔡德钧.一种面向运动的视频对象分割算法. *计算机学报*, 2000, 23(12)
- 7 Michael M C, Tekalp A M, Sezan M I, et al. Simultaneous Motion Estimation and Segmentation. *IEEE Trans. on Image Processing*, 1997(6): 1 326~1 333
- 8 黄波,杨勇,王桥,等.基于模糊聚类和时域跟踪的视频分割. *通信学报*, 2001, 22(12)
- 9 Moscheni F, Bhattacharjee S, Kunt M. Spatio-temporal Segmentation Based on Region Merging. *IEEE Trans. on*

- Pattern Analysis and Machine Intelligence, 1998, 20(9): 897~915
- 10 韩军, 熊璋, 孙文彦, 等. 自动分割及跟踪视频运动对象的一种实现方法. 中国图像图形学报, 2001, 6(8)
- 11 陈晓棠, 毕凌燕, 余英林, 等. 结合形态学和假设检验的视频对象分割. 计算机工程与应用, 2001(21): 102~105
- 12 Kuo C M, Hsieh C H. A New Mesh-based Temporal-Spatial Segmentation for Image Sequence. Computer Software and Applications, 2000. 395~400
- 13 Caplier A, Bonnaud L, Chassery J M. Robust Fast Extraction of Video Objects Combining Frame Differences and Adaptive Reference Image. Image Processing, 2001(2): 785~788
- 14 Isil C, Tekalp A M, Gokcetekin M H, et al. 2-D Mesh-based Video Object Segmentation and Tracking with Occlusion Resolution. Signal Processing, Image Communication, 2001(16): 949~962
- 15 Sebastien V, Magnin L, Prost R, et al. Active Mesh for Video Segmentation and Objects Tracking. IEEE Trans. on Image Processing, 2001(2): 77~80
- 16 Besson S J, Barlaud M. Video Object Segmentation Using Eulerian Region-based Active Contours. IEEE, 2001. 353~360
- 17 Chuang G, Ming-chieh L. Semantic Video Object Tracking Using Region-based Classification. IEEE ICIP' 98, Chicago, 1998. 643~647
- 18 Vigus S A, Bull D R, Canagarajah C N, et al. Video Object Tracking Using Region Split and Merge and a Kalman Filter Tracking Algorithm. Image Processing, 2001(1): 650~653
- 19 Park H W, Schoepflin T, Yogmin K, et al. Active Contour Model with Gradient Directional Information. IEEE Trans. on Circuits and Systems for Video Technology, 2001, 11(2): 252~256

第一作者简介: 沈未名, 教授。现主要从事图像与视频数据压缩和编码、计算机立体视觉等方面的研究开发工作。代表成果: 可视电话视频编解码子系统的设计和实现; 基于小波变换和神经网络的遥感影像数据压缩; 基于神经网络的智能整体影像匹配等。

E-mail: wnshen@public.wh.hb.cn

Video Objects Segmentation and Tracking in Video Sequences

SHEN Weiming¹ JIANG Liu¹ CHONG Yanwen¹

(1 Key Laboratory of Multimedia and Network Communication Engineering, Hubei Province

Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

Abstract: The ISO MPEG4 has attracted much attention recently for providing a standard solution for object-based coding and multimedia data access and manipulation. So content-based representation and coding of the visual information is currently becoming an extremely active research field. Object-based video coding can provide greater compression ratio and better quality of reconstructed images. Video objects segmentation is a key technology in object-based video coding and multimedia data access etc. This paper depicts and analyzes and compares the video objects automatic segmentation and tracking processes and methods. Existing problems and development prospect in this field are described.

Key words: video object; object segment; object tracking; optical flow; snake model; clustering analysis

About the first author: SHEN Weiming, professor, his major research orientations are image and video compression, computer 3D vision, etc.

E-mail: wnshen@public.wh.hb.cn

(责任编辑: 晓平)