

一种面向对象的视频分割算法¹

刘新春 周建宇* 邹谋炎

(中国科学院电子学研究所 北京 100080)

*(新加坡国立大学电子工程系 新加坡 119260)

摘要 该文讨论面向对象编码的视频分割算法,由于数学形态学工具能够很好地处理诸如大小、形状、对比度和连通性等对图像分割非常重要的特征,这种技术常用来对图像进行帧内分割,得到一些具有某种相似性的区域,然后利用运动信息进行区域合并。为了实现区域合并,该文提出一种新方法来判断局部运动,设计了一个分割细化步骤,对区域边界点进行再判决,可以得到更好的结果。实验结果表明,该文提出的方法对于平稳背景和运动背景中的视频对象分割都是有效的。

关键词 MPEG-4, 面向对象的视频分割, 数学形态学, 仿射变换

中图分类号 TN919.8

1 引言

作为第二代编码标准, MPEG4 最重要的一个特征就是视频数据的基于内容的表示^[1]。随着视频对象的引入,对象的自动分割就成为面向对象编码的先决条件了,但由于物理对象的一些常用的特征比如颜色、亮度和光流场等并不均匀,这就使面向对象的视频分割变得困难^[1]。

近来,结合空间信息和运动方法的视频分割算法取得了一些令人兴奋的进展^[2,3],一方面,利用图像的空间信息可以得到颜色和亮度上具有一致性的区域,另一方面,利用图像序列的运动信息能够实现区域合并。一般说来,运动信息更为重要。因为物理对象的运动往往具有一致性,并且与背景的运动不同。许多视频分割方法是利用光流场(Optical Flow Field, OFF)或者变化检测模板(Change Detection Masks, CDM)从背景中分离不同的视频对象,文献[4, 5]中的方法在利用运动信息时有些不同,然而,由于遮挡和孔径问题使得计算出来的运动信息往往不太精确,而且对噪声非常敏感。虽然文献[6, 7]中提出用形态学滤波器的方法来处理运动边缘区域像素的对象归属问题,但由于这类算法只用到了灰度信息,因而敏感于噪声,而且运算量很大。

我们的算法如下:第1步,利用形态学滤波器分割的方法把1帧图像分成几个不同的区域;第2步,用我们提出的运动判据进行区域合并,得到1个二维的代表不同视频对象位置的二值图像;第3步,根据边缘区域像素的运动信息对经过运动合并后的结果进一步细化。最后,我们给出了利用我们的算法对标准的视频测试图像进行分割的结果。

2 帧内分割

在文献[8]中所提出的分层次形态学分割方法具有极好的边缘保持特性,主要包含3个步骤:(1)利用形态学算子对图像进行简化;(2)标记提取;(3)利用分水线方法进行区域增长。

对图像进行简化的目的是为了分割变得更容易些,通过形态学连通算子的作用,图像会变成一些均匀亮度区域,同时能够很好地保证这些区域之间的边界与原始图像中对应位置的一致性。下一步就是给每一个均匀亮度区域分配一个标记,它可以是这个区域中的亮度值,但必须保证它的唯一性。然后让这些分类标记作为分水线算法的输入,进行区域增长,从而得到帧内分割的结果。

进行帧内分割以后,整幅图像被分解成许多具有不同标记的区域,因此,可以把图像看成由一系列具有不同标记的图像区域组成,即: $R = \sum_{i=1}^N R_i, i = 1, 2, \dots, N$ 。图1(c)和图2(c)是帧内分割的结果。帧内分割的结果作为我们下一步——区域合并的输入,就可以得到我们所需要的二值图像。

¹ 2000-01-28 收到, 2000-10-12 定稿

3 区域合并

通常, 运动信息是区域合并所要用到的重要判据, 虽然位移帧差 (Displaced Frame Difference, DFD) 可以作为区域合并的一个判据, 但当背景运动时, 我们发现它并不是一个好的选择. 因此, 基于对光流场和局部区域运动的估计, 我们提出一种新的运动合并判据. 首先, 我们给出区域合并的步骤如下:

第 1 步 计算整幅图像 R 的光流场;

第 2 步 估计 R 中各区域运动的仿射模型参数;

第 3 步 比较参数化运动矢量与光流场矢量, 若一个区域 R_i 中它们之间欧氏距离的平均值超过了一个门限, 就把这个区域看作一个模型失效 (Model Failure, MF) 区域;

第 4 步 标记模型失效区域, 并把那些参数几乎一致的区域合并成一个大的模型失效区域池 P ;

第 5 步 估计 P 中的仿射运动参数.

利用这种从上到下的算法, 整幅图像被剖分成几个独立的运动对象, 每一个对象都遵循一种参数化的运动模型. 我们首先采用文献 [2] 中的方法来计算光流场, 下面我们用 $(u(x, y), v(x, y))$ 表示像素 (x, y) 处的光流矢量.

我们用 6 变量仿射模型来描述图像上物体的运动:

$$\left. \begin{aligned} \hat{u}(x, y) &= a_1x + a_2y + a_3 \\ \hat{v}(x, y) &= a_4x + a_5y + a_6 \end{aligned} \right\} \quad (1)$$

通常, 一帧图像中有好几个物体在运动, 这就要求运动参数估计算法足够稳健. 我们采用文献 [3] 中的方法, 把运动参数估计问题转化为下面函数的最小化问题:

$$E(\Theta) = \sum_{X_i \in R} \rho(\text{DFD}_\Theta(X_i)) \quad (2)$$

其中 R 代表整幅图像, $\rho(\bullet)$ 为运动估计算子, Θ 为估计出的运动参数矢量 $(a_1, a_2, a_3, a_4, a_5, a_6)^T$, DFD_Θ 为位移帧差:

$$\text{DFD}_\Theta(X_i, t) = I(X_i, t) - I(X_i + d_\Theta(X_i), t + 1) \quad (3)$$

其中 I 为图像的亮度, X_i 为像素 i 的坐标, d_Θ 为像素 i 的运动矢量, 它可以由运动参数矢量 Θ 及 X_i 求得, 我们选择 Tukey 的双权函数 [3], 把运动估计问题转化为一个加权的乘方问题, 通过增量估计策略可以求得最佳的运动参数矢量 Θ .

利用 (1) 式我们可以计算每一个像素 i 的运动矢量 $(\hat{u}(x, y), \hat{v}(x, y))$. 我们选择运动矢量与光流矢量之间的欧氏距离作为运动判据:

$$\|d_\Theta - d_0\| = \sqrt{(\hat{u}(x, y) - u(x, y))^2 + (\hat{v}(x, y) - v(x, y))^2} \quad (4)$$

对于独立运动对象上的像素, $\|d_\Theta - d_0\|$ 值比较大; 相反, 对于背景上的像素, $\|d_\Theta - d_0\|$ 值比较小. 我们取一个区域上所有像素的 $\|d_\Theta - d_0\|$ 的平均值作为运动判据, 与一个门限进行比较即可得到所需的二值图像.

经过运动合并以后, 我们得到一个表示几个不同运动对象区域的标记图 Z : $Z(x, y) = l$, 其中 $l = 1, 2, \dots, M$, M 为分割出的对象数目, 对应于这些对象的运动参数矢量有 M 个, $\{\Theta_i, i = 1, 2, \dots, M\}$.

4 分割细化

经过运动合并后得到的分割图 Z 还需要进行分割细化以便得到好的基于对象的视频分割结果. 利用估计出的运动参数矢量集, 通过边缘像素的再归类, 从而实现分割细化. 为了对边缘像素进行再归类, 我们利用位移帧差作为判据. 分割细化的流程如下:

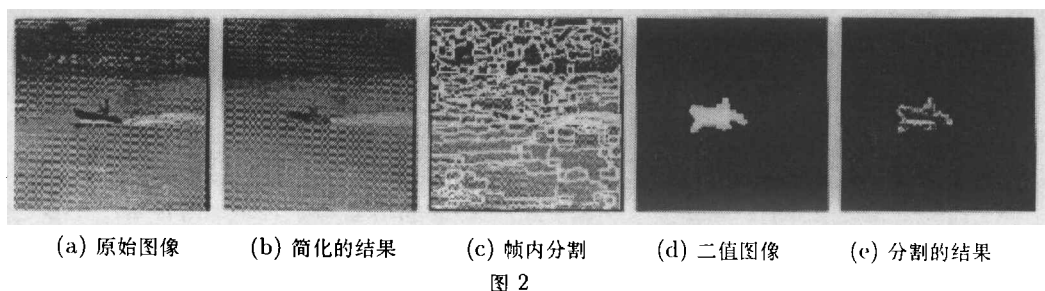
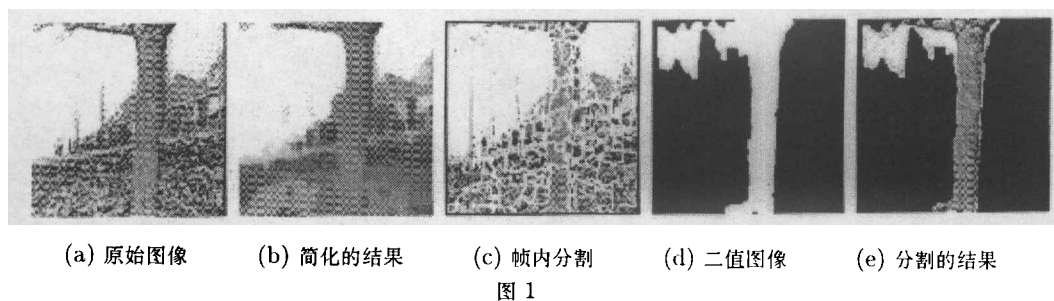
- 第 1 步 判断当前像素 p 是否为 Z 中的边缘像素, 若不是, 转向第 5 步;
 第 2 步 判断 Z 中以 p 为中心的小窗 w_p 内所包含所有对象的集合 C_p ;
 第 3 步 计算小窗 w_p 内的 $DFD_{\Theta_i}, i \in C_p$;
 第 4 步 把 p 归类到对象 k , 其 DFD_{Θ_k} 最小;
 第 5 步 选择扫描方向上的下一点作为当前点, 转向第 1 步。
 在实际应用中, 我们选择以 p 为中心的小窗 w_p 内位移帧差的平均值作为判据:

$$\overline{DFD}_p = \sqrt{\sum_{X_i \in w_p} DFD^2(X_i) / 9} \quad (5)$$

5 实验结果

利用本文的算法对 Flower Garden 和 Coastguard 进行分割的结果分别如图 1 和图 2 所示。Flower Garden 序列中背景不平稳, 并且其中的树有一较大的从右到左的运动, 图 1(a) 是序列的第一帧, 图 1(b) 是运用数学形态学 open-close 算子后的结果, 图 1(c) 是帧内分割的结果, 图 1(d) 是经过运动合并及分割细化后的结果, 图中白色部分代表运动对象, 图 1(e) 是最终的分割结果。可以看得出来, 有一部分天空被归类到树枝上去了, 这是由于加上了平滑性约束从而导致光流场计算不准确造成的。

图 2 是序列 Coastguard 的分割结果, 由于船比背景要小得多, 因而区域局部运动也较好估计, 图 2(e) 所得到的结果也比较理想。



6 结 论

实验的结果证明了我们提出的视频对象分割算法在从背景中分割出运动对象时是成功的, 它首先对图像进行帧内分割得到一些在灰度上具有一致性的区域, 然后根据运动判据进行区域合并。作为进一步的工作, 我们正在研究如何利用运动信息对序列中的运动对象进行跟踪。

致谢 衷心感谢朱继莲博士的大力帮助。

参 考 文 献

- [1] MPEG Video Group, MPEG-4 video verification model version 10.1, ISO/IECJTC1/SC29/WG11/W3464, Tokyo Meeting, March 1998.
- [2] P. Anandan, A computational framework and an algorithm for the measurement of visual motion, *Int. J. Computer Vision*, 1989, 2(2), 283-310.
- [3] J. M. Odobez, P. Bouthemy, Robust multiresolution estimation of parametric motion models, *Journal of Visual Comm. and Image Represent.*, 1995, 6(3), 348-365.
- [4] J. G. Choi, S. W. Lee, S. D. Kim, Spatio-temporal video segmentation using a joint similarity measure, *IEEE Trans on Circuits and Systems For Video Technology*, 1997, CAS VT-7(2), 279-286.
- [5] G. D. Borshukov, *et al.*, Motion segmentation by multistage affine classification, *IEEE Trans. on Image Processing*, 1997, IP-6(10), 1591-1594.
- [6] P. Salembier, A. Oliveras, L. Garrido, Antiextensive connected operators for image and sequence processing, *IEEE Trans. on Image Processing*, 1998, IP-7(4), 555-570.
- [7] T. Meier, K. N. Ngan, Segmentation and tracking of moving objects for content-based video coding, *IEE Proc-F*, 1999, 146(3), 144-150.
- [8] P. Salembier, M. Pardas, Hierarchical morphological segmentation for image sequence coding, *IEEE Trans. on Image Processing*, 1994, IP-3(5), 629-651.

VIDEO OBJECT SEGMENTATION FOR
CONTENT-BASED VIDEO CODING

Liu Xinchun Zhou Jianyu* Zou Mouyan

*(Institute of Electronics, Chinese Academy of Sciences, Beijing 100080, China)***(Dept. of Electrical Engineering, National University of Singapore, Singapore 119260)*

Abstract This paper deals with automatic object segmentation method for object-oriented video encoding. Since methods of mathematical morphology are capable of detecting geometrical features such as size, contrast, or connectivity which are important to image segmentation. Intra-frame segmentation is in first step performed using morphological tools. After obtaining different areas in intra-frame segmentation step, a few video object areas can be merged according to their motion parameters. Finally, a new technique for the detection of motion of a local area is proposed. A two-dimensional binary model can be obtained from the object of interest. By introducing a refinement step, the border pixels can be categorized into one of the neighboring objects so as to achieve a minimum predictive error that computed from its model parameters. Experimental results demonstrate a successful segmentation of objects located in a still or moving background.

Key words MPEG-4, Object-oriented video segmentation, Mathematical morphology, Affine transform

刘新春: 男, 1968 年生, 博士生, 研究方向为视频技术及其应用.

周建宇: 男, 1970 年生, 硕士, 研究方向为 ASIC 设计技术.

邹谋炎: 男, 1941 年生, 研究员, 博士生导师, 研究方向为图像处理及视频技术.