

基于视差和阈值分割的立体视频对象提取

安平¹⁾ 刘苏醒¹⁾ 高欣¹⁾ 张兆杨^{1),2)}

¹⁾(上海大学通信与信息工程学院,上海 200072) ²⁾(上海大学新型显示技术及应用集成教育部重点实验室,上海 200072)

摘要 视频对象分割和提取是编码、通信以及视频检索等基于内容视频处理中的关键问题,为了从只有单一全局运动、含有重叠多对象的立体视频序列中提取对象,提出了一种基于视差分析和阈值分割的对象提取方法。该方法首先用改进的区域匹配法进行立体视差估计,并通过合理减少匹配窗的运算量及根据视差特性设定搜索路径来加快匹配速度;然后针对图像中不同的对象分别采用迭代阈值法和自适应阈值法进行二次分割;最后从阈值分割结果中提取出各个对象。实验提取出的各深度层视频对象效果良好,表明该方法是一种有效的适用于全局运动的立体视频序列对象提取方法。

关键词 立体视频 对象提取 视差估计 阈值分割

中图分类号: TP391.41 TN941.1 **文献标识码**: A **文章编号**: 1006-8961(2006)11-1669-04

Stereoscopic Video Object Extraction Based on Disparity and Threshold

AN Ping¹⁾, LIU Su-xin¹⁾, GAO Xin¹⁾, ZHANG Zhao-yang^{1),2)}

¹⁾(School of Communication and Information Engineering, Shanghai University, Shanghai 200072)

²⁾(Key Laboratory of Advanced Displays and System Application, Ministry of Education, Shanghai 200072)

Abstract Object segmentation and extraction are vital tasks in many issues of content-based video processing. A video object segmentation algorithm is proposed in this paper based on disparity analysis and threshold segmentation for stereoscopic sequences including overlapped multi-objects with global motion. An improved area-based method is firstly adopted for disparity estimation by accelerating the matching processing. Then, to segment different objects in the scene, iterative threshold segmentation and self-adaptive threshold segmentation are respectively performed on the images, and the objects are extracted at last. Experimental results show that the proposed algorithm is an effective object extraction method suitable for stereoscopic sequences with unitary global motion.

Keywords stereoscopic video, object extraction, disparity estimation, threshold segmentation

1 引言

单视点视频对象提取主要利用灰度、边缘以及运动等信息进行分割^[1]。在立体视频中,可以获得单视频中所没有的视差或深度信息,由于深度信息接近于语义对象,十分有利于对象的分割^[2,3],因此对于立体视频来说,有更多的信息可用于分割。由于视差信息的利用,因此在立体视频对象分割中,可

用相对简单的方法解决在单视频分割中难以解决或用复杂方法才能解决的问题。

对于一类像机与场景之间有相对运动,但实际场景静止的立体视频序列,由于所有对象的运动均一致,因此基于运动分析的分割方法将失效。尤其当场景中有多个互遮挡对象时,用时空分割很难得到好的分割效果。为此,本文提出一种基于视差分析和阈值分割的立体视频分割方法,即在视差匹配的基础上,结合数学形态学处理,通过对视差图进行

基金项目:上海市自然科学基金项目(04ZR14056);上海市教委发展基金项目(04AB56);国家自然科学基金项目(60572127)

收稿日期:2006-06-30;改回日期:2006-08-01

第一作者简介:安平(1968~),女。2002年获上海大学通信与信息工程学院博士学位,现为上海大学通信与信息工程学院副教授。主要研究方向为图像及3维视频信号处理。已发表学术论文30多篇。E-mail:anping@staff.shu.edu.cn

阈值分割来得到语义视频对象。实验结果表明,对于只有单一全局运动、且对象之间重叠的立体图像序列,本文方法能够很好地分割出对象。

2 算法框架

图 1 为提出的基于视差和阈值分割算法的原理框架。本文算法主要由视差匹配、数学形态学处理、迭代阈值分割、自适应阈值分割等几部分组成。算法步骤为首先用改进的区域匹配法对立体图像对进行视差估计,同时采用减少匹配窗计算量、优化搜索路径的方法降低算法复杂度,并对遮挡区域进行补偿;然后,对所得到的视差图作数学形态学处理,消除视差图的内部空洞和不连续区域,以优化视差图的质量;接着针对图像中不同的视频对象分别采用全局迭代阈值法和局部自适应阈值法进行二次分割;最后在阈值分割的结果中提取视频对象。

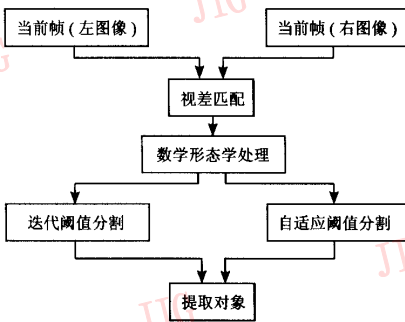


图 1 算法框图

Fig. 1 Algorithm framework

3 算法原理

3.1 视差估计

视差是指立体系统中图像对之间的几何位置差异,其中包括各摄像机的相对运动及场景中对象的运动所产生的差异,视差估计也是立体或 3 维视频处理中的核心问题。

视差估计的方法很多,其中区域法是一类简单有效的方法,它可以直接得到稠密的视差场^[4]。本文提出一种改进的基于区域的估计方法,即通过合理减少匹配窗口的运算量及根据视差特性设定适当的搜索路径来提高匹配速度。匹配误差采用累计绝对误差和(sum of absolute difference, SAD)准则, SAD 准则具有不需乘法运算、实现简单方便的优

点,其定义如下:

$$SAD(i, j) = \sum_{m=0}^M \sum_{n=0}^M |f_L(m, n) - f_R(m+i, n+j)| \quad (1)$$

其中, $f_L(m, n)$ 、 $f_R(m+i, n+j)$ 分别为左图像和右图像的像素值, (i, j) 为位移矢量。具体的改进措施如下:

(1) 匹配时,首先计算当前候选点匹配窗口中像素点的 SAD,如果该值已经大于预设的门限,则不必再对匹配窗口中余下的像素点进行计算,并丢弃此候选点。这样,只有正确的匹配点才需要对匹配窗口中的多数像素点进行计算,而对于大多数非匹配点而言,则只需计算匹配窗口中的少量像素点,从而大大减少了运算量。

(2) 在平行双目摄像系统中,由于右图像可近似看作是左图像中各像素点沿水平轴负向平移的结果,因此进行由左图像至右图像的立体匹配时,理论上视差在水平方向上的分量是负的。通过设定适当的搜索路径可进一步提高搜索速度:

① 先水平负向、再水平正向进行搜索;

② 考虑到两台摄像机并不能精确地处于同一水平面,还要对垂直方向进行搜索,因此先搜索当前扫描行,再搜索当前扫描行的上一行和下一行,依次类推,搜索范围一般设为 ± 2 Pixels 高度。

视差匹配后,还需对失配点进行遮挡检测和补偿。本文采用双向匹配的方法寻找遮挡点^[5],并用遮挡区域四邻域内的非遮挡点的视差值来补偿遮挡区域的视差,实验结果表明,效果良好。

3.2 视差图的形态学处理

为便于后续分割处理,需对视差图作后处理。本文采用数学形态学腐蚀、膨胀、开启和闭合等基本运算进行处理,以连接原本相关的区域或边界段和分离原本不接触的区域或填充对象内部存在的空洞。形态学处理之后,视差图呈现出层次明显的不同景深,以便于进一步的深度分割处理。

3.3 视差图的阈值分割

由于视差和景深成反比关系,因此对视差图进行阈值分割可以分割出不同深度层上的对象。对于在图像中面积比例较大的对象,宜采用全局迭代阈值分割法,而对于面积比例较小的对象,则采用自适应局部阈值分割方法较为合适。

3.3.1 迭代阈值分割

取图像灰度范围的中值作为初始阈值 T_0 , 设有 L 个灰度等级,求分割阈值的算法步骤如下:

(1) 求出图像中的最小和最大灰度值 z_{\min} 和

z_{\max} , 令阈值初值为

$$T_0 = \frac{z_{\min} + z_{\max}}{2} \quad (2)$$

(2) 根据阈值 T_0 将图像分割成目标和背景两部分, 并计算两部分的平均灰度值 \bar{z}_0 和 \bar{z}_B , 其中, 下角 O 代表 object, B 代表 background。

$$\bar{z}_0 = \frac{\sum_{z(i,j) < T_K} z(i,j) \times W(i,j)}{\sum_{z(i,j) < T_K} N(i,j)} \quad (3)$$

$$\bar{z}_B = \frac{\sum_{z(i,j) > T_K} z(i,j) \times W(i,j)}{\sum_{z(i,j) > T_K} W(i,j)} \quad (4)$$

式中, $z(i, j)$ 为像点 (i, j) 的灰度值, $W(i, j)$ 为 (i, j) 点的权重系数, 一般取 $W(i, j) = 1.0$ 。

(3) 计算新的阈值

$$T_{K+1} = \frac{\bar{z}_0 + \bar{z}_B}{2} \quad (5)$$

(4) 如果 $T_K = T_{K+1}$, 则结束, 否则令 $K = K + 1$, 转步骤(2)。

迭代结束时得到的 T_K 为最佳分割阈值。

3.3.2 自适应阈值分割

对于小目标对象, 由于目标在图像中所占比例较小, 因此用上述的全局最佳阈值不能将其和背景分开。为了获得正确的分割阈值, 必须提高目标灰度信息在待分割直方图中所占的比例。本文根据分割出的目标和背景灰度统计量为判断准则, 来对图像直方图进行多次分割, 以获得局部最佳阈值。其算法思想如下:

假设在图像中目标灰度比背景灰度高, 则可对原始图像用最大类间方差法^[6]进行分割, 得到目标区域和背景区域, 其原始图像灰度直方图中高于阈值的部分可看作目标区域的灰度统计。对于只有目标和背景组成的图像, 其灰度直方图可视作组合背景和目標像素灰度混合分布的概率密度函数, 且通常假定混合分布的两个分量 $p(i/0)$ 和 $p(i/1)$ 都是正态分布的, 其均值、标准差和先验概率分别为 $\mu_0, \mu_1, \sigma_0, \sigma_1, p_0$ 和 p_1 。其中, μ_0, μ_1 和 σ_0, σ_1 , 分别由式(6)和式(7)给出:

$$\begin{cases} \mu_0(t) = \sum_{z=0}^i zp_z/p_0(t) \\ \mu_1(t) = \sum_{z=i+1}^{L-1} zp_z/p_1(t) \end{cases} \quad (6)$$

$$\begin{cases} \sigma_0(t) = \left\{ \frac{\sum_{z=0}^i [z - \mu_0(t)]^2 p_z}{\sum_{z=0}^i p_z} \right\}^{1/2} \\ \sigma_1(t) = \left\{ \frac{\sum_{z=i+1}^{L-1} [z - \mu_1(t)]^2 p_z}{\sum_{z=i+1}^{L-1} p_z} \right\}^{1/2} \end{cases} \quad (7)$$

设灰度为 z 的像素数为 n_z, t 为灰度阈值, $z = 1, 2, \dots, t, \dots, L$ 。则图像的总像素数为 $N = \sum_{z=0}^{L-1} n_z$, 灰度为 z 的像素出现的概率为 $p_z = n_z/N$ 。不失一般性, 可认为目标和背景的灰度分布分得足够开时, 应满足

$$\mu_1 - \mu_0 > \alpha(\sigma_0 + \sigma_1) \quad (8)$$

式中, 参数 α 需根据目标和背景的灰度分布特性确定, 一般可在 2~3 之间取值。

若由某一阈值分割出的两部分, 其灰度分布的均值和标准差满足上述条件, 则认为该阈值可将目标和背景完全分开。相反, 若不能使式(8)成立, 则需对分割出的目标区域进行更精细的分割。随着待分割区域的缩小, 目标对象所占的比例增大, 最终可获得正确的分割结果。

为提高实时性, 可采用缩小阈值计算时的灰度取值范围的方法来减少计算量。由于目标在整个视场中所占的比例较小, 且目标的灰度值必然大于由大量暗背景和少量亮目标组成的图像的灰度均值。因此, 可将图像灰度均值 T 定为初始阈值, 这一方面可减少计算量, 以节省处理时间; 另一方面由于初始阈值去掉了大量的暗背景, 从而提高了目标对象在待分割区域中所占的比例。

3.4 对象的提取

根据上述阈值分割后得到的各深度层的分割掩模, 再用分割掩模对待分割图像进行与运算, 即可提取各深度层的视频对象。

4 实验结果及讨论

为验证本文算法的有效性, 本文对 Aqua 立体序列用本文算法进行了分割实验。该序列只有相机和场景间的全局运动, 场景中包括假山、大小鱼以及水草等处于不同景深的多个对象。各帧的分割效果相似, 限于篇幅, 这里仅列出第 1 帧的实验结果(如图 2 所示)。图 2(a)、图 2(b)为原始图对, 图 2(c)

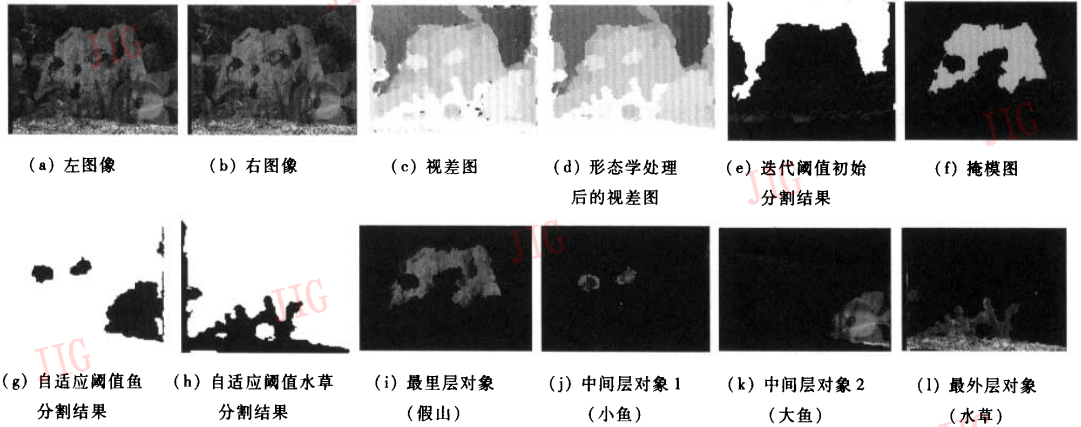


图 2 实验结果

Fig. 2 Experimental results

为求得的视差图,图 2(d)为经过形态学处理后的视差图。由图 2(d)可见,经形态学处理后,原视差图中的对象内部空洞和不连续区域得到较好的处理,有利于进一步的分割。

对于场景中的大对象——假山,可采用迭代阈值进行分割,其分割结果如图 2(e)所示。为了在提取对象时,能区分假山和与之重叠的其他对象(水草和小鱼等),可在图 2(e)的范围内,寻找出视差值大于阈值 T 的区域(即假山的范围),图 2(f)即为假山掩模图。

对于鱼和水草等小对象,可采用自适应阈值进行分割。其中图 2(g)为鱼的分割结果,图 2(h)为水草的分割结果。

最后提取的对象如图 2(i)~图 2(l)所示,其中图 2(i)为位于最里层的假山,图 2(j)为在山前面、属于中间层的小鱼,图 2(k)为在小鱼前面的大鱼部分,图 2(l)为位于图像最外层的水草部分。

由于 Aqua 序列只有全局运动信息,因此基于运动分析的常规分割方法无法分割出多个深度层上的对象,而本文采用视差阈值分割则可以有效提取出位于不同层次上的多对象。

5 结 论

本文提出一种基于视差和阈值分割的立体视频对象提取方法。该方法首先对原始的视差图,采用形态学处理优化其质量;然后针对图像中不同的对

象分别采用迭代阈值法和自适应阈值法进行二次分割;最后由阈值分割结果提取出各个对象。实验结果表明,本文方法能有效解决只有同步全局运动的立体视频序列的多对象分割问题,其在基于内容的立体或多视点视频处理中有良好的应用前景。

参考文献 (References)

- 1 Long Fu-hui, Feng Da-gan, Peng Han-chuan. *et al.* Extracting semantic video objects [J]. IEEE Computer Graphics and Applications, 2001, 21(1): 48 ~ 55.
- 2 Ntalianis K S, Doulamis N D, Doulamis A D, *et al.* A feature point based scheme for unsupervised video object segmentation in stereoscopic video sequences [A]. In: Proceedings of IEEE International Conference on Multimedia and Expo [C]. New York, USA, 2000, 3: 1543 ~ 1546.
- 3 Kim S H, Choi J H, Kim H B, *et al.* A new snake algorithm for object segmentation in stereo images [A]. In: Proceedings of 2004 IEEE International Conference on Multimedia and Expo [C], Sorrento, Italy, 2004, 1: 13 ~ 16.
- 4 An Ping, Zhang Zhao-yang, Ma Ran. Hierarchical MRF/GRF model based disparity estimation and segmentation for stereo images [J]. ACTA Electronic Sinica, 2003, 31(4): 597 ~ 601. [安平, 张兆杨, 马然. 应用分层 MRF/GRF 模型的立体图像视差估计及分割 [J]. 电子学报, 2003, 31(4): 597 ~ 601.]
- 5 Rainer J, Izquierdo E. An object-based system for stereoscopic viewpoint synthesis [J]. IEEE Transactions on Circuits and System for Video Technology, 1997, 7(5): 801 ~ 811.
- 6 Otsu N. A threshold selection method from gray-level histogram [J]. IEEE Transactions on Systems Man, and Cybernetics, 1979, 9(1): 62 ~ 66.