

文章编号:1007-2780(2010)04-0598-03

一种多视点视频中的对象提取方法

张 倩¹,吴妍菲¹,安 平^{1,2*},张兆杨^{1,2}

(1. 上海大学 通信与信息工程学院,上海 200072,E-mail: anping@shu.edu.cn;
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: 提出一种自由视点电视的对象提取方法,首先用 Birchfield 算法进行快速的视差估计方法得到一个较为平滑的视差图,然后对初始的分割结果用形态学滤波,实验结果证明了所提出方法的有效性和可靠性。

关 键 词: 多视点视频;分割;视差

中图分类号: TP391. 41 **文献标识码:** A

Multi-View Point Video Object Extraction Method

ZHANG Qian¹, WU Yan-fei¹, AN Ping^{1,2*}, ZHANG Zhao-yang^{1,2}

(1. School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China, E-mail: anping@shu.edu.cn;
2. Key Laboratory of Advanced Displays and System Application, Shanghai University, Shanghai 200072, China)

Abstract: A new object segmentation method was presented in this paper. Firstly, a disparity estimation method was presented, which get a fast estimation result. Then, the segmentation is obtained by Morphological filtering. The experimental results verify the validity and reliability of this method.

Key words: muli-view video; segmentation; disparity

1 引 言

视频处理的很多领域都用到了对象分割,如压缩、对象识别和视点合成等。传统的视频对象分割又称为单视点分割,其分割的视频由一个摄像机拍摄生成。随着多视点捕获技术的发展,多视点视频分割逐步发展成为一个重要的研究领域。

在多视点视频分割中视差是一个重要的依据,基于视差的分割又被称为深度分割。然而由于从立体视频中估计深度有很多种不同的模式,因此有必要对于视差分割采用一种不同的处理方式。大多数提出的方法同时利用视差和运动信息来实现分割^[1-4],只用视差而不用运动信息的分割很少见诸报道^[5]。

本文提出了一种新颖的多视点视频对象分割方法,利用实时的深度获取算法快速获取深度信息,进而可以得到初始的分割信息,然后再利用形态学滤波将初始分割结果进行优化。这种方法完全不利用运动信息,而是根据视差信息来完成分割。Birchfield 的视差估计方法能快速有效地获取视差信息,尤其对背景等无纹理区域能够取得较好的平滑视差图。分割的实验结果证实了所提出算法的有效性。

2 Birchfield 算法的基本原理

为了兼顾对象分割的时效性和精确性,采用 Birchfield 算法^[6-8]生成视差图。Birchfield 算法^[6,7]的优点是可以比较快速地得到相对可靠的

深度不连续点和无匹配点,缺点是使用该算法得到的深度图的精确性不够。该算法用动态规划对极线进行匹配,在获取深度不连续点和无匹配点的同时得到深度图,然后在相邻极线间对匹配结果进行校正。应注意的是,该算法在匹配过程中没有使用基于窗口的方法,也不对像素灰度值进行预处理,而是直接对单个像素点进行匹配,这也是该算法不能得到较精确匹配的原因之一。

Birchfield 算法有如下 3 个优点:

(1)采用采样不敏感的像素相似性判定准则,避免了图像采样中产生的问题;

(2)解决了大块无纹理区域的匹配问题,这个问题对于现存的许多立体匹配算法都是很大的挑战;

(3)对非相似点进行减枝,从而减少了算法的计算时间。

Birchfield 算法使用一个全局的代价函数表征匹配序列的可信度,为了节省运算时间使用一个经验公式作为全局代价函数:

$$r(M) = N_{\text{occ}}k_{\text{occ}} - N_m k_r + \sum_{i=1}^{N_m} d(x_i, y_i) \quad (1)$$

其中, k_{occ} 表示遮挡代价值, k_r 是匹配增益值, $d(x_i, y_i)$ 表示两点间的非相似度。 N_{occ} 、 N_m 则分别表示遮挡区和匹配区的数目。

其中非相似度可以定义如下:

$$d(x_i, y_i) = \min\{\bar{d}(x_i, y_i, I_L, I_R), \bar{d}(y_i, x_i, I_R, I_L)\} \quad (2)$$

I_L, I_R 分别表示左右视图像中一对对应极线的灰度函数,而 $\bar{d}()$ 定义如下:

$$\bar{d}(x_i, y_i, I_L, I_R) = \min_{y_i - 0.5 \leq y \leq y_i + 0.5} |I_L(x_i) - \hat{I}_R(y)| \quad (3)$$

其中, $\hat{I}_R$ 为右图中某条极线上各采样点进行线性插值的灰度函数。

进而可以用动态规划算法和之前定义的代价函数,使用搜索法沿着极线得到一个最佳的匹配序列。

3 形态学滤波

首先,用阈值来获得初始的分割掩模:

$$S(i, j) = \begin{cases} 0 & M(i, j) \geq \delta \\ 1 & \text{elsewhere} \end{cases} \quad (4)$$

其中, δ 是阈值,在实验中被设定为 12.5。 $M(i, j) \geq \delta$ 意味着该像素属于前景物体,反之则认为

是背景区域。

接着采用形态学滤波对掩模进行滤波。

数学形态学运算^[9]分为二值形态学和灰度形态学两类,这里采用灰度数学形态学。

数学形态学的运算主要有膨胀、腐蚀两个基本运算,以及由这些基本运算组合而成的其他运算。

设 A 为给定的灰度图像, B 为结构元素, A 被 B 的膨胀记为 $A \oplus B$, $\oplus$ 为膨胀算子,膨胀的定义为:

$$A \oplus B(x) = \max\{A(x - z) + B(z)\} \quad (5)$$

腐蚀定义为:

$$A \ominus B(x) = \min\{A(x + z) - B(z)\} \quad (6)$$

由式(5)和式(6)可知,膨胀扩充图像而腐蚀收缩图像。

数学形态学的闭运算是先对图像进行膨胀运算再腐蚀其结果,定义为:

$$A \cdot B = (A \oplus B(z)) \ominus A \quad (7)$$

4 实验结果与分析

在内存为 2 GB、主频为 2.66 GHz 的机器上,用本文提出的算法分割立体视频序列 Akko 前景中的对象。图 1(a)、图 1(b)分别为 Akko 序列第 48 和第 49 帧的视点,图 2 为 Birchfield 算法提取的视差图,图 3(a)和图 3(b)分别为初始的掩模和经过形态学滤波后的掩模,图 4 是最终的分割结果。

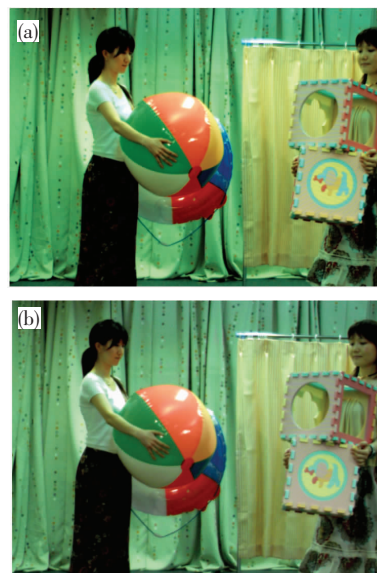


图 1 Akko 序列第 48(a)和第 49 帧(b)的视点
Fig. 1 No. 48 (a) and 49 (b) frames of Akko sequence



图 2 Birchfield 算法提取的视差图

Fig. 2 Disparity map obtained by Birchfield method

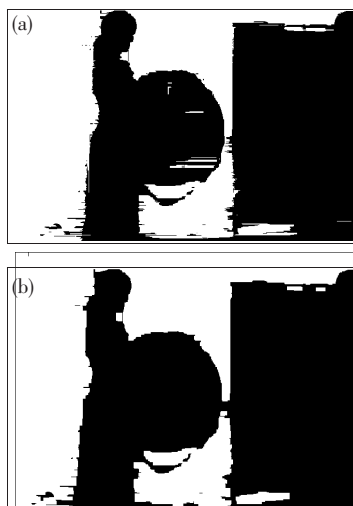


图 3 初始的掩模(a)和经过形态学滤波后的掩模(b)

Fig. 3 Initial mask (a) and the mask after mathematical morphology(b)



图 4 最终的分割结果

Fig. 4 Segmented result

由实验结果可知,本文方法可以将纹理复杂场景中的语义对象较好地提取出来,且分割所得前景对象的边缘也较准确平滑。由于尚无立体视频对象分割的客观评价准则,这里只给出主观评价。

5 结 论

提出了一种面向多视点视频的立体对象视频分割方法,可以快速提取视差场进行修正和分割,并用形态学算法进行优化,最终分割出对象。实验结果表明,所提出的算法兼顾了对象分割的时效性和精确性,对于纹理复杂的场景有很好的分割效果,可以获得与语义一致的对象。

参 考 文 献:

- [1] Francois E, Chupeau B. Depth-based segmentation [J]. *IEEE Trans. on CSVT*, 1997, 7:237-239.
- [2] Izquierdo E, Kruse S. Image analysis for 3D modeling, rendering, and virtual view generation [J]. *Computer Vision Image Understanding*, 1998, 71(2): 231-253.
- [3] Chang Yu-lin, Fang Chih-ying, Ding Li-fu, et al. Depth map generation for 2D-to-3D conversion by short-term motion assisted color segmentation[C]//*IEEE International Conference on Multimedia and Expo*, Beijing, China; IEEE, 2007:1958-1961.
- [4] Cigla C, Aydin Alatan A. Depth assisted object segmentation in multi-view video[C]//*3D TV Conference: The True Vision - Capture, Transmission and Display of 3D Video*, Istanbul, Turkey; IEEE, 2008:185-188.
- [5] Thakoor N, Jean G, Devarajan V. Multihypothesis prior for segmentation of stereo disparity [J]. *IEEE Signal Processing Lett.*, 2008, 15:613-616.
- [6] Birchfield S, Tomasi C. Depth discontinuities by pixel-to-pixel stereo[C]//*Sixth International Conference on Computer Vision*, Bombay, India; IEEE, 1998:1073-1080.
- [7] 陈胜勇,刘盛. 基于 opencv 的计算机视觉技术实现[M]. 北京:科学技术出版社, 2008.
- [8] Xu Xiu-bing, Xie Xu-dong, Dai Qiong-hai. Real-time 3D video synthesis from binocular stereo camera[C]//*3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video*, Istanbul, Turkey; IEEE, 2008:133-136.
- [9] Salembier P, Torres L, Meyer F, et al. Region based video coding using mathematical morphology [J]. *Proc. of IEEE*, 1995, 83(6):843-857.