

实时抠像技术在舞台环境中的实现

陈临强¹, 杨盼¹, 吴珏辉²

CHEN Lin-qiang¹, YANG Pan¹, WU Jue-hui²

1.杭州电子科技大学 图形图像研究所,杭州 310018

2.中国美术学院 新媒体系,杭州 310053

1.Institute of Graphics and Images Process, Hangzhou Dianzi University, Hangzhou 310018, China

2.Department of New Media, China Academy of Art, Hangzhou 310053, China

E-mail: yangpan116@sina.com

CHEN Lin-qiang, YANG Pan, WU Jue-hui. Implement real-time matting technology in stage environment. *Computer Engineering and Applications*, 2008, 44(16): 177-178.

Abstract: The real-time matting technology is widely used in the field Virtual Studios, Virtual Advertisement and Monitoring System. This paper aims to use the real-time matting technology in the stage environment with the method of OTSU algorithm, to implement the real-time interaction with actors and incarnate the interaction of New Media Electronic Arts. The result of the experiment shows that the method mentioned in the paper has the features of high speed and steady, which can fully fulfill the need of real-time interaction.

Key words: New Media Arts; real-time matting technology; OTSU algorithm

摘要:目前实时抠像技术广泛的应用于虚拟演播室、虚拟广告、监控系统等领域。通过 OTSU 最大类间方差法,把实时抠像技术应用于舞台环境中,实现了与演员实时互动,体现了新媒体电子艺术的互动性,进一步拓展了抠像技术的应用范围。实验结果表明,此方法对于类似舞台环境下的抠像,具有速度快、稳定性高等特性,可完全满足实时互动的需要。

关键词:新媒体艺术;实时抠像;OTSU 算法

DOI: 10.3778/j.issn.1002-8331.2008.16.054 文章编号: 1002-8331(2008)16-0177-02 文献标识码: A 中图分类号: TP391

1 前言

随着新媒体电子艺术的发展,舞台表演方式也越来越多样化,而其互动要求也越来越高。奥地利 Ars Electronica 电子艺术中心、荷兰 V2 多变媒体协会、美国 MIT 媒体实验室、德国 ZKM 艺术与媒体中心在这些方面做了比较深入的研究,而国内在这方面则相对较少。其中 Ars Electronica Center(AEC)有着“未来的技术博物馆”^[1]的美誉,拥有着如声音交互系统等多种多媒体交互技术^[2]。同样作为互动技术的实时抠像技术,目前广泛应用于虚拟演播室、虚拟广告、监控系统等领域。如以色列 Orad 公司的虚拟演播室和虚拟广告系统产品,其中的实时抠像技术,已成熟应用于电视节目播出当中。文献[3,4]则把抠像技术应用在视频监控方面。进一步拓展了实时抠像技术的应用范围,通过带门限的 OTSU 算法和舞台灯光场景布置,成功地把舞台演员从舞台环境中实时分割出来。

2 抠像算法的选择

在虚拟演播室内,传统的抠像技术有两种:亮度抠像和色度抠像^[5]。亮度抠像是把图像进行黑白灰度化后取其亮度信号作为键信号进行抠像。色度抠像主要使用颜色来进行抠像,一

般使用蓝色,也有使用红色、绿色的。色度抠像可以把图像中的一种颜色替换成另一路图像。

由于舞台环境的复杂性,无法跟虚拟演播室等室内固定环境相等同,尤其是舞台背景无法和演播室内的蓝背景相比拟,所以在算法的选择上,很难使用基于 HLS 等颜色通道进行色度抠像的算法,也不能使用背景差分法^[6]来实现。背景差分认为背景图像是不变的,在摄像机静止的情形下,首先为背景图像建立背景模型,然后通过当前帧与该模型相比较,获取目标。但是舞台的灯光变化,以及演员的动作的随机性都使得背景模型的稳定性大大降低。文献[7]提出了一种基于直方图统计的动态求取颜色背景模型,进行色度抠像的方法。这种方法虽然能较好地获取背景模型,但是色度抠像的阈值无法动态调整,在舞台灯光变化较明显时,效果不理想。本文采用了一种亮度抠像方法。OTSU 算法^[8]是在灰度直方图基础上用最小二乘法原理推导出来的,具有统计意义上的最佳分割阈值。OTSU 算法简单、速度快,对光线等具有较好的自适应能力,具有一定的鲁棒性,在灰度图像目标和背景的自动分割上具有广泛的应用。如文献[9]把 OTSU 法在多阈值分割上做了推广,但是其计算量过大,并且缺乏明确的物理意义。本文通过依靠舞台灯光和演员

作者简介:陈临强(1963-),男,副教授,主要研究领域为图像处理,计算机图形学;杨盼(1985-),男,硕士,主要研究领域为图像处理,计算机视觉;吴珏辉(1980-),男,硕士,主要研究领域为新媒体电子艺术。

收稿日期:2007-11-08 修回日期:2008-02-01

服装的配合来拉大灰度差,使得 OTSU 的动态分割阈值更加稳定。

3 OTSU 算法描述及流程图

OTSU 算法又叫最大类间方差法,它是一种自适应的阈值确定方法^[10],由日本学者大津于 1979 年提出。记图像 $I(x,y)$ 为灰度图,其根据分割阈值 T 分割后的二值化图像为 $g(x,y)$ 。则用转变公式表示:

$$g(x,y) = \begin{cases} 0 & I(x,y) \leq T \\ 1 & I(x,y) > T \end{cases}$$

将图像 $I(x,y)$ 表示为 L 级灰度,例如:1,2,...,L。在 i 级上的像素个数为 n_i ,总图像像素个数为 N 。图像 $I(x,y)$ 的直方图可以看作是图像的亮度分布。假设将图像根据可选的阈值 k 分为两类。 k 的值在 1 与 L 之间,那么灰度级别 k 刚好在这个分布的谷底,那么可以用如下公式分别表示两个类别:

$$\mu_1(k) = \frac{\sum_{i=1}^k i n_i}{\omega_1} \quad \mu_2(k) = \frac{\sum_{i=k+1}^L i n_i}{\omega_2}$$

$$\omega_1(k) = \sum_{i=1}^k n_i \quad \omega_2(k) = \sum_{i=k+1}^L n_i$$

则图像的全局值和方差可以表示为:

$$\mu_T = \mu(L) = \frac{\sum_{i=1}^L i n_i}{N} = \frac{\omega_1(k) \mu_1 + \omega_2(k) \mu_2}{\omega_1(k) + \omega_2(k)}$$

$$\sigma_T^2 = \frac{\sum_{i=1}^L (i - \mu_T)^2 n_i}{N}$$

两个类之间的方差可以表示为:

$$\sigma_B^2(k) = \omega_1(k) (\mu_1(k) - \mu_T)^2 + \omega_2(k) (\mu_2(k) - \mu_T)^2$$

由 k 值得到的类之间的距离用 η 表示。当 η 值为最大时,可选的阈值 k 即为所求。同时, η 也表示了该图像的信息测量度。

$$\eta(k) = \frac{\sigma_B^2}{\sigma_T^2}$$

由于舞台环境的特殊性,要考虑到演员进出舞台的过程变换情况。当舞台上没有演员时,使用 OTSU 算法分割会分割出很多噪点。本文设置了一个最低阈值,在演员没有进入舞台时预先统计训练得到。当 OTSU 的阈值 k 小于该门限阈值时,表示舞台上没有演员。

算法设计的具体流程见图 1。

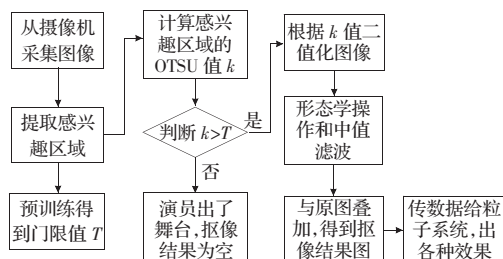


图 1 OTSU 抠像算法流程

4 实验实现与结果分析

实验舞台环境布置为地面铺黑幕,电脑灯打出一个演员跳

舞的大致范围圈,较柔和。其中电脑灯从上往下和从下往上各 4 个灯,这样打出的灯光较为柔和,可以有效地去除演员身上遮挡产生的阴影。硬件装配如下:Sony Z1C 小高清摄像机,Viewcast 的 Osprey210 视频采集卡,HP 的 Nx4400 工作站,CPU E6320,内存 2 G,显卡 8600GTS,系统为 Windows XP SP2。程序的图像处理方面借助 Intel 的 OpenCV (Open Source Computer Vision Library) 库。OpenCV 是 Intel 资助的开源计算机视觉库。它由一系列 C 函数和少量 C++ 类构成,实现了图像处理和计算机视觉方面的很多通用算法。由于 OpenCV 在 Windows 下视频采集采用 VFW 的方式,每秒只有 15 帧左右,无法达到每秒 25 帧的实时要求,所以本文在视频采集方面使用了 Microsoft 的 DirectShow 技术。DirectShow 视频采集采用 WDM 方式,每秒能达到 60 帧以上。同时采用 Ogre3D (Open-source Graphics Rendering Engine) 的粒子系统来实现互动效果。具体实现的开发环境如表 1。

表 1 开发环境

编译工具	视频采集	抠像实现	效果显示
VC2005	DirectShow	OpenCV	Ogre3D

本文的舞台演员实时抠像和位置的跟踪技术已经成功应用于 2007 年首届上海电子艺术节的开幕式音乐会上。图 2 为原图与抠像结果图的对比,可以清楚地看到本文的实时抠像效果还是非常理想、速度快,对光照亮度具有一定的鲁棒性。本文实验训练得到门限 T 的值为 15,只要电脑灯打的亮度在 70~200 之间,演员在舞台光圈范围内表演,就能得到比较完美的抠像结果。本文的算法也有一个局限性,当演员在舞台光圈边界上时,光照不均匀,会导致抠像结果不完整。

图 3 为更多的互动效果图。

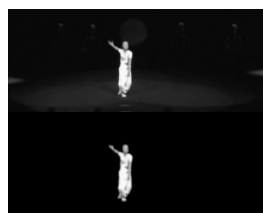


图 2 原图与抠像结果图

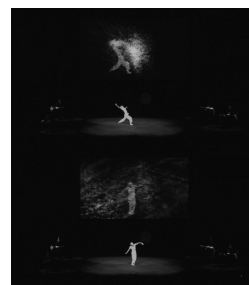


图 3 抠像及互动效果图

5 结论

本文成功地把实时抠像技术应用于舞台表演艺术上面,很好地实现了演员与画面的互动效果,体现了电子技术与艺术的完美结合。今后要对目前的算法做进一步改进,以提高算法的稳定性并扩大适用范围,得到三维的抠像数据,实现立体的互动效果。同时,进一步改善算法和设备,希望能够通过红外设备等硬件的改善,从而摆脱灯光范围限制的局限性。

参考文献:

[1] Janko S, Leopoldseder H, Stocker G. Ars electronica center: museum of the future [EB/OL]. [1996]. Ars Electronica Center, Linz, Austria, <http://www.aec.at/de/index.asp>.
 [2] Borchers J, Muhlhauser M. Musical design patterns: an example of a human-centered model of interactive multimedia [C] // Proceedings of