

文献编号:1000-582X(2003)03-0042-04

虚拟环境下的实时自助照相技术

邵桂芳,李祖枢,张 华

(重庆大学 智能自动化研究所,重庆 400044)

摘 要:虚拟现实技术作为面向21世纪的新技术,是计算机图形学、人机接口技术、传感器技术以及人工智能技术等交叉与综合的结果。针对虚拟现实技术在照相领域中的应用空白,提出了虚拟环境下的实时自助照相技术,讨论了该技术所涉及到的虚拟现实、视频截取等关键性问题,详细介绍了在系统开发中提出的一种新算法:基于对比度感知特征的图像融合技术。

关键词:虚拟环境;自助照相;视频截取;图像融合

中图分类号:TP391

文献标识码:A

传统的照相机,包括数码照相机都是以实际景物为背景,由他人为照相者拍摄,或由照相者启动自拍机构跑到相应位置,延时数秒后拍摄。照相者仅能以固定的背景和有限的角度出发进行拍摄,缺乏实时性、交互性和多样性。为此提出了基于蓝背景抠像技术的虚拟环境下的实时自助照相技术,它将自助式照相、可选虚拟背景下照相和多次拍摄的影像互补合成为动感照片的功能集成在一起,最大限度地发挥照相者的自主性,展示照相者的自我风采。该系统集娱乐性、趣味性、科学性于一体,充分体现了技术创新与效益的结合。

1 虚拟环境下实时自助照相技术的分类

目前虚拟环境下实时自助照相技术在人们生活中的应用,可分为3个阶段:

1)人工合成阶段。这个阶段是人类对虚拟环境下实时自助照相的一种憧憬,期望自己能够置身于某种环境或景物之中。因此采用了各种图像处理软件,将自己的照片通过专业的图像处理技术与期望的背景合成。其中最为优秀的当属美国 Adobe 公司的 Photoshop 软件,以及 Coreldraw 等软件。但其缺点是不具有实时性,且只有具有专业技术人员才能很好地进行上述操作。

2)非实时拍照阶段。虽然各种各样的虚拟环境

在此阶段已基本构成,但由于某些条件如图像的实时提取技术、虚拟场景与人像的实时合成显示技术等不成熟,人们仍然需借助一些专业的图像提取软件,才能得到满意的虚拟环境中的照片。在这个阶段的成熟系统有很多,如国内外服装业设计中的试装软件、婚纱软件、国内的《远潮婚纱摄影系统》等。此类产品一般属于虚拟系统在某一行业中的应用,用户仍不能在构建的虚拟环境中获取实时的效果。

3)实时拍照初级阶段。在该阶段内,照相者可以实时的在构建的虚拟环境中漫游,挑选自己喜欢的场景并可实时地看到自己在场景中的实际效果。之所以称它为初级阶段,是因为在当前的情况下,由于各种技术(图像处理技术、三维造型技术、硬件加速技术)尚不能达到系统的要求。这一初期阶段的代表产品有日本的“新拍立得”和韩国的“立拍得”,用户可以在设定的场景中看到实时的动态效果。

总而言之,如果将基于几何模型的计算机图形技术与基于实景图像的技术充分地结合,就有可能产生集交互性和实时显示质量于一身的虚拟环境。再结合一些先进技术,就可以构造极其完美的虚拟环境下的实时自助照相系统。

目前,此项技术在重庆大学智能自动化研究所与重

• 收稿日期:2002-10-12

基金项目:国家科技型中小企业技术创新基金资助项目(01C26225110588)

作者简介:邵桂芳(1978-),女,黑龙江哈尔滨人,重庆大学硕士研究生,研究方向为模式识别与图像处理。

庆龙运智能自动化技术有限公司合作开发下已经进入产品化阶段,并获国家发明专利一项,申请号:01107322.5;国家实用新专利一项,专利号:ZL01214640.4,专利权人,李祖枢。

2 系统中涉及的关键技术

2.1 虚拟现实技术

虚拟现实(Virtual Reality),又称作人造世界、远地呈现、机控宇宙、“灵境”、“幻真”等,它作为信息科学的一个分支,涉及计算机图形学、传感器技术、动力学、光学、人工智能及社会心理学等研究领域,是多媒体发展的更高境界。美国著名计算机图形专家 J. Foley 曾指出:虚拟现实或许是人机接口作为计算机设计的最后一个堡垒中最有意义的领域^[1]。虚拟现实系统具有3个特征:沉浸感(immersion)、交互性(interaction)、人工现实性(artificial reality)。

虚拟环境(Virtual Environment)是利用虚拟现实技术建立的,其视景是由三维计算机图形构成。早期,虚拟环境生成主要采用几何建模并涉及到光线及其表面特征,虽然投入很大,仍难以或几乎不能重现照片中所呈现的自然界中复杂纹理和微妙光照效应。因此,转向了基于图像的环境生成研究,取得了较好的结果。

虚拟背景或虚拟场景是一种相对的概念。针对由摄像机直接可以采集到的真实景物而言,它包含着两方面的含义:1)它是一种真实存在的景物图像;2)相对于前者,它是一种并不实际存在,且不具有实时性的景物图像。

2.2 视频截取

这里提到的视频是有相应声音效果的动态图像更替,是多幅单一图像的快速切换。视频处理所需要面对的一个关键问题是能否实时地分割处理图像,并完成动态显示的功能,这恰恰也是虚拟现实技术普遍面临的技术难题。现在国外最新的娱乐照相系统中,已有能够完成实时处理功能的系统,但是其制造成本和消费成本较高,不利于大范围普及。针对以上一些问题,笔者提出利用软件实现来减少硬件开销,降低成本。当然,如果用软件来实现它,则视频处理的循环周期是能否实现实时性的关键所在。因此从精简算法、合理编程上来考虑尽可能缩短运行周期。

利用蓝背景提取前景人物,首先要获取视频信号。前景图像的获取是通过 CCD 摄像头、视频捕捉卡来完

成的。视频捕捉卡是一块集视频捕捉、视频输出、VGA 显示器输出为一体的多功能卡。在 Video for Windows (简称 VFW)^[2] 出现之前,捕获数字视频是一项极其复杂的工作。而现在微软公司提供了功能强大的视频图像动态捕获软件 VFW,它可同步处理声音和图像。并且微软公司的 Visual C++ 自从 4.0 版就开始支持 VFW,这给视频捕获编程带来了很大的方便。

VFW 是微软公司于 1992 年推出的关于数字视频的一个软件包,它能使应用程序数字化并播放从传统模拟视频源得到的视频剪辑。其关键思想是播放时不需要专用硬件,为了解决数字视频数据量大的问题,要对数据进行压缩。它引进了一种叫 AVI 的文件标准,该标准未规定如何对视频进行捕获、压缩及播放,仅规定视频和音频该如何存储在硬盘上,在 AVI 文件中交替存储视频帧和与之相匹配的音频数据。

VFW 给程序员提供了 .VBX 和 AVICap 窗口类的高级编程工具,使程序员能通过发送消息或设置属性来捕获、播放和编辑视频剪辑。系统研究开发中采用了 AVICap,它支持实时的视频流捕获和单帧捕获,并提供对视频源的控制。同时,它也可将数字视频捕获到文件。AVICap 类集在显示视频时提供了 2 种模式:

1) 预览(Preview)模式。该模式使用 CPU 资源,视频帧先从捕获硬件传到系统内存,可通过操作内存来处理摄像头采集到的图像。在物理上,这种模式需要通过 VGA 卡在监视器上显示。

2) 叠加(Overlay)模式。该模式使用硬件叠加进行视频显示,叠加视频的显示不经过 VGA 卡,叠加视频的硬件将 VGA 的输出信号与其自身的输出信号合并,形成组合信号显示在计算机的监视器上。但目前只有部分视频捕捉卡才具有视频叠加能力。

比较以上两种模式的实用性,笔者采用预览模式来进行前景图像的采集。因为预览模式是通过传递消息机制来进行的,即每一帧图片进入系统时都要发送一个消息,便于控制。在实时自助照相系统中就通过事先登记回调函数来响应其消息,在回调函数中进行抠像、前景与背景合成、图像自适应平滑处理等工作。

2.3 基于对比度感知特征的图像融合技术^[3]

利用彩色分割技术(即扣像),很容易把人像部分从蓝背景中分离出来,有利于实时配上期望的背景图像。但如何使提取出来的人像和背景合成比较自然、逼真,满足人眼的视觉特征,是一个难点。大家知道,

人类视觉对图像的对比度很敏感,广义地讲,人体视觉对灰度对比与色差对比均有较强的感知与认知能力。传统视频领域中的键控技术的本质就是应用视频开关控制图像的“抠”和“填”。“抠”就是利用前景物体轮廓作为遮挡,控制电平将背景画面的颜色沿该轮廓线抠掉,使背景变成黑色。“填”就是将所要叠加的视频信号填到被抠掉的无图像区域中,而最终生成前景物体与叠加背景合成的图像。实践证明,传统键信号波形是前后沿很陡的矩形脉冲信号,在合成输出图像时前景与背景画面的分界处有锯齿、抖动和突变等现象,同时在分界处还有彩色闪烁和幕布色镶边等现象,使人感到生硬和不自然,缺乏真实感。

为了解决这一现象所带来的负面影响,在系统开发中提出了一种新算法:基于对比度感知特征的图像融合技术,其目的就是要消除由直接图像合成所造成的边缘对比度过大,也就是要在缝合边界恢复一个渐变的过渡区^[4]。通过在合成图像中添加合理的过渡区,能进一步弥补提像过程中因为参数波动所带来的人像边缘不平滑现象,提高整个算法的鲁棒性;另一方面,依据图像几何特征所形成的渐变区还能产生“渐入渐出”这样的艺术效果。

图像融合的思路是这样的:对采集进来的人像进行各种处理,如平滑滤波、阈值分割等,得到提取后的人像二值模版,利用所产生的二值模版,应用快速算法逐步地让人像“收缩”,由此得到人像外边缘部分的多个“剥层”,对这样的多个“剥层”赋予不同的加权系数,在合成图像的边缘部分根据这些“加权系数”计算虚拟背景与人像各自所占的比例。

对于上述思路,其中一个主要问题就是如何得到人像边缘部分的几个“剥层”。这里可以利用基于形态学的边缘抽取方法来得到。数学形态学^[5](Mathematical Morphological)是研究数字影像形态结构特征与快速并行处理方法的理论,通过对目标影像的形态变换实现结构分析和特征提取的目的。数学形态学以图像的形态特征为研究对象,它提供了一整套概念、变换和算法,用来描述图像的基本特征和基本结构,也就是描述图像中元素与元素、部分与部分之间的关系。关于数学形态学处理的几个基本概念,如结构元、击中、击不中、平移、腐蚀、膨胀、开闭运算,可参见文献[6]。

数学形态学算子的性能主要以几何方式进行刻画,传统的理论却以解析方程式的形式描述算子的性

能,而几何描述特点则更适合视觉信息的处理和分析。数学形态学的处理方法见图1。

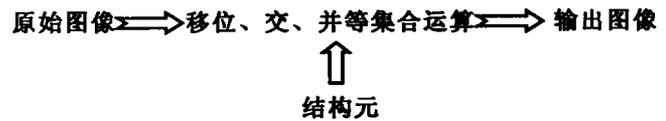


图1 数学形态学的方法

当一个图形被中心在其内部的结构元腐蚀时,所得结果是原图形小了一圈,也就是说腐蚀类似于收缩,如图2所示。

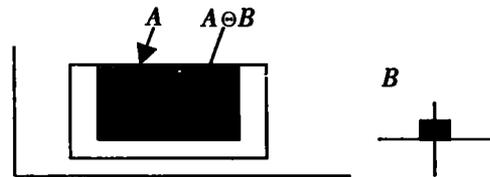


图2 腐蚀运算示意图

设 A 和 B 分别为二维欧氏空间 E^N 集合中的子集,设 A 为物体, B 为某种结构元(structure element)。集合 A 被集合结构元 B 腐蚀记作: $A \ominus B$, 其意义为:

$$A \ominus B = \{x: B+x \subset A\} \text{ 或 } A \ominus B = \cap \{A-b; b \in B\} \quad (1)$$

由式(1)知,用 B 对二值人像模版 A 的腐蚀操作相当于将图 A 左、右、下各移一个点距后所得3个图“相与”,设所得结果为 C ,则 C 即为原人像模版在右、左、上3个方向各收缩了一个像素的缩小二值人像模版。

根据以上结论,可以很容易地构造一个快速算法来实现人像二值模版的“收缩”处理。实际应用中,通过对人像进行连续3次腐蚀操作,得到3个渐变等级系数,再利用这3个渐变等级系数将人像与实际背景进行融合,所得图像效果甚佳。

3 实验结果及讨论

实验中,采用 Honeywell 的 CCD 摄像头采集人像,在 PC 机上对其作各种处理后得到提取后的人像二值模版。为了验证本文算法的效果,利用前面得到的人像二值模版进行了大量的实验,现给出一组实验结果对比如下:图3为原始人像图,图4为基于对比度感知特征的图像融合算法处理过的人像与实际背景的合成图像,图5是经过传统键控方法处理后得到的合成图像。实验结果表明,用基于对比度感知特征的图像融合算法来进行人像与背景的合成,图像自然、逼真。其

简单、精确、快速等优点使得本算法在电视会议、虚拟演播室、可视电话等领域有着广阔的应用前景。



图3 原始人像



图4 融合处理后的图像



图5 传统键控法合成图像

4 结束语

在各方面技术的支持下,系统基本满足了视频图像实时显示的要求。但在当前情况下,由于各种技术(图像处理技术、三维造型技术、硬件加速技术)尚不能达到系统的要求,要完成复杂场景中人像的实时显示,还有一定的困难。系统的后续研究工作将向着以下方面发展:人与场景的实时交互漫游,使人具有在场景中漫游时可以感受到与实际中的感受相同的感觉,提高场景的三维动态效果等。

参考文献

- [1] 张正喜,赵军军. 多媒体技术的发展[J]. 宝鸡文理学院学报(自然科学版),1998,98(3):51-53.
- [2] 谢勇. Visual C++ 6.0 实例精通[M]. 北京:科学出版社,2000. 386-409.
- [3] 张华. 基于人类视觉特征的实时人像提取及其在虚拟环境照相中的应用[D]. 重庆:重庆大学,2002.
- [4] 马向英,杜威,袁晓君,等. 基于图像的室内虚拟漫游系统[J]. 北京:中国图形图像学报,2001,6(1):86-91.
- [5] 龚炜,石青云,程民德. 数字空间中的数学形态学:理论及应用[M]. 北京:科学出版社,1997.
- [6] LEES U, CHUNGS Y. A Comparative Performance Study of Several Global Threshold Techniques for Segmentation[J]. Computer Vision, Graphics and Image Processing, 1990, 152:171-190.

Real-Time Self-Photographing Technology in the Virtual Environment

SHAO Gui-fang, LI Zu-shu, ZHANG Hua

(Institute of Intelligent Automation, Chongqing University, Chongqing 400044, China)

Abstract: As a new technology towards 21st century, the occurs of the virtual reality technology is a result of computer graphic, human-machine interface technology, sensor technology and artificial intelligence. The real-time self-photographing technology based on the virtual environment is a typical application of this virtual reality technique. A series of things has been discussed and studied, such as visual reality, video intercept and so on, and the real-time self-photographing system based on the technology of blue background image extracting. A new arithmetic is emphasized which is produced during study and exploitation. It is image fusion technique based on the contrast appreciate characteristic.

Keywords: virtual environment; self-photographing; video intercept; image fusion

(责任编辑 张 苹)