

虚拟广告系统关键技术研究

王建宇¹, 刘哲¹, 周献中¹, 吴慧中²(¹南京理工大学自动化系, 南京 210094; ²南京理工大学计算机系, 南京 210094)

摘要:近几年,虚拟广告系统在全球广播电视行业中得到广泛的应用。首先通过介绍虚拟广告系统的基本原理,结合作者实现的小型虚拟广告系统的具体环节,对系统中涉及的关键技术在理论上进行了探讨。然后详细介绍了作者实现的小型虚拟广告系统的软硬件开发环境,体系结构,以及在工程实践中采用的具体设计方案。最后对该系统性能进行了分析说明。

关键词:虚拟广告;摄像机标定;运动模糊;视频合成

文章编号:1004-731X(2005)07-1686-04

中图分类号:TP391.9

文献标识码:A

Research of key technology in virtual advertisement system

WANG Jian-yu¹, LIU zhe¹, ZHOU Xian-zhong¹, WU Hui-zhong²(¹Department of Automatic Control, Nanjing University of Science and Technology, Nanjing 210094, China;

²Department of Computer Science, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: The virtual advertisement systems have been used in the field of broadcast and television in the world. The key technology of the system was researched theoretically based on the principle of virtual advertisement and the small virtual advertisement system realized. The software and hardware developing environment, the architecture and the design scheme in projection of the small virtual advertisement system realized were introduced. The performances of system were analyzed.

Key words: virtual advertisement; camera calibration; motion blur; video composition

引言

在现代社会中,电视已成为最广泛的大众传播媒介和娱乐工具,其中广告形式也层出不穷。随着计算机视频技术的不断发展,虚拟广告技术应运而生。虚拟广告系统可广泛应用各种体育或文艺节目的直播中,将演播室制作的虚拟广告牌插入到赛场或表演场的空地上,或用虚拟广告牌替换掉场地上原有的广告牌,合成后可达到以假乱真的效果。利用虚拟广告系统可以增强广告的有效性,提高资源利用率。各种二维或三维的动画广告更能吸引观众的注意力。使用虚拟广告后,广告位置、广告尺寸不再有任何限制。而且,虚拟广告可插入到以前很难或根本无法利用的空间,如水面、沙滩、雪地或非常高的地方^[9]。该技术研究具有很高的应用价值。

1 基本原理

目前全球著名的虚拟广告系统是以色列的 Orad 公司的 IMAgINE 系统,它是目前唯一能替换或者插入篮球、足球、冰球、赛马等护栏广告牌的虚拟广告系统,具有许多独特功能。能处理快速摄像机运动并能在各种天气条件下工作。此外 EPSIS 现场虚拟广告系统也在全球有较高的使用率。美国 PRINCETON 等公司也推出了自己的虚拟广告系统。各

个厂家的系统虽然各有特点,但原理基本相同。原理结构如图 1 所示。

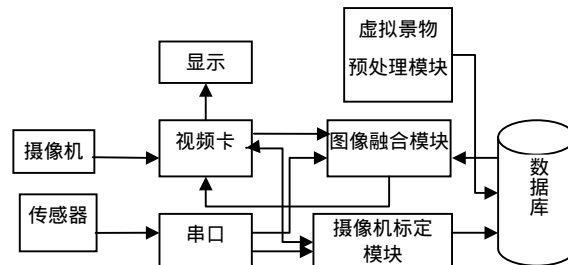


图 1 系统基本结构

从图 1 可以看出,在虚拟广告系统中就是将摄像机拍摄下来的景物作为背景,通过获取摄像机的摇移、俯仰、变焦、聚焦、摄像机空间位置等运动参数,计算出广告贴放的位置、大小、形状等信息去控制广告图像生成,使之生成与背景图像保持正确透视关系的广告图像,最后利用视频合成技术把广告图像叠加在背景上合成输出。

当然实际的虚拟广告系统还要复杂的多,其中还包括视频延时器、音频延时回放系统、信号回路切换台等环节。

2 关键技术

虚拟广告系统涉及的关键技术有摄像机标定、运动模糊、视频合成等。

2.1 摄像机标定

在图像采集中需要将客观世界的 3-D 场景投影到摄像机的 2-D 像平面上,最常用的成像变换是几何透视变换,其特点是随着 3-D 场景与摄像机之间的距离变化,像平面

收稿日期:2003-07-15

修回日期:2004-03-15

作者简介:王建宇(1964-),男,教授,研究方向为图形学、虚拟现实;
刘哲(1978-),男,研究生,研究方向为计算机视频处理;周献中(1962-),男,教授、博导,研究方向为计算机仿真;吴慧中,女,教授,博导,研究方向为计算机仿真,虚拟现实。

上的投影也发生变化,本系统采用 Pin-hole 透镜投影模型。要能正确反映实际场景,首先必须进行摄像机标定,摄像机参数一般分为外部参数(空间位置、方向,即旋转矩阵和平移矩阵)和内部参数(焦距、光心、纵横比、畸变系数等)。

对带畸变的 Pin-hole 模型,从客观场景到数字图像的成像变换由以下 4 步组成:

(1) 从世界坐标 (x_w, y_w, z_w) 到摄像机 3-D 坐标 (x, y, z) 的变换(假设景物为刚体):

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = R \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + T \quad (1)$$

(2) 从摄像机 3-D 坐标 (x, y, z) 到无失真像平面坐标 (x_u, y_u) 的变换(假设 $z \gg f$):

$$x_u = f \frac{x}{z}, y_u = f \frac{y}{z} \quad (2)$$

(3) 从无失真像平面坐标 (x_u, y_u) 到受镜头径向畸变影响而偏移的实际像平面坐标 (x_d, y_d) 的变换:

$$x_d = x_u (1 + kr^2)^{-1}, y_d = y_u (1 + kr^2)^{-1} \quad (3)$$

式中 $r^2 = x_u^2 + y_u^2$, k 为畸变系数。

(4) 从实际像平面坐标 (x_d, y_d) 到计算机图像(帧存)坐标的变换:

$$x_f = N_x x_d + x_c, y_f = N_y y_d + y_c \quad (4)$$

式中 (x_c, y_c) 为图像中心坐标, (N_x, N_y) 为像平面 x 方向和 y 方向的比例系数,亦即图像平面上单位距离上的像素点数。

本系统中考虑一种简化的形式,假定摄像机位置固定,仅考虑内部和畸变参数,经过对摄像机标定理论算法的研究,并结合本系统的具体要求,采取以下的技术措施进行摄像机标定。

(1) 用直径比 N_y/N_x 方法计算比例系数测定。

(2) 用变焦距法计算图像中心 (x_c, y_c) 。

(3) 用一幅含有 N 个共面特征点 (x_{wi}, y_{wi}) 的标定体图像,然后确定 N 个特征点的图像坐标 (x_{fi}, y_{fi}) ,

(4) 利用式(1)、(4)可推得:

$$\begin{bmatrix} y_i & -(y_f - y_c)/N_y \end{bmatrix} \begin{bmatrix} f \\ T_z \end{bmatrix} = (y_f - y_c)w_i/N_y$$

其中 $w_i = r_7 x_{wi} + r_8 y_{wi}$ (5)

根据式(5)对每一特征点列一方程,联立并求解这 N 个方程组,求出焦距 f 和平移矢量 T 的 T_z 分量。

(5) 用前面求出的值作初值,利用优化算法求出 f 、 k 的精确值。

2.2 运动模糊

运动模糊(Motion Blur)是景物图像移动的效果。摄像机的工作原理是很短的时间里把场景在胶片上曝光。如果在曝光的过程中,场景发生变化,就会产生模糊的画面。不难发现在视频流中的每一帧图像中的运动物体是存在模糊的,但是由于人眼的视觉暂留特性,使得人感觉是平滑自然的。在电视的体育报道中,也能看到类似的效果。因

为普通电视节目使用的是 25 帧/秒的摄像机,而体育节目则用的是 100 帧/秒的摄像机,从而提供了清晰的慢镜头回放。在虚拟广告中,大多数是以图象形式出现,图像“棱角分明”很注目,如果不做运动模糊的处理,在摄像机移动过程中同样存在跳跃感较强的问题。

图像运动模糊是景物运动造成的,对流方程是物理学中描述运动的一种方程,可用来分析景物的运动。一般摄像机移动可认为是匀速直线运动,故只需研究一维情况,不妨设运动在 x 方向,则描述匀速直线运动的对流方程为^[5]:

$$\left(\frac{\partial}{\partial t} + v \frac{\partial}{\partial x}\right) f = 0 \quad (6)$$

其中 v 是常数,变量 t 表示时间,而 $f = f(x, t)$ 是随时间改变的一维图像,式(6)描述了图像沿 x 轴方向速度为 v 的匀速直线运动。

设函数 $p(x, t)$ 表示随曝光时间 t 的不同,照相底片上形成的不同图像。底片感光密度与入射光强和曝光时间在一定范围内成正比,因此当聚焦正确时, $p(x, t)$ 是景物影像 $f(x, t)$ 的积分。由于 $f(x, t)$ 是时变函数, $f(x, t) = s(x - vt)$ 为等价变换。 $s(x, t)$ 因为影像叠置而模糊的影像。再对积分部分作变量代换 $x - vt$ 后,我们可得到描述图像运动模糊的表达式:

$$p(x, t) = \frac{1}{v} \int_{x-vt}^x s(\zeta) d\zeta = t \left(\frac{1}{vt} \int_{x-vt}^x s(\zeta) d\zeta \right) \quad (7)$$

这表明,图像运动模糊相当于在运动方向上对原始图像作邻域平均后再乘上曝光时间,对原始图像起平滑作用。取曝光时段为常量 T ,根据式(7)有:

$$p(x, T) = \frac{1}{v} \int_{x-vT}^x s(\zeta) d\zeta \quad (8)$$

在数字图像中应为离散求和的形式。记 $p(x, T)$ 的离散值为 $p(n)$, $s(x)$ 的离散值为 $s(n)$,式(8)变为:

$$p(n) = \frac{1}{v} \sum_{i=n-D}^n s(i) \quad (9)$$

此即运动模糊的离散形式,其中 $D = v \times T$ 是曝光期间景物移动的距离。对于 y 方向的运动雷同。因此其他角度方向的运动可以看作这两个方向运动的合成^[10]。

图 2a 所示系统所用的广告是大小是 500 × 80 BMP 格式标准位图。广告顺时针旋转 45°,向上、下、左、右四个水平方向运动,渲染 10 个像素的模糊效果如图 2b、2c、2d、2e 所示。

运动模糊在实时程序中是很难实现的,它要求把帧速提高几倍。原理简单但是实现起来耗费时间极大。所以作者采用在标定完成之后得到广告位图的旋转角度,然后按照这个角度事先渲染出广告位图上下左右四个运动方向的运动模糊效果,通过实时判断摄像机移动的方向来叠加对应方向渲染效果的广告,取得较好的效果。

2.3 视频对象提取

最后视频合成之前,首先应该完成视频对象提取,这样才能提高虚拟制作的真实感。虚拟广告系统就是所谓的“遮挡问题”。在实际应用中大多数仅对简单背景进行。简单背景提取是类似虚拟演播室中的蓝箱技术。设 Ca 为当前帧的

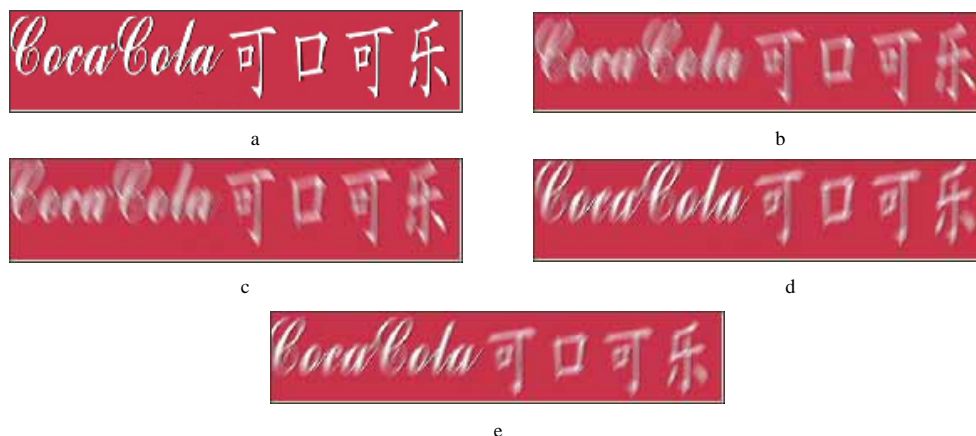


图2 运动模糊实验效果

颜色值, C_b 为背景的颜色值, C_a' 为视频对象提取后视频帧的颜色值, 进一步考虑理想中的简单背景并不存在, 背景其实是一个颜色范围, 设 T 为颜色的域值范围, 则有:

$$C_a' = \begin{cases} C_a, \dots & |C_a - C_b| > T \\ 0, \dots & |C_a - C_b| \leq T \end{cases} \quad (10)$$

对简单背景下的前景目标, 通过色差为特征进行提取。最简单的算法就是按欧氏距离度量的特征提取方法。假设单一背景 n 个像素点颜色均值为 $E(r_e, g_e, b_e)$, 这里

$$\begin{aligned} r_e &= (\sum_{i=0}^n r_i) / n; \quad g_e = (\sum_{i=0}^n g_i) / n; \\ b_e &= (\sum_{i=0}^n b_i) / n \end{aligned} \quad (11)$$

欧氏距离

$$d_i^2 = (r_i - r_e)^2 + (g_i - g_e)^2 + (b_i - b_e)^2 \quad (12)$$

设 R 为域值半径。当 $d_i^2 > R^2$ 时候, 我们认为这是前景目标上的点, 进行提取; 当 $d_i^2 \leq R^2$ 的时候, 我们认为这是单一背景上的点保留。

对于彩色图像的图像分割, 一般的分割方法都是基于 RGB 的彩色分量的计算, 但有时 RGB 分量不能确切体现图像中目标物体的特征, 而经 HSL 变换后的图像特征明显, 易于进行边缘检测、分割和目标识别处理。目前对于视频目标提取的算法很多, 具体采用什么算法, 要看具体的场合和条件。

2.4 视频合成与同步技术

虚拟广告系统的视频合成技术就是在摄像机拍摄的虚拟广告所在位置的真实背景中, 通过图像处理的办法对出现在该位置上的视频对象实时进行提取, 然后与计算机生成的虚拟广告合成一个有深度信息的画面。视频合成的过程也就是广告位置背景视频信号、虚拟广告和广告位置上视频对象这三层视频信号的同步叠加过程。

要使虚拟广告和背景天衣无缝的合成, 就要使得让背景视频信号与虚拟广告信号同步。因为当进行实时合成时, 虚拟广告信号的每一帧是与背景视频信号的对应帧合成。虚拟广告生成由于摄像机跟踪系统和实时生成都要耗费时间,

所以要让背景信号延时来保证与虚拟广告的不同步合成。在实际系统中, 一般都增加帧延时器来解决这个问题。

3 系统构成

一套典型虚拟广告系统的构成应该是由摄像机跟踪系统, 虚拟广告生成系统和视频合成系统三大部分组成。本系统实现的小型虚拟广告系统的构成基本相同, 只是虚拟广告生成部分是在 PC 机上实现的, 没有用专门的图形工作站。本系统采用的软硬件环境: PC 机为开发平台, Windows 2000 操作系统, 512M 内存, 广播级视频卡一块, 监视器一台, 角编码器 3 个, 广播级摄像机一台以及相应的云台和三角架等。图 3 就是本系统的结构框图。

3.1 摄像机跟踪系统

摄像机跟踪系统就是一个能够随时检测和提供摄像机运动参数和广告位置参数的系统, 是实现虚拟广告、背景图像完全同步联动、再现起初的透视关系的关键所在。目前常用的摄像机跟踪系统主要有编码器和图像识别方式。

图像识别方式是指通过摄像机拍摄的图像的形态识别和分析来确定摄像机的各种运动参数, 结合摄像机标定算法对场景中的特征点进行识别。它的优点也很明显, 免除了编码器方式繁琐的安装和调校。另外, 这种方式不需要对摄像机进行任何改造, 可直接使用原有的摄像机。但这种方式要对图像进行分析、计算时间长, 摄像机运动参数精度较低, 现场噪声很难估计, 从图像处理的角度来说比较难达到广播级的要求。正是由于这些考虑本系统采用的是编码器方式。

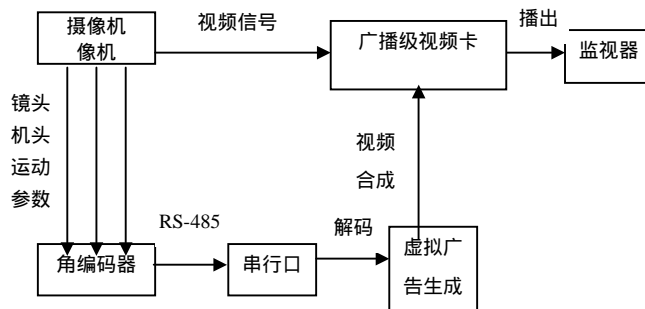


图3 系统结构图

本系统中主要对变焦、摇移、俯仰三个运动参数进行跟踪。在摄像机镜头和机头上分别安装不同的运动参数编码器,当变焦、摇移、俯仰位置发生变化时,编码器能实时检测出变化并将其编码输出。输出的3路数据(变焦、摇移、俯仰)通过1个高速的串行位置接口(SPI)混合成1路串行数据送给PC机。由于利用编码器方式检测摄像机运动参数,使本系统具有了处理速度快、精确度高、工作稳定等诸多优点。

3.2 视频合成系统

在本系统中,视频的采集和合成工作都是由用于直播系统的高端视频卡完成的,系统选用PAL制式。由于采用了DirectX相关技术,视频卡把一帧化成两场,以场为单位处理视频信号。并且它允许事先读入一副标准大小的位图,转化成与视频信号的奇偶场对应的两个Direct3D Surface,覆盖在视频信号上播出。通过开发视频卡的SDK可以获得每场视频信号和Surface数据区的指针,这样就可以人为的截取视频信号获取标定的特定场景图。实现时通过特征标志完成摄像机的标定。通过标定计算把处理好的广告位图拷贝到这层Surface相应的位置上,从而实现广告与背景视频合成

的功能。本系统采用的虚拟广告是一幅BMP位图,当然也可以将3DMAX等软件生成的广告动画文件或是其他广告序列图像合成进来。

4 系统分析与结论

在足球场上实验的效果如图4所示,4a、4b、4c、4d为本系统在不同位置加入广告的效果图。实现时首先通过初始标定模块完成摄像机内参数计算,为保证不会因为各个环节的实时计算耗时太多而造成播出丢帧、系统死机等问题,采取了多线程并行处理与自定义同步消息机制的模式。视频线程的速度是由视频卡保证,每秒50次。串口通讯线程基本上是视频线程速度的40-50倍,标定计算的线程是串口通讯线程速度的100倍左右,这样完全能够保证实时要求。

本系统实现的小型虚拟广告系统,广告定位准确,运动平滑自然,运动中广告位置基本无偏差,变焦时广告放缩采用运动模糊技术真实感强,达到广播级要求,设置广告位图的Alpha通道,使广告与背景具有融合效果,减少不真实感。

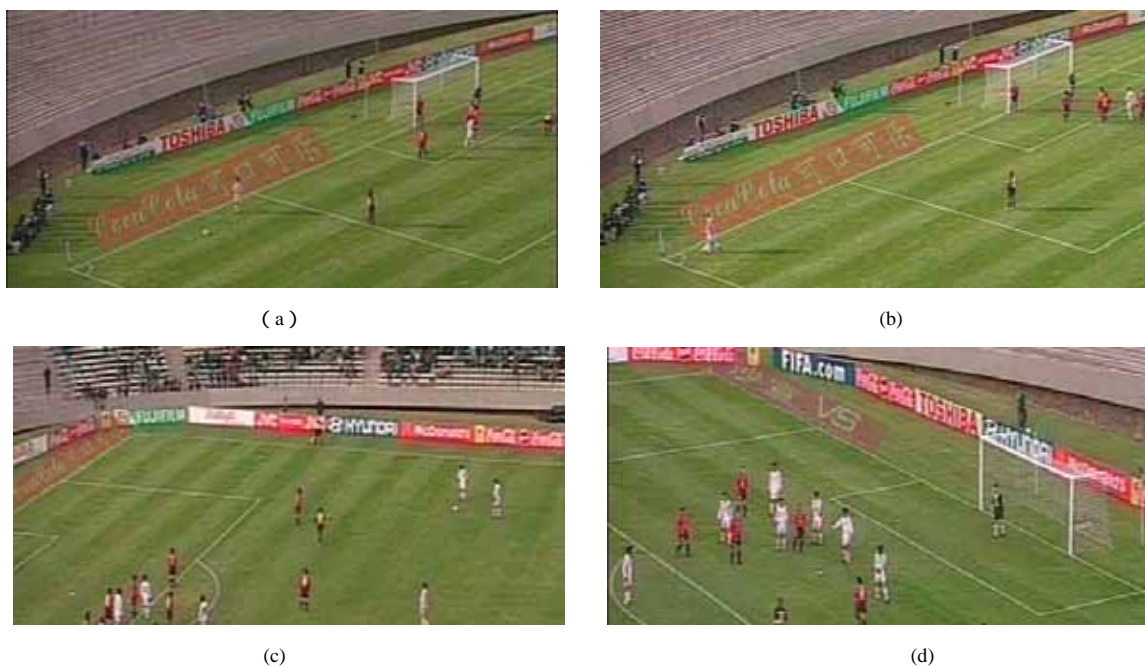


图4 实际效果图

参考文献:

- [1] Reimar K L, et al. Techniques of Calibration of the scale factor and image center for high accuracy 3-D machine vision metrology[J]. IEEE Trans, on PAMI 1988,10(5):713-720.
- [2] Juyang Weng, et al. Camera calibration with distortion models and accuracy evaluation[J]. IEEE Trans. On PAMI 1992,14(10): 965-980.
- [3] T.A.Clarke, J.G.Fryer. The Development of Camera Calibration Methods and Models[J]. Photogrammetric Record, 1998, 16(91): 51-66.
- [4] R.Y.Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using TV camera and lenses[J]. IEEE Journal of Automation, 1987,3(4):323-344.
- [5] COURANT R, HILBERT D. Methods of Mathemahcal Physicis (Vol. 1)[M]. NewYork: Wiley(Interscience), 1953.
- [6] 郑南宁.计算机视觉与模式识别[M]. 北京:国防工业出版社,1998. 22-31.
- [7] 章毓晋.图像工程(下册)——图像理解与计算机视觉[M]. 北京:清华大学出版社,2000.
- [8] 扬冰,李梦君,张茂军.虚拟空间会议系统中视频合成技术的研究与实现[J]. 小型微型计算机系统, 2001, 6(22): 683-686.
- [9] 李晓梅.虚拟制作在奥运报道的应用[J]. 广播电视与制作, 2002, 5: 27-28.
- [10] 郑捷文,蔡利栋.匀速直线运动模糊图像的恢复[J]. 暨南大学学报(自然科学版), 2002,6: 23-25.