

一种基于视觉的增强现实三维注册算法

施琦 王涌天 陈靖

(北京理工大学光电工程系, 北京 100081)

摘要 为了提高增强现实系统三维注册的精度和效率,提出了一种基于计算机视觉的增强现实三维注册算法,并已将其用于采用透射式头盔显示器的增强现实系统中。该算法有如下特点:①构架简单,实用性强,一般情况下只需4个平面标志物就可实现三维注册;②工作范围大,甚至可以应用到室外的增强现实系统中;③数值求解过程是线性过程,误差小,可以满足增强现实系统高精度三维注册的要求。另外,应用这种算法所需要作的图象处理计算量很小,其典型系统包括一个彩色 CCD 和几个不同颜色的标志点,由于很容易获得标志点的对应像面位置,而且不需要计算图象对,因此实时性好,是一般图形工作站和 PC 机上的增强现实系统进行实时注册的算法基础。

关键词 基于视觉三维注册 增强现实 计算机视觉

中图分类号: TP391.9 **文献标识码**: A **文章编号**: 1006-8961(2002)07-0679-05

Vision-Based Algorithm for Augmented Reality Registration

SHI Qi, WANG Yong-tian, CHENG Jing

(Dept. of Optoelectronic Engineering, Beijing Institute of Technology, Beijing 100081)

Abstract Instead of computely redacting the real world as a Virtual Reality(VR) system does, the user can see computer-generated 3D model superimposed on the real world with an Augmented Reality(AR) system. One of the key technologies of AR is the real-time 3D registration of objects in the real world. This paper describes a vision-based method for AR registration and its application in an AR system with See-Through Head Mounted Display (STHMD). The scheme has a simple structure, which employs one color CCD camera and several mark points with different colors. The real world coordinates of the markers are known, and their screen coordinates are obtained through color filtering of the input video stream. If all the markers are successfully tracked, a translation matrix and an orientation matrix can be uniquely determined using projective transform relations. The algorithm requires only linear equation solving with small calculation errors and little computation time consumption. Therefore it has good real-time performance characteristics and can be employed with either graphic workstations or ordinary PCs. The method can work at long ranges and is applicable to outdoor augmented reality systems.

Keywords Vision-based 3D registration, Augmented reality, Computer vision

0 引言

增强现实是从虚拟现实技术中分支出来的一个较新的研究领域,它是一种利用计算机产生的附加信息来对使用者所看到的真实世界景象进行增强或扩张的技术^[1,2]。增强现实与虚拟现实的区别在于,

它不是把真实世界与使用者隔离开,而是将计算机生成的虚拟物体或场景叠加到真实场景中,从而实现对现实的“增强”。在增强现实的环境中,使用者在看到周围真实环境的同时,也可以看到计算机产生的增强信息。这种增强的信息可以是在真实环境中与真实环境共存的虚拟物体,也可以是与存在的真实物体有关的非几何信息。在使用者眼里,真实物体

基金项目:国家杰出青年科学基金(60025513)

收稿日期:2001-04-10; 改回日期:2001-09-13

和虚拟物体是共存的,也是互相增强或者互为补充的.由于它具有很大的应用和发展潜力,因此正越来越引起人们的关注,众多大学和研究机构都在从事这方面的研究和探索.

三维环境注册技术不仅是实现增强现实的基础技术,也是保证增强现实系统性能的关键技术,因此一直是增强现实系统研究中的重要方面.增强现实系统利用附加的图形或文字信息来对周围真实世界的场景进行动态地增强.因为当转动和移动头部的时候,眼睛所看到的视野将随之变动,而且计算机生产的增强信息也应该随之做相应的变化,所以,增强现实系统必须能够实时地检测出回路中人的头部位置和视线指向,然后才能根据这些信息实时地确定所要添加的虚拟信息在真实空间坐标中的映射位置,并将这些信息实时地显示在图象的正确位置,这就是三维环境注册系统所要完成的任务.虽然利用电磁跟踪器^[3]或者超声波探测器^[4]也可以进行三维注册,但是这些注册方法的精度和使用范围都不能满足增强现实的需要,同时也容易受到外界干扰,因而通用性差;而常用的立体视觉注册方法^[5],一般需要两个以上的摄像机,还要通过立体图象对的分析来进行注册,由于这类方法往往算法复杂,而且需要作大量的图象处理运算,因此将使系统延时增加,实时性能下降.

1 算法原理及理论推导

增强现实系统的三维环境注册包含使用者头部(摄像机)的空间定位跟踪以及虚拟物体在真实空间中的定位两个方面的内容.

1.1 算法原理

为了说明本文的注册算法原理,这里采用4个坐标系来进行描述,如图1所示, x,y,z 是绝对空间坐标系(真实空间坐标系); ξ,η,ζ 表示虚拟空间坐标

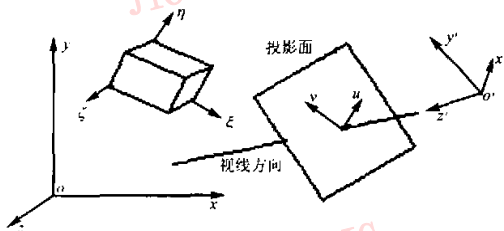


图1 虚拟物体投影注册算法原理图

系,此坐标系用来对所添加的虚拟空间进行几何描述; x',y',z' (左手坐标系)表示观察空间坐标系(摄像机坐标系); uv 是一个二维坐标系,表示投影图象平面坐标系,此坐标系的 $o'z'$ 轴与观察者视线方向重合.

由于所需要添加的虚拟物体在真实空间坐标系中的方位是由系统要完成的功能所决定的,即虚拟空间坐标系 $\xi\eta\zeta$ 与真实空间坐标系 xyz 的关系是已知的,所以虚拟空间坐标系 $\xi\eta\zeta$ 中三维虚拟物体的几何描述 (ξ,η,ζ) 可以变换为真实空间坐标系 xyz 中的几何描述 (x,y,z)

$$\begin{bmatrix} x & y & z & 1 \end{bmatrix} = \begin{bmatrix} \xi & \eta & \zeta & 1 \end{bmatrix} A \quad (1)$$

上式中, A 是两个坐标系之间已知的变换矩阵.

如果能够求出真实空间坐标系 xyz 与观察空间坐标系 $x'y'z'$ 之间的变换矩阵 B ,就可以将虚拟物体坐标系中三维虚拟物体的几何描述 (ξ,η,ζ) 变换为观察空间坐标系中的几何描述 (x',y',z')

$$\begin{bmatrix} x' & y' & z' & 1 \end{bmatrix} = \begin{bmatrix} \xi & \eta & \zeta & 1 \end{bmatrix} AB \quad (2)$$

此时,再将观察空间坐标系中的虚拟物体投影到图象平面坐标系 uv 中,就完成了增强现实系统的三维注册.

1.2 三维基准注册的算法推导

三维基准注册的任务是实现使用者头部(摄像机)的空间定位和跟踪,通过上面的分析可见,由于只要求出矩阵 B 就达到了目的,因此本文算法的基本步骤是:(1)已知标志点在场景中的坐标 (x,y,z) ;(2)测定标志点的投影坐标 (u,v) ;(3)利用透视投影方程建立图象平面点与场景点的联系;(4)解出变换矩阵 B .

为方便起见,在这里虚拟物体几何表达与真实物体几何表达的关系用如下普通的变换表达式^[6]

$$\begin{pmatrix} x' & y' & z' \end{pmatrix}^T = \mathbf{R} \begin{pmatrix} x & y & z \end{pmatrix}^T + \mathbf{T} \quad (3)$$

其中,

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

$$\mathbf{T} = (t_x, t_y, t_z)$$

从真实空间坐标系到观察空间坐标系的变换如下

$$\begin{aligned} x' &= r_{11}x + r_{12}y + r_{13}z + t_x \\ y' &= r_{21}x + r_{22}y + r_{23}z + t_y \\ z' &= r_{31}x + r_{32}y + r_{33}z + t_z \end{aligned} \quad (4)$$

在观察空间坐标系中,虽然场景点的位置是未知的,但是点在图象平面上的投影可以由透视投影方程来

确定

$$\frac{u}{F} = \frac{x'}{z'} = \frac{r_{xx}x + r_{xy}y + r_{xz}z + t_x}{r_{zx}x + r_{zy}y + r_{zz}z + t_z} \quad (5)$$

$$\frac{v}{F} = \frac{y'}{z'} = \frac{r_{yx}x + r_{yy}y + r_{yz}z + t_y}{r_{zx}x + r_{zy}y + r_{zz}z + t_z}$$

其中, F 为摄像机的焦距(这里假设摄像机的内部参数已知, 即 F 已知), 由联立式(4)和式(5)可知, 每一个标志点可以确定如下两个独立的约束方程

$$u(r_{zx}x + r_{zy}y + r_{zz}z + t_z) - F(r_{xx}x + r_{xy}y + r_{xz}z + t_x) = 0 \quad (6)$$

$$v(r_{zx}x + r_{zy}y + r_{zz}z + t_z) - F(r_{yx}x + r_{yy}y + r_{yz}z + t_y) = 0$$

显然, 由 6 个标志点可以产生 12 个方程, 从而可以解出变换矩阵 B 中的 12 个参数.

实际上, 如果使用旋转矩阵的规范化正交约束, 只需 4 个点就可以解出矩阵 B 中的 12 个参数. 例如在图 5 的应用系统中, 由于标志点都在同一个平面内, 4 个标志点的绝对坐标为 $(x_i, y_i, 0) (i=1, 2, 3, 4)$, 所以式(6)可以简化为

$$u(r_{zx}x + r_{zy}y + 1) - F(r_{xx}x + r_{xy}y + t_x) = 0 \quad (7)$$

$$v(r_{zx}x + r_{zy}y + 1) - F(r_{yx}x + r_{yy}y + t_y) = 0$$

其实, 由于 t_z 只是各参数的比例因子, 所以可以设 $t_z=1$. 同时由于式(7)中有 8 个参数, 而且由 4 个标志点可以得到 8 个方程, 因此解这个线性方程组就可以求出 8 个参数: $r_{xx}, r_{xy}, r_{xz}, r_{yx}, r_{yy}, r_{yz}, t_x$ 和 t_y .

因为旋转矩阵 R 是正交规范化矩阵, 所以 $R^T R = I$ 一定成立, 利用这一特性可以计算如下几个参数

$$r_{xx} = \sqrt{1 - r_{xz}^2 - r_{zy}^2}$$

$$r_{yy} = \sqrt{1 - r_{yz}^2 - r_{zx}^2}$$

$$r_{zz} = \sqrt{1 - r_{zx}r_{xz} - r_{zy}r_{yz}}$$

这样, 就完成了三维基准注册任务.

1.3 虚拟物体投影注册算法推导

完成了虚拟空间坐标系 $\xi\eta\zeta$ 到观察空间坐标系 $x'y'z'$ 的转换后, 下一步就是将虚拟物体投影到画平面坐标系中. 如图 2 所示, 用 OXY 表示要显示的画平面坐标系, D 表示摄像机中心至显示器的距离, 有一个虚拟物体图象点 A 在观察空间坐标系中的坐标为 (x'_a, y'_a, z'_a) , 本文的目的就是计算出 A 点在

画平面坐标系中的坐标 (X_0, Y_0) .

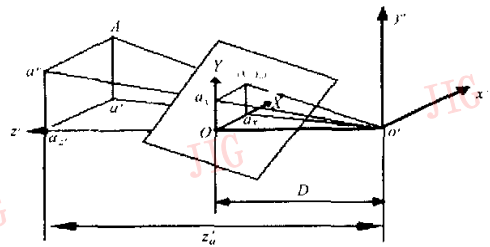


图 2 虚拟物体投影坐标变换

图中, a'_x 和 a'_y 是 A 在 $o'x'z'$ 和 $o'y'z'$ 上的正投影, 它们在画平面上的投影图象分别是 a_x 和 a_y , 由于

$$\Delta o'oa_x \approx \Delta o'a'_xa'_x, \Delta o'oa_y \approx \Delta o'a'_ya'_y$$

所以有

$$\frac{X_0}{D} = \frac{x'_a}{z'_a}, \quad \frac{Y_0}{D} = \frac{y'_a}{z'_a}$$

其中, x'_a, y'_a, z'_a 为 A 点在观察空间坐标系中的坐标(在基准注册中已经求出), 由此可以得到

$$X_0 = \frac{Dx'_a}{z'_a}, \quad Y_0 = \frac{Dy'_a}{z'_a}$$

这样, 就完成了对虚拟物体的注册.

1.4 立体显示问题

如果考虑到增强现实中需要立体显示, 则应对三维注册算法作适当修改. 从立体显示示意图(见图 3)可以看出, 立体显示时, 因为观察点的位置分别左移和右移 d 的距离, 所以需要将 $x'y'z'$ 坐标系沿 x 轴作平移变换, 以便最后得到图象点分别在左右眼平面坐标系中的坐标, 为此, 算法上需要将式(2)作修改, 即

$$\text{左眼} [x'_L \ y'_L \ z'_L \ 1] = [\xi \ \eta \ \zeta \ 1] ABA_L$$

$$\text{右眼} [x'_R \ y'_R \ z'_R \ 1] = [\xi \ \eta \ \zeta \ 1] ABA_R$$

其中,

左眼变换矩阵

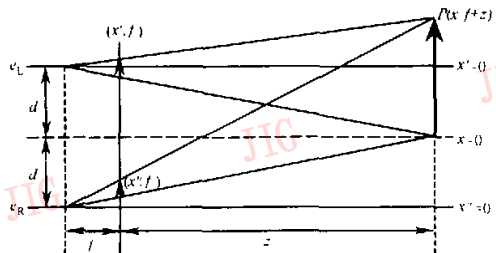


图 3 立体显示坐标变换

$$A_L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ d & 0 & 0 & 1 \end{bmatrix}$$

右眼变换矩阵

$$A_R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ d & 0 & 0 & 1 \end{bmatrix}$$

2 应用实例

笔者开发的一种新型掌上增强现实系统(如图4所示)就成功地应用了上述算法来进行三维注册,并且效果良好.在该系统中,是采用透射式HMD(头盔显示器)来在一块平板上增强显示一个虚拟物

体.具体做法是:首先在平板的4个角上放置4个不同颜色的标志点,并利用头盔上的一个彩色CCD摄像机实时拍摄真实场景;然后在三维注册时,则通过图象处理的方法提取出4个标志点的平面坐标,再用上述算法实现三维场景注册;最后图形工作站则根据该三维注册信息实时生成左、右眼虚拟物体图象对,并同步地送到左右眼的LCD显示器上,从而在用户面前呈现一个虚实结合的场景.

图5展示了几个在实验室里的AR系统上,以不同的角度试验获得的真实场景和复合场景,由图5(b)可见,一只虚拟的恐龙被添加到真实场景中.其中,真实场景只是一个黑色的箱子,戴上AR头盔的用户所看到的是已经添加了虚拟恐龙的复合场景,无论用户以何种角度观察,虚拟的恐龙好像就是站在箱子上,说明该系统所生成的复合场景是成功的.

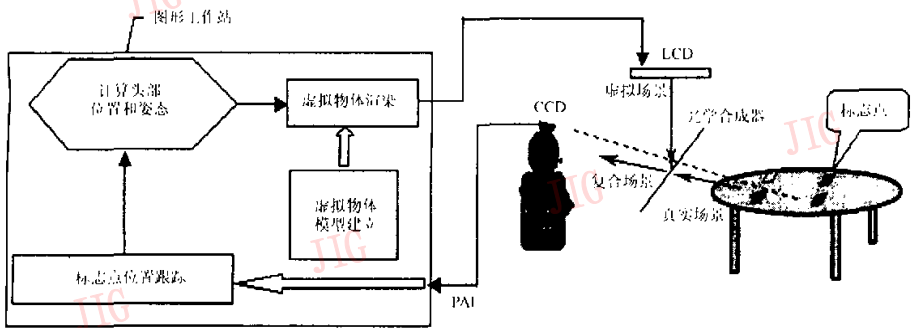


图4 注册算法应用实例

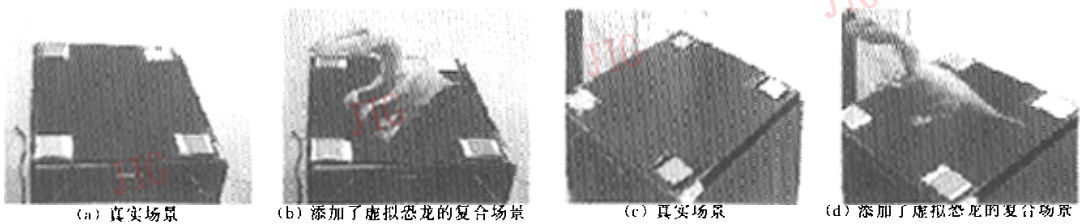


图5 不同角度的真实场景和添加了虚拟恐龙的复合场景

3 结论

与虚拟现实相比,由于在增强现实的环境注册时,不仅要实时地获得用户头部的位置姿态信息,而

且还要实现虚拟物体(由计算机生成的物体)与真实场景的准确融合,因此对注册有更高的要求.这样常用的电磁或超声波传感器位置跟踪法已经不能满足需要了,而两个摄像机的立体图象对注册法又计算量太大,为此,本文从理论上论证了一种基于视觉的

三维注册算法,由于这种算法不仅计算量小,通用性强,而且数值求解是线性过程,因此误差小。通过实际应用证明,该种算法具有良好的三维注册性能,可以满足增强现实的实时三维注册需要。

虽然本文从理论上证明了4个标志点就可以算出注册矩阵的12个参数,但是在实际的增强现实系统应用中,为了提高精度或扩大应用范围,应在一些关键的地方多加几个标志点。

相反,有时为了减小计算量和简化注册模型,可以只选用3个注册点^[7],因为这样做往往是为了忽略某些自由度上变量的改变。

在实际应用中,标志点不一定是实物点,比如可以用光束投影来获得标志点;另外,标志点也不一定像试验中那样是刻意加上的,例如对于行驶中的汽车,标志点可以从道路边的标志物中获得。

总之,基于视觉的三维注册方法是解决增强现实中三维注册问题最有前途的方法。

参考文献

- 1 Azuma R T. A survey of augmented reality[J]. *Teleoperators and Virtual Environments*, 1997, 4(2): 355~385.
- 2 汪成为, 高文, 王行仁. 灵境(虚拟现实)技术的理论、实现及应用[M]. 北京: 清华大学出版社, 1996: 5~17.
- 3 State A, Hirota G, Chen D T *et al*. Superior augmented reality registration by integrating landmark tracking and magnetic tracking[A]. In: Proc. SIGGRAPH'96[C]. New Orleans, 1996: 129~146.

- 4 Bajura M, Fuchs H, Ohbuchi R. Merging virtual objects with the real world: Seeing ultra sound imagery within the patient[A]. In: Proc. SIGGRAPH'92[C]. Chicago, 1992: 203~210.
- 5 Yokoya N, Takemura H, Okuma T *et al*. Stereo vision based video see-through mixed reality[A]. In: Mixed Reality: Merging Real and Virtual Worlds[C]. New York: Ohmsha-Springer Verlag, 1999: 131~145.
- 6 贾云得. 机器视觉[M]. 北京: 科学出版社, 2000: 181~205.
- 7 Simon G, Lepetit V, Berger M O. Computer vision methods for registration: Mixing 3D knowledge and 2D correspondences for accurate image composition[A]. In: Proceedings of IWAR'98[C], San Francisco, Calif, 1998: 58~62.



施琦 1976年生, 1999年获北京理工大学信息学院学士学位, 现为北京理工大学光电工程系研究生。主要研究方向是虚拟现实和增强现实技术。



王涌天 1957年生, 1986年4月获英国 Reading 大学物理系工学博士学位, 现任北京理工大学教授、博士生导师、信息工程学院学术委员会委员。主要研究领域为虚拟现实、增强现实系统及应用、光学设计和 CAD、衍射光学元件、新型光学系统和仪器。发表论文 80 余篇。

陈靖 1976年生, 现为北京理工大学光电工程系博士研究生。主要研究方向是虚拟现实和增强现实技术。